

# **ParA: a novel target for anti-tubercular drug discovery**

A thesis submitted by

**Shahista Yasmin Nisa**

BSc (Hons) Victoria University of Wellington, NZ

for the degree of

Doctor of Philosophy

in the

School of Biological Sciences,

Victoria University of Wellington,

New Zealand, 2010



**Primary supervisor**

Doctor Ronan O'Toole

Senior Lecturer

School of Biological Sciences

Victoria University of Wellington

Wellington, New Zealand

**Co-supervisor**

Doctor Joanna Kirman

Infectious Disease Group Leader

Malaghan Institute of Medical Research

Wellington, New Zealand

## Abstract

Tuberculosis (Tb) continues to be one of the world's greatest challenges in the public health arena. The current treatment for Tb entails a long duration of therapy making adherence to the whole course difficult. This has given rise to drug resistant strains of *Mycobacterium tuberculosis* which are posing a significant threat to Tb control strategies. To counteract this problem, there is an urgent need to develop alternative anti-tuberculous drugs which target processes that are critical for the growth and/or survival of this microbe.

To identify such targets in *M. tuberculosis*, I used comparative genomics and mutagenesis data to identify conserved essential genes as viable targets for the development of broad-spectrum antibiotics. In addition, I validated the essentiality of three cell division genes in *Mycobacterium smegmatis* using conditional antisense RNA expression under different culture conditions. Furthermore, I performed high-throughput screens (HTS) using a differential susceptibility assay against one of the validated targets to identify its cognate inhibitor(s). Lastly, I developed a novel biochemical assay of the target to validate the specificity of the inhibitors identified in the HTS and evaluated the potency of the inhibitors against *M. tuberculosis*.

This study identified 261 conserved putative essential genes as broad-spectrum targets. I hypothesized that antisense RNA expression of such genes will lead to its down-regulation and thereby affect the viability of the cells if these genes are essential. I also hypothesized that an essential gene will be required under all culture conditions. One gene, *parA*, demonstrated that it was essential under various culture conditions. This gene encodes for a protein which contain the conserved Walker A motif thus I theorized that it may possess ATPase activity. The results illustrated that the *M. tuberculosis* ParA protein possesses ATPase activity. This biochemical activity was used to validate two specific inhibitors of ParA, phenoxybenzamine and octoclotheptin, which were identified in the cell-based HTS. Kinetic studies suggest that phenoxybenzamine is a mixed inhibitor while octoclotheptin is a competitive inhibitor of ParA. This data is also supported

by *in silico* docking. Both these compounds show low minimum inhibitory concentrations in *M. smegmatis* under nitrogen starvation conditions.

In summary, this thesis illustrates that ParA is a viable target for anti-tubercular drugs. It demonstrates that ParA is an ATPase which has the potential to bind competitive and non-competitive inhibitors that can be exploited to target cell division in *M. tuberculosis*. Finally, this study presents phenoxybenzamine and octoclotheptin as inhibitors of ParA. In conclusion, these compounds can either be developed to increase potency or be used as reference structures to screen for more potent inhibitors of the enzyme.

## Acknowledgements

First and foremost, I would like to thank my supervisors, Dr Ronan O'Toole, for his guidance, support and for sharing his knowledge with me throughout my PhD and Dr Joanna Kirman, for proof-reading my thesis and for the great enthusiasm and motivation that every meeting with her brought.

I would like to acknowledge Dr David Ackerley for all his support throughout my PhD and his very keen and unduly devoted PhD student, Jeremy Owen, for the development of the brilliant green assay. Furthermore, I would like to acknowledge Prof. Brent Copp for analysing the hits from the HTS and Dr Joel Tyndell for all the *in silico* analyses. I would also like to thank Prof. Paul Atkinson for allowing me to use his high-throughput screening equipment and Chris Miller for training me with the high-throughput robots. Research Assistant Yee Syuen Low, greatly aided in accelerating the rate of my PhD, Mudassar Altaf for his BCG and MTB work and Rekha Murikoli for her toxicity work. In addition, I would like to acknowledge Ian Bassett, for his bactericidal assays and for putting his expert grasp of the English language to good use and proof-reading my thesis. I would also like to take this opportunity to thank everyone else in KK817 and in SBS who didn't really help me in any way during my PhD but went to great efforts just to put up with me.

I wouldn't have been able to have as great a time during my PhD without friends like Sandi Dempsey, for her limitless distractions and for introducing me to awesome people like Dave. And Gareth Prosser, for slowly but surely brinking insanity with me and for being the best insane office-mate ever. Also to all other friends that kept a tab on me while I was being a recluse, especially towards the end of my PhD.

Lastly, but not the least, I would like to thank my fiancé, Christian Schneider, for being my rock at the most difficult of times and for supporting me in every way he possibly could.

*for Christian Schneider*

## Table of Contents

<b>Abstract.....</b>	<b>iii</b>
<b>Acknowledgements.....</b>	<b>v</b>
<b>Table of Contents.....</b>	<b>vii</b>
<b>List of Figures.....</b>	<b>xiii</b>
<b>List of Tables.....</b>	<b>xvi</b>
<b>List of Abbreviations.....</b>	<b>xviii</b>
<b>List of Publications.....</b>	<b>xxii</b>

## **CHAPTER ONE: General Introduction.....1**

1.1 Preface.....	2
1.2 Global Burden of Tuberculosis.....	2
1.2.1 Tuberculosis in New Zealand.....	5
1.3 <i>Mycobacterium tuberculosis</i> .....	5
1.3.1 <i>Mycobacterium smegmatis</i> .....	7
1.4 <i>Mycobacterium tuberculosis</i> Pathogenesis.....	8
1.5 Tb Treatment.....	10
1.5.1 Mode of action (MOA) of first-line drugs.....	10
1.5.2 MDR-Tb.....	11
1.5.3 MOA of second-line drugs.....	12
1.5.4 XDR-Tb.....	12
1.5.5 Mechanisms of Drug Resistance.....	13
1.6 Tb Drug Development.....	14
1.6.1 Current Tb drugs in the pipeline.....	14
1.7 Genomics.....	15
1.7.1 The Mycobacterium Genome.....	17
1.7.2 Drug Targets: essential and conserved genes.....	18
1.7.3 Target Validation.....	19
1.8 Targeted Screening.....	21
1.9 Aims of this study.....	23

## **CHAPTER TWO: Materials and Methods.....24**

2.1 Bacterial Strains and Plasmids.....	25
2.1.1 Bacterial Strains.....	25
2.1.2 Plasmids.....	26

2.2 Media.....	28
2.2.1 <i>E. coli</i> .....	28
2.2.2 <i>M. smegmatis</i> .....	28
2.2.3 <i>M. tuberculosis</i> H37Ra and <i>M. bovis</i> BCG.....	29
2.2.4 Media supplements.....	29
2.3 Culture Conditions.....	30
2.4 Chemicals and Reagents.....	30
2.5 Enzymes.....	31
2.6 Oligonucleotides.....	31
2.7 Bioinformatics.....	33
2.7.1 Quantification of essential genes in <i>M. tuberculosis</i> H37Rv.....	33
2.7.2 Quantification of conserved essential genes in <i>M. tuberculosis</i> H37Rv.....	33
2.7.3 Orthologus of conserved essential genes in <i>M. smegmatis</i> mc <sup>2</sup> 155.....	33
2.7.4 Homology model of <i>M. tuberculosis</i> ParA.....	34
2.7.5 Ligand docking.....	34
2.8 DNA Isolation, Purification and Quantification.....	35
2.8.1 Purification of mycobacterial chromosomal DNA.....	35
2.8.1.1 Phenol/chloroform extraction.....	35
2.8.1.2 ZR fungal/bacterial DNA Kit™.....	36
2.8.2 Purification of plasmid DNA.....	36
2.8.2.1 Phenol/chloroform extraction.....	37
2.8.2.2 Zppy™ Plasmid Miniprep kit.....	37
2.8.3 Gel Purification of DNA.....	38
2.8.4 Purifying Salts and Enzymes from DNA.....	38
2.8.4.1 Nitrocellulose filters.....	39
2.8.4.2 DNA Clean and Concentrator Kit™.....	39
2.8.5 DNA Quantification.....	39
2.9 Genetic Manipulation.....	40
2.9.1 Polymerase Chain reaction (PCR).....	40
2.9.2 DNA Ligations.....	42
2.9.3 Transformations into <i>E. coli</i> .....	43
2.9.3.1 Preparation of Calcium competent Cells.....	44
2.9.3.2 Heat-shock Transformation.....	44
2.9.4 Transformations into <i>M. smegmatis</i> .....	45
2.9.4.1 Preparation of electro-competent cells.....	45
2.9.4.2 Electroporation.....	45
2.9.5 Identification of recombinant clones.....	46
2.9.5.1 Colony screening.....	46
2.9.5.2 PCR products in pCR®2.1.....	47
2.9.5.3 Orientation PCR.....	47
2.9.6 Restriction digests.....	48
2.9.7 Dephosphorylation.....	49
2.9.8 Agarose gel electrophoresis.....	50
2.9.9 DNA sequencing.....	50
2.9.10 Curing plasmids.....	50
2.10 Proteomics.....	51
2.10.1 Protein Over-expression Optimization.....	51



2.10.1.1	Expression of recombinant protein in <i>E. coli</i> LMG194...	51
2.10.1.2	Expression of recombinant protein in <i>E. coli</i> BL21 .....	52
2.10.2	Purification of recombinant proteins.....	53
2.10.3	Desalting Proteins.....	54
2.10.4	Protein quantification.....	54
2.10.5	SDS-PAGE.....	54
2.10.6	Polyclonal Antibody Production.....	55
2.10.7	Western Blots.....	56
2.11	<i>In vitro</i> ATPase Assay.....	57
2.11.1	Brilliant green assay.....	58
2.12	Antisense Assays.....	58
2.12.1	Growth curves.....	58
2.12.1.1	Rich media.....	59
2.12.1.2	Nutrient starvation media.....	60
2.12.2	Colony forming units.....	60
2.12.3	DAPI staining.....	60
2.13	Differential Susceptibility Assays.....	61
2.13.1.1	Assay set-up.....	62
2.13.1.2	Data analysis.....	63
2.13.1.3	Statistical analysis.....	63
2.14	Dose-response Assays.....	64
2.15	Stationary Cidal Assays.....	64
2.16	Cytotoxicity Assays.....	65

## **CHAPTER THREE: Target Selection.....66**

3.1	Introduction.....	67
3.1.1	Comparative genomics.....	67
3.1.2	Identification of novel targets using a genomics approach.....	68
3.2	Objectives of chapter three.....	70
3.3	Results.....	71
3.3.1	Quantification of essential genes in <i>M. tuberculosis</i> H37Rv.....	71
3.3.1.1	Genes mutated from the list published by Sassetti.....	71
3.3.2	Quantification of conserved essential genes in <i>M. tuberculosis</i> H37Rv.....	73
3.3.3	Orthologs of conserved essential genes in <i>M. smegmatis</i> .....	74
3.4	Discussion.....	75
3.4.1	Essential genes.....	75
3.4.2	Broad-spectrum targets.....	76
3.4.3	Comparison to other studies in <i>M. tuberculosis</i> .....	79
3.4.4	Uniqueness of this study.....	80
3.4.5	Summary.....	81

## **CHAPTER Four: Target Validation.....82**

4.1	Introduction.....	83
4.1.1	Approaches to target validation.....	83
4.1.1.1	Antisense Technology.....	84
4.1.1.2	Cell Division.....	87
4.2	Objectives of chapter four.....	87
4.3	Results.....	88
4.3.1	Construction of expression vectors.....	88
4.3.1.1	Construction of antisense vectors.....	90
4.3.2	Electroporation into <i>M. smegmatis</i> .....	93
4.3.3	Determination of optimal tetracycline concentration.....	94
4.3.4	Induction of antisense transcript.....	96
4.3.5	Antisense assays.....	97
4.3.5.1	Antisense expression under rich culture conditions (LBT)....	98
4.3.5.2	Antisense expression under carbon starvation (C <sup>-</sup> ) condition.....	101
4.3.5.3	Antisense expression under nitrogen starvation (N <sup>-</sup> ) conditions.....	104
4.3.6	Measuring protein levels.....	107
4.3.7	Characterization of <i>M. smegmatis</i> -pMind_ parA+parB_GFP.....	108
4.4	Discussion.....	114
4.4.1	Gene modulation.....	114
4.4.1.1	Addressing incomplete repression.....	116
4.4.1.2	Obg.....	117
4.4.1.3	ParA and ParB.....	119
4.4.1.3.1	Chromosome partitioning in mycobacteria.....	124
4.4.1.3.2	Chromosome partitioning in <i>B. subtilis</i> .....	125
4.4.1.3.3	Chromosome partitioning in <i>E. coli</i> .....	126
4.5	Summary.....	127

## **CHAPTER Five: Targeted Screening.....128**

5.1	Introduction.....	129
5.1.1	Parameters of a drug screen.....	129
5.1.2	Differential susceptibility assays.....	130
5.1.3	ATPases.....	132
5.2	Objectives of chapter five.....	133
5.3	Results.....	134
5.3.1	LOPAC Screen.....	135
5.3.1.1	LBT (LOPAC).....	135
5.3.1.2	C <sup>-</sup> (LOPAC).....	137
5.3.1.3	N <sup>-</sup> (LOPAC).....	139
5.3.2	Spectrum Collection.....	142
5.3.2.1	LBT (Spectrum Collection).....	142

5.3.2.2	C <sup>-</sup> (Spectrum Collection).....	144
5.3.2.3	N <sup>-</sup> (Spectrum Collection).....	145
5.3.3	NIH Diversity Set.....	146
5.3.3.1	LBT (NIH Diversity Set).....	146
5.3.3.2	C <sup>-</sup> (NIH Diversity Set).....	149
5.3.3.3	N <sup>-</sup> (NIH Diversity Set).....	151
5.4	Discussion.....	152
5.4.1	Goals of targeted screening.....	153
5.4.2	Biological aspects of the hits.....	154
5.4.3	Chemistry aspects of the hits.....	157
5.4.4	Summary.....	159

## **CHAPTER SIX: Hit Validation.....160**

6.1	Introduction.....	161
6.1.1	Validation assays.....	161
6.2	Objectives of chapter six.....	162
6.3	Results.....	163
6.3.1	Protein over-expression .....	163
6.3.2	Protein purification.....	166
6.3.3	Development of a ParA ATPase assay.....	167
6.3.4	Optimization of the biochemical assay.....	168
6.3.5	Determination of the kinetic properties of the ParA ATPase.....	170
6.3.6	ATPase inhibitors.....	172
6.3.7	Hit validation.....	173
6.3.8	Kinetic parameters of the inhibitors.....	177
6.3.9	Anti-mycobacterial properties of ParA inhibitors.....	177
6.3.10	Cytotoxicity of ParA inhibitors.....	179
6.3.11	<i>In silico</i> interaction between ParA and its inhibitors.....	180
6.4	Discussion.....	182
6.4.1	ParA ATPase.....	183
6.4.2	Inhibition of the ParA ATPase activity.....	185
6.4.3	Hit validation.....	187
6.5	Summary.....	192

## **CHAPTER SEVEN: General Discussion.....193**

7.1	General Discussion.....	194
7.1.1	Summary of experiments and their findings.....	194
7.1.2	Other findings.....	199
7.2	Critical Evaluation.....	200
7.2.1	Target validation technique.....	198

7.2.2	High-throughput assays.....	201
7.2.3	<i>In vitro</i> assay.....	201
7.3	Future Research Directions.....	202
7.3.1	Future goals.....	202
7.3.2	Improvements.....	203
7.4	Concluding Remarks.....	204

<b>References.....</b>	<b>205</b>
------------------------	------------

## **CHAPTER NINE: Appendices.....226**

9.1	Maps of plasmids used in this study.....	227
9.2	Essential Genes.....	228
9.2.1	List of Essential Genes published by Sassetti <i>et. al.</i> ....	228
9.2.2	Conserved Essential Genes.....	230
9.3	Cellular morphology of <i>M. smegmatis</i> strains under different culture conditions.....	239
9.3.1	Cell morphology of cultures in LBT.....	239
9.3.2	Cell morphology of cultures in C <sup>-</sup> media.....	240
9.3.3	Cell morphology of cultures in N <sup>-</sup> media.....	241
9.4	Cellular Morphology of <i>M. smegmatis</i> strains in HTS.....	242
9.4.1	Cell morphology in N <sup>-</sup> media with LOPAC hits.....	242
9.4.2	Cell morphology in N <sup>-</sup> media with Spectrum Collection hits....	243
9.4.3	Cell morphology in C <sup>-</sup> media with NIH Diversity Set hit.....	244
9.5	Media and Solutions.....	245
9.5.1	Hartmans-de Bont minimal media.....	245
9.5.2	ZYP-5052.....	246
9.5.3	Solutions.....	247
9.5.4	Buffers.....	249
9.5.5	SDS gels.....	251

## List of Figures

Figure 1.1: Global burden of Tb.....	4
Figure 1.2: The mycobacterial cell wall.....	7
Figure 1.3: Pathogenesis of Tb infection and disease.....	9
Figure 1.4: The <i>M. tuberculosis</i> H37Rv genome.....	17
Figure 1.5: Flow chart of target-based drug discovery.....	22
Figure 2.1: Plate map used to screen compound libraries.....	63
Figure 4.1: Gene expression.....	86
Figure 4.2: Orientation PCR of pMindGFP.....	89
Figure 4.3: Figure 4.3: pMindGFP vector.....	90
Figure 4.5: PCR product of genes cloned into pMind and pMindGFP.....	91
Figure 4.5: Gene arrangement of <i>parA</i> and <i>parB</i> .....	92
Figure 4.6: Electroporation of <i>M. smegmatis</i> .....	93
Figure 4.7: Determination of optical inducer concentration.....	95
Figure 4.8: Induction of antisense transcript.....	96
Figure 4.9: Growth curves of cultures in LBT.....	99
Figure 4.10: Number of viable cells in LBT.....	100
Figure 4.11: Growth curves of cultures in C <sup>-</sup> media.....	102
Figure 4.12: Number of viable cells in C <sup>-</sup> media.....	103
Figure 4.13: Growth curves of cultures in N <sup>-</sup> media.....	105
Figure 4.14: Number of viable cells in N <sup>-</sup> media.....	106
Figure 4.15: Western blot showing non-specific binding.....	108
Figure 4.16: PCR products of pMind6938+6939GFP.....	109
Figure 4.17: Electroporation of <i>M. smegmatis</i> .....	110
Figure 4.18: Cured <i>M. smegmatis</i> strain.....	110
Figure 4.19: Electroporation of cured <i>M. smegmatis</i> .....	111
Figure 4.20: Sequence alignment of 6938+6939.....	113
Figure 4.21: Conditional mutant system.....	117
Figure 4.22: The <i>parAB</i> genome region in mycobacteria.....	119
Figure 5.1: Graphical representation of LOPAC screens in LBT.....	135
Figure 5.2: Graphical representation of LOPAC screens in C <sup>-</sup> media.....	137
Figure 5.3: Graphical representation of LOPAC screens in N <sup>-</sup> media.....	139

Figure 5.4: Graphical representation of Spectrum Collection screens in LBT.....	142
Figure 5.5: Graphical representation of Spectrum Collection screens in C-media.....	144
Figure 5.6: Graphical representation of Spectrum Collection screens in N-media.....	145
Figure 5.7: Graphical representation of the NIH Diversity Set screens in LBT.....	146
Figure 5.8: Graphical representation of the NIH Diversity Set screens in C-media.....	149
Figure 5.9: Graphical representation of the NIH Diversity Set screens in N-media.....	151
Figure 6.1: Protein over-expression of ParA.....	164
Figure 6.2: Protein over-expression of ParB.....	165
Figure 6.3: Soluble fractions from different expression protocols.....	166
Figure 6.4: Purified protein.....	167
Figure 6.5: ParA ATPase activity.....	168
Figure 6.6: Effects of different parameters on ParA ATPase activity.....	169
Figure 6.7: Assay optimization for hit validation.....	170
Figure 6.8: Kinetic parameters of <i>M. tuberculosis</i> ParA ATPase.....	171
Figure 6.9: Effect of DMSO and vanadate on ATPase activity of ParA.....	172
Figure 6.10: Effect of NIH hits on the ParA ATPase activity.....	173
Figure 6.11: Effect on ATPase activity of ParA with varying concentrations NSC57850 and NSC70817.....	175
Figure 6.12: Effect on ATPase activity of ParA with phenoxybenzamine, fenoldopam, methiothepin and octoclotheptin.....	176
Figure 6.13: Molecular docking of ParA and its inhibitors.....	181
Figure 6.14: Walker A region of the ParA protein between different bacterial species.....	186
Figure 6.15: Structure of ATP and NSC57850.....	188
Figure 6.16: Structure of compounds tested in the ParA ATPase assays.....	191
Figure 9.1: Vectors used in this study.....	227
Figure 9.2: Cell morphology of cultures in LBT.....	239

Figure 9.3: Cell morphology of cultures in C <sup>-</sup> media.....	240
Figure 9.4: Cell morphology of cultures in N <sup>-</sup> media.....	241
Figure 9.5: Cell morphology of cultures in N <sup>-</sup> media with hits from LOPAC.....	242
Figure 9.6: Cell morphology of cultures in N <sup>-</sup> media with hits from Spectrum Collection.....	243
Figure 9.7: Cell morphology of cultures in C <sup>-</sup> media with hits from NIH Diversity Set.....	244

## List of Tables

Table 2.1: Bacterial strains used in this study.....	25
Table 2.2: Plasmids used in this study.....	26
Table 2.3: Antibiotics and/or chemicals used in this study.....	30
Table 2.4: Primers used in this study.....	32
Table 2.5: PCR cycles used for basic amplification.....	40
Table 2.6: Reagents and concentrations used with Taq DNA Polymerase.....	41
Table 2.7: Reagents and concentrations used with PCR Master Mix.....	41
Table 2.8: PCR annealing temperatures used per primer set.....	42
Table 2.9: Reagents and concentrations used for TA cloning.....	43
Table 2.10: Reagents and concentrations used for cohesive end ligation.....	43
Table 2.11: PCR cycles used for orientation screens.....	48
Table 2.12: Reagents and concentrations used for restriction digests .....	42
Table 2.13: Schedule for the production of ParA antibodies.....	56
Table 2.14: Concentration of reagents used in the ParA ATPase assays.....	58
Table 2.15: Compound library screens against antisense strains.....	61
Table 3.1: List of genes mutated from the 614 putative essential genes published by Sassetti <i>et. al</i> .....	72
Table 3.2: List of broad-spectrum targets in <i>M. tuberculosis</i> H37Rv.....	73
Table 3.3: Organisms in DEG.....	78
Table 4.1: Strains used for antisense assays.....	97
Table 5.1: Hits from LOPAC in LBT.....	136
Table 5.2: Hits from LOPAC in C <sup>-</sup> media.....	138
Table 5.3: Hits from LOPAC in N <sup>-</sup> media.....	140
Table 5.4: Hits from Spectrum Collection in LBT.....	143
Table 5.5: Hits from Spectrum Collection in C <sup>-</sup> media.....	144
Table 5.6: Hits from Spectrum Collection in N <sup>-</sup> media.....	145
Table 5.7: Hits from NIH Diversity Set in LBT media.....	147
Table 5.8: Hits from NIH Diversity Set in C <sup>-</sup> media.....	150
Table 5.9: Hits from NIH Diversity Set in N <sup>-</sup> media.....	151



Table 6.1: Kinetics parameters of ParA ATPase in the presence of inhibitors...	177
Table 6.2: MIC values of ParA inhibitors in mycobacterial cells.....	179
Table 6.3: SCC values in <i>M. smegmatis</i> under different culture conditions.....	179
Table 6.4: Cytotoxicity in HL-60 cells and mouse.....	179
Table 9.1: List of Essential Genes published by Sassetti <i>et. al</i> .....	228
Supplementary Table S1.....	230

## List of Abbreviations

ADP	adenosine diphosphate
Amp <sup>+</sup>	supplemented with ampicillin
AT	adenine-thymine
ATP	adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
<i>B. subtilis</i>	<i>Bacillus subtilis</i>
bp	base pairs
BSA	Bovine Serum Albumin
C <sup>-</sup>	Hartman-de Bonts carbon starvation media
CFA	Complete Freund's Adjuvant
CFU	colony forming units
DAPI	4',6-diamidino-2-phenylindole
ddH <sub>2</sub> O	double distilled water
DEG	Database of Essential Genes
DHFR	dihydrofolate reductase
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTPs	deoxynucleotide triphosphates
DOTS	Directly Observed Therapy, Short-course
dsDNA	double-stranded deoxyribonucleic acid
dsRNA	double-stranded ribonucleic acid
DTT	dithiothreitol
E	experimental
<i>E. coli</i>	<i>Escherichia coli</i>
eIF4H	eukaryotic translation initiation factor 4H
ETH	ethambutol
Fab	fatty acid biosynthesis
FCS	fetal calf serum
FDA	Food and Drug Administration
GC	guanine-cytosine
GFP	green fluorescent protein

GOLD	Genomes OnLine Database
HdeB	Hartman-de Bonts
HIV	human immunodeficiency virus
hrs	hours
Hyg <sup>+</sup>	supplemented with hygromycin
HTS	high-throughput screening
IC <sub>50</sub>	half maximal inhibitory concentration
ID	intradermally
IFA	Incomplete Freund's Adjuvant
INH	isoniazid
IPTG	isopropyl β-D-1-thiogalactopyranoside
Kan <sup>+</sup>	with kanamycin
kbp	kilo base pairs
K <sub>cat</sub>	catalytic constant
KEGG	Kyoto Encyclopedia of Genes and Genomes
K <sub>m</sub>	Michaelis constant
LB	Luria broth
LBT	Luria broth with tween 80
LD <sub>50</sub>	50% of the lethal dose
LOPAC	Library of Pharmacology Active Compounds
mAGP	mycolyl arabinogalactan-peptidoglycan
MCS	multiple cloning site
MIC	minimum inhibitory concentration
MOA	mode/mechanism of action
MDR-Tb	multi-drug resistant tuberculosis
<i>M. leprae</i>	<i>Mycobacterium leprae</i>
<i>M. smegmatis</i>	<i>Mycobacterium smegmatis</i>
<i>M. tuberculosis</i>	<i>Mycobacterium tuberculosis</i>
MTB	<i>Mycobacterium tuberculosis</i>
MTBC	<i>Mycobacterium tuberculosis</i> complex
mRNA	messenger ribonucleic acid
N <sup>-</sup>	Hartman-de Bonts nitrogen starvation media
NIH	National Institute of Health

NZ	New Zealand
O <sub>2</sub> <sup>-</sup>	Hartman-de Bonts oxygen starvation media
OADC	oleic acid, albumin, dextrose, catalase
OD	optical density
OD <sub>590</sub>	optical density at 590 nm
OD <sub>600</sub>	optical density at 600 nm
ORF	open reading frame
<i>P. aeruginosa</i>	<i>Pseudomonas aeruginosa</i>
PARB	poly(A) binding protein
PAS	<i>p</i> -aminosalicylic acid
PBS	phosphate buffered saline
PBST	phosphate buffered saline with tween 20
PCR	polymerase chain reaction
PDF	polypeptide deformylase
PVDF	polyvinylidene fluoride
P <sub>i</sub>	inorganic phosphate
<i>P. putida</i>	<i>Pseudomonas putida</i>
P-loop	phosphate binding loop
POA	pyrazinoic acid
PZA	pyrazinamide/pyrazinecarboxamide
PZAse	pyrazinamidase
QSAR	quantitative structure-activity relationship
RBS	ribosome binding site
RDD	rational drug design
RIF	rifampicin
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
RT	room temperature
SAR	structure-activity relationship
SBC	stationary bactericidal concentration
SC	subcutaneously
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis

sSNPs	synonymous single nucleotide polymorphism
<i>S. coelicolor</i>	<i>Streptomyces coelicolor</i>
<i>S. pneumonia</i>	<i>Streptococcus pneumonia</i>
SE	standard error
ssDNA	single-stranded deoxyribonucleic acid
TARGET	Tuberculosis Animal Research and Gene Evaluation Taskforce
TAE	Tris-acetate-EDTA
Tb	tuberculosis
Tc	tetracycline
TE	Tris-EDTA
TIGR	the Institute of Genomic Research
TraSH	Transposon Site Hybridization
tRNA	transfer RNA
UA	undetermined activity
<i>V. cholerae</i>	<i>Vibrio cholerae</i>
v/v	volume/volume
WHO	World Health Organisation
WT	wild type
w/v	weight/volume
XDT-Tb	extensively- drug resistant Tb
X-gal	5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside

## List of Publications

**Nisa S**, Blokpoel MCJ, Robertson BD, Tyndall J, O'Toole R (2010). Targeting the chromosome partitioning protein ParA in tuberculosis drug discovery. *Journal of Antimicrobial Chemotherapy*.

Miller CH, **Nisa S**, Dempsey S, Jack C, O'Toole R (2009). Modifying culture conditions in chemical library screening identifies alternative inhibitors of mycobacteria. *Antimicrobial Agents and Chemotherapy* 53(12): 5279-5283.

# **CHAPTER ONE**

## **General Introduction**

---

## 1.1 Preface

*Mycobacterium tuberculosis* (*M. tuberculosis*) is the fundamental etiologic agent for human tuberculosis and is the biggest killer due to bacterial infection in the world today [1]. This thesis attempted to identify the genes required by this organism to survive and validates one of these genes as being essential. It also characterizes its function and presents some potential inhibitors against this target.

This chapter commences with introducing *M. tuberculosis* as a human pathogen followed by the current treatment regimens for the multiple strains and their mechanisms of action. This is followed by a section on some of the drugs currently in the Tb pipeline. It also provides an in-depth review of the drug discovery process from a genomics point of view which is the focus of this thesis. Specific introduction relevant to the separate results chapters are included at the beginning of the respective chapters.

## 1.2 Global Burden of Tuberculosis

Tuberculosis, a disease which has afflicted humanity since ancient times still remains a major public health concern with an estimated one third of the world's population being infected with *M. tuberculosis* [2]. The earliest confirmation of the disease comes from skeletal evidence dating back to about 9,000 years from a Neolithic settlement in the Eastern Mediterranean [3]. Humans are the only known reservoir of this microbe [4]. According to the 2005 World Health Organisation (WHO) report, 8.8 million new cases and 1.75 million deaths occurred in 2003. These figures escalated to 9.24 million new cases and 1.7 million deaths in 2006 and 9.27 million new cases and 2 million deaths in 2007. The WHO has predicted that a failure to stop this epidemic will lead to about one billion people being newly infected, 200 million people developing Tb and at least 35 million deaths from the disease by 2020 [5-7].

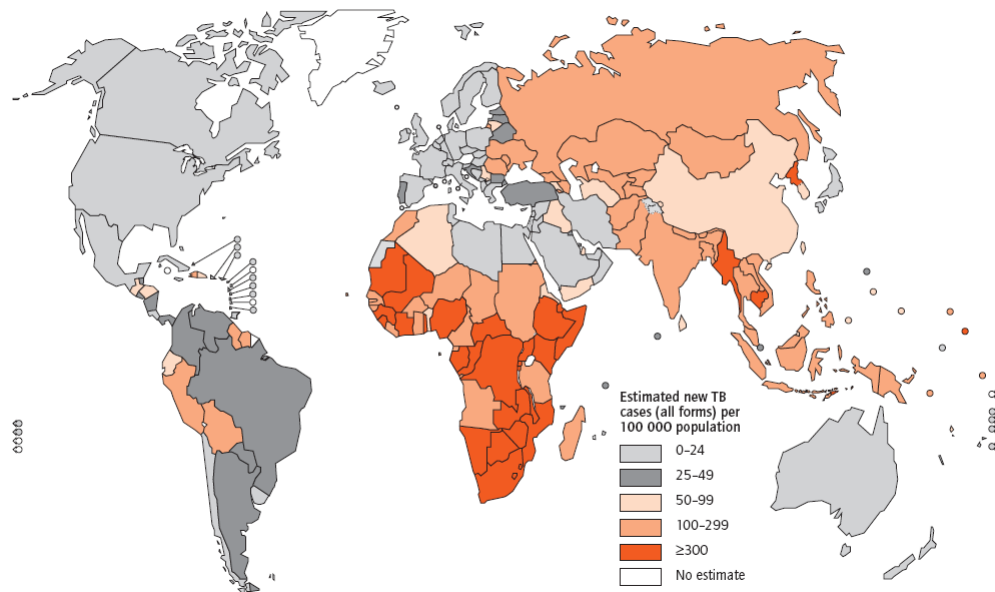
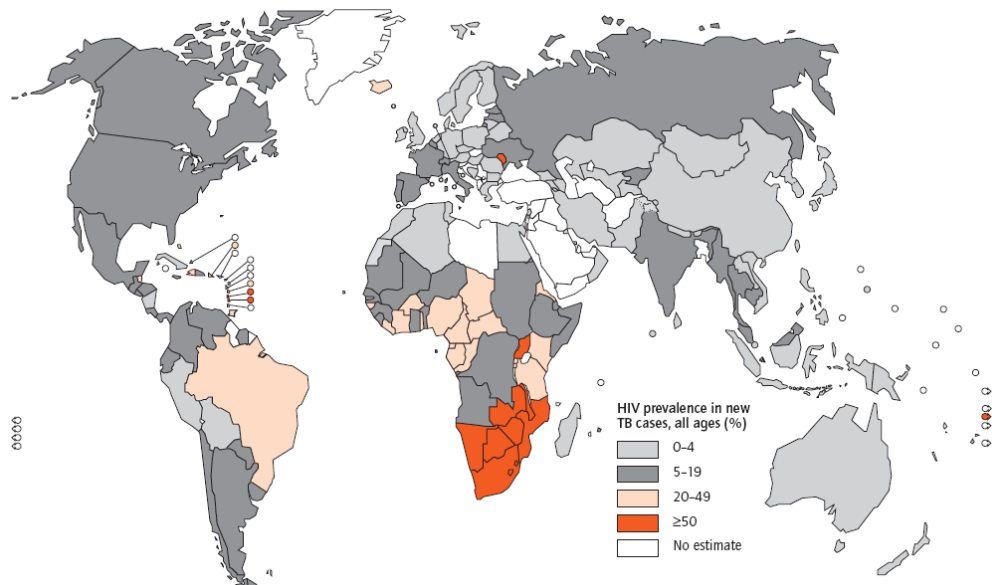


---

One of the major global burdens of Tb is due to human immunodeficiency virus (HIV). HIV can reactivate latent Tb infections (latency is discussed in section 1.4) and greatly hasten the progression of Tb infection into disease [8]. A person with a Tb infection has a 10% to 20% chance of developing disease in his lifetime; however, an HIV positive person co-infected with Tb has a greater than 10% chance of developing the disease within a year [8] and greater than 42% chance of death [9]. In 2007, nearly 14.8% of the new Tb cases and 23% of all deaths were in HIV positive people [7], an increase from the figures seen in 2006 [6]. The world's largest and fastest growing Tb epidemic is in Africa, infecting 79% of all HIV-positive cases (Figure 1.1 B) [7].

In 1993, the WHO declared Tb as a global emergency [10] and in 1994, launched the Directly Observed Therapy, Short-course (DOTS). DOTS is the most cost effective, highly efficient strategy for controlling Tb, whereby diagnosis and registration of every Tb patient is followed by supervised standardized treatment. Failure to follow the DOTS program stringently quickly raises the number of Tb cases. Evidence of this was seen in Eastern Europe. The socioeconomic burden placed by the dismantling of the Soviet Union weakened the public health system and thus the DOTS program causing a sudden increase in the Tb incidence rate [11].

In 2006, the Stop Tb Partnership launched the Global Plan to Stop Tb. This program aimed to halt or reverse the prevalence of Tb and death by 2015 and eliminate Tb by 2050 by ensuring effective access to diagnosis, treatment and cure for all Tb patients [12]. It is estimated that American, East Mediterranean and South-East Asian regions may achieve this goal of halving the prevalence of the disease and death by 2015 while the Western Pacific region may just be able to halve the prevalence of disease. The African and the European regions seem unlikely to be able to meet the target of halving either the prevalence or death at the current rate [7]. To achieve this goal, greater resources are needed to enhance and provide high quality DOTS program, reach those in need, especially the poor, improve health system and research for the development of new diagnostics, drugs and vaccines [7].

**A. Estimated numbers of new TB cases, 2008****B. Estimated HIV prevalence in new TB cases, 2008****Figure 1.1: Global burden of Tb**

**A)** Global burden of new Tb cases by country. There are 22 high burden countries with India, China, Indonesia, South Africa and Nigeria ranked among the highest. **B)** The global burden of Tb due to HIV epidemic, with the highest burden in South Africa. Reproduced with permission from WHO Report, 2009 [7].

### 1.2.1 Tuberculosis in New Zealand

From 1943 to 1988, the number of Tb cases in New Zealand (NZ) fell from 2,600 to 295 annually. However since then, there has been a resurgence of Tb with between 300 to 500 cases being reported per year [13]. From 1994 to 2004, NZ Tb rate has remained constant at about 10 per 100,000 people, slightly lower than the rate in United Kingdom (UK) but twice that of Australia, Canada and the United States of America (USA). This was distributed predominantly in the non-European and migrant population due to the increase in immigrants from high-risk countries, with 64.6% of patients being born overseas [14]. This increase in the incidence of Tb rates in the ethnic minorities is usually due to socio-economic pressures such as poverty and overcrowding causing a very high transmission rate from infected individuals. Fortunately, NZ is not a high-burden country as Tb rates here are not caused by HIV and/or multi-drug resistant Tb (section 1.5.2) [15].

### 1.3 *Mycobacterium tuberculosis*

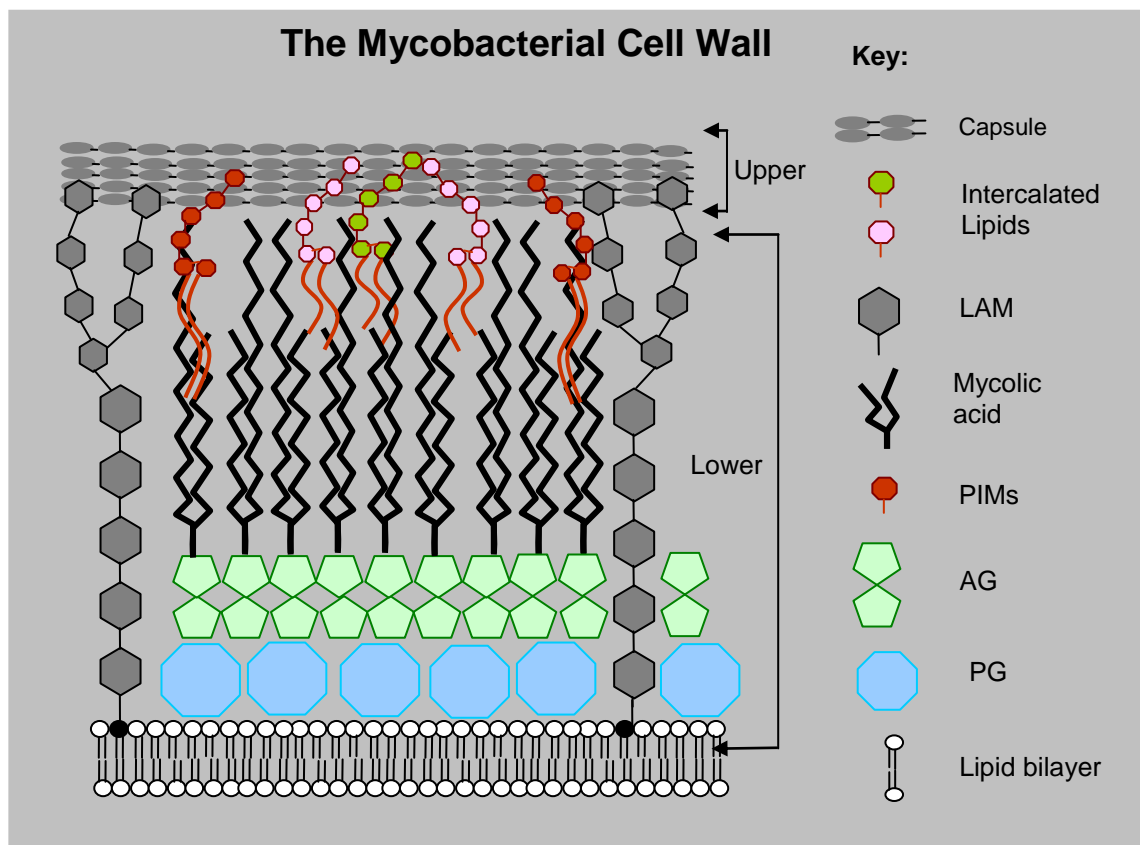
*M. tuberculosis* was first isolated by the German physician and scientist, Robert Koch in 1882 [16]. This microbe is an intracellular obligate aerobic actinomycete about 2-6  $\mu\text{m}$  long with a generation time of approximately 20-24 hours under optimal conditions [17, 18]. It forms tight, rope-like aggregates in liquid culture [19] and creamy irregular raised colonies with condensed filaments in the center in solid media [20] which can take up to 7 days or more to form [21].

The genus *Mycobacterium* can be divided into three groups; the slow-growing pathogenic mycobacteria that is the *M. tuberculosis* complex (MTBC); the slow-growing non-tuberculous mycobacteria and the fast growing non-tuberculous mycobacteria [22]. The MTBC consists of human pathogens: *Mycobacterium tuberculosis*, *Mycobacterium africanum*, *Mycobacterium canetti*, a rodent pathogen, *Mycobacterium microti*, and a wide host range pathogen:

---

*Mycobacterium bovis* [23]. Unlike most bacterial clonal complexes which have at least 1% synonymous single nucleotide polymorphisms (sSNPs), the MTBC exhibit between 0.01% to 0.03% sSNPs [24]. Although the members of the MTBC display such high genetic homogeneity, they exhibit very different epidemiology and host range thus differentiation between the members is necessary for correct treatment [25].

Another key difference that separates mycobacteria from other microbes is the presence of a highly complex cell wall consisting of a capsule, a core and an inner membrane (Figure 1.2). The capsule is made up of free lipids and mycolates such as the phosphatidylinositol mannosides (PIMs) and lipoarabinomannan (LAM). The core consists of peptidoglycan (PG) and mycolic acid connected by an arabinogalactan (AG) polysaccharide, collectively called the mycolyl arabinogalactan–peptidoglycan (mAGP) complex. And lastly the inner membrane which consists of a lipid bilayer [26]. The presence of such a large amount of lipid in the cell wall makes it acid-fast. This is evident when the cells are stained with Ziehl-Neelsen stain. They resist decolorization with acid alcohol and retain the primary carbol fuchsin stain giving it the characteristic red/pink colour.



**Figure 1.2: The mycobacterial cell wall**

The upper segment consists of the soluble capsule intercalated with lipids (phthiocerol dimycocerosate, cord factor/dimycolyltrehalose, sulfolipids and phosphatidylinositol mannosies) while the lower segment consists of the insoluble mAGP complex.

### 1.3.1 *Mycobacterium smegmatis*

*M. smegmatis*, first isolated from smegma [27], was discovered by Lustgarten in 1884 [28, 29]. This microbe is a fast growing opportunistic pathogen [30, 31] of the genus *Mycobacterium* (section 1.3). Like *M. tuberculosis*, it is a rod-shaped acid fast bacilli, however, its generation time is only approximately 2 to 3 hours [18] and it forms colonies within 3 to 4 days on solid media [32]. The commonly used laboratory strain, *M. smegmatis* mc<sup>2</sup>155, can also be efficiently transformed with foreign DNA and can undergo homologous recombination. *M. smegmatis* has also shown a 100% specificity to anti-Tb drugs and has the most similar profile to MDR-Tb compared to other fast growers [33]. For these reasons, in this thesis, *M.*

*smegmatis* has been used as a surrogate organism for studying mycobacterial biology [32-34] and the pathways that are conserved in *M. tuberculosis* at a much faster rate [35]. Significant results were verified in *M. tuberculosis*, the disease causing pathogen.

#### **1.4 *Mycobacterium tuberculosis* Pathogenesis**

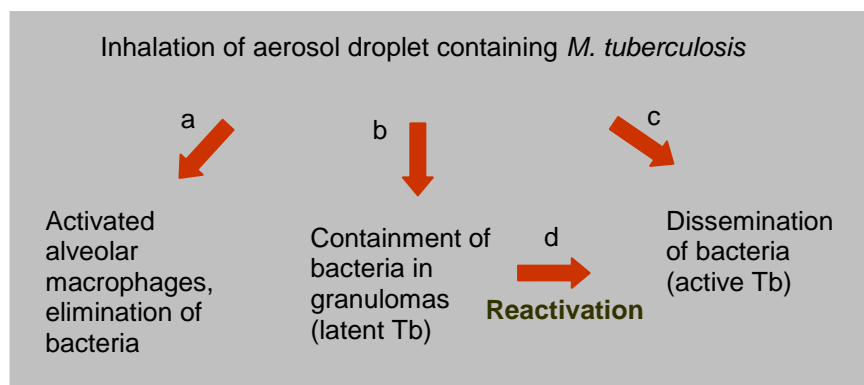
*M. tuberculosis* pathogenesis and its virulence factors are poorly understood and its intracellular location often confers protection from the immune system presenting a multifaceted interaction between the pathogen and the host immune response upon infection.

The most prevalent form of Tb is pulmonary tuberculosis [36] as the common route of infection is via inhalation of aerosol droplets containing the tubercle bacilli from an infected person [37]. Upon entering the lungs, the bacillus travels to the alveoli where it is phagocytosed by alveolar macrophages [38-40]. At this point, there can be several different outcomes of the infection (Figure 1.3) depending on the dose of bacteria upon infection and the immune status of the host [41]. The first scenario that can take place is the activation of the alveolar macrophages by T cells that have recognized mycobacterial antigens [39]. This causes the phagosomes within the macrophages to fuse with lysosomes [38] resulting in the elimination of the bacteria. However, the rate of this is not known and may be very minuscule [42] as complete eradication of the bacilli is very difficult to achieve.

The second scenario that can take place is a latent Tb infection. This occurs when the alveolar macrophage does not get activated and the bacilli reside within the phagosome. The bacteria change the intracellular environment of the phagosome by preventing its acidification, thus avoiding normal maturation and phagocytosis [38, 43]. The bacteria continue to divide in the macrophage until the macrophage lyses. This causes the macrophages in the surrounding tissue to

phagocytose the bacilli and this continues as a vicious cycle. It also causes the immune system to respond and try to eliminate the bacteria as the lysis of the macrophage causes mycobacterial antigen release. This battle between the immune system trying to annihilate the bacteria and the bacteria endeavoring to survive causes a caseating tubercle [38]. This can lead to the establishment of a granuloma. The granuloma is formed from the collection of infected macrophages, lymphocytes and fibroblasts developing a fibrous sheath and reducing the number of blood vessels penetrating it thus walling off the bacteria from the immune system. This is termed a latent Tb infection where there is no display of any clinical symptoms [4, 43, 44]. The host response can contain the latent Tb infection for life, however, any weakening of the immune system can lead to reactivation of Tb [45] thus greatly increasing the chance of clinical manifestations in HIV positive people (Figure 1.3).

The last scenario that can arise is infection followed by active disease. The bacillus multiplies in the lungs, spreads to the lymph nodes and can disseminate into the blood system to the rest of the body [41]. Dissemination of the bacilli in the blood vessel can cause disease manifestations on more distant organs like the kidneys, bones and brain.



**Figure 1.3: Pathogenesis of Tb infection and disease**

Inhalation of Tb bacilli can have 3 different outcomes. **a)** It can get eliminated by the immune system, **b)** it can be contained by the immune system in a latent form and **c)** it can lead directly to active disease or **d)** latent Tb gets reactivated by immune suppression later in life to cause active disease.

## 1.5 Tb Treatment

Streptomycin, discovered in 1944, was the first anti-tubercular drug to enter clinical use [46, 47]. Soon after this, other drugs used for the treatment of Tb were identified. These include *p*-aminosalicylic acid (PAS, 1946), isoniazid (INH, 1952), rifampicin (RIF, 1965), ethambutol (ETH, 1968) and pyrazinamide (PZA, 1970).

The current antibiotic regimen for the treatment of drug susceptible Tb consists of multi-drug therapy (MDT) of first-line drugs to reduce drug resistance and increase efficacy [48]. The treatment is divided into an initial 2 month intensive phase followed by a 4 month continuation phase. The initial therapy consists of RIF, INH and PZA together with an introductory phase of ETH and/or streptomycin [2, 47-50]. The continuation phase consists of RIF and INH, the two most potent first-line drugs used for the treatment of drug susceptible Tb.

### 1.5.1 Mode of Action (MOA) of first-line drugs

RIF targets the DNA-dependent RNA polymerase. It binds to the beta subunit, prevents transcription [51] and kills slowly metabolizing bacteria therefore, eliminates persistent bacilli [52]. INH inhibits the biosynthesis of mycolic acid, a major component of the mycobacterial cell wall (Figure 1.2) and is bactericidal against active bacteria in the pulmonary cavities thus, is crucial in early therapy [52]. Streptomycin targets the 30S ribosomal subunit and inhibits protein synthesis. It has no bactericidal activity *in vivo* but greatly limits bacterial growth therefore is given in the introductory phase [51]. ETH inhibits the arabinosyl transferases involved in cell-wall biosynthesis and PZA is used as bactericidal agent against *M. tuberculosis* but the exact functional mechanisms of PZA has not yet been elucidated [51]. PZA also works at a lower pH making it important in killing bacteria in the acidic caseous necrotic centre [52].



---

The six month treatment regimen (section 1.5) is a major breakthrough from the 18-24 month treatment in the early 1960's [47]. This therapy is administered via DOTS to ensure adherence and reduce the emergence of drug resistant Tb [48]. However, resource-limited countries with restricted health facilities and poor access to anti-tubercular drugs find the adherence to the 6 month regimen very difficult. This has given rise to multi-drug resistant Tb (MDR-Tb).

### **1.5.2 MDR-Tb**

MDR-Tb has been reported since the early days of MDT, and is thought to have developed due to ineffective drug therapies, inadequate drug supplies and failure to adhere to the treatment regimen [53]. MDR-Tb is described as a combined resistance to INH and RIF, the two main first-line anti-tubercular drugs [54, 55]. The WHO reported approximately 0.5 million cases of MDR-Tb in 2007 with 27 high-burden countries, 15 of which are from the Eastern European region, accounting for 85% of all cases. Of the 0.5 million MDR-Tb cases worldwide, 0.3 million were reported as new cases and 0.2 million among previously treated cases [7].

The WHO launched the DOTS-Plus project in 1999 for the management of MDR-Tb. MDR-Tb treatment is based on individual susceptibility thus requires reliable drug-susceptibility testing. The treatment for MDR-Tb is considerably longer than drug susceptible Tb and the drugs used have more side-effects and is less efficacious [54, 56, 57]. The second-line drugs used for therapy are aminoglycosides (kanamycin and amikacin), fluoroquinolones (ofloxacin and ciprofloxacin), polypeptides (capreomycin, viomycin and enviomycin), D-cycloserine and thionamides (ethionamide and prothionamide); with aminoglycosides and fluoroquinolones being the most potent of all the second-line drugs [58].

### 1.5.3 MOA of second-line drugs

Fluoroquinolones exert their effect by inhibiting DNA gyrase, which is required for DNA replication, repair and transcription. They also have the ability to penetrate macrophages and exert antimycobacterial activity intracellularly. Aminoglycosides inhibit protein synthesis by binding to the 30s ribosomal subunit as mentioned in section 1.5.1 for streptomycin, an aminoglycoside used as a first-line drug. The MOA of polypeptides have not yet been elucidated; however, they bind to the ribosome and prevent protein synthesis. D-cycloserine is an analog of D-alanine which inhibits alanine racemase, an essential enzyme for the synthesis of peptidoglycan that is required for cell wall biosynthesis [54].

Apart from the longer treatment regimen for MDR-Tb, it is also associated with a higher cost, making adherence to the treatment even more difficult than drug susceptible Tb. This has in turn given rise to extensively-drug resistant Tb (XDR-Tb).

### 1.5.4 XDR-Tb

In March 2006, the Centers for Disease Control and Prevention reported the emergence of XDR-Tb. XDR-Tb is resistant to the most potent first- and second-line drugs, that is, resistance to RIF, INH and 3 or more of the second-line drugs; one or more fluoroquinolones and one or more injectables (capreomycin and/or any aminoglycosides) [57]. Treatment for XDR-Tb is associated with higher toxicity and lower efficacy. The prevalence of XDR-Tb is unknown however 50 countries had reported XDR-Tb cases by mid 2008 with the largest number of cases reported in South Africa due to the HIV epidemic [59]. In 2006, a town in the province of Kwa-zulu Natal had 53 cases of XDR-Tb, 44 of which were tested for HIV and were HIV-positive and 52 died within 16 days after sputum sample collection [60]. XDR-Tb has become a major problem in the 21<sup>st</sup> century and there are insufficient drugs to target active XDR-Tb [61]. More recently, there

---

have been reports of super XDR-Tb, more commonly known as totally drug resistant Tb (TDR-Tb) [62].

### 1.5.5 Mechanisms of Drug Resistance

Drug resistance is a natural phenomenon and the mechanisms via which bacteria develop resistance are diverse and complex. *M. tuberculosis* resistance occurs due to its highly hydrophobic cell wall (Figure 1.2) that is impermeable to most drugs and the resistance determinants encoded by its genome [17, 52].

The resistance to RIF is due to mutations in the *rpoB* gene which is responsible for producing the beta subunit of the DNA-dependent RNA polymerase. However, there is a small fraction of resistant bacteria that does not exhibit this mutation suggesting there are other mechanisms of RIF resistance [51]. INH is a pro-drug that is activated by an enzyme called KatG and it inhibits an enzyme that is involved in fatty acid biosynthesis called InhA. Resistance to INH occurs due to mutations in the either *katG* or *inhA* genes [51]. Streptomycin resistance occurs due to mutations in the 16s rRNA genes *rrs* and *rpsL*, the gene encoding ribosomal protein S12 [63]. ETH resistance is acquired via the mutations in the *embB* gene that encodes for the arabinosyl transferases that is inhibited by ETH [64].

Resistance to fluoroquinolones is due to specific mutations in the *gyrA* and *gyrB* genes that encode for DNA gyrase A and B subunits respectively [65]. Aminoglycosides have the same mechanism of resistance as described for streptomycin. Since the MOA of polypeptides has not been characterized, the mechanism of resistance is not fully understood; however it also changes the 16s rRNA and thus may confer cross-resistance from streptomycin [51].

Despite the impact of global tuberculosis and the prevalence of drug resistance, no new classes of drugs have been approved for Tb treatment by the US Food and Drug Administration since 1972 [66]. There is an urgent need for

---

new and alternate drugs to target Tb with novel MOA to prevent cross-resistance from current anti-tubercular drugs [67].

## 1.6 Tb Drug Development

There are a few challenges impeding Tb drug development. Any new drug would have to offer an advantage over the current therapy [46]. Specifically, drugs that enable a shorter treatment regimen would be ideal. This would improve patient adherence, thereby, increasing the likelihood of elimination of the infection and decreasing the risk of the appearance of MDR strains [6]. A shorter drug course would also reduce the cost of treatment and therefore, enable the expansion of treatment programs to cover larger sections of a particular population. New drugs also need to be more effective and safe for people who are infected with resistant strains of *M. tuberculosis*. Any new drug should be able to treat latent Tb to help eradicate *M. tuberculosis* from the large number of people carrying the pathogen asymptomatically but who can potentially succumb to tuberculosis later in life [6]. And lastly, it should not elicit any drug-drug interactions, including with anti-retroviral drugs (ARVs) used for the treatment of HIV. There are currently a number of Tb drugs in the pipeline to try and achieve these goals.

### 1.6.1 Current Tb drugs in the pipeline

The TB alliance has accumulated the largest pipeline of potential new anti-tubercular drugs to try and prevent and reverse the number of Tb cases. One of the most promising of these drugs is moxifloxacin, a fluoroquinolone which has reached Phase III clinical trials. It targets DNA gyrase, an essential enzyme for bacterial survival [68, 69] and it has little interaction with cytochrome P450 enzymes that is involved in the metabolism of some of the ARVs [70]. Replacement of INH with moxifloxacin in the 6 month treatment regimen (section

1.5) has been shown to eliminate Tb infection in 4 months, decreasing the treatment regimen by 2 months [71].

Other drugs that have reached Phase II clinical trials are TMC207 and PA-824. TMC207 is a diarylquinoline with a novel MOA that specifically targets the bacterial ATP synthase [72], an enzyme essential for energy homeostasis [73, 74]. TMC207 has shown bactericidal activity against drug-susceptible, MDR-Tb and XDR-Tb with no cross-resistance to current first-line drugs [75-77]. It also eradicates dormant bacilli as effectively as replicating bacilli [74]. A combination of TMC207 with either the current first-line (section 1.5) or second-line (section 1.5.2) treatment regimen led to the successful eradication of infection within two months in drug-sensitive mouse model [78].

PA-824, a nitroimidazole with a novel mechanism of action [79], has shown activity in a number of MDR-Tb strains and has sterilizing effects *in vivo* [80]. In addition, it does not possess mutagenic features against genotoxicity studies, that is, non-deleterious effect on cellular genetics, and has no significant cytochrome P450 interactions [79]. A combination therapy of PA-824 with moxifloxacin and pyrazinamide cured the murine model of Tb faster than the current first-line regimen of rifampin, isoniazid, and pyrazinamide [81]. These drugs, if applicable to humans, can lead to the shortening of the current treatment regimen. The discovery of such novel antibacterial agents, their mechanism of action and their resistance determinants has been greatly accelerated by the advances in technology such as the availability of genomic data.

## 1.7 Genomics

Bacteria generally have a double-stranded circular DNA genome with sizes varying from 500 kilobase pairs (kbp) to 9,000 kbp [82, 83]. Current high-throughput sequencing technologies can sanction the assembly of a microbial genome within 24 hours [84]. The genome sequence is the starting point of most

---

genetic based scientific research and has many applications. The first and foremost advantage is gene prediction that is the identification of the genes contained in the genome by looking at the open reading frames (ORFs) [85]. However, sequence alone does not provide the functional annotation of a gene.

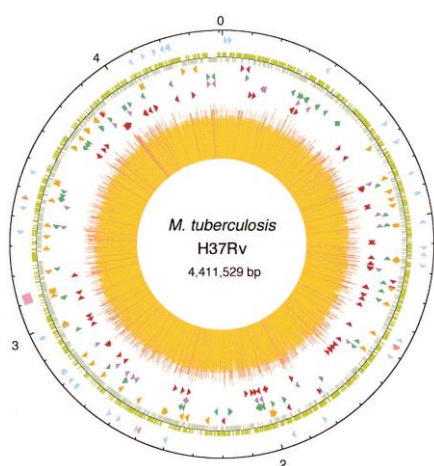
This leads to the second use of the genomic data that is genome annotation. The most common method used for genome annotation is via bioinformatics. Bioinformatics provides a means to study genes, proteins and their structure and function *in silico*. It also provides the foundation for rational design of experiments to test the *in silico* hypothesis. There are several different methodologies used for gene annotation. Comparative genomics uses sequence homology between genomes to explore the distribution, conservation and functions of genes based on their similarity [67], [86]. However, functional annotation is not always this simple as not all gene sequences are conserved [87]. These genes can be annotated via structural genomics as the structure of a protein is important for its function comprehension [88, 89]. Structural genomics can identify the catalytic and ligand-binding domains present in a protein. This can then be used to predict protein-protein and protein-nucleic acid interactions that can aid in functional annotation [90]. Nowadays, *ab initio* gene prediction algorithms are used to predict the protein-coding sequence *in silico* which is then verified by a combination of high-performance liquid chromatography and mass spectrometry [91].

Genome annotation can in turn lead to the identification of resistance determinants, virulence factors, new protein families and estimated metabolic aptitudes of an organism [92-94]. This can shed light on the genes that are critical for the survival of the bacteria. The most significant advantage of such genes in pathogenic bacteria is the potential to identify novel antimicrobial targets [95, 96]. State of the art second-generation genomics technologies are constantly being developed to exploit the genomic data for such studies [97].

### 1.7.1 The Mycobacterium Genome

The publication of the complete genome sequence of *M. tuberculosis* H37Rv has opened a whole window of opportunities for drug development. It is being used to identify, validate and exploit new targets to discover novel antimicrobials. The current anti-tubercular treatment regimen primarily targets cell wall biosynthesis, transcription and translation (section 1.5.1 and 1.5.3). The components of other essential mycobacterial processes have remained largely unexploited for anti-tubercular therapy mainly due to lack of information on the molecular mechanisms of these processes.

*M. tuberculosis* H37Rv is the most commonly used laboratory strain [98] due to its virulence in animal models, drug susceptibility and its ability to tolerate genetic manipulations. The *M. tuberculosis* H37Rv genome was first sequenced in 1998 and has since been revised [17, 99, 100]. It is a GC rich genome with 65.6% GC nucleotides and 34.4% AT nucleotides containing 4,411,529 base pairs [17]. This genome was reported to encode 4,057 genes, 4,007 encoding for proteins and 50 encoding RNA molecules [17]. This genomic data is regularly updated on the Tuberculist website (<http://genolist.pasteur.fr/TubercuList/help/current.html>) which currently reports 4,411,532 base pairs with 4,009 protein coding sequences.



**Figure 1.4: The *M. tuberculosis* H37Rv genome.** Reproduced from Cole *et al*, 1998 [17].

---

Since the sequencing of the *M. tuberculosis* H37Rv genome, about 52% of its genes have been assigned functions, most of which are based only on comparative genomics. Approximately 10% of the unknown genes are unique to *M. tuberculosis* [101]. There are approximately 250 genes for lipid metabolism as opposed to 50 in *Escherichia coli*. This is due to the complex cell wall structure of *M. tuberculosis* as discussed in section 1.3 [89, 102].

Other findings exclusive to the mycobacterial genome are the presence of repeat elements (IS6110 and IS1081) and regulatory genes such as two-component systems and eukaryotic-like serine/threonine kinases. About 10% of the genome is dedicated to the Pro-Glu (PE) and Pro-Pro-Glu (PPE) protein families with highly conserved N-terminus which currently have not been functionally annotated [17].

From the sequence data, it is predicted that *M. tuberculosis* is able to synthesize all its essential amino acids, vitamins and co-factors and has all the enzymes needed for respiration under aerobic and anaerobic conditions [17]. The generation and analysis of genomic data has led to the identification of a lot of potential drug targets.

### 1.7.2 Drug Targets: essential and conserved genes

The number of genes required to sustain cellular life under favorable conditions without external stress is known as the minimal gene set. The genes in the minimal gene set are deemed essential. Since essential genes are critical for the survival of an organism, they make excellent drug targets. Gene products essential for cellular life processes include those involved in transcription, translation, DNA replication, cell division, cell membrane and cell wall synthesis, DNA recombination and repair, bioenergetics, and macromolecule synthesis.

Some essential genes are also conserved among different bacteria. This signifies a selection pressure and may indicate their potential essentiality. A high



degree of gene conservation also suggests that mutations are not tolerated in these genes. Targeting such conserved genes will reduce the chances of developing resistance via genetic variability [67]. Genes that are essential and conserved in bacteria but not in eukaryotes are potential targets for broad-spectrum antibiotic development [103, 104]. Broad-spectrum antibiotics are clinically useful against different strains of the target organism [67] and important for economic and diagnostic reasons [105]. Genes that are specific to a small fraction of bacteria are potential targets for narrow-spectrum antibiotics [97]. By assessing multiple bacterial genomes, essential targets present in clinically relevant species can be identified based on their conservation. The identification of such targets has been greatly accelerated by *in silico* technologies and comparative genomics is often used to assess the spectrum of these targets between genomes [104-106].

However, identification of a conserved essential gene is not enough to deem it a prospective drug target. Interpretation of essential genes can change under different conditions [107] as microbes alter their gene expression to ensure its survival upon sensing a change in the environment [108] making certain genes only essential under certain conditions. For example, a gene essential *in vivo* under stress conditions may not be essential *in vitro* under optimal conditions. These are known as conditionally essential genes [107]. When selecting essential genes for drug targeting, it is important that the genes are required under all conditions. Thus the vast number of drug targets in any bacterial genomes requires a second level of precedence to select an apt target entailing intervention with molecular genetic techniques to validate potential targets [101].

### 1.7.3 Target Validation

The first step in targeted drug discovery is the identification and validation of drug targets. Target validation determines if the modulation of a target will lead to desirable clinical outcomes [109]. If the inhibition or disruption of a particular

---

target leads to non-viable cells, it is deemed an essential gene and thus, a valuable target.

There are a number of different methodologies employed for essentiality testing. One of the most widely used techniques is transposon mutagenesis. This involves the insertion of genetic elements containing a selectable marker and transposase that inserts randomly in the genome leading to the disruption of ORFs [67]. The advantage of this methodology is that it can be applied to whole genomes. The most successful of the techniques used is transposon site hybridization (TraSH) where pools of mutants are identified using DNA microarrays and genes that are unable to be mutated are deemed essential [110, 111]. This technique can also be applied to find conditionally essential genes (section 1.7.2) where mutants are recovered in some conditions but not all [107].

Other techniques used for establishing gene essentiality involve sequence directed mutagenesis which is especially useful for target validation as this can be used to characterize each target individually. These include plasmid insertion (gene disruption), allelic exchange (gene deletion; marked and unmarked) and directed evolution of target DNA (PCR based) [67]. Directed evolution is particularly useful to mutate genes in species that are naturally competent to uptake DNA. Genes in such species can be mutated *in vitro* using directed evolution and instantaneously introduced into the cells thus removing the cloning steps in creating mutants. However, the most fundamental experiments for target validations are target modulation studies. This involves conditional gene expression techniques like antisense expression (section 4.1.1.1) and microarray expression technologies [67].

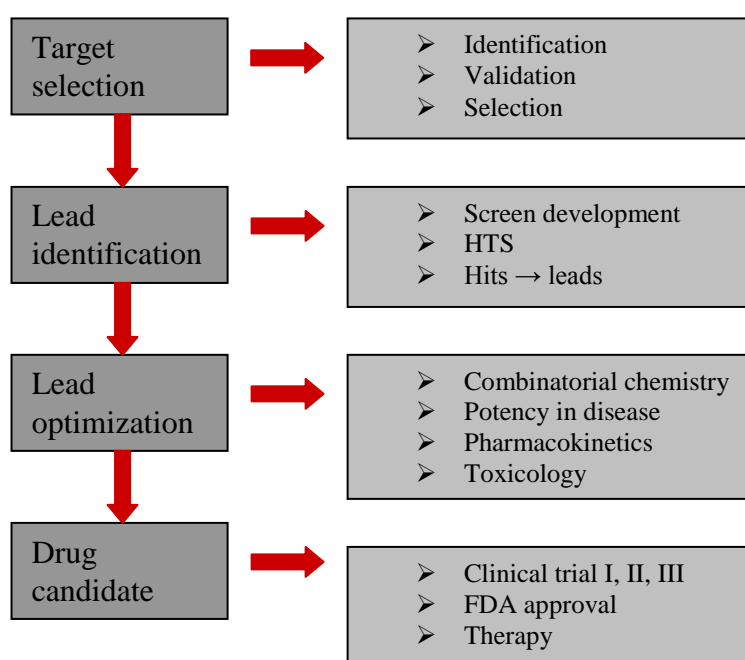
A key feature of drug discovery today is the ability to utilize high-throughput screening (HTS). This calls for target validation assays that are easy and reliable to accommodate in these HTS strategies.

## 1.8 Targeted Screening

Once a target is validated as being essential, it can be used for screening compounds which specifically inhibit its function. Such screening requires the development of an efficient *in vitro* HTS assay for the identification of hits. A hit in high throughput drug screening is defined as a drug that gives a positive result/desired phenotype that is being screened for in an assay. There are multiple assays for HTS. These include cell-free assays like functional and binding assays and whole-cell assays [104]. Cell-free functional assays require a known target and measure the biological activity whereas binding assays can be applied to an unknown target to elucidate the receptor, ligand or antibody binding [112]. Binding assays may also give an insight to the function of the unknown gene. Each assay has its own advantage but whole-cell assays are preferred as the permeability of the drug is ascertained [104] before resources are invested in exploring an inhibitor that may be impermeable [113]. However, compounds can affect cell growth in a variety of ways thus, a secondary assay is required to stipulate the specificity of the inhibitor to the target [97]. Specificity and biochemical information is achieved from cell-free based assays with an isolated target, making them vital post-screening assays [114].

More recently, there has been a merging of the targeted approach with the whole-cell based assay which can differentiate between specific and non-specific inhibition and enable the determination of the bactericidal activity of the drug and the druggability of the targets [105]. Such assays involve titrating the expression level of the target and identifying potential inhibitors based on the phenotype of the cell depending on the function/pathway that is inhibited [97]. These assays may provide some evidence of the function of unknown genes and additional information like the MOA of compounds [97]. Since these methods define the MOA, bioinformatics can be used to identify related targets in mammalian systems and the potential toxicity of the drug to mammalian cells which is crucial for anti-mycobacterial therapy, given the long treatment regimens. A targeted approach can also lead to rational drug design (RDD) using structure-activity relationship (SAR) [115].

In summary, there are many different aspects to be considered in the development of new anti-tubercular drugs (Figure 1.5). These features have to undergo numerous phases of validation and trials before a drug can enter clinical use making active research highly desirable to achieve a good pipeline of drugs as it is unlikely that one drug alone will conquer the disease. It is also inevitable that resistance will evolve over time due to selection pressure from any antibiotic thus requiring back-up drugs for the treatment of resistant Tb.



**Figure 1.5: Flow chart of target-based drug discovery**

Schematic representation of the processes involved and the outcomes achieved in target-based drug discovery. It includes the identification of a protein target whose inhibition is potentially going to result in the successful eradication of the microbe, an assay that will enable screening small-molecule compound libraries *in vitro* and *in vivo* for the identification of promising inhibitors in model systems and the determination whether the findings are extensible to humans; a complex and resource-extensive process.

## 1.9 Aims of this study

The purpose of this study was to identify inhibitory molecules against potential drug targets involved in cell division in the *M. tuberculosis* genome. The aims of this thesis to achieve the above mentioned objective were:

- A. To establish the number of conserved essential genes in the genome of *M. tuberculosis* H37Rv via bioinformatics and literature curation.
- B. To validate some of the conserved essential genes involved in cell division for their essentiality in the surrogate *M. smegmatis*.
- C. To run high-throughput library screens against various compound libraries with the validated target(s).
- D. To develop a biochemical assay of the target protein to test for the specificity of all significant hits from the high-throughput screens.

## **CHAPTER TWO**

### **Materials and Methods**

## 2.1 Bacterial Strains and Plasmids

### 2.1.1 Bacterial strains

The bacterial strains used in this study are presented in Table 2.1 together with their relevant genotype and phenotype. Stocks of all strains were made up in 20% glycerol and stored at -80°C.

**Table 2.1: Bacterial strains used in this study**

Strain	Relevant genotype and phenotype	Reference
<i>E. coli</i> DH5 $\alpha$	supE44 $\Delta$ lacU169 (F80 lacZ $\Delta$ M15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1	Promega, (Sydney, Australia)
<i>E. coli</i> LMG194	F- $\Delta$ lacX74 galE thi rpsL $\Delta$ phoA (Pvu II) $\Delta$ ara714 leu::Tn10 with Sm <sup>R</sup> , Tc <sup>R</sup>	Invitrogen, (Auckland, NZ)
<i>E. coli</i> BL21(DE3)	F <sup>-</sup> ompT gal dcm lon hsdS <sub>B</sub> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) $\lambda$ (DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])	Novagen, (Palmerston North, NZ)
<i>M. smegmatis</i> mc <sup>2</sup> 155	High-frequency transformation mutant of <i>M. smegmatis</i> ATCC 607	[116]
<i>M. bovis</i> BCG	Strain Pasteur 1173P2	[117]
<i>M. tuberculosis</i> H37Rv	Laboratory strain ATCC 27294	[118]
<i>M. tuberculosis</i> H37Ra	Strain ATCC 25177	[119]

### 2.1.2 Plasmids

All plasmids utilized in this study are listed in Table 2.2 together with their relevant characteristics. The plasmid maps of pCR<sup>®</sup>2.1, pBAD/HisB, pET-28a(+) and pMind are provided in Appendix 9.1.

**Table 2.2: Plasmids used in this study**

Plasmid	Characteristics	Reference
pCR <sup>®</sup> 2.1	TA cloning vector for direct ligation of PCR products, pUC origin, T7 promoter/primer site, LacZ $\alpha$ fragment, M13F and M13R priming site, Amp <sup>r</sup> , Kan <sup>r</sup> , 3.9 kb	Invitrogen
pBAD/HisB	Protein over-expression vector, pBR322 origin, <i>araBAD</i> promoter, <i>rrnB</i> transcription termination, <i>araC</i> gene, polyhistidine tag, MCS, Amp <sup>r</sup> , 4.1 kb	Invitrogen
pET-28a(+)	Protein over-expression vector, pBR322 origin, T7 promoter, T7 transcriptional start and terminator, His-Tag <sup>®</sup> , MCS, Kan <sup>r</sup> , 5.369 kb	Novagen
pMind	Conditional mycobacterium expression vector, CloliE1 origin, pAL5000 origin, inducible P <sub>tet</sub> , MCS, Kan <sup>r</sup> , Hyg <sup>r</sup> , 6.849 kb	[120]
pHS201	Mycobacterial- <i>E. coli</i> shuttle vector with a high-copy-number version of the pAL5000 mycobacterial replicon, BCG hsp60 promoter, <i>gfp</i> with a synthetic Shine-Delgarno sequence, Kan <sup>r</sup> , 5.813 kb	[121]
pYL GFP	BCG- <i>E. coli</i> shuttle vector, two constitutively expressed promoters (BCG hsp60 and hsp70), two MCS, Kan <sup>r</sup>	[122]
pSE100	Mycobacterial- <i>E. coli</i> shuttle vector, containing origins of replication for <i>E. coli</i> (pMB1) and mycobacteria (pAL5000) and a MCS flanked by two transcriptional terminators, contains P <sub>myc</sub> tetO, Hyg <sup>r</sup> , 5.538 kb	[123]
pMSC1s	Single-copy-integrating plasmid which inserts into L5 <i>attB</i> , <i>tetR</i> under P <sub>myc</sub> , Kan <sup>r</sup>	[123]
pSHIGH+hsp60	Mycobacterial- <i>E. coli</i> shuttle vector with a high-copy-number version of the pAL5000 mycobacterial replicon, BCG hsp60 promoter, <i>gfpmut2</i> with a synthetic Shine-Delgarno sequence, Kan <sup>r</sup> , 5.6 kb	[124]
pCR <sup>®</sup> 2.1_GFP	pCR <sup>®</sup> 2.1 vector cloned with a 0.750 kb PCR fragment encoding the <i>gfp</i> from pHS201, 4.65 kb	This work
pCR <sup>®</sup> 2.1_2357	pCR <sup>®</sup> 2.1 vector cloned with a 1.209 kb PCR fragment encoding the MSMEG_2357 gene from <i>M. smegmatis</i> ,	This work



---

	5.109 kb	
pCR <sup>®</sup> 2.1_ogb	pCR <sup>®</sup> 2.1 vector cloned with a 1.458 kb PCR fragment encoding the MSMEG_4623 gene from <i>M. smegmatis</i> , 5.358 kb	This work
pCR <sup>®</sup> 2.1_parB	pCR <sup>®</sup> 2.1 vector cloned with a 1.056 kb PCR fragment encoding the MSMEG_6938 gene from <i>M. smegmatis</i> , 4.956 kb	This work
pCR <sup>®</sup> 2.1_parA	pCR <sup>®</sup> 2.1 vector cloned with a 0.972 kb PCR fragment encoding the MSMEG_6939 gene from <i>M. smegmatis</i> , 4.872 kb	This work
pCR <sup>®</sup> 2.1_parA+parB	pCR <sup>®</sup> 2.1 vector cloned with a 2.087 kb PCR fragment encoding the MSMEG_6938 and MSMEG_6939 gene from <i>M. smegmatis</i> , 5.987 kb	This work
pCR <sup>®</sup> 2.1_parA_600	pCR <sup>®</sup> 2.1 vector cloned with the first 0.6 kb PCR fragment encoding the MSMEG_6939 gene from <i>M. smegmatis</i> , 4.5 kb	
pCR <sup>®</sup> 2.1_ParA	pCR <sup>®</sup> 2.1 vector cloned with a 1.044 kb PCR fragment encoding the Rv3918c gene from <i>M. tuberculosis</i> H37Rv, 4.944 kb	This work
pCR <sup>®</sup> 2.1_ParB	pCR <sup>®</sup> 2.1 vector cloned with a 1.035 kb PCR fragment encoding the Rv3917c gene from <i>M. tuberculosis</i> H37Rv, 4.935 kb	This work
pSE100_parA_600	Derivative of pSE100 with the first 0.6 kb PCR fragment encoding MSMEG_6939 gene from <i>M. smegmatis</i> cloned in the <i>SphI/NotI</i> sites	This work
pMindGFP	Derivative of pMind with a 0.750 kb PCR fragment encoding the <i>gfp</i> from pHS201 cloned in the <i>SpeI</i> site, 7.598 kb	This work
pMind2357	Derivative of pMind with a 1.209 kb PCR fragment encoding the MSMEG_2357 gene from <i>M. smegmatis</i> cloned in the antisense orientation in the <i>PacI</i> site, 8.058 kb	This work
pMind_ogb	Derivative of pMind with a 1.458 kb PCR fragment encoding the MSMEG_4623 gene from <i>M. smegmatis</i> cloned in the antisense orientation in the <i>NdeI/SpeI</i> sites, 8.307 kb	This work
pMind_parB_GFP	Derivative of pMindGFP with a 1.056 kb PCR fragment encoding the MSMEG_6938 gene from <i>M. smegmatis</i> cloned in the antisense orientation in the <i>NdeI/SpeI</i> sites, 8.654 kb	This work
pMind_parA_GFP	Derivative of pMindGFP with a 0.972 kb PCR fragment encoding the MSMEG_6939 gene from <i>M. smegmatis</i> cloned in the antisense orientation in the <i>EcoRV</i> site, 8.57	This work

---

---

	kb	
pMind_ <i>parA</i> + <i>parB</i> _GFP	Derivative of pMindGFP with a 2.087 kb PCR fragment encoding the MSMEG_6938 and MSMEG_6939 gene from <i>M. smegmatis</i> cloned in the antisense orientation in the <i>NdeI/SpeI</i> sites, 9.685 kb	This work
pBAD/HisB_ParA	Derivative of pBAD/HisB with a 1.044 kb PCR fragment encoding the Rv3918c gene from <i>M. tuberculosis</i> H37Rv cloned in the <i>BglII/HindIII</i> sites, 5.144 kb	This work
pBAB/HisB_ParB	Derivative of pBAD/HisB with a 1.035 kb PCR fragment encoding the Rv3917c gene from <i>M. tuberculosis</i> H37Rv cloned in the <i>BglII/HindIII</i> sites, 5.135 kb	This work
pET-28a(+)_ParA	Derivative of pET-28a(+) with a 1.044 kb PCR fragment encoding the Rv3918c gene from <i>M. tuberculosis</i> H37Rv cloned in the <i>BamHI/HindIII</i> sites, 6.413 kb	This work
pET-28a(+)_ParB	Derivative of pET-28a(+) with a 1.035 kb PCR fragment encoding the Rv3917c gene from <i>M. tuberculosis</i> H37Rv cloned in the <i>BamHI/HindIII</i> sites, 6.404 kb	This work

---

## 2.2 Media

### 2.2.1 *E. coli*

All *E. coli* strains were cultured in Luria-Bertani broth (LB broth) and 1.5% Luria-Bertani agar (LB agar) as per the manufacturers' instructions with media supplements as described in section 2.2.4 and culture conditions as described in section 2.3. All media was sterilized via autoclaving at 121°C for 30 minutes in a bench-top autoclave on a liquid cycle.

### 2.2.2 *M. smegmatis*

All *M. smegmatis* strains were cultured in LB broth, 1.5% LB agar or Hartman-de Bonts (HdeB) minimal media (Appendix 9.5.1) [125] with media supplements as described in section 2.2.4 and culture conditions as described in section 2.3.

### 2.2.3 *M. tuberculosis* H37Ra and *M. bovis* BCG

All *M. tuberculosis* H37Ra and *M. bovis* BCG strains were cultured in Middlebrook 7H9 broth and Middlebrook 7H10 agar. All media was supplemented with 0.1% (v/v) Tween 80 and 10% (v/v) oleic acid + albumin + dextrose + catalase enrichment (OADC; Fort Richard Laboratories) and cultured in conditions as described in section 2.3.

### 2.2.4 Media supplements

Tween 80: Stock solutions of Tween 80 were made to 10% with distilled and deionized water (ddH<sub>2</sub>O) and autoclaved at 121°C for 30 minutes in a bench-top autoclave on a liquid cycle. All *M. smegmatis* culture in rich media had a final concentration of 0.1% Tween 80 (LBT).

Glycerol: Stocks of glycerol were made up to 10%, 50% and 80% with ddH<sub>2</sub>O and autoclaved at 121°C for 30 minutes in a bench-top autoclave on a liquid cycle.

Antibiotics: All antibiotic stocks were made with ddH<sub>2</sub>O and filter-sterilized with a 0.22 µm Millipore filter and kept at -20°C except hygromycin solution which was bought at 50 mg/mL and stored at -4°C. The final concentrations of all the antibiotics used are listed in Table 2.3 unless otherwise stated.

Chemicals: Isopropyl-β-D-thiogalactopyranoside (IPTG) and L-arabinose stocks made with ddH<sub>2</sub>O were filtered-sterilized with a 0.22 µm Millipore filter. 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) stocks were made with 100% DMSO. All stocks were kept at -20°C. The final concentrations of all chemicals used are tabulated in Table 2.3 unless otherwise stated.

**Table 2.3: Antibiotics and/or chemicals used in this study**

<b>Antibiotics/ Chemicals</b>	<b>Stocks</b>	<b><i>E. coli</i></b>	<b><i>M. smegmatis</i></b>
Ampicillin (Amp)	100 mg/mL	200 µg/mL	-
Kanamycin (Kan)	50 mg/mL	50 µg/mL	50 µg/mL
Hygromycin (Hyg)	50 mg/mL	50 µg/mL	50 µg/mL
Tetracycline (Tc)	10 mg/mL	0.02 µg/mL	0.02 µg/mL
IPTG	100 mg/mL	100 µg/mL	-
X-gal	40 mg/mL	40 µg/mL	-

## 2.3 Culture Conditions

All bacterial strains were cultured (solid and liquid) at 37°C unless otherwise stated. All agar plates were incubated in a stationary plate incubator and all liquid cultures in flasks, polypropylene tubes and 96 well plates (Raylab) were aerated with orbital shaking at 200 rpm. All flasks were sterilized by autoclaving at 138°C on a heat/dry cycle with a cotton-wool bung covered with aluminum foil which was used as tops during incubations. All 96 well plates were sealed with plate seals (Interlab), wrapped in cling film and placed in a plastic bag to prevent evaporation from the wells during incubation [124].

## 2.4 Chemicals and Reagents

All chemicals and/or compounds were purchased from Sigma-Aldrich unless otherwise stated. IPTG, Xgal, dNTPs, DNA ladder, DNA loading buffer, DNA Molecular Marker and Protein Molecular Marker were procured from Bioline (Auckland, NZ).

## 2.5 Enzymes

All restriction endonucleases were purchased from Biolab (Auckland, NZ). T4 DNA ligase was procured from Invitrogen and Bioline, Taq DNA polymerase from Qiagen and PCR Master Mix from Bioline. Proteinase K, lysozyme, RNase A and DNase I was obtained from Sigma.

## 2.6 Oligonucleotides

Oligonucleotide primers were analyzed for restriction sites and designed using the software package Vector NTI Advance 10 (Invitrogen) and were synthesized by Invitrogen. The sequences of the primers used in this study are shown in Table 2.4. Each primer was resuspended in ddH<sub>2</sub>O as a stock solution at 100  $\mu$ M and working stocks were made from these at 10  $\mu$ M with ddH<sub>2</sub>O.

**Table 2.4: Primers used in this study**

<b>Primer Name</b>	<b>Sequence (5'- 3')</b>
<b>GFP</b>	
GFP_RBS_F	CTAG <u>ACTAGT</u> <b>TTTAAGAAGGAGATATACATATGAG</b> TAAAGGAGGA
GFP_R	CTAG <u>ACTAGT</u> TTTATTATTTGTATAGTTCATCCATGCC
<b>Genes from <i>M. smegmatis</i> mc<sup>2</sup>155</b>	
MSMEG_4623_F	CTAG <u>ACTAGT</u> ATGCCCCGGTTTGTGCGACCG
MSMEG_4623_R	CTAGCATATGTCACGAGTCCTCTCCTCCCG
MSMEG_6938_F	CTAG <u>ACTAGT</u> ATGAATCAGCCGGCACGCAA
MSMEG_6938_R	CTAGCATATGTTACTCGTTCTGGGCGCTCA
MSMEG_6939_F	CTAG <u>GATATC</u> ATGGGTCGGGTCAGAACAA
MSMEG_6939_R	CTAG <u>GATATC</u> CTACTGCTGGCGCGGCGGCG
MSMEG_6939_600F	CTAG <u>G</u> CATGCATGGGTCGGGTCAGAACAA
MSMEG_6939_600R	CTAG <u>GCGGCCG</u> CGTTGATGGTGAGCAGTCCGA
<b>Genes from <i>M. tuberculosis</i> H37Rv</b>	
ParA_F	CTAG <u>AGATCT</u> GTGAGTGCTCCGTGGGGCCC
ParA_R	CTAG <u>AAGCTT</u> TTCATGGTCGTCCCTTCGCGG
ParB_F	CTAG <u>AGATCT</u> ATGACCCAGCCGTCACGCAG
ParB_R	CTAG <u>AAGCTT</u> TTTACAGAGCGTCCCTGTGCA
<b>DNA from <i>M. tuberculosis</i> H37Rv</b>	
ParS_H37Rv_F	AGATCGATCGTCGCCGGGAT
ParS_H37Rv_R	GTCTCGAGAGCGGAGAATGT
<b>Sequencing primers</b>	
M13F	CCCAGTCACGACGTTGTAAAACG
M13R	AGCGGATAACAATTTACACACAGG
TetROF11	CTTCCGGTGGTGAGTCATAG
HygR11	AAGTAACAGGGATTCTTGTGTCAC
P118	ACGATCCGCGACGTGATGAT

All 5' restriction sites have been underlined and the ribosome binding site on GFP\_RBS\_F primer is in bold. The forward and reverse primers are denoted by an F or an R respectively at the end of the primer name.

## 2.7 Bioinformatics

### 2.7.1 Quantification of essential genes in *M. tuberculosis* H37Rv

Sasseti, *et al.* identified 614 putative essential genes (Appendix 9.2.1) in *M. tuberculosis* H37Rv under optimal growth conditions via high density transposon mutagenesis [126]. These genes were searched for in the literature and in the Tuberculosis Animal Research and Gene Evaluation Taskforce (TARGET) database (<http://webhost.nts.jhu.edu/target/>) at Johns Hopkins University. These comprehensive literature searches identified any genes that had subsequently been mutated and/or deemed non-essential. The genes that met this criterion were deleted from the list of 614 essential genes published by Sasseti to generate a new set of putative essential genes in *M. tuberculosis* H37Rv under optimal conditions.

### 2.7.2 Quantification of conserved essential genes in *M. tuberculosis* H37Rv

The translated sequences of all the genes identified in section 2.7.1 were obtained from the National Centre for Biotechnology Information (NCBI) protein database (<http://www.ncbi.nlm.nih.gov/sites/entrez?db=protein>) and submitted to the Database of Essential Genes (DEG; <http://tubic.tju.edu.cn/deg/>) for BLAST alignment to analyse if orthologs of these proteins were essential in other bacterial species. The presences of an orthologs in the DEG was determined using cut-off parameters of 20% identity over the entire length of the query protein and an E value of  $1e^{-20}$  or lower to obtain a list of conserved essential genes in *M. tuberculosis* H37Rv under optimal conditions.

### 2.7.3 Orthologs of conserved essential genes in *M. smegmatis* mc<sup>2</sup>155

Genes identified in section 2.7.2 were analysed by means of BLAST alignment at the amino acid level with the *M. smegmatis* mc<sup>2</sup>155 genome available at the

---

Institute for Genomic Research (TIGR; <http://tigrblast.tigr.org/cmr-blast/>) to determine if orthologs of the *M. tuberculosis* H37Rv proteins existed in the *M. smegmatis* genome.

#### **2.7.4 Homology model of the *M. tuberculosis* ParA**

A homology model of *M. tuberculosis* ParA (NCBI Reference Sequence: NP\_218434.2) was generated using modeler 9v7. This sequence was used as a BLAST search query of the PDB. Of the multiple hits obtained, the pdb2BEJ gave the lowest E value. Sequence alignment was carried out via the T-coffee (v7.38) webserver ([http://www.igs.cnrs-mrs.fr/Tcoffee/tcoffee\\_cgi/index.cgi](http://www.igs.cnrs-mrs.fr/Tcoffee/tcoffee_cgi/index.cgi)). Five models (residues G82-P341; magnesium ion and 4 coordinated water molecules) were generated using the alignment and the model with the lowest objective function was chosen for docking studies.

#### **2.7.5 Ligand docking**

All ligands (and isomers) were generated in SYBYL-X using the sketcher module. The ligands were energy minimized to a low energy starting conformation using the Tripos forcefield. Flexible ligand docking was carried out using Gold 4.1 using chemscore. Threonine 100 was designated as the centre of the ligand binding site with a radius of 12 Å. The metal ion was designated as octahedral.



---

## 2.8 DNA Isolation, Purification and Quantification

### 2.8.1 Purification of mycobacterial chromosomal DNA

Two methods were used for the isolation of *M. smegmatis* mc<sup>2</sup>155 and *M. tuberculosis* H37Rv chromosomal DNA: phenol/chloroform extraction (section 2.8.1.1) and ZR Fungal/Bacterial DNA Kit<sup>TM</sup> (Zymo Research- Nelson NZ, section 2.8.1.2).

#### 2.8.1.1 Phenol/chloroform extraction

This method was adapted and modified from Mycobacteria Protocols [127]. *M. smegmatis* mc<sup>2</sup>155 frozen stocks were streaked on LB agar plates and incubated (section 2.3) for 72 hours. After 72 hours, a loopful of colonies were inoculated into 3 mL LBT in 15 mL polypropylene tubes and cultured overnight (section 2.3). Approximately 1.5 mL of the overnight culture was pelleted by centrifugation at 13,000 x g for 5 minutes and frozen at -80°C for 4 hours (all subsequent centrifugation steps were carried out at 13,000 x g). After 4 hours, the pellet was thawed and resuspended in 400 µL TE buffer (Appendix 9.5.4) by vigorous vortexing. A 0.02 volume of 100 mg/ml lysozyme was added to the cells and incubated in a water bath at 37°C for 30 minutes. After incubation, a 0.1 volume of 10% SDS (w/v) and a 0.02 volume of 10 mg/mL proteinase K (Appendix 9.5.3) was added to the cell lysate and incubated in a 65°C water bath for 15 minutes. After 15 minutes, 400 µL of TE-buffered phenol was added to the lysate, mixed gently and centrifuged for 10 minutes. The residual phenol was removed from the upper aqueous layer by the addition of 400 µL of chloroform; which was mixed gently and again centrifuged for 10 minutes. A 2.5 volume of cold ethanol and 0.1 M sodium acetate (pH 5.2) were added to the aqueous layer in a fresh microfuge tube, mixed gently and centrifuged for 20 minutes to precipitate the DNA. The supernatant was removed and 70% cold ethanol was

added to the pellet and centrifuged for 5 minutes. The supernatant was removed again and the DNA was dried on a 37°C heating block, resuspended in 200 µL TE buffer and stored at -20°C.

#### **2.8.1.2 ZR fungal/bacterial DNA Kit™**

*M. smegmatis* mc<sup>2</sup>155 was cultured and pelleted (as described in section 2.8.1.1), to 50-100 mg (wet weight), washed once in 1x PBS (Appendix 9.5.4), resuspended in 750 µL Lysis Solution and added to ZR BashingBead™ Lysis tube. This was vortexed vigorously for 10 minutes and centrifuged at 13,000 x g for 1 minute. Approximately 400 µL of supernatant was transferred to Zymo-Spin™ IV Spin Filter collection tube and centrifuged at 7,000 x g for 1 minute. About 1,200 µL of Fungal/Bacterial DNA Binding buffer was added to the filtrate and this was transferred to Zymo-Spin™ IIC column in a collection tube and centrifuged at 10,000 x g for 1 minute (this step was carried out twice with 800 µL at a time and all subsequent centrifugation steps were carried out at 10,000 x g). The flow through was discarded and 200 µL of DNA Pre-Wash buffer was added to the Zymo-Spin™ IIC column and centrifuged for 1 minute. 500 µL Fungal/Bacterial DNA Wash buffer was added to the Zymo-Spin™ IIC column and again centrifuged for 1 minute. The Zymo-Spin™ IIC column was transferred to a fresh 1.5 mL microfuge tube and 100 µL of DNA Elution buffer was added to the column matrix then centrifuged for 30 seconds. The eluted DNA was stored at -20°C.

#### **2.8.2 Purification of plasmid DNA**

Two methods were used for the isolation of plasmid DNA in this study; phenol/chloroform extraction (section 2.8.2.1) and Zyppy™ Plasmid Miniprep Kit (Zymo Research, section 2.8.2.2).

### **2.8.2.1 Phenol/chloroform extraction**

Single colonies were picked from LB agar plates and inoculated in 2 mL LB broth in 15 mL polypropylene tubes with appropriate antibiotics (Table 2.3) and cultured overnight (section 2.3). A 1.5 mL of the overnight cultures were transferred to microfuge tubes and centrifuged at 13, 000 x g for 30 seconds (all subsequent centrifugation steps were carried out at 13, 000 x g). The pellets were resuspended in 100 µL cold Alkaline Lysis Solution I (Appendix 9.5.3) by vigorous vortexing. Approximately 200 µL of freshly prepared Alkaline Lysis Solution II (Appendix 9.5.3) was added to the bacterial suspensions and mixed via inverting the tube. 150 µL of cold Alkaline Lysis Solution III (Appendix 9.5.3) was added to the lysate, mixed via inverting the tubes, incubated on ice for 3 minutes and centrifuged for 5 minutes. One volume of phenol: chloroform (1:1) was added to the supernatant, mixed via vortexing then centrifuged for 2 minutes. This step was repeated with chloroform. The aqueous layer was transferred to fresh microfuge tubes and 2.5 volumes of cold 99% ethanol was added to these. The solution was mixed via vortexing and allowed to stand at room temperature for 2 minutes before it was centrifuged for 5 minutes. The supernatant was removed and all drops of fluid were removed by inverting the tubes on a Kimwipe. About 200 µL of cold 70% ethanol was added to the pellets, inverted several times and centrifuged for 2 minutes. The supernatant was removed and the tubes were stored at room temperature until all the ethanol evaporated. The nucleic acid was resuspended in 50 µL TE buffer (Appendix 9.5.4) and stored at -20°C.

### **2.8.2.2 Zyppy™ Plasmid Miniprep Kit**

Approximately 100 µL of 7x Lysis buffer was added to 600 µL of overnight cultures (as described in section 2.8.2.1) in 1.5 mL microfuge tubes and mixed via inverting the tubes several times. About 350 µL cold Neutralization buffer (with 100 µg/mL RNase A) was added to the lysate, mixed via inverting the tubes several times and centrifuged at 13, 000 x g for 2 minutes (all subsequent

centrifugation steps were carried out at 13, 000 x g for 30 seconds). The supernatant was transferred to Zymo-Spin<sup>™</sup> II columns in collection tubes and centrifuged, washed with 200 µL Endo-Wash buffer, centrifuged again and re-washed with 400 µL Zyppy<sup>™</sup> Wash buffer via centrifugation with the flow-through being discarded between each centrifugation step. The columns were transferred to a clean microfuge tubes and the plasmid was eluted with either 30 µL Zyppy<sup>™</sup> Elution buffer or 30 µL ddH<sub>2</sub>O by centrifugation and stored at -20°C.

### 2.8.3 Gel purification of DNA

DNA from agarose gels were purified using Zymoclean<sup>™</sup> Gel DNA Recovery Kit (Zymo Research). DNA fragments were excised out of the gel using a scalpel and transferred to pre-weighed microfuge tubes. 3 volumes of ADB buffer was added to each volume of agarose gel and incubated on a heating block at 37-55°C for 5-10 minutes until all the gel dissolved. The solution was transferred to Zymo-Spin<sup>™</sup> columns in collection tubes and centrifuged at 13, 000 x g for 30 seconds and washed twice with Wash Buffer by centrifugation (all centrifugation steps were carried out at 13, 000 x g for 30 seconds). The columns were transferred to new microfuge tubes and the DNA was eluted with 6-10 µL ddH<sub>2</sub>O by centrifugation and stored at -20°C.

### 2.8.4 Purifying salts and enzymes from DNA

Plasmid DNA extracted via phenol/chloroform extraction (section 2.8.2.1) was desalted on nitrocellulose membrane filters (Biolab; section 2.8.4.1) or run through the DNA Clean and Concentrator<sup>™</sup> column (Zymo Research, section 2.8.4.2). The DNA Clean and Concentrator<sup>™</sup> Kit was also used to clean PCR products (section 2.9.1) before ligation reactions (section 2.9.2) and enzymes from restriction digests (section 2.9.6). Plasmid DNA purified using Zyppy<sup>™</sup> Plasmid Miniprep Kit was not desalted as the flow-through via the column removed all salts.

#### **2.8.4.1 Nitrocellulose filters**

Nitrocellulose membrane filters (Biolab) were placed on petri dishes filled with ddH<sub>2</sub>O with the matt surface on the water and the gloss side up. Approximately 50 µL of phenol/chloroform extracted plasmid DNA was placed on top of the filter and was dialysed for 30 minutes to remove all salts. After 30 minutes, the plasmids were recovered from the filter (~ 48 µL), transferred to sterile microfuge tubes and stored at -20°C.

#### **2.8.4.2 DNA Clean and Concentrator Kit<sup>TM</sup>**

One volume DNA Binding buffer was added to DNA samples, mixed via vortexing and transferred to the Zymo-Spin<sup>TM</sup> columns in collection tubes. The tubes were centrifuged at 13,000 x g for 30 seconds (all centrifugation steps were carried out at 13,000 x g for 30 seconds). The flow-through was discarded and 200 µL of Wash buffer was added to the columns and centrifuged. This step was repeated once and the DNA was eluted out with 6-10 µL ddH<sub>2</sub>O via centrifugation into fresh microfuge tubes and stored at -20°C.

#### **2.8.5 DNA quantification**

All nucleic acid was quantified using NanoDrop<sup>®</sup> ND-100 Spectrophotometer (NanoDrop Technologies, Auckland, NZ) using NanoDrop<sup>®</sup> 3.1.0 software. The pedestal was cleaned and initialized with ddH<sub>2</sub>O and blanked with the solution the DNA was dissolved in (TE buffer, elution buffer or ddH<sub>2</sub>O). A 1 µL of the sample was placed on the pedestal and read at 260 nm in ng/µL.

## 2.9 Genetic Manipulation

### 2.9.1 Polymerase Chain Reaction (PCR)

PCRs were performed using eppendorf Mastercycler<sup>®</sup> gradient (Global Science) in PCR tubes using either Taq DNA polymerase (Table 2.6) or PCR Master Mix (Table 2.7). The annealing temperatures and the templates used for each primer set are given in Table 2.8. All PCR cycles were carried out as shown in Table 2.5 except for the amplification of MSMEG\_parA+parB with primers MSMEG\_6939\_F and MSMEG\_6938\_R which had an extension time of 120 seconds in the 30 x cycle.

**Table 2.5: PCR cycles used for basic amplification**

Cycle	Temperature	Time	No. of cycles
Initial denaturation	94°C	2 mins	1 x
Denaturation	94°C	30 sec	
Annealing	As shown in Table 2.8	30 sec	30 x
Extension	72°C	60 sec	
Final extension	72°C	10 mins	1 x

**Table 2.6: Reagents and concentrations used with Taq DNA Polymerase**

<b>Reagent</b>	<b>Stock conc.</b>	<b>Final conc.</b>
Primer_F	10 $\mu$ M	0.4 $\mu$ M
Primer_R	10 $\mu$ M	0.4 $\mu$ M
CoralLoad Buffer	10 x	1 x
Taq	250 U	1 U
MgCl <sub>2</sub>	25 mM	1.5 mM
DMSO	100%	5%
Template	100 – 200 ng/uL	40 – 80 pg

**Table 2.7: Reagents and concentrations used with PCR Master Mix**

<b>Reagent</b>	<b>Stock conc.</b>	<b>Final conc.</b>
Primer_F	10 $\mu$ M	0.4 $\mu$ M
Primer_R	10 $\mu$ M	0.4 $\mu$ M
BioMix Red	2 x	1 x
DMSO	100%	5%
Template	100 – 200 ng/uL	40 – 80 pg

**Table 2.8: PCR annealing temperatures used per primer set**

Primer Name	Anneal temp.	Template
MSMEG_6939_600F; MSMEG_6939_600R	55°C	<i>M. smegamitis</i> mc <sup>2</sup> 155
P118; MSMEG_6939_600R	40°C	pSE100_parA_600
GFP_RBS_F; GFP_R	54°C	pHS201
MSMEG_6938_F ; MSMEG_6938_R	53°C	<i>M. smegamitis</i> mc <sup>2</sup> 155
MSMEG_6939_F; MSMEG_6939_R	53°C	<i>M. smegamitis</i> mc <sup>2</sup> 155
MSMEG_2357_F; MSMEG_2357_R	52°C	<i>M. smegamitis</i> mc <sup>2</sup> 155
MSMEG_2509_F; MSMEG_2509_R	58°C	<i>M. smegamitis</i> mc <sup>2</sup> 155
MSMEG_4623_F; MSMEG_4623_R	56°C	<i>M. smegamitis</i> mc <sup>2</sup> 155
MSMEG_6327_F; MSMEG_6327_R	53°C	<i>M. smegamitis</i> mc <sup>2</sup> 155
ParA_F; ParA_R	57°C	<i>M. tuberculosis</i> H37Rv
ParB_F; ParB_R	55°C	<i>M. tuberculosis</i> H37Rv
ParS_H37Rv_F; ParS_H37Rv_R	45°C	<i>M. tuberculosis</i> H37Rv
TetROF11; HygR11	44°C	pMind*

\* The pMind vectors had various inserts that were screened (section 2.9.5.3) using different annealing temperatures which was suitable for both the forward and the reverse primers.

## 2.9.2 DNA ligations

Two different T4 DNA ligase were used in this study. TA cloning of PCR fragments (Table 2.9) was carried out using the T4 DNA ligase supplied with the pCR<sup>®</sup>2.1 TA Cloning Kit (Invitrogen). All other ligation reactions, that is, cohesive end ligations (Table 2.10), were carried out using T4 DNA ligase purchased from Bioline. All TA ligation reactions were ran with a negative control with no vector. All cohesive end ligation reactions had a negative control with no vector and a control with the digested but unligated vector to identify the percentage of false positives in the transformations (section 2.9.3.2).



**Table 2.9: Reagents and concentrations used for TA cloning**

<b>Reagent</b>	<b>Stock conc.</b>	<b>Final conc.</b>
T4 DNA Ligase	4 U/ $\mu$ L	0.4 U/ $\mu$ L
Ligation Buffer	10 x	1 x
pCR <sup>®</sup> 2.1 vector	25 ng/ $\mu$ L	25 ng
PCR product	-	50 - 100 ng

The above ligation reactions were prepared in PCR tubes and made-up to a total volume of 10  $\mu$ L with ddH<sub>2</sub>O and incubated in the eppendorf Mastercycler<sup>®</sup> gradient at 16°C overnight before transformations (section 2.9.3.2).

**Table 2.10: Reagents and concentrations used for cohesive end ligation**

<b>Reagent</b>	<b>Stock conc.</b>	<b>Final conc.</b>
T4 DNA Ligase	10 U/ $\mu$ L	0.5 U/ $\mu$ L
Ligation Buffer	10 x	1 x
ATP solution	10 x	1 x
Vector	10 – 20 ng/ $\mu$ L	10 ng
Insert	50 – 200 ng/ $\mu$ L	90 ng

The above ligation reactions were prepared in PCR tubes and made-up to a total volume of 10  $\mu$ L and incubated in the eppendorf Mastercycler<sup>®</sup> gradient at 16°C for 30 minutes before transformations (section 2.9.3.2).

### **2.9.3 Transformations into *E. coli***

*E. coli* was transformed with vectors and ligation reactions into calcium competent cells (section 2.9.3.1) with heat-shock (section 2.9.3.2).

### 2.9.3.1 Preparation of calcium competent cells

*E. coli* frozen stocks were streaked on LB agar plates and incubated overnight (section 2.3). Single colonies from the overnight plates were inoculated into 10 mL LB broth in 50 mL polypropylene tubes and cultured overnight (section 2.3). The overnight cultures were sub-cultured in 100 mL LB broth (1:10) in 250 mL flasks and cultured with vigorous shaking (section 2.3) until the optical density at 600 nm ( $OD_{600}$ ) reached  $\sim 0.5$ .  $OD_{600}$  of the cultures were measured in 10 mm path length cuvettes using a Thermo Helios  $\gamma$  Spectrophotometer (all cultures were diluted with the appropriate media to give an  $OD_{600}$  reading ranging between 0.05-0.5 for accurate reading in the spectrophotometer). These mid-logarithmic phase cells were transferred to 50 mL polypropylene tubes and stored on ice for 30 minutes. The cells were harvested by centrifugation at 4,000  $\times g$  for 10 minutes at 4°C (all subsequent centrifugation steps were carried out at 4,000  $\times g$  for 10 minutes at 4°C). The supernatant was decanted and the pellet was resuspended in 50 mL cold sterile 0.1 M  $CaCl_2$ . This was incubated on ice for 20 minutes and the cells were recovered by centrifugation. The cell pellets were resuspended in 2 mL cold sterile 0.1 M  $CaCl_2$  with 14% glycerol and 50  $\mu L$  aliquots of these were stored into microfuge tubes and were either used for heat-shock transformations (section 2.9.3.2) and/or stored at -80°C.

### 2.9.3.2 Heat-shock transformation

Calcium competent cells (section 2.9.3.1) were taken out of -80°C and thawed on ice for 5 minutes. Approximately 150  $\mu L$  cold sterile 0.1 M  $CaCl_2$  was added to the cells and 10  $\mu L$  of the ligation reactions (section 2.9.2) or 50 ng/ $\mu L$  of purified plasmid DNA (section 2.8.2) were added to the cells, incubated on ice for 20 minutes and heat-shocked at 42°C on a heating block for 90 seconds. The tubes were incubated on ice for 5 minutes after heat-shock before 800  $\mu L$  LB broth were added to the cells. These transformations were recovered in LB at 37°C for 60 minutes in a stationary incubator. After 60 minutes, the cells were spun at 13,000  $\times g$  for 30 seconds, 900  $\mu L$  of the supernatant was discarded and the pellet was

resuspended in the remaining 100  $\mu$ L media and plated on LB agar plates with appropriate antibiotics (Table 2.3).

#### **2.9.4 Transformations into *M. smegmatis***

*M. smegmatis* was transformed with vectors using electro-competent cells (section 2.9.4.1) via electroporation (section 2.9.4.2).

##### **2.9.4.1 Preparation of electro-competent cells**

*M. smegmatis* was cultured overnight as described in section 2.8.1.1 in 10 mL LBT in 50 mL polypropylene tubes. The overnight cultures were sub-cultured into 100 mL LBT (1:10) in 250 mL flasks until the OD<sub>600</sub> reached between 0.5 to 1.0. These mid-log phase cells were transferred to 50 mL polypropylene tubes and stored on ice for 20 minutes. The cells were harvested by centrifugation at 4,000 x g for 20 minutes at 4°C (all subsequent centrifugation steps were carried out at 4,000 x g for 20 minutes at 4°C). The supernatant was decanted and the pellet was resuspended in 50 mL cold sterile 10% glycerol, incubated on ice for 10 minutes and recovered by centrifugation. The cell pellet was resuspended in 20 mL cold 10% glycerol, recovered by centrifugation and resuspended in 4 mL 10% glycerol. Approximately 200  $\mu$ L of this suspension was aliquoted out into microfuge tubes and were either used for electroporations (section 2.9.4.2) and/or stored at -80°C.

##### **2.9.4.2 Electroporation**

Electro-competent cells (section 2.9.4.1) were taken out of the -80°C freezer and thawed on ice for 5 minutes. Approximately 200  $\mu$ L of cold sterile 10% glycerol was added to the cells before they were transferred into chilled 0.2 cm electrode gap electroporation cuvettes (Raylab, Auckland, NZ). Between 0.5-5  $\mu$ g purified, desalted plasmid DNA (section 2.8.2 and section 2.8.4 respectively) were added

to the cuvettes and incubated on ice for 2 hours. After incubation, the cuvettes were placed in the electroporation chamber and subjected to a single pulse of 2.5 kV, capacitance of 25  $\mu$ F and the pulse-controller resistance at 1000  $\Omega$  using a Gene Pulser Xcell<sup>™</sup> Electroporation System (Bio-Rad). After electroporation, the cells were transferred into microfuge tubes, 600  $\mu$ L LB broth was added to these and incubated at 37°C for 6-16 hours. After incubation, the cells were harvested by centrifugation at 13, 000 x g for 30 seconds, 900  $\mu$ L of supernatant was removed and the cells were resuspended in the remaining 100  $\mu$ L media and plated on LB agar plates with appropriate antibiotics (Table 2.3).

## **2.9.5 Identification of recombinant clones**

Recombinants clones were initially identified via colony screening (section 2.9.5.1). Clones with the correct inserts were cultured overnight (section 2.3) in LB with appropriate antibiotics (Table 2.3) from which plasmids were extracted (section 2.8.2).

### **2.9.5.1 Colony screening**

Colonies were picked from agar plates and used as templates in the PCR reactions with the appropriate primers (Table 2.4) and their respect annealing temperatures (Table 2.8) using the PCR protocols shown in section 2.9.5.3. Approximately 5  $\mu$ L PCR products were run on 1% agarose gels (section 2.9.8) and the presence of the correct size PCR product verified that the insert was present in the colony. Clones verified with inserts were prepared using either of the protocols in section 2.8.2 for further analysis.

### 2.9.5.2 PCR products in pCR<sup>®</sup>2.1

Identification of recombinant clones in pCR<sup>®</sup>2.1 vector (Appendix 9.1) was done via blue/white selection on LB agar plates supplemented with Amp<sup>+</sup>, Kan<sup>+</sup>, IPTG<sup>+</sup>, X-Gal<sup>+</sup> (refer to Table 2.3 for conc.). White colonies signified insertion of the PCR product. To verify this, single colonies were picked for colony screening (section 2.9.5.1) and also inoculated into 3 mL LB Amp<sup>+</sup>, Kan<sup>+</sup> broth in 15 mL polypropylene tubes to be cultured overnight for plasmid preparation (section 2.8.2), restriction analysis (section 2.9.6) and DNA sequencing (section 2.9.9) to confirm the insertion of the correct PCR product. Restriction analysis verified that the correct enzyme sites from the PCR primers (Table 2.4) had been integrated and agarose gel electrophoresis (section 2.9.8) was used to confirm that the correct size fragment was inserted before plasmids were sent for sequencing. DNA sequencing (section 2.9.9) was performed using primers M13F and M13R (Table 2.4) and it verified the exact nucleotide sequence of the insert.

### 2.9.5.3 Orientation PCR

Orientation PCRs were performed when DNA fragments containing the same restriction sites on both ends were used for cloning into the pMind vector, that is, GFP into the *SpeI* site, *parA* into the *EcoRV* site, and MSMEG\_2357 into the *PacI* site (Table 2.4). The orientations of the inserts were screened using an external primer designed from the sequence of the plasmid (TetROF11 and/or HygR11) and an internal primer designed from the sequence of the insert (Table 2.4).

All orientation PCR reactions were denatured for 2 minutes at 94°C followed by 15 cycles of denaturation for 30 seconds at 94°C, 30 seconds at annealing temperatures shown in Table 2.8 for the respective gene primers used and 60 seconds of extension at 72°C. This was followed by 15 cycles of denaturation for 30 seconds at 94°C, 30 seconds of annealing at 44°C for the external primer and

60 seconds of extension at 72°C, followed by a final extension for 10 minutes at 72°C (Table 2.11).

**Table 2.11: PCR cycles used for orientation screens**

Cycle	Temperature	Time	No. of cycles
Initial denaturation	94°C	2 mins	1 x
Denaturation	94°C	30 sec	15 x
Annealing	As shown in Table 2.8	30 sec	
Extension	72°C	60 sec	
Denaturation	94°C	30 sec	15 x
Annealing	44°C	30 sec	
Extension	72°C	60 sec	
Final extension	72°C	10 mins	1 x

### 2.9.6 Restriction digests

All restriction digests were performed with the appropriate New England Biolabs (NEB) buffers supplied with the enzyme. Appropriate buffers for double digests were searched in the Double Digest Finder, NEB (<http://www.neb.com/nebecomm/doubledigestcalculator.asp>). Approximately 1x BSA was added to some enzyme digests as recommended by the manufacturer.

All restriction digests (Table 2.12) for the analysis of insertions into vectors were made up to a total volume of 10  $\mu$ L with ddH<sub>2</sub>O and incubated at 37°C for 4 hours before running the digests on a 1% agarose gel (section 2.9.8).

All restriction digests performed to obtain cohesive end DNA fragments for ligation reactions (section 2.9.2) were incubated overnight at 37 °C.

Digests used to extract DNA fragments from vectors were made up to a total volume of 50 µL with ddH<sub>2</sub>O and incubated overnight at 37°C and run on a 1% agarose gel (section 2.9.8) to separate the extracted DNA fragment from the plasmid DNA before being gel purified (section 2.8.3). All vectors that were digested to ligate DNA fragments into were made up to a total volume of 20 µL with ddH<sub>2</sub>O and put through the DNA Clean and Concentrator<sup>™</sup> column (Zymo Research, Section 2.8.4.2) after the overnight incubation to purify the DNA of enzymes and salts from the buffers.

**Table 2.12: Reagents and concentrations used for restriction digests**

Reagent	Stock conc.	Final conc.
Restriction Enzyme	10, 000 U/µL	0.5 U/µL
NEBuffer	10 x	1 x
BSA	10 x	1 x
Vector	50 – 200 ng/µL	25 - 100 ng

### 2.9.7 Dephosphorylation

Vectors digested with the same restriction site for ligation of DNA fragments were first dephosphorylated using calf intestinal phosphatase (CIP). Digested vectors were clean and concentration (section 2.8.4.2) and quantified (section 2.8.5). Approximately 10 units of CIP was incubated with 1 µg of DNA at 37°C for 4 hours. After 4 hours the reaction was clean and concentrated again and used in ligation reactions (section 2.9.2)

### 2.9.8 Agarose gel electrophoresis

1% agarose gels were prepared using 1g of agarose dissolved in 100 mL 1 x TAE buffer (Appendix 9.5.4) via microwaving. Approximately 100 mL of 1% agarose was mixed with 10mg/ml ethidium bromide, poured into a gel cast with combs and set at RT for 30 minutes. 5  $\mu$ L of DNA ladder and DNA with 5x DNA loading buffer (1: 4) were loaded on the gel and ran at 100-120 V for 60-90 minutes. Gels were visualized using a High Performance UV transilluminator. Gel photos were taken with an Olympus C-5060 Wide Zoom digital camera and analyzed using UVP LS (Life Sciences) Software.

### 2.9.9 DNA sequencing

DNA sequencing reactions were sent to The Allan Wilson Centre for Molecular Ecology and Evolution Genome Services (Palmerston North, NZ) in PCR tubes containing 300 ng of plasmid DNA, 3.2 pmol of the appropriate primer and ddH<sub>2</sub>O to make the reaction up to 15 $\mu$ L. The chromatograms received from the sequencing reactions were analyzed using Vector NTI Advance 10 software (Invitrogen).

### 2.9.10 Curing plasmids

*M. smegmatis*-pMind\_parA+parB\_GFP colonies were picked from LB Kan<sup>+</sup>, Hyg<sup>+</sup> plates and cultured overnight in 10 mL LBT (section 2.3) in 50 mL polypropylene tubes. These were subcultured into 100 mL LBT in 250 mL flasks, grown overnight, subcultured again (1:10) into 100 mL LBT in 250 mL flasks and grown overnight. Serial dilutions of the last overnight cultures were plated out onto LB agar plates. Singles colonies were picked from these serial dilution plates and patched onto LB and LB Kan<sup>+</sup>, Hyg<sup>+</sup> plates. Colonies that lost resistance were used as cured strains.



## 2.10 Protein Over-expression, Purification and Quantification

### 2.10.1 Protein over-expression optimization

Protein over-expressions were carried out in *E. coli* LMG194 with the pBAD/HisB vector and in *E. coli* BL21 (DE3) with the pET-28a(+) vector. Both these vectors were cloned with the wild-type *parA* and *parB* genes from *M. tuberculosis* H37Rv under the transcriptional control of the arabinose promoter for the pBAD/HisB vector and the T7 promoter for the pET-28a(+) vector (Table 2.2).

#### 2.10.1.1 Expression of recombinant protein in *E. coli* LMG194

*E. coli* LMG194 expression strains were overexpressed in various conditions to try to obtain soluble Par proteins. *E. coli* LMG194 transformed with pBAD/HisB\_LacZ was used as a positive control.

Expression strains were taken from the -80°C freezer and streaked onto LB agar Amp<sup>+</sup> plates and incubated overnight (section 2.3). Single colonies were picked from these plates and cultured into 3 mL LB Amp<sup>+</sup> overnight (section 2.3). Approximately 50 µL of the overnight cultures were subcultured into 10 mL LB Amp<sup>+</sup> in 50 mL polypropylene tubes and subjected to the conditions below:

- I. The first experiment was carried out at 37°C. The subcultured cells were grown to an OD<sub>600</sub> of 0.5 and induced with 0.2% L-arabinose for 4 hours. (Initial experiments with varying concentrations of L-arabinose (0.0002% - 0.2%) at different time points was carried out to determine the concentration of L-arabinose and the time required for optimal over-expression).

- 
- II. The second experiment was carried out exactly as described in I above except after the OD<sub>600</sub> reached around 0.5, the cultures were cooled to 18°C and induced overnight with 0.2% L-arabinose at 18°C.
- III. The third experiment was carried out in high salt media. The starter cultures were grown overnight at 37°C in LB with 0.5 M NaCl, subcultured as described above in LB containing 0.5 M NaCl and 1 mM betaine, grown till the OD<sub>600</sub> reached around 0.5 and induced overnight at 37°C with 0.2% L-arabinose.
- IV. The fourth experiment was carried out as described in III above with a few modifications. The subcultured cells were grown till the OD<sub>600</sub> reached 0.9, heatshocked at 47°C in a waterbath, induced with 0.2% L-arabinose for 20 minutes in the 47°C waterbath before it was transferred to 25°C for overnight expression [128].
- V. The fifth experiment was carried out in auto-inducing media ZYP-5052 (Appendix 9.5.2) [129]. Overnight starter cultures were subcultured in ZYP-5052 and grown overnight at 37°C in the absence of the L-arabinose.

#### **2.10.1.2 Expression of recombinant protein in *E. coli* BL21**

*E. coli* BL21(DE3) expression strains were cultured and expressed similar to *E. coli* LMG194 expression strains (section 2.10.1.1) except all media had Kan instead of Amp and all cultures were induced with 0.5 mM IPTG instead of 0.2% L-arabinose except in autoinducing media which did not require an inducer. The fifth condition mentioned above (section 2.10.1.1) with ZYP-5052 auto-induction media gave the most soluble protein. This was scaled up to 500 mL over-expression cultures for protein isolation.

### 2.10.2 Purification of recombinant protein

Approximately 500 mL of overexpressed cultures (section 2.10.1.2) were pelleted in 50 mL polypropylene tubes and were lysed via resuspending the pellet in:

1. 20 mL Bug Buster<sup>®</sup> Protein Extraction reagent (Novagen) for 20 minutes on a rocking platform at room temperature (RT).
2. 20 mL cold 1x bind buffer (Novagen) and French pressed (Thermo Scientific) three times at 1000 psi.
3. 30 mL cold 1x bind buffer\* (Novagen) and sonicated (Ultrasonic processor W-380) for 3 minutes in one second bursts with a power output ranging from 50-70% in tubes packed with crushed ice.

The cell lysate was transferred to 50 mL polypropylene Sorvall tubes and the cellular debris was removed from the soluble fraction via centrifugation at 14,000 in a Sorvall RC-5C centrifuge (Thermo Scientific) for 30 minutes at 4°C. The supernatant was column purified using the Ni-NTA His•Bind<sup>®</sup> Resin and His•Bind<sup>®</sup> Purification kit (Novagen).

Approximately 3 mL His•Bind<sup>®</sup> resin was added to the chromatography column and settled via gravity. The column was washed with 4.5 mL ddH<sub>2</sub>O with the peristaltic pump flow rate set at 3 mL/minute throughout the purification process. This was followed by 8 mL of 1x charge buffer, 4 mL of 1x bind buffer, 20 mL of soluble fraction, 20 mL of 1x bind buffer, 8 mL of 1x wash buffer and 9 mL of 1x elution buffer. The eluate was collected in 1.5 mL fractions and ran on SDS gels (section 2.10.5). The most concentrated fraction was desalted (section 2.10.3) and used in the biochemical assays (section 2.11).

\* ParA protein immediately precipitated upon elution from the column with elution buffer. To prevent this, 10% glycerol (v/v) was used in all buffers for purification of ParA.

### 2.10.3 Desalting Proteins

All protein was desalted using 5 mL HiTrap™ desalting columns (GE Healthcare). The desalting column was clamped and attached to a 50 mL syringe and equilibrated with 25 mL 50 mM Tris-HCl pH 7.5 (Appendix 9.5.4) at a flow rate of 3 mL/min. Approximately 1.5 mL of purified protein fraction was put through the column and the protein was eluted out in 2 mL 50 mM Tris-HCl pH 7.5 and stored at 4°C.

Once again all buffers used for desalting ParA had 10% glycerol (v/v). Desalted ParA was stored in 50% glycerol at -20°C.

### 2.10.4 Protein quantification

All protein was quantified using the *DC* Protein Assay (Bio-Rad). Four dilutions (0.25-1.5 mg/mL) of BSA was prepared in 50 mM Tris-HCl pH 7.5 as standards. 5 µL of proteins and standards were aliquoted into a 96 well plate in triplicate. Approximately 25 µL of reagent A and 200 µL of reagent B was added to each well containing protein samples or standards respectively and incubated at RT for 15 minutes. This was read at 750 nm.

### 2.10.5 SDS-PAGE

About 4.5 mL of 12% separating gel (Appendix 9.5.5) was poured between gel plates (Bio-Rad), covered with 700 µL 100% isopropanol and set at RT for 30 minutes. After 30 minutes the isopropanol was decanted off and 4% stacking gel (Appendix 9.5.5) was poured on top of the separating gel, the combs put in place and set at RT for 45 minutes. The combs were removed after 45 minutes. The wells were washed gently with water to remove unpolymerized acrylamide and the apparatus was placed in the gel tank (Bio-Rad) and filled with 1x SDS running buffer (Appendix 9.5.4).

Approximately 1 mL bacterial pellets were resuspended into 1 mL of 2% SDS. About 20  $\mu$ L of these samples and/or 20  $\mu$ L of soluble fraction or purified protein were added to 10  $\mu$ L 3x sample buffer (2:1) (Appendix 9.5.4) and boiled at 95°C in the thermocycler (eppendorf Mastercycler<sup>®</sup> gradient) for 10 minutes before being loaded onto the gel. Approximately 10  $\mu$ L (small wells) or 20  $\mu$ L (big wells) of samples were loaded in the wells with 5  $\mu$ L protein molecular weight marker (also boiled with sample buffer) and run at 200 V for 60 minutes.

Once the bromophenol blue ran off the end of the gels, the gels were carefully removed from between the glass plates and stained for 30 minutes in coomassie stain (Appendix 9.5.3) on a rocking platform at RT. After 30 minutes, the gels were rinsed 3 times in water and destained (Appendix 9.5.3) for 2 hours with one change of destain after an hour. After 2 hours, the gels were rinsed in water and left overnight in water for complete destaining. The gels were scanned with a Personal Densitometer SI (Molecular Dynamics) and the images were processed using Image J (NIH).

#### **2.10.6 Polyclonal Antibody Production**

Approximately 5mL of 1 mg/mL purified ParA protein (section 2.10.2) was sent to the Pocono Rabbit Farm and Laboratory Inc. (PRF&L, Pennsylvania, USA) to raise polyclonal antibodies against ParA in a SPF white female rabbit. The rabbit was initially immunized with Complete Freund's Adjuvant (CFA) followed by Incomplete Freund's Adjuvant (IFA) for all subsequent injections with 50  $\mu$ g to 200  $\mu$ g of immunogen (Table 2.13). The immunogen was diluted to 0.25 mL with sterile saline and combined with 0.25 mL of the adjuvant, mixed thoroughly to form a stable emulsion before it was either injected beneath the skin of the rabbit (subcutaneously (SC)) in the area around the shoulders or intradermally (ID) into the hind haunches near the rear legs for a maximum of two times. Blood was collected from the central ear artery with a 21-gauge needle and allowed to clot and retract at 37°C overnight. The serum was decanted and clarified by

centrifugation at 2, 500 rpm for 15 minutes. The schedule of immunization and bleeds are shown in Table 2.13.

**Table 2.13: Schedule for the production of ParA antibodies**

<b>Days</b>	<b>Procedure – immunogen used (µg)</b>	<b>Adjuvant description</b>
0	Preimmune Bleed – 4 mL	-
0	Inject ID – 200 µg	CFA
14	Inject ID – 100 µg	IFA
28	Inject SC – 100 µg	IFA
42	Small bleed - 4 mL	-
56	Inject SC – 50 µg	IFA
70	Small bleed - 4 mL	-
84	Inject SC – 50 µg	IFA
98	Final Bleed - 4 mL	-

### **2.10.7 Western Blotting**

SDS-PAGE was run as described in section 2.10.5 except Amersham ECL Plex™ fluorescent rainbow marker (GE Healthcare) was used instead of the standard protein molecular weight marker. Amersham Hybond-LFP membrane and blotting paper (GE Healthcare) were cut to the size of the gel. The membrane was activated in methanol for 30 seconds before being equilibrated in the transfer buffer (Appendix 9.5.4) with the gel for 20 minutes. All blotting paper and foam pads were soaked in transfer buffer before assembly. The blotting sandwich was assembled in transfer buffer with forceps from the negative electrode with a foam pad followed by 3 pieces of blotting paper, gel, PVDF membrane, 3 pieces of blotting paper and another foam pad. All air bubbles were rolled off with a roller and the assembly was closed, placed in the tank with an ice pack, filled with transfer buffer and ran at 4°C for 2.5 hours at 37 V. All western blot apparatus used in this experiment were from Bio-Rad.

After transfer, the membrane was rinsed in 1x PBS (Appendix 9.5.3) and blocked with 3% low-fat milk in 1x PBS for 1 hour at RT. All wash steps and incubation periods were carried out at RT on a rocking platform. After an hour, the blocked membrane was washed twice for 5 minutes in 1x PBST (Appendix 9.5.3) and incubated with a 1:1000 dilution of the polyclonal antibody (section 2.10.6) in 1x PBST for 2 hours. After incubation with the primary antibody, the membrane was washed four times for 5 minutes in 1x PBST and incubated with a 1:500 dilution of Amersham ECL™ Plex goat-anti-rabbit IgG Cy3 (GE Healthcare) in 1x PBST for 1-2 hours. After incubation with the secondary antibody, the membrane was washed four times for 5 minutes with 1x PBST followed by two final washes with 1x PBS for 5 minutes. The transferred fluorescent labelled protein was scanned with a Fujifilm FLA-5100 imaging system (Alphatec) using ImageReader FLA-5000 series V1.0 software.

## **2.11 *In vitro* ATPase Assay**

The biochemical assay performed with the purified ParA protein was modified and adapted from various published sources [130-132] to identify the time frame and the protein concentration that resulted in a linear relationship with respect to the ParA activity. The reaction for the assay was set-up as shown in Table 2.14 and made up to a total volume of 100  $\mu$ L with ddH<sub>2</sub>O in 96 well plates. All assays were incubated at 37°C unless otherwise stated. The hydrolysis of ATP in the reaction was determined by quantifying the amount of inorganic phosphate (P<sub>i</sub>) released using the brilliant green assay, a modified protocol of the malachite green assay from McQuade *et. al.*, 2009 [133] (section 2.11.1).

**Table 2.14: Concentration of reagents used in the ParA ATPase assays**

Reagents	Initial conc.	Final conc.
Tris-HCl pH7.5	1 M	50 mM
DTT	100 mM	1 mM
MgCl <sub>2</sub>	100 mM	1 mM
KCl	1 M	50 mM
BSA	10 mg/mL	0.1 mg/mL
ATP	10 mM	0 – 1 mM
ParA	*	1 µg

\* The initial concentration of ParA varied in different preps

### 2.11.1 Brilliant green assay

The release of P<sub>i</sub> upon ATP hydrolysis was measured using a colourmetric assay. Approximately 20 µL of Reagent A (4.2% ammonium molybdate in 4 M sulphuric acid) was added to the ATPase assay and mixed briefly. This was followed by the addition of 20 µL Reagent B (0.045% brilliant green in ddH<sub>2</sub>O) and again mixed briefly. The reaction was incubated at RT for 2-5 minutes and read at 650 nm.

## 2.12 Antisense Assays

### 2.12.1 Growth curves

All antisense expression vectors created via insertion of various genes into the pMind vector in the antisense orientation under the transcriptional control of the tetracycline-inducible promoter (Table 2.2) were electroporated into *M. smegmatis* (section 2.9.4.2) and subjected to growth assays.



Growth assays were conducted in rich (LB broth), carbon and nitrogen starvation medias (HdeB minimal media; Appendix 9.5.1) with appropriate media supplements (section 2.2.4) and culture conditions (section 2.3) in 96 well plates in a total volume of 100  $\mu$ L. The promoter was induced with a titration of tetracycline from 0- 40ng/mL and growth was monitored via measuring the OD<sub>590</sub> of the cultures using the Wallac EnVision 2102 Multilabel Reader and Wallac EnVision Manager software (version 1.12) (Perkin Elmer). Induction of the promoter and antisense expression was confirmed via measuring the relative fluorescence in strains that contained a *gfp* sequence downstream of the antisense sequence (Table 2.2) using the multilabel reader with an excitation wavelength of 485 nm and an emission wavelength of 510 nm.

*M. smegmatis* strain hosting pMindGFP vector was used as controls in the antisense assays. The plate was divided into 2 parts; columns 1-6 with no Tc and columns 7-12 with 20 ng/mL Tc (the optimal concentration for promoter induction).

#### **2.12.1.1 Rich media**

*M. smegmatis* colonies transformed with the antisense vectors were picked from LB Kan<sup>+</sup>, Hyg<sup>+</sup> agar plates (refer to Table 2.3 for conc.) and cultured overnight (section 2.3) in 5 mL LBT Kan<sup>+</sup>, Hyg<sup>+</sup> until they reached mid-logarithmic phase with OD<sub>600</sub> ranging from 1 to 1.5. The mid-log phase cultures were sub-cultured in LBT Kan<sup>+</sup>, Hyg<sup>+</sup> and Kan<sup>+</sup>, Hyg<sup>+</sup>, Tc<sup>+</sup> with the starting OD<sub>590</sub> of 0.01 in 96 well plates as described in section 2.12.1 and monitored for OD<sub>590</sub>, the number of CFU (section 2.12.2), and/or fluorescence. Cells were also analyzed with 4',6-diamidino-2-phenylindole (DAPI) staining (section 2.12.3) at 144 hours.

### 2.12.1.2 Nutrient starvation media (Appendix 9.5.1)

*M. smegmatis* colonies transformed with the antisense vectors were picked from LB Kan<sup>+</sup>, Hyg<sup>+</sup> agar plates and cultured overnight in 5 mL HdeB oxygen starvation media (O<sub>2</sub><sup>-</sup>) with Kan<sup>+</sup>, Hyg<sup>+</sup> until they reached mid-logarithmic phase with OD<sub>600</sub> ranging from 0.5 to 1.0. The mid-log phase cultures were sub-cultured in carbon and nitrogen starvation (C<sup>-</sup> and N<sup>-</sup>) medias with Kan<sup>+</sup>, Hyg<sup>+</sup> and Kan<sup>+</sup>, Hyg<sup>+</sup>, Tc<sup>+</sup> with a starting OD<sub>590</sub> of 0.05 in 96 well plates as described in section 2.12.1 and monitored for OD<sub>590</sub>, the number of CFU (section 2.12.2) and/or fluorescence. Cells were also analyzed with DAPI staining (section 2.12.3) at 72 hrs.

### 2.12.2 Colony forming units (CFU's)

Cells were removed at 0, 24, 72 and 144 hours of the growth assays to calculate the CFU/mL of the cultures. About 20 µL of cells were serially diluted with 180 µL 0.15% NaCl and 100 µL of various serial dilutions were plated on LB agar plates. These plates were incubated at 37°C for 72 hours and colonies were counted from plates that had between 50 -150 colonies. CFU/mL was calculated using the formula:

$$\text{CFU/mL} = \frac{\text{number of colonies (n)} \times \text{dilution factor of plate (10}^n\text{)}}{\text{volume of culture on plate (0.1 mL)}}$$

### 2.12.3 DAPI staining

All cells were stained with ViaGram™ Red<sup>+</sup> Bacterial Gram Stain and Viability Kit (Molecular Probes). Approximately 50 µL of cells were spun down at 13, 000 x g for 1 minute (all subsequent centrifugation steps were carried out at 13, 000 x g for 1 minute), the supernatant was removed and the cells were washed with 50 µL BSA-saline solution (0.25% BSA: 0.15 M NaCl). This was spun down and resuspended in fresh BSA-saline solution. About 2.5 µL of 2 mg/mL WGA

conjugate was added to the cells and incubated at RT for 10 minutes, spun down and resuspended in BSA-saline solution. Approximately 2.5  $\mu$ L of DAPI stain: SYTOX green working solution in water (1:1:20) was added to the resuspended cells and incubated again at RT for 10 minutes. About 10  $\mu$ L the stained cells were transferred onto slides, covered with a cover slip and observed under oil immersion using the Olympus Provis AX70 microscope. Cell lengths were measure using the calibrated eye piece graticule.

## 2.13 Differential Susceptibility Assays

Differential susceptibility assays were carried out with *M. smegmatis* strains transformed with vectors pMind\_parA\_GFP and pMind2357. These strains were screened against 3 compound libraries, LOPAC (Library of Pharmacologically Active Compounds; Sigma Aldrich), Spectrum Collection (MicroSource Discovery Systems, Inc.) and NIH Diversity Set (NIH) in three different culture conditions, LBT, C<sup>-</sup> and N<sup>-</sup> (Table 2.15). All assays were conducted in triplicate.

**Table 2.15: Compound library screens against anitsense strains**

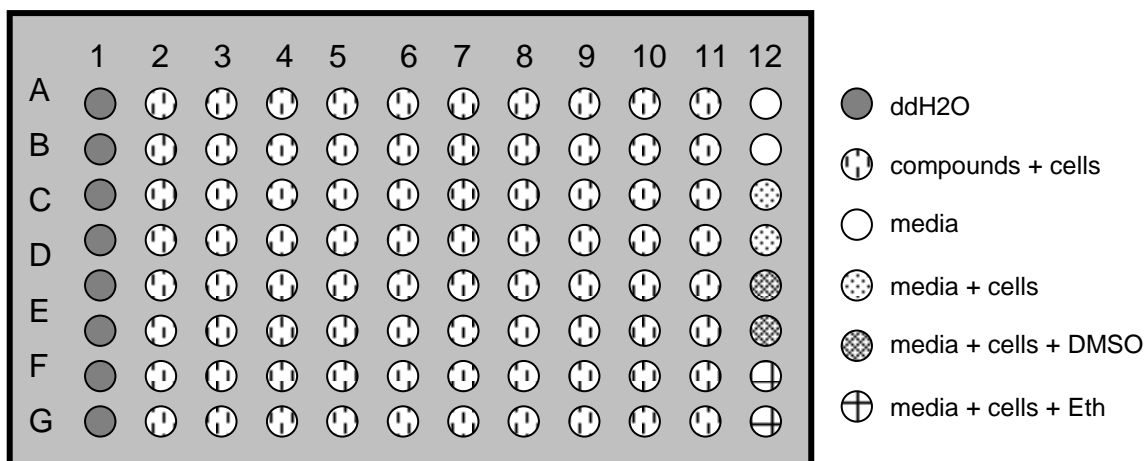
Library	Strain	LBT	C <sup>-</sup>	N <sup>-</sup>
LOPAC	<i>M. smegmatis</i> _pMind_parA_GFP	3x	3x	3x
	<i>M. smegmatis</i> _pMind2357	3x	3x	3x
Spectrum Collection	<i>M. smegmatis</i> _pMind_parA_GFP	3x	3x	3x
	<i>M. smegmatis</i> _pMind2357	3x	3x	3x
NIH	<i>M. smegmatis</i> _pMind-parAGFP	3x	3x	3x
	<i>M. smegmatis</i> _pMind2357	3x	3x	3x

### 2.13.1.1 Assay set-up

All screens were performed in 96 well plates in a total volume of 100  $\mu$ L as described previously [124]. Column 1 was filled with 100  $\mu$ L ddH<sub>2</sub>O to minimize evaporation from the plates during incubation and column 12 had duplicates of various controls, that is, negative control (media), cell growth control, solvent control and positive control (ethambutol at 50  $\mu$ M). All compound libraries were resuspended in DMSO thus DMSO was used as solvent control in all assays (Figure 2.1).

Approximately 50  $\mu$ L of media with Kan<sup>+</sup>, Hyg<sup>+</sup> Tc<sup>+</sup> (Table 2.3) was transferred to columns 2 to 12. Exactly 2  $\mu$ L of library compound from 1 mM working plates were dispensed into wells from column 2 to column 11 using the Cybi-Well liquid handler (CyBio) covering 80 compounds per plate with a final concentration of 20  $\mu$ M per well.

About 20 mL cultures of the strains to be screened were grown overnight in 50 mL polypropylene tubes to mid-log phase in either LBT Kan<sup>+</sup>, Hyg<sup>+</sup> for LBT assays or in HdeB O<sub>2</sub><sup>-</sup> Kan<sup>+</sup>, Hyg<sup>+</sup> for C<sup>-</sup> and N<sup>-</sup> assays (section 2.3). The overnight cultures were subcultured with the appropriate media, that is, LBT, C<sup>-</sup> and N<sup>-</sup> with Kan<sup>+</sup>, Hyg<sup>+</sup> Tc<sup>+</sup> to get a starting OD<sub>600</sub> of 0.15 in 10 mm path length cuvettes (Thermo Helios  $\gamma$  Spectrophotometer). Approximately 50  $\mu$ L of the subcultured cells were aliquoted into all columns wells with cells (Figure 2.1). The plates were sealed and incubated as described in section 2.3 for 72 hours. OD<sub>590</sub> and GFP fluorescence measurements were taken after the incubation period as described in section 2.12.1.



**Figure 2.1: Plate map used to screen compound libraries**

### 2.13.1.2 Data Analysis

All data from the library screens were compiled and analysed in excel worksheets. Wells that exhibited a greater than 40% inhibition in the *M. smegmatis*\_pMind\_parA\_GFP strain without having an effect on the *M. smegmatis*\_pMind2357 control strain were taken as hits and the cells from these wells were analyzed via microscopy (Olympus Provis AX70) for GFP and with DAPI stain (section 2.12.3). These cells were compared to the corresponding wells of *M. smegmatis*\_pMind2357 strain.

### 2.13.1.3 Statistical analysis

All differential susceptibility assays used for HTS of compound library had their Z factors calculated [134] to evaluate the quality of the assays where the sample was cells grown in 2% DMSO with 20 ng/mL Tc giving maximum signal and the control was the antibiotic control giving minimal signal using the formula:

$$Z = 1 - \frac{3 (\text{SD of sample} + \text{SD of control})}{|\text{mean of sample} - \text{mean of control}|}$$

## 2.14 Dose-response Assays

Compounds that were validated as inhibitors of ParA in the biochemical assay (section 2.11) were analyzed for their IC<sub>50</sub> and MIC values in *M. smegmatis*-pSHIGH\_hsp60 in LBT, C<sup>-</sup> and N<sup>-</sup> medias (supplemented with Kan<sup>+</sup>) and in *M. bovis* BCG-pYL GFP (supplemented with Kan<sup>+</sup>) and *M. tuberculosis* H37Ra in Middlebrook 7H9 broth (section 2.2.4). *M. smegmatis*-pSHIGH\_hsp60 strains were cultured (section 2.3) overnight in appropriate medias. *M. bovis* BCG-pMOD12\_hsp70 and *M. tuberculosis* H37Ra were cultured for 16 days. All cells were diluted to a starting OD<sub>600</sub> of 0.05 and 50 µL of these were dispensed into 50 µL of two-fold serially diluted drugs in the appropriate media. All assays were set-up in 96 well plates in a total volume of 100 µL and incubated at 37°C for 4 days for *M. smegmatis* and 14 days for *M. bovis* BCG and *M. tuberculosis* H37Ra. IC<sub>50</sub> and MIC values were determined from OD<sub>590</sub> and/or GFP readings. All dose-response assays were performed in triplicate and all data were analyzed with SigmaPlot 11 (SYSTAT) using four-parameter logistic standard curve analysis to determine IC<sub>50</sub> and MIC values.

## 2.15 Stationary Bactericidal Assays

Stationary bactericidal assays were carried out with *M. smegmatis*-pSHIGH+hsp60 in LBT, C<sup>-</sup> and N<sup>-</sup> medias supplemented with Kan<sup>+</sup>. All LBT assays had starter cultures grown in LBT Kan<sup>+</sup> for 8 days. After 8 days, the cultures were diluted down to an OD<sub>600</sub> of 0.5 with PBS. All starvation assays had starter cultures either grown in C<sup>-</sup> or N<sup>-</sup> media for 10 days. After 10 days, the cells were spun down and resuspended in PBS to an OD<sub>600</sub> of 0.5. 50 µL of these cells were used in standard two-fold dilutions in PBS for dose-response assays in a total volume of 100 µL. All plates were exposed to drugs for 96 hours at 37°C. After 96 hours, viability of the cells was measured by adding 30 µL of 0.02% (w/v) resazurin. This was further incubated for another 24 hours at 37°C. After 24 hours, the assay plates

---

were read with the Wallac EnVision 2102 Multilabel Reader at excitation wavelength of 530 nm and emission wavelength of 590 nm. All bactericidal assays were performed in triplicate and all data were analyzed with SigmaPlot 11 (SYSTAT) using four-parameter logistic standard curve analysis to determine IC<sub>50</sub> and MIC values.

## **2.16 Cytotoxicity Assays**

Cytotoxicity assays were carried out using cancer cell line HL-60. The cells were maintained in RPMI supplemented with 10% fetal calf serum (FCS) and penicillin-streptomycin at 100:10:1. All assays were set-up in 96 well plates in a total volume of 100  $\mu$ L. 50  $\mu$ L of 10, 000 cells per well were dispensed into two-fold serially diluted drugs and cultured at 37°C in a 5% CO<sub>2</sub> incubator for 48 hours. Cell viability was measured by adding 10  $\mu$ L of Alamar blue reagent. This was further incubated for 4 hours at 37°C and read at OD<sub>750</sub> and/or fluorescence at excitation wavelength of 570nm and emission wavelength of 585 nm. All cytotoxicity assays were performed in triplicate and all data were analyzed with SigmaPlot 11 (SYSTAT) using four-parameter logistic standard curve analysis to determine IC<sub>50</sub> values.

## **CHAPTER THREE**

### **Target Selection**



### 3.1 Introduction

The availability of whole genome sequences has led to a new era in drug development, with bioinformatics often being the primary step in the drug discovery programme today [135, 136]. *In silico* technologies are routinely used for the identification of novel targets by predicting the putative functions [137, 138] and/or the metabolic activity of these targets [67, 135, 139]. This is a more rapid and inexpensive method compared to the traditional methods of target validation using a series of experiments, that is, cell-based assays and loss-of-function phenotypes using knock-out mutants. These experiments are still required at later stages of target validation [135] but the targets identified via *in silico* technologies save a lot of time and money before the validation experiments are carried out [67, 140]. This chapter reviews potential drug targets identified via bioinformatics analyses.

#### 3.1.1 Comparative genomics

*Haemophilus influenzae* Rd KW20 was the first bacterial genome sequenced by The Institute of Genomic Research (TIGR) in 1995 [141]. Since then, genome sequencing data has increased exponentially with approximately 2,605 different bacterial genomes being sequenced, most of these publicly available at the Genomes OnLine Database (GOLD; <http://www.genomesonline.org/gold.cgi>) as well as at the individual sequencing centres. The availability of these genomic data has led to genome mining, where genes are assigned functions based on sequence homology and conservation between genomes. It also revealed some of the differences between genomes. This is particularly relevant to pathogenic bacteria as these differences may shed light on the variety of mechanisms which are associated with their unique virulence and clinical manifestations.

It is now widely accepted that bacterial genomes contain a core set of genes encoding proteins that have essential cellular functions [67, 142, 143]. These genes can be easily and rapidly identified using comparative genomics.

Such genes present a set of novel broad-spectrum antimicrobial targets as they are vital to the survival of the microbe. However, ideal targets have to fulfil certain parameters making them difficult to ascertain, that is, they should have a known function so a biochemical assay can be developed to be used in the HTS [135, 144]. Most importantly ideal targets should be absolutely imperative for the organisms survival [104, 145] so targeting it will result in the elimination of the microbe to prevent recurrence of disease after therapy [67].

### **3.1.2 Identification of novel targets using a genomics approach**

There is a great need to identify new antimicrobials with novel MOA [146] and/or different cellular targets to prevent cross-resistance from existing antibiotics [104]. To date there has been no novel antitubercular drugs introduced for therapy in over 3 decades [110] reflecting the amount of work that is required for antitubercular drug discovery. This has led the drug discovery industry to move from selecting targets that are well-characterized to target-driven discovery of lead compounds [140, 147]. Currently, there are approximately only 40 bacterial targets that are being exploited for drug targeting leaving a potentially vast number of novel targets from other essential components to be discovered [148].

With target-based drug discovery, one of the major limitations is target selection [104]. Genome sequencing and *in silico* technology have provided a new dimension to the target selection paradigm. These technologies have not yet produced any new FDA approved antibiotics but have led to the identification of inhibitors to specific targets for lead optimization. It takes at least 12-15 years for antibiotic candidates to go through drug development to FDA approval [109]. Since most pathogenic genome data has become publicly available from 2000 onwards, it is unrealistic to expect many antibiotics from target-based drug discovery yet; however, the dawn for this era may very well be near. Another reason for the low number of FDA approved antibiotics from target-based drug discovery is due to the lack of investment in the field after the discovery of most antibiotics in the mid 1960s [105].

Some genomics-driven targets that had led to the identification of novel inhibitors are:

- The inhibition of polypeptide deformylase (PDF). Comparative genomics revealed that PDF is conserved in a broad-spectrum of pathogenic bacteria [135] and genomic technologies showed it to be an essential enzyme for protein biosynthesis and maturation [149]. PDF's are inhibited by actinonin, a naturally occurring antibacterial agent [135, 150]. This information of the lead compound has been used for rational drug design to synthesize derivatives of actinonin to increase potency and the spectrum of target organisms. Two of its derivatives, LBM415 [151] and BB-83698 [152], have reached phase III clinical trials [153].
- The inhibition of dihydrofolate reductase (DHFR). DHFR produces folate cofactors that are required for the synthesis of purines and some amino acids [154]. Genome sequences of DHFR have been used to produce over 100 crystal structures of the protein that has aided in the identification of its inhibitors. They are inhibited by diaminopyrimidines and the structural information has been used to synthesize a derivative of diaminopyrimidines, iclaprim, which has shown potency against trimethoprim, methicillin, vancomycin resistant pathogens and has undergone phase III clinical trials [155, 156].
- The inhibition of various enzymes involved in bacterial fatty acid biosynthesis (FAB) pathway, an essential metabolic process for bacterial survival [157]. Genomics led to the identification of all the enzymes involved in this pathway in clinically important pathogens and this information has been used to develop inhibitors via various rational drug development schemes:
  - diazaborines and triclosan - inhibitor of the enoyl-acyl carrier protein (ACP) reductase (FabI) [158]
  - cerulenin - inhibitor of beta-ketoacyl-ACP synthase I and II (FabB and FabF)
  - thiolactomycin - inhibitor of FabB, FabF and beta-ketoacyl-ACP synthase III (FabH)
  - platensimycin – inhibitor of FabF and FabB [159, 160]

Thus genomics and molecular techniques have revolutionized the drug discovery process and can lead to the identification of a collection of novel validated molecular targets. This, coupled with HTS can produce a range of structurally diverse compounds as leads to counteract disease and drug resistance.

### 3.2 Objectives of chapter three

The goal of this chapter was to prioritize potential broad-spectrum drug targets in *M. tuberculosis*.

There are various strategies that can be employed for prioritizing drug targets which are dependent upon the type of targets being selected for and the parameters set out for their identification [161]. The type of targets being selected for in this study was broad-spectrum targets. The aims to achieve this were:

- I. To identify all genes required for growth under optimal (nutrient-rich) conditions, experimental data from Sassetti *et al.* [162]
- II. To determine if the genes required for growth under optimal conditions are essential
- III. To distinguish if the identified essential genes were conserved and essential in other microbes by comparative genomics.

The genes identified in this study are referred to as putative essential targets as they will require further experimentation to support their essentiality in *M. tuberculosis* H37Rv. This list does not include genes that may be essential for latency, infection and/or genes that are specific to mycobacteria.

### 3.3 Results

#### 3.3.1 Quantification of essential genes in *M. tuberculosis* H37Rv

The foremost criteria for selecting essential genes in this chapter is based on the work by Sassetti, *et al.* [162]. The study by Sassetti identified 614 putative essential genes (Appendix 9.2) in *M. tuberculosis* H37Rv under optimal growth conditions via high density transposon mutagenesis [126]. Since the publication of the work by Sassetti, 58 of the 614 proposed essential genes have been mutated and deemed non-essential by other researchers. A comprehensive literature search was conducted to identify all the genes that have been considered non-essential from the genes identified by Sassetti. These genes are given in Table 3.1 together with their functional group and gene annotations.

##### 3.3.1.1 Genes mutated from the list published by Sassetti

One of the primary sources of *M. tuberculosis* mutants is the TARGET database at Johns Hopkins University. Of the 614 genes mentioned in the work by Sassetti, 48 were published as being able to sustain transposons in the TARGET database. Besides the TARGET database, other research groups have also managed to mutate some of the 614 genes via transposon mutagenesis and site-directed mutagenesis.

Consequently, after the literature search, a total of 58 genes were excluded from the 614 putative essential genes required for growth under optimal conditions (Table 3.1). This reduced the number of putative essential genes to 556. These remaining 556 genes were put through comparative genomics analysis (section 3.3.2) to identify the number of conserved essential genes required by *M. tuberculosis* for growth under optimal conditions.

**Table 3.1: List of genes mutated from the 614 putative essential genes published by Sassetti *et. al.***

Functional category	No. of genes mutated	H37Rv Annotations
a)		
Amino acid metabolism	6	Rv0118c, Rv0337c, Rv1005c, Rv1187, Rv1188, Rv1248c
Energy metabolism	2	Rv1448c, Rv2702
Lipid metabolism	5	Rv0400c, Rv0404, Rv2967c, Rv3139, Rv3564
Co-factor metabolism	3	Rv0425c, Rv1391, Rv3111
Transcription	5	Rv0339c, Rv0668, Rv1708, Rv3198c, Rv3201c
DNA replication/repair	2	Rv0949, Rv1420
Cell division	1	Rv2748c
Cell structure	2	Rv3793, Rv0112
Membrane translocation	6	Rv0411c, Rv0450c, Rv1348, Rv1456c, Rv2041c, Rv2833
Molecular chaperones/proteases	3	Rv0384c, Rv2110c, Rv3417c
Regulatory	2	Rv1027c, Rv3132c
Unknown functions	11	Rv0127, Rv0311, Rv0588, Rv0755c, Rv0875c, Rv0955, Rv1303, Rv1753c, Rv2739c, Rv2837c, Rv3660c
b)		
Amino acid biosynthesis	1	Rv2855
Co-factor metabolism	3	Rv2338c, Rv3109, Rv3112
Transcription	1	Rv0894
Cell structure	1	Rv3597c
Unknown functions	3	Rv1490, Rv3113, Rv3343c
c)		
Regulatory	1	Rv0982 [163]

a) - genes mutated via transposon mutagenesis at Johns Hopkins University

b) - genes mutated via transposon mutagenesis by McAdam *et. al.* [164]c) - genes mutated via site-directed mutagenesis by Pang *et. al.* [163, 165]

### 3.3.2 Quantification of conserved essential genes in *M. tuberculosis* H37Rv

The aim of this study was to identify a set of broad-spectrum targets thus gene conservation was taken into account. To isolate conserved genes that might be essential in *M. tuberculosis* H37Rv, I performed comparative genomics with the 556 genes from section 3.3.1 against the genes available at the Database of Essential Genes [146] using the criteria outlined in section 2.7.2. This database contains essential genes from 14 prokaryotes including *M. tuberculosis* H37Rv (Table 3.3).

Protein sequences of the 556 putative essential genes were used for protein-protein BLAST alignment against the DEG proteins. This confirmed 261 genes as fulfilling the cut-off (20% identity) of conserved genes in this study. Table 3.2 shows the list of conserved essential genes as prioritized drug targets in *M. tuberculosis* H37Rv together with their functional group.

**Table 3.2: List of broad-spectrum targets in *M. tuberculosis* H37Rv**

Functional category	No. of conserved essential genes	H37Rv Annotations
Amino acid metabolism	30	Rv0422c <sup>#</sup> , Rv0500, Rv1293 <sup>#</sup> , Rv1464 <sup>#</sup> , Rv1601 <sup>+</sup> , Rv1603 <sup>+</sup> , Rv1609, Rv1612 <sup>+</sup> , Rv1652 <sup>#</sup> , Rv1655 <sup>#</sup> , Rv1659, Rv1850 <sup>+</sup> , Rv2121c <sup>+</sup> , Rv2192c <sup>+</sup> , Rv2220, Rv2231c <sup>+</sup> , Rv2537c <sup>+</sup> , Rv2538c <sup>#</sup> , Rv2539c <sup>#</sup> , Rv2540c <sup>+</sup> , Rv2753c <sup>#</sup> , Rv2987c <sup>+</sup> , Rv2995c, Rv2996c <sup>#</sup> , Rv3025c, Rv3356c, Rv3708c, Rv3709c <sup>+</sup> , Rv3778c <sup>#</sup> , Rv3799c
Nucleotide metabolism	12	Rv0772, Rv0780, Rv0808, Rv1017c, Rv1381, Rv1389, Rv1699, Rv2883c, Rv2977c <sup>+</sup> , Rv3411c, Rv3752c <sup>#</sup> , Rv3913 <sup>#</sup>
Lipid metabolism	18	Rv0562, Rv0675, Rv0973c, Rv1018c, Rv1193 <sup>#</sup> , Rv1350, Rv1661, Rv1662, Rv2182c, Rv2500c, Rv2682c, Rv3140, Rv3285, Rv3332 <sup>#</sup> , Rv3398c, Rv3581c <sup>#</sup> , Rv3582c <sup>#</sup> , Rv3800c <sup>#</sup>
Co-factor metabolism	14	Rv0510, Rv0524, Rv0542c <sup>#</sup> , Rv0555 <sup>#</sup> , Rv0558 <sup>#</sup> , Rv1409 <sup>+</sup> , Rv1416 <sup>+</sup> , Rv2217, Rv2218, Rv2225 <sup>+</sup> , Rv3602c <sup>+</sup> , Rv3608c <sup>+</sup> , Rv3609c, Rv3713
Energy metabolism	38	Rv0526 <sup>#</sup> , Rv0527 <sup>#</sup> , Rv0529, Rv0462, Rv1023, Rv1092c <sup>#</sup> , Rv1098c, Rv1308, Rv1309, Rv1310, Rv1311 <sup>+</sup> , Rv1326c, Rv1392, Rv1436, Rv1437, Rv1438, Rv1449c, Rv1465 <sup>+</sup> , Rv1617, Rv1622c <sup>+</sup> , Rv1631 <sup>#</sup> , Rv1695 <sup>#</sup> , Rv1713, Rv1712 <sup>+</sup> , Rv1714 <sup>#</sup> , Rv2215, Rv2364c, Rv2421c, Rv2447c, Rv2786c, Rv3043c, Rv3255c <sup>#</sup> , Rv3257c, Rv3396c, Rv3436c, Rv3441c <sup>#</sup> , Rv3559c, Rv3634c

DNA replication/repair	13	Rv0001, Rv0005 <sup>#</sup> , Rv0058 <sup>#</sup> , Rv1547 <sup>+</sup> , Rv1629 <sup>#</sup> , Rv1701 <sup>#</sup> , Rv2343c <sup>#</sup> , Rv2986c <sup>#</sup> , Rv3014c, Rv3048c, Rv3051c, Rv3646c <sup>#</sup> , Rv3721c <sup>#</sup>
Transcription	12	Rv0667 <sup>#</sup> , Rv1297 <sup>#</sup> , Rv1390 <sup>+</sup> , Rv2444c <sup>+</sup> , Rv2511, Rv2703 <sup>#</sup> , Rv2710, Rv2841c, Rv2925c <sup>#</sup> , Rv3246c <sup>#</sup> , Rv3457c, Rv3907c
Translation	63	Rv0041, Rv0640, Rv0651, Rv0682, Rv0683 <sup>#</sup> , Rv0684, Rv0685, Rv0700, Rv0701, Rv0702, Rv0704, Rv0705, Rv0707, Rv0708, Rv0710 <sup>#</sup> , Rv0714, Rv0715, Rv0716, Rv0718, Rv0719, Rv0720, Rv0721, Rv1007c, Rv1014c, Rv1292, Rv1298, Rv1299, Rv1406, Rv1630, Rv1641, Rv1649, Rv1650 <sup>#</sup> , Rv1689 <sup>#</sup> , Rv2130c, Rv2357c, Rv2441c, Rv2442c, Rv2448c, Rv2534c, Rv2555c <sup>#</sup> , Rv2580c, Rv2614c, Rv2727c <sup>#</sup> , Rv2839c, Rv2845c, Rv2882c, Rv2889c <sup>#</sup> , Rv2890c, Rv2904c, Rv2906c <sup>#</sup> , Rv2992c <sup>#</sup> , Rv3009c, Rv3011c, Rv3105c <sup>#</sup> , Rv3336c, Rv3443c, Rv3458c <sup>#</sup> , Rv3459c, Rv3462c, Rv3579c <sup>#</sup> , Rv3580c, Rv3598c, Rv3834c
Cell division	9	Rv0014c <sup>#</sup> , Rv0015c, Rv0284 <sup>#</sup> , Rv2150c, Rv2154c, Rv2440c, Rv3610c, Rv3817c, Rv3918c
Cell structure	20	Rv0993 <sup>+</sup> , Rv1315 <sup>+</sup> , Rv2152c <sup>+</sup> , Rv2153c <sup>*</sup> , Rv2155c <sup>+</sup> , Rv2156c <sup>*</sup> , Rv2157c <sup>*</sup> , Rv2163c, Rv2165c <sup>#</sup> , Rv2245, Rv2246, Rv2361c, Rv2746c, Rv2881c, Rv2981c <sup>*</sup> , Rv3032, Rv3423c <sup>*</sup> , Rv3464, Rv3627c <sup>#</sup> , Rv3712 <sup>#</sup>
Membrane translocation	13	Rv0732 <sup>#</sup> , Rv1349, Rv1463, Rv2093c, Rv2397c, Rv2398c <sup>#</sup> , Rv2399c <sup>#</sup> , Rv2477c, Rv2832c, Rv2846c, Rv2916c, Rv3240c, Rv3781 <sup>+</sup>
Molecular chaperones/proteases	8	Rv0352 <sup>#</sup> , Rv0440, Rv1223, Rv2373c <sup>#</sup> , Rv2457c, Rv2460c, Rv3418c, Rv3596c
Unknown functions	11	Rv1229c <sup>#</sup> , Rv1461 <sup>#</sup> , Rv1462 <sup>#</sup> , Rv1711 <sup>#</sup> , Rv2166c <sup>#</sup> , Rv2188c, Rv2869c <sup>#</sup> , Rv2897c <sup>#</sup> , Rv3177 <sup>#</sup> , Rv3422c, Rv3921c

Further information on the known or predicted function of the individual conserved essential genes is given in the supplementary Table S1 (Appendix 9.2.2). Genes marked with a hash symbol (#) were also identified as targets by Raman *et al.* [140], genes marked with a plus sign (+) were identified as targets by Anishetty *et al.* [166] and genes marked with an asterisk (\*) were identified as targets by both the above mentioned studies (section 3.4.3).

### 3.3.3 Orthologs of conserved essential genes in *M. smegmatis*

The model organism used in this study for experimentation was *M. smegmatis* mc<sup>2</sup>155; therefore orthologs of all the genes in Table 3.2 were searched for and identified in *M. smegmatis*. The *M. smegmatis* gene annotations corresponding to the *M. tuberculosis* genes are also provided in supplementary Table S1 (Appendix 9.2.2).



### 3.4 Discussion

This chapter highlights the importance of target identification in the drug discovery process. It presents a list of broad-spectrum candidates in *M. tuberculosis* generated by comparative genomics. This list consists of conserved essential genes thus providing a starting point for drug discovery for broad-spectrum antibiotics.

#### 3.4.1 Essential genes

The best targets are genes that are essential for the cellular maintenance of the microbe [104]. Thus the first step in identifying potential targets in *M. tuberculosis* in this study was determining gene essentiality. The gold standard for identifying essential genes is to determine if the gene can be knocked out [167]. Thus, in accordance with the first and second condition discussed in the aims of this chapter, we looked at the data presented by Sassetti *et. al.* [162]. This study identified 614 essential genes in *M. tuberculosis* via high density transposon mutagenesis and transposon site hybridization (TraSH) that could not be knocked out under optimal growth conditions [107].

The justification for looking at genes under optimal conditions was to uncover the inventory of housekeeping genes that is responsible for essential biosynthetic, metabolic and regulatory processes, which may be all that is required by a cell under optimal conditions when all the mandatory nutrients are provided. If this is the case and these genes are indispensable under favourable conditions, the likelihood of them being essential under all growth conditions, that is, nutrient deprived conditions, is high as opposed to conditionally essential genes (section 1.7.2). Targeting conditionally essential genes may not be as effective in therapy as constitutively essential genes because if the metabolic state of the cell changes, the conditional genes may not be required leaving the drug without a target and rendering therapy useless. Apart from genes required for growth, it would also be ideal to have a set of targets for the latent bacilli where

the bacteria is not growing actively, however, latency is very poorly understood and latent bacilli do not cause active disease therefore for the scope of this study, the emphasis was on targeting genes that are essential for growth.

The study by Sasseti *et. al.* also compared the 614 essential genes identified under optimal conditions against the genome sequence of *Mycobacterium leprae* and found that 78% of these genes were retained in *M. leprae* [162]. This has huge implications as *M. leprae* has been found to have undergone reductive evolution, shedding a lot of its genes and only keeping the genes that are critical for its survival. This has led the *M. leprae* genome to be about half the size of the *M. tuberculosis* genome [168]. Because of this genome reduction, only 40% of the *M. tuberculosis* genes are conserved in *M. leprae* [169]. Thus a 78% conservation of the 614 genes in *M. leprae* identified as being essential in *M. tuberculosis* strengthens the hypothesis that these genes may very well be essential in *M. tuberculosis* [137].

Since the focus of this study was in the identification of broad-spectrum targets and in accordance with the third condition of this chapter, a comparative analysis of the genes identified as being essential in *M. tuberculosis* (section 3.3.1) against other more divergent bacterial species (Table 3.3) available at DEG was performed.

### 3.4.2 Broad-spectrum targets

Protein sequences between *M. tuberculosis* and the microbes available at DEG (Table 3.3) were compared as these are more analogous than primary sequence due to codon degeneracy [147]. The rationale for comparing the putative protein sequences of essential mycobacterium genes against the protein sequences of essential genes in other bacterial species was to identify the ubiquitous essential genes as potential broad-spectrum targets.

Several genome comparison studies have revealed that a very small number of genes maybe required for a pathogen to survive [67, 170]. Mushegian *et al.* compared the smallest bacterial genome, *Mycoplasma genitalium* with another bacterial genome, *Haemophilus influenzae* [170]. This study found that over 70% of the genes, that is, a total of 240 genes, in the minimal gene set derived from the sequences of *M. genitalium* and *H. influenzae* are conserved, [111, 170]. These two genomes are relatively small; however in this study we found 261 genes that were common between 14 different bacterial species (Table 3.2). This follows a pattern in that the number of conserved genes does not vary much between different genome sizes. However, several studies tried to validate the genes identified as being essential by Mushegian *et al.* [171-173] and found that there is a high level of redundancy due to non-orthologous gene displacement. This phenomenon is portrayed when the same function is fulfilled by an related or paralogous protein in different organisms [174]. This places emphasis on the need for target validation.

The importance of target validation was demonstrated in another study where 347 conserved genes were being validated for essentiality and as potential broad-spectrum targets in *Streptococcus pneumoniae* via high-throughput gene disruption. Of the 347 conserved genes, only 113 were presumed as being essential as they were unable to grow after gene disruption [161]. Comparative genomics studies like these have shown that there might be approximately 250 to 350 essential broad-spectrum targets [67, 161]. However, these figures may be an over estimation as target validation experiments demonstrate that not all genes identified as being essential is actually essential as was shown for *S. pneumoniae*. Validation studies may also disprove some of the genes shown as being essential in DEG (Table 3.3) as these organisms were analysed using whole genome analysis and such analysis techniques are prone to some degree of error. This was demonstrated in this study with *M. tuberculosis* where some of the genes identified as being essential by Sassetti *et. al* via a whole genome analysis were in fact found to be non-essential by other researchers (Table 3.1).

**Table 3.3: Organisms in DEG**

Organism	No. of essential genes	Reference
<i>Acinetobacter baylyi</i> ADP1	499	[175]
<i>Bacillus subtilis</i> 168	271	[176]
<i>Escherichia coli</i> MG1655	712	[177, 178]
<i>Francisella novicida</i> U112	392	[179]
<i>Haemophilus influenzae</i> Rd KW20	642	[171]
<i>Helicobacter pylori</i> 26695	323	[180]
<i>Mycobacterium tuberculosis</i> H37Rv	614	[162]
<i>Mycoplasma genitalium</i> G37	381	[181]
<i>Mycoplasma pulmonis</i> UAB CTIP	310	[182]
<i>Pseudomonas aeruginosa</i> UCBPP-PA14	335	[183]
<i>Salmonella typhimurium</i> LT2	230	[184, 185]
<i>Staphylococcus aureus</i> N315	302	[1, 186, 187]
<i>Streptococcus pneumoniae</i>	244	[161]
<i>Vibrio cholerae</i>	5*	[188]

\* The 5 genes identified as being essential in *V. cholerae* is not a true representative of all the essential genes present in this organism as this study only characterized a total of 16 mutants.

Another study that used the same criteria as defined in our study found 306 essential genes in *Pseudomonas aeruginosa* [189]. This is slightly higher than what was identified in this study with *M. tuberculosis* however, their cut-offs for identifying homologs in the DEG database was lower than the cut-offs set in the current study.

This current study also performed a comparative analysis of the 261 conserved essential targets from *M. tuberculosis* H37Rv (Table 3.2) with the *M. smegmatis* mc<sup>2</sup>155 genome and found that all these proteins were present in *M. smegmatis* with high amino acid identity (supplementary Table S1- Appendix 9.2.2). This provides additional justification that these genes may be required for essential cellular processes in mycobacteria.

### 3.4.3 Comparison to other studies in *M. tuberculosis*

Irrespective of the criteria chosen to identify targets by any researcher, the end goal is to achieve a list of targets whose impairment will lead to the eradication of the microbe [95]. A good example of using different criteria to generate distinct target lists is given in a study by Hasan *et. al.* where three different lists of prioritized drug targets for *M. tuberculosis* were generated, each conformed to a different criterion [137]. This study used all of the MTB genes as targets whereas our study gives a more focussed list of targets where some genes were eliminated as potential targets. Thus, the list of genes from Table 3.2 was compared to other studies that followed a similar method of identifying targets where genes not falling under the selected criteria were excluded.

This analysis revealed that about 34% of the genes identified in this study were also established as targets by another research group [140]. This study used a number of different filters to ours to eliminate genes at various stages of analysis. It also looked at genes that were specific to *M. tuberculosis* and related to persistence and resistance as well as broad-spectrum targets and came up with a total of 451 high-confidence targets. In contrast, our analysis only identified 261. However, this is to be expected as our list only contains broad-spectrum targets, addition of mycobacterial specific and/or persistence genes would most certainly increase this figure.

A second study that looked at potential drug targets in *M. tuberculosis* confirmed 13% of the genes identified in this study as targets [166]. However, this study by Anishetty *et. al.* only performed a comparative analysis of all the enzymes in *M. tuberculosis* from the Kyoto Encyclopedia of Genes and Genomes (KEGG) metabolic pathway database against proteins from humans. In doing so, they discounted any potential target that is not an enzyme or may be an enzyme but is not yet characterised in the KEGG database. This may account for the low percentage of comparable targets between thier study with our study.

It is also worth noting that some of the genes identified in our study (Table 3.2) already had their cognate inhibitors identified and are currently used for Tb treatment. Rv0667 (RpoB) is the known target for rifampicin [190], one of the main first-line anti-tubercular drug and Rv0683 (RpsL) is a known target for streptomycin [191], another first-line anti-tubercular drug. Rv2981c (Ddl) and Rv3423c (Alr) are known targets for D-cycloserine, a second-line Tb drug [192] and Rv0005 (GyrB) is the known target to fluoroquinolones [68], another second-line drug. The fact that this method identified already known targets of currently used anti-tubercular drugs proved that the method used in this study extracted creditable targets for future drug development.

#### **3.4.4 Uniqueness of this study**

There are a few ideas in this study that deviate from the routinely used bioinformatics checklist for target identification. Firstly, the targets identified in this study were not compared for human analogues as it is common to find some degree of homology between bacterial enzymes against the eukaryotic genome [67]. Important as this is, there have been reported cases where despite a significant homology between the eukaryotic and prokaryotic proteins, drugs did not have the same level of toxicity. For example the staphylococcal gene encoding isoleucyl-tRNA synthetase is 38% identical to its human homolog, falling well within the range to inhibit both prokaryotic and eukaryotic proteins however; mupirocin, an analog of isoleucine which competitively inhibits amino acid activation by isoleucyl tRNA synthetase and thus inhibits protein synthesis [193], has been successfully used as an antibiotic for many years [193, 194]. This is due to two amino acid substitutions in the active site weakening the mupirocin interaction in humans and preventing the inhibition of the mammalian enzyme [195]. This proves that antibiotic selectivity cannot be based on protein sequence homology alone and single amino acid substitutions can have significant changes to antibiotic binding affinities [67]. An additional example of this is given by the inhibition of bacterial DHFR but not the highly similar human enzyme by trimethoprim and its derivatives. [196, 197].

---

Another common feature used in identifying targets is comparing it to metabolic and/or structural databases. However, with this method, unusual proteins underrepresented in the databases or proteins of unknown functions are eliminated. This results in missing out in the identification of completely novel functions that may constitute potential new antibacterial targets with unique MOA and/or new classes of antibiotics, something that we are in the hunt for.

### 3.4.5 Summary

Notably, bioinformatics is routinely used for the discovery, selection and prioritization of drug targets to combat disease. This chapter presents a focussed list of broad-spectrum targets for *M. tuberculosis* identified via bioinformatics with support from experimental data by various researchers. It also illustrates the functional classes of these genes. This provides the basis for selecting unique targets and/or pathways to identify antibiotics with novel mode of action to overcome current resistant mechanisms.

Such step-wise prioritization and elimination of genes using simple biological criteria is an efficient way to reduce the number of genes to an experimentally manageable level. These genes can be validated in more traditional low-throughput methods and the results can be used to make informed decision on where investment should be made for drug development.

## **CHAPTER FOUR**

### **Target Validation**



## 4.1 Introduction

Recent drug discovery has been driven largely by genomics-based approaches. The primary barrier to drug discovery prior to the availability of genomic data was the small number of potential drug targets that were characterized. Genomics has removed this target identification barrier from the few characterized genes to the whole genome [198]. The targets identified via *in silico* genomics have to be validated to determine if their alteration is going to have a therapeutic effect [109]. This includes the exploration of target gene function which in turn can be used to develop suitable assays for high-throughput antimicrobial screening. These rational target-based screening programs combine genomic data with genome-based technologies to isolate and/or validate potential targets [67]. One of the most commonly used methods for modulating target protein function is blocking gene expression at the RNA level, a method that was implemented in this study.

### 4.1.1 Approaches to target validation

Essential and conserved genes make excellent drug targets as their conservation signifies some canonical function required for growth or survival as opposed to species specific roles [114]; however, target selection requires genetic manipulation for validation. There are various strategies applied for target validation [161, 171, 185, 186, 199] (section 1.7.3 and Table 3.3), all of which use the genomic data that has become available to analyse the essential processes required for growth.

Gene disruption validates gene essentiality where loss of function is associated with loss of viability [67] but this technique does not give the functional annotation of the targets as mutation of essential genes is lethal to cells [200]. Some gene disruption techniques can also be misleading as they may have polar effects on the downstream genes if the target is present in a polycistronic operon. If the downstream genes are essential, such conditions will lead to a null

phenotype that is not related to the target. This would require further genetic manipulation to overcome, thus making it unsuitable for genome scale high-throughput analysis and inferior to other techniques like allelic mutagenesis [67, 201].

Since lethal mutants cannot be analysed further, controlled gene expression studies have become a useful tool for studying essential genes [185]. Gene expression studies allows the regulation of the target where it can be titrated up or down, giving more perspective on gene functionality based on the phenotype achieved from these titrations [202]. Such studies can expose the cellular pathways of genes with unknown functions and most importantly these studies can subsequently be used in HTS to find inhibitors against the target.

Gene regulation studies include the incorporation of inducible promoters used for direct modulation of target gene expression and antisense technology for indirect modulation of target gene expression (section 4.1.1.1). With direct gene modulation, conditional mutants can be created either by replacing the native promoter with an inducible promoter or an inducible promoter can be used to control a complementary copy of the gene while the gene is removed from the genome. Both of these systems can be used to regulate the level of gene expression in a dose-dependent manner with respect to the inducer [114]. The most commonly used of these inducible systems are based on bacterial response regulators that bind to specific operator sequences which can be modulated to induce or repress gene expression with specific inducing molecules [203].

#### **4.1.1.1 Antisense Technology**

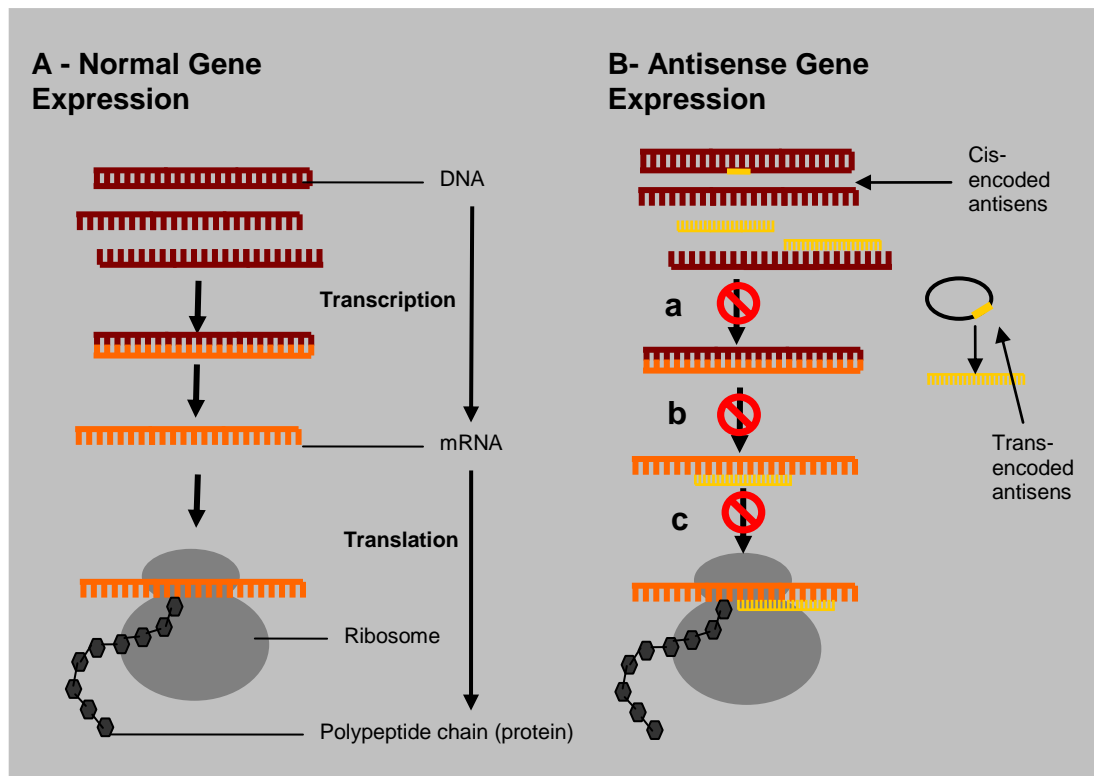
The key feature of molecular biology consists of genetic information being transcribed from DNA into RNA which is then translated into proteins (Figure 4.1A). Antisense technology exploits these two essential processes. Antisense gene regulation is a sequence-specific inactivation of protein production without affecting the gene. In nature, there are cis-encoding and trans-encoding antisense

---

molecules that regulate gene expression [204]. There are three different mechanisms via which antisense inhibition works in bacterial cells (Figure 4.1B). The most common and natural mechanism involves binding of the antisense molecule to complementary messenger RNA (mRNA) forming a duplex. This prevents the translation of the protein and leads to the degradation of the double-stranded RNA (dsRNA) by dsRNA-specific-RNases [205]. The second mechanism involves the dsRNA covering the ribosome recognition sites and causing steric hindrance at the ribosome and thus preventing translation [204]. The third mechanism involves the formation of DNA:RNA heterodimers which are degraded by RNases [204, 205].

Antisense inhibition provides a means to down-regulate gene expression giving rise to loss of function phenotype thus this technology can also be used for functional annotation. However, the efficiency of loss of function depends on the length, structure and intracellular concentration of the antisense molecule [206]. This technique has been used on single essential genes for functional annotations in a number of microbes including the *hisD* gene in *M. smegmatis*, the *ahpC* gene in *M. bovis*, the *sgp* gene in *Streptococcus mutans* and the *srrAb* in *Staphylococcus aureus* [200, 207-209]. It has also been used with a whole genome for the identification of the conditional essential gene expression profile in *Staphylococcus aureus* [185, 186]. Successful use of antisense technology has also been applied to identifying inhibitors of specific targets in HTS [159]

The molecular and genetic functions of many essential genes have not yet been studied as deletion mutants are difficult to study making conditional mutants a necessity for the study of essential genes [123]. Antisense is a commonly used technique for such studies as it removes a lot of laborious lab work which is associated with creating genome mutations. Therefore, this study utilized antisense RNA expression to validate some of the targets identified via the bioinformatics analysis for essentiality and viability.



**Figure 4.1: Gene expression**

**A)** Normal gene expression in a prokaryotic cell. Double-stranded DNA (dsDNA) duplex unfolds and gets transcribed into mRNA and the mRNA gets translated into protein at the ribosome. **B)** Mechanisms of gene regulation using the expression of antisense RNA. This can be either achieved using a cis-encoded antisense sequence or a trans-encoded antisense sequence a) a vector containing the antisense sequence of the gene of interest is integrated at a pre-determined site in the genome to develop a conditional mutant. This system uses the normal cellular machinery and the inducible promoter from the integrated vector to transcribe antisense RNA. This is complementary to the sense strand of the DNA and binds to ssDNA upon replication thus preventing transcription and mRNA formation which in turn leads to decrease in protein production. b) The antisense molecule is provided in trans in a vector; once again the vector uses cellular machinery to transcribe antisense RNA which is complementary to the mRNA thus binds to it forming dsRNA and prevents translation and protein production as dsRNA is unstable and gets degraded. c) dsRNA that reaches the ribosome is unable to fully transcribe as it encounters steric hindrance, again preventing translation thus leading to a decrease in protein production.

#### 4.1.1.2 Cell division

The focus of this thesis is in validating drug targets that are essential in the cell division process. A step-by-step progression of cell division involves DNA condensation and replication, chromosome partitioning to opposite poles of the cells, Z-ring assembly and septum formation at midcell followed by cytokinesis to generate daughter cells with identical genetic material [210-212]. This is a complex process necessitating the need for spatial and temporal coordination of an array of proteins however little is known about the factors regulating these processes in mycobacteria [210, 212-215]. Thus it was hypothesized that elucidating the fundamental genes involved in cell division may present novel targets with alternate MOA to current Tb treatment as cell division proteins are underexploited targets for drug discovery.

### 4.2 Objectives of chapter four

The goal of this chapter was to validate some of the conserved essential genes identified in chapter 3 that were involved in the cell division process as potential anti-tubercular targets. This would be achieved by assessing if down-regulation of these genes has any effect on the viability of the cells.

It was hypothesized that decreasing an essential gene product will affect a cells survivability as an essential gene product signifies that it is absolutely vital for the cell [167].

To explore this hypothesis, 3 different *M. smegmatis* genes identified as being essential for cell division were selected. These were cloned into a plasmid in the antisense orientation under a tetracycline inducible promoter upstream of a GFP sequence to develop a trans-encoded antisense RNA expression system. All strains were subjected to growth assays under different culture conditions and monitored for:

- 
- GFP expression; to ensure antisense induction
  - Optical density (OD); to monitor growth rate
  - CFU; to monitor the number of separable and viable cells
  - Cell morphology; to observe any discernable phenotypic difference upon antisense induction

The outcomes from the above investigations were used to discern if the putative essential genes tested in this work were viable candidates for targeted drug screening.

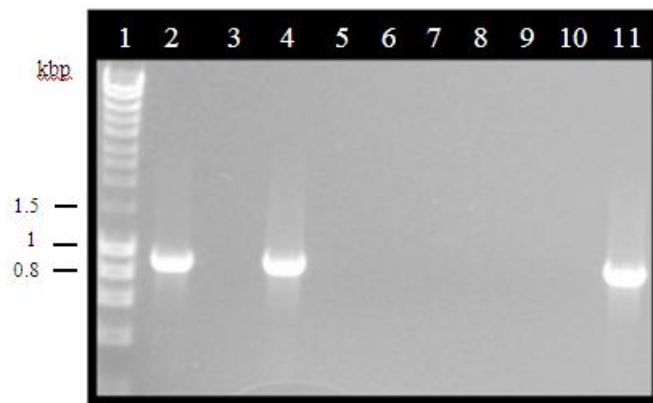
## 4.3 Results

### 4.3.1 Construction of expression vectors

All PCR products were cloned into the pCR<sup>®</sup>2.1 TA cloning vector (Appendix 9.1A) and transformed into *E. coli* DH5 $\alpha$  cells. Transformants were selected and plasmids were isolated from these clones and sequenced. The correct sequences were then digested out of pCR<sup>®</sup>2.1 using appropriate restriction enzymes and cloned into equally restricted pMind (Appendix 9.1D).

This study utilized the tetracycline inducible vector, pMind (Appendix 9.1D) for antisense gene expression. pMind uses a tetracycline-regulated efflux system isolated from a *Corynebacterium glutamicum* plasmid. It contains a transcriptional repressor (TetR) and an operator region (*tetO*) containing the promoter sequence and repressor binding site [120]. The repressor protein sterically prevents transcription by binding to the operator sequence between the inducible promoter and the gene product in the absence of the inducer (tetracycline). Transcription can be de-repressed by the addition of tetracycline which binds to the repressor and prevents it from binding to the operator.

This vector was used as described by Blokpoel *et. al.* [120] with a few modifications. To monitor expression from the promoter upon induction with tetracycline, a *gfp* sequence was amplified from the pHS201 vector [121] using the primers GFP\_RBS\_F and GFP\_R (Table 2.4) and cloned into the *Spe*I site (MCS) to act as a reporter. Since both the forward and reverse primers had the same enzyme sites, an orientation PCR was carried out to ensure that the *gfp* sequence had been inserted in the correct direction using the primers TetROF11 and GFP\_R (Figure 4.2; section 2.9.5.3). Clones with the correct inserts were isolated and this plasmid was called pMindGFP. This was electroporated into *M. smegmatis* and used as a control in later experiments.

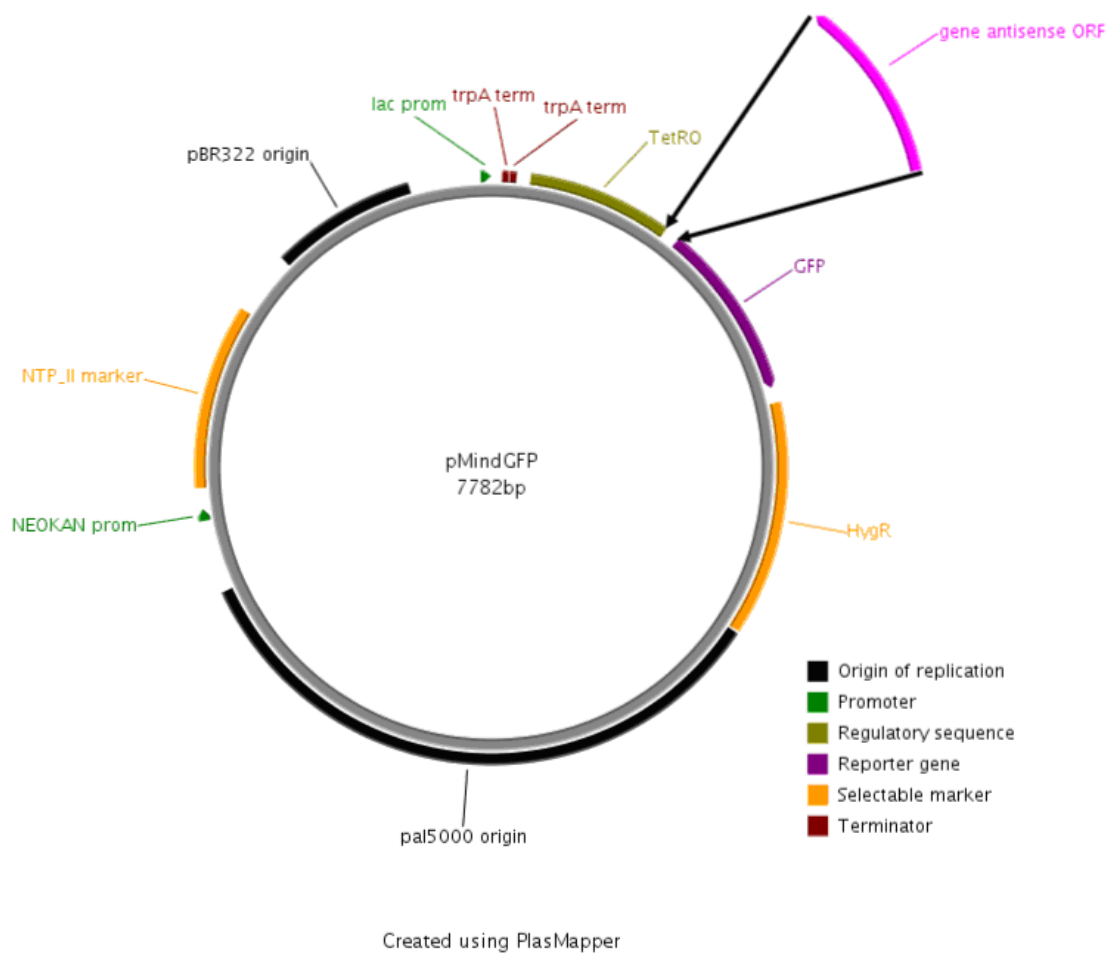


**Figure 4.2: Orientation PCR of pMindGFP**

1% agarose gel showing PCR product of pMindGFP amplified with primers TetROF11 and GFP\_R with lane 1 showing HyperLadder I and lanes 2, 4 and 11 showing positive insertion of the *gfp* sequence in the correct orientation. The *gfp* sequence is ~ 750 bp. The insert shown in the gel is ~ 850 bp since a 100 bp sequence was added onto the PCR product by the TetROF11 primer as this is an external primer, that is, from the plasmid outside of the GFP sequence (Figure 4.4).

#### 4.3.1.1 Construction of antisense vectors

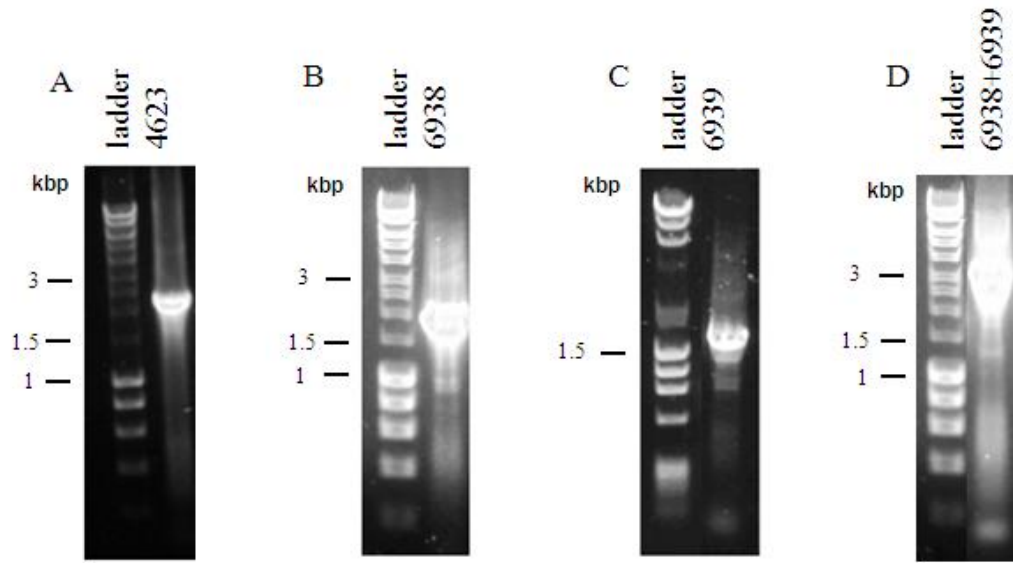
All genes of interest were amplified from *M. smegmatis* mc<sup>2</sup>155 genomic DNA using their respective primers (Table 2.4) and sub-cloned in the antisense orientation in either pMind or pMindGFP from pCR<sup>®</sup>2.1 (section 4.3.1). The inserts in the vectors were verified with various primer sets and verified plasmids were electroporated in *M. smegmatis*.



**Figure 4.3: pMindGFP vector**

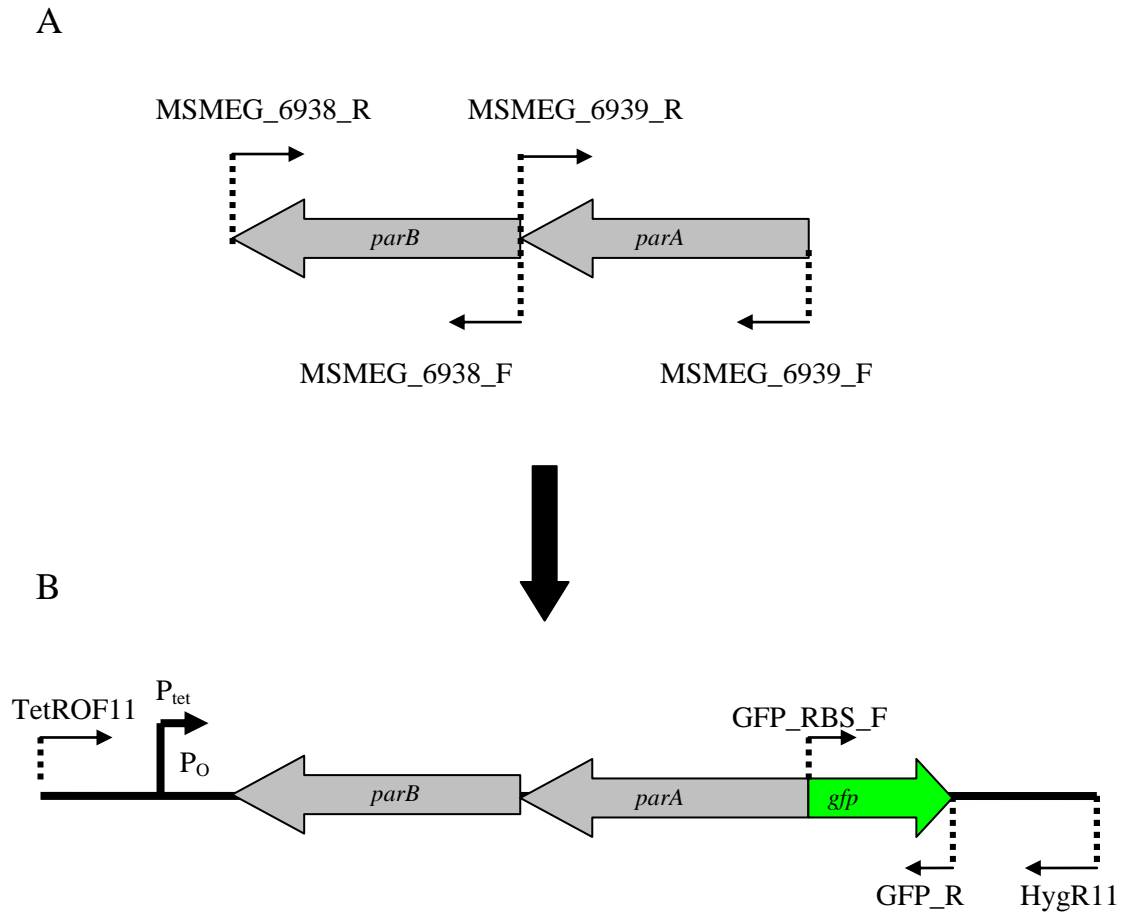
pMindGFP vector showing the placement of the ORF for the target gene in the antisense orientation (pink) upstream of the GFP reporter (purple). Vector created using PlasMapper [216].





**Figure 4.4: PCR product of genes cloned into pMind and pMindGFP**

1% agarose gels showing DNA ladders and PCR products derived from *M. smegmatis* that were cloned in the antisense orientation into the pMind and pMindGFP vectors. The PCR products are slightly larger than the gene size as the external primer and the GFP gene added more sequence to the *M. smegmatis* gene products. **A)** PCR product from pMind\_ogb with primer set TetROF11 and MSMEG\_4623\_F (gene size ~ 1.5 kbp). **B)** PCR product from pMind\_parB\_GFP with primer set MSMEG\_6938\_R and GFP\_R (gene size ~ 1 kbp). **C)** PCR product from pMind\_parA\_GFP with primer set MSMEG\_6939\_R and GFP\_R (gene size ~ 1 kbp). **D)** PCR product from pMind\_parA+parB\_GFP with primer set MSMEG\_6938\_R and GFP\_R (combined gene size of set *parA* and set *parB* ~ 2 kbp). These were also verified with external primers TetROF11 and HygR11 (data not shown) and with DNA sequencing. The reason for using the same primer sets for the verification of the *parB* and MSMEG\_*parA*+*parB* fragments was because of the gene arrangement of *parA* and *parB*. Figure 4.5 illustrates this gene arrangement and the position of all the primers. The cloning sites used for each of these fragments are given in Table 2.2.

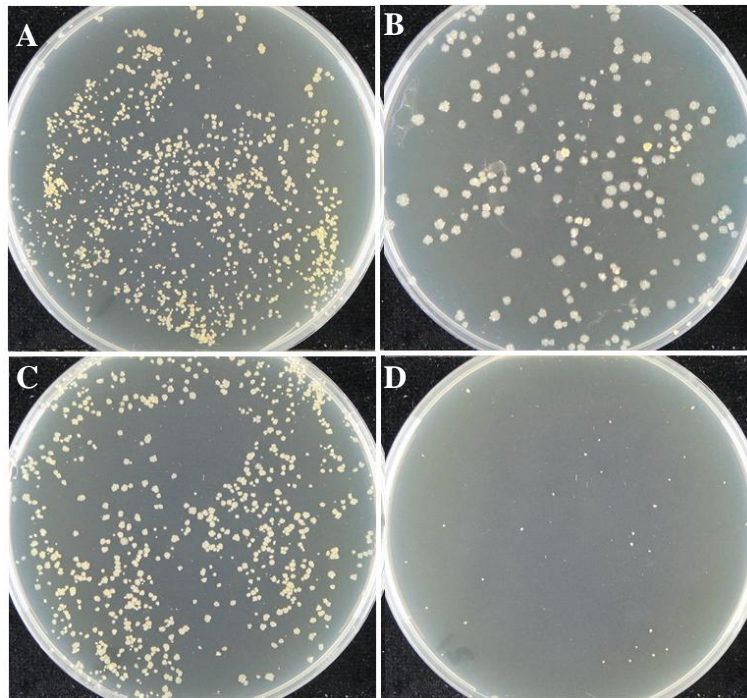


**Figure 4.5: Gene arrangement of *parA* and *parB***

**A)** This diagram illustrates the gene arrangement of *parA* and *parB* as in the mycobacterium genome and the primer binding sites of the various primers used in this study for the gene amplification, PCR screening and DNA sequencing. **B)** This diagram shows the two genes (*parA* and *parB*) cloned in the antisense orientation in the pMind vector under the tetracycline inducible promoter ( $P_{tet}$ ) upstream of the operator sequence ( $P_o$ ). The *gfp* sequence was cloned downstream of the gene sequences. It also shows the primer binding sites for GFP (GFP\_RBS\_F and GFP\_R) and the external primers used for orientation PCR (TetROF11 and GFP\_R) and DNA sequencing (TetROF11 and HygR11). This vector was called pMind\_*parA+parB*\_GFP.

### 4.3.2 Electroporation into *M. smegmatis*

All the above constructs were electroporated in *M. smegmatis* (section 2.9.4.2), plated on agar plates with appropriate antibiotics for transformant selection and incubated at 37°C. All plates had 1-3 mm colonies after 72 hours except for the plate containing *M. smegmatis* electroporated with the pMind\_*parA*+*parB*\_GFP vector. Figure 4.6 shows the difference between the electroporation plates of strains that had the vectors pMindGFP, pMind\_*parB*\_GFP, pMind\_*parA*\_GFP and pMind\_*parA*+*parB*\_GFP. All strains created were used in antisense expression assays except *M. smegmatis*-pMind\_*parA*+*parB*\_GFP. This strain was further characterized (section 4.3.7).



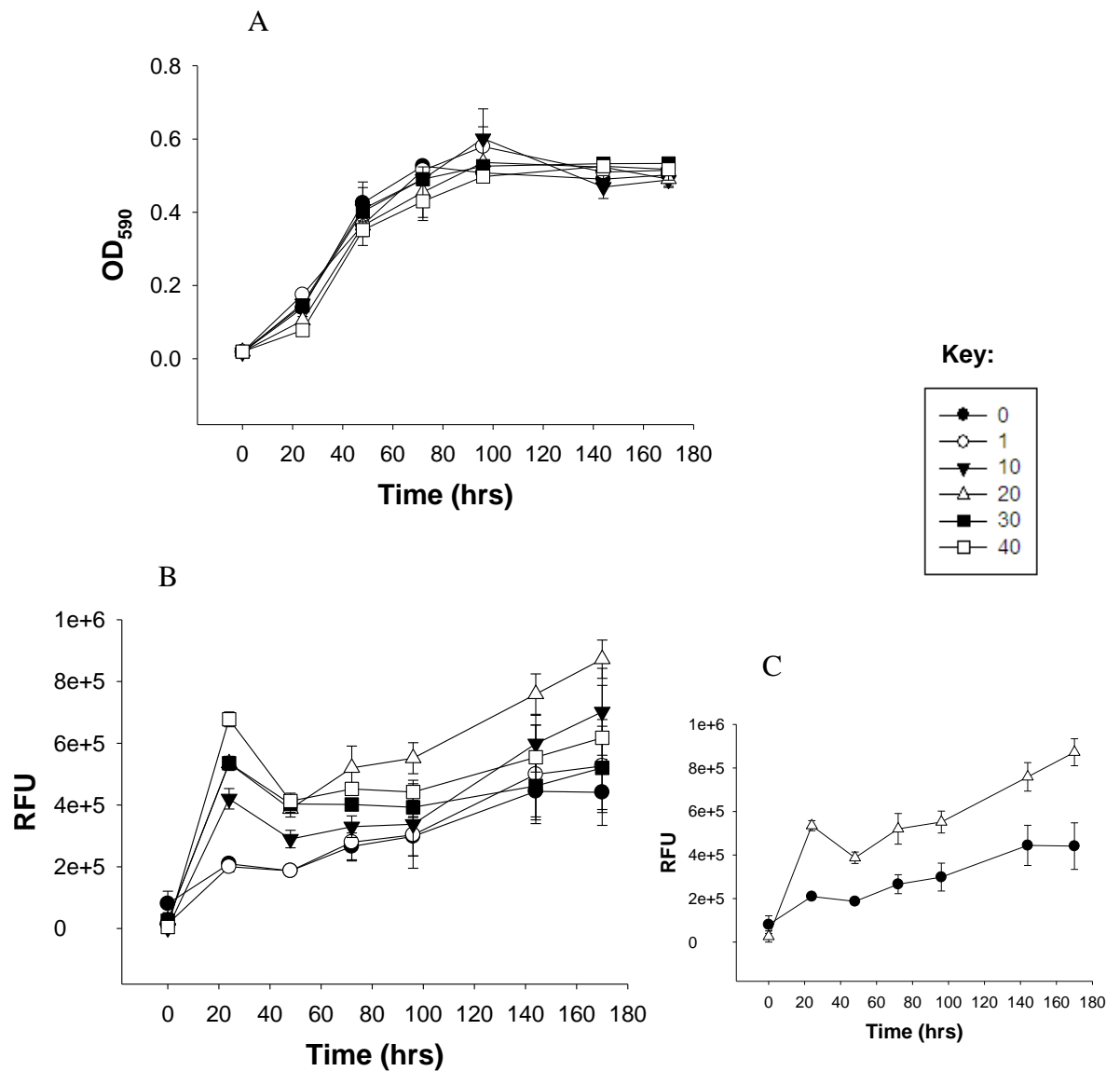
**Figure 4.6: Electroporation of *M. smegmatis***

These plate photos were taken at 72 hours showing *M. smegmatis* electroporated with A) pMindGFP, B) pMind\_*parB*\_GFP, C) pMind\_*parA*\_GFP and D) pMind\_*parA*+*parB*\_GFP. The growth of the *M. smegmatis*-pMind\_*parA*+*parB*\_GFP strain was severely inhibited when both *parA* and *parB* genes were placed in the antisense orientation in the same vector. This was performed in triplicate.

### 4.3.3 Determination of optimal tetracycline concentration

Before the strains created in this study could be used in antisense RNA expression assays, the optimal inducer concentration had to be established. This was determined by conducting growth assays of *M. smegmatis*-pMindGFP in rich media (LBT) (section 2.12.1.1). A titration of 0–40 ng/mL of Tc was used in this assay and the OD<sub>590</sub> and GFP emission were monitored every 24 hours (Figure 4.7).

As can be seen in Figure 4.7A, there were no obvious differences between the OD<sub>590</sub> of the cultures containing different concentrations of Tc indicating that Tc concentration as high as 40 ng/mL could be used in the assays however, the GFP signal dropped at high Tc concentrations (Figure 4.7B). The GFP signal was used to determine the concentration of Tc to be used for antisense induction as it is a more sensitive measure than OD. From these results, 20 ng/mL of Tc was considered as the optimal concentration to achieve maximum expression from the Tet promoter and was used for all future experiments unless otherwise stated.

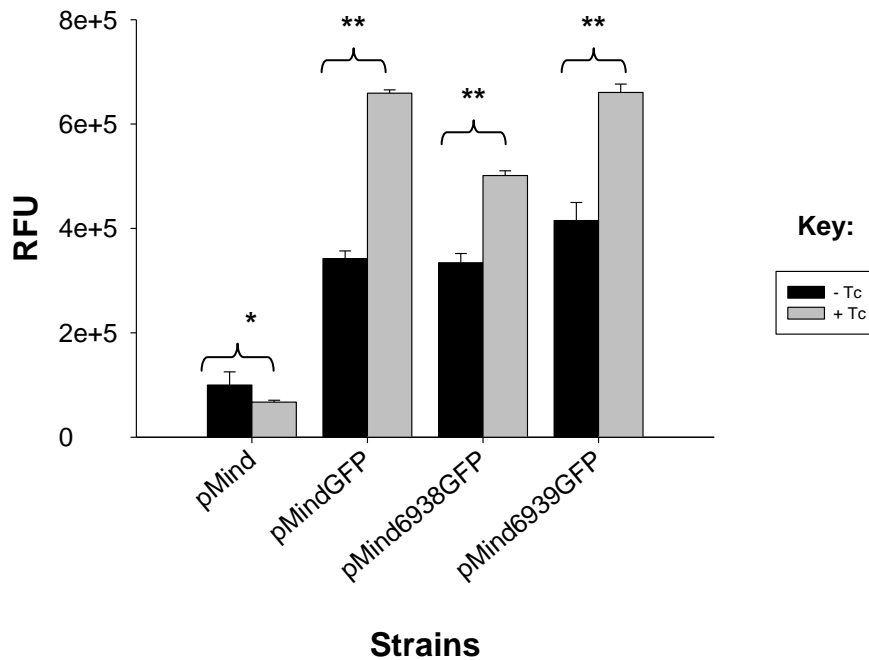


**Figure 4.7: Determination of optical inducer concentration**

Growth curve of *M. smegmatis*-pMindGFP in LBT with varying concentrations of Tc at different time points plotted against **A**) OD<sub>590</sub> and **B**) relative fluorescence units (RFU) and **C**) RFU difference between 0 ng/mL and 20 ng/mL. Statistical analysis using a two-tailed T-test showed an insignificant difference in fluorescence with all concentrations of Tc compared to 0 ng/mL except 20 ng/mL ( $p = 0.05$ ; Figure 4.7C). Results from four independent experiments with triplicates always showed 20 ng/mL of Tc with the highest GFP expression level. All results are representative of mean values of triplicates  $\pm$  SE.

#### 4.3.4 Induction of antisense transcript

To prove that the antisense transcript was being transcribed by the promoter, all strains harbouring the *gfp* sequence were cultured in LBT with or without Tc for 72 hours. The *gfp* sequence was cloned downstream of the antisense fragments (Figure 4.4) thus expression of GFP would indicate that the antisense transcript was being produced. *M. smegmatis*-pMind was used as a negative control. As can be seen from Figure 4.7, there was incomplete repression of the promoter in the absence of the inducer. However, there was increase in GFP expression upon induction of the promoter.



**Figure 4.8: Induction of antisense transcript**

This graph shows *M. smegmatis* electroporated with various pMind vectors (indicated on the graph) in the absence (■) and presence (□) of tetracycline. The strains with the pMind vector showed no change in fluorescence with or without Tc. All the strains with the pMindGFP vector show some fluorescence in the absence of the tetracycline (incomplete repression of the promoter) but this increased upon the addition of tetracycline showing that there is expression being induced from the promoter. \* represents p value > 0.05 and \*\* represents p value < 0.05. All results are representative of mean values of triplicates ± SE.

### 4.3.5 Antisense assays

Growth assays with different culture conditions were carried out (section 2.12.1) with the various strains created in this study (Table 4.1) to determine if down regulation of the conserved essential genes had any effect on cell growth, viability and morphology. All cultures were analysed by measuring the OD every 24 hours for 11 days. CFU and cell morphology were analysed at time 0 and at late exponential phase as the cultures were entering stationary phase. Two independent experiments were carried out for each assay; however, results are shown from just one experiment. All statistical analysis was carried out on data from the late exponential phase when the number of CFU were determined and cell morphology was analysed. All data is compared to the *M. smegmatis*-pMindGFP control.

**Table 4.1: Strains used for antisense assays**

Strains	Annotation
<i>M. smegmatis</i> -pMindGFP	<i>gfp</i>
<i>M. smegmatis</i> -pMind_ <i>obg</i>	<i>obg</i>
<i>M. smegmatis</i> -pMind_ <i>parB</i> _GFP	<i>parB</i>
<i>M. smegmatis</i> -pMind_ <i>parA</i> _GFP	<i>parA</i>

This table shows all the strains that were used in the antisense expression assays and the annotations used to denote these strains in the graphs.

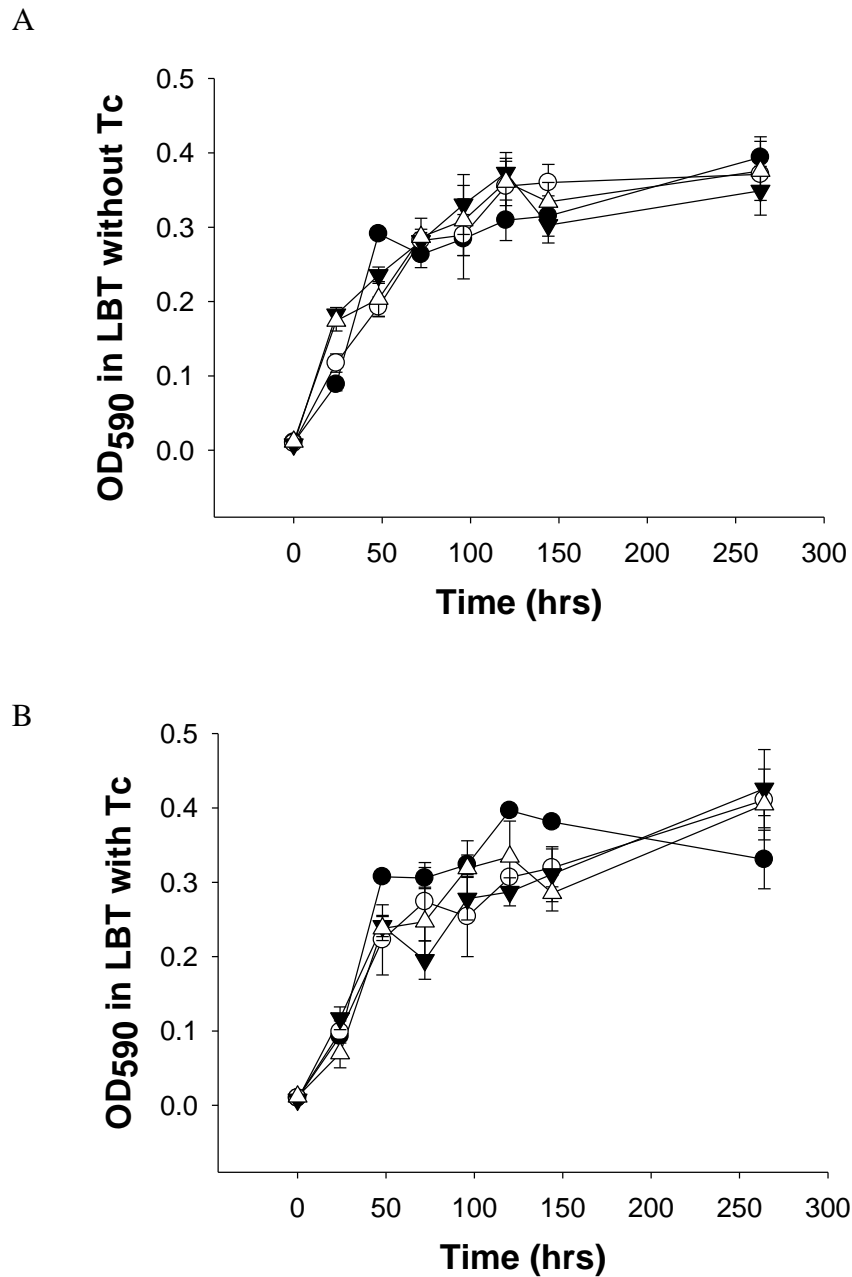
#### 4.3.5.1 Antisense expression under rich culture conditions (LBT)

The above mentioned strains (Table 4.1) were cultured in rich media (LBT) with and without Tc. Analysis of results were carried out at 144 hours with all cultures supplemented with Tc. At 144 hours in the 1<sup>st</sup> experiment, there was approximately a 25% growth inhibition seen with all cultures compared to *M. smegmatis*-pMindGFP control (Figure 4.9A). In the 2<sup>nd</sup> experiment, only the *M. smegmatis*-pMind\_ *parA*\_GFP cultures showed a 33% inhibition (data not shown). These results were verified with CFU data (Figure 4.10).

In the 1<sup>st</sup> experiment, the *M. smegmatis*-pMindGFP control cultures had  $2.31 \times 10^{12}$  CFU/mL compared to *M. smegmatis*-pMind\_ *obg* cultures which had  $3.41 \times 10^{10}$  CFU/mL, a 67 fold decrease (Figure 4.10B). There was no significant difference in the number of CFU in the *M. smegmatis*-pMind\_ *parB*\_GFP cultures however; *M. smegmatis*-pMind\_ *parA*\_GFP cultures had  $7.83 \times 10^9$  CFU/mL, a 295 fold decrease compared to the control (Figure 4.10B). In the 2<sup>nd</sup> experiment, there was a decrease in the number of CFU in the *M. smegmatis*-pMind\_ *obg* (100 fold) and the *M. smegmatis*-pMind\_ *parA*\_GFP (400 fold) cultures at 144 hours with no difference in the number of CFU in the *M. smegmatis*-pMind\_ *parB*\_GFP cultures compared to the control cultures (data not shown).

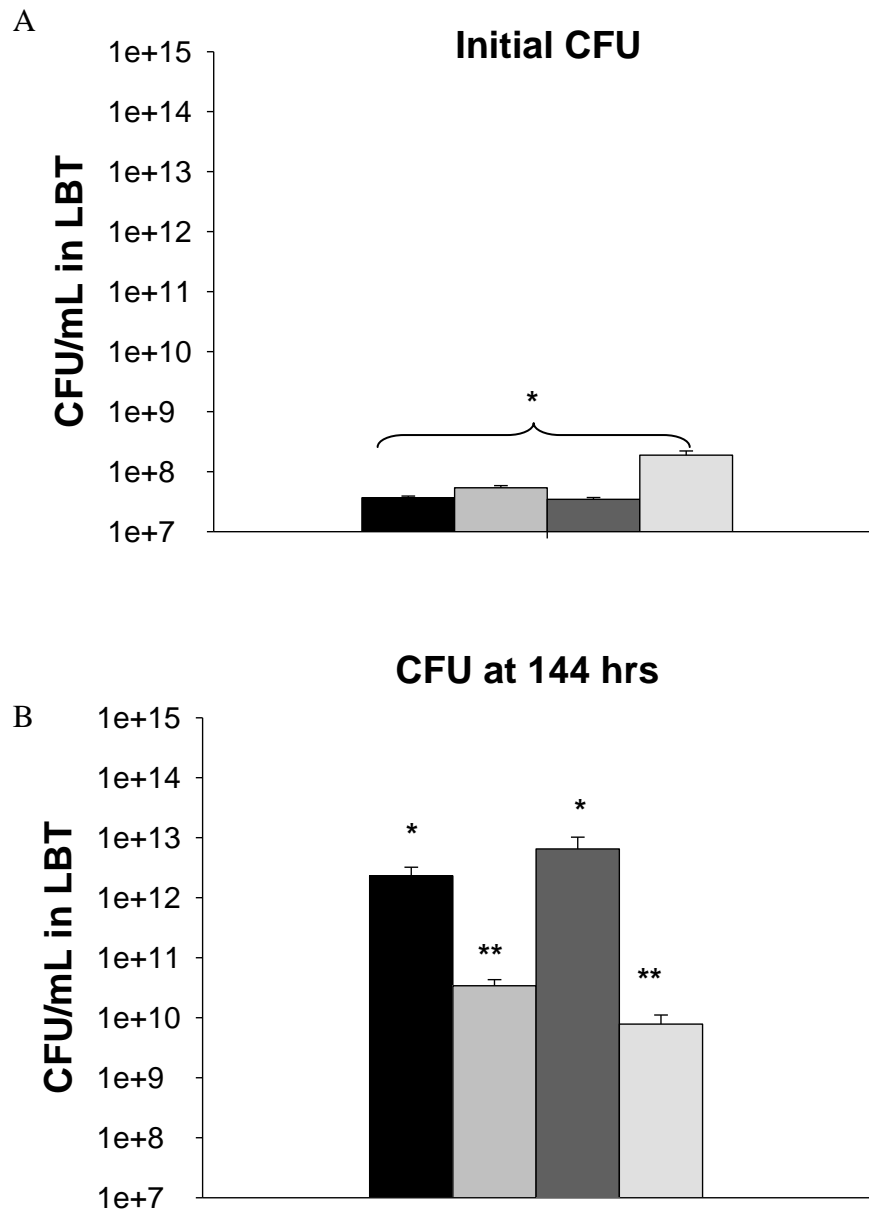
Consequently, expressing antisense RNA of the various genes in rich media gave varying results with respect to OD and the number CFU. All strains showed very little inhibition with OD; however there were significant reductions in the number of CFU with two genes, *obg* and *parA* with *parA* having the most significant effect. All strains were also analysed for any morphological differences with DAPI stain but there were no significant changes in the morphology of the cells between all strains (Appendix 9.3.1).





**Figure 4.9: Growth curves of cultures in LBT**

Growth curves of *M. smegmatis* strains electroporated with vectors GFP (●), *obg* (○), *parB* (▼) and *parA* (Δ) cultured in LBT **A**) without Tc and **B**) with 20 ng/mL Tc. As can be seen from these two graphs, there were no differences between the strains without antisense induction. There was a slight reduction in OD with antisense induction at late exponential phase however; upon reaching stationary phase these differences were not significantly maintained. All results are representative of mean values of quintuplicates  $\pm$  SE.



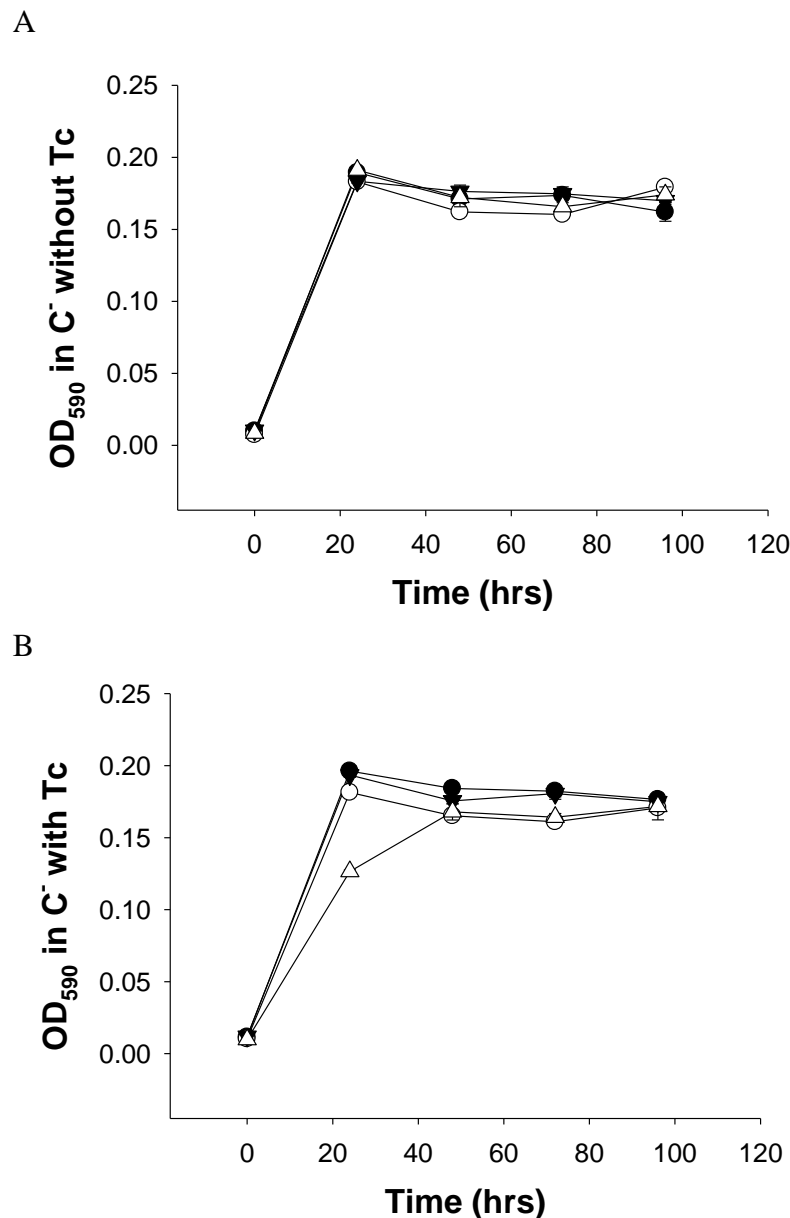
**Figure 4.10: Number of viable cells in LBT**

The above graphs show the CFU/mL of *M. smegmatis* strains electroporated with vectors GFP (■), *obg* (□), *parB* (■) and *parA* (□) cultured in LBT with 20 ng/mL Tc at **A**) 0 hours and **B**) 144 hours (late exponential phase, refer to Figure 4.9). At time 0, all cultures had approximately the same number of cells except for *parA* which had a slightly higher inoculum however this difference was insignificant. At 144 hours, there was ~ a 2 log reduction in the number of CFU between the control strain and the antisense strains *M. smegmatis*-pMind\_*obg* and *M. smegmatis*-pMind\_*parA*\_GFP. \* represents p value > 0.06 and \*\* represents p value < 0.03.

#### 4.3.5.2 Antisense expression under carbon starvation (C<sup>-</sup>) conditions

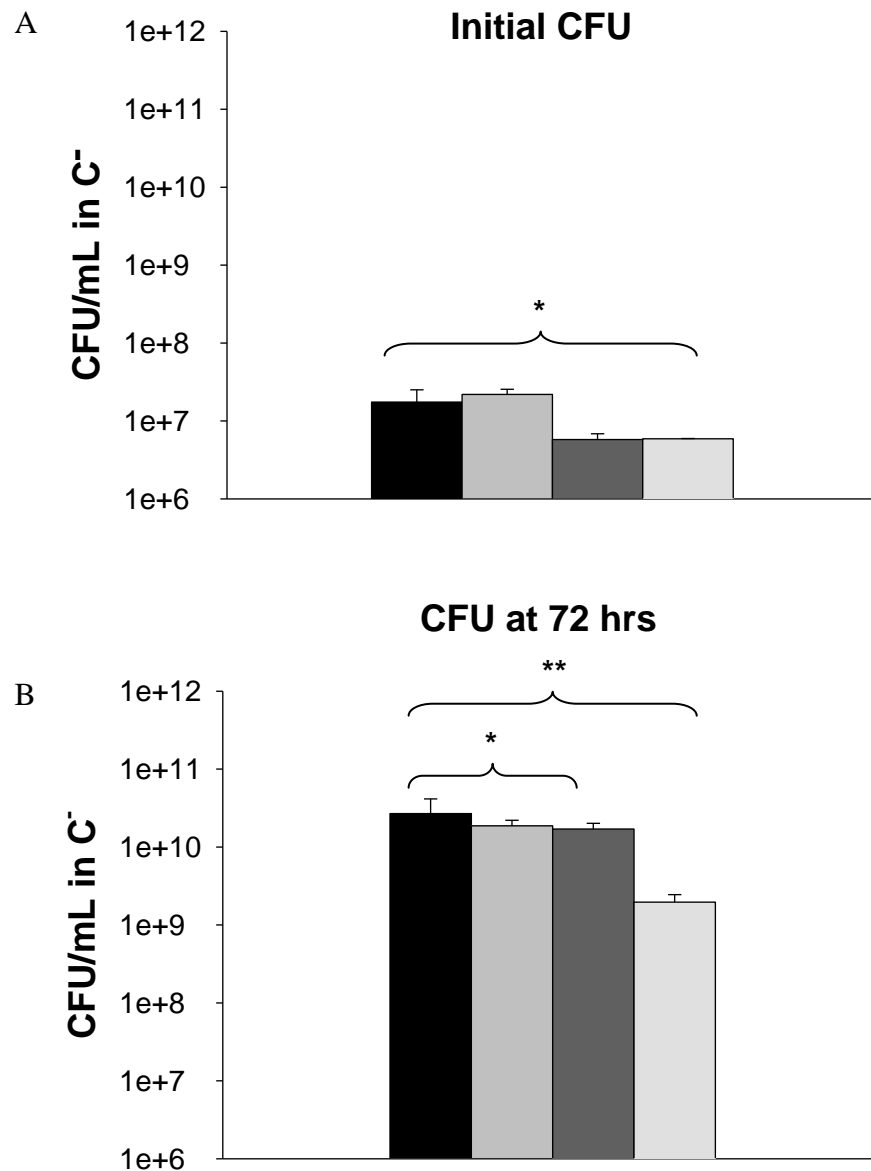
The above experiment was carried out in carbon limited media with and without Tc. Analysis of results was carried out at 72 hours with all cultures supplemented with Tc. At 72 hours in the 1<sup>st</sup> experiment, only the OD of the *M. smegmatis*-pMind\_*parA*\_GFP cultures showed a 10% inhibition compared to the *M. smegmatis*-pMindGFP control (Figure 4.11B). *M. smegmatis*-pMind\_*parA*\_GFP cultures also took twice as long to reach the stationary phase OD compared to all other cultures (Figure 4.11B). This was consistent with what was seen in the 2<sup>nd</sup> experiment; *M. smegmatis*-pMind\_*parA*\_GFP cultures exhibited a 24% inhibition with a longer exponential phase (data not shown). These results were verified with CFU data.

The CFU data from the 1<sup>st</sup> experiment showed no change in the number of CFU between all cultures and the control except for the *M. smegmatis*-pMind\_*parA*\_GFP cultures (Figure 4.12B). *M. smegmatis*-pMind\_*parA*\_GFP cultures showed approximately a 14 fold reduction the number of CFU. This was confirmed in a 2<sup>nd</sup> experiment which only showed a decrease in the number of CFU with the *M. smegmatis*-pMind\_*parA*\_GFP cultures (data not shown). The cells were also analysed for morphological difference with DAPI stain at 72 hours (Appendix 9.3.2). There were no differences in the cellular morphology between all cultures except for the cells from the *M. smegmatis*-pMind\_*parA*\_GFP cultures. These were ~ 3  $\mu$ M smaller than the cells in the control culture.



**Figure 4.11: Growth curves of cultures in C<sup>-</sup> media**

Growth curves of *M. smegmatis* strains electroporated with vectors GFP (●), *obg* (○), *parB* (▼) and *parA* (Δ) cultured in C<sup>-</sup> media **A**) without Tc and **B**) with 20 ng/mL Tc. There were no significant differences between all the strains without antisense induction; however upon induction the *M. smegmatis*-pMind\_*parA*\_GFP strain took twice as long to reach the stationary phase OD compared to all other strains. It can also be seen that *M. smegmatis*-pMind\_*parA*\_GFP and *M. smegmatis*-pMind\_*obg* cultures were maintained at a slightly lower OD in the early stationary phase compared to the 2 control strains however, this difference was not significant once the cells entered late stationary phase. All results are representative of mean values of quintuplicates  $\pm$  SE.



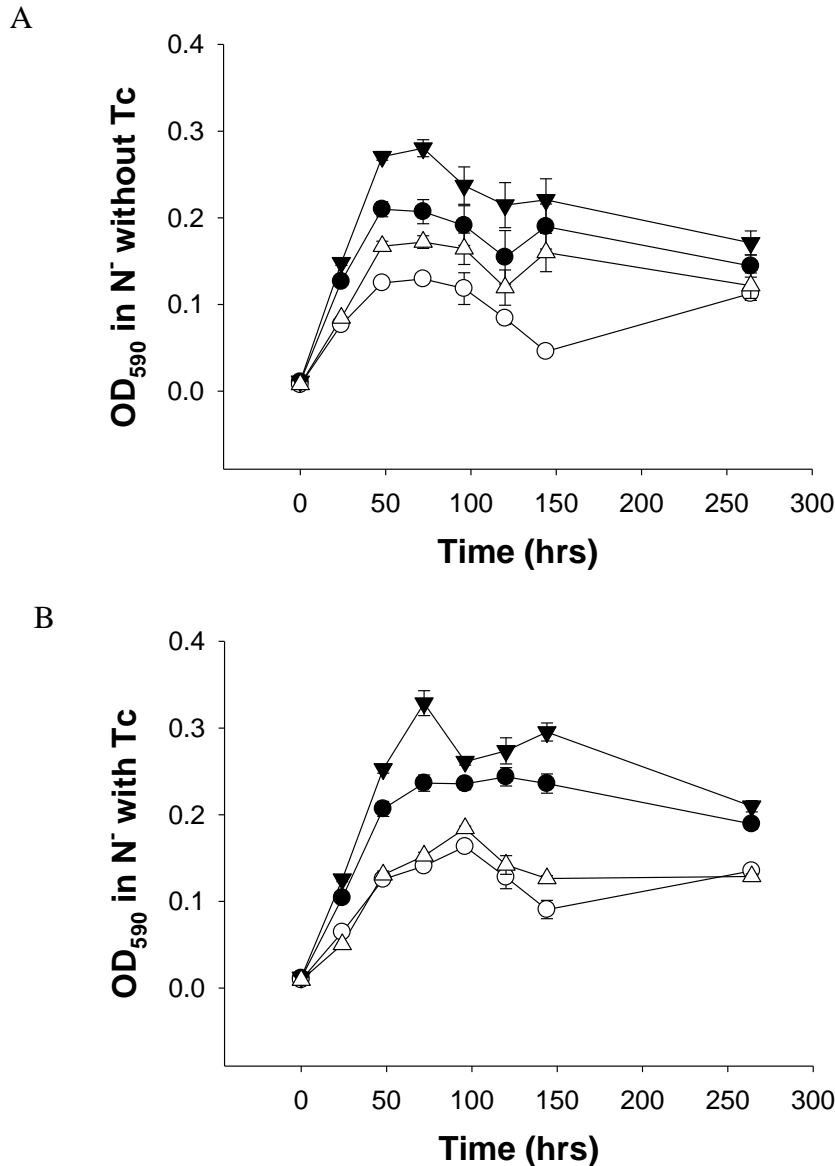
**Figure 4.12: Number of viable cells in C<sup>-</sup> media**

The above graphs show the CFU/mL of *M. smegmatis* strains electroporated with vectors GFP (■), *obg* (□), *parB* (■) and *parA* (□) cultured in C<sup>-</sup> media with 20 ng/mL Tc at **A**) 0 hours and **B**) 72 hours. At time 0 all cultures had approximately the same number of cells however at 72 hours, the CFU for the antisense strain *M. smegmatis*-pMind\_*parA*\_GFP was approximately 1 log lower than all other strains which is comparable with what was observed in the growth curve by OD. \* represents p value > 0.1 and \*\* represents p value < 0.05.

#### 4.3.5.3 Antisense expression under nitrogen starvation (N<sup>-</sup>) conditions

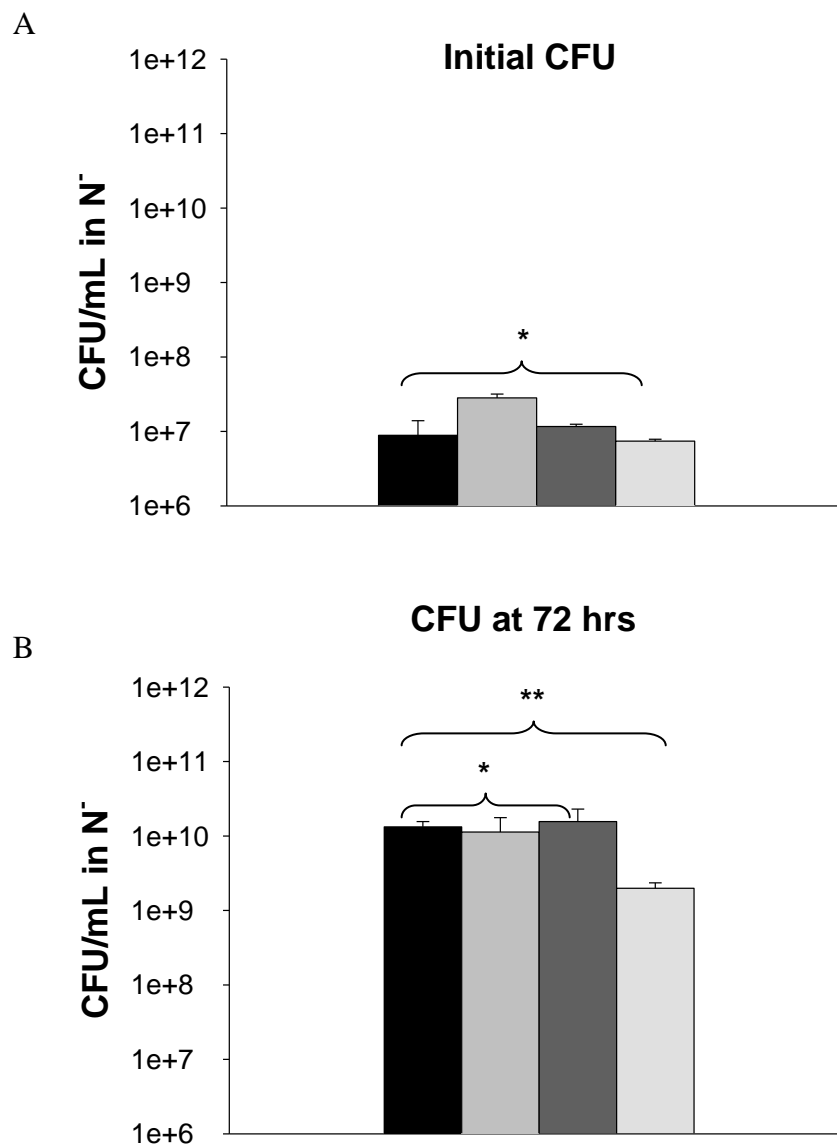
The above experiment was carried out in nitrogen limited media with and without Tc. Analysis of results was carried out at 72 hours. Growth experiments in N<sup>-</sup> media showed the most pronounced effect on the OD of the cells (Figure 4.13). In the absence of antisense induction, the basal level of antisense expression of all essential genes from incomplete repression of the promoter had some effect on the cells (Figure 4.13A). However, upon antisense induction this effect was much more pronounced (Figure 4.13B). *M. smegmatis*-pMind\_ogb cultures were inhibited 44% while *M. smegmatis*-pMind\_parA\_GFP cultures were inhibited 32% compared to the *M. smegmatis*-pMindGFP cultures at 72 hours. The growth of *M. smegmatis*-pMind\_parB\_GFP cultures were approximately 40% enhanced compared to the *M. smegmatis*-pMindGFP cultures. The above experiment was repeated and results compared. Both experiments showed consistent results. These results were verified with CFU data (Figure 4.14).

The number of CFU of the cultures in N<sup>-</sup> media was not consistent with the OD data. Even though cultures of all the strains showed an effect with OD compared to the OD of the control strain (Figure 4.13), the number of CFU were approximately the same as the control for all cultures except the *M. smegmatis*-pMind\_parA\_GFP cultures (Figure 4.14B). These cultures showed a 6.7 fold decrease in the number of CFU. This was confirmed in a 2<sup>nd</sup> experiment which also showed an effect with OD with all cultures consistent with the 1<sup>st</sup> experiment but only a decrease in the number of CFU with the *M. smegmatis*-pMind\_parA\_GFP cultures (data not shown). Expressing antisense RNA of the different genes under N<sup>-</sup> media gave varying results with respect to OD and the number CFU. All cultures exhibited a difference with OD; however there were no significant differences in the number of CFU between all cultures except for the cultures expressing antisense RNA of *parA*. The cells were also analysed for morphological difference with DAPI stain at 72 hours. All cells were slightly elongated in N<sup>-</sup> media and *M. smegmatis*-pMind\_ogb and *M. smegmatis*-pMind\_parA\_GFP cells had irregularly located foci along the length of the cells (Appendix 9.3.3).



**Figure 4.13: Growth curves of cultures in N<sup>-</sup> media**

Growth curves of *M. smegmatis* strains electroporated with vectors GFP (●), *obg* (○), *parB* (▼) and *parA* (Δ) cultured in N<sup>-</sup> media **A**) without Tc and **B**) with 20 ng/mL Tc. Under nitrogen starvation, there was some growth inhibition of *obg* and *parA* antisense strains without induction. Statistical analysis using a two-tailed T-test showed a non-significant difference without induction (p value > 0.03). This difference was more pronounced after induction showing a significant difference between the control and the *obg* and *parA* antisense strains. Statistical analysis using a two-tailed T-test gave a p value < 0.0008 for *M. smegmatis*-pMind\_ *obg* and *M. smegmatis*-pMind\_ *parA*\_GFP. *M. smegmatis*-pMind\_ *parB*\_GFP antisense control strain grew better than all strains, even the control, p value = 0.001. All results are representative of mean values of quintuplicates ± SE.



**Figure 4.14: Number of viable cells in  $N^-$  media**

The above graphs show the CFU/mL of *M. smegmatis* strains electroporated with vectors GFP (■), *obg* (□), *parB* (■) and *parA* (□) cultured in  $N^-$  media with 20 ng/mL Tc at **A**) 0 hours and **B**) 72 hours (late exponential phase, refer to Figure 4.13). At time 0 all cultures had approximately the same number of cells. However at 72 hours, the number of CFU for the antisense strains *M. smegmatis*-pMind\_*parA*\_GFP was approximately 1/3 a log lower than all other strains. The CFU data did not concur with the OD data with *M. smegmatis*-pMind\_*obg* which showed a reduction upon antisense induction in two independent experiments. \* represents  $p$  value  $> 0.3$  and \*\* represents  $p$  value  $< 0.004$ .

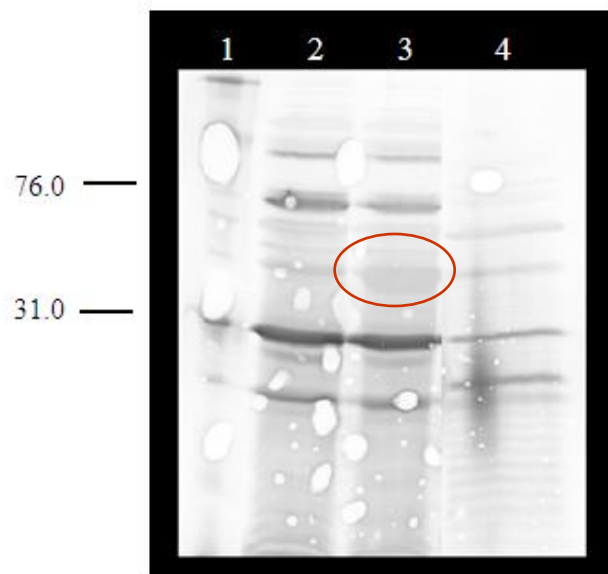


#### 4.3.6 Measuring protein levels

Antisense induction of *parA* decreased the viability of the cells under all culture conditions tested. Thus it was hypothesized antisense induction of *parA* led to a down-regulation of ParA that caused this effect. To test this hypothesis, antibodies were raised against the ParA protein to measure the level of protein present in the cell lysates with and without expression of *parA* antisense RNA. Purified ParA protein was sent to the Pocono Rabbit Farm and Laboratory Inc. for antibody production (section 2.10.6).

Upon receipt of the antibodies, western blots were performed with various dilutions (section 2.10.7) however no bands resolved at ~ 40 kDa, the expected size of the ParA protein (data not shown). There was however a prominent non-specific band resolved at 31 kDa and several other non-specific bands of other molecular weight.

To identify if there was not enough protein in the lysate to be detected by the antibodies or if the antibodies raised were non-specific, the ParA protein was over-expressed in *E. coli* and the cell lysate was run on SDS-PAGE and used for western blotting. This western blot illustrated that the antibodies raised were non-specific as there was an abundant of overexpressed protein transferred on to the membrane (Figure 4.15). Thus the level of ParA protein down-regulated upon antisense expression of *parA* could not be measured in this study.



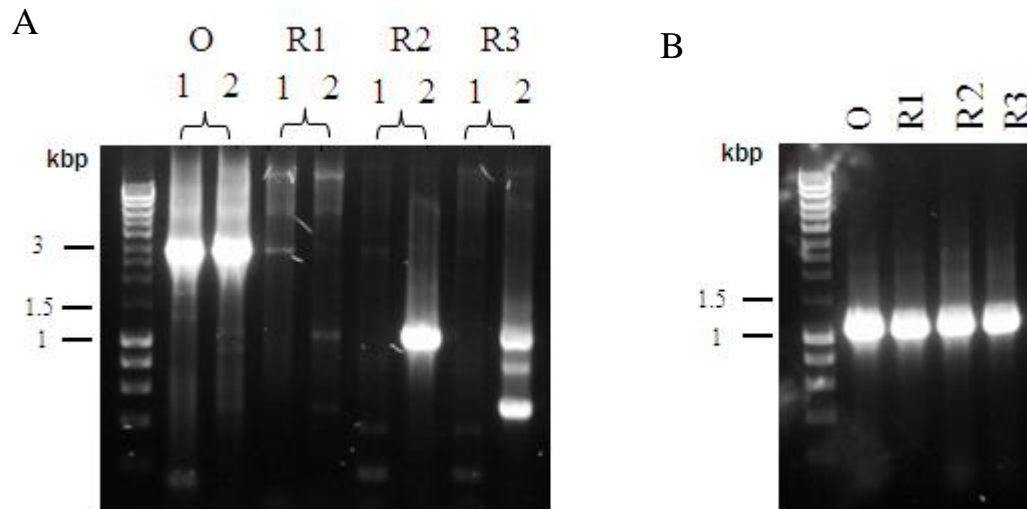
**Figure 4.15: Western blot showing non-specific binding**

The above Western blot contains: lane 1- a protein molecular marker, lane 2- cell lysates of *E. coli*-pBAD/HisB, lane 3- cell lysate of over-expressed *E. coli*-pBAD/HisB\_ParA and lane 4- purified ParA protein. As can be seen in lane 3 (red circle), there was sufficient ParA protein (over-expressed) transferred on the membrane however, it did not bind the antibodies but showed non-specific binding. Since the ParA antibodies showed non-specific binding, it was unable to be used in determining the level of ParA upon antisense expression in the *M. smegmatis* cells.

#### 4.3.7 Characterization of *M. smegmatis*-pMind\_*parA*+*parB*\_GFP

*M. smegmatis*-pMind\_*parA*+*parB*\_GFP colonies took at least 3 weeks to grow sufficiently to be picked from plates to be cultured in broth (Figure 4.6). After 3 weeks, the colonies were cultured in LBT Kan<sup>+</sup>Hyg<sup>+</sup>. These cultures grew overnight to an OD<sub>600</sub> of ~ 0.3. This was unexpected and it appeared that these strains had mutated that allowed it to overcome the growth defect that was experienced by the cells when grown on solid media.

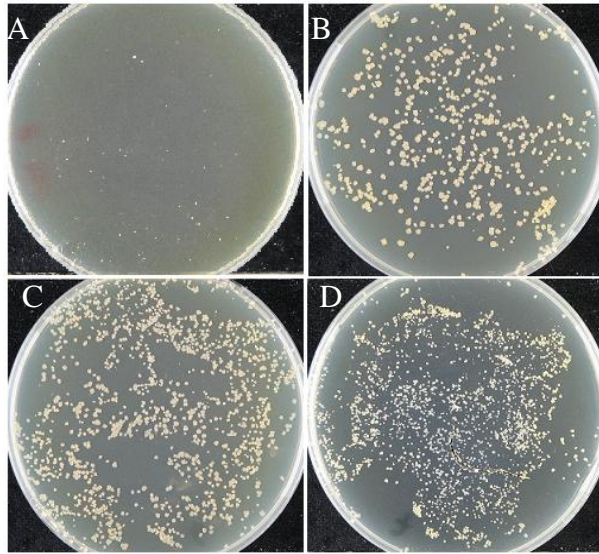
In order to identify where such mutation(s) might have occurred, the pMind\_*parA*+*parB*\_GFP vector was isolated from the *M. smegmatis* cells to detect if the mutation(s) were extrachromosomal. The original pMind\_*parA*+*parB*\_GFP vector electropoarted into *M. smegmatis* was called O and the vectors recovered from *M. smegmatis* after electoporations were called R1, R2 and R3. The inserts in these vectors were amplified with various primer sets (Figure 4.16) and sequenced. The vectors were also re-electroporated into WT *M. smegmatis* (Figure 4.17) and cured *M. smegmatis* strains (Figure 4.19), that is, the strains that had the pMind\_*parA*+*parB*\_GFP vector eliminated from the cells (Figure 4.18).



**Figure 4.16: PCR products of pMind\_*parA*+*parB*\_GFP**

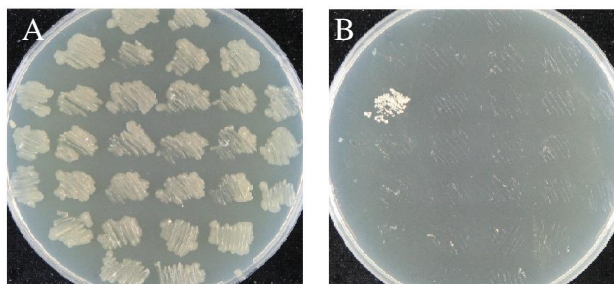
1% agarose gels showing PCR products of pMind\_*parA*+*parB*\_GFP vectors O\*, R1\*, R2\* and R3\*. **A)** All lanes labelled 1 contain PCR products from the primer set TetROF11 and GFP\_R and all lanes labelled 2 contain PCR products from the primer set HygR11 and MSMEG\_6938\_R. The original vector (O) shows ~ 3 kbp insert that is to be expected from the 1.056 kb *parB* fragment, 0.972 kb *parA* fragment and 0.750 kb GFP fragment together however, recovered vectors (R1, R2 and R3) showed non-specific binding with these primer sets and none of the PCR's gave the correct size insert band. **B)** PCR products from primer set GFP\_RBS\_F and GFP\_R of vector O, R1, R2 and R3 showing the correct size insert of GFP in all 4 vectors.

\*O = original vector, R1 = recovered vector 1, R2 = recovered vector 2 and R3 = recovered vector 3.



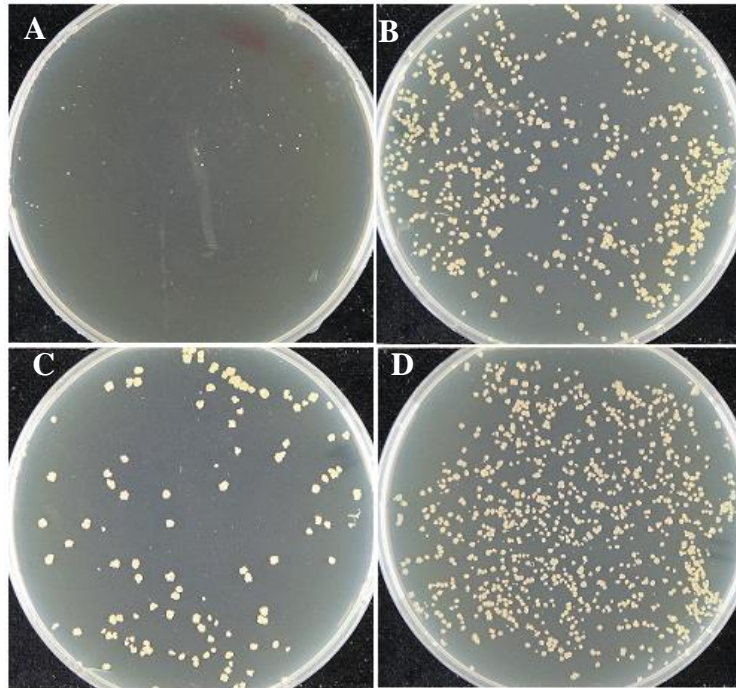
**Figure 4.17: Electroporation of *M. smegmatis***

These plate photos show *M. smegmatis* electroporated with pMind\_*parA*+*parB*\_GFP vectors **A)** O\* **B)** R1\* **C)** R2\* and **D)** R3\*. The growth of the strain with original pMind\_*parA*+*parB*\_GFP (O) exhibited inhibition as reported in Figure 4.6 however; cells electroporated with the recovered vectors (R1, R2 and R3) overcame this difficulty implying that these vectors had mutated. This was performed in triplicate and all plates exhibited similar results. \*O = original vector, R1 = recovered vector 1, R2 = recovered vector 2 and R3 = recovered vector 3.



**Figure 4.18: Cured *M. smegmatis* strain**

*M. smegmatis*-pMind\_*parA*+*parB*\_GFP colonies grown in LBT and plated on LB agar as described in section 2.9.10 were patched onto **A)** LB **B)** LB Kan<sup>+</sup>Hyg<sup>+</sup> plates. The loss of antibiotic resistance signified that the strains were cured of the pMind\_*parA*+*parB*\_GFP vector.



**Figure 4.19 Electroporation of cured *M. smegmatis***

These plate photos show cured *M. smegmatis* strain electroporated with pMind\_*parA*+*parB*\_GFP vectors A) O\* B) R1\* C) R2\* and D) R3\*. This was performed in triplicate with 3 cured strains, all exhibiting the same result.

\*O = original vector, R1 = recovered vector 1, R2 = recovered vector 2 and R3 = recovered vector 3.

As can be seen from Figure 4.19, the cured cells were phenotypically similar to the WT cells (Figure 4.17) when electroporated with the same vectors where the original vector caused growth inhibition and the recovered vectors overcame this defect. This suggests that the mutation(s) arising due to the presence of the antisense fragment of *parA* and *parB* together in one vector were not chromosomal but extrachromosomal. All four vectors (O, R1, R2 and R3) were sent for DNA sequencing to verify this. The primers TetROF11, MSMEG\_6938\_F, MSMEG\_6939\_R and HygR11 (Figure 4.5) were used for sequencing. The two internal primers were used as the expected sequence was ~ 3 kbp; too long to give legible sequence for the whole length by the external primers alone. The original pMind\_*parA*+*parB*\_GFP vector (O) gave the expected

---

sequence of *parA*, *parB* and *gfp* that was cloned into the pMind vector with the 4 primers (data not shown).

All three recovered vectors (R1, R2 and R3) gave the correct *gfp* sequence that was originally cloned with the HygR11 primer (data not shown). This was also confirmed by PCR (Figure 4.16B) where primers GFP\_RBS\_F and GFP\_R gave the expected product size of *gfp*. The TetROF11 primer gave ~ 100 bp that read into the end of the *parB* gene before numerous mutations were found (Figure 4.20). The MSMEG\_6938\_F and MSMEG\_6939\_R primers gave illegible sequences with R1, R2 and R3. This could be explained by the fact that most of the sequences past the last 100 bp of the *parB* gene were mutated so failed to get recognised by these primers.

Since electroporating the vector hosting the sequence of *parA* and *parB* in the antisense orientation gave rise to mutations in the sequence, all *M. smegmatis* antisense strains were cultured overnight and the antisense vectors were recovered and sequenced from these to identify if the original sequences of the cloned genes underwent mutations like the *parA+parB* gene fragment. All other sequenced vectors showed that the genes originally cloned in were maintained without mutations after being electroporated into *M. smegmatis*.

```

6938n6939GFP  -----
R1-TetRO_F11  C N N N C N N N T  T N T N N A T A C T  T T A N C T T T A G  A T A A A G T G A C  T G C T C C C T A C  T C T C A T C T G A  C T G C T C G C T A  C T C T C A T C G T
R2-TetRO_F11  - N N N N C N N G N  N C N N N A T A C T  T T A T C T T - A G  A T A A A G T G A C  T G C T C G C T A C  T C T C A T C T G A  C T G C T C G C T A  C T C T C A T C G T
R3-TetRO_F11  -- N N C N N G T N  A T C N G A T A C T  T T A T C T T - A G  A T A A A G T G A C  T G C T C G C T A C  T C T C A T C T G A  C T G C T C G C T A  C T C T C A T C G T
Identity      -----

6938n6939GFP  -----
R1-TetRO_F11  G G A A T C C T G A  C A G G A T C C C T  T A A G A A C T G C  A G T C C A T A T G  T T A C T C G T T C  T G G G C G C T C A  T C A G C T C G A C  G A T G C G C T G C
R2-TetRO_F11  G G A A T C C T G A  C A G G A T C C C T  T A A G A A C T G C  A G T C C A T A T G  .....
R3-TetRO_F11  G G A A T C C T G A  C A G G A T C C C T  T A A G A A C T G C  A G T C C A T A T G  .....
Identity      -----
*****

6938n6939GFP  A G A T C G T C C A  C G G A C C C G A A  C T C G A C G A C G  A T C T T G C C C T  T G C G T T T G C C  G A G G C T C A C G  G T G A C G C G C G  T G T C G A A C G C
R1-TetRO_F11  .....
R2-TetRO_F11  .....
R3-TetRO_F11  .....
Identity      *****

6938n6939GFP  C G T C G A G A G C  C G C T C G G C A -  - - A C G T C C T G  G A G A C C G G C  A T C T G G A T C G  G C T T A C G G C G  C G G G G C A G G C  T G T G T C T T G C
R1-TetRO_F11  A A C T A T A . C T  . A . A . A A T G T  A T . . A . . A .  . . A G A . A A A .  . A A A . A . . G .  A . . C . A A . T T  A A C T T . . A A A  T . A . A . A C A A
R2-TetRO_F11  .....
R3-TetRO_F11  A A C T A T A . C T  . A . A . A A T G T  A T . . A . . A .  . . A G A . A A A .  . A A A . A . . G .  A A . C . A A . T T  A A C T T . G A A A  T . A . A . A C A A
Identity      * * * * *

6938n6939GFP  C T T C A C G G T T  C G C G A G C G T C  A C G G C T T C T T  C C G T G G C C C G  G A C C G A C A G C  C C T T C G G C G A  C G A T G C G C G C  C G C G A G T T C C
R1-TetRO_F11  . A . T G A A C A .  G . - A . . . . . T  - . A A . . A G C A  - - - - . A . . A T  T . T . A . G G A A  A T - - A C T . C .  - A T . . - . . A  T . G C C C . G T .
R2-TetRO_F11  .....
R3-TetRO_F11  . A . T G A A . A .  G . - A . . . . . T  - . A A . . A G C A  - - - - . A . . A T  T . T . A . . . A A  A T - - A C T . C .  - A T . . - . . A  T . G C C C . G T .
Identity      * * * * *

6938n6939GFP  T C C T G G G C C T  C C G G T C C G C C  C T C G A G C G C G  A G G A G C G C C C  G C G C G T G G C C  G G C C G A A A G T  A C A C C C G C C G  C G A C C C G C C G
R1-TetRO_F11  C T T . T A C . A G  A . A A C . A T T A  . C T . T C . A . A  C A A T C T . . .  - - - - - T T T  C T A A A G . T C C  C A . . G A A A A .  A . . G A . C A . A
R2-TetRO_F11  C A T A T . A G T A  A A . . A G G A G A  A C T T T T . A . T  G . A G T T . T .  C A A T - - - T . T  T . T T . . . T T A  G A T G G T . A T  T T . A T G . G . A
R3-TetRO_F11  C T T . T A C . A .  A . A A C . A T T A  . C T . T C . A . A  C A A T C T . . .  - - - - - T T T  C . A A A G . T C C  C A . . G A A A A .  A . . G A . C A . A
Identity      * * * * *

6938n6939GFP  C T G C A C C G C G  A T G G G C A G A C  G C A G C A G C C G  G A T C A T G T T C  G A G A T G A G C G  G C C G G G A G C G  A C C G A T G C G C  G C C G C G A G T T
R1-TetRO_F11  T G . T C . T T . T  T G A . T T T . T A  A . . C . T . . T .  . G - - - - - A T T . C A C A T .  . A T . . . T - .  . A . T . . A A A  A T A A T A . A C .
R2-TetRO_F11  . A A A T T T T . T  - - - . T . . . T G  . A G A G G . T G A  A G - - - - - . T . . . . C A A C  A T A C . . . A - A  . . T T . C C . T T  A A A T T T . T .
R3-TetRO_F11  T G . T C . T T . T  T G A . T T T . T A  A . . . T . . T .  . G - - - - - A T T . C A C A T .  . A T . . . T - .  . A . T . . A A A  A T A A T A . A C .
Identity      * * * * *

6938n6939GFP  C G T C G T G G G T  G A C G C C G A A C  T C G T C G A G C A  G C T G C T G A T A  G G C C G A C G C C  T C T T C G A G T G  G G T T C A A C T G  G A C G C G G T G G
R1-TetRO_F11  - - - - A G T . C A  T G . C T G A C . .  A A . A A T C C . T  . . A . . T C A C  . A . . . - T A T T  G A . . . . G A . .  A T . C . T . . G C  . . G C . T . C . .
R2-TetRO_F11  G - - . A C T A C .  . G A A A A C T . .  C T . . T C C A T G  . . C A A C A C . T  . T . A C - T A . T  . T C G . . T A . .  . T - - . T T . A A  T G . T T T . C . A
R3-TetRO_F11  - - - - A G T . C A  T G . . T G A C . .  A A . A A T C C . T  . T . A . . T C . C  . A . . . - T A T T  G A . . . . G A . .  A T . C . T . . G C  . . G C . T . C . .
Identity      * * * * *

```

**Figure 4.20: Sequence alignment of *parA*+*parB***

The above is an alignment of the original *parA*+*parB*\_GFP sequence with approximately 600 bp of the sequences obtained from the TetROF11 primer from the vectors R1, R2 and R3. The first 100 bp is only present in R1, R2 and R3 as this is the sequence from the pMind vector from the TetROF11 primer (Figure 4.5); *parA*+*parB*\_GFP only contains sequences of *parA*, *parB* and *gfp*. The second 100 bps that is perfectly aligned with the *parA*+*parB*\_GFP sequence is the end of the *parB*. Everything downstream of the perfectly aligned 100 bps has undergone extreme mutations. These mutations do not seem to follow any particular pattern so the individual residues that might be crucial for these genes could not be identified.

‘\*’ denotes bases that are identical in all vectors, ‘-’ denotes deleted bases and ‘.’ denotes bases that are identical to the original sequence.



## 4.4 Discussion

This chapter aimed to validate some of the homologues of the genes identified as potential broad-spectrum targets in chapter 3 in the surrogate *M. smegmatis* using conditional gene silencing with antisense RNA. Of particular interest to this study was validating as yet uncharacterized proteins that may be involved in cell division to generate novel targets for anti-tubercular therapy. The two cell division genes identified as potential broad-spectrum targets in this study, *ftsZ* and *ftsH* (supplementary Table S1), have extensive research being invested on them to characterize these proteins and identify their inhibitors and their mechanisms of inhibition [217-223] and thus were discounted in this study.

For the scope of this project, I looked at other genes that are known or thought to be associated with cell division in *M. tuberculosis*. The conserved essential gene list presented nine such genes (Table 3.2). Three of these genes, Rv2440c, Rv3917c and Rv3918c, believed to be involved in chromosome partitioning, were chosen to be evaluated in this study.

### 4.4.1 Gene modulation

The system used for target validation in this study was gene modulation via a tetracycline-inducible vector pMind. This was used for conditional antisense RNA expression however there was one drawback with this system. This vector has been reported to exhibit incomplete repression of the promoter by the repressor in the absence of the inducer, showing approximately a 10 fold higher luminescence when uninduced compared to a vector that lacked the *tetRO* region [120]. This was also observed in this study when GFP expression was detected without induction of the promoter upon placement of the *gfp* sequence under its control (Figure 4.8). There are a number of reasons why this vector may be leaky. Firstly, the *tetA* promoter may be a very strong mycobacterial promoter and the binding of TetR to *tetO* is not able to effectively suppress transcription of GFP. Alternatively, there could be background GFP expression due to readthrough from



---

promoter sequences within the plasmid backbone. Thus this system was thoroughly evaluated before being used in the target validation experiments to ensure that there were no repercussions due to partial expression under repressed conditions. Firstly, all the vectors constructed and electroporated in *M. smegmatis* were isolated from *M. smegmatis* and sequenced to ensure the maintenance of the antisense sequence. Secondly, it was demonstrated that there was an increase in GFP expression from the promoter upon induction with tetracycline validating that this system could be used for target modulation. Thirdly, only freshly electroporated cells were used in the experiments to prevent any acquisition of mutations over time in these strains. This is a commonly used technique to prevent suppressor mutations [224]. This system had also been successfully used previously to down regulate essential genes [120]. Another more tightly regulated system was also attempted in this study but this method proved unsuccessful in generating the mutants (section 4.4.1.1).

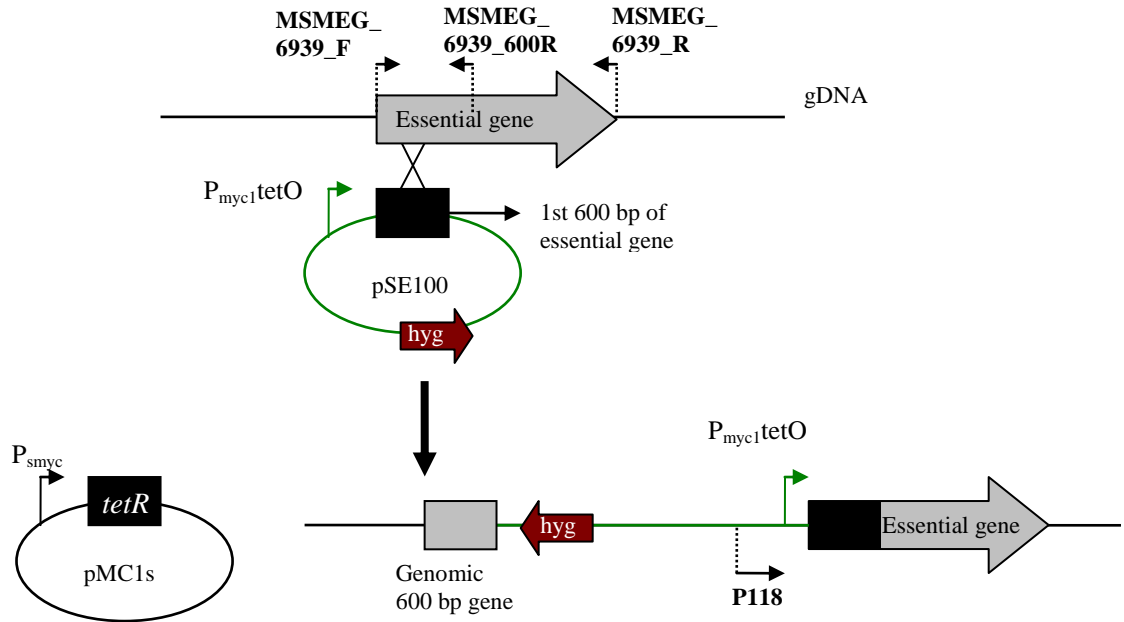
GFP expression was measured to monitor expression from the promoter upon induction however it was not used for the growth analysis of the cultures even though it is a more sensitive measure than OD. This was because we were looking at genes involved in cell division and it has been known that perturbing proteins essential to this process gives filamentous phenotype [215]. These elongated cells will emit higher fluorescence compared to regular sized cells which may not necessarily correlate with an increase in the number of cells. While OD gave an approximate measure of the number of cells present in the cultures, this was not used as an absolute measure either for the same reason mentioned above; elongated cells giving greater absorbance. CFU were used to give the most accurate measure of the number of viable cells present in the cultures as this has been used as the gold standard for proof of cell viability. All CFU data presented in this study is exclusive to strains in the presence of tetracycline. The number of CFU without tetracycline is not fully reflective of the CFU in the absence of the antisense RNA as there is some expression from the promoter at all times (Figure 4.8). All data was compared against the *M. smegmatis*-pMindGFP control with tetracycline to determine if there were any differences upon antisense induction of an essential gene. Section 4.4.1.2 and 4.4.1.3 discusses the genes that were tested

in this work together with their roles and their potential contribution to the cell division process.

#### 4.4.1.1 Addressing incomplete repression

An important factor that needs to be addressed for any inducible gene regulation system is its ability to be tightly regulated. Since gene expression from the pMind vector was not very tightly regulated, a different vector was looked at for gene regulation, pSE100 (Figure 4.21) [225]. This system creates conditional mutants using a suicide vector that contains approximately 600 bp of the gene of interest from the start codon, downstream of the TetR-responsive promoter,  $P_{myc}tetO$ . This undergoes homologous recombination in the chromosome and replaces the native gene promoter with  $P_{myc}tetO$ . A second vector, pMC1s with the promoter  $P_{smyc}$  is used to constitutively express TetR which binds to the operator region and prevents transcription of the gene from  $P_{myc}tetO$ . Upon induction, the inducer binds to repressor and relieves transcription repression allowing gene expression.

This conditional knockout system of controlling gene expression was attempted with *parA* where the first 600 bp of the gene was amplified using primers MSMEG\_6939\_F, MSMEG\_6939\_600R and cloned into pSE100. The insertion was verified with primers P118 and MSMEG\_6939\_600R and DNA sequencing (data not shown). This vector, pSE100\_*parA*\_600, was electroporated into *M. smegmatis*. Numerous attempts were made to acquire mutants that had undergone homologous recombination with no success thus this approach was abandoned.



**Figure 4.21: Conditional mutant system** (adapted from Ehrt *et. al.*) [123]

This diagram illustrates the conditional knock-out system that uses the pSE100 suicide plasmid to replace the native promoter with  $P_{myc1tetO}$  which is controlled by TetR. It also shows primer binding sites for MSMEG\_6939\_600F, MSMEG\_6939\_600R, MSMEG\_6939\_R and P118 (Table 2.4).

#### 4.4.1.2 Obg

MSMEG\_4623 is the *M. smegmatis* homologue of Rv2440c (Appendix 9.2.2), a ribosome associated GTPase of the obg subfamily of proteins that is essential for bacterial viability [226, 227]. The Obg subfamily of protein is one of the P-loop small GTPases that is highly conserved in many organisms from bacteria to human and is named after its best investigated member - the Obg protein of *Bacillus subtilis*. It is implicated to be involved in cell growth, morphological differentiation, chromosome partitioning, replication initiation and stress response, however the explicit function of the gene is not known [226, 228]. In *B. subtilis* the obg protein is co-transcribed in a bicistronic operon with SpoOB

---

(sporulation protein) and is essential during exponential growth [229]. It also activates transcription factors that induce the general stress regulon in response to environmental stress [230]. In *E. coli*, the temperature sensitive mutant of the Obg homologue had been found to be involved either directly or indirectly in chromosome partitioning and underwent cell elongation [231]. Similarly, in *Vibrio harveyi*, an Obg-like GTP-binding protein is involved in regulation of chromosomal functions where mutant cells are filamentous with unpartitioned chromosome; a phenotype which is associated with impaired chromosome partitioning [232]. In *M. tuberculosis*, this gene is co-expressed with various other genes which are associated with the ribosome and protein biosynthesis (data obtained from the NIAID BioHealthBase BRC online at <http://www.biohealthbase.org>).

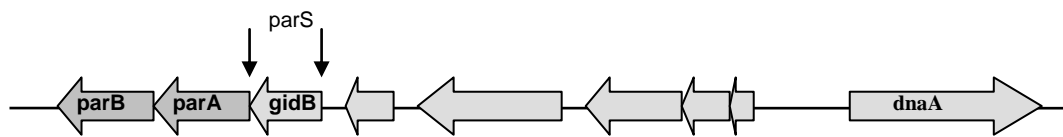
In this work, antisense RNA of the *M. smegmatis obg* GTPase was expressed under various culture conditions. In the rich media, there were a significantly lower number of viable cells at late exponential phase (Figure 4.10). This may imply that *obg* is required for exponential growth in *M. smegmatis* as down-regulating it in rich media affected its viability. In nitrogen starvation media, there was a significantly low OD but this was not reflected in the number of CFU. However, the cells did display filamentous cellular morphology with irregularly distributed foci along the length of the cell (Appendix 9.3.3) giving an indication that this gene may be required for chromosomal partitioning in *M. smegmatis*; however this would have to be confirmed with further experimentation.

GTP-binding proteins play a role in diverse cellular processes and, as mentioned above, the Obg protein is involved in various cell cycle regulations thus the effects upon down-regulating Obg may arise from a combination of different actions. Since Obg is associated with protein biosynthesis in mycobacteria, down-regulation of this gene may lead to a down-regulation of the co-expressed protein biosynthesis genes. If this is the case and Obg is required in exponential growth as was indicated in rich culture conditions, such a situation will render it non-essential when metabolism is down-regulated. This could

explain why there was no difference in the number of cells under nutrient starvation conditions as the cells may have sensed the environmental stress and slowed down metabolism, a common response to starvation conditions [108] thus, not requiring the Obg protein as it aborted exponential growth.

#### 4.4.1.3 ParA and ParB

MSMEG\_6939 is the *M. smegmatis* homologue of Rv3918c (Appendix 9.2.2), a probable chromosome partitioning protein called ParA. In mycobacterium, like in most other bacterial species, the *parA* gene is located near the chromosomal origin of replication (*oriC*) in an operon with another gene, *parB* (MSMEG\_6938 in *M. smegmatis* and Rv3917c in *M. tuberculosis*) and *cis*-acting *parS* sites which are located upstream of the *parAB* operon [165, 233-235].



**Figure 4.22: The *parAB* genome region in mycobacteria**

The diagram shows the genome arrangement of the *parAB* region in mycobacteria around the *oriC* region. The arrows at the start sites of *parA* and *gidB* indicate the *parS* sequences that bind ParB [165, 234].

The chromosomal encoded *par* system has been studied based on the model of the P1 plasmid [236, 237]. Plasmid partitioning is an active process that requires the separation and maintenance of replicated DNA at subcellular locations [238]. This energy is derived from the ATPase activity of ParA which has a nucleotide binding domain. ParA is recruited to the partition nucleoprotein complex (ParB-*parS*) via its interaction with ParB which stimulates ATP hydrolysis [237]. ParB binds the centromere-like *parS* sites which in turn binds membrane bound

proteins [239, 240]. This process aligns the newly replicated plasmid towards the cell poles to ensure equipartitioning during division.

The binding of the *parABS* for accurate partitioning has been demonstrated in a number of different bacterial species however they do not all function in the same manner [130]. In *P. aeruginosa*, the ParA protein is not essential but is required for faithful chromosome segregation in the exponential growth phase and is down-regulated in the stationary phase and deletion of ParA leads to anucleated and filamentous cells [224]. In *P. putida* however, the deletion of the *parAB* genes show no defects in rich media but had anucleate cells in minimal media [241]. The ParAB proteins have been shown to be essential for chromosomal partitioning in *Caulobacter crescentus* [242], where ParB binds to *parS* sequence and initiates cell division by regulating ParA ATPase activity [243, 244].

*Streptomyces coelicolor*, the best characterized member of the actinomycete genus has a complex chromosomal partitioning system where multigenomic aerial hyphae is converted into chains of unigenomic spores requiring synchronous segregation of multiple chromosomes [165, 245]. Deletion of either *parA* or *parB* gene in *S. coelicolor* does not affect the growth but causes deficiency in DNA partitioning. Elimination of ParA affected chromosome segregation and septation in *S. coelicolor* as ATP-bound ParA is crucial for interaction with ParB and thus chromosomal segregation which is followed by septation (section 4.4.1) [165]. ParA protein has also been reported to act as a transcriptional repressor on the *par* operon [246]. ATP-bound ParA is thought to stimulate chromosomal segregation by associating with the ParB-*parS* nucleoprotein complex whereas ADP-bound ParA binds to operator sequence and acts as a transcriptional repressor [244, 246, 247].

The activity of ParAB has not yet been characterized in *M. tuberculosis* however both genes have been shown to be indispensable in *M. tuberculosis* H37Rv [126]. In *M. smegmatis* however, *parB* has been designated non-essential where deletion of the *parB* gene only caused a longer lag phase and produced anucleate cells [165]. Jakimowicz *et. al.* also showed that ParB binds to two *oriC*-proximal *parS* sites for the formation of the nucleoprotein complex [165].

In this study, antisense RNA of both MSMEG\_6938 (ParB) and MSMEG\_6939 (ParA) were expressed under various culture conditions and monitored. Upon expression of ParB antisense, there was no effect in the growth, the number of CFU or cellular morphology of the cells in any culture conditions compared to the *M. smegmatis*-pMindGFP control. This work commenced before the publication by Jakimowicz *et. al* that demonstrated the ParB was not essential in *M. smegmatis*[165]. The data in this study confirmed what was demonstrated by Jakimowicz *et. al*. It revealed that down-regulation of ParB had no effect on cell viability demonstrating that ParB was not essential for cell division in *M. smegmatis*.

Expression of ParA antisense however, had various effects under all culture conditions tested. In rich media, the number of CFU was approximately 2-4 log lower (Figure 4.10B). Under carbon starvation conditions, cells that expressed ParA antisense were slower in reaching the exponential phase and took twice as long to reach stationary phase compared to all other strains (Figure 4.11B). The CFU in carbon starvation for ParA antisense expressing cells were also lower (Figure 4.12B) and the cells were smaller in size compared to all other strains (Appendix 9.3.2). Decreased in cell length has been associated with cell division in older cultures where biosynthetic process slow down and division starts before the cells have doubled their starting length [248]. Exhibition of this phenotype only in the ParA antisense expressing strain may indicate cells going into stationary phase due to stress other than carbon starvation, that is, potential decrease in the ParA protein concentration. In nitrogen starvation conditions, antisense expression of ParA affected the growth, CFU and cellular morphology of the cells (Figure 4.13, 4.14 and Appendix 9.3.3 respectively). The growth and CFU were lower than the control and the cells had multiple foci along it when stained with DAPI stain, a phenotype that is usually associated with improper chromosome partitioning.

This study demonstrated that ParA is essential for cell division in *M. smegmatis* and plays a far more essential role in chromosome partitioning than

ParB. ParA may be required in exponential phase growth as it had the lowest number of CFU under rich media where the cells are actively dividing. The growth of the *parA* antisense expressing strain was also compromised in the exponential phase under carbon starvation conditions (Figure 4.11B) further supporting that ParA may be essential for exponential growth in *M. smegmatis*. This was also demonstrated in another study that quantified mRNA levels of *parA* and *parB* and found that the mRNA levels for both these genes reduced from the exponential phase to late exponential and the levels were very low at stationary phase [234]. This may explain why the difference in the OD seen at the early growth phases under antisense inducing conditions was not significantly maintained at late stationary phase as the ParA protein was not highly required thus decreasing its level did not have a significant effect (Figure 4.9B and 4.11B). It was also shown by Casart *et. al.* that *parA* levels were maintained at a higher level at all growth phase than *parB* in *M. smegmatis* [234].

It has been established that mycobacteria slow-down transcription, energy metabolism, lipid biosynthesis and cell division upon nutrient starvation [249]. A study by Betts *et. al.* also showed that ParA and ParB were down-regulated upon nutrient starvation in *M. tuberculosis*; ParA 2 fold more than ParB. This demonstrates that these proteins are not as highly required in nutrient starvation conditions when cell division has slowed down compared to rich growth conditions when cells are actively dividing. Upon expression of ParA antisense under nutrient limitation conditions in this study, the number of CFU was lower than the controls. However, this was not as significant as it was shown under rich culture conditions as cells sense nutrient deprivation by their stress sensors and start to shut down/slow metabolism [249]. This would prevent them from undergoing division as actively as in rich media thus the cells would not require ParA in very high quantities. Therefore, these strains would not be expected to exhibit a very significant difference in CFU under starvation conditions upon decreasing the ParA protein concentration which is consistent with what was seen in this study. Since ParA is already down-regulated twice as much as ParB under nutrient depriving conditions [249], a further down-regulation with antisense



RNA may lead the cells to show a more severe effect which may explain why the effects upon antisense expression of ParA were discernable compared to ParB.

Interestingly, expression of *parA* and *parB* antisense together had severe effects on the cells. This was shown by the *M. smegmatis*-pMind\_*parA+parB*\_GFP strain which had the antisense of both the genes placed in one vector (Figure 4.6). The cells were unable to sustain the basal level of antisense expression from the promoter in the absence of the inducer and underwent suppressor mutations to overcome this (Figure 4.20). This was not totally surprising as expression of *parA* antisense showed defects; however down-regulation still leaves some molecules to carry out its functional process. Further down-regulation of the protein it may bind to (ParB) for functional activity would reduce the activity of ParA to a minuscule level leading to exaggerated phenotypic effects. The fact that this was not tolerated by the cells further confirms the essentiality of ParA.

The above section discussed some of the putative roles of the various genes down-regulated in this work. Most hypotheses are based on genes that are co-expressed as these are likely to have a functional relationship. Transcriptional co-regulation can sometimes indicate interaction between proteins and the association of regulatory patterns of known proteins with unknown proteins can suggest a linkage [198] especially in the case of physically interacting proteins, proteins belonging to the same pathway or genes demonstrating synthetic lethality. Currently, it is not known what genes specifically affect the growth of mycobacteria, however; it is unlikely that one single gene will be able to control the growth rate as the pathways that drive this process are interconnected to other pathways, that is, signalling pathways that will try and prevent cellular demise [210]. Thus, affecting one gene will most likely display pleiotropic effects.

Further experimentation is required to elucidate the precise role of these genes as there are a lot of factors that can influence how a cell behaves in whole cell-based assays with stresses like nutrient starvation and gene down-regulation. Nevertheless, this study sheds some light on the genes tested and whether they are essential under the conditions tested. The only gene that showed an effect upon

antisense RNA induction under all culture conditions was *parA*. It had a lower CFU than the control strain at late exponential phase in all culture conditions when the initial number of cells for all strains was the same showing that *parA* may be essential for growth in *M. smegmatis*.

#### 4.4.1.3.1 Chromosome partitioning in mycobacteria

Bacterial genome segregation is tightly coupled with replication where the newly synthesized origin separates towards the opposite poles to form daughter cells [165]. This process, while crucial, is not very well understood. This study together with data presented by other recent studies indicates that the Par proteins are involved in this process. Results from this chapter demonstrate that ParB is not essential in *M. smegmatis* which was also demonstrated by Jakimowicz *et. al.* [165]. However, Jakimowicz *et. al.* established that while ParB is not essential, absence of ParB created by deletion mutants lead to significant growth inhibition and segregation defect where 10% of *M. smegmatis* cells were anucleate. Recent report of another actinomycete, *C. glutamicum* showed that deletion mutants of either ParA or ParB in *C. glutamicum* also resulted in DNA-free cells [250]. Their study also showed that ParB localises to the poles which is indistinguishable from the localization of the origin. The origin is known to have multiple *parS* sites which has been shown to bind to ParB *in vitro* [250] and thus may indicate that ParB-*parS* colocalize to the poles together. ParB localization to the poles has also been reported in *M. smegmatis* [251] and *M. smegmatis* ParB binds to two *oriC*-proximal *parS* sites for the formation of the nucleoprotein complex. This binding of ParB to *parS* has been proven in plasmid partitioning (section 4.4.1.3) and since ParB is able to bind to *parS* in *M. smegmatis*, it may behave in a similar manner [165]. Both *C. glutamicum* and *M. smegmatis* ParB is occasionally seen at the midcell location and *C. glutamicum* ParB associates with the FtsZ protein which is present at the cytokinetic ring at the midcell. It is hypothesized that the ParB-*parS* complex is directed to midcell as this will be the position of the new poles.

Donovan *et. al.* also showed that ParB is able to interact with ParA and ParA has subcellular localization in *C. glutamicum* [250]. It is hypothesized that ParA binds to the nucleocomplex and pulls it to the pole which requires the ATPase activity of ParA. Similar results were also seen in *C. crescentus* [252]. *M. smegmatis* ParA is found at the poles, at quarter-cell and at midcell position and is believed to oscillate from pole to pole as is seen for Soj (ParA homolog) in *B. subtilis* and *C. crescentus* [251, 252]. This study demonstrated that ParA may be essential in *M. smegmatis* as antisense expression of ParA affected the viability of the cells under all culture conditions and resulted in elongated cells with multiple foci under nitrogen limited conditions.

#### 4.4.1.3.2 Chromosome partitioning in *B. subtilis*

*B. subtilis* also encodes homologues of the ParA and ParB protein, Soj and Spo0J respectively [238]. Unlike most bacterial species, inactivation of *soj* does not produce any chromosomal defects on its own however, it is required for chromosomal partitioning in conjunction with the structural maintenance of chromosome (SMC) protein where the absence of both these genes leads to a increase in anucleate cell production [238]. Interestingly, Spo0J mutants produce a 100 fold more annucleate cells than wild type [253]. Spo0J binds to the *parS* sites [254] and this nucleoprotein is required for the efficient recruitment of the SMC complex (condensin) to the origin [255, 256]. The condensin consists of SMC homodimer binding to the ScpA kleisin and ScpB [255]. The recruitment of the condensin to the origin by Spo0J-parS organizes the origin region. Defects in equipartitioning of DNA have been attributed to the loss of the SMC–ScpA–ScpB complex [238] thus in *B. subtilis*, Soj/Spo0J and SMC/ScpAB complex are required for proper chromosome partitioning [255]. However, Soj has been found to regulate the initiation of replication by regulating DnaA [257]. DnaA recognizes the origin and forms a complex that allows loading of the helicase for DNA replication. This study by Murray and Errington hypothesize that the effects on sporulation seen with Soj mutants may be secondary due its effect on DNA replication via DnaA [257].

#### 4.4.1.3.3 Chromosome partitioning in *E. coli*

The *E. coli* genome does not contain homologues of the Par proteins but encodes an SMC-like condensation complex consisting of 3 subunits, MukB, MukE and MukF. MukB has an N-terminal domain with ATP/GTP-binding (ATPase and GTPase activities) and C-terminal domain with DNA-binding activity [258]. It forms a complex with MukE, a kleisin subunit, via binding to MukF [259]. MukF exhibits sequence homology to ScpA while MukE is similar to ScpB [260]. These proteins are believed to be essential for chromosome partitioning in *E. coli* [261] as deletion of any of the genes leads to annucleate cells [262]. MukB has been reported to colocalize to the poles with the origin region of the chromosome thus MukB may be responsible for organisation the origin [262]. Interestingly, MukB interacts with topoisomerase IV and promotes it relaxation activity *in vitro* [261]. Similar to what was postulated in *B. subtilis*, this may indicate that the chromosome partitioning defects seen may be a result of chromosome replication defect.

A majority of bacterial species encode either of the SMC complex (SMC/ScpA/ScpB or MukBEF). *M. tuberculosis* contains a homolog of these two proteins (Rv2922c) that has not been evaluated in relation to DNA replication and/or chromosome partitioning. Interestingly, it has been shown that ParA is co-expressed with translation elongation factors, beta subunit of the DNA-directed RNA polymerases and transcription termination factor Rho (<http://www.biohealthbase.org>) thus a down-regulation of ParA may also lead to the down-regulation of these crucial factors involved in DNA replication which will in turn cause a decrease in cell division and is not directly related to chromosome partitioning. Looking at the relationship between this and the Par proteins may shed more light on this vital process in mycobacteria.

---

## 4.5 Summary

The aim of this chapter was to validate the essentiality of some of the putative essential genes involved in the cell division process via antisense RNA expression as antisense-knock down can be used to demonstrate gene product essentiality [148]. An important aspect of target validation is to identify genes that are essential under all growth conditions [105]. This is of significance as bacterial cells continuously monitor the availability of nutrients and/or the presence of exogenous stress and have developed complex regulatory networks that enable quick responses [263, 264]. These responses include differential expression of transcriptional regulators that leads to modulation of gene expression allowing the bacteria to adapt to the various stimuli [263].

This study took homologues of three putative essential *M. tuberculosis* genes that may be involved in chromosome partitioning (*obg*, *parA* and *parB*) to test in *M. smegmatis*. Of the three genes tested, only *parA* demonstrated that it may be essential for the viability of the cells under the various culture conditions tested; nutrient rich, carbon starvation and nitrogen starvation conditions. It also exhibited a null phenotype upon antisense expression with *parB* which is believed to encode the protein that may aid in the ATPase activity of the protein encoded by *parA*. Down-regulation of *obg* only showed an effect under rich culture conditions but not under nutrient limitation conditions demonstrating that it is not vital for the viability of the cells at all times.

The final goal was to identify a gene that may be required under all conditions as a potential target for screening for anti-tubercular drugs. This chapter illustrated that *parA* may be a viable target for drug screening.

## **CHAPTER FIVE**

### **Targeted Screening**

## 5.1 Introduction

Broad cell-based screens have led to the identification of a number of antibiotics. However, a major draw-back of these screens is the lack of information on the biochemical targets of the antibiotics identified. This creates a series of problems downstream and has resulted in the drug discovery industry moving towards targeted drug discovery [46, 67, 115, 147, 198].

### 5.1.1 Parameters of a drug screen

As discussed in section 1.8, there are various different assays that can be used for target-based HTS and this calls for a number of different parameters to be evaluated for the selection of the best screening method. The most obvious of these is the antimicrobial activity of the molecules. For a molecule to display antibacterial effects, it has to overcome membrane permeability, evade efflux pumps, prevent degradation in the cell and reach its target in the active form [95]. Cell permeability is particularly a concern with antimycobacterial compounds as mycobacteria are known to have a complex cell wall (Figure 1.2) making them impermeable to most antibiotics [265]. These attributes are overlooked in cell-free biochemical screens and hits require further modification involving expensive medicinal chemistry to accommodate these measures [97]. This makes cell-based assays a superior choice over cell-free assays for initial screening. Another advantage of such whole-cell assays is the ability to screen for compounds in multiple culture conditions [124]. Minimal media conditions are an appropriate choice since *M. tuberculosis* dwells in granulomas where there is believed to be a limitation of nutrients [266].

A significant factor that needs to be considered with a targeted screen is the selectivity of the compounds to the target. This aspect can be quickly and easily addressed using bioinformatics and genomics where similar targets in other microbes and mammals can be identified. Known targets can also be used in mode of action studies and can be used to uncover heritable resistance determinants.

Both of these are difficult to elucidate in broad cell-based assays when the target is unknown. This requires a lot more post-screening labour [147] thus making target based assays a preferred option.

By far, the most important feature of targeted screens in today's drug development program is ability to use the MOA in SAR analysis. This can proceed onto rational drug design using combinatorial chemistry as most antimicrobials in use today have undergone some semi-rational optimization [198]. The lead optimization of weak ligands using SAR can greatly increase the potency and can also be modified with respect to specificity [105]. This again, cannot be done when the target is unknown.

### **5.1.2 Differential susceptibility assays**

The goal of any HTS is to identify a molecule that will influence the overall fitness of the microbe. This requires the bioactive molecule to inhibit the expression of a target gene(s). This in turn will affect the activity of the protein encoded by the gene and thus have an effect on the whole organism. Such discoveries are greatly accelerated using chemogenomics which combines the effect of compounds on biological targets together with modern genomics technologies. This includes knockdown experiments where there is a reduced level of target protein in the cell which is further reduced upon exposure to specific inhibiting molecules [267]. Such modular assays using bacterial genetics offer a number of advantages. Firstly, these assays make targets hypersensitive and screening more effective as the hits isolated from these screens would otherwise be missed. Secondly, once a modulation technique is optimized, the same procedure can be used on many different targets providing a technically simple, inexpensive and permanent tool as the engineered strains can be used in multiple screens. And lastly, such assays offer a means to identify inhibitors of uncharacterized targets which do not have a biochemical assay. It also aids in MOA studies of the compound as modulation of the target should relate to variations on the minimum inhibitory concentration and may produce phenotypic



changes that may assist in its functional elucidation [67]. Thus cellular responses upon exposure to small molecules can validate a therapeutic target before the direct biochemical function of the target is fully clarified [268].

One such modulation technique which has demonstrated its success in drug screening is the expression of antisense RNA (section 4.1.1.1). This procedure requires little genetic manipulation compared to mutants created via homologous recombination and provides an invaluable genetic tool [269]. Additionally, this controlled gene expression system can be transferred to test the activity of active compounds *in vivo*.

Below are some examples of inhibitors that have been verified using antisense RNA expression system:

- The MOA of the polypeptide deformylase (PDF) in *Staphylococcus aureus* was confirmed using a TetR-regulated system that displayed increased susceptibility to a PDF specific inhibitor upon partial inhibition of PDF expression [269].
- Increased susceptibility of *ileS* and *fabI* to mupirocin and triclosan, their respective inhibitors, was demonstrated upon antisense RNA expression [270].
- Platensimycin, inhibitor of the enzyme FabF protein involved in the fatty acid biosynthesis pathway and a novel broad-spectrum antibiotic, was discovered by differential sensitivity assay [271]

Keeping in mind the above discussed parameters of a HTS, a cell-based targeted approach was employed to screen for inhibitors against *parA*, the gene encoding a Walker type ATPase.

### 5.1.3 ATPases

ATPases function in a broad variety of biochemical process. They catalyse the hydrolysis of adenosine triphosphate (ATP) molecules into adenosine diphosphate (ADP) molecules which releases an inorganic phosphate ( $P_i$ ). The dephosphorylation of the high energy phosphate bond releases energy which is then used to drive other chemical reactions. The most common type of ATPases comprise of the conserved Walker motifs, the form seen in ParA according to sequence homology. The Walker motifs contain the characteristic phosphate binding loop (P-loop) [272]. The P-loop is the mononucleotide binding fold and is also conserved in GTPases [273, 274]. Since the P-loop is so highly conserved, the ATPase active site has not been a very attractive target for drug discovery. However, as their structures have been further elucidated, it has been found that there are different classes of enzymes that have the P-loop and they have diverse nucleotide binding sites [275]. This has made it appealing to the drug discovery industry to design inhibitors that will only target specific types of ATPases.

There can be two different modes of actions for ATPase inhibitors; it can either bind to the active site and prevent ATP from binding (competitive inhibitors and/or adenine mimics) or it can bind to some other site on the ATPase that changes the conformation of the active site thus preventing ATP from binding to it (non-competitive inhibitors) [276]. Competitive inhibitors have to be highly selective however; various ATPases inhibitors have been designed with specificity to the target ATPases, that is, DNA gyrases and protein kinases inhibitors [277-279]. Such work demonstrates that targeting ParA is a plausible strategy for identifying antimicrobials to combat disease.

## 5.2 Objectives of chapter five

The main aim of this chapter was to find molecules that will target the cell division process in mycobacteria.

As discussed in chapter 4, ParA indicated that it might be essential for cell division in *M. smegmatis* under all the culture conditions tested in this work making it a viable target for drug screening. I hypothesized that any molecule affecting ParA will result in hypersensitivity and may exhibit phenotypic changes that are consistent with inhibiting cell division genes, that is, filamentous cell morphology [215, 221].

In order to test this hypothesis and discover such molecules, *M. smegmatis*-pMind\_*parA*\_GFP strain was subjected under antisense inducing conditions to differential susceptibility assays in conjunction with compound libraries under various culture conditions. My main objectives were to:

- Run HTS of *M. smegmatis*-pMind\_*parA*\_GFP with a reference strain, both under antisense inducing conditions in various culture conditions
- Isolate compounds that exhibited hypersensitivity in the *M. smegmatis*-pMind\_*parA*\_GFP strain only
- Perform microscopy to identify if the compounds exhibiting hypersensitivity also displayed filamentous cell morphology in the *M. smegmatis*-pMind\_*parA*\_GFP strain compared to the reference strain.

### 5.3 Results

All HTS were performed in triplicate with a control strain *M. smegmatis*-pMind2357. *M. smegmatis*-pMind2357 contains the MSMEG\_2357 gene in the antisense orientation under the tetracycline promoter of pMind. MSMEG\_2357 is a cysteine desulfurase, homologue of which is an essential gene in *M. tuberculosis* (Rv3025c; Supplementary Table S1- Appendix 9.2.2). This strain was closest in profile to *M. smegmatis*-pMind\_*parA*\_GFP in growth assays with respect to OD and CFU (data not shown) and MSMEG\_2357 is not involved in cell division thus serving as an ideal reference strain in the HTS.

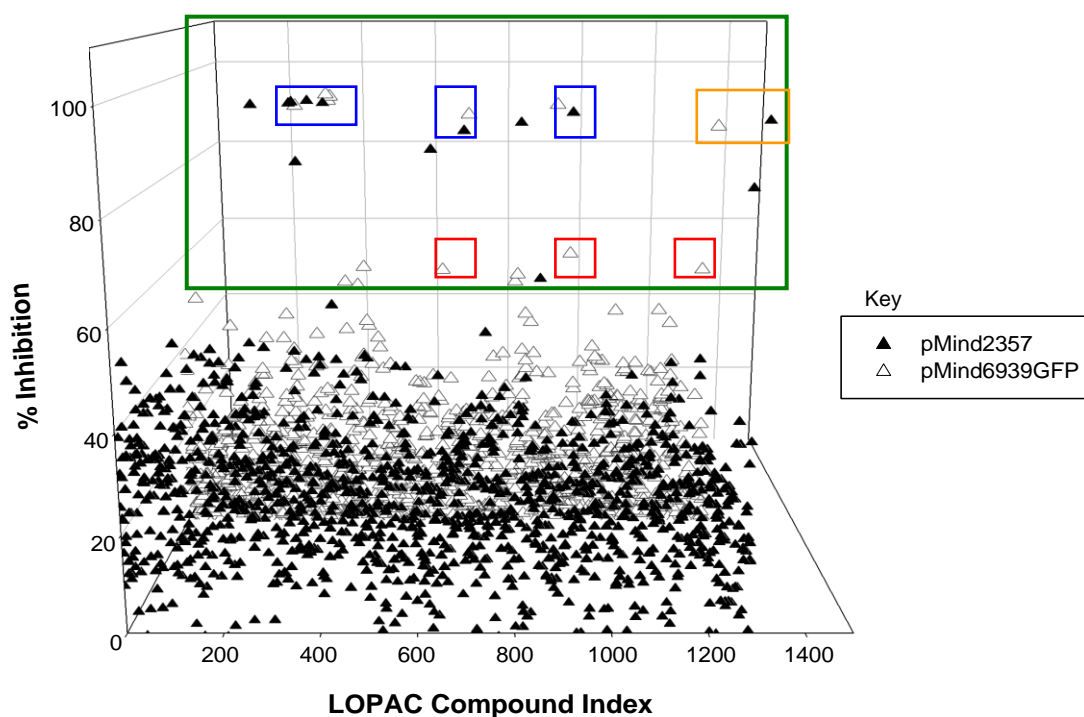
All assays were set-up as described in section 2.13.1.1 under antisense inducing condition with 20 ng/mL Tc in LBT, carbon starvation (C<sup>-</sup>) and nitrogen starvation (N<sup>-</sup>) conditions. For the purpose of this thesis, only hits from the *M. smegmatis*-pMind\_*parA*\_GFP were analysed. Hits were also compared to wild type strains of *M. smegmatis* [124] to eliminate molecules that were not picked up in the reference strain. All wells that exhibited a 40% or more inhibition in at least two of the 3 screens were analyzed.

This chapter presents all the hits obtained through the various HTS screens and discusses their potential as inhibitors of ParA. A total of 69 hits were identified from screening the *M. smegmatis*-pMind\_*parA*\_GFP strain against the compound libraries under three different culture conditions.

### 5.3.1 LOPAC Screen

LOPAC (Library of Pharmacological Active Compounds) has 1280 compounds representing all major target classes. There were a total of 37 hits against *M. smegmatis*-pMind\_ *parA*\_GFP from screening LOPAC; nine in LBT (Table 5.1), four in C<sup>-</sup> (Table 5.2) and twenty-four in N<sup>-</sup> (Table 5.3) conditions.

#### 5.3.1.1 LBT (LOPAC)

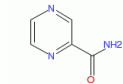
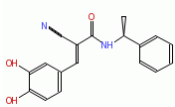
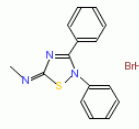
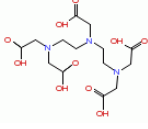
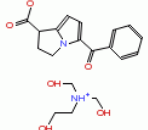
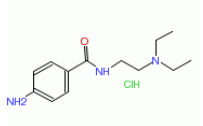
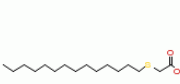
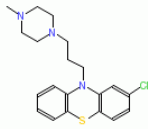
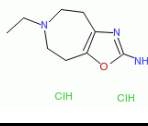


**Figure 5.1: Graphical representation of LOPAC screens in LBT**

This graph shows the average percentage inhibition of *M. smegmatis*-pMind2357 and *M. smegmatis*-pMind\_ *parA*\_GFP strains in LBT against LOPAC. The green box indicates all compounds that exhibited over 40% inhibition. The orange box indicates the positive control ethambutol; the blue boxes show inhibition in both strains and the red boxes show some molecules that exhibited inhibition in the target strain. Error bars not shown for clarity.

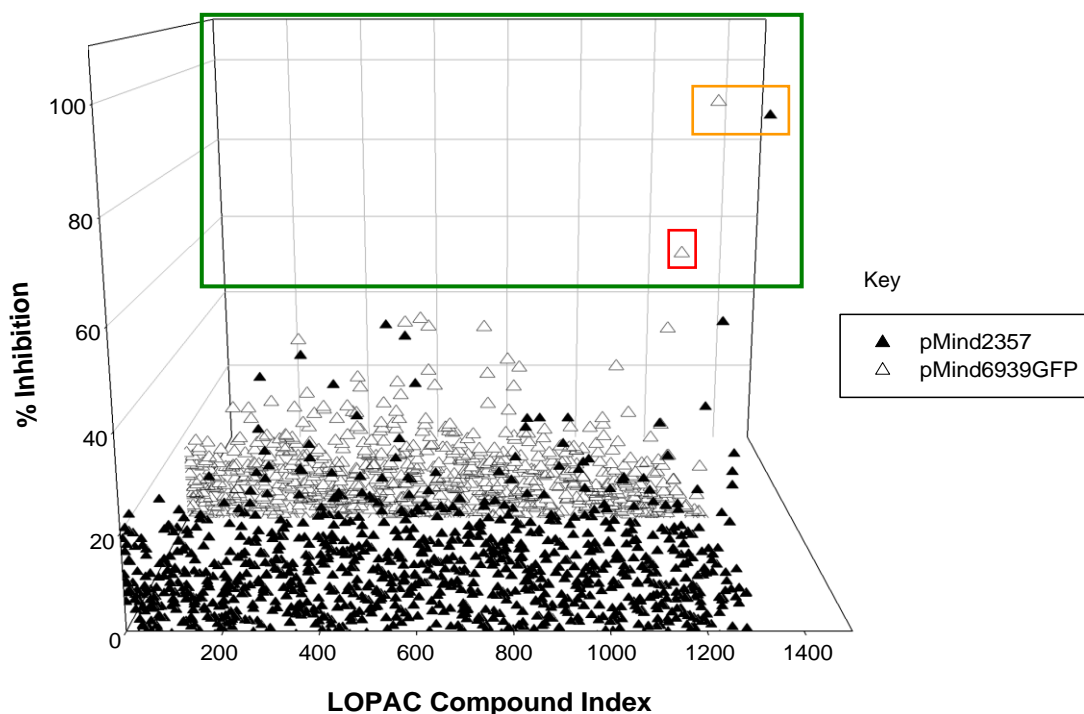
**Table 5.1: Hits from LOPAC in LBT**

Nine HTS hits against ParA from LOPAC in rich culture conditions.

Cpd No.	Cpd name	Compound structure	MW (g/mol)	Description/ Action
1	Pyrazinecarboxamide		123.12	Antibiotic; tuberculostatic
2	Tyrphostin AG 698		308.34	EGFR protein tyrosine kinase inhibitor
3	SCH-202676 hydrobromide		348.27	Allosteric agonist and antagonist of G protein coupled receptors (GPCRs)
4	Diethylenetriaminepentaacetic acid		393.35	Zinc-dependent proteases carboxypeptidase A and aminopeptidase N inhibitor; chelating agent
5	Ketorolac tris salt		376.41	COX (enzyme) inhibitor. Non-steroidal anti-inflammatory (NSAID) drug
6	Procainamide hydrochloride		271.79	Na <sup>+</sup> channel blocker and Class I anti-arrhythmic
7	Tetradecylthioacetic acid		288.50	Transcription agonist-Peroxisome proliferator-activated receptor (PPAR)-alpha agonist
8	Prochlorperazine dimaleate		606.10	Dopamine antagonist-Antipsychotic agent; used in the treatment of spastic gastrointestinal disorders
9	B-HT 933 dihydrochloride		254.16	Selective alpha2 adrenoceptor agonist

All compounds in Table 5.1 exhibited more toxicity against the ParA antisense expressing strain compared to the reference strain. Cells from the wells containing these compounds were analysed under a fluorescence microscope for any discernable morphological changes. This analysis revealed that albeit showing a difference in the optical density between the two strains, there were no significant differences in the cellular morphology.

### 5.3.1.2 C<sup>-</sup> (LOPAC)

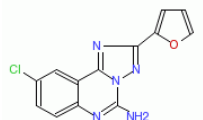
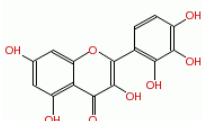
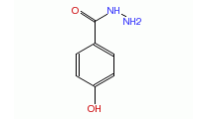
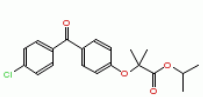


**Figure 5.2: Graphical representation of LOPAC screens in C<sup>-</sup> media**

This graph shows the average percentage inhibition of *M. smegmatis*-pMind2357 and *M. smegmatis*-pMind<sub>parA</sub>GFP strains in C<sup>-</sup> media against LOPAC. The green box indicates all compounds that exhibited over 40% inhibition. The orange box indicates the positive control (ethambutol) and the red box shows inhibition in the target strain. Error bars not shown for clarity.

**Table 5.2: Hits from LOPAC in C<sup>-</sup> media**

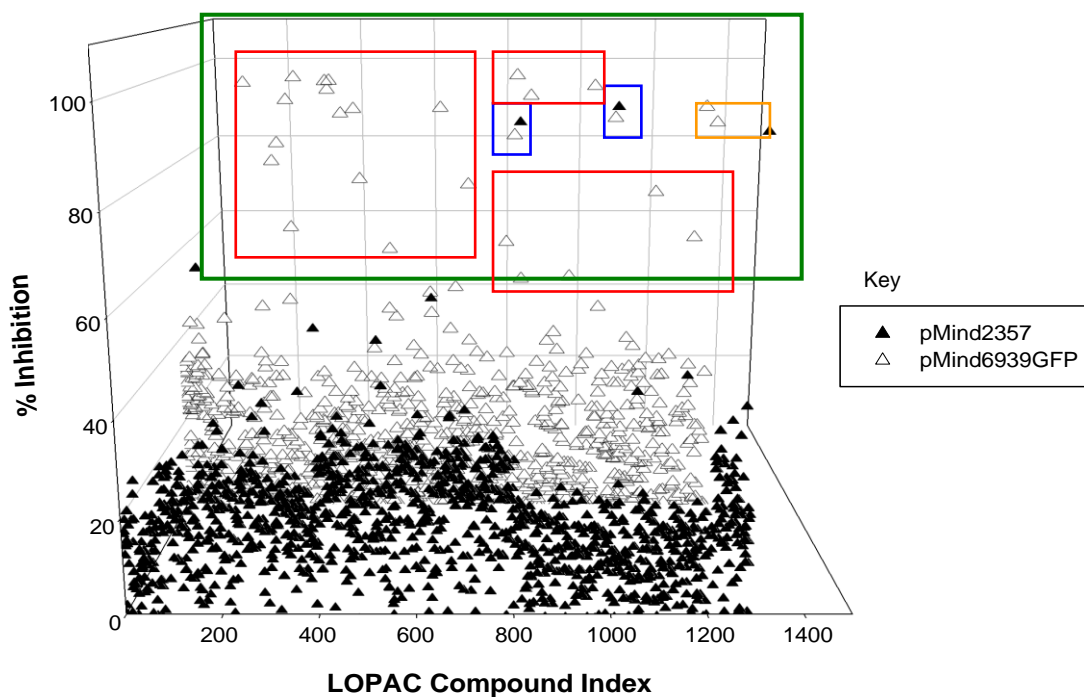
Four HTS hits against ParA from LOPAC under carbon starvation conditions.

Cpd No.	Compound name	Compound structure	MW (g/mol)	Description/ Action
10	CGS-15943		285.69	Highly potent, non-selective A1 adenosine receptor antagonist
11	Myricetin		318.24	Casein Kinase II inhibitor
12	4-Hydroxybenzhydrazide		152.15	Phenylalanine ammonia-lyase inhibitor; used to measure reducing sugar levels
13	Fenofibrate		360.84	Transcription agonist-PPAR-alpha receptor agonist; lipid regulating drug

ParA antisense expressing strains containing the compounds shown in Table 5.2 showing greater sensitivity in carbon starvation were analysed under the fluorescence microscope and were found to have no phenotypic effects compared to the control cells containing the same compounds.



### 5.3.1.3 N<sup>-</sup> (LOPAC)



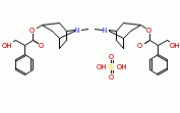
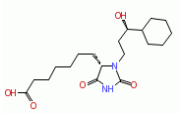
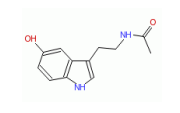
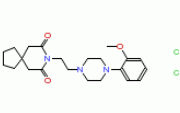
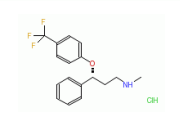
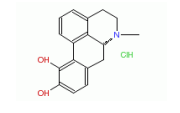
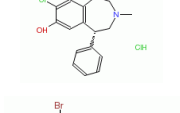
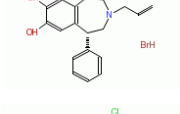
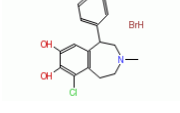
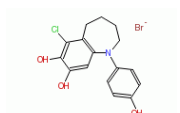
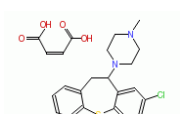
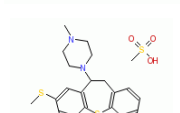
**Figure 5.3: Graphical representation of LOPAC screens in N<sup>-</sup> media**

This graph shows the average percentage inhibition of *M. smegmatis*-pMind2357 and *M. smegmatis*-pMind<sub>parA</sub>-GFP strains in N<sup>-</sup> media against LOPAC. The green box indicates all compounds that exhibited over 40% inhibition. The orange box indicates the positive control ethambutol, the blue boxes shows inhibition in both strains and the red boxes shows inhibition in the target strain. Error bars not shown for clarity.

**Table 5.3: Hits from LOPAC in N<sup>-</sup> media**

Twenty-four HTS hits against ParA from LOPAC under nitrogen starvation conditions.

Cpd No.	Compound name	Compound structure	MW (g/mol)	Description/ Action
14	Xanthine amine congener		428.50	A1 Adenosine receptor antagonist
15	N6-Phenyladenosine		343.34	Potent A1 adenosine receptor agonist
16	Amsacrine hydrochloride		429.93	DNA topoisomerase II inhibitor
17	Anthrapyrazolone; 1,9-Pyrazoloanthrone		220.23	Selective c-Jun N-terminal kinase (c-JNK) inhibitor.
18	ODQ		187.16	Potent and selective NO-sensitive guanylyl cyclase inhibitor
19	Palmitoyl-DL-Carnitine chloride		436.08	Long-chain acylcarnitine; modulator of PKC activation; intermediate in mitochondrial fatty acid oxidation
20	Diphenhydramine hydrochloride		291.82	H1 Histamine receptor antagonist
21	Sulindac sulfone		372.42	Enzyme inhibitor- Non-steroidal anti-inflammatory compound (NSAID); metabolite of sulindac
22	AA-861		326.44	5-lipoxygenase inhibitor
23	AL-8810		402.51	Prostaglandin F2alpha analog with selective antagonist effects at the prostaglandin F2alpha (FP) receptor
24	Chloroquine diphosphate		515.87	Antimalarial; anti-amoebic; anti-inflammatory; binds strongly to double-stranded DNA
25	(±)-Nipecotic acid		129.16	GABA uptake inhibitor

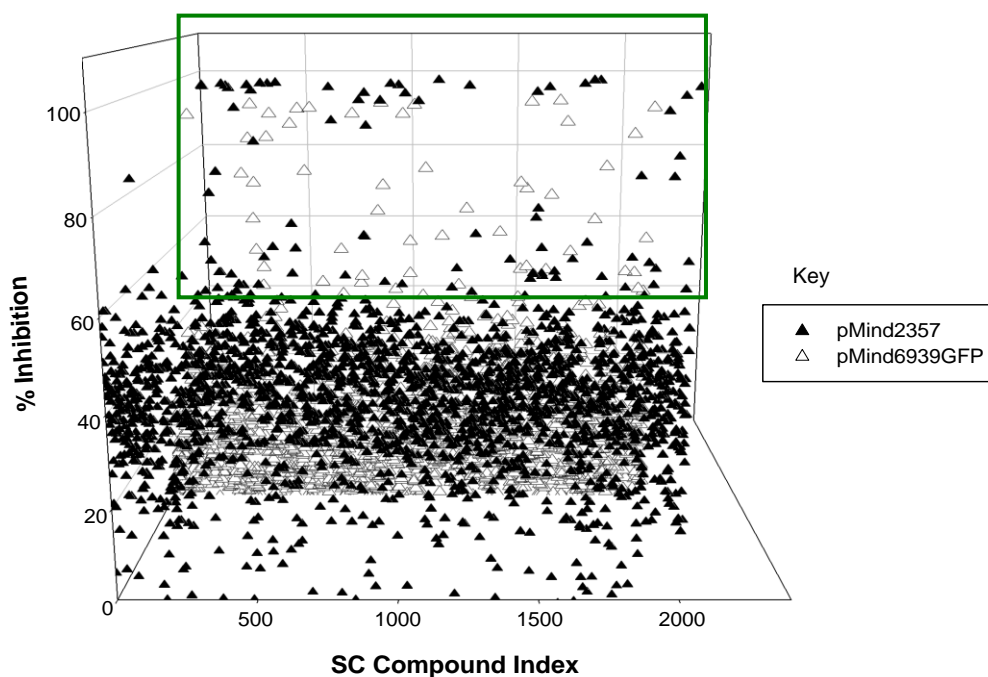
26	Atropine sulfate		676.83	Muscarinic receptor	acetylcholine antagonist
27	BW 245C		368.48	Potent, prostanoid receptor	selective DP agonist
28	N-Acetyl-5-hydroxytryptamine		218.26	Immediate precursor of melatonin	
29	BMY 7378 dihydrochloride		458.43	Partial 5-HT1A receptor agonist; alpha1D adrenoceptor antagonist	serotonin
30	R(-)-Fluoxetine hydrochloride		345.80	Selective serotonin reuptake inhibitor	
31	Apomorphine hydrochloride hemihydrate		303.79	Non-selective dopamine receptor agonist	
32	R(+)-SCH-23390 hydrochloride		324.25	Selective D1 dopamine receptor antagonist	
33	R(+)-6-Bromo-APB hydrobromide		455.19	D1 Dopamine receptor agonist	
34	SKF 83565 hydrobromide		419.15	Atypical D1 dopamine receptor agonist; displays antagonist activity in vitro and agonist activity in vivo	
35	Fenoldopam bromide		386.68	Peripheral D1 dopamine receptor agonist	
36	(±)-OctoclothePIN maleate		460.98	D2 Dopamine receptor antagonist; serotonin receptor antagonist	
37	Methiothepin mesylate		452.66	5-HT1 Serotonin receptor antagonist; blocks serotonin autoreceptors	

The cells from the hit wells against ParA in  $N^-$  media (Table 5.3) were analyzed via a fluorescence microscope and compared to the control cells. Six out of the twenty-five compounds exhibited filamentous morphology. These were amsacrine hydrochloride, AA-861, R(-)-fluoxetine hydrochloride, fenoldopam bromide, (±)-octoclothePIN maleate and methiothepin mesylate (Appendix 9.4.1).

### 5.3.2 Spectrum Collection

The Spectrum Collection has 2000 compounds with a wide range of structures having various biological activities with 50% representing drugs, 30% representing natural products and 20% representing other non-drug bioactive compounds. There were a total of seven hits against *M. smegmatis*-pMind\_*parA*\_GFP from screening the Spectrum Collection; five in LBT (Table 5.4), one in C<sup>-</sup> (Table 5.5) and one in N<sup>-</sup> (Table 5.6) conditions. As can be seen from the graphs (Figure 5.6, 5.7 and 5.8), the lack of hits in this library owes to the fact that most of the hits came up in the reference strain as well.

#### 5.3.2.1 LBT (Spectrum Collection)

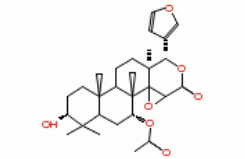
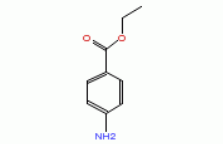
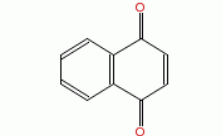
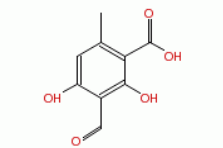
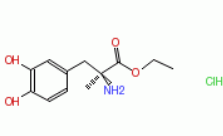


**Figure 5.4: Graphical representation of Spectrum Collection screens in LBT**

This graph shows the average percentage inhibition of *M. smegmatis*-pMind2357 and *M. smegmatis*-pMind\_*parA*\_GFP strains in LBT against Spectrum Collection. The green box indicates all compounds that exhibited over 40% inhibition. Error bars not shown for clarity.

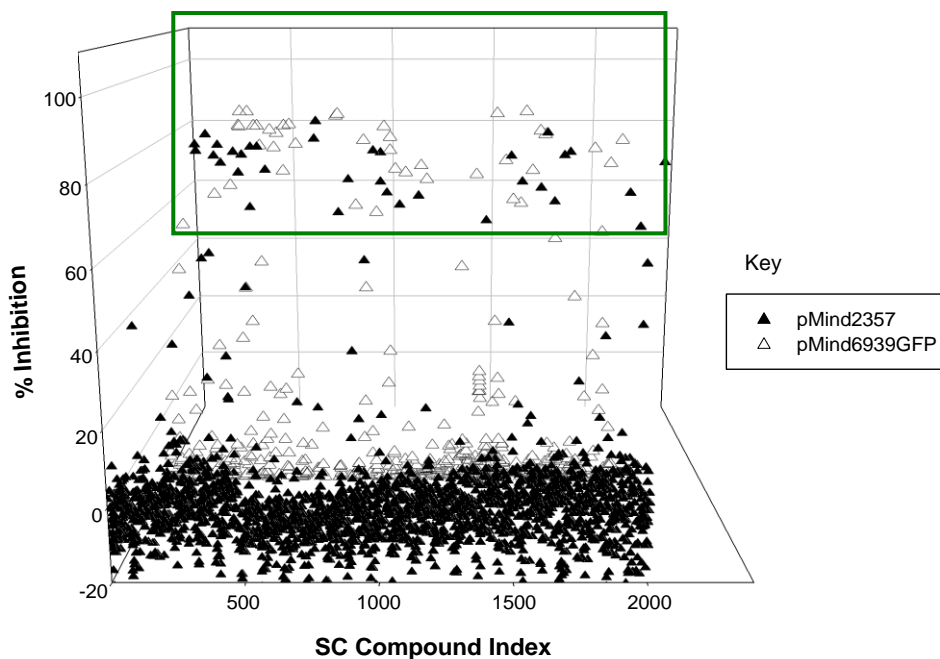
**Table 5.4: Hits from Spectrum Collection in LBT**

Five HTS hits against ParA from Spectrum Collection under rich culture conditions. UA = undetermined activity; E = experimental

Cpd No.	Compound name	Compound structure	MW (g/mol)	Description/ Action
38	alpha-Dihydrogedunol		486.61	Derivative, UA
39	Benzocaine		165.19	anesthetic (topical)
40	1,4-Naphthoquinone		158.16	coal tar product, E
41	Haematommic Acid		196.16	UA
42	Methyldopate Hydrochloride		275.73	Antihypertensive; Adrenergic alpha-Agonists

All cells from the hit wells were analysed for morphological differences and compared to the control cells with the same compound. There were no differences in morphology between the two strains with any of the compounds shown in Table 5.4.

### 5.3.2.2 C<sup>-</sup> (Spectrum Collection)



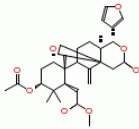
**Figure 5.5: Graphical representation of Spectrum Collection screens in C<sup>-</sup> media**

This graph shows the average percentage inhibition of *M. smegmatis*-pMind2357 and *M. smegmatis*-pMind<sub>parA</sub>\_GFP strains in C<sup>-</sup> media against Spectrum Collection. The green box indicates all compounds that exhibited over 40% inhibition. Error bars not shown for clarity.

**Table 5.5: Hits from Spectrum Collection in C<sup>-</sup> media**

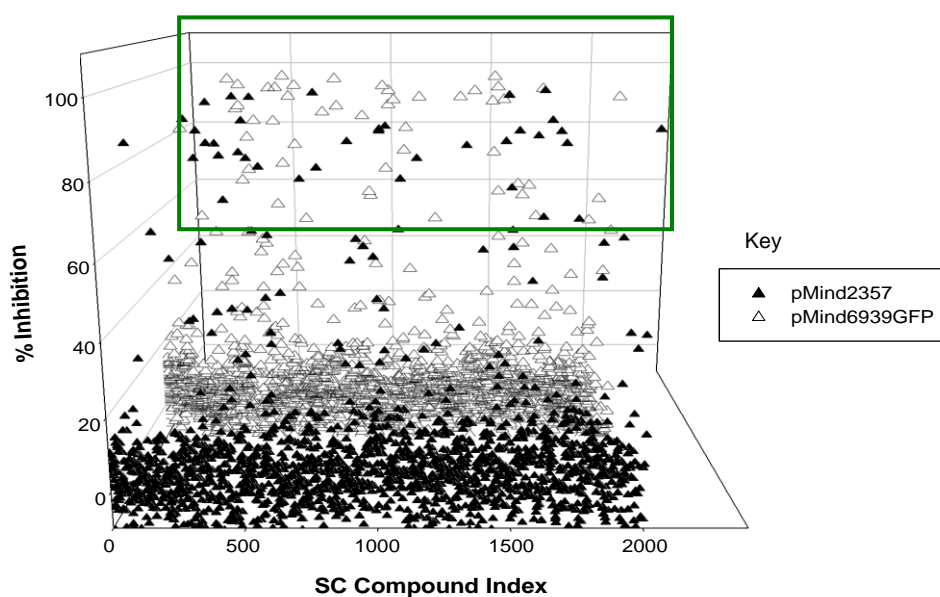
One HTS hit against ParA from Spectrum Collection under carbon starvation conditions.

UA = undetermined activity

Cpd No.	Compound name	Compound structure	MW (g/mol)	Description/ Action
43	3beta-Acetoxydeoxyangolensic Acid, Methyl Ester		514.62	UA

The hit from Spectrum Collection under carbon starvation condition against ParA did not display the morphological change that was expected upon perturbing cell division.

### 5.3.2.3 N<sup>-</sup> (Spectrum Collection)

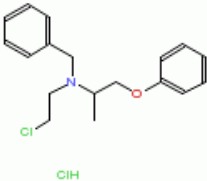


**Figure 5.6: Graphical representation of Spectrum Collection screens in N<sup>-</sup> media**

This graph shows the average percentage inhibition of *M. smegmatis*-pMind2357 and *M. smegmatis*-pMind<sub>parA</sub>GFP strains in N<sup>-</sup> media against Spectrum Collection. The green box indicates all compounds that exhibited over 40% inhibition. Error bars not shown for clarity.

**Table 5.6: Hits from Spectrum Collection in N<sup>-</sup> media**

One HTS hit against ParA from Spectrum Collection under nitrogen starvation conditions.

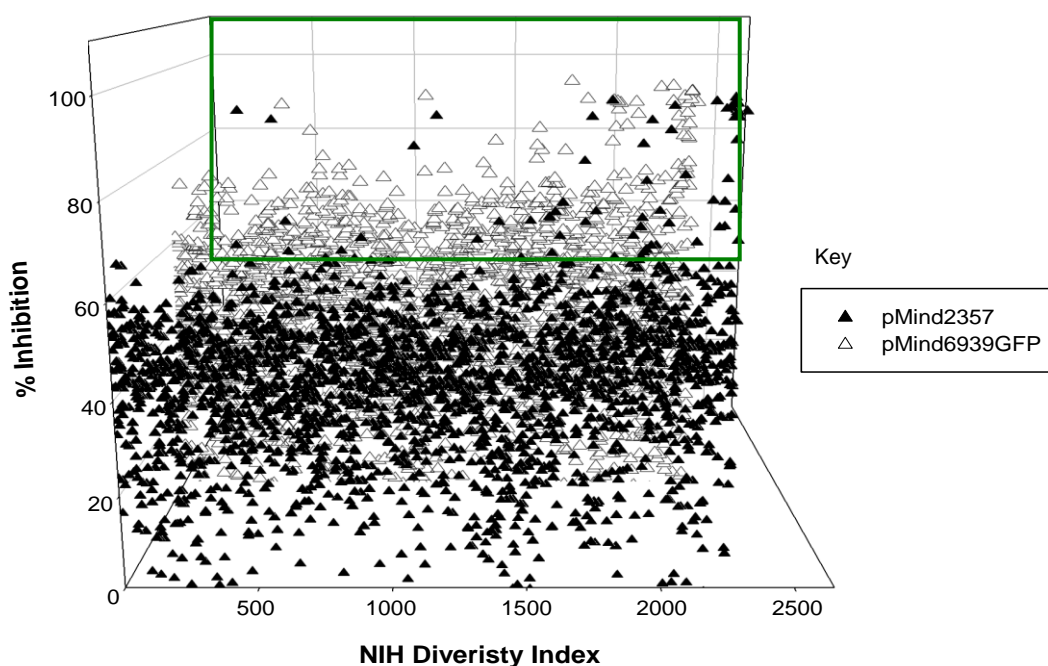
Cpd No.	Compound Name	Compound structure	MW (g/mol)	Description/ Action
44	Phenoxybenzamine Hydrochloride		340.287	Non-selective alpha-adrenergic antagonist [280]

The hit against ParA from the Spectrum Collection in N<sup>-</sup> media displayed the filamentous morphology that is expected when cells are inhibited from proper cell division. This was not seen in the control strain (Appendix 9.4.2).

### 5.3.3 NIH Diversity Set

The NIH Diversity Set consists of 2400 compounds with a broad range of structures derived from about 140,000 compounds. There were a total of twenty-five hits against *M. smegmatis*-pMind\_*parA*\_GFP from screening the NIH Diversity Set; twenty-one in LBT (Table 5.7), three in C<sup>-</sup> (Table 5.8) and one in N<sup>-</sup> (Table 5.9) conditions.

#### 5.3.3.1 LBT (NIH Diversity Set)



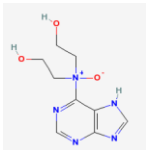
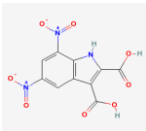
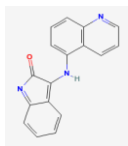
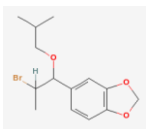
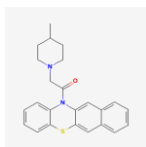
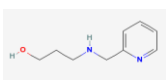
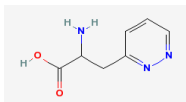
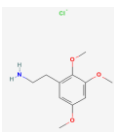
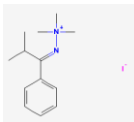
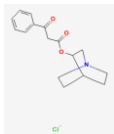
**Figure 5.7: Graphical representation of the NIH Diversity Set screens in LBT**

This graph shows the average percentage inhibition of *M. smegmatis*-pMind2357 and *M. smegmatis*-pMind\_*parA*\_GFP strains in LBT against the NIH Diversity Set. The green box indicates all compounds that exhibited over 40% inhibition. Error bars not shown for clarity.



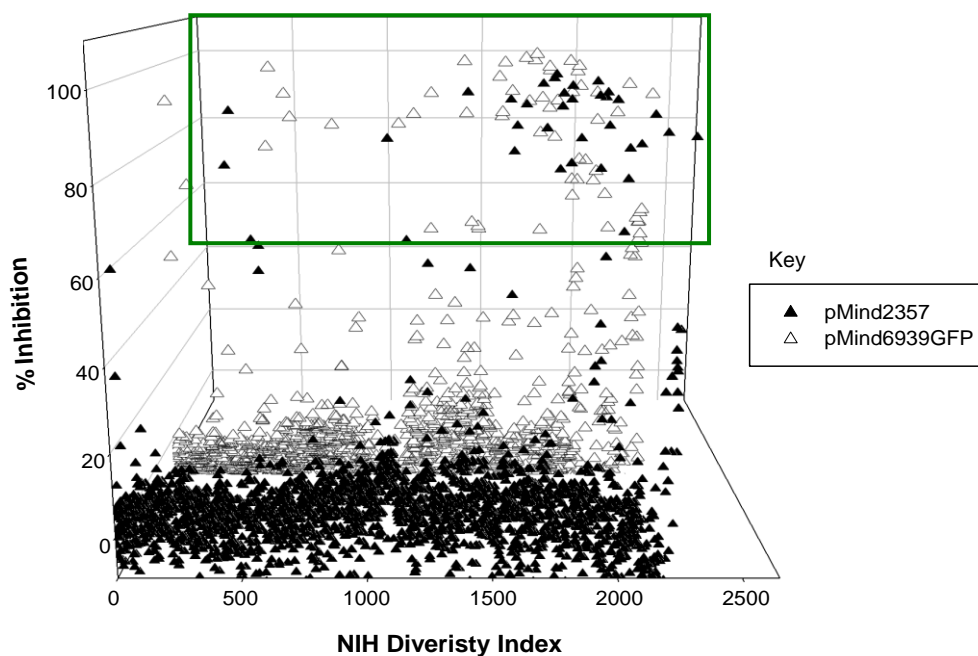
**Table 5.7: Hits from NIH Diversity Set in LBT media**

Twenty-one HTS hits against ParA from the NIH Diversity Set under rich culture conditions. UA = undetermined activity

Cpd No.	Compound Name/ ID	Compound structure	MW (g/mol)	Description/ Action
45	NSC57850		239.23	Perturbs 3 pathways: cholesterol metabolism, proteasome and oxidative phosphorylation
46	NSC131615		295.16	Targets human cytochrome P450
47	NSC117198		273.28	Targets human cytochrome P450
48	NSC195151		315.2	Targets human cytochrome P450
49	NSC116490		388.52	Regulation of actin cytoskeleton, GTPase inhibitor activity, transcription regulation
50	NSC65650		166.22	UA
51	3-Pyridazinealanine		167.165	Targets voltage-gated K <sup>+</sup> channel (human)
52	NSC53069		246.71	Nucleotide binding, affects ATP binding, GTPase and serine/threonine kinase activity
53	NSC138617		332.22	Targets voltage-gated K <sup>+</sup> channel (human)
54	NSC83391		308.78	Targets voltage-gated K <sup>+</sup> channel (human)



### 5.3.3.2 C<sup>-</sup> (NIH Diversity Set)

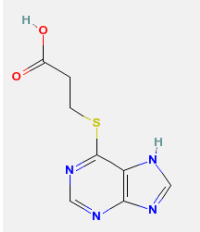
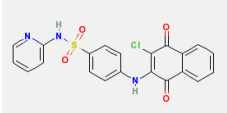
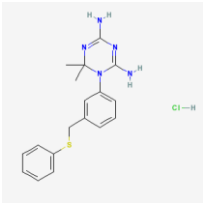


**Figure 5.8: Graphical representation of the NIH Diversity Set screens in C<sup>-</sup> media**

This graph shows the average percentage inhibition of *M. smegmatis*-pMind2357 and *M. smegmatis*-pMind<sub>parA</sub>GFP strains in C<sup>-</sup> media against the NIH Diversity Set. The green box indicates all compounds that exhibited over 40% inhibition. Error bars not shown for clarity.

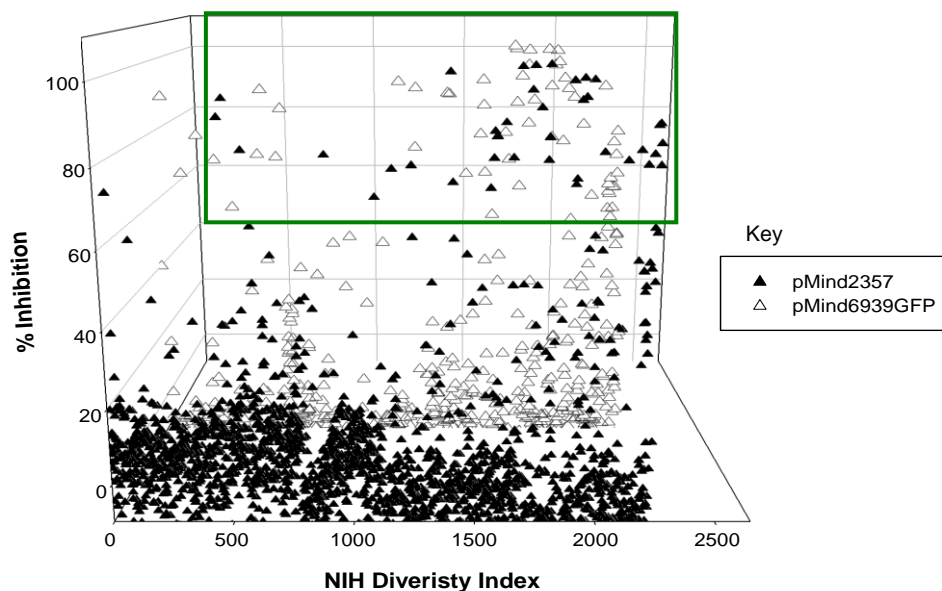
**Table 5.8: Hits from NIH Diversity Set in C<sup>-</sup> media**

Three HTS hits against ParA from the NIH Diversity Set under carbon starvation conditions. UA = undetermined activity.

Cpd No.	Compound Name/ ID	Compound structure	MW (g/mol)	Description/ Action
66	NSC22242		224.24	UA
67	NSC45382		439.871	Targets human cytochrome P450, prostaglandin E receptor 2, Hiv-2 Reverse Transcriptase, RNase HI
68	NSC368891		375.918	UA

Cellular morphology of all hits against ParA in C<sup>-</sup> from the NIH Diversity Set (Table 5.8) were analysed under a fluorescence microscope. Of the three hits, only compound 66 showed the expected phenotype (Appendix 9.4.3).

### 5.3.3.3 N<sup>-</sup> (NIH Diversity Set)

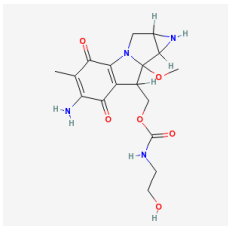


**Figure 5.9:** Graphical representation of the NIH Diversity Set screens in N<sup>-</sup> media

This graph shows the average percentage inhibition of *M. smegmatis*-pMind2357 and *M. smegmatis*-pMind<sub>parA</sub>-GFP strains in N<sup>-</sup> media against the NIH Diversity Set. The green box indicates all compounds that exhibited over 40% inhibition. Error bars not shown for clarity.

**Table 5.9: Hits from NIH Diversity Set in N<sup>-</sup> media**

One HTS hit against ParA from the NIH Diversity Set under nitrogen starvation conditions. UA = undetermined activity.

Cpd No.	Compound Name/ ID	Compound structure	MW (g/mol)	Description/ Action
69	Mitomycin C		378	antibiotic isolated from the broth of <i>Streptomyces caespitosus</i> , selectively inhibits DNA synthesis, cellular RNA and protein synthesis are also suppressed

## 5.4 Discussion

The aim of this chapter was to identify small molecules as hits against ParA that may potentially be developed as ParA inhibitors. This was performed using a whole-cell differential susceptibility assay. This assay is based on reverse chemical genetics approach with a gene driven analysis based on phenotypic changes and/or sensitivity created by the small molecules upon modulation of the target protein using antisense RNA.

The first thing to take into consideration in this case is the justification for using ParA as a target. There are several arguments that support ParA as a plausible target. Firstly, as discussed in chapter 3, ParA cannot sustain transposon mutagenesis [126] and it appears to be required for growth under various culture conditions (chapter 4). Secondly, it is involved in an essential cellular pathway, that is, cell division [130, 234, 236, 238-240, 246, 247, 282-284] and thirdly, it is an enzyme [243, 282, 283, 285]. There are a number of advantages to using enzymes as targets. They have been widely used as targets for antibacterials and it is easy to develop functional assays for them [286]. Some even consider all bacterial enzymes to be potential drug targets [287]. Most eminently though, enzymes interact with small ligands and this increases the likelihood of them binding small molecule inhibitors [148]. Furthermore, ParA is an ATPase (section 5.1.3). It is believed that inhibitors of some classes of enzymes are more effortlessly identified than others [46]. This makes ParA, an ATPase, an optimal target as ATPase inhibitors are well established [288-292]. As discussed in section 4.4.1.3, ParA may bind to another protein ParB. This interaction can also be exploited for drug discovery as inhibiting protein-protein interaction has lead to successful drugs [293]. Lastly, ParA has not been previously targeted by any antibiotics thus exploitation of this unique target may reveal compounds with a novel mechanism of action.

### 5.4.1 Goals of targeted screening

Current HTS employed by large pharmaceutical and biotechnology companies involves screening for hundreds of thousands and even millions of compounds [294, 295]. However, such screens are expensive and require the development of cost-effective ways of screening. With this in mind, this study implemented a so-called ‘smart screening’ strategy [296]. This involved screening a relatively small but diverse set of compounds (between 2000 to 5000) in an attempt to learn what to screen next.

The LOPAC library provided a diverse set of compounds as it contains active compounds from an assortment of biological arenas. It also contains some gold standards with well characterised mode of actions for assay validation. Since LOPAC is so extensively used, it was ideal to use in the case of ParA as this is a unique target increasing the chances of excluding the inhibitors of known targets and picking out novel hits as the success of exploiting drugs with known targets is extremely low [112]. The Spectrum Collection also fell in the category of structurally diverse compounds (section 5.3.2). Lastly, the NIH diversity Set included a range of structures which were derived from 3-point pharmacophores as a means to enhance the overall presentation of chemical diversity. The Spectrum Collection and the NIH Diversity Set also contain experimental bioactives and natural products increasing the potential of identifying novel structures against ParA. The resulting structures from the overall screening of the three libraries is not target biased, but are unique, providing useful starting structures for *in silico* screening of an uncharacterised target.

Finally, apart from the libraries used, the assays used for HTS are also coming under scrutiny for the generation of high quality hits and the removal of false positives from the hit lists. It is worth mentioning that no one screening strategy is void of artefacts, thus the successful development of any antibiotic requires more than one assay [112]. The method used for initial screening and hit identification in this study was a differential susceptibility assay (section 5.1.2). The reasons for using a screen as such are discussed in section 5.1.1. There is also

a caveat to this assay which can be used as an advantage. As attractive and common as it is to target a single protein, this concept can be very deceiving. It is a known fact that genes do not function in solitude. They form networks that interact with exquisite complexity [297] and it has been shown through transcription profiling that microbes undergo global alteration in transcription regulation upon sensing a change in its environment [298-302]. This sheds a different light on things, as a target may be essential but inhibiting the target with a molecule can change the transcription profile and perhaps even render the target non-essential under these conditions. Inhibiting the target can also lead to the upregulation of genes with similar functionality to overcome its handicap. Thus the changes in the network connectivity by molecules modulating the target gene can potentially uncover biological redundancy in the system [303] and it can also lead to the upregulation of other druggable targets. Thus this assay can validate a target with respect to functional connectivity with other genes. Identification of other genes in the same pathway that causes synthetic lethality can be used to exploit ParA further for drug targeting. Such experiments has been successfully conducted in *M. tuberculosis* with other genes [304]. Additionally, these targets can be subjected to combination drug therapy to exhibit more potent MIC as has been shown for meropenem with clavulonate [305].

#### **5.4.2 Biological aspects of the hits**

For the purpose of hit identification in this study, a cut-off of 40% inhibition was chosen. A 40% cut-off was decided upon from the data output of the HTS. All compounds exhibiting greater than 90% inhibition at 20  $\mu$ M (screening concentration) were detected in the WT assay (scope of other work) so inhibition lower than 90% was screened for. An initial cut-off of 50% was used. Most compounds exhibited 50-60% inhibition in at least 2 of the 3 screens. However, often data from 1 of the 3 screens had a 45-49% inhibition, just falling below the 50% cut-off. A greater variability than a standard assay was expected in these screen as the inhibition was dependent on two factors, the amount of antisense produced and the concentration of drug. While the concentration of drug would



have been nearly identical in all 3 screens (dispensed with liquid handler), the amount of antisense could be highly variable thus after looking at the data, the cut-off was dropped to 40% so potential hits or weak inhibitors would not be missed out. All hits were then compared to the reference strain to exclude molecules that were common to sensitive strains expressing antisense RNA of an essential gene and to distinguish hits that were specific to *M. smegmatis*-pMind\_*parA*\_GFP. This cut-off is lower than what would normally be used however; the level of reduction of the target protein in the cell is unknown. It is also not known what level of protein is required for the cells to grow normally without displaying any discernable physiological characteristics. Thus using a low cut-off allowed us to pick out hits that may be weak inhibitors of the target and do not display a very high sensitivity due to some target protein still in the cell. It should also be kept in mind that growth inhibition assays are not as sensitive as *in vitro* biochemical assays [112] and many good leads only show a moderate effect upon initial screening and requires chemical modification before being considered as successful leads [295].

A major disadvantage of a whole-cell assay is that it does not only identify inhibitors that are specific to the target; there are often nonstoichiometric hits that lack any SAR relationship or seemingly affect unrelated proteins [306]. There are several different reasons why there are such promiscuous inhibitions. The most obvious of these is that some compounds have alkylating or acylating properties that covalently react with functional groups in the target protein [307]. Ambiguous false positives like these can be more rigorously analysed via pharmacological studies such as dose-response curves which can either confirm the false positives or identify them as potential leads. Compounds can also act on other proteins in the same pathway thus giving similar phenotype for different cellular targets. Another reason that has been proposed for false positives is due to the formation of aggregates by self-association of small molecular weight compounds at micromolar concentration [306, 308, 309]. These aggregates have been shown to sequester and inhibit many different enzymes [308]. Consequently, like most screening hit lists, the list in this study is populated with such enzyme inhibitors. Some examples of these include kinase inhibitors (compound numbers

2, 11, 17, 18, 21 and 52) lipoxygenase inhibitors (compound numbers 22 and 62), COX inhibitor (compound number 5) and lyase inhibitor (compound number 12).

Another non-specific hit to ParA that was identified in this screen was amsacrine hydrochloride (compound number 16). This is a eukaryotic DNA topoisomerase II inhibitor that has shown antibacterial activity [310]. Topoisomerases acts by binding to DNA and amsacrine inhibits its activity by trapping the covalently bound enzyme-DNA complex. Such complexes can obstruct cellular processes like DNA replication and transcription, the predecessors of cell division. This, combined with expressing antisense RNA of ParA is bound to have some impact on the cell cycle which would explain why there was filamentous cell morphology observed with this drug (Appendix 9.4.1). Even though it has a different cellular target, it is involved in the same pathway as ParA and if the pathway is essential for the survival of the bacterial cell, inhibiting any steps in the pathway will display discernable phenotypic characteristics. An additional example of a non specific hit includes mytomycin C (compound number 69). It is a potent DNA crosslinker that has shown activity against *M. bovis* BCG [311]. It inhibits RNA and protein synthesis and may show a decrease in optical density when another process preventing the same pathway is perturbed. It is also toxic against mammalian cells.

A surprising hit was pyrazinecarboxamide (compound number 1); also known as PZA, a first-line drug used for Tb therapy (section 1.5). It is a nicotinamide analog prodrug which is converted to its active form pyrazinoic acid (POA) by an enzyme called pyrazinamidase (PZase). There are various hypotheses as to what its mode of action might be [312, 313] however it may not have a single target. Its MIC varies from 8 µg/mL to 60 µg/mL in *M. tuberculosis* depending on the assay and culture conditions and it is believed to be active against semidormant bacilli but it has no activity against other mycobacterium species [312, 313]. *M. smegmatis* is highly resistant to PZA (MIC > 2 000 µg/mL) [314] due to the presence of a very efficient POA efflux mechanism [315]. In this assay, upon expression of ParA antisense, *M. smegmatis* exhibited approximately a 40% inhibition at 20 µM, that is, 2.463 µg/mL of PZA, much lower than the

normal 2 000 µg/mL. It is unlikely that ParA is the molecular target for PZA however down-regulating a gene involved in something as important as cell division may have caused a down-regulation of other cellular processes including the efflux mechanism of POA resulting in sensitivity at a low concentration. Another nicotinamide analog, 3-pyridazinealanine (compound 51), was also found to be a hit in this assay supporting the hypothesis that the POA efflux mechanism is down-regulated upon antisense expression of ParA. Interestingly both of these hits were picked up in rich media in different libraries but were not identified in screens conducted under nutrient limitation conditions.

There were some hits that could be potential competitive inhibitors of ParA. These include adenosine receptor antagonists (compound numbers 10, 14 and 15), all of which mimicked the nitrogenous base of the ATP molecule, the substrate for ParA. Other hits that can also be considered competitive inhibitors were molecules that bind to nucleotide binding proteins (compound numbers 45, 49 and 55). Most of these have non-selective activity thus it is improbable that they are exhibiting inhibition specifically against ParA. There were also some nucleotide binding molecules (compound numbers 52 and 59) that were picked up as hits. These molecules do not target ParA but bind to its substrate. Since the level of ParA may be reduced due to the expression of ParA antisense RNA, further reducing its substrate is likely to have detrimental effects on the cells.

Lastly there are hits that have obvious toxicity against mammalian cells and are not appropriate compounds to pursue. These includes molecules that target actin (compound number 49), molecules that target mammalian enzymes (compound numbers 16, 17, 46, 47, 48, 56, 60 and 67) and/or known toxins (compound numbers 39 and 40).

### 5.4.3 Chemical aspects of the hits

As discussed above, most HTS hits contain false positives. False positives related to biology are easily detected as they are not reproducible in subsequent screens.

False positive related to chemistry are harder to discern as the positive signal is related to the chemical structure being screened and are reproducible. Thus a hit list has to be closely examined by a medicinal chemist to identify potential leads from such false positives as it requires an understanding of the pharmacophore and its interaction with the target structure [306, 307]. This is important as the process of hit-to-lead go through chemical series that takes into account the biological properties and the quantitative structure-activity relationship (QSAR) of a compound.

Of the 69 hits identified in these screens, there are certain compounds that are not very good medicinal chemistry targets or are difficult to perform SAR with. These include compound numbers 7, 10, 12, 19, 22, 38, 53, 63, 64 and 68. There are other compounds that have a complicated chemistry associated with them, another drawback from a medicinal chemistry point of view. These included compound numbers 24, 30, 31, 42, 43 and 55. Other compounds that looked plausible were compounds with similar structures, that is, compound numbers 32 & 33 and compound numbers 36 & 37. However compounds 32 & 33 have stereochemistry associated with them and compounds 36 & 37 are difficult to synthesize. Thus all hits were re-evaluated to determine their potential as new drug leads.

After the analysis of the biological activity and the chemical structure of the 69 hits that were identified in the screens, only ~ 20 compounds seemed to be likely leads. Some potential leads include compound numbers 4, 6, 9, 14, 15, 20, 28, 35, 44, 45, 53, 54 and 66. Compound 35, fenoldopam bromide, also gave elongated cell morphology (Appendix 9.4.1) when exposed at 20  $\mu$ M in nitrogen limitation media compared to the controls. Compound 44, phenoxybenzamine hydrochloride gave similar results with elongated cells (Appendix 9.4.2) in nitrogen starvation conditions. Compound 66, NSC22242 gave elongated cells in carbon starvation media (Appendix 9.4.3).

---

#### 5.4.4 Summary

In summary, this chapter has identified potential hits against ParA in a cell-based assay using HTS. This method preserved the interaction of the target protein with other macromolecules that are critical for cellular function and displayed compounds as ineffective in the presence of competitive parameters. From all the screens against the compound libraries, a total from 5680 compounds were screened in 3 different culture conditions. This collectively identified 69 hits that occurred repeatedly in the target strain but not in the WT or reference strains. After taking into consideration the biological functions and chemical structures of the 69 hits, approximately 20 compounds were identified as good hits to pursue. However, it is important to note that this was a preliminary screen and the hits identified in these screens have to be validated as compounds can display sensitivity due to an unpremeditated MOA in a cell-based assay. It would be ideal to have a secondary functional assay to distinguish that the sensitivity exhibited by the compounds is due to the inhibition of the specific target before it can be considered as a lead against ParA.

## **CHAPTER SIX**

### **Hit Validation**

## 6.1 Introduction

All hits identified as active in whole-cell assays have to be authenticated for specificity with respect to their targets to be considered as viable leads. The mode of inhibition provides confidence to inferring the SAR of the compound since this requires the physical binding of the molecule to the target. Such information can lead to rational drug design through combinatorial chemistry efforts to further embellish the inhibitory properties of the compound to generate novel antibiotics. The assays used for targeted screening in chapter 5 did not establish a direct relationship between the inhibitory molecules against ParA. This chapter validated some of the hits identified in chapter 5 as potential leads against ParA using a biochemical assay.

### 6.1.1 Validation assays

Specificity of bioactive compounds against particular targets from a cell-based assay is usually determined using cell-free *in vitro* assays (section 1.8). The follow-up of HTS with biophysical and enzymological studies are referred to as validation assays which isolates the quality hits from the false-positives [295].

The development of a successful biochemical assay has to comply with multiple factors. Firstly, such an assay requires the expression and purification of the target protein in a soluble and active form. Acquisition of active protein is followed by the development and optimization of a functional assay to determine the best working conditions for the target [316]. However, this is not always plausible if the protein has no known function or conserved domains that can potentially be used to derive its hypothetical function. In such cases, affinity-based technologies are used to determine the specificity of compounds against the target. Affinity-based assays just require the binding of the compound to the target protein however it does not impart the target functionality [268]. Once SAR is observed, successful compounds are subsequently tested for cytotoxicity in mammalian cells before a compound is regarded as a good lead [46].

---

Biochemical assays can also be utilized for kinetic studies. Such studies lead to an understanding of the basic principles of the enzyme catalysed reaction [317]. The fundamental concept of enzyme kinetics is conveyed by the Michaelis-Menton equation which is derived from the assumption that a reaction forms an intermediate enzyme-substrate complex:  $[E] + [S] \leftrightarrow [ES] \leftrightarrow [E] + [P]$  where E, S and P represents enzyme, substrate and product respectively [318]. The Michaelis-Menton equation can be used to measure the catalytic mechanism of an enzyme. The catalytic constants point towards the physiological concentrations of the reactants as enzymes generally work within or just above the Michaelis constants [317]. This can in turn be used to find out the mechanism via which a drug inhibits the enzyme activity.

## 6.2 Objectives of chapter six

This chapter focused on developing and optimizing a biochemical assay for ParA. This assay was subsequently used to validate the hits obtained in the HTS to distinguish between primary and off-target hits.

Given that the ParA protein contains the conserved Walker motif that is present in ATPases, I hypothesized that *M. tuberculosis* ParA may possess ATPase activity. I also hypothesized that ParA may require the ParB protein and *parS* DNA sites to stimulate its ATPase activity as has been found in other systems (section 4.4.1.3).

To investigate these hypotheses, the *M. tuberculosis* ParA and ParB proteins were cloned into expression vectors and soluble proteins were purified for the functional assay. The main aims of this chapter were to:



- 
- Determine an optimal protocol for isolating soluble Par proteins and develop an *in vitro* biochemical assay for ParA
  - Test for the specificity of the inhibitors identified in chapter 5 in the *in vitro* assay and determine the types of inhibitions
  - Elucidate MIC values of ParA inhibitors in *M. smegmatis*, *M. bovis* BCG and *M. tuberculosis*
  - Perform *in silico* analyses of the interactions between the validated inhibitors and ParA.

## 6.3 Results

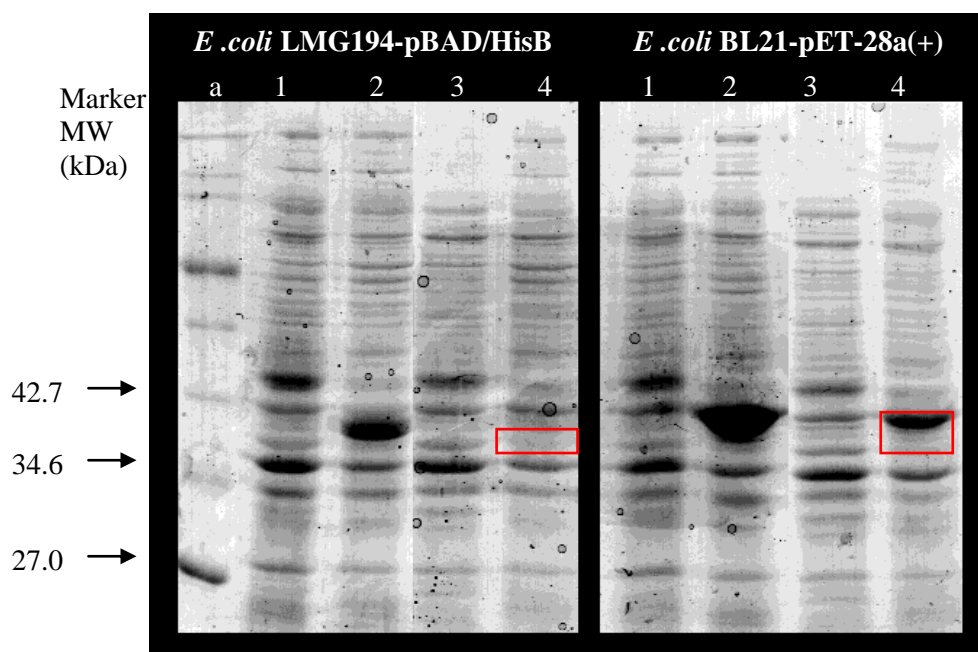
Both *M. tuberculosis* *parA* (Rv3918c) and *parB* (Rv3917c) genes were amplified from the *M. tuberculosis* H37Rv genomic DNA and cloned into expression vectors pBAD/HisB and pET-28a(+) (Table 2.2). These vectors were transformed into *E. coli* LMG194 and *E. coli* BL21 (DE3) respectively to express and purify recombinant proteins for the development of a biochemical assay.

### 6.3.1 Protein over-expression

A high yield of recombinant protein was required for the biochemical assay and subsequent hit validation thus various different expression protocols were used to determine the optimal conditions required for isolating a moderate amount of soluble protein (section 2.10.1).

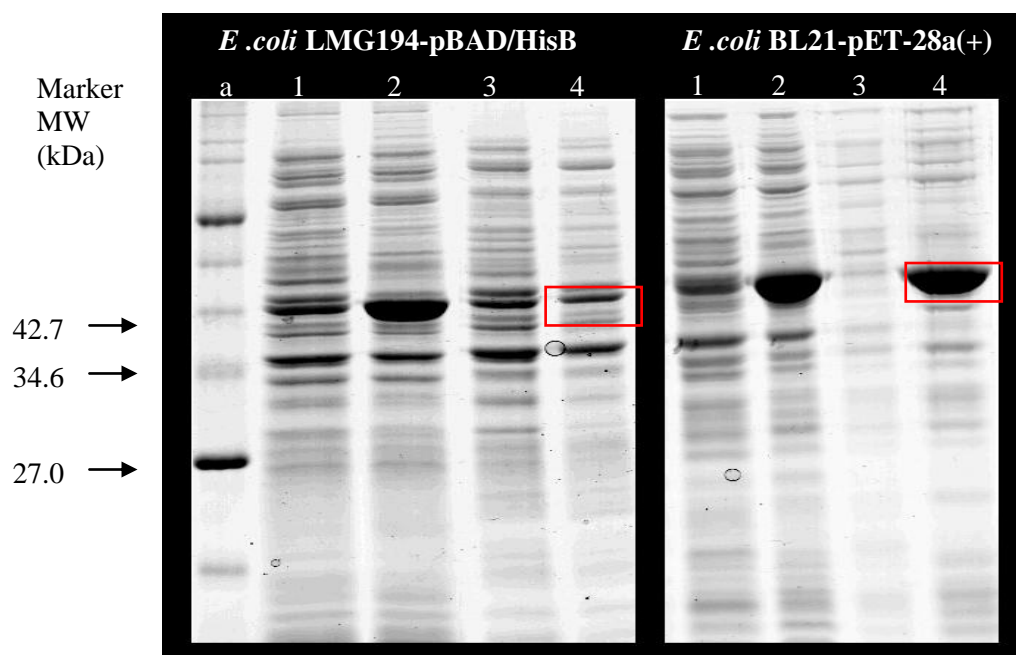
Expression with pBAD/HisB vector gave a low yield of soluble protein at 37°C (Figure 6.1 and 6.2). The *araBAD* promoter also did not express at lower temperatures or significantly increase the level of expression upon osmotic stress or heat-shock (data not shown); both of which has shown to increase the expression of recombinant proteins with the T7 promoter in the pET system [128].

Consequently, both the *parA* and *parB* genes were subcloned from the pBAD/HisB vector into the pET-28a(+) vector and subjected to all the conditions stated in section 2.10.1.1. The ZYP-5052 auto-induction media [129] significantly increased the amount of soluble recombinant protein being expressed (Figure 6.1 and 6.2) thus this protocol was used for the purification of recombinant ParA and ParB protein for all biochemical assays.



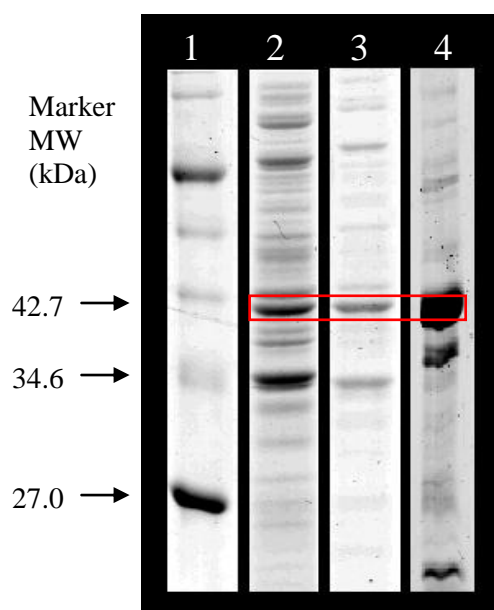
**Figure 6.1: Protein over-expression of ParA**

SDS-PAGE showing protein over expression of ParA at 37°C for 4 hours with *E. coli* LMG194-pBAD/HisB\_ParA induced with L-arabinose and *E. coli* BL21-pET-28a(+)\_ParA induced with IPTG (section 2.10.1.1, protocol I). Gel lanes contain **a**) protein molecular weight marker, lane **1**) uninduced cell lysate, lane **2**) induced cell lysate, lane **3**) soluble uninduced cell lysate, lane **4**) soluble induced cell lysate. All lanes have the same amount of protein loaded. *E. coli* BL21-pET-28a(+)\_ParA has more soluble protein present compared to *E. coli* LMG194-pBAD/HisB\_ParA (lanes 4- red box) thus *E. coli* BL21-pET-28a(+) was used for the expression and purification of recombinant protein. The molecular weight of ParA without the His-Tag is 37.5 kDa.



**Figure 6.2: Protein over-expression of ParB**

SDS-PAGE showing protein over expression of ParB at 37°C for 4 hours with *E. coli* LMG194-pBAD/HisB\_ParB induced with L-arabinose and *E. coli* BL21-pET-28a(+)\_ParB induced with IPTG (section 2.10.1.1, protocol I). Gel lanes contain **a**) protein molecular weight marker, lane **1**) uninduced cell lysate, lane **2**) induced cell lysate, lane **3**) soluble uninduced cell lysate, lane **4**) soluble induced cell lysate. All lanes have the same amount of protein loaded. *E. coli* BL21-pET-28a(+)\_ParB has more soluble protein present compared to *E. coli* LMG194-pBAD/HisB\_ParB (lanes 4- red box) thus *E. coli* BL21-pET-28a(+) was used for the expression and purification of recombinant protein. The molecular weight of ParB without the His-Tag is 37.02 kDa.

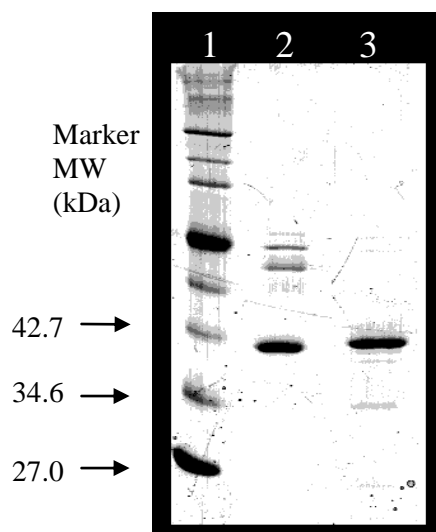


**Figure 6.3: Soluble fractions from different expression protocols**

SDS-PAGE showing: lane **1**) protein molecular weight marker, lane **2**) soluble fraction from *E. coli* LMG194-pBAD/HisB\_ParA lysate after induction at 37°C, lane **3**) soluble fraction from *E. coli* BL21-pET-28a(+)\_ParA after induction at 37°C and lane **4**) soluble fraction from *E. coli* BL21-pET-28a(+)\_ParA in auto-induction media at 37°C. Lane 4 contains half the amount of protein as in lanes 2 and 3. However, the amount of soluble ParA protein in lane 4 is still considerable higher compared to lanes 2 and 3 (red box).

### 6.3.2 Protein purification

Recombinant protein was purified as stated in section 2.10.2. All proteins purified for the biochemical assays were lysed via sonication as the Bug Buster<sup>®</sup> Protein Extraction reagent promoted the formation of inclusion bodies that rendered most protein insoluble (data not shown). French pressing resulted in the loss of some lysate when transferring it into the chamber making sonication the optimal method for lysing the cells.

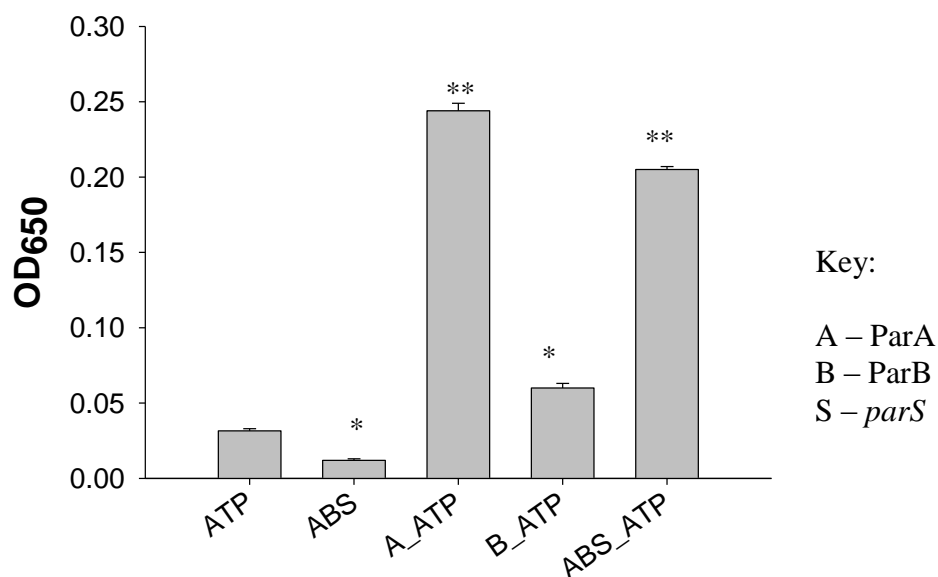


**Figure 6.4: Purified protein**

SDS-PAGE showing: lane 1) protein molecular weight marker, lane 2) purified and desalted ParA protein and lane 3) purified and desalted ParB protein.

### 6.3.3 Development of a ParA ATPase assay

An initial biochemical assay was set-up to measure ATP hydrolysis with reagents as stated in section 2.11 except  $\text{MgCl}_2$  was used at a final concentration of 10 mM and the reaction contained 3  $\mu\text{g}$  ParA protein, 3  $\mu\text{g}$  ParB protein and 1  $\mu\text{g}$  *parS* DNA (amplified from the *M. tuberculosis* genome). The reaction was started with the addition of ATP at saturating concentrations (1 mM) and incubated at 37°C for one hour. After 1 hour the amount of  $\text{P}_i$  released from the reactions was measured using the brilliant green assay (section 2.11.1) to determine if there was any ATP hydrolysed in the reactions. The results showed that there was ATP hydrolysis upon addition of ParA, no ATP hydrolysis upon addition of ParB and no further stimulation of ATPase activity upon the addition of ParB and *parS* to ParA compared to the activity exhibited by ParA alone (Figure 6.5). This assay showed that *M. tuberculosis* ParA possessed ATPase activity. It also demonstrated that the *M. tuberculosis* ParA ATPase did not require the presence of the ParB-*parS* complex to catalyse and/or stimulate this reaction thus only ParA protein was used for the functional assay.



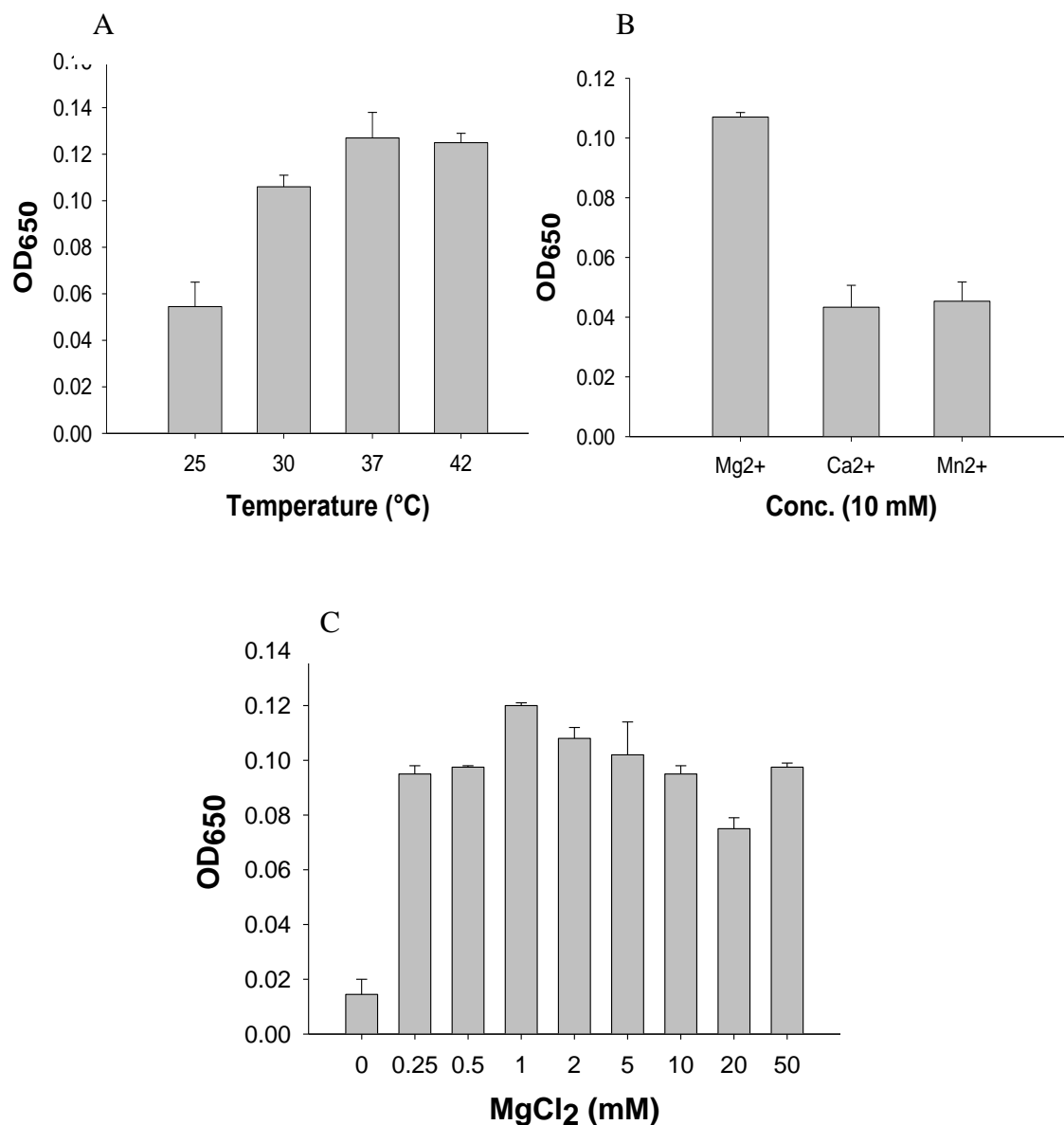
**Figure 6.5: ParA ATPase activity**

This figure shows ATP hydrolysis in the presence of ParA and ParB protein. The first bar shows spontaneous ATP hydrolysis in the reaction buffer in the absence of any protein. The second bar demonstrates that there are no contaminating phosphates in the reaction reagents or protein preparations. The last three bars show ATP hydrolysis in the presence of ParA, ParB and ParA:ParB:*parS* respectively. Results are representative of mean values of duplicates  $\pm$  SE (n=2). \* represent p values  $> 0.07$  and \*\* represent p values  $< 0.008$ .

### 6.3.4 Optimization of the biochemical assay

The influence of other parameters on the ATPase activity of ParA was investigated at saturating ATP concentration (1 mM) to find the ideal conditions for this enzyme. The reactions were set-up as indicated in Table 2.14 except for the variation of the parameter being tested. All assays contained 3  $\mu$ g ParA protein and were incubated at 37°C for one hour unless otherwise stated. The ParA ATPase activity was tested at different temperatures (Figure 6.6A). This showed a reduced activity at lower temperature (25°C and 30°C) with optimal activity at higher temperatures (37°C and 42°C).  $Mg^{2+}$  was found to stimulate the ATPase

activity of ParA the most compared to various other divalent cations (Figure 6.6B). There was no ATPase activity in the absence of  $\text{Mg}^{2+}$  with very little difference in the level of activity with  $\text{Mg}^{2+}$  concentration ranging from 0.25-50 mM (Figure 6.6C).

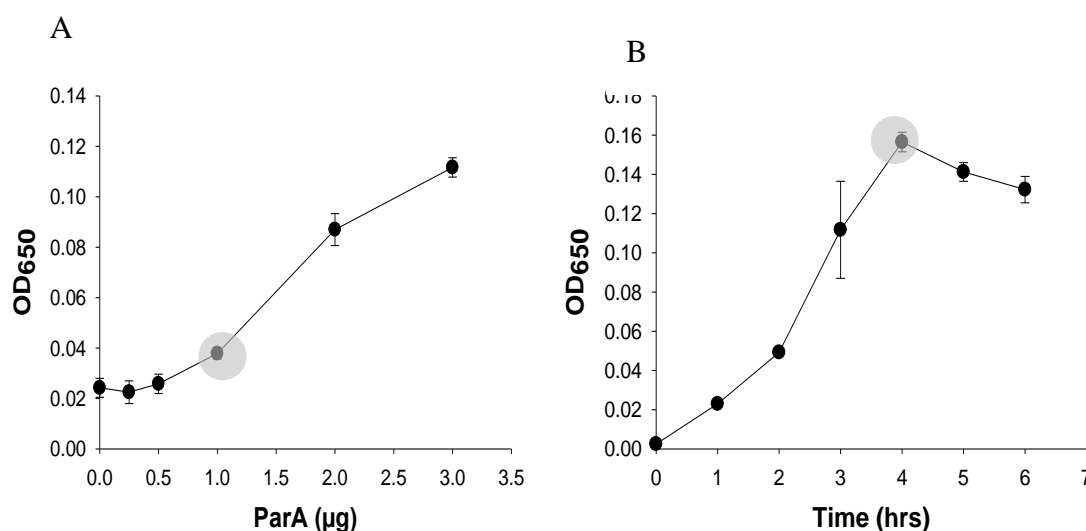


**Figure 6.6: Effects of different parameters on ParA ATPase activity**

ATPase activity of ParA **A)** in different temperatures, **B)** with different divalent cations and **C)** with varying  $\text{Mg}^{2+}$  concentrations. All experiments were carried out in triplicate and the data is normalised against spontaneous ATP decay. Results are representative of mean values of triplicates  $\pm$  SE.

### 6.3.5 Determination of the kinetic properties of the ParA ATPase

Once the constituents and assay conditions of the functional assay were determined (Figure 6.5 and 6.6), the assay was optimized for hit validation to find the parameters required to give the maximum signal with the minimum amount of protein. A titration of protein from 0.25  $\mu\text{g}$  to 3  $\mu\text{g}$  was assayed to detect the lowest amount of protein required for detectable activity in one hour (Figure 6.7A). From this assay it was established that 1  $\mu\text{g}$  of ParA was required for detectable activity at 37°C in an hour. Approximately 1  $\mu\text{g}$  of ParA was then assayed over different time points to find the time point that gave the maximum signal (Figure 6.7B). From these assays it was established that 1  $\mu\text{g}$  of protein gave maximum signal after 4 hours at 37°C.



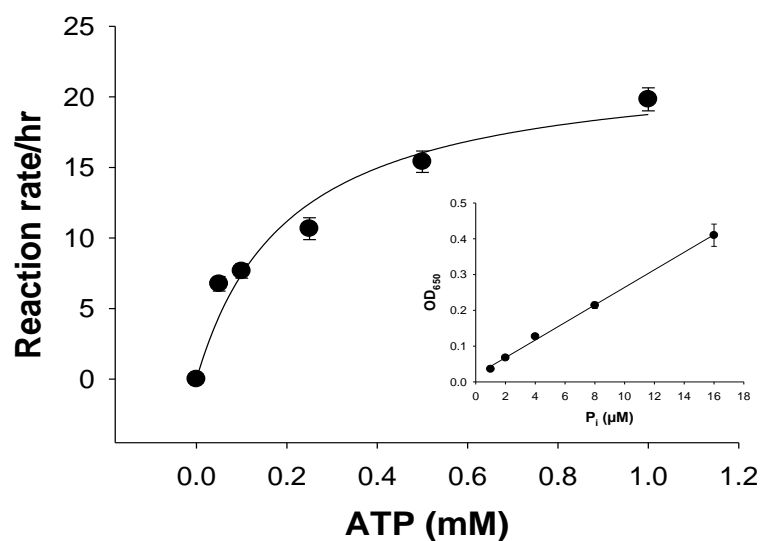
**Figure 6.7: Assay optimization for hit validation**

**A)** ATPase activity measured against varying concentrations of ParA at 37°C for 1 hour ( $n=3$ ). This shows a linear correlation between the levels of activity to the concentration of protein from 0.5-3  $\mu\text{g}$ . The minimum amount of protein required to exhibit activity after 1 hour at 37°C was 1  $\mu\text{g}$  (grey bubble). **B)** ATPase activity with 1  $\mu\text{g}$  of ParA at 37°C measured at different time points showing a linear correlation between the levels of activity against time from 2-4 hours ( $n=2$ ). Maximal signal was achieved at 4 hours (grey bubble). All experiments were carried out in duplicate and all data points are normalised against spontaneous ATP decay. Results are representative of mean values of duplicates  $\pm$  SE.



The information from Figure 6.7 was used to set up assays to determine the kinetic parameters of ParA. The rate of the ATPase reaction was determined by measuring the amount of  $P_i$  released. The amount of  $P_i$  released in the reactions was calculated by constructing a standard curve with different concentrations of  $P_i$  against  $OD_{650}$ . The standard curve was generated in the ATPase reaction buffer with different concentrations of pyrophosphate incubated with 3 U/mL inorganic pyrophosphatase at RT for 45 minutes (Invitrogen). After 45 minutes the concentration of  $P_i$  in the reaction was determined using the brilliant green assay (section 2.11.1) and the data was analysed with linear regression.

The kinetic parameters of ParA were determined by incubating 1  $\mu$ g of ParA with varying concentrations of substrate (ATP ranging from 0 - 1 mM) for 3 hours at 37°C. Non-linear regression analysis of this data gave a  $K_m$  of  $0.2 \pm 0.07$  mM and a  $K_{cat}$  of  $22.57 \pm 2.64$ /hr (means  $\pm$  SE) (Figure 6.8).

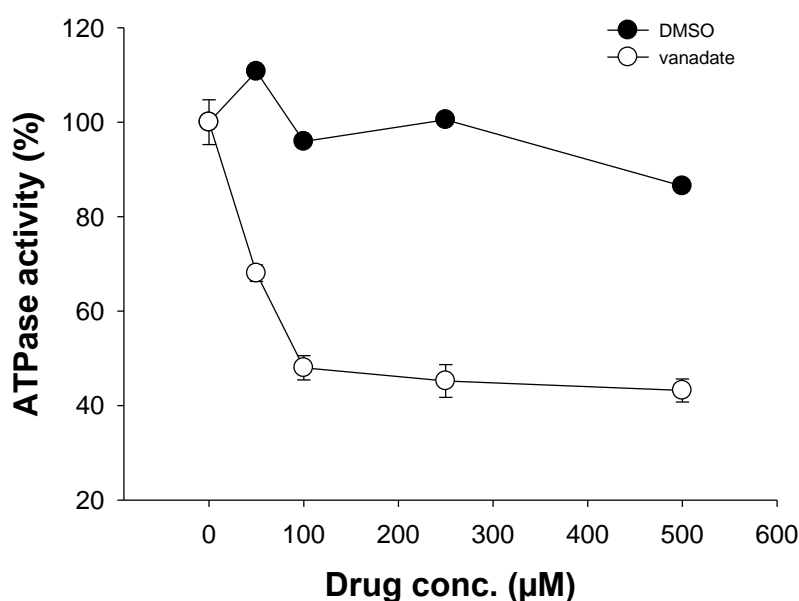


**Figure 6.8: Kinetic parameters of *M. tuberculosis* ParA ATPase**

Saturation curve showing the relationship between the concentration of substrate (ATP) and the rate of reaction ( $n=10$ ). The data have been fitted to the Michaelis-Menten equation. The insert graph shows the amount of  $P_i$  released versus  $OD_{650}$  ( $R^2 = 0.998$  and the line equation is  $y = 0.0263x$ ). All experiments were carried out in duplicate. Results are representative of mean values of duplicates  $\pm$  SE.

### 6.3.6 ATPase inhibitors

The optimized ParA ATPase assay was used to test the specificity of the compounds identified as having an effect on the *M. smegmatis*-pMind\_*parA*-GFP strain (chapter 5). All compounds tested in this study were made up in DMSO. Therefore the effect of DMSO in the biochemical assay was investigated. DMSO up to 20% had no significant effect on the ATPase activity (Figure 6.9). The effect of an ATPase inhibitor, sodium metavanadate ( $\text{NaVO}_3$ - referred to as vanadate from this point forward) was also tested on the ParA ATPase activity. An assay with 1  $\mu\text{g}$  ParA was set-up and incubated for 3 hours at 37°C with saturating ATP concentration (1 mM) and varying vanadate concentrations. This assay showed that vanadate inhibited the ATPase activity of ParA (Figure 6.9).

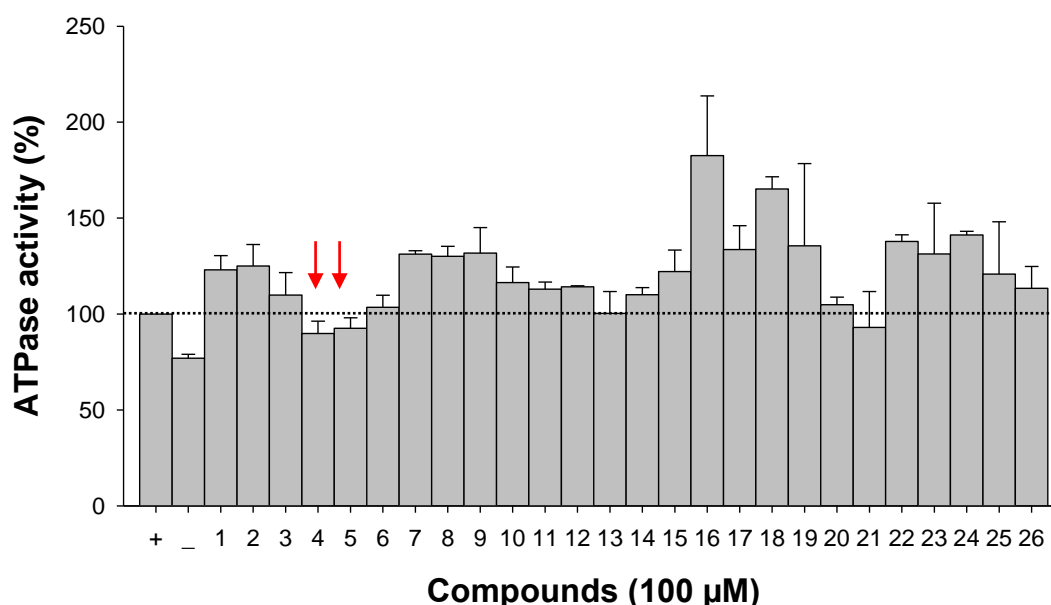


**Figure 6.9: Effect of DMSO and vanadate on ATPase activity of ParA**

This graph (●) shows no significant inhibition of the ATPase activity with DMSO,  $p$  value  $> 0.7$ . The second graph (○) shows inhibition of ATPase activity upon the addition of 50  $\mu\text{M}$  and 100  $\mu\text{M}$  sodium metavanadate to ~ 70% and 40% respectively. All concentrations after 100  $\mu\text{M}$  maintained ~ 40% activity ( $n=3$ ). All data points are normalised against spontaneous ATP decay and all experiments were carried out in duplicate. Results are representative of mean values of duplicates  $\pm$  SE.

### 6.3.7 Hit validation

Initial hit validation assays were carried out with the compounds from the NIH Diversity Set as these were readily available in the laboratory. All 26 hits obtained from the cell-based assays in the three culture conditions (chapter 5- section 5.3.3) were tested at 100  $\mu\text{M}$  to identify if these compounds specifically affected the ATPase activity of ParA (Figure 6.10). The assay was carried out with 1  $\mu\text{g}$  ParA protein in the presence of 50  $\mu\text{M}$  ATP and incubated at 37°C for 3 hours. ParA without any compound was used as a positive control (100% activity) and ParA with 100  $\mu\text{M}$  sodium metavanadate was used as a negative control.



**Figure 6.10: Effect of NIH hits on the ParA ATPase activity**

This graph exhibits the ATPase activity of ParA upon addition of 100  $\mu\text{M}$  of the 26 compounds obtained as hits from the NIH Diversity Set in the whole-cell assay. '+' indicates positive control (uninhibited ParA) and '-' indicates negative control (ParA with 100  $\mu\text{M}$   $\text{NaVO}_3$ ). Compound numbers 4 and 5 (indicated by red arrows) exhibited ~10% inhibition however; these were statistically not significant,  $p$  value  $> 0.1$ . All data is normalised against spontaneous ATP decay. Results are representative of mean values of duplicates  $\pm$  SE.

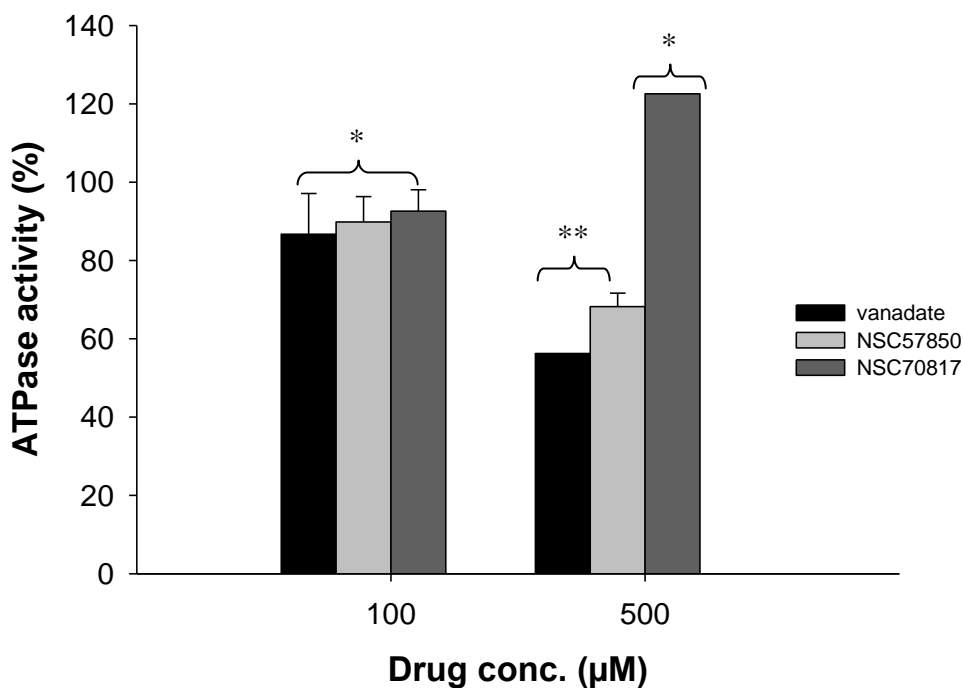
Compound 4 in Figure 6.10 represent NSC57850, found to exhibit inhibition in the cell-based assay under nutrient rich conditions. It was also discussed as an optimal hit to pursue in section 5.4.3 after considering its biological roles and chemical structure. Compound 5 in Figure 6.10 represents NSC70817, also found to exhibit inhibition in the cell-based assay under nutrient rich conditions. This compound is cytotoxic as it targets human cytochrome P450 (PubChem).

Even though the inhibition exhibited by these two compounds were insignificant, it was hypothesized that these drugs may show activity against ParA at higher concentrations as they showed some evidence of inhibition at 100  $\mu\text{M}$ . Thus the assay was repeated with 500  $\mu\text{M}$  of drugs. This illustrated that NSC57850 specifically inhibited the ATPase activity of ParA at 500  $\mu\text{M}$  while NSC70817 had no effect (Figure 6.11). However, due to a limited supply of compound, NSC57850 could not be tested further in this work.

Since NSC57850 could not be acquired, other hits that could be bought in were validated from the LOPAC and Spectrum Collection. However, these compounds had to be prioritized as they were not readily available in the laboratory. Two compounds, phenoxybenzamine hydrochloride (phenoxybenzamine) and fenoldopam bromide (fenoldopam), that were considered as good hits to pursue in chapter 5 (section 5.4.3) were procured to be tested in the biochemical assay. The results showed that increasing the concentration of phenoxybenzamine progressively inhibited the activity of the ParA ATPase (Figure 6.12A). Increasing concentrations of fenoldopam from 0  $\mu\text{M}$  to 100  $\mu\text{M}$  inhibited the ATPase activity of ParA to  $\sim 60\%$  but there was no further inhibition at 250  $\mu\text{M}$  (Figure 6.12A). Greater than 250  $\mu\text{M}$  of phenoxybenzamine and fenoldopam could not be tested as it precipitated in the reaction and interfered with the colour development of the brilliant green assay (data not shown).

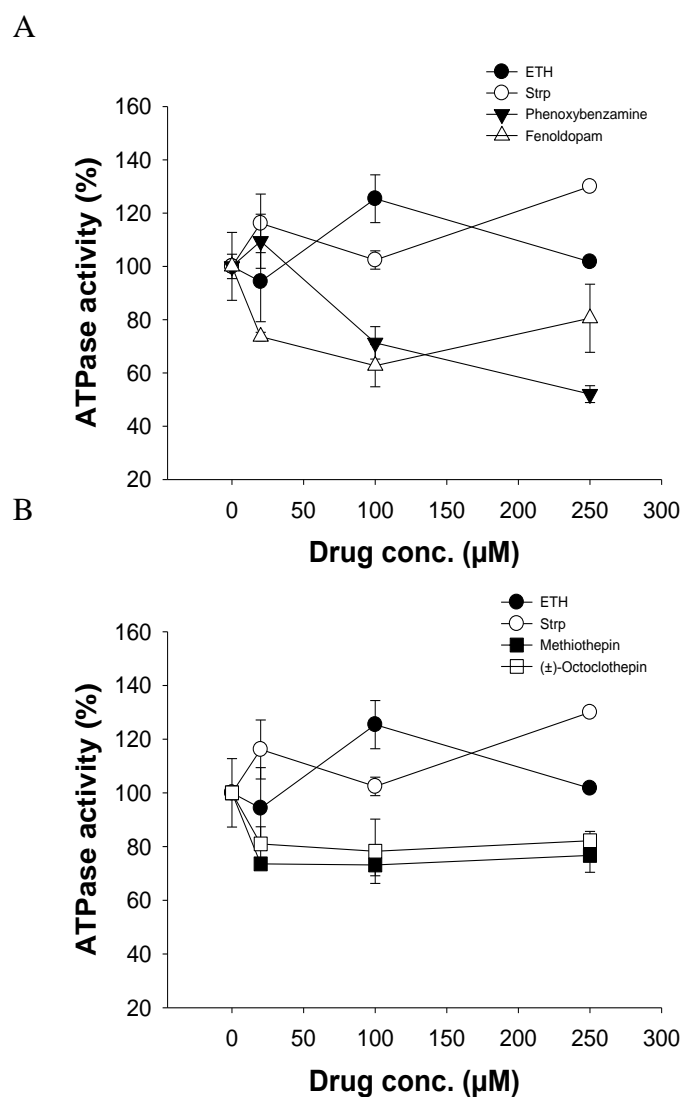
Another two compounds, methiothepin mesylate (methiothepin) and ( $\pm$ )-octoclotheptin maleate (octoclotheptin), that had similar structures were also purchased to be tested. Addition of about 50  $\mu\text{M}$  of methiothepin and

octoclothepein exhibited ~ 20% inhibition of the ParA ATPase activity but this did not show greater inhibition with increasing concentrations of drugs up to 250  $\mu\text{M}$  and the activity was maintained at 80% (Figure 6.12B).



**Figure 6.11: Effect on ATPase activity of ParA with varying concentrations NSC57850 and NSC70817**

This graph shows the effect of vanadate (■), NSC57850 (□) and NSC70817 (■) on the ATPase activity of ParA at 100  $\mu\text{M}$  and 500  $\mu\text{M}$ . Vanadate and NSC57850 showed that it inhibited the ParA activity at 500  $\mu\text{M}$ . Due to a shortage in compounds, this experiment was carried out only with two different concentrations in duplicate. All data is normalised against spontaneous ATP decay. Results are representative of mean values of duplicates  $\pm$  SE. \* represent p values  $> 0.1$  and \*\* represent p values  $< 0.03$ .



**Figure 6.12: Effect on ATPase activity of ParA with phenoxybenzamine, fenoldopam, methiothepin and octoclotheptin**

Effect on the ATPase activity of ParA with **A**) phenoxybenzamine (▼), fenoldopam (Δ), **B**) methiothepin (■) and octoclotheptin (□). Both these graphs show that standard Tb drugs, ETH and streptomycin; have no significant effect on the ATPase activity of ParA. All data points are normalised against spontaneous ATP decay and all experiments were carried out in duplicate (n=2). Results are representative of mean values of duplicates ± SE.

### 6.3.8 Kinetic parameters of the inhibitors

To determine the type of inhibition exerted by the inhibitors tested in Figure 6.12, the kinetic parameters of the ParA ATPase assay was investigated in the presence and absence of the inhibitors at various substrate concentrations. The assay was carried out with 1  $\mu$ g ParA with ATP concentration ranging from 0.05 - 1 mM. These reactions were incubated for 3 hours at 37°C. Non-linear regression analysis of this data was used to determine the Michaelis constant ( $K_m$ ) and catalytic constant ( $K_{cat}$ ) of the reactions in the presence of the inhibitors (Table 6.1, means  $\pm$  SE). These results showed that sodium metavanadate, fenoldopam bromide, methiothepin mesylate and ( $\pm$ )-octoclotheptin maleate exhibit competitive inhibition while phenoxybenzamine hydrochloride exhibit a mixed inhibition.

**Table 6.1: Kinetics parameters of ParA ATPase in the presence of inhibitors**

Compound	Conc. ( $\mu$ M)	$K_{cat}$ /hr	$K_m$ (mM)
ParA (uninhibited)	-	$22.57 \pm 2.64$	$0.2 \pm 0.07$
Sodium metavanadate	500	$29.76 \pm 4.16$	$0.86 \pm 0.21$
Phenoxybenzamine hydrochloride	250	$6.38 \pm 0.59$	$0.44 \pm 0.09$
Fenoldopam bromide	250	$29.19 \pm 3.03$	$0.48 \pm 0.1$
Methiothepin mesylate	250	$25.25 \pm 2.37$	$0.41 \pm 0.09$
( $\pm$ )-Octoclotheptin maleate	250	$21.01 \pm 0.71$	$0.45 \pm 0.02$

### 6.3.9 Anti-mycobacterial properties of ParA inhibitors

Three of the compounds (phenoxybenzamine hydrochloride, fenoldopam bromide and ( $\pm$ )-octoclotheptin maleate) were tested for their MIC values in *M. smegmatis*, *M. bovis* BCG and *M. tuberculosis* H37Ra to test their anti-mycobacterial activity.

Dose-response assays (section 2.14) against *M. smegmatis* was conducted in three different culture conditions; nutrient rich (LBT), carbon starvation (C<sup>-</sup>)

---

and nitrogen starvation (N<sup>-</sup>). MIC values of these assays were determined on day 4 for rich culture conditions and day 3 for nutrient starvation conditions (Table 6.2). MIC values of *M. bovis* BCG and *M. tuberculosis* H37Ra were only determined in rich media on day 14.

Of the three compounds tested, octoclotheptin displayed the best MIC values in all three microbes with the lowest MIC in N<sup>-</sup> conditions. Phenoxybenzamine displayed varying MIC values in *M. smegmatis* under different culture conditions with the highest MIC in rich culture conditions and lowest under N<sup>-</sup> conditions. It also gave significantly high values in the slow growers. Fenoldopam had a high MIC under all culture conditions in *M. smegmatis* and in the slow-growers.

These compounds were also tested for their stationary bactericidal activity (SBC- section 2.15) in *M. smegmatis* under nutrient rich, C<sup>-</sup> and N<sup>-</sup> conditions. Stationary bactericidal concentrations were determined after 4 days (Table 6.3). The SBC for fenoldopam could not be determined under any culture conditions as it was beyond the concentrations tested in these assays. Phenoxybenzamine exhibited a high but consistent SBC values under all culture conditions while octoclotheptin displayed a lower SBC in nutrient rich compared to the nutrient starved conditions.

Since fenoldopam bromide displayed very high MIC values in the slow-growers and very high SBC values in *M. smegmatis*, it was excluded from further analysis in this study.



**Table 6.2: MIC<sub>90</sub> values of ParA inhibitors in mycobacterial cells**

	Fenoldopam ( $\mu\text{M}$ )	Phenoxybenzamine ( $\mu\text{M}$ )	Octoclotheptin ( $\mu\text{M}$ )
<i>M. smegmatis</i> : Rich media	500	$\geq 1\ 000$	61.16
Carbon starvation	$\geq 1\ 000$	500.7	78.92
Nitrogen starvation	500	132.2	38.2
<i>M. bovis</i> BCG: Rich media	$> 1\ 000$	$\geq 1\ 000$	37.88
<i>M. tuberculosis</i> H37Ra: Rich media	$> 1\ 000$	$\geq 1\ 000$	56.88
<i>M. tuberculosis</i> H37Rv: Rich media	unknown	$> 1\ 000$	34.7
<i>M. tuberculosis</i> CDC1551: Rich media	unknown	$> 1\ 000$	17.4

**Table 6.3: SBC values in *M. smegmatis* under different culture conditions**

Condition	Fenoldopam ( $\mu\text{M}$ )	Phenoxybenzamine ( $\mu\text{M}$ )	( $\pm$ )- Octoclotheptin ( $\mu\text{M}$ )
LBT	$> 500$	$\geq 500$	62.5
C <sup>-</sup>	$> 500$	$\geq 500$	250
N <sup>-</sup>	$> 500$	$\geq 500$	250

### 6.3.10 Cytotoxicity of ParA inhibitors

IC<sub>50</sub> values of phenoxybenzamine and octoclotheptin were also determined in HL-60 cells (Table 6.4). These were compared to the 50% of the lethal dose (LD<sub>50</sub>) in mouse (published on ChemIDplus Advanced: <http://chem.sis.nlm.nih.gov/chemidplus/chemidheavy.jsp>).

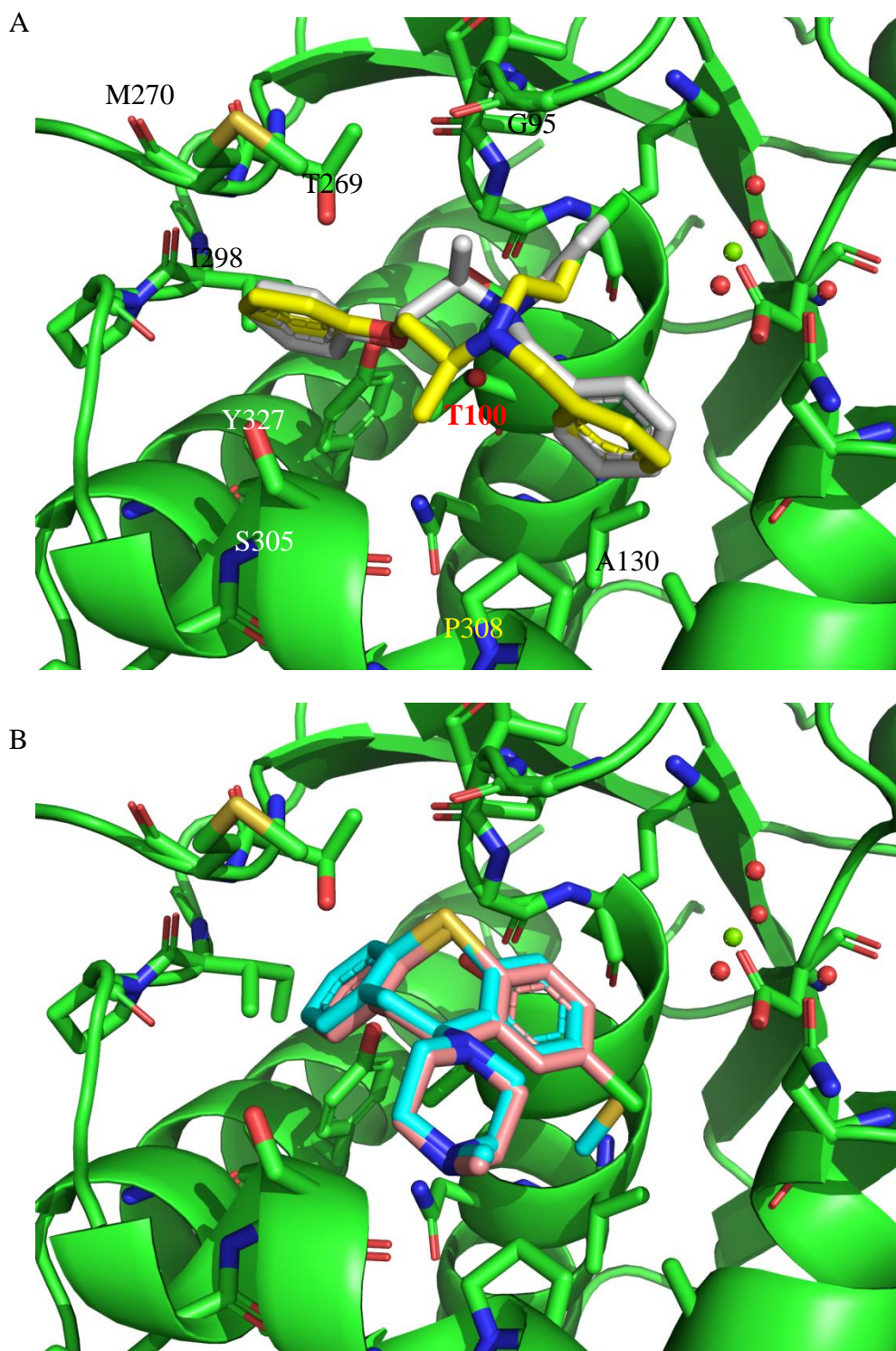
**Table 6.4: Cytotoxicity in HL-60 cells and mouse**

	Streptomycin	Phenoxybenzamine	( $\pm$ )-Octoclotheptin
HL-60 cells (IC <sub>50</sub> )	$> 250\ \mu\text{M}$	$108.76 \pm 3.88\ \mu\text{M}$	$9.97 \pm 7.72\ \mu\text{M}$
Mouse (LD <sub>50</sub> - oral administration)	500 mg/kg [319]	1535 mg/kg [320]	78 mg/kg [321]

### 6.3.11 *In silico* interaction between ParA and its inhibitors

A homology model of *M. tuberculosis* ParA was generated based on the template structure of *T. thermophilus* SOJ [322]. This was used to dock in (R)- and (S)-phenoxybenzamine, methiothepin and octoclotheptin. Both isomers of phenoxybenzamine form a hydrogen bond between the oxygen atom and the OH of threonine 100 (Figure 6.13A). The benzylamine ( $C_6H_5CH_2N-R_2$ ) binds to a relatively hydrophobic pocket lined by alanine130, proline308 and threonine99 (data not shown). The remainder of the binding site is lined by threonine269, methionine270, isoleucine198, tyrosine327 and serine305. There is different conformation observed between crystal structures of SOJ with and without a ligand. The helix which includes threonine 100 is formed at an earlier position than the liganded form. This information tells that the phenoxybenzamine can fit into the ligand binding site of the models however; it does not provide any information its ability to inhibit this enzyme.

Methiothepin and octoclotheptin docked to the ligand binding site in an almost identical fashion with the chlorine and the methyl sulfane group binding to the hydrophobic pocket centred at alanine130 (Figure 6.13B). In all three molecules, an aromatic ring occupies a similar position to the adenine ring of ADP.



**Figure 6.13: Molecular docking of ParA and its inhibitors**

**A)** Both (R) and (S) isomers of phenoxybenzamine docked into the binding site of the ParA homology model (colours: nitrogen - blue, oxygen - red, sulfur - green, chlorine - green, carbons - variable; ParA - Green ribbon, (R) and (S) isomers white and yellow carbons respectively). **B)** Methiothepin and octoclotheptin (cyan and brown carbons respectively) docked into the ligand binding site. All figures were generated using Pymol.

## 6.4 Discussion

The goal of this chapter was to validate the inhibitors identified in the cell-based assay (chapter 5) against ParA. This was achieved by designing a secondary functional assay for the ParA protein. The *M. tuberculosis* proteins were used for the development of the biochemical assay even though the initial cell-based assays were performed with *M. smegmatis*. *M. smegmatis* has been proven to be a suitable surrogate as a primary screen to shortlist compounds for testing against *M. tuberculosis* [33]. However, using *M. tuberculosis* proteins for the functional assay would not only validate the hits from the cell-based assay as being specific to the ParA protein, it would deem them relevant as potential leads for anti-tubercular treatment as it is specifically against the protein of the disease causing microbe.

The first obstacle in developing the functional assay was obtaining a moderate amount of soluble protein. Bacterial proteins are commonly over-expressed in *E. coli* expression strains. Typically only 30–40% of *M. tuberculosis* proteins are obtained in soluble form when expressed in *E. coli*; that is after the exclusion of membrane proteins [323, 324]. *E. coli* LMG194 with pBAD/HisB vector and *E. coli* BL21 (DE3) with pET-28a(+) vector was used for the over-expression of recombinant protein in various conditions as stated in section 2.10.1. From these experiments it was concluded that the *E. coli* LMG194 strain with pBAD/HisB vector did not give high levels of recombinant *M. tuberculosis* Par proteins (0.05 mg/mL from a 500 mL culture) compared to the *E. coli* BL21 strain with the pET-28a(+) vector (0.5 mg/mL from a 500 mL culture) (Figure 6.1 and 6.2). The best condition to obtain the soluble fractions of the Par proteins was in the *E. coli* BL21 (DE3) strain in auto-induction media, ZYP-5052 (Appendix 9.5.2) (Figure 6.3). Thus all proteins used for the biochemical assay in this work were expressed in *E. coli* BL21 with the pET-28a(+) vector in the ZYP-5052 media.

Once the optimal protocol was selected for the production of soluble recombinant protein, it was batch purified from 500 mL cultures. The ParB

protein was purified as stated in section 2.10.2. The ParA protein immediately precipitated upon eluting it off the column when purified from the *E. coli* BL21 (DE3) strain. Some proteins are extremely sensitive to solution conditions and under non-optimal conditions, can alter their structure. Such a solution consists of molecules that are folded, misfolded or unfolded. In all instances, proteins can aggregate which is favored by crowding, for example, over-expression and this shows up as a precipitate. This was not seen when ParA was purified from the *E. coli* LMG194 strains as it had low concentrations of proteins. Some proteins are prone to aggregation at high concentrations [325]. The ParA plasmid partitioning protein has also been reported to be insoluble at high concentrations [247]. Approximately 10% glycerol was used as an osmolyte and was added to all buffers (from lysis buffer to elution buffer) in the purification of ParA which prevented its precipitation as glycerol has been shown to stabilize the native structure of proteins [326]. The desalted ParA protein was kept in 50% glycerol at -20° and used in the biochemical assay. The addition of glycerol did not affect the ATPase activity when added up to 15% (data not shown).

#### 6.4.1 ParA ATPase

As discussed in section 4.4.1.3, the ParA plasmid partitioning protein possesses ATPase activity which is stimulated by the binding of the ParB-*parS* nucleoprotein complex [132, 239, 282]. In *M. smegmatis*, it has been shown that the ParB protein binds to *oriC*-proximal *parS* sequences and this is believed to be assisted by the ParA protein [165]. In this study, it was found that expressing the antisense RNA of *parA* and *parB* genes together could not be tolerated in *M. smegmatis* (section 4.3.2).

Thus it was hypothesized that like the plasmid partitioning ParA, *M. tuberculosis* ParA may also possess ATPase activity. It was also hypothesized that this activity could potentially be stimulated with the ParB-*parS* nucleoprotein complex. To test these hypothesis, various concentrations of ParA and ParB protein with *parS* DNA was added to the *in vitro* ATPase assay (section 2.11)

similar to the work carried out by Jensen and Gerdes [132] and the activity was measured. Most enzyme reactions are measured by the formation of the product. As discussed in section 5.1.3, ATPases hydrolyse ATP to ADP and  $P_i$ . Thus the presence of ATPase activity in this work was determined by measuring the amount of  $P_i$  released using the brilliant green assay as stated in section 2.11.1.

From this work, it was determined that the *M. tuberculosis* ParA protein possesses ATPase activity. It was also established that the *M. tuberculosis* ParA activity is not further activated either by the addition of the ParB protein or the *parS* DNA (Figure 6.5). This is similar to what was seen with the ParA ATPase activity in *C. crescentus* where the ATPase activity was not dependent on cooperative binding of ParB [243]. Thus for the functional assay, only the ParA protein was used.

The effects of some other factors on the ParA ATPase activity were also tested. The assay was incubated at varying temperatures (25°C-42°C). This demonstrated that lower temperatures did not give optimal activity. Not surprisingly, maximal activity was observed at 37°C and this level of activity was maintained up to 42°C (Figure 6.6A). The importance of several divalent cations was also tested on this assay. This revealed that  $Mg^{2+}$  was the best metal cofactor for ParA activity while  $Ca^{2+}$  and  $Mn^{2+}$  stimulated the activity to a lower extent (Figure 6.6B). The optimal concentration of  $Mg^{2+}$  required for the assay was determined by measuring the ATPase activity in the presence of varying concentrations of  $MgCl_2$ . In the absence of  $Mg^{2+}$ , there was no ATPase activity with maximal activity at 1 mM. However, increasing the concentration of  $Mg^{2+}$  up to 50 mM did not significantly alter the level of ATPase activity.

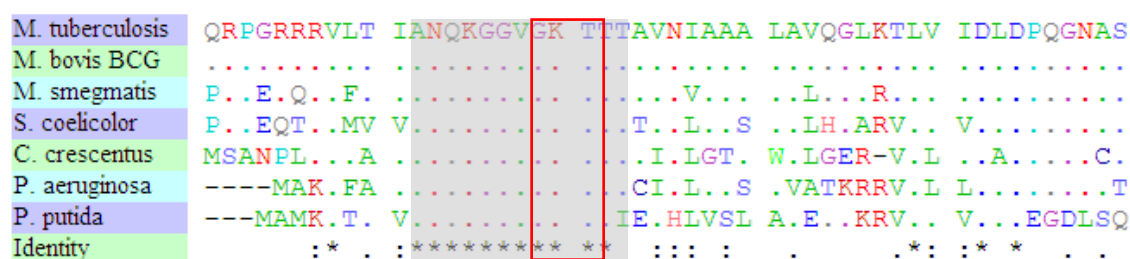
Once the biochemical assay for ParA was developed, it was optimised for inhibition assays, that is, optimal incubation times and enzyme concentrations within the linear range of the assay. This was established so a low concentration of potential inhibitor will be able to exert its effect as there will not be an abundance of protein in the assay. Assays were carried out with varying concentrations of protein and measurements at different time points were taken

(Figure 6.7). From these assays, it was determined that the optimal signal was achieved with 1  $\mu$ g of ParA incubated at 37°C for 4 hours. It showed a linear increase in the ATPase activity with increasing enzyme concentration and time for at least 4 hours indicating that the enzyme was stable under these assay conditions. To determine the steady state conditions of this assay, the kinetic parameters of the enzyme were measured by incubating 1  $\mu$ g of ParA protein for 3 hours at 37°C with varying substrate concentrations. The graph of the rate of ATP hydrolysis against ATP concentration showed that *M. tuberculosis* ParA ATPase obeyed Michaelis-Menton kinetics (Figure 6.8). This showed that the reaction rate increased asymptotically as the substrate concentration increased. However, at high substrate concentration, the rate of the reaction was independent of the substrate concentration as all active sites became saturated and the reaction turnover rate became constant. Once the key kinetic parameters of the ParA ATPase were characterized, this assay was used to test for inhibitors against ParA. These inhibitors were compounds that were hits from the cell-based assays against the ParA antisense RNA expressing strain that were being validated for specificity and potency against ParA.

#### 6.4.2 Inhibition of the ParA ATPase activity

Before the ParA functional assay was used for target validation, it was tested with a known ATPase inhibitor. There are no known inhibitors against ParA as it is a novel target however; sequence analysis shows that it possesses the conserved Walker A ATP binding domain (P-loop) with the signature GKT/S residues (Figure 6.1). This Walker motif possesses the critical phosphate-binding lysine (K) that binds to the  $\beta$  and  $\gamma$  phosphates of ATP [327]. Some such ATPases has been shown to be inhibited by vanadate [328, 329] where vanadate esters resemble the substrates and bind to the enzyme as substrate analogues [330]. This forms an enzyme-inhibitor (EI) complex and prevents substrate binding. This work tried to determine if the ParA ATPase activity was similarly affected by vanadate.

Since vanadate binds to the active site, its presence is likely to decrease the affinity of substrate binding which is typical of a competitive inhibitor. This would decrease the rate of the reaction thus increasing the  $K_m$ . However, increase in substrate concentration in such a reaction would inevitably out-compete the inhibitor binding to the enzyme thus the  $k_{cat}$  should remain unchanged. These kinetic parameters were investigated in the ParA ATPase assay with vanadate. Vanadate showed competitive inhibition with respect to ParA against ATP. There was an increase in  $K_m$  of uninhibited ParA to ParA inhibited with 500  $\mu$ M sodium metavanadate while the  $K_{cat}$  remained approximately the same.



Sequence alignment of the ATP binding motif of the Walker A region present in the ParA proteins between different bacterial species. Identical residues present in all sequences are marked with an asterisk (\*) with universally conserved GKT/S signature region marked in the red rectangle. The GenBank accession numbers for the proteins analysed are: *M. tuberculosis*, CAA16230.2; *M. bovis* BCG, YP\_980053.1; *M. smegmatis*, ABK74964.1; *S. coelicolor*, AAF16005.1; *C. crescentus*, AAB51267.1; *P. aeruginosa*, YP\_794041 and *P. putida*, YP\_534787.1.

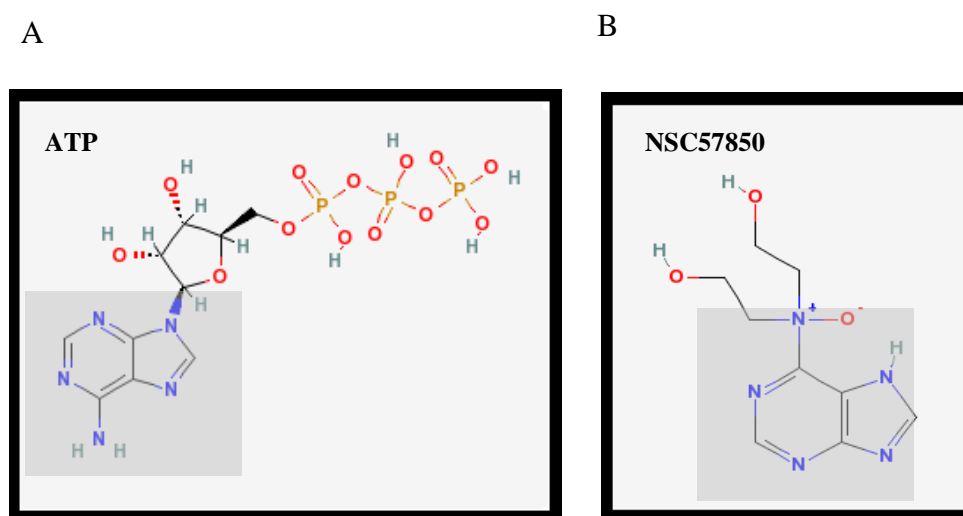


### 6.4.3 Hit validation

Given that a biochemical assay was successfully developed for the ParA protein and its activity effectively inhibited by a P-loop inhibitor, vanadate (Figure 6.9), the assay was subsequently used for the validation of the hits isolated from the cell-based assays. All hits from the NIH Diversity set were tested in the ATPase assay at 100  $\mu$ M as all these compounds were readily available. Some of these have been disregarded as good compounds to pursue due to them being cytotoxic or not having ideal chemical structures for medicinal chemistry. However, these were still tested in order to discover if any of the chemical structures proved to be a potent inhibitor of the ParA protein as this is a novel target with no drugs known to specifically inhibit its activity. It was hypothesized that while these suboptimal drugs may not serve as effective therapeutics, they could potentially be used to identify the chemical structures that the ParA protein can bind to. These structures can then be used to build a model with data from high- to low-activity compound and be used to screen analogues *in silico*.

Sodium vanadate was used as a positive control in the inhibition assays as it is a general Walker A ATPase inhibitor shown to exhibit activity against ParA (Figure 6.9). It was hypothesized that other compounds that may be weak inhibitors of ParA would not display as high an inhibition as vanadate. From the 26 compounds tested, none of the compounds potently inhibited the ParA ATPase activity at 100  $\mu$ M. However, two compounds (NSC57850 and NSC70817) showed approximately a 10% inhibition of the ParA ATPase activity at 100  $\mu$ M (Figure 6.10). Both compounds were retested in the ATPase assay at 500  $\mu$ M. It was hypothesized that if either of these compounds were inhibiting ParA, this would result in a further inhibition of the ATPase activity in a dose-dependent manner. This assay showed that 500  $\mu$ M of NSC57850 decreased the ParA ATPase activity to approximately 70% while NSC70817 had no effect. NSC57850 is known to bind to nucleotide binding proteins and perturbs cholesterol metabolism, pathways involving ATPases and proteasome oxidative phosphorylation involving ATP synthase (BioInfoBank Institute). NSC57850 has a similar structure to the adenine moiety of ATP (Figure 6.15) thus it was

hypothesized that it may inhibit the ATPase activity of ParA by competing for the active site. However, due to a lack of this compound, this could not be tested. From these results, it can be concluded that except for NSC57850, the phenotypic changes observed from the rest of the 25 compounds from the NIH Diversity sets that were hits in the cell-based assay, that is, decreased in OD and/or elongated cell morphology under antisense inducing conditions, were due to another MOA.



**Figure 6.15: Structure of ATP and NSC57850**

A) Molecular structure of ATP, the substrate for ParA. The adenine base is shaded in grey. B) Molecular structure of the compound NSC57850. This compound has a similar structure to the adenine base of the ATP molecule (shaded in grey) and showed activity against ParA (Figure 6.11).

The hits from the LOPAC and Spectrum Collection had to be prioritized and purchased as these compounds were not readily available in the laboratory. The hits that were discussed as being suboptimal in chapter 5 were excluded from further analysis. Four compounds were purchased from the LOPAC and Spectrum Collection: fenoldopam bromide, phenoxybenzamine hydrochloride, methiothepin mesylate and ( $\pm$ )-octoclotheptin maleate. These compounds showed a decrease in OD and elongated cell morphology under nitrogen starvation conditions specifically in the ParA antisense RNA expressing strain.

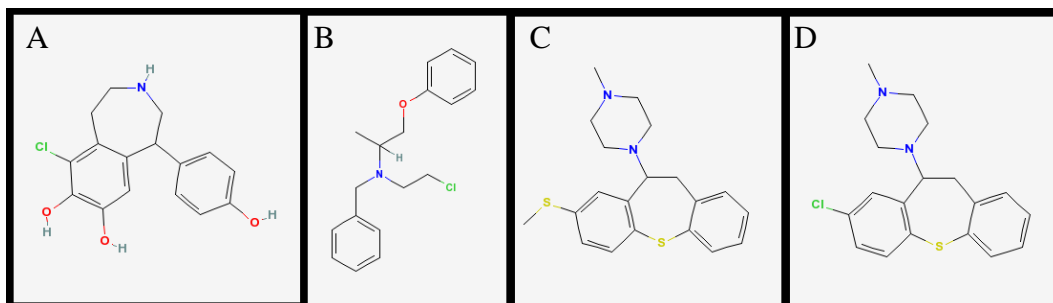
All four compounds were tested at varying concentrations along with 50  $\mu$ M ATP and 1  $\mu$ g of ParA. These assays were incubated at 37°C for 3 hours and the formation of  $P_i$  was measured to determine if there was any inhibition in the activity upon addition of the compounds. These assays were performed together with two standard Tb drugs, ethambutol and streptomycin, as controls (Figure 6.12). The results showed that the ATPase activity of ParA was not affected by other drugs, that is, ETH and streptomycin.

Fenoldopam bromide showed an initial decrease in the ParA ATPase activity to 70% at 50  $\mu$ M and 60% at 100  $\mu$ M. However, further increasing the fenoldopam bromide concentration did not increase this inhibition (Figure 6.12A). Fenoldopam is a selective dopamine D1 receptor agonist that is used as a vasodilator for treating emergency hypertension [331]. It has also been found to inhibit the *B. subtilis* Sfp phosphopantetheinyl transferase, the proliferation of several strains of *Plasmodium falciparum* and the HSP90 protein in humans (PubChem Bioassays). Inhibition of the HSP90 protein by fenoldopam is interesting as like ParA, the HSP90 protein also possesses a nucleotide binding pocket [332] and may act with a similar mechanism of action. However, the mode of HSP90 inhibition by fenoldopam has not yet been elucidated. It was hypothesized that if fenoldopam was binding to the active site, it would show a competitive type inhibition. Kinetic analysis with this compound showed a competitive mode of inhibition with respect to ParA (Table 6.1). It showed that the substrate binding affinity of ParA decreased in the presence of the drug (increase in  $K_m$ ) which was overcome by high concentrations of substrate ( $K_{cat}$  approximately the same). Thus the inhibitory effect of fenoldopam bromide was investigated on mycobacterial cells. This showed that fenoldopam bromide exhibited very high  $IC_{50}$  and  $MIC_{90}$  values in *M. smegmatis*, *M. bovis* BCG and *M. tuberculosis* H37Rva (Table 6.2). It also showed no stationary bactericidal activity thus it was disregarded as a good inhibitor to pursue and was not analysed any further. However, its structure can be used to find similar compounds that may inhibit ParA and have better anti-mycobacterial properties.

Phenoxybenzamine hydrochloride was tested in a similar assay to fenoldopam. It inhibited the ATPase activity of ParA up to 50% at 250  $\mu\text{M}$  in a dose-dependent manner (Figure 6.12A). Phenoxybenzamine (trade name diabenzyline) is an alpha adrenergic receptors antagonists [333] that is used as an anti-hypertensive. It is also used for the treatment of benign prostatic hyperplasia [334] and hypoplastic left heart syndrome [333]. From the molecular structure of phenoxybenzamine hydrochloride (Figure 6.16), it appears that it may not bind to ParA active site. This would classify phenoxybenzamine hydrochloride as a non-competitive inhibitor. This is of significant interest as it presents a number of advantages over a competitive inhibitor that binds to the active site. Firstly, it will provide a novel MOA for drugs. Secondly, it will provide high selectivity as these drugs will not bind to the conserved region. Thirdly, a non-competitive inhibitor will not be affected by the concentration of substrate; that is, an increase in substrate concentration will not necessitate an increase in inhibitor concentration to have an effect [335]. Thus a low concentration of a potent inhibitor can still exhibit its effect at high substrate concentration. The type of inhibition mediated by phenoxybenzamine was elucidated using kinetics analysis which showed a mixed inhibition, that is decrease in the  $K_m$  and  $K_{cat}$  (Table 6.1). This suggests that phenoxybenzamine binds to the enzyme and enzyme-substrate complex thus making it a non-competitive inhibitor. *In silico* docking of phenoxybenzamine hydrochloride with a ParA model supports this data showing that it can bind to the substrate binding pocket. This may affect substrate binding via steric hinderance. While phenoxybenzamine hydrochloride exhibited high  $IC_{50}$  and  $MIC_{90}$  values in the three mycobacterial species, it was not very cytotoxic in mammalian cells thus was considered a reasonable hit to pursue as further medicinal chemistry may led to a more potent analogue with a novel mode of inhibition of ParA.

The last two compounds that were tested in this work were methiothepin mesylate and ( $\pm$ )-octoclothePIN maleate. Both compounds had a very similar pattern of inhibition which was not very surprising as their molecular structures are very similar (Figure 6.16). Both compounds inhibited the ParA ATPase activity up to  $\sim 75$ -80% upon addition of 50  $\mu\text{M}$  of drug but this inhibition was not increased upon increasing the concentration of the drugs (Figure 6.12B),

similar to the pattern of inhibition seen with vanadate (Figure 6.9), however, not as potent.



**Figure 6.16: Structure of compounds tested in the ParA ATPase assays**

**A)** Fenoldopam, a haloalkylamine, is a benzazepine derivative; **B)** phenoxybenzamine consists of two well spaced aromatic rings separated by a 3-carbon spacer; **C)** methiothepin and **D)** (±)-octoclotheptin with very similar structures.

Methiothepin is a non-selective 5-HT receptor antagonist [336, 337] that is used as an antipsychotic [338]. Other targets of methiothepin in humans include glucocerebrosidase, tumor protein p53, cytochrome P450 polypeptides, mitogen-activated protein kinase and potassium channels. Methiothepin also inhibits proliferation of several strains *P. falciparum* and yeast (PubChem Bioassays). It has been found to be inhibitory against *M. tuberculosis* H37Rv in a high-throughput screening by the Southern Research Institute however; its molecular target in *M. tuberculosis* is unknown (PubChem Bioassays). (±)-Octoclotheptin is a neuroleptic, also used as an antipsychotic and has potent antidopaminergic and antiserotonin actions [339]. Similar to methiothepin, other targets of (±)-octoclotheptin in humans include tumor protein p53, cytochrome P450 polypeptides and it also inhibits *P. falciparum* proliferation. Kinetics analysis of both these drugs showed a competitive mode of inhibition where the substrate binding affinity of ParA decreased in the presence of the drugs but this was overcome by high concentrations of substrate. This was confirmed with the *in silico* analysis where both methiothepin and octoclotheptin docked to the ligand

binding site (Figure 6.13B). (±)-OctoclothePIN maleate displayed the most potent anti-mycobacterial effects.

From these results phenoxybenzamine, methiothepin and octoclothePIN may be good drugs to pursue where these compounds can be further modified to increase their potency against mycobacteria. The lowest MIC was exhibited by (±)-octoclothePIN maleate in all mycobacterial species tested. Phenoxybenzamine only exhibited a low MIC under nutrient limited conditions in *M. smegmatis*. Nutrient limitation conditions could not be tested against the slow growers as the conditions are not yet optimized however, it is worth noting that all three species: *M. smegmatis*, *M. bovis* BCG and *M. tuberculosis* H37Ra, exhibited the same MIC under rich culture conditions (Table 6.2). Thus it is hypothesized that the slow-growers may show a similar MIC values under nutrient starvation conditions.

## 6.5 Summary

The work presented in this chapter demonstrates that *M. tuberculosis* ParA possesses ATPase activity. It also demonstrates that this activity is not stimulated by the ParB-*parS* nucleoprotein complex. The hydrolysis of ATP by ParA followed the Michaelis–Menten kinetics, yielding a  $K_m$  and  $K_{cat}$  of  $0.2 \pm 0.07$  mM and  $22.57 \pm 2.64$ /hr respectively. The recombinant enzyme also showed decreasing activity upon addition of a P-loop inhibitor, vanadate. To date there are no inhibitors of this enzyme described, creating the potential to exploit this as a novel anti-tubercular target to disrupt cell division.

The ATPase assays was used to verify inhibitors of ParA. Kinetics analysis shows that ParA has novel competitive (methiothepin and octoclothePIN) and non-competitive (phenoxybenzamine) inhibitors. This data is further supported by *in silico* analysis. These inhibitors represent a promising starting point for lead optimization. This study illustrates that the mycobacterial ParA warrant further drug discovery efforts.

## **CHAPTER SEVEN**

### **General Discussion**

## 7.1 General Discussion

The emergence of drug resistance in disease causing pathogens has led to a search for novel drug targets and their cognate inhibitors to identify structurally diverse antibiotics to combat disease. To overcome resistance, existing antibiotics have been extensively modified; however, efficacious modifications are reaching their limit thus steadily escalating the need for new classes of antibiotics [114].

Most drugs in current clinical use have been designed without the knowledge of their targets due to a lack of high-throughput technologies for target identification [140]. Nowadays, with the availability of whole genome sequences and high-throughput screening technologies, novel targets and their inhibitors can be more easily identified. The discovery of novel therapeutics requires the integration of various tools, information and approaches making it a resource-intensive process [340]. This calls for a multidisciplinary approach requiring biological, chemical and technological interventions that is organised in numerous phases; from the proposal of a target to the launch of the treatment [295].

This dissertation took a comprehensive approach towards mycobacterial drug discovery. It went from identifying potential drug targets, validating them as being required for growth under various culture conditions and identifying potential lead chemical structures against one of the validated targets.

### 7.1.1 Summary of experiments and their findings

The primary step in this research project was to find an ideal drug target in *M. tuberculosis* (section 3.1.1). To fulfil the criteria of such a target, the first thing that was considered was gene essentiality, a crucial factor in any drug discovery program [67, 137, 189, 341]. A study by Sassetti *et. al.* identified 614 genes in *M. tuberculosis* H37Rv that were unable to be recovered with transposon knockouts (Appendix 9.2.1) [126]. This served as a hypothesis that these genes may be essential, however, reputable proof of essentiality is only demonstrated by genetic manipulation of individual genes as whole genome analysis may have other



factors mitigating the failure to recover a mutant (section 4.1.1) [342]. Another factor that was considered in this study was gene conservation. There were two reasons for looking at gene conservation. Firstly, genes conserved across several bacterial species may indicate that they are essential for their viability and secondly highly conserved genes resist mutations thus targeting such genes may prevent the rise of resistance (section 1.7.2). Thus, with these two factors in mind, this study identified a total of 261 potential broad-spectrum targets in *M. tuberculosis* (Table 3.2). These targets belonged to various functional groups. The functional group that was undertaken for further analysis in this study was cell division. Cell division is an underexploited essential cellular process. It is an attractive pathway for drug targeting as it is required for the multiplication and viability of the organism and thus, is vital for infection [343]. Most bacterial cell division proteins do not have their catalytic activity determined [343]. There were nine putative essential cell division genes identified in this thesis. Three of these genes, Rv2440c (*obg*), Rv3917c (*parB*) and Rv3918c (*parA*), were chosen to be validated for their essentiality and potential as drug targets in the surrogate *M. smegmatis* (section 1.3.1).

Essentiality testing in this study was conducted using the expression of antisense RNA in an inducible system. This system determines gene essentiality via the potential decrease in target protein concentration due to the expression of antisense RNA of the gene encoding the protein (section 4.1.1.1). There are several reasons for using this method. Apart from providing a means to judge gene essentiality, it also provided a cost-effective assay for HTS and it can be easily transferred into an *in vivo* study. Homologues of the three *M. tuberculosis* genes: *obg*, *parA* and *parB* respectively, were expressed in *M. smegmatis* in the antisense orientation under different culture conditions (section 4.3.5). Testing the viability of the cells under various culture conditions was used to discriminate between essential and conditionally essential genes (section 1.7.2). This was to reveal an ideal target that was required under all conditions.

These experiments demonstrated that *obg* may be a conditionally essential gene as it only exhibited a low number of viable cells under rich culture

conditions upon antisense induction (Figure 4.10). All other culture conditions had the same number of viable cells as the control strain (Figure 4.12 and 4.14). *parB* has previously been demonstrated to be a non-essential gene in *M. smegmatis* [165]. This study supported the data showing no difference in the number of viable cells upon antisense expression of *parB* compared to the control strain under all culture conditions tested (Figure 4.10, 4.12 and 4.14). *parA* demonstrated that it may be required for cell viability under nutrient rich, carbon limitation and nitrogen limitation conditions. Antisense expression of *parA* exhibited a lower number of viable cells at late exponential phase in all three media compared to the control strain (Figure 4.10, 4.12 and 4.14 respectively).

Apart from being required for cell viability under different cultures conditions, the results from these experiments indicated that ParA was not as highly required under nutrient starvation conditions as under rich culture conditions. This is not surprising as ParA is a cell division protein and therefore is likely to be required under optimal respiration conditions. Nutrient limitation leads to a global shift in gene expression and a gradual shutdown of respiration to a basal level [249]. This was demonstrated in another study by Betts *et. al.* in *M. tuberculosis* where Rv3918c expression was down-regulated upon nutrient starvation. From these data, it was theorized that ParA may be required for exponential growth in mycobacteria. This theory is supported in a study by Casart *et. al.* [234]. Their study demonstrated that the levels of *parA* mRNA decreased between the exponential phase to stationary phase in *M. smegmatis* and *M. bovis* BCG.

Despite the transcriptional down-regulation of *parA* under nutrient starvation conditions, it was still required for the viability of the cells. Thus, ParA was chosen as a possible candidate for targeted drug screening. Nutrient limitation conditions are reflective of the environment in granulomas [344] thus inhibitory molecules against such a target would also be of significance for treating Tb bacilli under nutrient limitation conditions.

Target-driven approaches are directed at an individual gene/protein; however, the ultimate target is the whole organism. Thus targeted drug screening in this study was performed using a differential susceptibility assay (section 5.1.2) employing antisense RNA expression. This assay captured the target in a broader biological context. This screen also evaluated the druggability of the target in the whole cell, that is, the ability of the target to be modulated with small molecules [345]. A total of 5680 compounds were screened in three different culture conditions (section 5.3) as culture conditions affect the activity profile of compounds [124]. From the HTS, a total of 69 hits were identified; however, only approximately 0.35% of these were deemed as suitable hits to pursue. The purpose of this primary screening was to identify a number of hit compounds to be tested in a secondary assay to determine whether any lead compounds could be identified.

A secondary biochemical assay was developed to demonstrate the specificity of the identified bioactives. This functional assay was developed using the *M. tuberculosis* ParA protein encoded by Rv3918c. This assay was developed with the hypothesis that ParA may possess ATPase activity due to the conserved Walker A motif (Figure 6.14). It was also hypothesized that ParA may require another protein, ParB (encoded by Rv3917c), for its activity as in the plasmid partitioning system (section 4.4.1.3). The results from this study illustrated that the *M. tuberculosis* ParA possesses ATPase activity. It also demonstrated that ParA did not require other factors to stimulate its activity. Kinetic analysis of the enzyme established that the ATPase activity of ParA follows Michaelis-Menten kinetics.

This biochemical assay was tested against a known ATPase inhibitor, sodium metavanadate. This revealed that the *M. tuberculosis* ParA ATPase activity was capable of being inhibited by vanadate in a competitive manner. The assay was consequently used to screen the hits obtained in the cell-based assays. This revealed five potential inhibitors of ParA: NSC57850, fenoldopam, phenoxybenzamine, methiothepin and octoclotheptin. From the chemical structure of NSC57850 (Figure 6.15B), it appeared that it might be a competitive inhibitor

of ParA; however, due to limited supply of this compound it could not be tested with further experimentation. The modes of inhibition of the other compounds were determined by kinetic analyses. The results indicate that fenoldopam, methiothepin and octoclotheperin exhibit competitive inhibition while phenoxybenzamine exhibit a mixed inhibition of the ParA ATPase. These results were supported with *in silico* docking with a ParA model. Thus this study identified various compounds that affected the ATP binding ability of ParA while phenoxybenzamine was the only compound that was identified as affecting the ATPase activity per se.

Phenoxybenzamine, fenoldopam, and octoclotheperin were further evaluated for their MIC values in growth assays with *M. smegmatis*, *M. bovis* BCG and various strains of *M. tuberculosis* (Table 6.2) and SBC values in *M. smegmatis* (Table 6.3). From these experiments, it was deduced that phenoxybenzamine and octoclotheperin may be useful compounds to pursue.

Phenoxybenzamine had the highest inhibition in the *in vitro* ATPase assay (Figure 6.12A) and was found to be active under nitrogen starvation conditions in the HTS. This was reflected in the cell-based assays with *M. smegmatis* where the MIC value was high in nutrient rich media and low in nutrient starvation medias (C<sup>-</sup> and N<sup>-</sup>) with the lowest MIC value under N<sup>-</sup> conditions. However, phenoxybenzamine displayed high SBC values in all culture conditions. Octoclotheperin displayed a significantly lower MIC than all other compounds tested. Similar to phenoxybenzamine, it was picked up as a hit under N<sup>-</sup> conditions which was again reproduced in growth assay with *M. smegmatis*, giving a high MIC value with nutrient rich and C<sup>-</sup> media but a low MIC value in N<sup>-</sup> media. Interestingly, octoclotheperin gave the lowest SBC in rich media. These results indicate that *M. smegmatis* is more resistant to octoclotheperin in starvation conditions at stationary phase but susceptible in starvation conditions in the exponential phase. Such a compound may not be capable of treating persistent bacilli as it is believed that nutrient starvation is reflective in the persistent state of *M. tuberculosis* [249]. However, the MIC and SBC values for octoclotheperin in the starvation conditions were still lower than that of phenoxybenzamine.

Apart from having good anti-mycobacterial activity, it is critical to demonstrate that these leads have no overt toxicity associated with them. Thus phenoxybenzamine and octoclotheptin were tested for the cytotoxicity in HL-60 cells and compared against already published mouse LD<sub>50</sub> values (Table 6.4). This demonstrated that while octoclotheptin was potent against mycobacteria, it was also very toxic against the mammalian cells. However, it is important to note that despite its toxicity, octoclotheptin is already in clinical use for other purposes so it can still be tolerated in humans, making it a tractable lead compound. On the other hand, while phenoxybenzamine had a high MIC and SBC, it also has a relatively high mouse LD<sub>50</sub> value, especially when compared to streptomycin, a current anti-tubercular drug used for therapy, making it a good lead. These leads can now be used in another round of screening of related structures and SAR studies to develop new anti-tubercular drugs to target cell division.

### 7.1.2 Other findings

One subject that has so far not been discussed in this thesis is the phenotype seen with the *M. smegmatis*-pMind\_*parA*+*parB*\_GFP strain (Figure 4.6). Upon creation of this strain, it was found that the basal level of antisense expression of *parA* and *parB* simultaneously (from incomplete repression of the promoter) had a severe effect on the growth of the cells. It was hypothesized that ParB maybe required for ParA ATPase activity. Since ParA was found to be essential, it was deduced that down-regulating a protein that may be required for its activity, as well as the ParA protein, could not be tolerated by the cells. However, subsequent experiments demonstrated that ParB was not required for ParA ATPase activity.

Another possible explanation for the requirement of ParA and ParB together is that they may represent a regulatory pair *in vivo*. The ADP bound ParA has been discussed as a repressor for the *parAB* operon in other systems [244, 247]. In the P1 plasmid, the ParA repressor activity is stimulated by ParB; however, ParB has no effect on *par* gene expression on its own [247]. In *C. crescentus*, ParB regulates the activity of ParA by modulating its nucleotide-

bound state and acts as a nucleotide exchange factor [243]. This study illustrated that the ParAB proteins in *M. smegmatis* are required together. I hypothesize that the ParB protein may have a similar role as discussed above in regulating the repressor effect of ParA. A careful analysis of ParA-ParB interactions has to be executed to determine the exact role of these proteins in mycobacteria.

## 7.2 Critical Evaluation

### 7.2.1 Target validation technique

The target validation experiments in this study were carried out using an inducible antisense RNA expression system. Sophisticated as it is to use a regulatable system for gene expression, such a system never leads to zero gene product [286]. Of the three genes tested in this study, only antisense expression of *parA* showed a significant effect in all culture conditions tested thus was regarded as being essential for growth under these conditions. *parB* has previously been demonstrated to be non-essential which was validated in this study, however, *obg* could be an essential gene but the methods used in this study may have been inadequate to distinguish this. For example, if this protein is not required at very high concentrations in the cell, decreasing the amount of protein by antisense expression may not have reduced the amount of protein to a level that was sufficiently low to exhibit significant differences in cell viability. This is of significance in the persistent state, where metabolites are required at a much lower concentrations as a result of slow growth [115]. Therefore, before a gene is regarded as non-essential, it is imperative that a knock-out mutant is created and failure to grow the mutant in the absence of the gene be used as an absolute measure of its essentiality; something that was not performed in this study.

Another limitation to using genetics to validate drug targets is gene redundancy. In such situations, multiple genes may be individually dispensed although they can be inhibited by the same drug. An example of this are the rRNAs which are encoded by multiple genes that can be individually inactivated

---

without a substantial effect on cell growth [346]. However, aminoglycosides demonstrate that rRNA are valid targets [347].

### 7.2.2 High-throughput assays

The assay used for targeted HTS in this thesis while theoretically sound, in retrospect, may not have been optimal. This was deduced from the functional assay with the isolated protein. This *in vitro* assay demonstrated that vanadate, an ATPase inhibitor, only exhibited a 30% inhibition of the ParA ATPase activity at 50  $\mu$ M (Figure 6.9). The concentration of compounds used for screening in the HTS with the cell-based assay was 20  $\mu$ M and the inhibition cut-off used to discern hits was 40%. While this method lead to identification of some inhibitors of ParA, weak inhibitors of ParA would not have been picked out in these screens.

### 7.2.3 *In vitro* assay

The HTS in this thesis was carried out using *M. smegmatis* but the validation assay was performed with the *M. tuberculosis* protein. This is a logical progression as *M. tuberculosis* is eventually the target organism; however, this study did not investigate if the ParA of *M. smegmatis* and *M. tuberculosis* were functionally equivalent. If they are not a 100% equivalent, some of the hits from the HTS would be specific to *M. smegmatis* and would have no effect on the *M. tuberculosis* protein which may be one of the reasons why there were a high number of false positives (Figure 6.10).

## 7.3 Future Research Directions

As with any short-term research project, this work has generated a number of future goals and improvements that can be implemented in the search for ParA inhibitors.

### 7.3.1 Future goals

The compounds isolated from this study have to undergo further development in order to be regarded as viable anti-tubercular drugs. Some of the aims required to achieve this include:

- Firstly, the crystallization of the ParA protein as it is the shape of a target that determines the activity of a compound with respect to the molecular interaction between the target and the inhibitor. The 3-dimensional structure of the protein can be conclusively used to confirm the inhibitors identified in this study [348] and their mode of inhibition.
- Secondly, using the crystal structure of the enzyme-inhibitor complexes to understand the forces required for its formation [349]. This information can in turn be used for SAR [350] where the structure of the target and the bound ligands can be used to design analogue compounds [157]. This can allow rapid design of molecules with increased activity [105, 157] as the inhibitors identified in this study are required at very high concentrations and does not lead to 100% inhibition of ParA. Creation and screening an analog library can be quickly analysed using various computational methods which will greatly reduce the cost associated with drug development [351].
- Thirdly, testing all potent analogues in an *in vivo* system to determine their toxicity and efficacy to predict if these compounds have any clinical applications.



### 7.3.2 Improvements

There are also some improvements that can be applied to the HTS of ParA inhibitors.

- In order to minimize the number of false positives isolated against ParA from the cell-based assays, a secondary screen should be developed in a 1536 well plate format with the *in vitro* assay. Such an assay would decrease the reaction volume thereby increasing the drug concentration per reaction. This will either confirm the hits from the cell-based assays as potent inhibitors of ParA or as non-specific hits before further work is invested in the hits isolated from the primary screen.
- The use of the 3-D structure of the protein for high-throughput docking or *in silico* screens using different global optimization algorithms that can minimize the ligand–protein intermolecular and ligand intramolecular interaction energies to prioritize compounds to be tested *in vitro* [352]. This can allow a large number of drugs to be screened without huge costs associated with it.
- Finally, the results from all the screens can be used to create a hypothesis as to the chemotypes ParA binds to. This information can be used to screen the natural extract libraries to identify novel classes of compounds that inhibit ParA.

## 7.4 Concluding Remarks

The use of genomic data, molecular techniques and high-throughput screening, as discussed throughout this thesis, has presented a number of conclusions regarding the mycobacterial chromosome partitioning protein ParA.

- A. The ParA protein is essential for cell viability in *M. smegmatis* under nutrient rich, carbon starvation and nitrogen starvation conditions (chapter 4). It is inferred that ParA may also be essential for viability in *M. tuberculosis* as this protein is highly conserved, arranged in a similar operon in both species and is unable to sustain transposon insertions.
- B. The *M. tuberculosis* ParA protein possesses ATPase activity. This activity is not enhanced by the addition of other factors (ParB protein and *parS* DNA- chapter 6).
- C. *M. tuberculosis* ParA ATPase activity can be inhibited by a P-loop inhibitor, sodium metavanadate.
- D. Kinetic studies indicate that the *M. tuberculosis* ParA ATPase activity is capable of being inhibited by competitive and non-competitive inhibitor binding *in vitro* (chapter 6).
- E. Preliminary results indicate that *M. smegmatis* can be used as a model to identify specific inhibitors of the *M. tuberculosis* ParA protein (chapter 5 and 6).

In summary, this thesis demonstrates that ParA can be used as an effective target to find therapeutics to contribute towards an active and robust pipeline of anti-tubercular drugs. The compounds identified in this study now require considerable effort to be expended in detailed quantitative structure-activity relationships to define the minimal structures required for their activity. Ultimately, a combination of functional, structural and biophysical analyses will allow us to achieve this. This may lead to the development of novel ParA inhibitors.

## References

1. Andersen, P., *TB vaccines: progress and problems*. Trends Immunol, 2001. **22**(3): p. 160-8.
2. Guy, E.S. and A. Mallampalli, *Review: Managing TB in the 21st century: existing and novel drug therapies*. Ther Adv Respir Dis, 2008. **2**(6): p. 401-8.
3. Hershkovitz, I., et al., *Detection and Molecular Characterization of 9000-Year-Old Mycobacterium tuberculosis from a Neolithic Settlement in the Eastern Mediterranean*. PLoS ONE, 2008. **3**(10): p. e3426.
4. Andersen, P., et al., *The prognosis of latent tuberculosis: can disease be predicted?* Trends in Molecular Medicine, 2007. **13**(5): p. 175-182.
5. WHO, W.H.O., *Global Tuberculosis Control; Surveillance, Planning, Financing*. 2005.
6. WHO, W.H.O., *Global Tuberculosis Control; SURVEILLANCE, PLANNING, FINANCING*. 2008.
7. WHO, W.H.O., *Global Tuberculosis Control 2009: EPIDEMIOLOGY STRATEGY FINANCING*. 2009.
8. Corbett, E.L., et al., *The Growing Burden of Tuberculosis: Global Trends and Interactions With the HIV Epidemic*. Arch Intern Med, 2003. **163**(9): p. 1009-1021.
9. Bucher, H.C.G., Lauren E.; Guyatt, Gordon H.; Sudre, Philippe; Naef, Marcel; Sendi, Pedram; Battegay, Manuel, *Isoniazid prophylaxis for tuberculosis in HIV infection: a meta-analysis of randomized controlled trials*. AIDS, 1999. **13**(4): p. 501-507.
10. de Colombani, P., et al., *European framework to decrease the burden of TB/HIV*. Eur Respir J, 2004. **24**(3): p. 493-501.
11. Dye, C., *Global epidemiology of tuberculosis*. The Lancet, 2006. **367**(9514): p. 938-940.
12. WHO, W.H.O., *The Stop TB strategy*. 2006.
13. Howie, S., et al., *Tuberculosis in New Zealand, 1992-2001: a resurgence*. Arch Dis Child, 2005. **90**(11): p. 1157-1161.
14. Dilip Das, M.B., Lester Calder, *Tuberculosis epidemiology in New Zealand: 1995–2004*. The New Zealand Medical Journal, 2006. **119**(1243).
15. Dilip Das, M.B., Kamallesh Venugopal, Susan McAllister, *Why the tuberculosis incidence rate is not falling in New Zealand*. The New Zealand Medical Journal, 2006. **119**(1243).
16. Koch, R., *Aetiologie der Tuberculose*. Berliner Klinische Wochenschrift 1882. **19**: p. 221-230.
17. Cole, S.T., et al., *Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence*. Nature, 1998. **393**(6685): p. 537-544.
18. Rajagopalan, M., et al., *Mycobacterium smegmatis dnaA region and autonomous replication activity*. J. Bacteriol., 1995. **177**(22): p. 6527-6535.

19. Attorri, S., S. Dunbar, and J.E. Clarridge, III, *Assessment of Morphology for Rapid Presumptive Identification of Mycobacterium tuberculosis and Mycobacterium kansasii*. J. Clin. Microbiol., 2000. **38**(4): p. 1426-1429.
20. Fregnan, G.B. and D.W. Smith, *Description of various colony forms in mycobacteria*. J. Bacteriol., 1962. **83**(4): p. 819-827.
21. Shinnick, T.M. and R.C. Good, *Mycobacterial taxonomy*. European Journal of Clinical Microbiology & Infectious Diseases, 1994. **13**(11): p. 884-901.
22. H, T., *Bacteriology of mycobacteria: taxonomic and morphological characteristics*. Nippon Rinsho. , 1998. **56**(12): p. 3001-7.
23. Gillespie, S.H., *Tuberculosis: evolution in millennia and minutes*. Biochemical Society Transactions, 2007. **035**(5): p. 1317-1320.
24. Gutierrez, M.C., et al., *Ancient Origin and Gene Mosaicism of the Progenitor of Mycobacterium tuberculosis*. PLoS Pathog, 2005. **1**(1): p. e5.
25. Parsons, L.M., et al., *Rapid and Simple Approach for Identification of Mycobacterium tuberculosis Complex Isolates by PCR-Based Genomic Deletion Analysis*. J. Clin. Microbiol., 2002. **40**(7): p. 2339-2345.
26. Brennan, P.J., *Structure, function, and biogenesis of the cell wall of Mycobacterium tuberculosis*. Tuberculosis, 2003. **83**(1-3): p. 91-97.
27. , B.S., Leila Mendonça-Lima, Brigitte Gicquel, Jean-Marc Reyrat., *The Erp protein is anchored at the surface by a carboxy-terminal hydrophobic domain and is important for cell-wall structure in Mycobacterium smegmatis*. FEMS Microbiology Letters, 2004. **231**(2): p. 191-196.
28. Lustgarten, S., *Ueber spezifische Bacillen in syphilitischen Krankheitsprodukten*. Wiener Medizinische Wochenschrift 1, 1884.
29. Alvarez and E. T, *Recherches sur le bacille de Lustgarten*. Arch Physiol Normal Pathol, 1885: p. 303–321.
30. Sevrin, A. and A.C. Reboli, *Disseminated Mycobacterium smegmatis Infection Associated With an Implantable Cardioverter Defibrillator*. Infectious Diseases in Clinical Practice, 2009. **17**(5): p. 349-351  
10.1097/IPC.0b013e31819b8a9b.
31. Pierre-Audigier, C., et al., *Fatal disseminated Mycobacterium smegmatis infection in a child with inherited interferon  $\gamma$  receptor deficiency*. Clinical infectious diseases, 1997. **24**(5): p. 982-984.
32. Surbhi Gupta, D.C., *Stress Responses in Mycobacteria*. IUBMB Life, 2005. **57**(3): p. 149-159.
33. Chaturvedi, V., et al., *Evaluation of Mycobacterium smegmatis as a possible surrogate screen for selecting molecules active against multi-drug resistant Mycobacterium tuberculosis*. The Journal of General and Applied Microbiology, 2007. **53**(6): p. 333-337.
34. Wang, R. and E.M. Marcotte, *The Proteomic Response of Mycobacterium smegmatis to Anti-Tuberculosis Drugs Suggests Targeted Pathways*. Journal of Proteome Research, 2008. **7**(3): p. 855-865.
35. Tyagi, J.S. and D. Sharma, *Mycobacterium smegmatis and tuberculosis*. Trends in Microbiology, 2002. **10**(2): p. 68-69.
36. Smith, I., *Mycobacterium tuberculosis pathogenesis and molecular determinants of virulence*. Clin Microbiol Rev, 2003. **16**(3): p. 463-96.

- 
37. Mehta, P.K., et al., *Entry and intracellular replication of Mycobacterium tuberculosis in cultured human microvascular endothelial cells*. Microbial Pathogenesis, 2006. **41**(2-3): p. 119-124.
  38. Clark-Curtiss, J.E. and S.E. Haydel, *Molecular genetics of mycobacterium tuberculosis pathogenesis*. Annual Review of Microbiology, 2003. **57**(1): p. 517.
  39. Saunders, B.M. and A.M. Cooper, *Restraining mycobacteria: Role of granulomas in mycobacterial infections*. Immunol Cell Biol, 2000. **78**(4): p. 334-341.
  40. Mehta, P., et al., *Comparison of in vitro models for the study of Mycobacterium tuberculosis invasion and intracellular replication*. Infect. Immun., 1996. **64**(7): p. 2673-2679.
  41. Harries, A.D. and C. Dye, *Tuberculosis*. Annals of Tropical Medicine and Parasitology, 2006. **100**: p. 415-431.
  42. Kaufmann, S.H.E., *How can immunology contribute to the control of tuberculosis?* Nat Rev Immunol, 2001. **1**(1): p. 20-30.
  43. Russell, D.G., *Who puts the tubercle in tuberculosis?* Nat Rev Micro, 2007. **5**(1): p. 39-47.
  44. Riska, P.F. and S. Carleton, *Latent tuberculosis: Models, mechanisms, and novel prospects for eradication*. Seminars in Pediatric Infectious Diseases, 2002. **13**(4): p. 263-272.
  45. Flynn, J.L. and J. Chan, *Tuberculosis: Latency and Reactivation*. Infect. Immun., 2001. **69**(7): p. 4195-4201.
  46. Duncan, K., *Progress in TB drug development and what is still needed*. Tuberculosis, 2003. **83**(1-3): p. 201-207.
  47. Onyebujoh, P., et al., *Treatment of tuberculosis: present status and future prospects*. Bulletin of the World Health Organization, 2005. **83**: p. 857-865.
  48. Chan, E.D. and M.D. Iseman, *Current medical treatment for tuberculosis*. BMJ, 2002. **325**(7375): p. 1282-1286.
  49. Frieden, T.R., et al., *Tuberculosis*. The Lancet, 2003. **362**(9387): p. 887-899.
  50. Spigelman, Melvin K., *New Tuberculosis Therapeutics: A Growing Pipeline*. The Journal of Infectious Diseases, 2007. **196**(s1): p. S28-S34.
  51. Brennan, P.J. and D. B. Young, *Handbook of Anti-Tuberculosis Agents*. Tuberculosis, 2008. **88**(2): p. 85-170.
  52. Gillespie, S.H., *Evolution of Drug Resistance in Mycobacterium tuberculosis: Clinical and Molecular Perspective*. Antimicrob. Agents Chemother., 2002. **46**(2): p. 267-274.
  53. Espinal, M.A., *The global situation of MDR-TB*. Tuberculosis, 2003. **83**(1-3): p. 44-51.
  54. Di Perri, G. and S. Bonora, *Which agents should we use for the treatment of multidrug-resistant Mycobacterium tuberculosis?* J. Antimicrob. Chemother., 2004. **54**(3): p. 593-602.
  55. Sharma, S.K. and A. Mohan, *Multidrug-Resistant Tuberculosis\**. Chest, 2006. **130**(1): p. 261-272.
  56. Gardner, C.A., T. Acharya, and A. Pablos-Méndez, *The Global Alliance for Tuberculosis Drug Development--Accomplishments and Future Directions*. Clinics in Chest Medicine, 2005. **26**(2): p. 341-347.

- 
57. Dye, C., *Doomsday postponed? Preventing and reversing epidemics of drug-resistant tuberculosis*. Nat Rev Micro, 2009. **7**(1): p. 81-87.
  58. Wing Wai YEW, C.C.L., *Management of multidrug-resistant tuberculosis: Update 2007*. Respiriology, 2008. **13**(1): p. 21-46.
  59. WHO, W.H.O., *The Global MDR-TB and XDR-TB response plan 2007-2008*. 2007.
  60. Singh, J.A., R. Upshur, and N. Padayatchi, *XDR-TB in South Africa: No Time for Denial or Complacency*. PLoS Med, 2007. **4**(1): p. e50.
  61. Jones, K.D.J., T. Hesketh, and J. Yudkin, *Extensively drug-resistant tuberculosis in sub-Saharan Africa: an emerging public-health concern*. Transactions of the Royal Society of Tropical Medicine and Hygiene, 2008. **102**(3): p. 219-224.
  62. Velayati AA, M.M., Farnia P, Tabarsi P, Ghanavi J, Ziazarifi AH, Hoffner SE, *Emergence of new forms of totally drug-resistant tuberculosis bacilli: super extensively drug-resistant tuberculosis or totally drug-resistant strains in Iran*. Chest, 2009. **136**(2): p. 420-425.
  63. Springer, B., et al., *Mechanisms of Streptomycin Resistance: Selection of Mutations in the 16S rRNA Gene Conferring Resistance*. Antimicrob. Agents Chemother., 2001. **45**(10): p. 2877-2884.
  64. Amin, A.G., et al., *EmbA is an essential arabinosyltransferase in Mycobacterium tuberculosis*. Microbiology, 2008. **154**(1): p. 240-248.
  65. Kocagoz, T., et al., *Gyrase mutations in laboratory-selected, fluoroquinolone-resistant mutants of Mycobacterium tuberculosis H37Ra*. Antimicrob. Agents Chemother., 1996. **40**(8): p. 1768-1774.
  66. Nguyen, L. and C.J. Thompson, *Foundations of antibiotic resistance in bacterial physiology: the mycobacterial paradigm*. Trends in Microbiology, 2006. **14**(7): p. 304-312.
  67. Chalker, A.F. and R.D. Lunsford, *Rational identification of new antibacterial drug targets that are essential for viability using a genomics-based approach*. Pharmacology & Therapeutics, 2002. **95**(1): p. 1-20.
  68. VON GROLL, A., et al., *Fluoroquinolone resistance in Mycobacterium tuberculosis and mutations in gyrA and gyrB*. Antimicrob. Agents Chemother., 2009: p. AAC.00287-09.
  69. Mdluli, K. and Z. Ma, *Mycobacterium tuberculosis DNA gyrase as a target for drug discovery*. Infect Disord Drug Targets, 2007. **7**(2): p. 159 - 168.
  70. A., W., *The role of cytochrome P450 in antiretroviral drug interactions*. Expert Opin Drug Metab Toxicol, 2007. **3**(4): p. 583-598.
  71. Wang JY, W.J., Tsai TH, Hsu CL, Yu CJ, Hsueh PR, Lee LN, Yang PC., *Adding moxifloxacin is associated with a shorter time to culture conversion in pulmonary tuberculosis*. Int J Tuberc Lung Dis, 2010. **14**(1): p. 65-71.
  72. Haagsma, A.C., et al., *Selectivity of TMC207 towards Mycobacterial ATP Synthase Compared with That towards the Eukaryotic Homologue*. Antimicrob. Agents Chemother., 2009. **53**(3): p. 1290-1292.
  73. Andries, K., et al., *A Diarylquinoline Drug Active on the ATP Synthase of Mycobacterium tuberculosis*. Science, 2005. **307**(5707): p. 223-227.

- 
74. Koul, A., et al., *Diarylquinolines Are Bactericidal for Dormant Mycobacteria as a Result of Disturbed ATP Homeostasis*. Journal of Biological Chemistry, 2008. **283**(37): p. 25273-25280.
  75. Lounis, N., et al., *R207910 (TMC207): A new antibiotic for the treatment of tuberculosis* Med Mal Infect. **In Press, Corrected Proof**.
  76. Diacon, A.H., et al., *The Diarylquinoline TMC207 for Multidrug-Resistant Tuberculosis*. N Engl J Med, 2009. **360**(23): p. 2397-2405.
  77. Huitric, E., et al., *In Vitro Antimycobacterial Spectrum of a Diarylquinoline ATP Synthase Inhibitor*. Antimicrob. Agents Chemother., 2007. **51**(11): p. 4202-4204.
  78. Lounis, N., et al., *Combinations of R207910 with Drugs Used To Treat Multidrug-Resistant Tuberculosis Have the Potential To Shorten Treatment Duration*. Antimicrob. Agents Chemother., 2006. **50**(11): p. 3543-3547.
  79. Stover, C.K., et al., *A small-molecule nitroimidazopyran drug candidate for the treatment of tuberculosis*. Nature, 2000. **405**(6789): p. 962-966.
  80. Lenaerts, A.J., et al., *Preclinical Testing of the Nitroimidazopyran PA-824 for Activity against Mycobacterium tuberculosis in a Series of In Vitro and In Vivo Models*. Antimicrob. Agents Chemother., 2005. **49**(6): p. 2294-2301.
  81. Nuermberger, E., et al., *Powerful Bactericidal and Sterilizing Activity of a Regimen Containing PA-824, Moxifloxacin, and Pyrazinamide in a Murine Model of Tuberculosis*. Antimicrob. Agents Chemother., 2008. **52**(4): p. 1522-1524.
  82. Trevors, J., *Genome size in bacteria*. Antonie van Leeuwenhoek, 1996. **69**(1): p. 293-303.
  83. Ochman, H. and L.M. Davalos, *The Nature and Dynamics of Bacterial Genomes*. Science, 2006. **311**(5768): p. 1730-1733.
  84. Ioerger, T.R. and J.C. Sacchettini, *Structural genomics approach to drug discovery for Mycobacterium tuberculosis*. Current Opinion in Microbiology, 2009. **12**(3): p. 318-325.
  85. Kunin, V., et al., *A Bioinformatician's Guide to Metagenomics*. Microbiol. Mol. Biol. Rev., 2008. **72**(4): p. 557-578.
  86. Koonin, E.V., A.R. Mushegian, and K.E. Rudd, *Sequencing and analysis of bacterial genomes*. Current Biology, 1996. **6**(4): p. 404-416.
  87. Reeves, G.A., D. Talavera, and J.M. Thornton, *Genome and proteome annotation: organization, interpretation and integration*. Journal of The Royal Society Interface, 2009. **6**(31): p. 129-147.
  88. Eisenstein, E., et al., *Biological function made crystal clear -- annotation of hypothetical proteins via structural genomics*. Current Opinion in Biotechnology, 2000. **11**(1): p. 25-30.
  89. Baker, E., *Structural genomics as an approach towards understanding the biology of tuberculosis*. Journal of Structural and Functional Genomics, 2007. **8**(2): p. 57-65.
  90. Brenner, S.E., *A tour of structural genomics*. Nat Rev Genet, 2001. **2**(10): p. 801-809.
  91. Ansong, C., et al., *Proteogenomics: needs and roles to be filled by proteomics in genome annotation*. Brief Funct Genomic Proteomic, 2008. **7**(1): p. 50-62.

- 
92. Fournier, P.-E., M. Drancourt, and D. Raoult, *Bacterial genome sequencing and its use in infectious diseases*. The Lancet Infectious Diseases, 2007. **7**(11): p. 711-723.
  93. Saier, M.H., Jr., *Genome Sequencing and Informatics: New Tools for Biochemical Discoveries*. Plant Physiol., 1998. **117**(4): p. 1129-1133.
  94. Frishman, D. and H.-W. Mewes, *Genome-based structural biology*. Progress in Biophysics and Molecular Biology, 1999. **72**(1): p. 1-17.
  95. Monaghan, R.L. and J.F. Barrett, *Antibacterial drug discovery--Then, now and the genomics future*. Biochemical Pharmacology, 2006. **71**(7): p. 901-909.
  96. Allsop, A.E., *Bacterial genome sequencing and drug discovery*. Current Opinion in Biotechnology, 1998. **9**(6): p. 637-642.
  97. Loferer, H., et al., *Integrated bacterial genomics for the discovery of novel antimicrobials*. Drug Discovery Today, 2000. **5**(3): p. 107-114.
  98. Betts, J.C., et al., *Comparison of the proteome of Mycobacterium tuberculosis strain H37Rv with clinical isolate CDC 1551*. Microbiology, 2000. **146**(12): p. 3205-3216.
  99. Camus, J.-C., et al., *Re-annotation of the genome sequence of Mycobacterium tuberculosis H37Rv*. Microbiology, 2002. **148**(10): p. 2967-2973.
  100. Gold, B., et al., *Identification of a copper-binding metallothionein in pathogenic mycobacteria*. Nat Chem Biol, 2008. **4**(10): p. 609-616.
  101. Williams, K.J. and K. Duncan, *Current Strategies for Identifying and Validating Targets for New Treatment-Shortening Drugs for TB*. Current Molecular Medicine, 2007. **7**(3): p. 297-307.
  102. Manabe, Y.C., J. A. M. Dannenberg, and W.R. Bishai, *What we can learn from the Mycobacterium tuberculosis genome sequencing projects*. The International Journal of Tuberculosis and Lung Disease, 2000. **4**: p. S18-S23.
  103. Tatusov, R.L., E.V. Koonin, and D.J. Lipman, *A Genomic Perspective on Protein Families*. Science, 1997. **278**(5338): p. 631-637.
  104. Pathania, R.a.B., ED, *Small and lethal: searching for new antibacterial compounds with novel modes of action*. Biochem Cell Biol, 2008. **86**(2): p. 111-115.
  105. Hughes, D., *Exploiting genomics, genetics and chemistry to combat antibiotic resistance*. Nat Rev Genet, 2003. **4**(6): p. 432-441.
  106. Moir, D.T., et al., *Genomics and Antimicrobial Drug Discovery*. Antimicrob. Agents Chemother., 1999. **43**(3): p. 439-446.
  107. Sassetti, C.M., D.H. Boyd, and E.J. Rubin, *Comprehensive identification of conditionally essential genes in mycobacteria*. Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(22): p. 12712-12717.
  108. Boshoff, H.I., et al., *The transcriptional responses of M. tuberculosis to inhibitors of metabolism: Novel insights into drug mechanisms of action*. J. Biol. Chem., 2004: p. M406796200.
  109. Showalter, H.D.H. and W.A. Denny, *A roadmap for drug discovery and its translation to small molecule agents in clinical development for tuberculosis treatment*. Tuberculosis, 2008. **88**(Supplement 1): p. S3-S17.



- 
110. Duncan, K. and C.E. Barry, *Prospects for new antitubercular drugs*. Current Opinion in Microbiology, 2004. **7**(5): p. 460-465.
  111. Stephens, C.M. and M.T. Laub, *Microbial Genomics: All That You Can't Leave behind*. Current Biology, 2003. **13**(14): p. R571-R573.
  112. Donadio, S., et al., *Targets and assays for discovering novel antibacterial agents*. Journal of Biotechnology, 2002. **99**(3): p. 175-185.
  113. Miesel, L., J. Greene, and T.A. Black, *Genetic strategies for antibacterial drug discovery*. Nat Rev Genet, 2003. **4**(6): p. 442-456.
  114. Payne, D.J., et al., *Drugs for bad bugs: confronting the challenges of antibacterial discovery*. Nat Rev Drug Discov, 2007. **6**(1): p. 29-40.
  115. Balganes, T.S., P.M. Alzari, and S.T. Cole, *Rising standards for tuberculosis drug development*. Trends in Pharmacological Sciences, 2008. **29**(11): p. 576-581.
  116. S. B. Snapper, R.E.M., S. Mustafa, T. Kieser and W. R. Jacobs Jr, *Isolation and characterization of efficient plasmid transformation mutants of Mycobacterium smegmatis*. Molecular Microbiology, 1990. **4**(11): p. 1911-1919.
  117. Brosch, R., et al., *Genome plasticity of BCG and impact on vaccine efficacy*. Proceedings of the National Academy of Sciences, 2007. **104**(13): p. 5596-5601.
  118. Steenken, W., Jr., W.H. Oatway, Jr., and S.A. Petroff, *BIOLOGICAL STUDIES OF THE TUBERCLE BACILLUS: III. DISSOCIATION AND PATHOGENICITY OF THE R AND S VARIANTS OF THE HUMAN TUBERCLE BACILLUS (H37)*. J. Exp. Med., 1934. **60**(4): p. 515-540.
  119. Wang S.Y., Z.H.J., Lv L.D., Wang B.F., Zhang X.L., Pu S.Y., Zhu G.F., Wang H.H., Zhao G.P., Zhang Y, *Complete genomic sequence of Mycobacterium tuberculosis strain H37Ra, a non-pathogenic variant closely related to the well-characterized pathogenic strain H37Rv*. GenBank, 2007.
  120. Blokpoel, M.C.J., et al., *Tetracycline-inducible gene regulation in mycobacteria*. Nucl. Acids Res., 2005. **33**(2): p. e22-.
  121. Bourn, W.R., et al., *Creation and characterisation of a high-copy-number version of the pAL5000 mycobacterial replicon*. Tuberculosis, 2007. **87**(6): p. 481-488.
  122. Luo, Y., et al., *A novel method for monitoring Mycobacterium bovis BCG trafficking with recombinant BCG expressing green fluorescent protein*. Clin. Diagn. Lab. Immunol., 1996. **3**(6): p. 761-768.
  123. Ehrt, S. and D. Schnappinger, *Controlling gene expression in mycobacteria*. Future Microbiology, 2006. **1**(2): p. 177-184.
  124. Miller, C.H., et al., *Modifying culture conditions in chemical library screening identifies alternative inhibitors of mycobacteria*. Antimicrob. Agents Chemother., 2009: p. AAC.00803-09.
  125. Hartmans, S., and J. A. M. De Bont., *The genus Mycobacterium—nonmedical*. The prokaryotes, a handbook on the biology of bacteria: ecophysiology, isolation, identification, application, 1992: p. 1215–1237.
  126. Christopher M. Sassetti, D.H.B., Eric J. Rubin., *Genes required for mycobacterial growth defined by high density mutagenesis*. Molecular Microbiology, 2003. **48**(1): p. 77-84.

- 
127. Parish, T. and A.C. Brown, *Mycobacteria Protocols - Methods in Molecular Biology*, 2nd Ed. 2008.
  128. Oganessian, N., et al., *Effect of osmotic stress and heat shock in recombinant protein overexpression and crystallization*. Protein Expression and Purification, 2007. **52**(2): p. 280-285.
  129. Studier, F.W., *Protein production by auto-induction in high-density shaking cultures*. Protein Expression and Purification, 2005. **41**(1): p. 207-234.
  130. Pratto, F., et al., *Streptococcus pyogenes pSM19035 requires dynamic assembly of ATP-bound ParA and ParB on parS DNA during plasmid segregation*. Nucleic Acids Res, 2008. **36**(11): p. 3676-89.
  131. Hirano, M. and T. Hirano, *ATP-dependent aggregation of single-stranded DNA by a bacterial SMC homodimer*. EMBO J, 1998. **17**(23): p. 7139-7148.
  132. Jensen, R.B. and K. Gerdes, *Partitioning of plasmid R1. The ParM protein exhibits ATPase activity and interacts with the centromere-like ParR-parC complex*. Journal of Molecular Biology, 1997. **269**(4): p. 505-513.
  133. McQuade, T.J., et al., *A nonradioactive high-throughput assay for screening and characterization of adenylation domains for nonribosomal peptide combinatorial biosynthesis*. Analytical Biochemistry, 2009. **386**(2): p. 244-250.
  134. Zhang, J.-H., T.D.Y. Chung, and K.R. Oldenburg, *A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays*. J Biomol Screen, 1999. **4**(2): p. 67-73.
  135. McDevitt, D. and M. Rosenberg, *Exploiting genomics to discover new antibiotics*. Trends in Microbiology, 2001. **9**(12): p. 611-617.
  136. Yang, Y., S.J. Adelstein, and A.I. Kassir, *Target discovery from data mining approaches*. Drug Discovery Today, 2009. **14**(3-4): p. 147-154.
  137. Hasan, S., et al., *Prioritizing Genomic Drug Targets in Pathogens: Application to Mycobacterium tuberculosis*. PLoS Computational Biology, 2006. **2**(6): p. e61.
  138. Chakhaiyar, P. and S.E. Hasnain, *Defining the Mandate of Tuberculosis Research in a Postgenomic Era*. Medical Principles and Practice, 2004. **13**(4): p. 177-184.
  139. Terstappen, G.C. and A. Reggiani, *In silico research in drug discovery*. Trends in Pharmacological Sciences, 2001. **22**(1): p. 23-26.
  140. Raman, K., K. Yeturu, and N. Chandra, *targetTB: A target identification pipeline for Mycobacterium tuberculosis through an interactome, reactome and genome-scale structural analysis*. BMC Systems Biology, 2008. **2**(1): p. 109.
  141. Fleischmann, R., et al., *Whole-genome random sequencing and assembly of Haemophilus influenzae Rd*. Science, 1995. **269**(5223): p. 496-512.
  142. Medini, D., et al., *Microbiology in the post-genomic era*. Nat Rev Micro, 2008. **6**(6): p. 419-430.
  143. Gil, R., et al., *Determination of the Core of a Minimal Bacterial Gene Set*. Microbiol. Mol. Biol. Rev., 2004. **68**(3): p. 518-537.
  144. Pucci, M.J., *Use of genomics to select antibacterial targets*. Biochemical Pharmacology, 2006. **71**(7): p. 1066-1072.

- 
145. Yoneyama, H. and R. Katsumata, *Antibiotic Resistance in Bacteria and Its Future for Novel Antibiotic Development*. Bioscience, Biotechnology, and Biochemistry, 2006. **70**(5): p. 1060-1075.
  146. Apfel, C.M., et al., *Peptide Deformylase as an Antibacterial Drug Target: Target Validation and Resistance Development*. Antimicrob. Agents Chemother., 2001. **45**(4): p. 1058-1064.
  147. Mills, S.D., *When will the genomics investment pay off for antibacterial discovery?* Biochemical Pharmacology, 2006. **71**(7): p. 1096-1102.
  148. Bumann, D., *Has nature already identified all useful antibacterial targets?* Current Opinion in Microbiology, 2008. **11**(5): p. 387-392.
  149. Meinnel, T., Y. Mechulam, and S. Blanquet, *Methionine as translation start signal: A review of the enzymes of the pathway in Escherichia coli*. Biochimie, 1993. **75**(12): p. 1061-1075.
  150. Chen, D.Z., et al., *Actinonin, a Naturally Occurring Antibacterial Agent, Is a Potent Deformylase Inhibitor*. Biochemistry, 2000. **39**(6): p. 1256-1262.
  151. Fritsche, T.R., et al., *Comparative Antimicrobial Characterization of LBM415 (NVP PDF-713), a New Peptide Deformylase Inhibitor of Clinical Importance*. Antimicrob. Agents Chemother., 2005. **49**(4): p. 1468-1476.
  152. Ramanathan-Girish, S., et al., *Pharmacokinetics in Animals and Humans of a First-in-Class Peptide Deformylase Inhibitor*. Antimicrob. Agents Chemother., 2004. **48**(12): p. 4835-4842.
  153. Fieulaine, S., et al., *The Crystal Structure of Mitochondrial (Type 1A) Peptide Deformylase Provides Clear Guidelines for the Design of Inhibitors Specific for the Bacterial Forms*. J. Biol. Chem., 2005. **280**(51): p. 42315-42324.
  154. Hawser, S., S. Lociuero, and K. Islam, *Dihydrofolate reductase inhibitors as antibacterial agents*. Biochemical Pharmacology, 2006. **71**(7): p. 941-948.
  155. Schneider, P., S. Hawser, and K. Islam, *Iclaprim, a novel diaminopyrimidine with potent activity on trimethoprim sensitive and resistant bacteria*. Bioorganic & Medicinal Chemistry Letters, 2003. **13**(23): p. 4217-4221.
  156. Sincak, C.A. and J.M. Schmidt, *Iclaprim, a Novel Diaminopyrimidine for the Treatment of Resistant Gram-Positive Infections*. Ann Pharmacother, 2009. **43**(6): p. 1107-1114.
  157. Nie, Z., et al., *Structure-Based Design, Synthesis, and Study of Potent Inhibitors of  $\hat{F}^2$ -Ketoacyl-acyl Carrier Protein Synthase III as Potential Antimicrobial Agents*. Journal of Medicinal Chemistry, 2005. **48**(5): p. 1596-1609.
  158. Lu, H. and P.J. Tonge, *Inhibitors of FabI, an Enzyme Drug Target in the Bacterial Fatty Acid Biosynthesis Pathway*. Accounts of Chemical Research, 2008. **41**(1): p. 11-20.
  159. Wang, J., et al., *Platensimycin is a selective FabF inhibitor with potent antibiotic properties*. Nature, 2006. **441**(7091): p. 358-361.
  160. Balemans, W., et al., *Essentiality of FASII pathway for Staphylococcus aureus*. Nature. **463**(7279): p. E3-E3.

- 
161. Thanassi, J.A., et al., *Identification of 113 conserved essential genes using a high-throughput gene disruption system in Streptococcus pneumoniae*. Nucl. Acids Res., 2002. **30**(14): p. 3152-3162.
  162. Sassetti, C., D. Boyd, and E. Rubin, *Genes required for mycobacterial growth defined by high density mutagenesis*. Molecular Microbiology, 2003. **48**: p. 77 - 84.
  163. Pang, X., et al., *Evidence for complex interactions of stress-associated regulons in an mprAB deletion mutant of Mycobacterium tuberculosis*. Microbiology, 2007. **153**(4): p. 1229-1242.
  164. McAdam, R.A., et al., *Characterization of a Mycobacterium tuberculosis H37Rv transposon library reveals insertions in 351 ORFs and mutants with altered virulence*. Microbiology, 2002. **148**(10): p. 2975-2986.
  165. Jakimowicz, D., et al., *Characterization of the mycobacterial chromosome segregation protein ParB and identification of its target in Mycobacterium smegmatis*. Microbiology, 2007. **153**(12): p. 4050-4060.
  166. Anishetty, S., M. Pulimi, and G. Pennathur, *Potential drug targets in Mycobacterium tuberculosis through metabolic pathway analysis*. Computational Biology and Chemistry, 2005. **29**(5): p. 368-378.
  167. Wei, J.-R. and E.J. Rubin, *The many roads to essential genes*. Tuberculosis, 2008. **88**(Supplement 1): p. S19-S24.
  168. Vissa, V. and P. Brennan, *The genome of Mycobacterium leprae: a minimal mycobacterial gene set*. Genome Biology, 2001. **2**(8): p. reviews1023.1 - reviews1023.8.
  169. Cole, S., et al., *The decaying genome of Mycobacterium leprae*. Lepr Rev, 2001. **72**: p. 387-398.
  170. Mushegian, A.R. and E.V. Koonin, *A minimal gene set for cellular life derived by comparison of complete bacterial genomes*. Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(19): p. 10268-10273.
  171. Akerley, B.J., et al., *A genome-scale analysis for identification of genes required for growth or survival of Haemophilus influenzae*. Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(2): p. 966-971.
  172. Eisenreich, W., et al., *Biosynthesis of isoprenoids via the non-mevalonate pathway*. Cell Mol Life Sci, 2004. **61**(12): p. 1401 - 1426.
  173. Arigoni, F., et al., *A genome-based approach for the identification of essential bacterial genes*. Nat Biotech, 1998. **16**(9): p. 851-856.
  174. Koonin EV, M.A., Bork P, *Non-orthologous gene displacement*. Trends in Genetics, 1996. **12**(9): p. 334-336.
  175. de Berardinis, V., et al., *A complete collection of single-gene deletion mutants of Acinetobacter baylyi ADP1*. Mol Syst Biol, 2008. **4**.
  176. Kobayashi, K., et al., *Essential Bacillus subtilis genes*. Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(8): p. 4678-4683.
  177. Gerdes, S.Y., et al., *Experimental Determination and System Level Analysis of Essential Genes in Escherichia coli MG1655*. J. Bacteriol., 2003. **185**(19): p. 5673-5684.
  178. Baba, T., et al., *Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection*. Mol Syst Biol, 2006. **2**.

- 
179. Gallagher, L.A., et al., *A comprehensive transposon mutant library of Francisella novicida, a bioweapon surrogate*. Proceedings of the National Academy of Sciences, 2007. **104**(3): p. 1009-1014.
  180. Salama, N.R., B. Shepherd, and S. Falkow, *Global Transposon Mutagenesis and Essential Gene Analysis of Helicobacter pylori*. J. Bacteriol., 2004. **186**(23): p. 7926-7935.
  181. Glass, J.I., et al., *Essential genes of a minimal bacterium*. Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(2): p. 425-430.
  182. Christopher T. French, et al., *Large-scale transposon mutagenesis of Mycoplasma pulmonis*. Molecular Microbiology, 2008. **69**(1): p. 67-76.
  183. Liberati, N.T., et al., *An ordered, nonredundant library of Pseudomonas aeruginosa strain PA14 transposon insertion mutants*. Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(8): p. 2833-2838.
  184. Karin Knuth, et al., *Large-scale identification of essential Salmonella genes by trapping lethal insertions*. Molecular Microbiology, 2004. **51**(6): p. 1729-1744.
  185. Ji, Y., et al., *Identification of Critical Staphylococcal Genes Using Conditional Phenotypes Generated by Antisense RNA*. Science, 2001. **293**(5538): p. 2266-2269.
  186. R. Allyn Forsyth, et al., *A genome-wide strategy for the identification of essential genes in Staphylococcus aureus*. Molecular Microbiology, 2002. **43**(6): p. 1387-1400.
  187. SOO, K.K., et al., *Screening of Essential Genes in Staphylococcus aureus N315 Using Comparative Genomics and Allelic Replacement Mutagenesis*. J Microbiol Biotechnol, 2006. **16**: p. 623-632.
  188. Judson, N. and J. Mekalanos, *TnAraOut, a transposon-based approach to identify and characterize essential bacterial genes*. Nat Biotechnol. , 2000. **18**(7): p. 740-745.
  189. Sakharkar, K., M. Sakharkar, and V. Chow, *A novel genomics approach for the identification of drug targets in pathogens, with special reference to Pseudomonas aeruginosa*. In Silico Biol, 2004. **4**(3): p. 355 - 360.
  190. Telenti, A., et al., *Detection of rifampicin-resistance mutations in Mycobacterium tuberculosis*. The Lancet, 1993. **341**(8846): p. 647-651.
  191. Sreevatsan, S., et al., *Characterization of rpsL and rrs mutations in streptomycin-resistant Mycobacterium tuberculosis isolates from diverse geographic localities*. Antimicrob. Agents Chemother., 1996. **40**(4): p. 1024-1026.
  192. Feng, Z. and R.G. Barletta, *Roles of Mycobacterium smegmatis D-Alanine:D-Alanine Ligase and D-Alanine Racemase in the Mechanisms of Action of and Resistance to the Peptidoglycan Inhibitor D-Cycloserine*. Antimicrob. Agents Chemother., 2003. **47**(1): p. 283-291.
  193. Hodgson, J.E., et al., *Molecular characterization of the gene encoding high-level mupirocin resistance in Staphylococcus aureus J2870*. Antimicrob. Agents Chemother., 1994. **38**(5): p. 1205-1208.
  194. Patel, JeanÂ B., RachelÂ J. Gorwitz, and JohnÂ A. Jernigan, *Antimicrobial Resistance: Mupirocin Resistance*. Clinical Infectious Diseases, 2009. **49**(6): p. 935-941.

- 
195. Nakama, T., O. Nureki, and S. Yokoyama, *Structural Basis for the Recognition of Isoleucyl-Adenylate and an Antibiotic, Mupirocin, by Isoleucyl-tRNA Synthetase*. Journal of Biological Chemistry, 2001. **276**(50): p. 47387-47393.
  196. Wyss, P.C., et al., *Novel Dihydrofolate Reductase Inhibitors. Structure-Based versus Diversity-Based Library Design and High-Throughput Synthesis and Screening*. Journal of Medicinal Chemistry, 2003. **46**(12): p. 2304-2312.
  197. OEFNER, C., A. D'ARCY, and F.K. WINKLER, *Crystal structure of human dihydrofolate reductase complexed with folate*. European Journal of Biochemistry, 1988. **174**(2): p. 377-385.
  198. Rosamond, J. and A. Allsop, *Harnessing the Power of the Genome in the Search for New Antibiotics*. Science, 2000. **287**(5460): p. 1973-1976.
  199. Song, J.-H., et al., *Identification of Essential Genes in Streptococcus pneumoniae by Allelic Replacement Mutagenesis*. Mol. Cells, 2005. **19**(3): p. 365-374.
  200. Wilson, T., et al., *Antisense RNA to *ahpC*, an oxidative stress defence gene involved in isoniazid resistance, indicates that *AhpC* of *Mycobacterium bovis* has virulence properties*. Microbiology, 1998. **144**(10): p. 2687-2695.
  201. Payne, D.J., et al., *Bacterial fatty-acid biosynthesis: a genomics-driven target for antibacterial drug discovery*. Drug Discovery Today, 2001. **6**(10): p. 537-544.
  202. Dziadek, J., et al., *Conditional expression of *Mycobacterium smegmatis* *ftsZ*, an essential cell division gene*. Microbiology, 2003. **149**(6): p. 1593-1603.
  203. Ramos, J.L., et al., *The TetR Family of Transcriptional Repressors*. Microbiol. Mol. Biol. Rev., 2005. **69**(2): p. 326-356.
  204. Good, L., *Translation repression by antisense sequences*. Cellular and Molecular Life Sciences (CMLS), 2003. **60**(5): p. 854-861.
  205. Woodford, N., D.W. Wareham, and on behalf of the UK Antibacterial Antisense Study Group, *Tackling antibiotic resistance: a dose of common antisense?* J. Antimicrob. Chemother., 2009. **63**(2): p. 225-229.
  206. Hebert, C.G., J.J. Valdes, and W.E. Bentley, *Beyond silencing -- engineering applications of RNA interference and antisense technology for altering cellular phenotype*. Current Opinion in Biotechnology, 2008. **19**(5): p. 500-505.
  207. Tanya Parish, N.G.S., *Development and use of a conditional antisense mutagenesis system in mycobacteria*. FEMS Microbiology Letters, 1997. **154**(1): p. 151-157.
  208. Baev, D., R. England, and H.K. Kuramitsu, *Stress-Induced Membrane Association of the Streptococcus mutans GTP-Binding Protein, an Essential G Protein, and Investigation of Its Physiological Role by Utilizing an Antisense RNA Strategy*. Infect. Immun., 1999. **67**(9): p. 4510-4516.
  209. Pragman, A.A., Y. Ji, and P.M. Schlievert, *Repression of Staphylococcus aureus SrrAB Using Inducible Antisense *srrA* Alters Growth and Virulence Factor Transcript Levels*. Biochemistry, 2007. **46**(1): p. 314-321.

- 
210. Hett, E.C. and E.J. Rubin, *Bacterial Growth and Cell Division: a Mycobacterial Perspective*. Microbiol. Mol. Biol. Rev., 2008. **72**(1): p. 126-156.
  211. Rothfield, L.I. and S.S. Justice, *Bacterial Cell Division: The Cycle of the Ring*. Cell, 1997. **88**(5): p. 581-584.
  212. Hett, E.C., et al., *A Mycobacterial Enzyme Essential for Cell Division Synergizes with Resuscitation-Promoting Factor*. PLoS Pathog, 2008. **4**(2): p. e1000001.
  213. Chauhan, A., et al., *Interference of Mycobacterium tuberculosis cell division by Rv2719c, a cell wall hydrolase*. Molecular Microbiology, 2006. **62**(1): p. 132-147.
  214. Datta, P., et al., *Interaction between FtsZ and FtsW of Mycobacterium tuberculosis*. Journal of Biological Chemistry, 2002. **277**(28): p. 24983-24987.
  215. Slayden, R.A., D.L. Knudson, and J.T. Belisle, *Identification of cell cycle regulators in Mycobacterium tuberculosis by inhibition of septum formation and global transcriptional analysis*. Microbiology, 2006. **152**(6): p. 1789-1797.
  216. Dong, X., et al., *PlasMapper: a web server for drawing and auto-annotating plasmid maps*. Nucl. Acids Res., 2004. **32**(suppl\_2): p. W660-664.
  217. Prajapati, R.S., T. Ogura, and S.M. Cutting, *Structural and functional studies on an FtsH inhibitor from Bacillus subtilis*. Biochimica et Biophysica Acta (BBA) - General Subjects, 2000. **1475**(3): p. 353-359.
  218. Halder, S., A.B. Datta, and P. Parrack, *Probing the Antiprotease Activity of {lambda}CIII, an Inhibitor of the Escherichia coli Metalloprotease HflB (FtsH)*. J. Bacteriol., 2007. **189**(22): p. 8130-8138.
  219. Cutting, S., et al., *SpoVM, a small protein essential to development in Bacillus subtilis, interacts with the ATP-dependent protease FtsH*. J. Bacteriol., 1997. **179**(17): p. 5534-5542.
  220. Haydon, D.J., et al., *An Inhibitor of FtsZ with Potent and Selective Anti-Staphylococcal Activity*. Science, 2008. **321**(5896): p. 1673-1675.
  221. Wang, J., et al., *Discovery of a Small Molecule That Inhibits Cell Division by Blocking FtsZ, a Novel Therapeutic Target of Antibiotics*. J. Biol. Chem., 2003. **278**(45): p. 44424-44428.
  222. Czaplewski, L.G., et al., *Antibacterial alkoxybenzamide inhibitors of the essential bacterial cell division protein FtsZ*. Bioorganic & Medicinal Chemistry Letters, 2009. **19**(2): p. 524-527.
  223. Leung, A.K.W., et al., *Crystallization of the Mycobacterium tuberculosis cell-division protein FtsZ*. Acta Crystallographica Section D, 2000. **56**(12): p. 1634-1637.
  224. Lasocki, K., et al., *Deletion of the parA (soj) homologue in Pseudomonas aeruginosa causes ParB instability and affects growth rate, chromosome segregation, and motility*. J Bacteriol, 2007. **189**(15): p. 5762-72.
  225. Guo, X.V., et al., *Silencing Essential Protein Secretion in Mycobacterium smegmatis by Using Tetracycline Repressors*. J. Bacteriol., 2007. **189**(13): p. 4614-4623.

- 
226. Bang, W., et al., *AtObgC, a plant ortholog of bacterial Obg, is a chloroplast-targeting GTPase essential for early embryogenesis*. Plant Molecular Biology, 2009. **71**(4): p. 379-390.
227. Maddock, J., et al., *Identification of an essential Caulobacter crescentus gene encoding a member of the Obg family of GTP-binding proteins*. J. Bacteriol., 1997. **179**(20): p. 6426-6431.
228. Zhang, S. and W.G. Haldenwang, *Guanine nucleotides stabilize the binding of Bacillus subtilis Obg to ribosomes*. Biochemical and Biophysical Research Communications, 2004. **322**(2): p. 565-569.
229. Kok, J., K.A. Trach, and J.A. Hoch, *Effects on Bacillus subtilis of a conditional lethal mutation in the essential GTP-binding protein Obg*. J. Bacteriol., 1994. **176**(23): p. 7155-7160.
230. Scott, J.M. and W.G. Haldenwang, *Obg, an Essential GTP Binding Protein of Bacillus subtilis, Is Necessary for Stress Activation of Transcription Factor sigma B*. J. Bacteriol., 1999. **181**(15): p. 4653-4660.
231. Kobayashi, G., S. Moriya, and C. Wada, *Deficiency of essential GTP-binding protein ObgE in Escherichia coli inhibits chromosome partition*. Molecular Microbiology, 2001. **41**(5): p. 1037-1051.
232. Slominska, M., et al., *Impaired chromosome partitioning and synchronization of DNA replication initiation in an insertional mutant in the Vibrio harveyi cgtA gene coding for a common GTP-binding protein*. Biochem. J., 2002. **362**(3): p. 579-584.
233. Yamaichi, Y. and H. Niki, *Active segregation by the Bacillus subtilis partitioning system in Escherichia coli*. Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(26): p. 14656-14661.
234. Casart, Y., et al., *par genes in Mycobacterium bovis and Mycobacterium smegmatis are arranged in an operon transcribed from "SigGC" promoters*. BMC Microbiology, 2008. **8**(1): p. 51.
235. Hayes, F., *The partition system of multidrug resistance plasmid TP228 includes a novel protein that epitomizes an evolutionarily distinct subgroup of the ParA superfamily*. Molecular Microbiology, 2000. **37**(3): p. 528-541.
236. Bartosik, A.A., et al., *ParB of Pseudomonas aeruginosa: interactions with its partner ParA and its target parS and specific effects on bacterial growth*. J Bacteriol, 2004. **186**(20): p. 6983-98.
237. Hayes, F. and D. Barillà, *Assembling the bacterial segrosome*. Trends in Biochemical Sciences, 2006. **31**(5): p. 247-250.
238. Lee, P.S. and A.D. Grossman, *The chromosome partitioning proteins Soj (ParA) and Spo0J (ParB) contribute to accurate chromosome partitioning, separation of replicated sister origins, and regulation of replication initiation in Bacillus subtilis*. Mol Microbiol, 2006. **60**(4): p. 853-69.
239. Bouet, J.-Y. and B.E. Funnell, *P1 ParA interacts with the P1 partition complex at parS and an ATP-ADP switch controls ParA activities*. EMBO J, 1999. **18**(5): p. 1415-1424.
240. Surtees, J.A. and B.E. Funnell, *Plasmid and chromosome traffic control: how ParA and ParB drive partition*. Curr Top Dev Biol, 2003. **56**: p. 145-80.



- 
241. Lewis, R.A., et al., *Chromosome loss from par mutants of Pseudomonas putida depends on growth medium and phase of growth*. Microbiology, 2002. **148**(2): p. 537-548.
242. Mohl, D.A. and J.W. Gober, *Cell Cycle-Dependent Polar Localization of Chromosome Partitioning Proteins in Caulobacter crescentus*. Cell, 1997. **88**(5): p. 675-684.
243. Easter, J., Jr. and J.W. Gober, *ParB-stimulated nucleotide exchange regulates a switch in functionally distinct ParA activities*. Mol Cell, 2002. **10**(2): p. 427-34.
244. Figge, R.M., J. Easter, and J.W. Gober, *Productive interaction between the chromosome partitioning proteins, ParA and ParB, is required for the progression of the cell cycle in Caulobacter crescentus*. Mol Microbiol, 2003. **47**(5): p. 1225-37.
245. Kim, H.-J., et al., *Partitioning of the Linear Chromosome during Sporulation of Streptomyces coelicolor A3(2) Involves an oriC-Linked parAB Locus*. J. Bacteriol., 2000. **182**(5): p. 1313-1320.
246. Bignell, C. and C.M. Thomas, *The bacterial ParA-ParB partitioning proteins*. J Biotechnol, 2001. **91**(1): p. 1-34.
247. Davey, M.J. and B.E. Funnell, *Modulation of the P1 plasmid partition protein ParA by ATP, ADP, and P1 ParB*. J Biol Chem, 1997. **272**(24): p. 15286-92.
248. Thanky, N.R., D.B. Young, and B.D. Robertson, *Unusual features of the cell cycle in mycobacteria: Polar-restricted growth and the snapping-model of cell division*. Tuberculosis, 2007. **87**(3): p. 231-236.
249. Betts, J., et al., *Evaluation of a nutrient starvation model of Mycobacterium tuberculosis persistence by gene and protein expression profiling*. Mol Microbiol, 2002. **43**(3): p. 717 - 731.
250. Donovan, C., et al., *Subcellular Localization and Characterization of the ParAB System from Corynebacterium glutamicum*. J. Bacteriol. **192**(13): p. 3441-3451.
251. Maloney, E., M. Madiraju, and M. Rajagopalan, *Overproduction and localization of Mycobacterium tuberculosis ParA and ParB proteins*. Tuberculosis, 2009. **89**(Supplement 1): p. S65-S69.
252. Ptacin, J.L., et al., *A spindle-like apparatus guides bacterial chromosome segregation*. Nat Cell Biol. **12**(8): p. 791-798.
253. Ireton, K., N.W. Gunther, 4th, and A.D. Grossman, *spo0J is required for normal chromosome segregation as well as the initiation of sporulation in Bacillus subtilis*. J. Bacteriol., 1994. **176**(17): p. 5320-5329.
254. Lee, P.S., et al., *Effects of the Chromosome Partitioning Protein Spo0J (ParB) on oriC Positioning and Replication Initiation in Bacillus subtilis*. J. Bacteriol., 2003. **185**(4): p. 1326-1337.
255. Sullivan, N.L., K.A. Marquis, and D.Z. Rudner, *Recruitment of SMC by ParB-parS Organizes the Origin Region and Promotes Efficient Chromosome Segregation*. Cell, 2009. **137**(4): p. 697-707.
256. Gruber, S. and J. Errington, *Recruitment of Condensin to Replication Origin Regions by ParB/Spo0J Promotes Chromosome Segregation in B. subtilis*. Cell, 2009. **137**(4): p. 685-696.
257. Murray, H. and J. Errington, *Dynamic Control of the DNA Replication Initiation Protein DnaA by Soj/ParA*. Cell, 2008. **135**(1): p. 74-84.

- 
258. Chen, N., et al., *ATP-Induced Shrinkage of DNA with MukB Protein and the MukBEF Complex of Escherichia coli*. J. Bacteriol., 2008. **190**(10): p. 3731-3737.
259. Fennell-Fezzie, R., et al., *The MukF subunit of Escherichia coli condensin: architecture and functional relationship to kleisins*. EMBO J, 2005. **24**(11): p. 1921-1930.
260. Ku, B., et al., *Crystal structure of the MukB hinge domain with coiled-coil stretches and its functional implications*. Proteins: Structure, Function, and Bioinformatics. **78**(6): p. 1483-1490.
261. Hayama, R. and K.J. Mariani, *Physical and functional interaction between the condensin MukB and the decatenase topoisomerase IV in Escherichia coli*. Proceedings of the National Academy of Sciences.
262. Danilova, O., et al., *MukB colocalizes with the oriC region and is required for organization of the two Escherichia coli chromosome arms into separate cell halves*. Molecular Microbiology, 2007. **65**(6): p. 1485-1492.
263. Pribylova, R., P. Kralik, and I. Pavlik, *Oligonucleotide Microarray Technology and its Application to Mycobacterium avium subsp. paratuberculosis Research: A Review*. Molecular Biotechnology, 2009. **42**(1): p. 30-40.
264. Geiman, D.E., et al., *Differential Gene Expression in Response to Exposure to Antimycobacterial Agents and Other Stress Conditions among Seven Mycobacterium tuberculosis whiB-Like Genes*. Antimicrobial Agents and Chemotherapy, 2006. **50**(8): p. 2836-2841.
265. Trias, J., V. Jarlier, and R. Benz, *Porins in the cell wall of mycobacteria*. Science, 1992. **258**(5087): p. 1479-1481.
266. Murphy, D.J. and J.R. Brown, *Novel drug target strategies against Mycobacterium tuberculosis*. Current Opinion in Microbiology, 2008. **11**(5): p. 422-427.
267. Holton, S.J., Weiss, M.S, Tucker, P.A, Wilmanns M, *Structure-based approaches to drug discovery against tuberculosis*. Curr Protein Pept Sci. , 2007. **8**(4): p. 365-375.
268. Zhu, Z. and J. Cuozzo, *High-Throughput Affinity-Based Technologies for Small-Molecule Drug Discovery*. J Biomol Screen, 2009: p. 1087057109350114.
269. Yang, J., L. Zheng, and Y. Ji, *Confirmation of the mode of action of an antibacterial inhibitor using regulated antisense RNA*. World Journal of Microbiology and Biotechnology, 2006. **22**(3): p. 299-303.
270. Yin, D., et al., *Identification of antimicrobial targets using a comprehensive genomic approach*. Pharmacogenomics, 2004. **5**(1): p. 101-113.
271. Singh, S.B., et al., *Isolation, Structure, and Absolute Stereochemistry of Platensimycin, A Broad Spectrum Antibiotic Discovered Using an Antisense Differential Sensitivity Strategy*. Journal of the American Chemical Society, 2006. **128**(36): p. 11916-11920.
272. Chene, P., *ATPases as drug targets: learning from their structure*. Nat Rev Drug Discov, 2002. **1**(9): p. 665-673.
273. ED, B., *Conserved P-loop GTPases of unknown function in bacteria: an emerging and vital ensemble in bacterial physiology*. Biochem Cell Biol., 2005. **83**(6): p. 738-746.

- 
274. Lee, D.C. and Z. Jia, *Emerging structural insights into bacterial tyrosine kinases*. Trends in Biochemical Sciences, 2009. **34**(7): p. 351-357.
275. Vetter IR, W.A., *Nucleoside triphosphate-binding proteins: different scaffolds to achieve phosphoryl transfer*. *Q Rev Biophys*, 1999. **32**(1): p. 1-56.
276. Taylor, S.S. and E. Radzio-Andzelm, *Protein kinase inhibition: natural and synthetic variations on a theme*. Current Opinion in Chemical Biology, 1997. **1**(2): p. 219-226.
277. Burlison, J.A., et al., *Novobiocin: Redesigning a DNA Gyrase Inhibitor for Selective Inhibition of Hsp90*. Journal of the American Chemical Society, 2006. **128**(48): p. 15529-15536.
278. Corbett, K.D. and J.M. Berger, *Structural basis for topoisomerase VI inhibition by the anti-Hsp90 drug radicicol*. Nucl. Acids Res., 2006. **34**(15): p. 4269-4277.
279. Cohen, P., *Protein kinases - the major drug targets of the twenty-first century?* Nat Rev Drug Discov, 2002. **1**(4): p. 309-315.
280. Jr, M.A.I. and G. Kizelshteyn, *Treatment of Complex Regional Pain Syndrome Type I With Oral Phenoxybenzamine: Rationale and Case Reports*. Pain Practice, 2008. **8**(2): p. 125-132.
281. Meshnick, S.R., *Artemisinin: mechanisms of action, resistance and toxicity*. International Journal for Parasitology, 2002. **32**(13): p. 1655-1660.
282. Davis, M.A., et al., *The P1 ParA protein and its ATPase activity play a direct role in the segregation of plasmid copies to daughter cells*. Mol Microbiol, 1996. **21**(5): p. 1029-36.
283. Hunding, A., G. Ebersbach, and K. Gerdes, *A mechanism for ParB-dependent waves of ParA, a protein related to DNA segregation during cell division in prokaryotes*. J Mol Biol, 2003. **329**(1): p. 35-43.
284. Norris, V., et al., *DNA replication termination in Escherichia coli parB (a dnaG allele), parA, and gyrB mutants affected in DNA distribution*. J Bacteriol, 1986. **168**(2): p. 494-504.
285. Ebersbach, G., et al., *Regular cellular distribution of plasmids by oscillating and filament-forming ParA ATPase of plasmid pB171*. Mol Microbiol, 2006. **61**(6): p. 1428-42.
286. Black, M.T. and J. Hodgson, *Novel target sites in bacteria for overcoming antibiotic resistance*. Advanced Drug Delivery Reviews, 2005. **57**(10): p. 1528-1538.
287. Galperin, M.Y. and E.V. Koonin, *Searching for drug targets in microbial genomes*. Current Opinion in Biotechnology, 1999. **10**(6): p. 571-578.
288. Benziane, B., et al., *AMP-activated protein kinase activator A-769662 is an inhibitor of the Na<sup>+</sup>,K<sup>+</sup>-ATPase*. Am J Physiol Cell Physiol, 2009: p. ajpcell.00010.2009.
289. Sobota, J.A., et al., *Inhibitors of the V0 subunit of the vacuolar H<sup>+</sup>-ATPase prevent segregation of lysosomal- and secretory-pathway proteins*. J Cell Sci, 2009. **122**(19): p. 3542-3553.
290. Wigle, T.J. and S.F. Singleton, *Directed molecular screening for RecA ATPase inhibitors*. Bioorganic & Medicinal Chemistry Letters, 2007. **17**(12): p. 3249-3253.

- 
291. Flatman, R.H., et al., *Simocyclinone D8, an Inhibitor of DNA Gyrase with a Novel Mode of Action*. Antimicrob. Agents Chemother., 2005. **49**(3): p. 1093-1100.
292. Drlica, K., *Mechanism of fluoroquinolone action*. Current Opinion in Microbiology, 1999. **2**(5): p. 504-508.
293. Whitty, A. and G. Kumaravel, *Between a rock and a hard place?* Nat Chem Biol, 2006. **2**(3): p. 112-118.
294. Coma, I., et al., *Process Validation and Screen Reproducibility in High-Throughput Screening*. J Biomol Screen, 2009. **14**(1): p. 66-76.
295. Goodnow, J.R.A., *Hit and lead identification: Integrated technology-based approaches*. Drug Discovery Today: Technologies, 2006. **3**(4): p. 367-375.
296. MFM Engels, P.V., *Smart screening: Approaches to efficient HTS*. Current Opinion in Drug Discovery & Development, 2001. **4**: p. 275-283.
297. Tong, A.H.Y., et al., *Global Mapping of the Yeast Genetic Interaction Network*. Science, 2004. **303**(5659): p. 808-813.
298. Bandow, J.E., et al., *Proteomic Approach to Understanding Antibiotic Action*. Antimicrob. Agents Chemother., 2003. **47**(3): p. 948-955.
299. Ng, W.-L., et al., *Transcriptional Regulation and Signature Patterns Revealed by Microarray Analyses of Streptococcus pneumoniae R6 Challenged with Sublethal Concentrations of Translation Inhibitors*. J. Bacteriol., 2003. **185**(1): p. 359-370.
300. Utaida, S., et al., *Genome-wide transcriptional profiling of the response of Staphylococcus aureus to cell-wall-active antibiotics reveals a cell-wall-stress stimulon*. Microbiology, 2003. **149**(10): p. 2719-2732.
301. Ren, D., et al., *Differential Gene Expression To Investigate the Effect of (5Z)-4-Bromo- 5-(Bromomethylene)-3-Butyl-2(5H)-Furanone on Bacillus subtilis*. Appl. Environ. Microbiol., 2004. **70**(8): p. 4941-4949.
302. Qiu, J., et al., *Global gene expression profile of Yersinia pestis induced by streptomycin*. FEMS Microbiology Letters, 2005. **243**(2): p. 489-496.
303. Dixon, S.J. and B.R. Stockwell, *Identifying druggable disease-modifying gene products*. Current Opinion in Chemical Biology, 2009. **In Press, Corrected Proof**.
304. Kalscheuer, R., et al., *Self-poisoning of Mycobacterium tuberculosis by targeting GlgE in an [alpha]-glucan pathway*. Nat Chem Biol, 2010. **6**(5): p. 376-384.
305. Hugonnet, J.-E., et al., *Meropenem-Clavulanate Is Effective Against Extensively Drug-Resistant Mycobacterium tuberculosis*. Science, 2009. **323**(5918): p. 1215-1218.
306. Habig, M., et al., *Efficient Elimination of Nonstoichiometric Enzyme Inhibitors from HTS Hit Lists*. J Biomol Screen, 2009: p. 1087057109336586.
307. Rishton, G.M., *Reactive compounds and in vitro false positives in HTS*. Drug Discovery Today, 1997. **2**(9): p. 382-384.
308. McGovern, S.L., et al., *A Common Mechanism Underlying Promiscuous Inhibitors from Virtual and High-Throughput Screening*. Journal of Medicinal Chemistry, 2002. **45**(8): p. 1712-1722.
309. McGovern, S.L., et al., *A Specific Mechanism of Nonspecific Inhibition*. Journal of Medicinal Chemistry, 2003. **46**(20): p. 4265-4272.

- 
310. Denny, W.A. and L.P.G. Wakelin, *Kinetic and Equilibrium Studies of the Interaction of Amsacrine and Anilino Ring-substituted Analogues with DNA*. Cancer Res, 1986. **46**(4\_Part\_1): p. 1717-1721.
311. Peh, H.L., et al., *In Vitro Activities of Mitomycin C against Growing and Hypoxic Dormant Tubercle Bacilli*. Antimicrob. Agents Chemother., 2001. **45**(8): p. 2403-2404.
312. Blanchard, J.S., *Molecular Mechanisms of Drug Resistance in Mycobacterium Tuberculosis*. Annual Review of Biochemistry, 1996. **65**(1): p. 215-239.
313. Ashok Rattan, A.K., and Nishat Ahmad, *Multidrug-Resistant Mycobacterium tuberculosis: Molecular Perspectives*. Emerging Infectious Diseases, 1998. **4**(2): p. 195-209.
314. Zhang, Y. and D. Mitchison, *The curious characteristics of pyrazinamide: a review*. The International Journal of Tuberculosis and Lung Disease, 2003. **7**: p. 6-21.
315. Zhang, Y., et al., *Role of Acid pH and Deficient Efflux of Pyrazinoic Acid in Unique Susceptibility of Mycobacterium tuberculosis to Pyrazinamide*. J. Bacteriol., 1999. **181**(7): p. 2044-2049.
316. Sidoli, C., et al., *Biochemical assay development for drug discovery: a sequential optimization from protein expression to enzymatic activity*. Microbial Cell Factories, 2006. **5**(Suppl 1): p. P58.
317. Cleland, W.W., *Enzyme Kinetics*. Annual Review of Biochemistry, 2003. **36**(1): p. 77-112.
318. WILKINSON, J.H., *Enzyme kinetics and its relevance to enzyme assay*. J. clin. Path. , 1971. **4**: p. 14-21.
319. *Yakkyoku. Pharmacy*. 1987. **38**: p. 1493.
320. *Drug Dosages in Laboratory Animals - A Handbook*, Rev. ed. 1973: p. 195.
321. *Collection of Czechoslovak Chemical Communications*. 1983. **48**: p. 144.
322. Leonard, T.A., P.J. Butler, and J. Lowe, *Bacterial chromosome segregation: structure and DNA binding of the Soj dimer [mdash] a conserved biological switch*. EMBO J, 2005. **24**(2): p. 270-282.
323. Korepanova, A., et al., *Expression of membrane proteins from Mycobacterium tuberculosis in Escherichia coli as fusions with maltose binding protein*. Protein Expression and Purification, 2007. **53**(1): p. 24-30.
324. Bashiri, G., et al., *Expression, purification and crystallization of native and selenomethionine labeled Mycobacterium tuberculosis FGD1 (Rv0407) using a Mycobacterium smegmatis expression system*. Protein Expression and Purification, 2007. **54**(1): p. 38-44.
325. Bondos, S.E. and A. Bicknell, *Detection and prevention of protein aggregation before, during, and after purification*. Analytical Biochemistry, 2003. **316**(2): p. 223-231.
326. Gekko, K. and S.N. Timasheff, *Mechanism of protein stabilization by glycerol: preferential hydration in glycerol-water mixtures*. Biochemistry, 1981. **20**(16): p. 4667-4676.
327. Mitchell, M.S. and V.B. Rao, *Novel and deviant Walker A ATP-binding motifs in bacteriophage large terminase-DNA packaging proteins*. Virology, 2004. **321**(2): p. 217-221.

- 
328. Pezza, R.J., et al., *Vanadate inhibits the ATPase activity and DNA binding capability of bacterial MutS. A structural model for the vanadate-MutS interaction at the Walker A motif*. Nucl. Acids Res., 2002. **30**(21): p. 4700-4708.
329. Cremo, C.R., G.T. Long, and J.C. Grammer, *Photocleavage of myosin subfragment 1 by vanadate*. Biochemistry, 1990. **29**(34): p. 7982-7990.
330. Percival, M.D., K. Doherty, and M.J. Gresser, *Inhibition of phosphoglucomutase by vanadate*. Biochemistry, 1990. **29**(11): p. 2764-2769.
331. Murphy, M.B., C. Murray, and G.D. Shorten, *Fenoldopam -- A Selective Peripheral Dopamine-Receptor Agonist for the Treatment of Severe Hypertension*. N Engl J Med, 2001. **345**(21): p. 1548-1557.
332. Taldone, T., W. Sun, and G. Chiosis, *Discovery and development of heat shock protein 90 inhibitors*. Bioorganic & Medicinal Chemistry, 2009. **17**(6): p. 2225-2235.
333. Guzzetta, N.A., *Phenoxybenzamine in the Treatment of Hypoplastic Left Heart Syndrome: A Core Review*. Anesth Analg, 2007. **105**(2): p. 312-315.
334. CAINE, M., S. PERLBERG, and S. MERETYK, *A Placebo-controlled Double-blind Study of the Effect of Phenoxybenzamine in Benign Prostatic Obstruction*. British Journal of Urology, 1978. **50**(7): p. 551-554.
335. Swinney, D.C., *Biochemical mechanisms of drug action: what does it take for success?* Nat Rev Drug Discov, 2004. **3**(9): p. 801-808.
336. Pauwels, P.J., *5-HT<sub>1B/D</sub> receptor antagonists*. General Pharmacology: The Vascular System, 1997. **29**(3): p. 293-303.
337. Zhukovskaya, N.L. and J.F. Neumaier, *Clozapine downregulates 5-hydroxytryptamine<sub>6</sub> (5-HT<sub>6</sub>) and upregulates 5-HT<sub>7</sub> receptors in HeLa cells*. Neuroscience Letters, 2000. **288**(3): p. 236-240.
338. Newman-Tancredi, A., et al., *Agonist and antagonist actions of antipsychotic agents at 5-HT<sub>1A</sub> receptors: a [<sup>3</sup>H]GTP[ $\gamma$ S] binding study*. European Journal of Pharmacology, 1998. **355**(2-3): p. 245-256.
339. Metysová J, M.J., Dlabac A, Kazdová E, Valchár M., *Pharmacological properties of a potent neuroleptic drug octoclotheptin*. Acta Biol Med Ger, 1980. **39**(6): p. 723-740.
340. Wang, S., et al., *Tools for target identification and validation*. Current Opinion in Chemical Biology, 2004. **8**(4): p. 371-377.
341. Liu, M., et al., *Conserved Fungal Genes as Potential Targets for Broad-Spectrum Antifungal Drug Discovery*. Eukaryotic Cell, 2006. **5**(4): p. 638-649.
342. Kana, B.D. and V. Mizrahi, *Molecular genetics of Mycobacterium tuberculosis in relation to the discovery of novel drugs and vaccines*. Tuberculosis, 2004. **84**(1-2): p. 63-75.
343. Lock, R.L. and E.J. Harry, *Cell-division inhibitors: new insights for future antibiotics*. Nat Rev Drug Discov, 2008. **7**(4): p. 324-338.
344. Sneller, M.C., *Granuloma formation, implications for the pathogenesis of vasculitis*. Cleveland Clinic Journal of Medicine, 2002. **69**(Suppl 2): p. SII40.
345. Owens, J., *Determining druggability*. Nat Rev Drug Discov, 2007. **6**(3): p. 187-187.

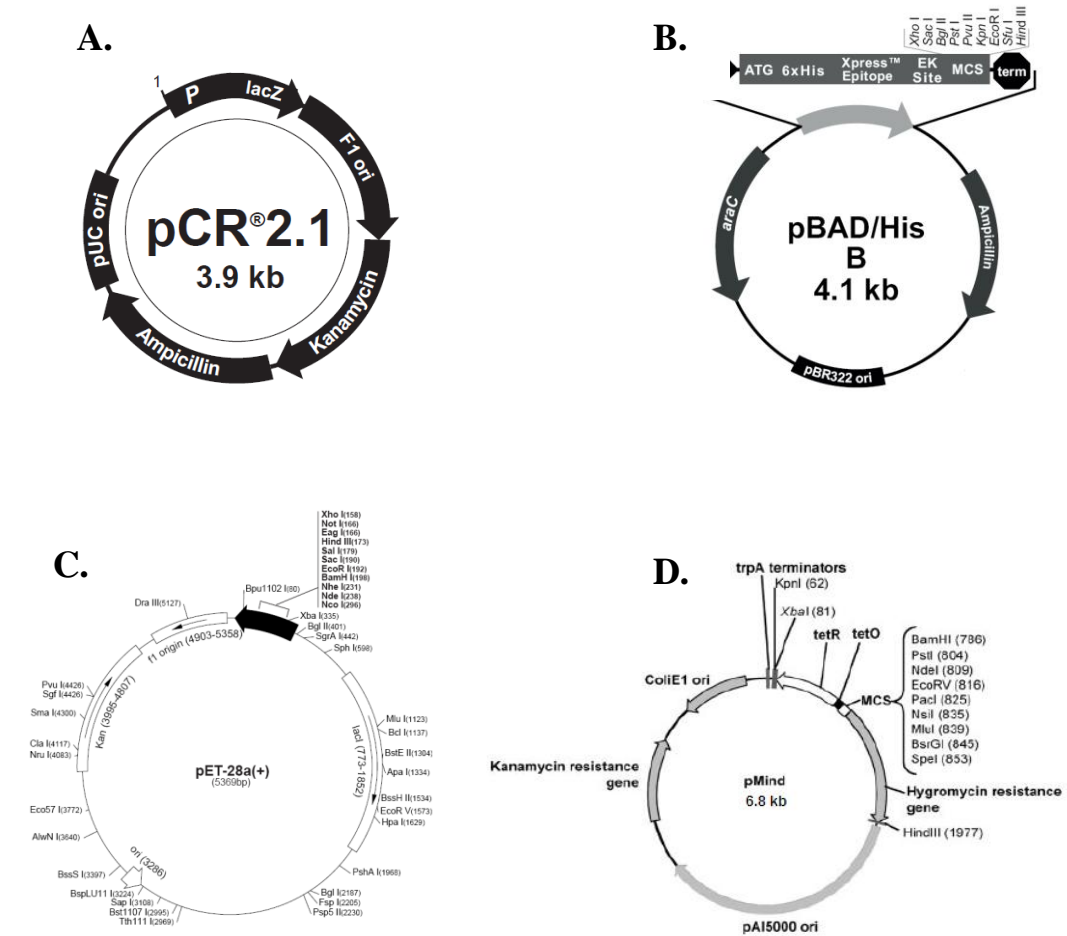
- 
346. Asai, T., et al., *An Escherichia coli strain with all chromosomal rRNA operons inactivated: Complete exchange of rRNA genes between bacteria*. Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(5): p. 1971-1976.
  347. Vicens, Q. and E. Westhof, *RNA as a Drug Target: The Case of Aminoglycosides*. ChemBioChem, 2003. **4**(10): p. 1018-1023.
  348. Chico, L.K., L.J. Van Eldik, and D.M. Watterson, *Targeting protein kinases in central nervous system disorders*. Nat Rev Drug Discov, 2009. **8**(11): p. 892-909.
  349. Hajduk, P.J., J.R. Huth, and S.W. Fesik, *Druggability Indices for Protein Targets Derived from NMR-Based Screening Data*. Journal of Medicinal Chemistry, 2005. **48**(7): p. 2518-2525.
  350. Nadendla, R., *Molecular modeling: A powerful tool for drug design and molecular docking*. Resonance, 2004. **9**(5): p. 51-60.
  351. Kortagere, S., M.D. Krasowski, and S. Ekins, *The importance of discerning shape in molecular pharmacology*. Trends in Pharmacological Sciences, 2009. **30**(3): p. 138-147.
  352. Davies, J.W., M. Glick, and J.L. Jenkins, *Streamlining lead discovery by aligning in silico and high-throughput screening*. Current Opinion in Chemical Biology, 2006. **10**(4): p. 343-351.

## **CHAPTER NINE**

### **APPENDICES**



## 9.1 Maps of plasmids used in this study



**Figure 9.1: Vectors used in this study**

A) pCR®2.1 TA cloning vector used for direct ligation of PCR fragments, B) pBAD/HisB expression vector used for over-expressing proteins in *E. coli* LMG194, C) pET28a(+) expression vector used for over-expressing proteins in *E. coli* BL21 (DE3) and D) tetracycline inducible vector used for conditional gene expression in *M. smegmatis*.

## 9.2 Essential Genes

### 9.2.1 List of Essential Genes published by Sassetti *et al.*

**Table 9.1: List of Essential Genes published by Sassetti *et al.***

List of 614 genes identified as essential for the growth and viability of *M. tuberculosis* H37Rv via high density transposon mutagenesis under optimal conditions by Sassetti *et al.*, 2003 [126].

Rv designator	Rv designator	Rv designator	Rv designator	Rv designator	Rv designator	Rv designator	Rv designator	Rv designator
Rv0001	Rv0527	Rv0900	Rv1310	Rv1630	Rv2196	Rv2539c	Rv2992c	Rv3458c
Rv0005	Rv0528	Rv0946c	Rv1311	Rv1631	Rv2200c	Rv2540c	Rv2995c	Rv3459c
Rv0014c	Rv0529	Rv0949	Rv1315	Rv1641	Rv2201	Rv2552c	Rv2996c	Rv3462c
Rv0015c	Rv0540	Rv0951	Rv1326c	Rv1649	Rv2210c	Rv2554c	Rv2999	Rv3464
Rv0041	Rv0541c	Rv0952	Rv1327c	Rv1650	Rv2211c	Rv2555c	Rv3001c	Rv3465
Rv0058	Rv0542c	Rv0955	Rv1339	Rv1652	Rv2215	Rv2580c	Rv3002c	Rv3490
Rv0060	Rv0553	Rv0957	Rv1342c	Rv1653	Rv2217	Rv2582	Rv3003c	Rv3538
Rv0085	Rv0555	Rv0973c	Rv1347c	Rv1654	Rv2218	Rv2611c	Rv3009c	Rv3559c
Rv0086	Rv0556	Rv0982	Rv1348	Rv1655	Rv2219	Rv2612c	Rv3011c	Rv3564
Rv0102	Rv0557	Rv0993	Rv1349	Rv1656	Rv2220	Rv2613c	Rv3014c	Rv3579c
Rv0112	Rv0558	Rv1002c	Rv1350	Rv1658	Rv2221c	Rv2614c	Rv3018c	Rv3580c
Rv0118c	Rv0562	Rv1005c	Rv1380	Rv1659	Rv2225	Rv2623	Rv3025c	Rv3581c
Rv0127	Rv0588	Rv1007c	Rv1381	Rv1661	Rv2229c	Rv2673	Rv3029c	Rv3582c
Rv0189c	Rv0607	Rv1011	Rv1382	Rv1662	Rv2231c	Rv2682c	Rv3030	Rv3593
Rv0205	Rv0627	Rv1014c	Rv1383	Rv1663	Rv2235	Rv2697c	Rv3031	Rv3596c
Rv0208c	Rv0628c	Rv1017c	Rv1388	Rv1683	Rv2242	Rv2698	Rv3032	Rv3597c
Rv0224c	Rv0635	Rv1018c	Rv1389	Rv1689	Rv2244	Rv2700	Rv3042c	Rv3598c
Rv0226c	Rv0636	Rv1023	Rv1390	Rv1695	Rv2245	Rv2702	Rv3043c	Rv3602c
Rv0227c	Rv0638	Rv1024	Rv1391	Rv1697	Rv2246	Rv2703	Rv3048c	Rv3604c
Rv0228	Rv0640	Rv1025	Rv1392	Rv1699	Rv2247	Rv2710	Rv3051c	Rv3607c
Rv0236c	Rv0651	Rv1026	Rv1406	Rv1701	Rv2251	Rv2726c	Rv3101c	Rv3608c
Rv0282	Rv0652	Rv1027c	Rv1409	Rv1708	Rv2260	Rv2727c	Rv3105c	Rv3609c
Rv0283	Rv0666	Rv1092c	Rv1415	Rv1711	Rv2268c	Rv2739c	Rv3109	Rv3610c
Rv0284	Rv0667	Rv1094	Rv1416	Rv1712	Rv2283	Rv2746c	Rv3111	Rv3611
Rv0285	Rv0668	Rv1098c	Rv1420	Rv1713	Rv2319c	Rv2747	Rv3112	Rv3625c
Rv0286	Rv0674	Rv1122	Rv1436	Rv1714	Rv2325c	Rv2748c	Rv3113	Rv3627c
Rv0289	Rv0675	Rv1126c	Rv1437	Rv1715	Rv2338c	Rv2753c	Rv3131	Rv3634c
Rv0290	Rv0682	Rv1128c	Rv1438	Rv1730c	Rv2343c	Rv2754c	Rv3132c	Rv3635
Rv0291	Rv0683	Rv1133c	Rv1446c	Rv1738	Rv2357c	Rv2786c	Rv3135	Rv3646c
Rv0292	Rv0684	Rv1166	Rv1448c	Rv1753c	Rv2361c	Rv2812	Rv3137	Rv3648c
Rv0311	Rv0685	Rv1169c	Rv1449c	Rv1828	Rv2364c	Rv2813	Rv3139	Rv3658c
Rv0312	Rv0697	Rv1177	Rv1456c	Rv1832	Rv2365c	Rv2817c	Rv3140	Rv3660c
Rv0334	Rv0700	Rv1182	Rv1457c	Rv1850	Rv2373c	Rv2827c	Rv3177	Rv3666c
Rv0335c	Rv0701	Rv1186c	Rv1459c	Rv1963c	Rv2378c	Rv2828c	Rv3198c	Rv3669
Rv0337c	Rv0702	Rv1187	Rv1461	Rv2013	Rv2382c	Rv2830c	Rv3201c	Rv3679
Rv0338c	Rv0703	Rv1188	Rv1462	Rv2026c	Rv2386c	Rv2832c	Rv3240c	Rv3708c
Rv0347	Rv0704	Rv1193	Rv1463	Rv2041c	Rv2391	Rv2833c	Rv3243c	Rv3709c
Rv0350	Rv0705	Rv1194c	Rv1464	Rv2050	Rv2392	Rv2837c	Rv3244c	Rv3710
Rv0351	Rv0706	Rv1201c	Rv1465	Rv2093c	Rv2397c	Rv2839c	Rv3245c	Rv3712
Rv0352	Rv0707	Rv1202	Rv1466	Rv2109c	Rv2398c	Rv2841c	Rv3248c	Rv3713
Rv0357c	Rv0708	Rv1208	Rv1475c	Rv2110c	Rv2399c	Rv2845c	Rv3255c	Rv3721c
Rv0372c	Rv0709	Rv1215c	Rv1476	Rv2111c	Rv2400c	Rv2846c	Rv3257c	Rv3722c
Rv0383c	Rv0710	Rv1223	Rv1477	Rv2121c	Rv2412	Rv2855	Rv3264c	Rv3752c
Rv0384c	Rv0714	Rv1229c	Rv1479	Rv2122c	Rv2418c	Rv2856	Rv3265c	Rv3778c
Rv0399c	Rv0715	Rv1232c	Rv1480	Rv2130c	Rv2421c	Rv2869c	Rv3267	Rv3781
Rv0400c	Rv0716	Rv1248c	Rv1485	Rv2138	Rv2437	Rv2881c	Rv3271c	Rv3782
Rv0404	Rv0718	Rv1254	Rv1490	Rv2145c	Rv2438c	Rv2882c	Rv3275c	Rv3790
Rv0410c	Rv0719	Rv1274	Rv1539	Rv2150c	Rv2439c	Rv2883c	Rv3281	Rv3791
Rv0411c	Rv0720	Rv1278	Rv1547	Rv2151c	Rv2440c	Rv2889c	Rv3285	Rv3792
Rv0412c	Rv0721	Rv1284	Rv1594	Rv2152c	Rv2441c	Rv2890c	Rv3332	Rv3793
Rv0415	Rv0722	Rv1285	Rv1595	Rv2153c	Rv2442c	Rv2897c	Rv3336c	Rv3795
Rv0416	Rv0732	Rv1286	Rv1599	Rv2154c	Rv2444c	Rv2903c	Rv3341	Rv3799c
Rv0417	Rv0736	Rv1292	Rv1600	Rv2155c	Rv2447c	Rv2904c	Rv3343c	Rv3800c
Rv0422c	Rv0744c	Rv1293	Rv1601	Rv2156c	Rv2448c	Rv2906c	Rv3356c	Rv3801c
Rv0423c	Rv0755c	Rv1294	Rv1602	Rv2157c	Rv2457c	Rv2907c	Rv3372	Rv3802c
Rv0425c	Rv0772	Rv1295	Rv1603	Rv2163c	Rv2460c	Rv2908c	Rv3396c	Rv3805c
Rv0429c	Rv0780	Rv1296	Rv1605	Rv2164c	Rv2477c	Rv2916c	Rv3398c	Rv3806c
Rv0430	Rv0781	Rv1297	Rv1606	Rv2165c	Rv2500c	Rv2925c	Rv3411c	Rv3808c
Rv0440	Rv0788	Rv1298	Rv1609	Rv2166c	Rv2507	Rv2926c	Rv3417c	Rv3834c
Rv0450c	Rv0798c	Rv1299	Rv1610	Rv2174	Rv2510c	Rv2927c	Rv3418c	Rv3858c
Rv0462	Rv0803	Rv1300	Rv1611	Rv2178c	Rv2511	Rv2967c	Rv3422c	Rv3859c

---

Rv0479c	Rv0808	Rv1301	Rv1612	Rv2179c	Rv2515c	Rv2968c	Rv3423c	Rv3907c
Rv0500	Rv0811c	Rv1303	Rv1613	Rv2182c	Rv2516c	Rv2969c	Rv3427c	Rv3913
Rv0509	Rv0817c	Rv1305	Rv1614	Rv2188c	Rv2524c	Rv2971	Rv3436c	Rv3915
Rv0510	Rv0824c	Rv1306	Rv1617	Rv2192c	Rv2533c	Rv2977c	Rv3441c	Rv3917c
Rv0511	Rv0875c	Rv1307	Rv1622c	Rv2193	Rv2534c	Rv2981c	Rv3443c	Rv3918c
Rv0524	Rv0883c	Rv1308	Rv1626	Rv2194	Rv2537c	Rv2986c	Rv3455c	Rv3921c
Rv0525	Rv0884c	Rv1309	Rv1629	Rv2195	Rv2538c	Rv2987c	Rv3457c	Rv3923c
Rv0526	Rv0894							

---

## 9.2.2 Conserved Essential Genes

Supplementary Table S1

Rv annotation	MSMEG annotation	Species conserved in	%ID; MTB vs MSMEG	Mycobacterium tuberculosis functional annotation
Rv0001	MSMEG_6947	<i>Bacillus subtilis</i>	77	chromosomal replication initiation protein (dnaA)
Rv0005	MSMEG_0005	<i>Bacillus subtilis</i>	80	DNA topoisomerase IV subunit B (gyrB)
Rv0014c	MSMEG_0028	<i>Streptococcus pneumoniae</i>	65	Transmembrane serine/threonine-protein kinase B (pknB)
Rv0015c	MSMEG_0030	<i>Streptococcus pneumoniae</i>	87	Transmembrane serine/threonine-protein kinase A (pknA)
Rv0041	MSMEG_6917	<i>Bacillus subtilis</i>	76	leucyl-tRNA synthetase (leuS)
Rv0058	MSMEG_6892	<i>Bacillus subtilis</i>	66	replicative DNA helicase (dnaB)
Rv0284	MSMEG_0617	<i>Escherichia coli</i>	72	Possible conserved membrane protein (FtsK)
Rv0352	MSMEG_0711	<i>Staphylococcus aureus</i>	71	Probable chaperone protein (dnaJ1)
Rv0422c	MSMEG_0825	<i>Haemophilus influenzae</i>	60	phosphomethylpyrimidine kinase (thiD)
Rv0440	MSMEG_0880	<i>Bacillus subtilis</i>	84	chaperonin (groEL)
Rv0462	MSMEG_0903	<i>Staphylococcus aureus</i>	81	dihydrolipoamide dehydrogenase (lpd)
Rv0500	MSMEG_0943	<i>Escherichia coli</i>	65	pyrroline-5-carboxylate reductase (proC)
Rv0510	MSMEG_0953	<i>Escherichia coli</i>	77	prophobilinogen deaminase (hemC)
Rv0524	MSMEG_0969	<i>Escherichia coli</i>	71	glutamate-1-semialdehyde aminotransferase (hemL)
Rv0526	MSMEG_0971	<i>Bacillus subtilis</i>	70	Possible thioredoxin protein
Rv0527	MSMEG_0972	<i>Helicobacter pylori</i>	61	Possible cytochrome C-type biogenesis protein (ccdA)
Rv0529	MSMEG_0974	<i>Bacillus subtilis</i>	66	Possible cytochrome C-type biogenesis protein (ccsA)
Rv0542c	MSMEG_1062	<i>Bacillus subtilis</i>	60	O-succinylbenzoic acid--CoA ligase (menE)
Rv0555	MSMEG_1109	<i>Bacillus subtilis</i>	75	2-oxoglutarate decarboxylase (menD)
Rv0558	MSMEG_1115	<i>Escherichia coli</i>	82	ubiquinone/menaquinone biosynthesis methyltransferase (menH)
Rv0562	MSMEG_1133	<i>Bacillus subtilis</i>	67	Probable polyprenyl-diphosphate synthase GRCC1 (grcC1)
Rv0640	MSMEG_1346	<i>Staphylococcus aureus</i>	85	50S ribosomal protein L11 (rplK)
Rv0651	MSMEG_1364	<i>Staphylococcus aureus</i>	76	50S ribosomal protein L10 (rplJ)
Rv0667	MSMEG_1367	<i>Bacillus subtilis</i>	89	DNA-directed RNA polymerase beta subunit (rpoB)
Rv0675	MSMEG_1390	<i>Staphylococcus aureus</i>	75	enoyl-CoA hydratase (echA5)
Rv0682	MSMEG_1398	<i>Helicobacter pylori</i>	99	30S ribosomal protein S12 (rpsL)

---

Rv0683	MSMEG_1399	<i>Bacillus subtilis</i>	96	30S ribosomal protein S7 (rpsG)
Rv0684	MSMEG_6535	<i>Staphylococcus aureus</i>	37	elongation factor EF-2 (fusA1)
Rv0685	MSMEG_1401	<i>Escherichia coli</i>	83	elongation factor Tu (tuf)
Rv0700	MSMEG_1435	<i>Bacillus subtilis</i>	98	30S ribosomal protein S10 (rpsJ)
Rv0701	MSMEG_1436	<i>Escherichia coli</i>	88	50S ribosomal protein L3 (rplC)
Rv0702	MSMEG_1437	<i>Staphylococcus aureus</i>	82	50S ribosomal protein L4 (rplD)
Rv0704	MSMEG_1439	<i>Staphylococcus aureus</i>	92	50S ribosomal protein L2 (rplB)
Rv0705	MSMEG_1440	<i>Staphylococcus aureus</i>	97	30S ribosomal protein S19 (rpsS)
Rv0707	MSMEG_1442	<i>Haemophilus influenzae</i>	97	30S ribosomal protein S3 (rpsC)
Rv0708	MSMEG_1443	<i>Bacillus subtilis</i>	93	50S ribosomal protein L16 (rplP)
Rv0710	MSMEG_1445	<i>Bacillus subtilis</i>	93	30S ribosomal protein S17 (rpsQ)
Rv0714	MSMEG_1465	<i>Staphylococcus aureus</i>	80	50S ribosomal protein L14 (rplN)
Rv0715	MSMEG_1466	<i>Bacillus subtilis</i>	87	50S ribosomal protein L24 (rplX)
Rv0716	MSMEG_1467	<i>Bacillus subtilis</i>	83	50S ribosomal protein L5 (rplE)
Rv0718	MSMEG_1469	<i>Bacillus subtilis</i>	96	30S ribosomal protein S8 (rpsH)
Rv0719	MSMEG_1470	<i>Staphylococcus aureus</i>	87	50S ribosomal protein L6 (rplF)
Rv0720	MSMEG_1471	<i>Bacillus subtilis</i>	88	50S ribosomal protein L18 (rplR)
Rv0721	MSMEG_1472	<i>Salmonella typhimurium</i>	76	30S ribosomal protein S5 (rpsE)
Rv0732	MSMEG_1483	<i>Escherichia coli</i>	82	preprotein translocase (secY)
Rv0772	MSMEG_5852	<i>Haemophilus influenzae</i>	73	phosphoribosylamine--glycine ligase (purD)
Rv0780	MSMEG_5841	<i>Haemophilus influenzae</i>	77	phosphoribosylaminoimidazole-succinocarboxamide synthase (hemH)
Rv0808	MSMEG_5800	<i>Streptococcus pneumoniae</i>	84	amidophosphoribosyltransferase (purF)
Rv0973c	MSMEG_5493	<i>Bacillus subtilis</i>	64	Probable acetyl-/propionyl-coenzyme A carboxylase alpha chain (accA2)
Rv0993	MSMEG_5471	<i>Escherichia coli</i>	75	Probable UTP-glucose-1-phosphate uridylyltransferase (galU)
Rv1007c	MSMEG_5441	<i>Bacillus subtilis</i>	70	methionine--tRNA ligase (metG)
Rv1014c	MSMEG_5432	<i>Haemophilus influenzae</i>	77	peptidyl-tRNA hydrolase (pth)
Rv1017c	MSMEG_5427	<i>Bacillus subtilis</i>	87	ribose-phosphate pyrophosphokinase (prsA)
Rv1018c	MSMEG_5426	<i>Streptococcus pneumoniae</i>	68	Probable UDP-N-acetylglucosamine pyrophosphorylase (glmU)
Rv1023	MSMEG_5415	<i>Bacillus subtilis</i>	82	phosphopyruvate hydratase (eno)
Rv1092c	MSMEG_5252	<i>Escherichia coli</i>	83	pantothenate kinase (coaA)
Rv1098c	MSMEG_5240	<i>Haemophilus influenzae</i>	82	fumarate hydratase (fumC)
Rv1193	MSMEG_5114	<i>Bacillus subtilis</i>	63	acyl-CoA synthase (fadD36)

---

Rv1223	MSMEG_5070	<i>Helicobacter pylori</i>	71	Probable serine protease (htrA)
Rv1229c	MSMEG_5068	<i>Salmonella typhimurium</i>	73	Probable mrp-related protein (mrp)
Rv1292	MSMEG_4959	<i>Bacillus subtilis</i>	81	arginyl-tRNA synthetase (argS)
Rv1293	MSMEG_4958	<i>Helicobacter pylori</i>	80	Probable diaminopimelate decarboxylase (lysA)
Rv1297	MSMEG_4954	<i>Staphylococcus aureus</i>	91	transcription termination factor Rho (rho)
Rv1298	MSMEG_4951	<i>Haemophilus influenzae</i>	80	50S ribosomal protein L31 (rpmE)
Rv1299	MSMEG_4950	<i>Bacillus subtilis</i>	76	peptide chain release factor 1 (prfA)
Rv1308	MSMEG_4938	<i>Streptococcus pneumoniae</i>	78	ATP synthase subunit A (atpA)
Rv1309	MSMEG_4937	<i>Streptococcus pneumoniae</i>	74	ATP synthase subunit C (atpG)
Rv1310	MSMEG_4936	<i>Streptococcus pneumoniae</i>	94	ATP synthase subunit B (atpD)
Rv1311	MSMEG_4935	<i>Streptococcus pneumoniae</i>	81	ATP synthase subunit epsilon (atpC)
Rv1315	MSMEG_4932	<i>Staphylococcus aureus</i>	88	UDP-N-acetylglucosamine 1-carboxyvinyltransferase (murA)
Rv1326c	MSMEG_4918	<i>Haemophilus influenzae</i>	81	glycogen branching enzyme (glgB)
Rv1349	MSMEG_6553	<i>Salmonella typhimurium</i>	67	Probable drugs-transprot transmembrane ATP-binding protein ABC transporter (MdlB)
Rv1350	MSMEG_2206	<i>Bacillus subtilis</i>	77	3-ketoacyl-(acyl-carrier-protein) reductase (fabG)
Rv1381	MSMEG_3044	<i>Staphylococcus aureus</i>	83	dihydroorotase (pyrC)
Rv1389	MSMEG_3051	<i>Helicobacter pylori</i>	75	guanylate kinase (gmk)
Rv1390	MSMEG_3053	<i>Streptococcus pneumoniae</i>	79	DNA-directed RNA polymerase omega subunit (rpoZ)
Rv1392	MSMEG_3055	<i>Bacillus subtilis</i>	87	S-adenosylmethionine synthetase (metK)
Rv1406	MSMEG_3064	<i>Escherichia coli</i>	67	methionyl-tRNA formyltransferase (fmt)
Rv1409	MSMEG_3067	<i>Escherichia coli</i>	64	Probable bifunctional riboflavin biosynthesis protein (ribG)
Rv1416	MSMEG_3073	<i>Escherichia coli</i>	74	riboflavin synthase subunit beta (ribH)
Rv1436	MSMEG_3084	<i>Escherichia coli</i>	87	glyceraldehyde-3-phosphate dehydrogenase (gap)
Rv1437	MSMEG_3085	<i>Bacillus subtilis</i>	71	phosphoglycerate kinase (pgk)
Rv1438	MSMEG_3086	<i>Streptococcus pneumoniae</i>	80	triosephosphate isomerase (tpi)
Rv1449c	MSMEG_3103	<i>Bacillus subtilis</i>	77	transketolase (tkt)
Rv1461	MSMEG_3122	<i>Bacillus subtilis</i>	94	hypothetical protein Rv1461 (component of SufBCD complex)
Rv1462	MSMEG_3123	<i>Bacillus subtilis</i>	78	hypothetical protein Rv1462 (SufB)
Rv1463	MSMEG_3124	<i>Bacillus subtilis</i>	85	Probable conserved ATP-binding protein ABC transporter (FeS assembly)
Rv1464	MSMEG_3125	<i>Staphylococcus aureus</i>	83	Probable cysteine desulfurase (csd)
Rv1465	MSMEG_3126	<i>Bacillus subtilis</i>	83	Possible nitrogen fixation related protein (IscU)

---

Rv1547	MSMEG_3178	<i>Escherichia coli</i>	84	DNA polymerase III subunit alpha (dnaE1)
Rv1601	MSMEG_3207	<i>Haemophilus influenzae</i>	81	imidazoleglycerol-phosphate dehydratase (hisB)
Rv1603	MSMEG_3209	<i>Haemophilus influenzae</i>	85	Probable phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase (hisA)
Rv1609	MSMEG_3217	<i>Salmonella typhimurium</i>	80	anthranilate synthase component I (trpE)
Rv1612	MSMEG_3220	<i>Streptococcus pneumoniae</i>	84	tryptophan synthase subunit beta (trpB)
Rv1617	MSMEG_3227	<i>Staphylococcus aureus</i>	84	pyruvate kinase (pykA)
Rv1622c	MSMEG_3232	<i>Escherichia coli</i>	60	Probable integral membrane cytochrome D ubiquinol oxidase (subunit II) cydB
Rv1629	MSMEG_3839	<i>Staphylococcus aureus</i>	77	DNA polymerase I (polA)
Rv1630	MSMEG_3833	<i>Haemophilus influenzae</i>	84	30S ribosomal protein S1 (rpsA)
Rv1631	MSMEG_2076	<i>Bacillus subtilis</i>	44	dephospho-CoA kinase/unknown domain fusion protein (coaE)
Rv1641	MSMEG_3793	<i>Bacillus subtilis</i>	86	translation initiation factor IF-3 (infC)
Rv1649	MSMEG_3778	<i>Bacillus subtilis</i>	80	phenylalanyl-tRNA synthetase alpha subunit (pheS)
Rv1650	MSMEG_3777	<i>Salmonella typhimurium</i>	65	phenylalanyl-tRNA synthetase beta subunit (pheT)
Rv1652	MSMEG_3776	<i>Escherichia coli</i>	71	N-acetyl-gamma-glutamyl-phosphate reductase (argC)
Rv1655	MSMEG_3773	<i>Staphylococcus aureus</i>	69	acetylornithine aminotransferase (argD)
Rv1659	MSMEG_3769	<i>Haemophilus influenzae</i>	71	argininosuccinate lyase (argH)
Rv1661	MSMEG_0408	<i>Bacillus subtilis</i>	37	Probable polyketide synthase (pks7)
Rv1662	MSMEG_0408	<i>Bacillus subtilis</i>	43	Probable polyketide synthase (pks8)
Rv1689	MSMEG_3758	<i>Escherichia coli</i>	70	tyrosyl-tRNA synthetase (tyrS)
Rv1695	MSMEG_3750	<i>Escherichia coli</i>	83	inorganic polyphosphate/ATP-NAD kinase (ppnK)
Rv1699	MSMEG_3746	<i>Bacillus subtilis</i>	90	CTP synthetase (pyrG)
Rv1701	MSMEG_3744	<i>Escherichia coli</i>	78	tyrosine recombinase (xerD)
Rv1711	MSMEG_3740	<i>Escherichia coli</i>	75	hypothetical protein Rv1711 (RsuA)
Rv1712	MSMEG_3739	<i>Bacillus subtilis</i>	61	cytidylate kinase (cmk)
Rv1713	MSMEG_3738	<i>Staphylococcus aureus</i>	72	GTP-binding protein (engA)
Rv1714	MSMEG_3607	<i>Staphylococcus aureus</i>	28	Probable oxidoreductase (fabG)
Rv1850	MSMEG_3625	<i>Staphylococcus aureus</i>	76	Urease alpha subunit (ureC)
Rv2093c	MSMEG_3886	<i>Helicobacter pylori</i>	66	Probable Sec-independent protein translocase transmembrane protein (tatC)
Rv2121c	MSMEG_4180	<i>Haemophilus influenzae</i>	84	Probable ATP phosphoribosyltransferase (hisG)
Rv2130c	MSMEG_4215	<i>Escherichia coli</i>	51	hypothetical protein Rv2030c
Rv2150c	MSMEG_4222	<i>Bacillus subtilis</i>	78	cell division protein (ftsZ)

Rv2152c	MSMEG_4226	<i>Haemophilus influenzae</i>	70	UDP-N-acetylmuramate--L-alanine ligase (murC)
Rv2153c	MSMEG_4227	<i>Salmonella typhimurium</i>	63	N-acetylglucosaminyl transferase (murG)
Rv2154c	MSMEG_4228	<i>Salmonella typhimurium</i>	54	FtsW-like protein (ftsW)
Rv2155c	MSMEG_4229	<i>Bacillus subtilis</i>	60	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate synthetase (murD)
Rv2156c	MSMEG_4230	<i>Bacillus subtilis</i>	74	phospho-N-acetylmuramoyl-pentapeptide-transferase (murX)
Rv2157c	MSMEG_4231	<i>Streptococcus pneumoniae</i>	59	Probable UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate- D-alanyl-D-alanyl ligase (murF)
Rv2163c	MSMEG_4233	<i>Salmonella typhimurium</i>	75	Probable penicillin-binding membrane protein (pbpB)
Rv2165c	MSMEG_4235	<i>Streptococcus pneumoniae</i>	63	S-adenosyl-methyltransferase (mraW)
Rv2166c	MSMEG_4236	<i>Mycoplasma genitalium</i>	89	hypothetical protein Rv2166c
Rv2182c	MSMEG_4248	<i>Bacillus subtilis</i>	73	1-acylglycerol-3-phosphate O-acyltransferase (plsC)
Rv2188c	MSMEG_4253	<i>Haemophilus influenzae</i>	67	hypothetical protein Rv2188c (rfaG)
Rv2192c	MSMEG_4258	<i>Haemophilus influenzae</i>	65	anthranilate phosphoribosyltransferase (trpD)
Rv2215	MSMEG_4283	<i>Escherichia coli</i>	72	dihydrolipoamide acyltransferase (sucB)
Rv2217	MSMEG_4285	<i>Haemophilus influenzae</i>	76	lipoyltransferase (lipB)
Rv2218	MSMEG_4286	<i>Staphylococcus aureus</i>	84	lipoyl synthase (lipA)
Rv2220	MSMEG_4290	<i>Salmonella typhimurium</i>	84	Glutamine synthetase (glnA1)
Rv2225	MSMEG_4298	<i>Helicobacter pylori</i>	71	3-methyl-2-oxobutanoate hydroxymethyltransferase (panB)
Rv2231c	MSMEG_3206	<i>Haemophilus influenzae</i>	27	hypothetical protein Rv2231c (cobC)
Rv2245	MSMEG_4327	<i>Streptococcus pneumoniae</i>	83	3-oxoacyl-(acyl carrier protein) synthase (kasA)
Rv2246	MSMEG_4328	<i>Streptococcus pneumoniae</i>	79	3-oxoacyl-(acyl carrier protein) synthase (kasB)
Rv2343c	MSMEG_4482	<i>Bacillus subtilis</i>	76	DNA primase (dnaG)
Rv2357c	MSMEG_4485	<i>Mycoplasma genitalium</i>	80	glycyl-tRNA synthetase (glyS)
Rv2361c	MSMEG_4490	<i>Escherichia coli</i>	76	Long (C50) chain Z-isoprenyl diphosphate synthase
Rv2364c	MSMEG_4493	<i>Bacillus subtilis</i>	79	GTP-binding protein (era)
Rv2373c	MSMEG_4504	<i>Staphylococcus aureus</i>	71	PROBABLE CHAPERONE PROTEIN (dnaJ2)
Rv2397c	MSMEG_4530	<i>Staphylococcus aureus</i>	79	Probable sulfate-transport ATP-binding protein ABC transporter (cysA1)
Rv2398c	MSMEG_4531	<i>Escherichia coli</i> MG1655	73	Probable sulfate-transport integral membrane protein ABC transporter (cysW)
Rv2399c	MSMEG_4532	<i>Haemophilus influenzae</i>	73	Probable sulfate-transport integral membrane protein ABC transporter (cysT)
Rv2421c	MSMEG_4581	<i>Bacillus subtilis</i>	76	nicotinic acid mononucleotide adenylyltransferase (nadD)
Rv2440c	MSMEG_4623	<i>Staphylococcus aureus</i>	66	Probable GTP1/OBG-family GTP-binding protein (obg)



---

Rv2441c	MSMEG_4624	<i>Helicobacter pylori</i>	91	50S ribosomal protein L27 (rpmA)
Rv2442c	MSMEG_4625	<i>Bacillus subtilis</i>	88	50S ribosomal protein L21 (rplU)
Rv2444c	MSMEG_4626	<i>Haemophilus influenzae</i>	69	Possible ribonuclease E (rne)
Rv2447c	MSMEG_4232	<i>Streptococcus pneumoniae</i>	29	Probable folypolyglutamate synthase protein (folC)
Rv2448c	MSMEG_4630	<i>Bacillus subtilis</i>	78	valyl-tRNA synthetase (valS)
Rv2457c	MSMEG_4671	<i>Salmonella typhimurium</i>	95	ATP-dependent protease ATP-binding subunit (clpX)
Rv2460c	MSMEG_4672	<i>Haemophilus influenzae</i>	88	ATP-dependent Clp protease proteolytic subunit (clpP2)
Rv2477c	MSMEG_4700	<i>Haemophilus influenzae</i>	84	Probable macrolide-transport ATP-binding protein ABC transporter
Rv2500c	MSMEG_4715	<i>Salmonella typhimurium</i>	73	Possible acyl-CoA dehydrogenase (fadE19)
Rv2511	MSMEG_4724	<i>Haemophilus influenzae</i>	79	oligoribonuclease (orn)
Rv2534c	MSMEG_3035	<i>Staphylococcus aureus</i>	80	elongation factor P (efp)
Rv2537c	MSMEG_1922	<i>Helicobacter pylori</i>	44	3-dehydroquinate dehydratase (aroD)
Rv2538c	MSMEG_3033	<i>Haemophilus influenzae</i>	77	3-dehydroquinate synthase (aroB)
Rv2539c	MSMEG_3031	<i>Haemophilus influenzae</i>	77	shikimate kinase (aroK)
Rv2540c	MSMEG_3030	<i>Haemophilus influenzae</i>	77	chorismate synthase (aroF)
Rv2555c	MSMEG_3025	<i>Salmonella typhimurium</i>	78	alanyl-tRNA synthetase (alaS)
Rv2580c	MSMEG_2976	<i>Staphylococcus aureus</i>	75	histidyl-tRNA synthetase (hisS)
Rv2614c	MSMEG_2931	<i>Streptococcus pneumoniae</i>	79	threonyl-tRNA synthetase (thrS)
Rv2682c	MSMEG_2776	<i>Salmonella typhimurium</i>	77	1-deoxy-D-xylulose-5-phosphate synthase (dxs1)
Rv2703	MSMEG_2758	<i>Bacillus subtilis</i>	87	RNA polymerase sigma factor (sigA)
Rv2710	MSMEG_2752	<i>Staphylococcus aureus</i>	93	RNA polymerase sigma factor (sigB)
Rv2727c	MSMEG_2734	<i>Escherichia coli</i>	58	tRNA delta(2)-isopentenylpyrophosphate transferase (miaA)
Rv2746c	MSMEG_2692	<i>Streptococcus pneumoniae</i>	61	Probable PGP synthase (pgsA3)
Rv2753c	MSMEG_2684	<i>Bacillus subtilis</i>	75	dihydrodipicolinate synthase (dapA)
Rv2786c	MSMEG_2653	<i>Escherichia coli</i>	76	hypothetical protein Rv2786c (ribF)
Rv2832c	MSMEG_0508	<i>Staphylococcus aureus</i>	54	Probable Sn-glycerol-3-phosphate transport ATP-binding protein (ugpC)
Rv2839c	MSMEG_2628	<i>Bacillus subtilis</i>	90	translation initiation factor IF-2 (infB)
Rv2841c	MSMEG_2625	<i>Bacillus subtilis</i>	83	transcription elongation factor (nusA)
Rv2845c	MSMEG_2621	<i>Haemophilus influenzae</i>	73	prolyl-tRNA synthetase (proS)
Rv2846c	MSMEG_2619	<i>Haemophilus influenzae</i>	71	Possible integral membrane efflux protein (efpA)
Rv2869c	MSMEG_2579	<i>Escherichia coli</i>	61	Probable conserved transmembrane protein

---

Rv2881c	MSMEG_2543	<i>Salmonella typhimurium</i>	59	Probable integral membrane phosphatidate cytidyltransferase (cdsA)
Rv2882c	MSMEG_2541	<i>Bacillus subtilis</i>	80	ribosome releasing factor (frr)
Rv2883c	MSMEG_2540	<i>Streptococcus pneumoniae</i>	83	uridylate kinase (pyrH)
Rv2889c	MSMEG_2520	<i>Bacillus subtilis</i>	70	elongation factor Ts (tsf)
Rv2890c	MSMEG_2519	<i>Staphylococcus aureus</i>	93	30S ribosomal protein S2 (rpsB)
Rv2897c	MSMEG_2509	<i>Haemophilus influenzae</i>	66	hypothetical protein Rv2897c
Rv2904c	MSMEG_2440	<i>Bacillus subtilis</i>	80	50S ribosomal protein L19 (rplS)
Rv2906c	MSMEG_2438	<i>Streptococcus pneumoniae</i>	82	tRNA (guanine-N(1)-)-methyltransferase (trmD)
Rv2916c	MSMEG_2430	<i>Bacillus subtilis</i>	81	Probable signal recognition particle protein (ffh)
Rv2925c	MSMEG_2418	<i>Bacillus subtilis</i>	78	ribonuclease III (rncS)
Rv2977c	MSMEG_2398	<i>Escherichia coli</i>	59	thiamine monophosphate kinase (thiL)
Rv2981c	MSMEG_2395	<i>Bacillus subtilis</i>	76	D-alanylalanine synthetase (ddlA)
Rv2986c	MSMEG_2389	<i>Bacillus subtilis</i>	85	Probable DNA-binding protein HU homolog (hupB)
Rv2987c	MSMEG_2388	<i>Haemophilus influenzae</i>	86	isopropylmalate isomerase small subunit (leuD)
Rv2992c	MSMEG_2383	<i>Bacillus subtilis</i>	72	glutamyl-tRNA synthetase (gltS)
Rv2995c	MSMEG_2379	<i>Haemophilus influenzae</i>	77	3-isopropylmalate dehydrogenase (leuB)
Rv2996c	MSMEG_2378	<i>Staphylococcus aureus</i>	76	Probable D-3-phosphoglycerate dehydrogenase (serA1)
Rv3009c	MSMEG_2367	<i>Staphylococcus aureus</i>	81	aspartyl/glutamyl-tRNA amidotransferase subunit B (gatB)
Rv3011c	MSMEG_2365	<i>Staphylococcus aureus</i>	77	glutamyl-tRNA amidotransferase subunit A (gatA)
Rv3014c	MSMEG_2362	<i>Salmonella typhimurium</i>	79	DNA ligase (ligA)
Rv3025c	MSMEG_2357	<i>Escherichia coli</i>	74	Probable cysteine desulfurase (iscS)
Rv3032	MSMEG_0933	<i>Saccharomyces cerevisiae</i>	27	Possible transferase
Rv3043c	MSMEG_4437	<i>Staphylococcus aureus</i>	82	Probable cytochrome C oxidase polypeptide I (ctaD)
Rv3048c	MSMEG_1033	<i>Mycoplasma genitalium</i>	93	ribonucleotide-diphosphate reductase beta subunit (nrdF2)
Rv3051c	MSMEG_1019	<i>Mycoplasma genitalium</i>	92	ribonucleotide-diphosphate reductase alpha subunit (nrdE)
Rv3105c	MSMEG_2086	<i>Bacillus subtilis</i>	75	peptide chain release factor 2 (prfB)
Rv3140	MSMEG_2080	<i>Escherichia coli</i>	81	Probable acyl-coA dehydrogenase (fadE23)
Rv3177	MSMEG_6831	<i>Haemophilus influenzae</i>	37	Possible peroxidase
Rv3240c	MSMEG_1881	<i>Bacillus subtilis</i>	78	Probable preprotein translocase (secA1)
Rv3246c	MSMEG_1874	<i>Bacillus subtilis</i>	84	Two component sensory transduction transcriptional regulatory protein (mtrA)
Rv3255c	MSMEG_1836	<i>Saccharomyces cerevisiae</i>	74	Probable mannose-6-phosphate isomerase (manA)

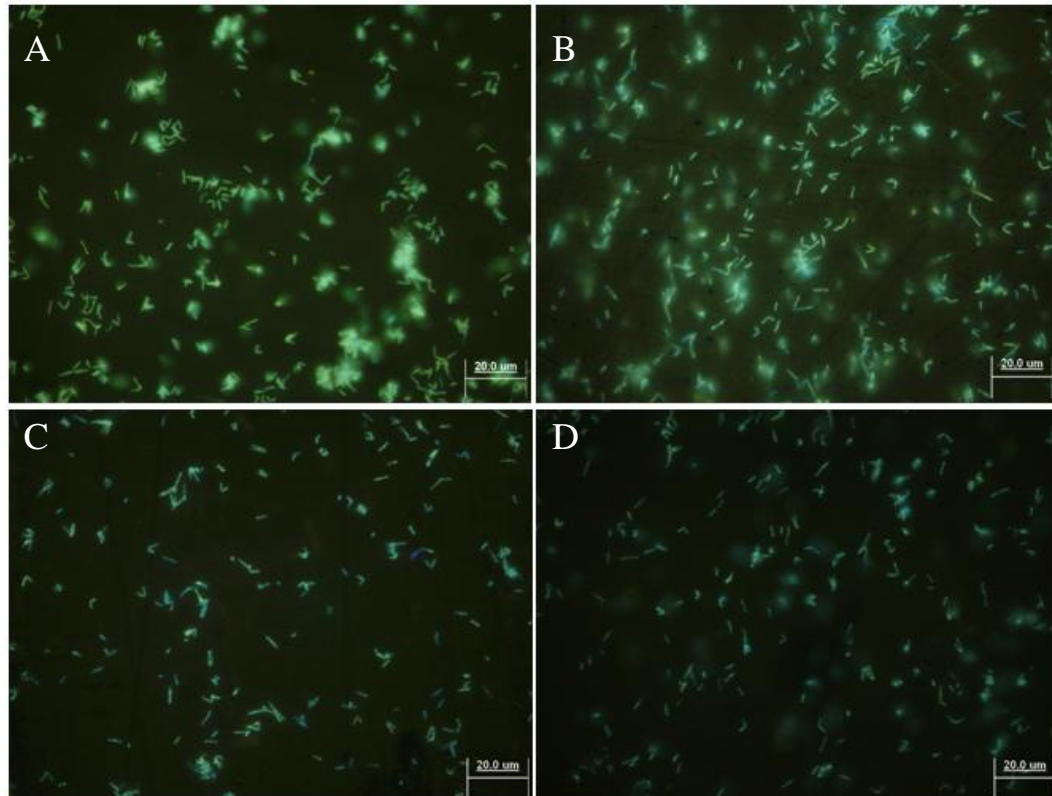
Rv3257c	MSMEG_1834	<i>Bacillus subtilis</i>	80	Probable phosphomannomutase (pmmA)
Rv3285	MSMEG_1807	<i>Bacillus subtilis</i>	81	Probable bifunctional protein acetyl-/propionyl-coenzyme A carboxylase (accA3)
Rv3332	MSMEG_2119	<i>Haemophilus influenzae</i>	47	Probable n-acetylglucosamine-6-phosphate deacetylase (nagA)
Rv3336c	MSMEG_1657	<i>Escherichia coli</i>	79	tryptophanyl-tRNA synthetase (trpS)
Rv3356c	MSMEG_1647	<i>Bacillus subtilis</i>	82	Probable bifunctional protein (fold)
Rv3396c	MSMEG_1610	<i>Staphylococcus aureus</i>	83	bifunctional GMP synthase/glutamine amidotransferase protein (guaA)
Rv3398c	MSMEG_4240	<i>Haemophilus influenzae</i>	31	Probable multifunctional geranylgeranyl pyrophosphate synthetase (idsA1)
Rv3411c	MSMEG_1602	<i>Bacillus subtilis</i>	77	inositol-5-monophosphate dehydrogenase (guaB2)
Rv3418c	MSMEG_1582	<i>Bacillus subtilis</i>	99	co-chaperonin (groES)
Rv3422c	MSMEG_1577	<i>Bacillus subtilis</i>	63	hypothetical protein Rv3422c
Rv3423c	MSMEG_1575	<i>Bacillus subtilis</i>	61	alanine racemase (alr)
Rv3436c	MSMEG_1568	<i>Escherichia coli</i>	85	D-fructose-6-phosphate amidotransferase (glmS)
Rv3441c	MSMEG_1559	<i>Bacillus subtilis</i>	67	Probable phospho-sugar mutase/ mrsA protein homolog (mrsA)
Rv3443c	MSMEG_1556	<i>Escherichia coli</i>	80	50S ribosomal protein L13 (rplM)
Rv3457c	MSMEG_1524	<i>Bacillus subtilis</i>	90	DNA-directed RNA polymerase alpha subunit (rpoA)
Rv3458c	MSMEG_1523	<i>Haemophilus influenzae</i>	83	30S ribosomal protein S4 (rpsD)
Rv3459c	MSMEG_1522	<i>Bacillus subtilis</i>	92	30S ribosomal protein S11 (rpsK)
Rv3462c	MSMEG_1519	<i>Bacillus subtilis</i>	98	translation initiation factor IF-1 (infA)
Rv3464	MSMEG_1512	<i>Haemophilus influenzae</i>	83	dTDP-glucose 4,6-dehydratase (rmlB)
Rv3559c	MSMEG_6011	<i>Bacillus subtilis</i>	78	short chain dehydrogenase
Rv3579c	MSMEG_6073	<i>Mycoplasma genitalium</i>	61	Possible tRNA/rRNA methyltransferase
Rv3580c	MSMEG_6074	<i>Escherichia coli</i>	78	cysteinyl-tRNA synthetase (cysS1)
Rv3581c	MSMEG_6075	<i>Haemophilus influenzae</i>	64	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (ispF)
Rv3582c	MSMEG_6076	<i>Haemophilus influenzae</i>	58	2-C-methyl-D-erythritol 4-phosphate cytidyltransferase (ispD)
Rv3596c	MSMEG_6091	<i>Staphylococcus aureus</i>	88	Probable ATP-dependent protease ATP-binding subunit (clpC1)
Rv3598c	MSMEG_6094	<i>Bacillus subtilis</i>	74	lysyl-tRNA synthetase (lysS)
Rv3602c	MSMEG_6097	<i>Helicobacter pylori</i>	60	pantoate--beta-alanine ligase (panC)
Rv3608c	MSMEG_6103	<i>Streptococcus pneumoniae</i>	63	Dihydropteroate synthase 1 (folP1)
Rv3609c	MSMEG_6104	<i>Streptococcus pneumoniae</i>	86	GTP cyclohydrolase I (folE)
Rv3610c	MSMEG_6105	<i>Staphylococcus aureus</i>	76	Membrane-bound protease- cell division protein (ftsH)

---

Rv3627c	MSMEG_6113	<i>Haemophilus influenzae</i>	56	hypothetical protein Rv3627c
Rv3634c	MSMEG_6142	<i>Haemophilus influenzae</i>	63	UDP-glucose 4-epimerase (galE1)
Rv3646c	MSMEG_6157	<i>Haemophilus influenzae</i>	81	DNA topoisomerase I (topA)
Rv3708c	MSMEG_6256	<i>Bacillus subtilis</i>	78	aspartate-semialdehyde dehydrogenase (asd)
Rv3709c	MSMEG_6257	<i>Helicobacter pylori</i>	84	aspartate kinase (ask)
Rv3712	MSMEG_6276	<i>Streptococcus pneumoniae</i>	64	Possible ligase
Rv3713	MSMEG_6277	<i>Streptococcus pneumoniae</i>	83	Possible cobyrinic acid synthase (cobQ2)
Rv3721c	MSMEG_6285	<i>Staphylococcus aureus</i>	72	DNA polymerase III subunits gamma and tau (dnaZX)
Rv3752c	MSMEG_6327	<i>Escherichia coli</i>	80	Possible cytidine/deoxycytidylate deaminase
Rv3778c	MSMEG_6363	<i>Bacillus subtilis</i>	70	Possible aminotransferase
Rv3781	MSMEG_6366	<i>Bacillus subtilis</i>	88	Probable O-antigen/lipopolysaccharide transporter (rfbE)
Rv3799c	MSMEG_6391	<i>Haemophilus influenzae</i>	81	Probable propionyl-CoA carboxylase beta chain 4 (accD4)
Rv3800c	MSMEG_6392	<i>Streptococcus pneumoniae</i>	70	Polyketide synthase (pks13)
Rv3834c	MSMEG_6413	<i>Bacillus subtilis</i>	71	seryl-tRNA synthetase (serS)
Rv3907c	MSMEG_6926	<i>Streptococcus pneumoniae</i>	80	Probable poly(A) polymerase (pcnA)
Rv3913	MSMEG_6933	<i>Streptococcus pneumoniae</i>	67	Probable thioredoxin reductase (trxB2)
Rv3917c	MSMEG_6938	<i>Staphylococcus aureus</i>	76	Probable chromosome partitioning protein (parB)
Rv3918c	MSMEG_6939	<i>Helicobacter pylori</i>	79	Probable chromosome partitioning protein (parA)
Rv3921c	MSMEG_6942	<i>Haemophilus influenzae</i>	72	putative inner membrane protein translocase component YidC

### 9.3 Cellular morphology of *M. smegmatis* strains under different culture conditions

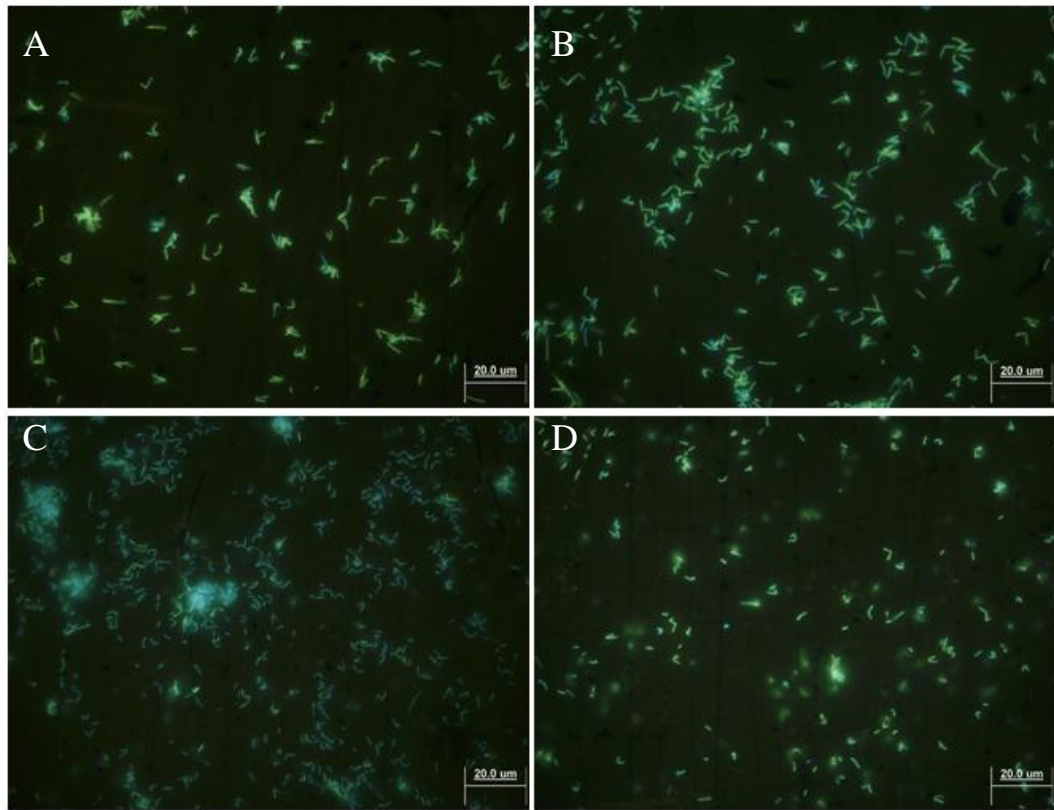
#### 9.3.1 Cell morphology of cultures in LBT



**Figure 9.2: Cell morphology of cultures in LBT**

The above pictures show the cell morphology of all the strains cultured in LBT with 20 ng/mL Tc stained with DAPI at 144 hours. **A)** *M. smegmatis*-pMindGFP **B)** *M. smegmatis*-pMind\_parB\_GFP **C)** *M. smegmatis*-pMind\_obg and **D)** *M. smegmatis*-pMind\_parA\_GFP. There was no detectable difference in the cell morphology between the different strains. All cells were approximately 2-6  $\mu\text{m}$  long.

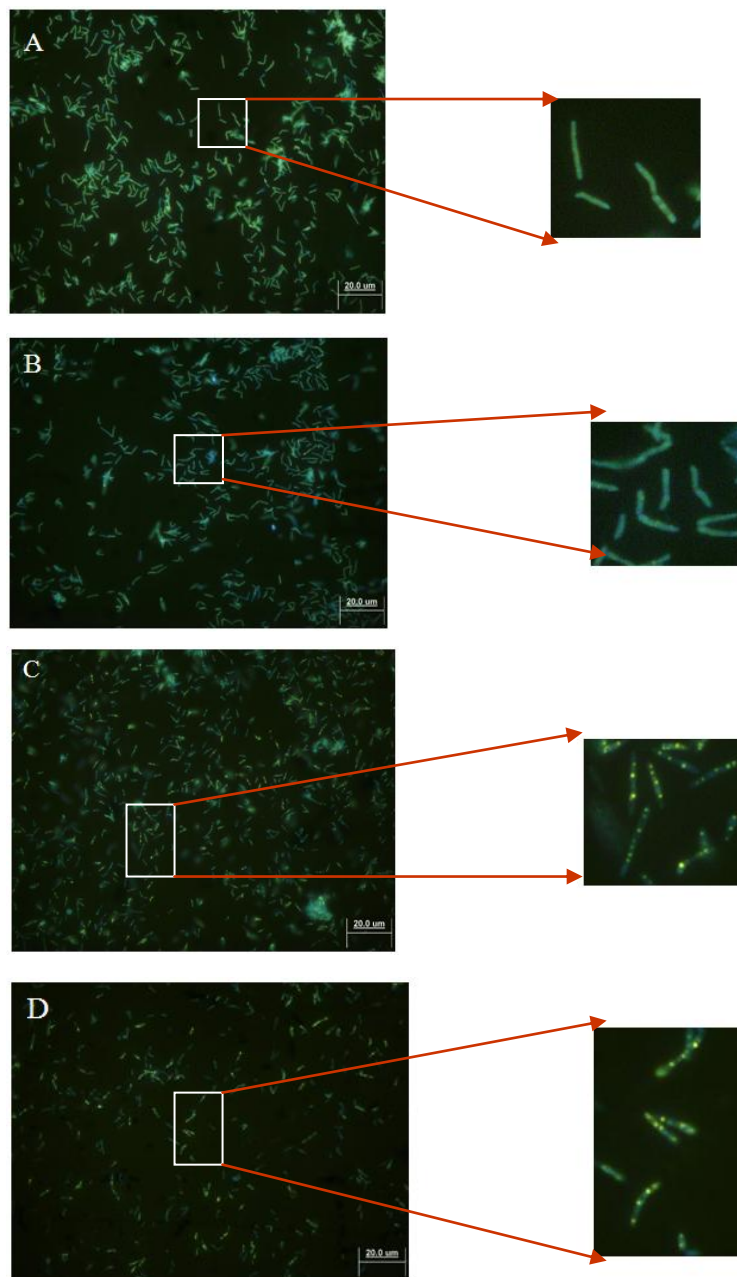
### 9.3.2 Cell morphology of cultures in C<sup>-</sup> media



**Figure 9.3: Cell morphology of cultures in C<sup>-</sup> media**

The above pictures show the cell morphology of all the strains cultured in C<sup>-</sup> media with 20 ng/mL Tc stained with DAPI at 72 hours. **A)** *M. smegmatis*-pMindGFP **B)** *M. smegmatis*-pMind<sub>parB</sub>\_GFP **C)** *M. smegmatis*-pMind<sub>obg</sub> and **D)** *M. smegmatis*-pMind<sub>parA</sub>\_GFP. All strains had cells ~ 4-7 μm long except *M. smegmatis*-pMind<sub>parA</sub>\_GFP which formed slightly smaller rods ~ 2-4 μm long.

### 9.3.3 Cell morphology of cultures in N<sup>-</sup> media

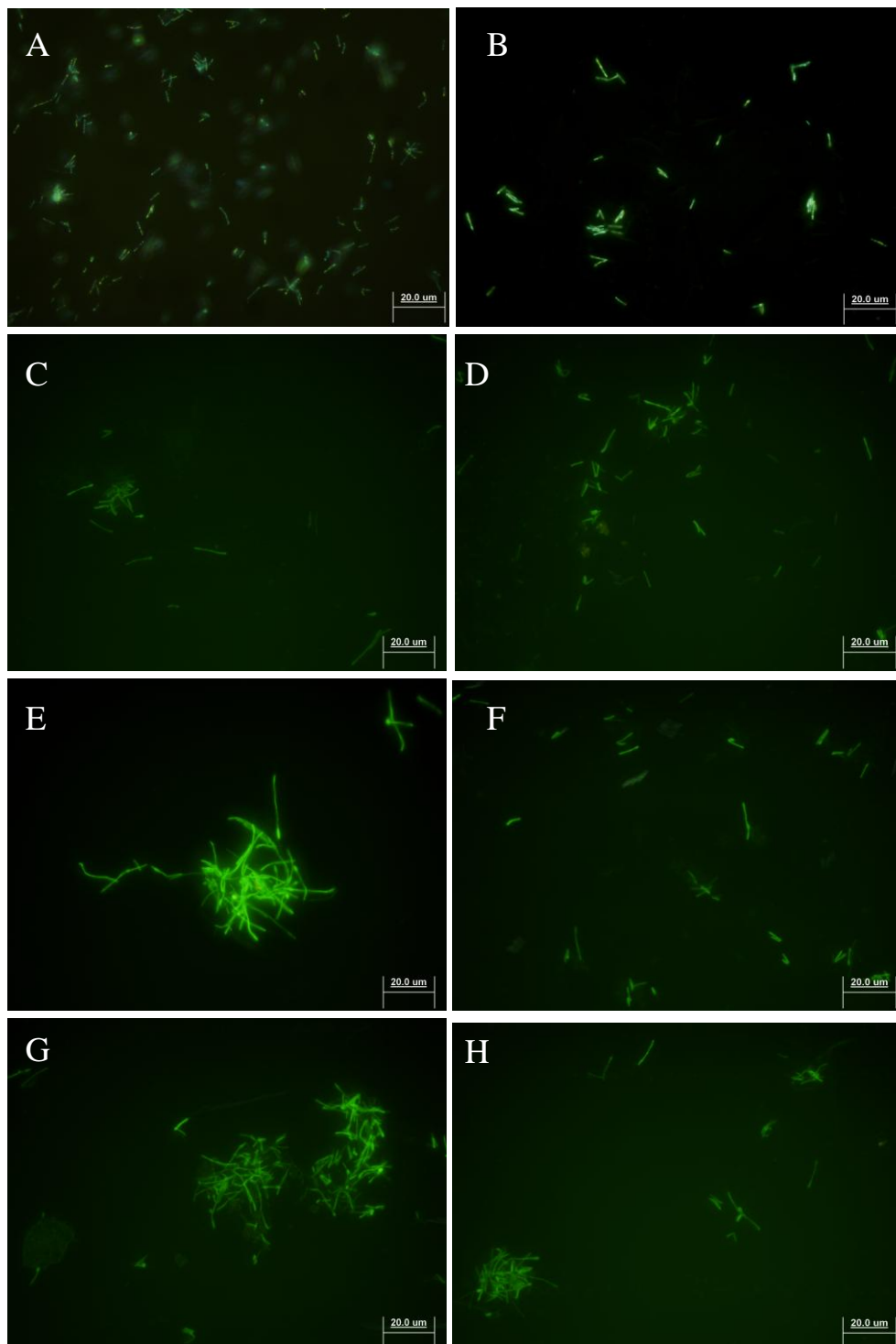


**Figure 9.4: Cell morphology of cultures in N<sup>-</sup> media**

The above pictures show the cell morphology of strains cultured in N<sup>-</sup> media with 20 ng/mL Tc stained with DAPI at 72 hours of **A)** *M. smegmatis*-pMindGFP, **B)** *M. smegmatis*-pMind<sub>parB</sub>-GFP, **C)** *M. smegmatis*-pMind<sub>obg</sub> and **D)** *M. smegmatis*-pMind<sub>parA</sub>-GFP. The cells were slightly elongated in nitrogen starvation media, ~ 6-11 μm in length. *M. smegmatis*-pMind<sub>obg</sub> and *M. smegmatis*-pMind<sub>parA</sub>-GFP cultures had irregularly located foci along the length of the cells.

## 9.4 Cellular Morphology of *M. smegmatis* strains in HTS

### 9.4.1 Cell morphology in $N^-$ media with LOPAC hits

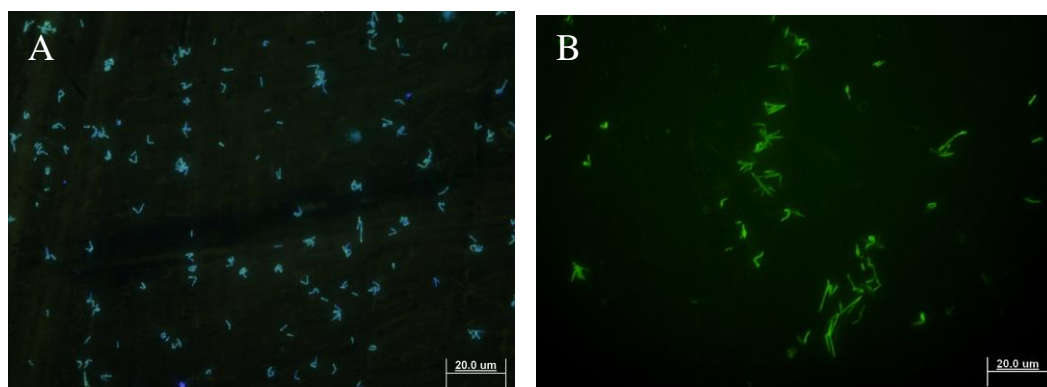


**Figure 9.5:** Cell morphology of cultures in  $N^-$  media with hits from LOPAC



All pictures were taken after 72 hours incubation in  $N^-$  media with 20 ng/mL Tc when screening for compounds in LOPAC. **A)** *M. smegmatis*\_pMind2357 and **B)** *M. smegmatis*\_pMind\_parA\_GFP cells without any compounds and stained with DAPI. Both these strains show cells ~ 6-10  $\mu$ M in length. GFP picture of *M. smegmatis*\_pMind\_parA\_GFP cells in the presence of 20  $\mu$ M **C)** amsacrine hydrochloride showing cells ~ 15-20  $\mu$ M in length, **D)** AA-861 showing cells ~ 7-17  $\mu$ M in length, **E)** R-(-)-fluoxetine hydrochloride showing cells ~ 20-25  $\mu$ M in length, **F)** fenoldopam bromide showing cells ~ 10-16  $\mu$ M in length, **G)** ( $\pm$ )-octoclothePIN maleate showing cells ~ 15-20  $\mu$ M in length and **H)** methiothepin mesylate showing cells ~ 15-20  $\mu$ M in length. Cells of *M. smegmatis*\_pMind2357 with the above mentioned compounds showed no difference compared to the cells shown in A) (data not shown).

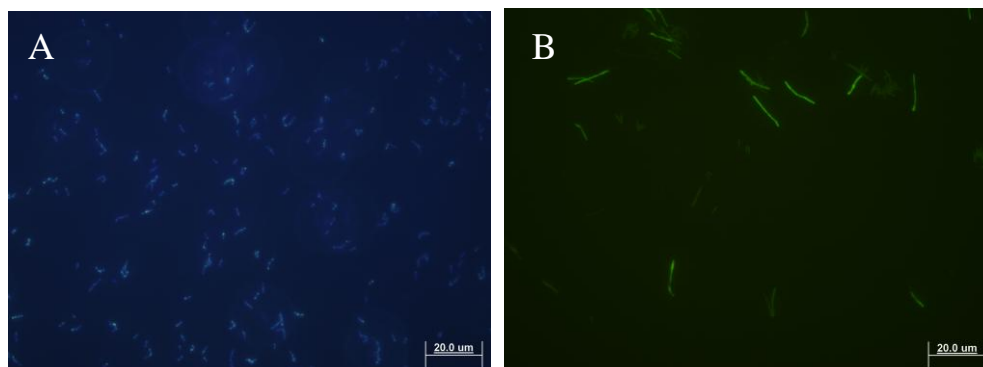
#### 9.4.2 Cell morphology in $N^-$ media with Spectrum Collection hits



**Figure 9.6: Cell morphology of cultures in  $N^-$  media with hits from Spectrum Collection**

Pictures were taken after 72 hours incubation in  $N^-$  media with 20 ng/mL Tc when screening for compounds in the Spectrum Collection. Both pictures show cells in 20  $\mu$ M phenoxybenzamine hydrochloride with **A)** *M. smegmatis*\_pMind2357 cells stained with DAPI (2  $\mu$ M – 5  $\mu$ M) and **B)** *M. smegmatis*\_pMind\_parA\_GFP cells (5  $\mu$ M – 15  $\mu$ M).

### 9.4.3 Cell morphology in C<sup>-</sup> media with NIH Diversity Set hit



**Figure 9.7: Cell morphology of cultures in C<sup>-</sup> media with hits from NIH Diversity Set**

Pictures were taken after 72 hours incubation in C<sup>-</sup> media with 20 ng/mL Tc when screening for compounds in NIH Diversity Set. Both pictures show cells in 20  $\mu$ M of compound NSC22242 with **A)** *M. smegmatis*\_pMind2357 stained with DAPI, cells ~ 2-5  $\mu$ M in length and **B)** GFP image of *M. smegmatis*\_pMind\_*parA*\_GFP, cells ~ 10-17  $\mu$ M in length.

## 9.5 Media and Solutions

### 9.5.1 Hartmans-de Bont minimal media [125]

#### 100x Metal Salts Solution

○ $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$	0.1 g
○ $\text{CoCl}_2 \cdot 2\text{H}_2\text{O}$	0.04 g
○ $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$	0.02 g
○ EDTA	1 g
○ $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$	0.5 g
○ $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$	10 g
○ $\text{MnCl}_2 \cdot 2\text{H}_2\text{O}$	0.1 g
○ $\text{NaMoO}_4 \cdot 2\text{H}_2\text{O}$	0.02 g
○ $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$	0.2 g

#### 100x Phosphates

○ $\text{K}_2\text{HPO}_4$	155 g
○ $\text{NaH}_2\text{PO}_4$	85 g

#### 100x Nitrogen

○ $(\text{NH}_4)_2\text{SO}_4$	200 g
--------------------------------	-------

The above solutions were made to 1 L with deionized water in Schott bottles, sterilized by autoclaving for 30 minutes at 121°C on a liquid cycle and stored at 4°C and used for up to 12 months.

#### HdeB Base

○ $\text{H}_2\text{O}$	945 mL
○ 100x Metal salts	10 mL
○ 10% Tween 80	5 mL

---

Oxygen Starvation ( $O_2^-$ ) Media

- HdeB base 960 mL
- 100x Nitrogen 10 mL
- 100x Phosphate 10 mL
- 10% Glycerol 20 mL

Carbon Starvation ( $C^-$ ) Media

- HdeB base 960 mL
- 100x Nitrogen 10 mL
- 100x Phosphate 10 mL
- 10% Glycerol 8 mL
- $H_2O$  12 mL

Nitrogen Starvation ( $N^-$ ) Media

- HdeB base 960 mL
- 100x Nitrogen 0.1 mL
- 100x Phosphate 10 mL
- 10% Glycerol 20 mL
- $H_2O$  9.9 mL

All solutions were sterilized by autoclaving for 30 minutes at 121°C on a liquid cycle and stored at RT and used within a week.

### 9.5.2 ZYP-5052

ZY

- N-Z amine AS or tryptone 10 g
- Yeast extract 5 g

ZY was made up to 919 mL with deionized water in Schott bottles, sterilized by autoclaving for 30 minutes at 121°C on a liquid cycle and stored at 4°C.

20x NPS (pH ~ 6.75)

(NPS = 100 mM PO<sub>4</sub>, 25 mM SO<sub>4</sub>, 50 mM NH<sub>4</sub>, 100 mM Na, 50 mM K)

- (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 66
- KH<sub>2</sub>PO<sub>4</sub> 136 g
- 1 M Na<sub>2</sub>HPO<sub>4</sub> 142 g

50x 5052

- Glycerol 250 g (w/v)
- Glucose 25 g
- α-lactose 100 g

The above two solutions were made up to 1 L with deionized water in Schott bottles, sterilized by autoclaving for 30 minutes at 121°C on a liquid cycle and stored at 4°C.

ZYP-5052

- ZY 919 mL
- 1 M MgSO<sub>4</sub> 1 mL
- 100x metal salts\* 10 mL
- 50x 5052 20 mL
- 20x NPS 50 mL

\* Metal salts was prepared as in HdeB minimal media (section 8.5.1). All the solutions were added in the order given to avoid precipitation. MgSO<sub>4</sub> was filter sterilized.

**9.5.3 Solutions**

Proteinase K – 1 mg lyophilized powder was dissolved in 1 mM Tris (pH 8.0) and 1.5 mM calcium acetate and stored at -20°C.

---

Alkaline Lysis Solution I

- 1 M Glucose 5 mL
- 1 M Tris-Cl (pH 8.0) 2.5 mL
- 0.5 M EDTA (pH 8.0) 2 mL

Alkaline Lysis Solution I was made up to 100 mL with deionized water, autoclaved for 30 minutes at 121°C on a liquid cycle and stored at 4°C.

Alkaline Lysis Solution II

- 0.2 M NaOH
- 1% SDS (w/v)

Alkaline Lysis Solution II was always freshly prepared before use.

Alkaline Lysis Solution III

- 5 M Potassium acetate 60.0 mL
- Glacial Acetic Acid 11.5 mL
- H<sub>2</sub>O 28.5 mL

Alkaline Lysis Solution III comprised 3 M potassium and 5 M acetate. This was prepared via autoclaved for 30 minutes at 121°C on a liquid cycle and stored at 4°C.

Destain

- Methanol 400 mL
- Acetic acid 100 mL

Destain was made up to 1 L with with deionized and stored at RT.

---

### Coomassie Stain

- Coomassie Brilliant Blue      1.25 g
- Methanol                              223 mL
- Acetic acid                            45 mL

Coomassie stain was made up to 400 mL with deionized and stored at RT. Brilliant blue was dissolved in methanol before the addition of acetic acid and water.

### **9.5.4 Buffers**

All buffers were kept at RT

#### 1 M Tris-HCl

- Tris base                              121.1 g

Tris base was dissolved in 800 mL deionized water, the pH was adjusted by adding concentrated HCl and the volume made up to 1 L.

#### 0.5 M EDTA (pH 8.0)

- EDTA                                    186.2 g

EDTA was dissolved in 800 mL deionized water, the pH was adjusted with 10 M NaOH with vigorous stirring with a magnetic stirrer and the volume made up to 1 L.

#### 10x Tris EDTA (TE) buffer

- 100 mM Tris-Cl (pH 7.4)
- 10 mM EDTA (pH 8.0)

This was sterilized by autoclaving for 30 minutes at 121°C on a liquid cycle.

---

10x Phosphate-buffered Saline (PBS)

- NaCl 80 g
- KCl 2 g
- Na<sub>2</sub>HPO<sub>4</sub> 14.4 g
- KH<sub>2</sub>PO<sub>4</sub> 2.4 g

This was dissolved in 800 mL water, the pH was adjusted to 7.4 by adding concentrated HCl and the volume made up to 1 L. The solution was sterilized by autoclaving for 30 minutes at 121°C on a liquid cycle. The 10x stocks of PBS was diluted to 1x PBS with ddH<sub>2</sub>O before use.

10x PBST

- 10x PBS (as described above)
- 1% Tween 80

10x stocks of PBST was diluted to 1x PBST with ddH<sub>2</sub>O before use.

50x TAE buffer

- Tris base 242 g
- Glacial acetic acid 57.1 mL
- 0.5 M EDTA (pH 8.0) 100 mL

This was made up to 1 L with ddH<sub>2</sub>O, stored at RT and diluted down to 1x with ddH<sub>2</sub>O before use.

10x Transfer Buffer

- Tris base 60.4 g
- Glycine 288 g

This was made up to 2 L with ddH<sub>2</sub>O, stored at RT and diluted down to 1x with ddH<sub>2</sub>O and 10% methanol (v/v) before use.



10x SDS Running buffer

- |             |       |
|-------------|-------|
| ○ Tris base | 0.2 g |
| ○ Glycine   | 44g   |
| ○ SDS       | 1 g   |

This was made up to 1 L with ddH<sub>2</sub>O, stored at RT and diluted down to 1x with ddH<sub>2</sub>O before use.

2X Sample Buffer

- |                         |        |
|-------------------------|--------|
| ○ 0.5 M Tris-HCl pH 6.8 | 2.5 mL |
| ○ Glycerol              | 2 mL   |
| ○ 10% SDS (w/v)         | 4 mL   |
| ○ 2-mercaptoethanol     | 1 mL   |
| ○ 1.5% Bromophenol Blue | 1 mL   |

This was made up to 20 mL with ddH<sub>2</sub>O. The solution was mixed with a stir bar and stored at -20°C as 1 mL aliquots in 1.5 mL microfuge tubes.

**9.5.5 SDS gels**12% separating gel (makes 50 mL)

- |                         |          |
|-------------------------|----------|
| ○ 40% acrylamide        | 14.6 mL  |
| ○ 2% bis-acrylamide     | 8 mL     |
| ○ 1.5 M Tris-HCl pH 8.8 | 12.5mL   |
| ○ 10% SDS (w/v)         | 0.5 mL   |
| ○ ddH <sub>2</sub> O    | 14.15 mL |

for 5 mL:

- |           |          |
|-----------|----------|
| ○ 10% APS | 0.05 mL  |
| ○ TEMED   | 0.006 mL |

---

4% stacking gel (makes 25 mL)

- |                       |         |
|-----------------------|---------|
| ○ 40% acrylamide      | 2.4 mL  |
| ○ 2% bis-acrylamide   | 1.3 mL  |
| ○ 1 M Tris-HCl pH 6.8 | 3.15 mL |
| ○ 10% SDS (w/v)       | 2.5 mL  |
| ○ ddH <sub>2</sub> O  | 15.4 mL |

for 2.5 mL

- |           |           |
|-----------|-----------|
| ○ 10% APS | 0.025 mL  |
| ○ TEMED   | 0.0042 mL |

Both gels were made with all reagents mixed together except APS and TEMED and kept at 4°C. APS and TEMED were added (as indicated above) just before the gels were poured.

