THE INFLUENCE OF SMALL RNAs ON THE PHYSIOLOGY OF *MYCOBACTERIUM TUBERCULOSIS*

By

Tawanda Kennedy Zvinairo

17066786

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Medical Sciences (Molecular Biology) at Stellenbosch University

Department of Biomedical Sciences,
Faculty of Medicine and Health Sciences,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

Promoter: Prof TC Victor

Co-Promoters: Dr Lynthia Paul and Dr Elizabeth Streicher

December 2015

DECLARATION

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

December 2015

SUMMARY

The role of bacterial small RNA (sRNA), *i.e.* RNA species between 50-500bp in size, in virulence, pathogenesis and drug resistance is gaining interest. In some bacterial species, it had been shown to play a crucial role in bacterial transcriptional and post-transcriptional regulation. sRNAs from various pathogenic bacteria were shown to modulate bacterial responses to the host and environment. In *Mycobacterium tuberculosis*, the causative agent of tuberculosis, more than 1000 sRNA species have been identified already; but the role of these sRNA in pathogenesis, virulence and stress responses is not well studied.

Central dogma suggests that drug resistance in *M. tuberculosis* is associated with mutations in specific genes. However, a number of clinical drug resistant isolates do not harbour mutations in these genes, implicating other factors such as unknown mutations, as well as altered regulation of these resistance genes. Prediction of resistance, using molecular methods, can therefore be inaccurate in cases where known mutations are absent. In cases where known drug-resistance associated mutations are absent, mutations in other genes that regulate such resistance-associated genes might influence drug resistance. Growing evidence, in other bacteria and *M. tuberculosis*, hints at a role for mutations in intergenic regions and sRNAs species to play a role in bacterial growth and drug sensitivity. In light of this we hypothesised that mutations in sequences encoding sRNA or in sRNA target sequences influence the phenotype of *M. tuberculosis* clinical isolates.

Using previously identified sRNA genes; we screened a genomic bank of clinical *M. tuberculosis* isolates for the presence of mutations in these sRNA encoding genes. A large number of isolates showed mutations in genes encoding for sRNAs. Furthermore, over-expression of sRNA using the plasmid pMV306 in *Mycobacterium smegmatis* showed

differences in growth indicating that the presence of the extra copies of the three sRNA (mcr3, ASpks and mpr6) had a phenotypic effect on the bacterium. Overexpressed sRNAs did not affect the bacterial drug resistance phenotypes, although this requires further investigation before concluding the effect of sRNAs on drug resistance. We successfully modified a method to extract and purify sRNAs from *Mycobacterium* species, clean enough to perform Real Time Polymerase Chain Reaction even with small amounts. However challenges were faced in terms of quantification. Another challenge that still remains is obtaining reference genes specifically for sRNAs as we currently have none.

OPSOMMING

Die rol van klein ribonukleïnsure (m.a.w RNS spesies van ongeveer 50-250bp in grootte) in bakteriële virulensie, patogenese en antibiotika weerstandigheid word al hoe meer bevraagteken. 'n Rol vir hierdie nukleinsure in transkripsie en post-transkripsie regulering was voorheen gewys in verskeie bakteriële studies, waar dit gedemonstreer was dat hierdie RNA spesies n rol speel vir die bakterieë om aan te pas in die gasheer se omgewing [1–3]. Meer as 1000 klein RNS spesies is voorheen in *Mycobacterium tuberculosis* (die bakterie wat tuberkulosis veroorsaak) geïdentifiseer, maar die rol van hierdie RNA in patogenese, virulensie en stress reaksies is nie bekend nie.

Antibiotika weerstandigheid in *M. tuberculosis* word tans geassosieer met mutasies in spesifieke gene. Daar is wel n aantal weerstandige isolate waar hierdie bekende mutasies heeltemal afwesig is, wat suggereer dat ander rolspelers aanleiding kan gee to middelweerstandigheid. Byvoorbeeld, veranderde regulering van transkripsie patrone van gene (wat n bekende rol in weerstandigheid het) mag ook aanleiding gee tot weerstandigheid, maar sulke alternatiewe meganismes is nog nie goed ondersoek in die bakterium nie. Dis belangrik om al die rolspelers te identifiseer, want bestaande molekulere diagnostiese tegnieke fokus slegs op bekende gene; dus sal weerstandigheid gemis word in isolate waar bekende mutasies afwesig is en slegs molekulere tegnieke gebruik word. Die potensiële assosiasie van klein RNS in tuberkulose antibiotika weerstandigheid is voorheen in n paar studies gemaak. In lig van hierdie studies, is dit voorspel dat mutasies in klein RNA kan aanleiding gee tot verandering in die sensitiwiteit teenoor antibiotika in *M. tuberculosis*.

Vir hierdie studie het ons n genoom bank, wat bestaan uit individuele genome van kliniese *M.tuberculosis* isolate, geanaliseer vir die teenwoordigheid van mutasies in klein RNS. Daar

was spesifiek gefokus op die klein RNS spesies wat in vorige studies met antibiotika weerstandigheid geassosieer was. Hierdie bio-informatiese analise het mutasies in klein RNS spesies in n groot aantal weerstandige stamme geïdentifiseer. Hierdie mutasies was nie in sensitiewe isolate gevind nie, Om die rol van spesieke RNS spesies te ondersoek, was rekombinante plasmiede geskep wat bestaan het uit spesifieke klein RNS spesies van *M. tuberculosis* en die plasmied pMV306. Hierdie rekombinante was getransformeer in *Mycobacterium smegmatis*. Die teenwoordigheid van hierdie *M. tuberculosis* klein RNS kopieë in *M.smegmatis* het n negatiewe impak gehad op groei, en dui aan dat hierdie RNA spesies, naamlik mcr3, ASpks and mpr6, n potensiele belangrike rol het in die fenotipe van mikobakterieë het. Die ekstra kopieë het nie veranderinge veroorsaak in sensitiwiteit van *M.smegmatis* teenoor die antibiotika moksifloksasien en kanamisien nie, hoewel meer studies gedoen moet word voordat definitiewe konklusies gemaak kan word.

In die finale deel van die studie, is n metode ontwerp om klein RNS op n makliker,vinner manier te isoleer van mikobakterieë. Hierdie metode was suksesvol aangewend om DNA-vry, hoë kwaliteit RNS, beide groter RNA en klein RNS spesies te isoleer. Die klein RNS was goeie kwaliteit, DNA-vry en kon omskep word in DNA met retrotranskripsie. Laasgenoemde DNA kon ook gebruik word in verder polymerase kettingreaksies. Dit het dus potential vir kwantitatiewe studies om die regulering van klein RNS te studeer.

ACKNOWLEDGEMENTS

I would like to acknowledge and thank the following, without which this work would not have been achieved:

- God for the strength and guidance
- Dr Lynthia Paul (Supervisor), Dr Elizabeth Streicher (co-supervisor) and Prof Tommie Victor (promoter)for their support, advice and guidance throughout this study
- My colleagues and friends within the department
- My family and friends for their support
- The National Research Foundation and the Department of Biomedical Sciences for financial support.

LIST OF ABBREVIATIONS

sRNA	Small Ribonucleic Acid
M. tb	*Mycobacterium tuberculosis*
TB	Tuberculosis
DR	Drug resistant
DR-TB	Drug resistant tuberculosis
MDR	Multi-Drug Resistant
XDR	Extensively Drug Resistant
PCR	Polymerase Chain Reaction
DNA	Deoxyribonucleic Acid
miRNA	Micro RNA
E.coli	*Escherichia coli*
tRNA	Total RNA
pri-miRNA	Primary micro RNA
pre-miRNA	Precursor miRNA
asRNA	Antisense RNA
snRNPs	Small nuclear ribonucleoproteins
RBPs	RNA binding proteins
PAP 1	Pol (A) polymerase 1
PNP	Poly nucleotide phosphorylase
RNase E	Ribonuclease E
rRNA	Ribosomal RNA
snRNA/U-RNA	small nuclear RNA
snoRNA	small nucleolar RNA
mRNA	Messenger RNA

RBS	Ribosomal Binding Site
BCG	Bacillus Calmette–Guérin
HIV	Human Immuno-deficiency Virus
SNP	Single-Nucleotide Polymorphisms
MIC	Minimum Inhibitory Concentration
IGR	Intergenic region
EMB	Ethambutol
RIF	Rifampicin
STR	Streptomycin
OFX	Ofloxacin
ETH	Ethionamide
KAN	Kanamycin
CPM	Capreomycin
INH	Isoniazid
PNA	Peptide-nucleic acids
ADC	Albumin dextrose catalase
BSA	Bovine serum albumin
MOPS	3-(N-morpholino) propane Sulphonic Acid
TAE	Tris base, acetic acid and EDTA Buffer
dNTP	Deoxynucleotide
UV	Ultra Violet
ZN	Ziehl Neelsen
WT	wild type
µl	microliters
LB	Luria broth

TetR	Tetracycline repressor
TetO	Tetracycline operator
M. smegmatis	*Mycobacterium smegmatis*
RTM	Room temperature
DNase	Deoxyribonuclease
qPCR	Quantitative PCR
cDNA	Complementary DNA

TABLE OF CONTENTS

CHAPTER 1

GENERAL INTRODUCTION

Background, Problem statement, aims and hypothesis

1.1 BACKGROUND:

Mycobacterium tuberculosis (hereafter referred to as *M. tb*) is the causative agent of tuberculosis (TB). South Africa has the third-largest TB burden in the world with a reported incidence of 948 cases per 100 000 population annually[1]. The evolution of drug resistant (DR) strains of *M. tb* is a major hindrance to the treatment and eradication of TB. Not only must new drugs be developed as a result of the old ones becoming ineffective, but measures should be implemented to ensure more effective use of the existing and the prospective drugs. Treatment options and outcomes of DR-TB are not optimised and patients with contagious and highly resistant TB (Multi-Drug Resistant - MDR and Extensively Drug resistant - XDR) strains even being sent home, mainly because they have exhausted all available treatment options [2,3]. These therapeutically destitute persons, as well as an overburdened health care system that cannot cope with the number of TB patients requiring hospitalisation further promotes community-based spread of the disease.

DR in *M. tb* is associated with mutations in specific genes, but more studies to characterise the regulation of DR and resistance genes are needed. Current molecular diagnostic methods to screen for drug resistance include only known targets. Prediction of resistance with the aid of only molecular methods can therefore be misleading in cases where known mutations are absent. More studies are therefore needed to investigate other causative bacterial factors such as unknown mutations, and/or altered regulation of these resistance genes. Understanding regulation of these genes could be a key factor in curbing DR-TB. One aspect of these regulatory factors includes gene products such as regulatory proteins and/or small RNA.

Small RNA (sRNA) species have been described in various pathogenic bacteria and were shown to modulate bacterial responses to the host and to the environment [4]. Understanding the role played by sRNAs in gene regulation could broaden our view of the mechanisms involved in TB infection, particularly the poorly understood mechanisms involved in drug resistance and

latency. This will also improve diagnosis of drug resistance and the development of new and more efficient drug therapies for the disease. A recent study found significant association between the presence of mutations in sRNA genes and intergenic regions with drug resistance in *M. tb*. Although functional studies are lacking, such a significant association suggests that these mutations potentially play a role either in drug resistance itself, or the evolution thereof, or potentially could aid the survival of multidrug-resistant bacteria where various genes are mutated and therefore render functionally impaired products.

1.2 PROBLEM STATEMENT:

Knowledge of the all the mechanisms involved in TB drug resistance is lacking, especially in isolates that lack known mutations associated with drug resistance. Growing evidence suggests a role for mutations in intergenic regions and sRNAs in gene regulation, but more studies are needed to understand its role in DR-TB. Binding of a sRNA to a target at post-transcriptional level can either promote or hinder translation, possibly of genes associated with drug resistance. More studies therefore need to be done to provide an insight into how small RNA influences drug metabolisms and phenotype in *M. tb*.

1.3 HYPOTHESIS:

Mutations in genes encoding small RNA affect regulatory pathways that may alter the phenotype of *M. tb* clinical isolates, including their sensitivity to anti-tuberculosis drugs.

1.4 OVERALL AIM:

To characterise the role of sRNA in *M. tb*, with the main focus on their potential role in drug resistance.

1.5 CLINICAL IMPLICATIONS

- To broaden our understanding of drug resistance mechanisms in *M. tb* resulting in the improvement of diagnostic methods and the development of more efficient drug therapies against tuberculosis.

1.6 OBJECTIVES AND EXPERIMENTAL APPROACH

Objective 1: To establish if single nucleotide polymorphisms (SNPs) occur in sRNA genes in clinical isolates with various phenotypes.

Method:

i) Literature search and bio-informatics analysis of published sRNAs from *M. tb* in order to select candidates.

ii) Screening for mutations in candidate sRNAs in the genomes of a large set of clinical isolates of different strain lineages. Additionally, whole genome sequences from published *M. tb* genomes will be examined for the presence of SNPs in the sRNA genes.

Objective 2: To characterize the expression of sRNAs identified in objective 1 (mcr3 and ASpks) after exposure of *M. tb* to anti-TB drugs (kanamycin, moxifloxacin).

Method:

i) A method to extract sRNAs from mycobacteria, without the use of cumbersome agarose gel purification methods, will be devised and optimized using a combination of known commercial available kits; the method will be optimized in *M. smegmatis*.

ii) Actively growing, log phase cultures of *M. tb* will be exposed to TB drugs and sRNAs will be extracted using the method in (i).

iii) The levels of specific sRNAs in antibiotic exposed and unexposed samples will be compared using quantitative reverse transcription PCR.

Objective 3: To characterise the effect of sRNAs on the phenotype of *M. smegmatis.*

Method:

i) Candidate genes will be cloned and expressed in *M. smegmatis*, using plasmid pMV306, as well as in a plasmid construct with the TET on/off controllable promoter.

ii) Transformants will be characterised using growth curves to assess growth phenotype and drug sensitivity tests.

CHAPTER 2

LITERATURE REVIEW

BACTERIAL SMALL RNAs, THEIR DISCOVERY, ROLE AND POTENTIAL USE AGAINST PATHOGENS SUCH AS MYCOBACTERIUM TUBERCULOSIS

2.1 INTRODUCTION

Prokaryotes, eukaryotes and viruses all have evolved numerous strategies to control gene expression, which allows for increased versatility and adaptability to changes in the environment. Gene regulation can occur at the level of DNA transcription, post-translationally and post-translationally. Effectors such as small non-coding RNAs (sRNAs) can provide additional knowledge on the mechanisms of gene regulation. The discovery of sRNAs as effectors of post-transcriptional regulation adds to our understanding of the complex processes involved in gene regulation. However, large knowledge gaps still exist regarding the roles of sRNAs in bacteria particularly in Mycobacterial species.

In this review we summarize some of the key strategies employed by bacteria in regulating gene expression via sRNA molecules. We detail gene regulation by sRNAs, with an emphasis on *Mycobacterium tuberculosis (M. tb)*, the pathogen responsible for the global tuberculosis (TB) epidemic. We also explore the possibility that sRNAs can contribute to drug resistance in clinically relevant pathogens, such as *M. tb*, and discuss the potential to use sRNAs or their targets in antimicrobial therapy.

2.2 THE DISCOVERY AND SIGNIFICANCE OF SMALL RNA MOLECULES

sRNAs range in length from 50 to 500 nucleotides [5]. The significance of sRNAs was first appreciated in plants, fungi, protozoa and metazoan animals in which they were referred to as micro RNA (miRNA) and were shown to function in transcriptional and post-transcriptional gene regulation [6–8]. miRNA was comprehensively described in 1993[6], but did not receive widespread recognition as a distinct component of gene regulation with conserved functionality until the 21st century.

sRNAs are important regulators because of their small size which translates to a quicker response to stimuli. Their small size means they can be rapidly synthesised in large quantities, easily locate their targets and are quickly degraded [9]. This allows the bacteria to rapidly and

efficiently integrate and respond to multiple environmental changes in a chronologically efficient way.

Although discoveries of sRNAs date back to the 1980s, their significance and role as bacterial regulatory components were only appreciated more recently. An understanding of how sRNAs function in gene regulation provides researchers with knowledge of how bacteria control their biology/physiology [10] for instance during oxidative stress, phage development, bacterial virulence, developmental control and is also important in understanding drug resistance and persistence [11]. One of the first significant discoveries of a sRNA in bacteria with regulatory activities in bacteria was made in 1981, with the description of an approximately 108 nucleotide sRNA linked to the inhibition of *ColE1* RNA primer formation by an antisense base pairing mechanism of a plasmid-specific RNA [12,13]. In 1983 another sRNA, approximately 70 nucleotides long, was described and linked to inhibition of transposase translation by pairing with the transposase mRNA [14]. The first chromosomally encoded sRNA, *MicF (approximately 174 nucleotides long), was discovered in 1984 in *Escherichia coli*. It blocks translation of a major outer membrane porin *OmpF* [15]. However some of the early sRNAs turned out to be proteins commonly referred to as transcription factors and this halted research into sRNA.

The recent rediscovery of sRNAs as regulatory elements has resurrected the interest into sRNAs as regulators. Subsequently, many of the already described sRNAs have been found to be conserved in closely related bacterial species [16,17]. To date, several sRNA molecules have been discovered in *E. coli* and other bacteria [18–25].

**** Small RNAs are not written in italicised format as do genes.***

2.3 METHODOLOGY TO IDENTIFY SMALL RNAs

Technical advances offer the potential to comprehensively identify further sRNAs, which will advance our understanding of these important molecules. Experimental approaches such as direct labelling and sequencing initiated the discovery of sRNAs. An example of direct labelling and sequencing involves metabolically labelling total RNA (tRNA) derived from bacteria, followed by gel fractionation with P-orthophosphate. Putative sRNAs would be visible as small single bands on the gels. This method is advantageous particularly in identifying sRNAs that are up-regulated in specific stress conditions (reviewed by Aluvia et al) [26].

Due to their relatively small sizes, sRNAs are difficult to identify by genetic screening and other molecular methods. However, given the growing availability of whole genome sequence data supported by experimental verification, it is now possible to identify a large number of sRNAs that could have been missed by more targeted genetic screening methods alone.

Various computational prediction methodologies are used to identify sRNAs. *Comparative genomics*, one of the most frequently used methods, involves exploiting the similarities seen between genome structures/sequences and using these to find and predict function of new sRNAs to a significant accuracy [27]. "Orphan" transcriptional signals such as promoters and rho-independent terminators found in the intergenic regions are some of the parameters used in designing sRNA computational identification tools [28]. Some sRNAs are found in multiple species, and in some cases their function is conserved. For example, the sRNA MicF exerts similar effects on *ompF* in *Yersinia* and *E.coli*, that is, inhibition of translation of the *ompF* mRNA [29]. This similarity in function could be exploited to predict sRNA function in different bacteria.

The *ab initio* method is another widely used computational method which takes advantage of the occurrence of similarities in protein/RNA-coding nucleotide sequences termed "signals".

Some of these "signals" include promoter sequences, transcription factor binding sites, specific sRNA structures, di/tri-nucleotide preferences, existence of preferred sequences such as GC content. RNAGENiE is an example of a tool that uses the *ab initio* method in which it predicts sRNAs by identifying commonly shared structures such as double helices, UNCG tetra loops, GNRA tetra loops, tetra loop receptors, uridine turns, adenosine platforms etc. [30,31]. Other tools include the Glimmer system [32–34] and GeneMark [35–37]. The *ab initio* method is less suited for more complex (e.g. eukaryotic) organisms.

The biogenesis of sRNA transcripts is not well understood. Studies of eukaryotic sRNAs (miRNAs) have provided evidence of secondary processing of miRNAs after transcription. This includes formation of a primary miRNA (pri-miRNA) which is then cleavage by protein assisted enzymes (Microprocessor complexes) into precursor miRNA (pre-miRNA). The pre-miRNA is transported into the cytoplasm where further cleaving occurs to produce mature, functional miRNA [38–41]. There is currently no evidence showing similar or related mechanisms in bacterial sRNAs. This should be further investigated in order to provide further understanding of gene regulation leading to new drug targets and/or non-chemical but rather RNA based therapeutics.

2.4 CLASSIFICATION OF SMALL RNAs

Small RNA molecules have been associated with regulation of bacterial gene expression both at transcript and protein level. Depending on the nature of the molecule being targeted, sRNAs can be divided into two broad classes. Those that base pair with their target mRNA, often termed base-pairing or antisense RNAs, form sRNA-mRNA complexes that modify translation and stability of target mRNA [42]. The second class consists of sRNAs that bind and modify the function of metabolic proteins [43]including regulatory proteins by antagonizing their functions [44]. The base pairing or antisense RNA (asRNA) have been the most and extensively studied

group of sRNAs and have been linked to regulatory responses as a result of environmental changes.

The base pairing group of prokaryotic sRNAs can be further grouped into two sub-classes. The first of these, *cis-encoded* sRNA species, are encoded in the same DNA region but on the opposite strand to the target DNA coding strand as discrete molecules. They have extensive potential to base pair with their target due to full base pair complementarity and they are diffusible molecules. The second class, *trans-encoded* sRNAs are encoded on another chromosomal location and are partially complementary to their target mRNAs. Unlike *cis*-encoded sRNAs, trans-encoded sRNAs exert their function in "trans" as diffusible molecules and can have multiple targets [44,45]. These sRNAs have been associated mostly with negative regulation of translation and/or mRNA stability [5].

Another broad classification categorizes sRNAs according to those that interact with the Sm-like chaperone protein, Hfq, and those that do not. Sm proteins form part of specific small nuclear riboproteins (nnRNPs) involved in the processing of pre-mRNAs to mature mRNAs; as a major part of the spliceosome.. The Hfq chaperone molecule forms complexes with the A/U-rich regions of sRNAs providing a stable platform for sRNA-mRNA interaction [46].

2.4 SMALL RNA-mRNA INTERACTIONS ARE ENHANCED BY NUCLEIC ACID BINDING PROTEINS LIKE HFQ

Prokaryotes have acquired a cascade of RNA binding proteins (RBPs) as part of post-transcriptional regulation [47]. Hfq is one such example, and is an abundant bacterial RNA binding protein [48], which is implicated in various crucial regulatory processes. Hfq in *E.coli* was originally shown to regulate its own synthesis by binding to the 5' UTR region of its own mRNA, thereby inhibiting the formation of a translation initiation complex (**Figure 1**). This protein assists in the antisense interaction of mRNAs and sRNAs by providing a stable platform

for the association to occur [46]. The Hfq mediated association of the sRNA and the target mRNA occurs via two RNA binding sites on Hfq, one site binds sRNA whilst the other site connects to the mRNA target. Hfq thus acts as a chaperone that aids the binding of sRNA molecules to their target mRNA molecules [49]. This property can be exploited to isolate sRNAs from various microorganisms, including mycobacteria [50].

The Hfq chaperone protein is mostly required by trans-encoded sRNAs to stabilise sRNA-target mRNA complex formation probably because of the short, imperfect base pairing interactions with target mRNA associated with these sRNAs [51]. The Hfq chaperone proteins not only stabilize sRNA-mRNA complexes but also can protect them from degradation [52]. However in *E.coli* it was found that Hfq bound to sRNA and its target, affecting the stability of mRNA by interfering with RNase E degradation [53,54]. Several research studies have shown associations between Hfq and proteins such as Pol(A) polymerase 1 (PAP 1), Poly nucleotide phosphorylase (PNP) and RNase E narrowing the possible functions of the sRNAs involved [55–58]. Hfq-dependent sRNA regulation has also been reported in the gram-positive bacterium *Listeria monocytogenes* [59,60].

Hfq protein has been implicated in the regulation of virulence in gram negative bacteria, such as *Klebsiella pneumoniae* [61,62]. It has also been shown to regulate the σ-mediated general stress response in *E. coli* [63]. The latter study showed that the Hfq molecule regulates the σ^{32}-mediated cytoplasmic heat shock response and secondly that it is fundamental for long-term adaptation of σ^{32} to chronic chaperone overexpression [63]. Hfq also acts as a host factor required for Qβ phage replication [48,64](involves the Qβ replicase complex used by viruses to hijack host translation elongation factors in order for it to multiply[65]).

It has been suggested that sRNAs are important role-players in stress-induced adaptive responses and are therefore a key part of bacterial pathogenesis [66]. A complete understanding of how sRNAs function as regulators of gene expression is lacking, although a few studies have

characterized sRNA candidates in terms of their target and effect. These studies have shown that the majority of sRNAs are antisense RNAs that function through base pairing mechanisms[67]. The most studied sRNAs, the *"trans*-encoded" sRNAs, are transcribed from intergenic regions of DNA, and pair with imperfect complementarity to their targets. As such these sRNAs are sometimes associated with having multiple targets [44] and it has been hypothesised that due to their weak association with target mRNA these sRNAs often require the aid of the RBP, Hfq [11]. Some of the sRNAs that have been shown to associate with the Hfq chaperone are listed in **Table 2.1**.

Table 2.1: sRNAs that have been shown to associate with the Hfq chaperone.

sRNA	Target	Reference
DicF	Inhibits the cell division gene *ftsZ* in *E.coli*	68
DsrA	Regulates transcription, by means of silencing H-NS nucleoid-associated protein, and promotes translation efficiency of the stress sigma factor *rpoS* in *E.coli*. Genes affected: *rbsD, argR, ilvI, hns* and *rpoS.*	69,70,71,72,73
FnrS	Regulates genes mostly associated with aerobic metabolism/ response to oxidative stress in *E.coli.* Genes affected: *adhP, cydD, mqo*	74,75
GadY	Stabilizes *gadX*, and acid response transcriptional regulator, mRNA in *E.coli.*	76
GcvB	This sRNA is found in a range of bacteria. Genes been shown to be targets: *Oppa* and *DppA* (transport oligopeptides and dipeptides respectively), *gltL, argT, stm, livJ, brnQ, sstT and cycA* (involved in amino acid uptake), *ilvC, gdhA, thrL* and *sera* (amino acid biosynthesis) and *PhoPQ* (magnesium homeostasis)	77,78,79,80,81
IsrJ	Affects the efficiency in which the virulence-associated effector proteins translocate into non-phagocytic cells in *Salmonella typhimurium*	82
MicA/SraD	Targets the *OmpA* hindering transcription by occluding the ribosomal binding site in *E.coli*.	83,84
MicC	Regulates the expression level of the porin protein ompC by interacting with the *ompC* mRNA in *E.coli*.	85
MicF	Is involved in regulation of stress response by controlling the of the outer membrane porin gene *ompF* in *E.coli* and related bacteria.	86
OmrA/OmrB/RygA/RygB /SraE	Is a family of sRNAs that negatively regulate several outer membrane protein genes: *cirA, CsgD, fecA, fepA* and *ompT.* Discovered in *E.coli*.	87,88,89

OxyS	Is involved in oxidative stress response in *E.coli*, it's been shown to regulate as many as 40 genes, *fhlA* being the best example.	90
Qrr	Involved in the regulation of quorum sensing in *Vibrio* species	91,92
RprA	Regulates translation of the sigma factor *rpoS* by occluding the ribosomal binding site. It also been shown to repress the protein coding genes *csgD and ydaM* (biofilm formation)	93,94,95
RybB	Increases the rate of degradation of *omp* mRNAs in response to stress.	96
RydC	Regulate the *yejABEF* mRNA that produces and ABC transporter protein.	97
RyeB	Identified in *E.coli*, *RyeB* is thought to work in a concerted manner with SraC/RyeA sRNA because the 2 sequences overlap	98,10
CyaR	*E.coli* derived, represses the porin OmpX in *Salmonella*	99
RyhB/SraI	Down regulates set of iron-dependant and iron-storing proteins	100,25
SgrS	Is activated during glucose-phosphate response and associated with intracellular accumulation of glucose-6-phosphate assisting cells in recovering from glucose-phosphate stress by repressing the *ptsG* mRNA translation.	101,102,103,104
Spot 42	Regulates the galactose operon by binding to the *galK* gene it also affects DNA polymerase 1	105,106
SraH/ArcZ/RyhA	In *Salmonella* it has been shown to regulate the expression of the protein involved in serine uptake, *sdaCB* and *tpx* (involved in oxidative response).	107,108
GlmZ/SraJ	Positively aids in the transcription of *GlmS* mRNA.	109,110,111,112,
SroB/MicM/RybC	A study implicates this sRNA in negative regulation of the outer membrane protein YbfM by binding to the *ybfM* mRNA. It also regulates the DpiA/DpiB two-component system.	113,114
SroC	Found in several Enterobacterial species and function is unknown	(130)

2.5 sRNAs AFFECT RNA STABILITY, RNASE ACTIVITY AND RNA MODIFICATION

Prokaryotic mRNA is less stable than eukaryotic mRNA, allowing prokaryotic cells to rapidly adapt to environmental changes [47]. In addition to transcriptional regulation, gene regulation can also occur at the level of control of RNA stability and degradation, both which affects translation processes. Part of this mRNA instability is attributed to the activity of endo-ribonucleases (RNases). RNases, some of which are part of toxin-antitoxin systems, cleave

RNA at specific sites. They are an important part of the gene regulation machinery operating at the post-transcriptional level and are involved in various aspects of RNA metabolism including processing, maturation and degradation of mRNA [116]. EndoRNases initialise RNA degradation by cleaving RNA into small fragments which are further degraded by exoRNases [47]. An example in *M. tb* is the *rnc*–encoded RNase III enzyme that functions in ribosomal RNA (rRNA) processing, mRNA maturation and degradation, small nuclear RNA (snRNA/U-RNA) and small nucleolar RNA (snoRNA) processing, and RNA interference [117].

M. tb RNase is homologous to *E.coli* RNases, though their mechanism of action differs. In *M. tb*, RNaseE digests a smaller portion of the A/U rich sequences of mRNA transcripts [118]. It also has a very slow mRNA turnover rate with a half-life of more than 9 minutes meaning that *M. tb* has a slower repression rate compared to other bacteria. Observations that support this hypothesis include the presence of stabilizing secondary structures on transcripts as well as the high GC content of mycobacteria [119]. This could in part explain the relatively slow growth of *M. tb* in comparison to other *Mycobacterium* species.

sRNAs play a role in the turnover of mRNA by guiding RNase E to its cleavage site and also by allosterically activating the enzyme. An example is MicC, a 109 nucleotide sRNA that regulates expression of the *Salmonella* outer membrane protein OmpD porin [120]. Such sRNAs most probably exist in *M. tb* given the existence of an RNase E homologue [118,121].

When sRNAs are bound to their target mRNAs, they either activate or inhibit translation depending on the position at which they are binding and they also affect stability of the mRNA [122]. This interaction is also influenced by chaperones such as Hfq as mentioned earlier. Most of these Hfq-assisted sRNAs have multiple targets and only a few of these targets have been linked to the corresponding sRNA [123].

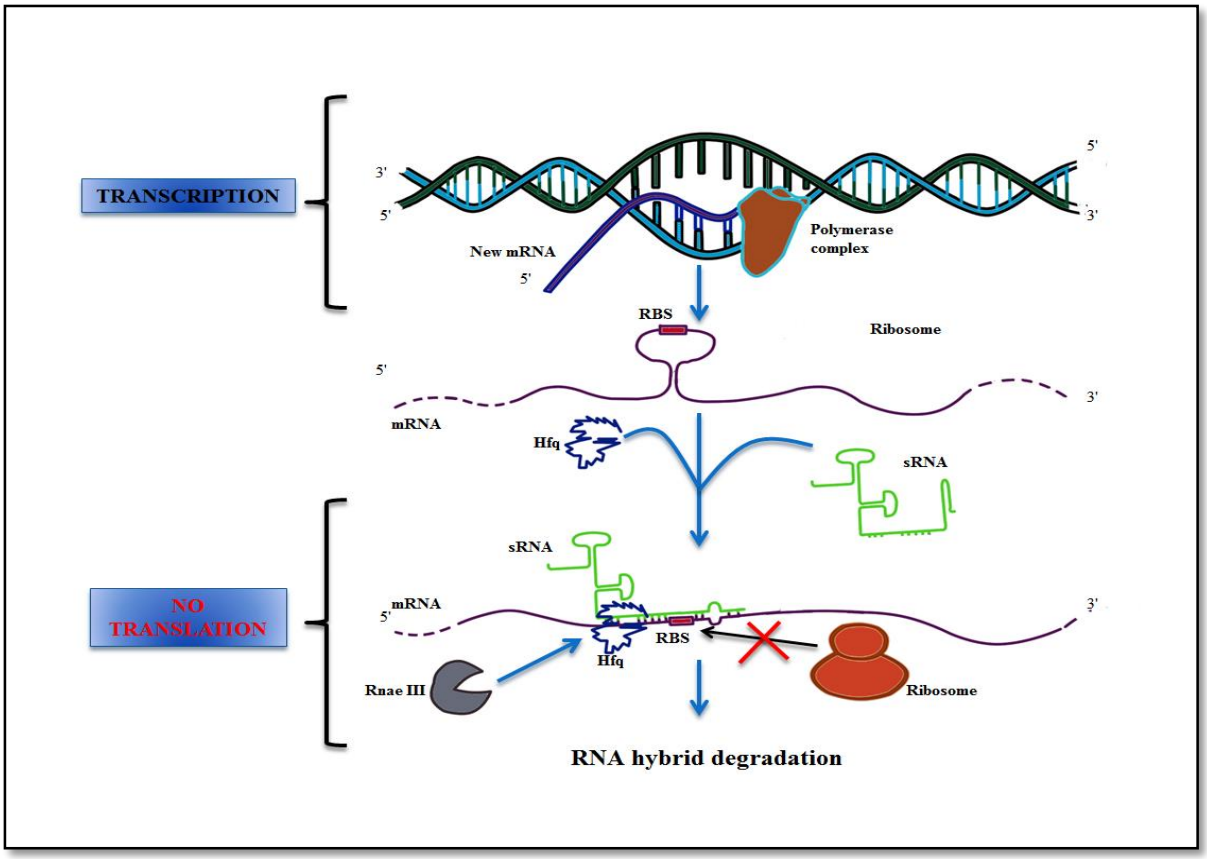

Figure 2.1 The Role of sRNA and Hfq in repressing translation of mRNA. sRNA and Hfq hybridizes to mRNA RBS, blocking ribosome and leading to mRNA degradation.

Interactions with the Hfq protein can affect sRNA target in a number of ways. In some cases binding of Hfq to target mRNA/sRNA opens up the hair pin secondary structure often formed by mRNA (**Figure 2.1)**. This facilitates the dynamics of sRNA-mRNA complex formation by de-sequestering occluded sequences necessary for complementation between a sRNA and its target mRNA thereby promoting translation [124]. This improves interactions between a sRNA and target mRNA through Hfq-Hfq interactions [56,125]. Degradation is another possible consequence of the Hfq/sRNA/mRNA complex [120]. This complex could be a Ribosomal Binding Site (RBS) occluding factor rendering the mRNA inactive and prone to RNase activity (**Figure 2.2**) [126]. In some cases the sRNA-mRNA complex results in increased RBS accessibility thereby increasing/promoting translation as shown in **Figure 2.2** [72].

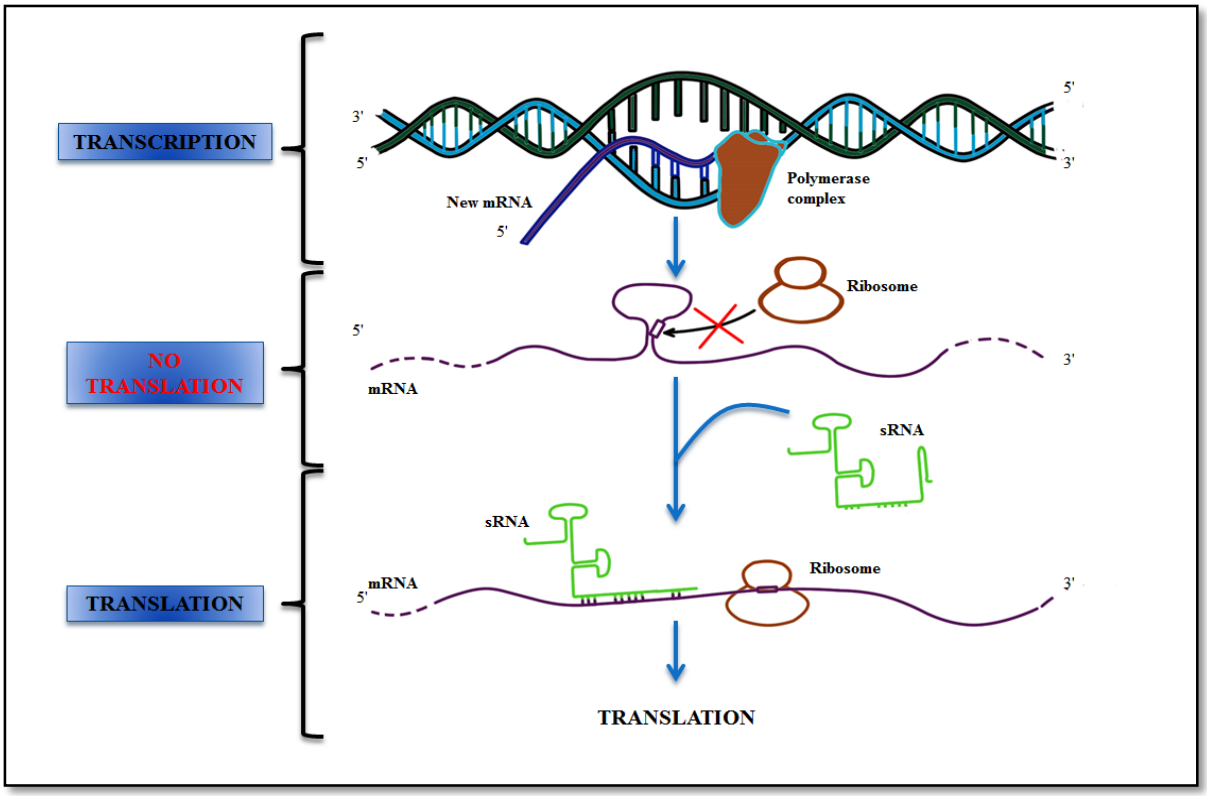

Figure 2.2 The role of sRNAs in promoting translation of mRNAs. mRNA forms a stem loop secondary structure around the ribosomal binding site (RBS) blocking the ribosome. The sRNA–mRNA hybrid formed opens up the mRNA secondary structure, exposing the RBS thereby promoting translation of the mRNA.

A large knowledge gap still exists regarding sRNAs regulation, signals for their synthesis and degradation and their importance in human pathogens. It is also not clear how the Hfq-sRNA-mRNA complexes are dissociated to pave way for degradation molecules like the RNases and hence more needs to be done to understand sRNAs as regulatory components.

It has been shown that during normal transcription, sRNAs are continuously degraded but degradation is halted upon addition of rifampicin and sRNAs [127]. These results therefore suggest that the binding of sRNAs to their targets is a signal for their degradation. Rifampicin is an anti-bacterial drug that inhibits RNA polymerase activity and consequently halts bacterial transcription [128]. It is a major part of TB therapy hence its effects on sRNAs deserve attention.

The sRNA RNAIII was shown not to require the aid of Hfq suggesting that some sRNA activity can proceed even in the absence of this molecule [129]. RNAIII is a 514 nucleotide sRNA with 14 stem-loop foldings. It was originally extracted from *Staphylococcus aureus*, where it is

implicated in controlling synthesis of virulence factors [130]. RNAIII has at least 3 known targets, *hla*, *spa*, and *rot* mRNAs, with *hla* encoding alpha-haemolysin which plays a role in the bacteria's pathogenesis processes [131]. In *hla* mRNA the 5' end is folded, resulting in occlusion of the RBS. RNAIII sRNA 5' binds to the *hla* 5' end of the mRNA, exposing the RBS and allows translation to take place. *spa* gene encodes for a 56 kDa MSCRAMM surface protein of *S. aureus*, it is also known as Protein A [132]. The formation of a RNAIII-*spa* mRNA hybrid has the opposite effect on translation. It base pairs with the *spa* mRNA at the RBS region, forming a complex that has a high affinity for degradation enzymes; thus promoting the degradation of the *spa* mRNA as illustrated by **Figure 2.3** [133]. RNAIII sequesters the *rot* mRNA in a similar manner as *spa* mRNA [134].

A unique and interesting sRNA mechanism of function without the aid of Hfq is one portrayed by CsrA (or RsmA) sRNAs. CrsA is a family of protein regulators negatively regulating translation in *E.coli* by binding and sequestering the RBS. The CsrA sRNAs form secondary structures (stem-loops) that mimic the mRNA RBS, thereby activating translation of the CrsA (RsmA) targets (**Figure 2.3**) [135].

It is of interest to note that in a study by Ramos *et al* they showed that even the Hfq post-transcriptional processes can also be regulated by sRNAs [136].

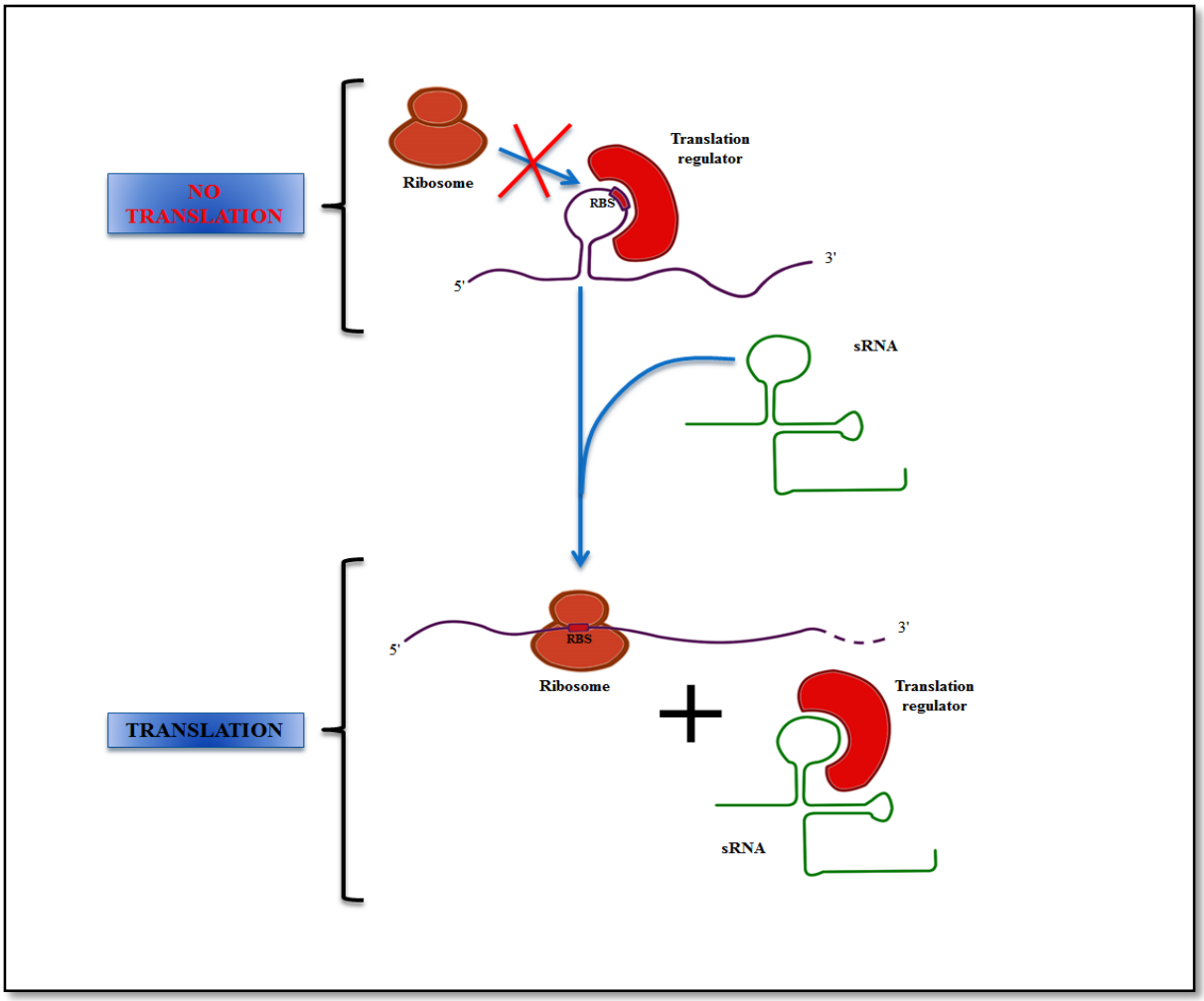

Figure 2.3: sRNA mediated activation of translation through its interaction with translation regulators. Translation regulators bind the RBS on mRNA making it inaccessible to ribosome binding; and blocking translation. Binding of sRNA to the translation regulator releases the mRNA enabling translation to occur.

2.6 SMALL RNAs AND THEIR ROLE IN *MYCOBACTERIUM* SPECIES

M. tb is one of the most prominent pathogenic bacterial species in the world, however a knowledge gap exists regarding the role of sRNAs in *M. tb* pathogenesis, virulence, stress responses and drug sensitivity. Understanding the mechanisms of regulation in *M. tb* by sRNAs will broaden our understanding of TB infection, particularly the poorly understood mechanisms involved in drug resistance, persistence and latency.

The existence of various sRNAs in *M. tb*, *Mycobacterium bovis* (BCG) and *Mycobacterium smegmatis* has been shown by a number of groups [42,66,137–140]. It was shown that many sRNAs in *M. tb* are stress-induced, suggestive of a role in adaptation to a hostile host environment [66].

Arnvig and Young showed that F6 significantly slowed down growth , B11 also resulted in slow bacterial growth with cells showing elongation narrowing the possible functions of B11 to cell wall synthesis, and/or cell division [66]. However Miotto and colleagues showed that most antisense sRNAs bind preferentially to genes implicated in two component systems and membrane activity, suggesting that sRNAs not only regulate stress responses but are also involved in the regulation of normal metabolic processes in *M. tb* [138]. DiChiara and co-workers identified 37 sRNAs in *M. bovis* BCG; 20 homolgoues of these were predicted and experimentally confirmed in *M. tb*, and 17 in *M. smegmatis*. Eight sRNAs were conserved only in BCG and *M. tb*, thus one can predict potential involvement in virulence, as *M. smegmatis* is considered non-pathogenic [137].

Li and colleagues showed that some sRNAs are growth-phase dependent. They identified 12 trans-encoded and 12 cis-encoded sRNAs in *M. smegmatis*, which they showed to be differentially expressed at exponential phase in comparison to stationary phase. This suggests that sRNAs also play a role in physiology and growth [50].

In a study by Tsai and colleagues, 17 sRNAs were identified in *M. smegmatis*, of which 9 homologues were found in *M. bovis* and 4 in *M. tb* [140]. The roles of these sRNAs need further characterisation.

2.7 CURRENT KNOWLEDGE OF sRNAs AND THEIR INFLUENCE IN *M.TUBERCULOSIS*

Drug resistance is one of the major obstacles in curbing the global TB pandemic. Not only do it require more resources which could have been used to treat drug susceptible TB but also successful treatment outcomes are low, particularly in HIV co-infected patients [141,142]. Drug resistance mechanisms are not comprehensively understood. A few genes of clinical significance have been associated with drug resistance and accepted popular dogma suggests that mutations in these genes give rise to drug resistance. However, there is evidence that the emergence of XDR-TB is the result of accumulating resistance associated with both known

and unknown single-nucleotide polymorphisms (SNPs) rather than of mutations in one or a few resistance associated genes. This might explain anomalies such as different drug Minimum Inhibitory Concentration (MIC) when the same DR mutations are present, as shown by. In this study nsSNPS in 72 genes, 28 intergenic regions (IGRs), 11 SNPs and 12 IGR SNPs (**Table 2.2**) were uniquely associated with drug-resistant strains of *M. tb*, and were absent in drug-sensitive isolates. These observations include both novel and known factors of drug resistance [143]. However it is not clear how these nsSNPs affects drug-sensitive *M. tb* and therefore wet bench experiments are required to test this hypothesis. It is important to note that some of the intergenic regions associated with drug resistance in this study encode for sRNAs [143] and in the recent view of sRNAs acting as regulatory elements this deserves more attention.

Table 2.2 IGR SNPs uniquely found in drug resistant *M. tuberculosis* isolates [143]

IGR	Flanking genes	Base change/mutation in IGR
Rv1482c-Rv1483	Hypothetical protein—fabG1	C→T
Rv1482c-Rv1483	Hypothetical protein—fabG1	C→T
Rv1482c-Rv1483	Hypothetical protein— fabG1	T→A
Rv2416c-Rv2417c	*eis*—hypothetical protein	C→T
Rv2427c-Rv2428	*proA—ahpC*	C→T
Rv2427c-Rv2428	*proA—ahpC*	G→A
Rv2754c-Rv2755c	*thyX—hsdS.1*	C→T
Rv2754c-Rv2755c	*thyX—hsdS.1*	G→A
Rv3185-Rv3186	transposase—transposase	T→A
Rv3185-Rv3186	transposase—transposase	T→A
Rv3793-Rv3794	**embC—embA*	G→C
Rv3793-Rv3794	**embC—embA*	C→AG

* *embA* and *embC* together with *embB* are part of gene cluster of *Mycobacterium tuberculosis* involved in resistance to ethambutol.

In other bacterial species there are sRNAs that have been associated with drug resistance. In a study by Yu and Schneiders, four sRNAs were linked to trigrecycline/tetracycline resistance in *Salmonella enterica serovar typhimiurium* [144]. Some *Staphylococcus* infections cannot be treated as a result of the emergence of glycopeptide resistant strains of *Staphylococcus aureus (S. aureus).* Howden and colleagues identified 409 putative sRNAs in *S. aureus* using RNAseq after exposure to four drugs (vancomycin, trigecycline/tetracycline, linezolinal and ceftobipile) [145]. Recently, the sRNA Sprx(RsaOR) has been shown to control a regulator that is involved in *S. aureus* resistance to glycopeptides [146]. These results are evidence that sRNAs can be involved in bacterial response to antibiotics and therefore play a role in anti-bacterial resistance. However, such in-depth studies lack in mycobacteria.

Some of the described Mycobacterial sRNAs are associated with known drug resistance genes. An example is the Mcr3 sRNA which is located upstream of the *M. tb rrs* gene, known to be involved in aminoglycoside resistance [147]. This sRNA is encoded by a region that includes one of the two promoters known to drive *rrs* transcription [137,139]. However, a functional link between this sRNA and aminoglycoside resistance has yet to be established.

What role does the genes flanking or co-transcribed with sRNAs play in *M. tb*? A number of studies have identified mutations in *pks12* in drug resistant isolates [148]. Whether it plays a role in drug resistance is however not clear. In *Mycobacterium* the gene *pks12* is involved in lipid metabolism particularly in the biosynthesis and translocation of cell wall surface exposed lipids controlling permeability. This gene has been associated with intrinsic resistance to a range of antibiotics in the *Mycobacterium avium* complex [149]. Mutations in the *pks12* gene were shown to induce drug susceptibility. It was also shown that the *pks12* homologue in *M. tb* H37RV also play a role in intrinsic drug resistance though with a narrow drug spectrum and moderate effect [149]. Matsunaga and colleagues also showed that deletion of *pks12* in *M. avium* increases drug susceptibility, an observation repeated in *M. bovis* BCG [150]. A clue as to how this gene might

influence susceptibility to drugs perhaps might be found in its coding region, wherein a sequence for a sRNA, ASpks, was found. Its functions are not known but given its coding region, which is within the pks12, it could have regulatory effects on this gene and ultimately play a role in intrinsic drug resistance.

A second sRNA is located between the genes *embC* and *embA* [42]. These genes together with *embB* are essential in *M. tb* as they are required in cell wall biosynthesis; they encode for the synthesis of arabinogalactan and lipoarabinomannan whose biosynthesis is the target of the anti-TB drug Ethambutol (EMB) [151].

More sRNAs directly associated with DR-genes exist and given that the full spectrum of DR-genes is not entirely known and that not all sRNAs have been identified and assigned function, various aspects of drug resistance mechanisms in *M. tb* remain unexplained. Elucidating the link between sRNAs such as Mcr3 and ASpks, and drug sensitivity could translate into novel therapeutic approaches.

Another observation is that base changes in the intergenic region (IGR) upstream of drug resistance genes results in low-level drug resistance. Examples of this are the IG regions upstream of *eis, inhA* and *katG*. Base changes in promoter upstream of the *eis* gene are known to be involved in low-level resistance to the second-line drug kanamycin, while mutations in IG regions upstream of *inhA* and *katG* promoter are known to result in low-level INH resistance. These three examples illustrate how mutations in intergenic regions could affect drug resistance in *M. tb*. It is therefore worthwhile to explore how mutations in other IGRs affect drug sensitivity. Another example is the mutation in IGR between *furA* and *katG*, described by Ando and co-workers, and which results in down-regulation of the *furA-katG* mRNA transcript. Ultimately this results in INH resistance [152]. While it is not known if sRNAs play a role in the regulation of these genes, these examples make it clear that mutations in

IGRs, which could encode sRNAs or sRNA targets, could alter drug sensitivity patterns in *M. tb*.

An alternative way in which knowledge of sRNAs could be exploited is suggested by the observation that the first line drug rifampicin affects sRNA degradation. As this drug is used as one of the first line combination regiment in TB therapy, it is important to know if sRNAs in TB are affected in a similar manner. The regulation of sRNA transcription, in response to antibiotic therapy had already been demonstrated in other pathogens such as *S. aureus*, thus it is plausible that the same is true in *M. tb* [145].

Another interesting observation, from Zhang's group, is that one gene or IGR could be associated with resistance to more than one anti-TB drug, while multiple genes/IGRs could be involved in resistance to one drug (**Table 2.3**) [143]. In the case of IGRs encoding sRNAs the latter agrees with the view that sRNAs can have multiple targets [123]. This could prove to be very efficient and also cost effective when identifying new drug targets as one drug will have multiple targets.

Hfq was shown to reduce persistence in *E.coli* by means of regulating genes involved in persistence. Deletion of *Hfq* under persister conditions induced expression of the genes *ybfM* and *dppA/oppA*, and overexpression of these genes led to a 28-fold and 12-fold increase in persistence respectively [153]. It is important to note that these two genes are strongly regulated by the sRNAs, MicM and GcvB, respectively [78,113]. Deletion of *hfq* also repressed a number of genes and included among these genes is the *micC* gene encoding and OmpC translation regulator sRNA MicC [153] (also found in *Salmonella* as mentioned earlier). This is particularly important in SigH as pathogenic bacteria such as *M. tb* also have the ability to enter a state of persistence thereby evading antibacterial therapy and host immune response.

Table 2.3 IGRs associated with specific anti-TB drugs [143]

DRUG	IGRs showing coherent association with drug resistance
Rifampicin & Isoniazid	Rv2754c-Rv2755c hypothetical protein--putative septation inhibitor protein
	Rv3185-Rv3186 transposase--transposase
	Rv3260c-Rv3261 *whiB2—fbiA*
Streptomycin	Rv1194c-Rv1195 hypothetical protein -- *PE13*
	Rv2764c-Rv2765 *hyA*--hypothetical protein
	Rv3185-Rv3186 transposase--transposase
	Rv3260c-Rv3261 *whiB2--fbiA*
	Rv3862c-Rv3863 *whiB6*--hypothetical protein
Ethambutol	Rv1080c-Rv1081c *greA*--hypothetical protein
	Rv1302-Rv1303 *rfe*--hypothetical protein
	Rv1347c-Rv1348 hypothetical protein--Rv1348
	Rv2733c-Rv2734 Rv2733c--hypothetical protein
	Rv2764c-Rv2765 *hyA*--hypothetical protein
	Rv3462c-Rv3463 *infA*--hypothetical protein
Ofloxacin	Rv1080c-Rv1081c *greA*--hypothetical protein
	Rv1816-Rv1817 transcriptional regulatory protein--hypothetical protein
	Rv3651-Rv3652 hypothetical protein--*PE_PGRS60*
Ethionamide	Rv0010c-Rv0011c hypothetical protein--putative septation inhibitor protein
	Rv2340c-Rv2341 *PE_PGRS39--lppQ*
	Rv3260c-Rv3261 *whiB2--fbiA*
	Rv3862c-Rv3863 *whiB6*--hypothetical protein
Kanamycin	Rv1042c-Rv1043c S like-2 transposase--hypothetical protein
	Rv1816-Rv1817 transcriptional regulatory protein--hypothetical protein
	Rv1900c-Rv1901 *lipJ--cinA*
	Rv2208-Rv2209 *cobS*--integral membrane protein
	Rv3210c-Rv3211 hypothetical protein—*rhlE*
Capreomycin	Rv0878c-Rv0879c *PPE13*--trans membrane protein
	Rv0920c-Rv0921 transposase--resolvase
	Rv2068c-Rv2069 *blaC--sigC*
	Rv2208-Rv2209 *cobS*--integral membrane protein
	Rv2764c-Rv2765 *hyA*--hypothetical protein
	Rv3765c-Rv3766 Rv3768c--hypothetical protein
	Rv3796-Rv3797 hypothetical protein--*fadE35*

The significance of persister cells and recurrence of disease, e.g. as seen in TB, is not well understood and therefore investigating the role played by sRNAs in persistence is worthwhile. Also the link between persister cells and latent infections is not well comprehended and could benefit from exploring this new avenue of sRNAs.

More still needs to be done in *M. tb* to decode the nature of sRNA activity. It is also noteworthy that a functional homologue for the sRNA chaperone Hfq, which thus far seems to be a major role player in the mechanism of action of sRNAs, thus far eludes identification in Mycobacteria.

2.8 COULD sRNAs BE USEFUL FOR ANTIMICROBIAL THERAPY?

RNases are of interest for their potential use in antimicrobial therapy. It is already known that human RNases 3 and 7, have antimicrobial activity against a wide range of Gram-positive and Gram-negative bacteria [154–157]. In Pulido's study, they investigated the effects of these two RNases as antimicrobials against *Mycobacterium vaccae,* a non-pathogenic and rapidly growing *Mycobacterium* species model. Their results showed total growth inhibition induced by both RNases citing new avenues for drug development [158]. This shows the potential of RNases as antimicrobial effectors. Technologies such as peptide nucleic acids, which is a combination of nucleic acid and amino acids are already being investigated as potential ways to target and incapacitate sRNA molecules and thereby be used in antibacterial therapy [159]. The potential usefulness of peptide-nucleic acids (PNA), to interfere with growth had been demonstrated in *M. smegmatis* where depletion of *InhA* with the aid of PNA resulted in a growth defect phenotype [160]. Innovation of using sRNAs targeted to drug resistance genes could therefore be useful to achieve similar effects, helping to enhance the bactericidal effect of current TB drugs.

Artificial sRNAs have been designed and experimented on in *E.coli*. sRNA transcripts with target recognition sequences as short as 10nt have been shown to induce gene silencing [161]. It

is therefore important to create sRNAs that can distinguish between human transcripts and also bacterial flora from target transcripts. Also, method of delivery of such compounds is important since they still have to cross the cell membrane barrier and upon entry they could be subjected to degradation by RNases and also effects of molecular chaperones are to be considered. Secondary structures formed once the artificial sRNA has entered the cell is important as some structures such as bulges and internal loops weakening the effects of the sRNA-mRNA duplex [161].

If transcription of a drug resistance-associated gene could be affected by a sRNA mediated antimicrobial effector, it would render the bacteria more susceptible to several first and second-line anti-TB drugs. Conversely, sRNA regulation of drug resistance genes could also enhance resistance to antimicrobial compounds. For example, it was shown, in *S. aureus*, that the SprX sRNA influences sensitivity to glycopeptides by down-regulation of the SpoVG protein which is known to affect sensitivity to glycopeptides [146]. Thus, the possibility exists that targeting of sRNAs could be used to achieve the opposite, that is, enhance sensitivity to a drug [146]. Over the years it has been noted that the introduction of new drugs has been coupled with evolution of new mutations in the bacteria rendering the new drugs once again ineffective. Ineffective use of drugs is one of the reasons for the constant evolution of drug resistance in *M. tb* and therefore policies have to be implemented to protect current drugs before the future antibiotics can be introduced.

Another reason why regulation by sRNAs might be a useful target for developing novel antimicrobial strategies is that many virulence genes, antimicrobial drug targets and metabolic pathways in *M. tb* and *M. bovis* are operonically encoded. Targeting operons is highly efficient as one drug could be used to disrupt transcription of multiple genes which share a promoter. sRNAs act as regulators and therefore could be exploited to induce such inhibitory effect in

organisms. An example is the *mce* operon (*mce1-mce4)* of *M. tb* which is a virulence-associated operon (important for bacterial entry into mammalian cells) [162]

2.9 CONCLUSION

Gene regulation by sRNAs, and its potential involvement in drug resistance development, is under investigation in bacteria. It is not exactly clear how the different sRNAs in *M. tb* and other micro-organisms find their target(s) and this review has shown the significance these sRNAs have on bacterial survival. Exploring these sRNAs may reveal hidden aspects regarding host-pathogen interactions for instance during latent TB infection. Latent TB, of with which one-third of the world's population is infected, is one of the major obstacles in curbing TB infection [163]. Identification of novel regulatory mechanisms associated with latent TB can lead to better TB diagnosis and treatment. Much about drug resistance mechanisms has been uncovered using traditional methods. However, new avenues need to be investigated to fully understand these mechanisms and identify novel drug targets and also improve or revive the efficacy of existing drugs. This review highlights how sRNAs potentially could be exploited as antimicrobials by manipulating their regulatory effects on gene expression. They also can be used to aid the current drugs by altering the regulation mechanisms of DR-genes thereby rendering drug resistant species susceptible to the old drugs. It is therefore worthwhile to explore this avenue further.

CHAPTER 3

MATERIALS AND METHODS

3.1. PREPARATION OF BACTERIAL STOCKS, CULTURE MEDIA AND CHEMICALS

Strain selection

Mycobacterium smegmatis (accession: NC_008596.1 GI: 118467340) was used as a model organism for some aspects of this study. This bacterium is an acid-fast bacterium of the genus *Mycobacterium*. It is generally regarded as a non-pathogenic strain and can cause disease mainly in immune-compromised persons [164]. It is fast growing (compared to other mycobacteria) and shares many characteristics with *M. tb* including the complex cell wall structure and more than 2000 protein homologues, making it an ideal model for *M. tb* studies. *M. smegmatis* is widely used as a model to study genes and proteins of pathogenic members of the mycobacteria. *M. smegmatis* strain mc^2155 was chosen for our cloning studies. *M. tb* strain H37Rv (accession: AL123456.3 GI: 444893469) was used in the sRNA expression experiments as well as for preparation of genomic DNA to use as template in PCR reactions. *E.coli* strain DH5α (NZ_JRYM01000004.1 GI: 817646645) was used in the cloning experiments.

Media and chemicals

Luria Broth (LB) media was prepared by dissolving 20g of LB lyophilised powder (Sigma-Aldrich) in 1000ml distilled H_2O (dH_2O) and sterilised by autoclaving. To make LB agar plates 15g of Bacteriological agar powder (Sigma-Aldrich) was added to the LB broth before autoclaving. Appropriate drugs were added after cooling media/agar to a temperature of approximately 60^oC.

7H9-ADC broth, used to culture mycobacteria, was made by mixing the following: 4.7g lyophilised Middlebrook 7H9 broth base (BD Diagnostics, USA), 900ml dH_2O, 4ml of 50% glycerol and 2.5ml 20% Tween-80. The mixture was autoclaved and 100ml of Middlebrook

albumin-dextrose-catalase (ADC) growth supplement was added after the autoclaved broth was cooled down. The resulting mixture was filtered and stored at 4ºC. The ADC used in making the 7H9-ADC broth was made mixing 25g of BSA powder, 10g glucose, 750p.l catalase (20mg protein/ml, Sigma-Aldrich. Co. LLC) and 500ml H_2O. The mixture is stirred for 2-3 hours on a magnetic stirrer, after which it is filtered (TPP Filtermax vacuum filters, Zellkultur und Labortechnologie Switzerland) and aliquoted, under sterile conditions, into 50ml tubes

The relevant antibiotics depended on the particular experiments. Ampicillin was prepared as 100mg/ml stocks by dissolving the powder in deionised water, followed by filter sterilisation. Hygromycin B (Thermo-Fischer Scientific, USA) was purchased as a liquid, in a concentration of 50mg/ml. Final drug concentrations used were 100µg/ml for ampicillin and Hygromycin B at 100µg/ml for *E.coli* experiments, or 50µg/ml for *M. smegmatis*. The drugs isoniazid (INH) and kanamycin (KAN) were made by dissolving 100mg of the appropriate powder, in double distilled H_2O (ddH_2O). The flouroquinolones, Ofloxacin (OFL) and Moxifloxacin (MOX), were prepared by dissolving100mg of the powder in 3ml alkaline water (1N NaOH in 7ml of dH_2O). Antibiotics were sterilised by filtration through a 0.22µm filter. For both drugs a stock concentration of 10mg/ml was made in a total volume of 10ml and was stored at -20ºC until use.

Preparation of stock cultures and competent cells

M. smegmatis cryo-frozen stock cultures were obtained from stocks in our laboratory. These were inoculated into 10ml 7H9-ADC media. Cultures were incubated overnight at 37ºC in an incubator with a rotating platform in order to enable proper aeration of cultures. Thereafter 1ml of each pre-culture was sub-cultured into 100ml 7H9-ADC media and incubated at 37ºC until the optical density at 600nm (OD_{600nm}, as measured in a light spectrophotometer) of ~1 or slightly greater is reached. Multiple 200µl aliquots were taken from this culture, mixed with 10% glycerol and stored at -80ºC until use. *M. tb* H37Rv stock cultures were used in a similar

manner to prepare frozen stocks, except that each set of cultures took approximately one week to reach an OD of ~1, and cultures were left in a stationary position during incubation, and not shaken on a rotating platform. All *M. tb* culturing was performed in the biosafety level 3 laboratory of the department.

Preparation of electro-competent cells

i) ***E.coli* chemically competent cells**

Chemically competent *E.coli* cells where prepared, as previously described [165], by inoculating 1ul of *E.coli* DH5α cells into 5ml of LB broth. Cultures were incubated at 37˚C overnight (~16 hours), with shaking for good aeration. The overnight culture was diluted 1:100 (5ml: 500ml), in fresh LB media, and incubated at 37˚C again until OD_{600nm} = 0.3-0.4. These log phase cells were then transferred into 50ml centrifuge tubes and centrifuged, in a temperature regulated centrifuge (Eppendorf centrifuge 5810R by Eppendorf® USA) set to 4ºC; at speed 4000rpm for 10 min. Cells were resuspended in 50ml of ice-cold $MgCl_2$ and incubated on ice for 30 minutes. Centrifugation was repeated, at 4˚C for 10 minutes, after which cells were resuspended in 5ml ice-cold buffer (0.05 mM $CaCl_2$-15% glycerol). These chemically competent cells were aliquoted, 200µl, in pre-cooled 1.5ml micro centrifuge tubes and stored at -80˚C.

ii) ***M. smegmatis* mc^{2}155 electro-competent cells**

A pre-culture was prepared by inoculating 5ml of 7H9-ADC with 100µl of a glycerol stock and incubated at 37°C with shaking. After this culture had grown to late log phase, that is actively growing cells, we inoculated 1ml of the culture into 100ml of 7H9-ADC broth. The culture was incubated until it had reached an OD_{600nm} of ~1 or slightly greater. Ziehl Neelsen-staining (ZN-stain) was done according to a previously describe procedure [166] on sample cells to ensure no contamination occurred during growth. Cells (in 50ml aliquots) were harvested by low speed centrifugation (4000rpm) for 10min at 4ºC. All steps were performed on ice and in the

temperature controlled centrifuge. Supernatants were discarded and cells resuspended carefully first in a small amount of 10% glycerol (~3ml) by pipetting up and down. This was done to avoid the creation of bubbles and flocculation during subsequent spins. After the cells were resuspended, the volume was made up to 100ml and the harvesting procedure repeated as follows: 50ml (2nd wash), 20ml (3rd wash) and 10ml (final wash). After the final wash, cells were resuspended in approximately 3-5ml of pure water, whereafter aliquoted into 400ul for electroporation purposes [167–169].

iii) **Gel electrophoresis reagents and equipment.**

50x TAE (Tri-Acetate-EDTA, Thermo Fisher Scientific USA) buffer was prepared as previously described, and used at a 1x TAE concentration for DNA gel electrophoreses[170].

For denaturing RNA gel electrophoresis, agarose gels were prepared in 1x 3-(N-morpholino) propane sulfonic acid (MOPS) Buffer and denaturing gel conditions using formamide and formalin was used, as previously described [171].

Nucleic acids were visualised by the incorporation of ethidium bromide [4µl (of a 10mg/ml stock solution) per 250ml agarose gel], and visualised under UV light at 240nm. When PCR products were used for cloning, bands were visualised with long range UV light or with a blue light transilluminator (Syngene). DNA bands were excised and purified from gel slices using the Zymoclean DNA gel recover kit (Zymoresearch, USA).

3.2 ANALYSIS OF *M. SMEGMATIS* GROWTH AND DRUG SENSITIVITY

Firstly, a *M. smegmatis* growth experiment was done to assess the growth pattern of the *Mycobacterium* under the conditions needed for the antibiotic sensitivity assays and transcription assays to be done. Starter cultures were prepared by inoculating 200µl aliquots (from frozen stocks) into 10ml 7H9 broth and culturing at 37ºC for 18-20 hours. Starter cultures were used to inoculate 200ml of 7H9 culture media, adding appropriately calculated

volumes to achieve the same starting (time= 0) optical density for all biological replicate cultures. Optical densities were assessed at 600nm (OD_{600nm}) in a spectrophotometer (Ultrospec 4051, LKB Biochrom). Cultures were grown at a temperature of 37ºC, with aeration in a shaker incubator (Orbital shaker incubator LM530, Yihder.Co.LTD), and the optical densities monitored for a period of 2 days. OD_{600nm} readings were taken every 3 hours by mixing 500µl of culture with 500µl of 7H9 in a disposable cuvette. The resulting values were multiplied by 2 to obtain the OD_{600nm} reading of cultures. The mean values and standard errors were calculated from three independent biological experiments and used to plot growth curves.

3.3 DETERMINATION OF MINIMUM INHIBITORY CONCENTRATION (MIC) OF *M. SMEGMATIS*

To assess what concentration of various TB drugs would be lethal and sub-lethal the *M. smegmatis* strain mc^2155 was exposed to various concentrations of INH, OFL, MOX and KAN. For this purpose, the bacterium was grown in 7H9-ADC as described for the growth curve until an optical density representing the mid-log phase of growth. At this point, the cultures were split into 20ml cultures and the antibiotics added to obtain various concentrations, **Table 3.1**.

Table 3.1 Concentrations tested to obtain sub-lethal drug concentrations for exposure experiments

Drug	Concentration of drug (µg/ml)				
Ofloxacin	0.00	1.00	1.50	2.00	3.00
Kanamycin	0.00	0.20	0.50	0.80	1.40
Moxifloxacin	0.00	0.03	0.05	0.10	0.20
Isoniazid	0.00	10.00	18.00	24.00	30.00

Aliquots of 1ml each were taken from cultures at different time-points to assess cell death. For this purpose, serial dilutions were made (10^{-1} – 10^{-7}) from the aliquots and 10 µl aliquots (from serial dilutions) spotted onto antibiotic-free LB Agar (**Figure 5.4).** The growth, or lack thereof,

on these plates, after two days of growth at 37°C, were used to calculate MICs as reported in **Table 5.4.**

3.4 BIO-INFORMATIC ANALYSIS OF *M. TUBERCULOSIS* sRNA GENES AND INTERGENIC REGIONS

Previous studies have identified putative sRNAs in various mycobacteria (see literature review). The data from these studies, combined with published literature on mycobacterial sRNAs was used to search for potential associations between drug resistant *M. tb* clinical isolates and published sRNAs. For this purpose, an established genomic database (whole genome data derived from clinical *M. tb* isolates) was scrutinised for the presence of known sRNA genes and intergenic regions that were significantly associated with drug resistance, as defined in a previous unpublished study. The strains in this genomic database represent a cohort of South African TB patients in the Western and Eastern Cape provinces, ranging from drug susceptible strains to XDR strains. This bank is hosted by the Department of Biomedical Sciences Division of Molecular Biology at Stellenbosch University.

An in-house python script which was developed by Ruben van der Merwe (unpublished data) which closely follows the GATK protocol (genome analysis toolkit best practices https://www.broadinstitute.org/gatk/guide/best-practices) was used to screen for SNPs in previously identified sRNA regions. The *M. tb* strain H37Rv (accession: AL123456.3 GI: 444893469) was used as the reference genome in this analysis. The reads were obtained from the Illumina platform and three mappers (BWA, Novoalign, and SMALT) were used [172,173]. Only SNPs appearing on all platforms were regarded as high confidence SNPs. Additional data obtained included strain family, drug resistance patterns of each isolate etc.

3.5 EXPRESSION OF *M. TUBERCULOSIS* sRNA FRAGMENTS IN *M. SMEGMATIS* AND INVESTIGATION OF ITS EFFECT ON PHENOTYPE

Selection of sRNA candidates

Relevant DNA sequences of sRNA fragments were obtained using Tuberculist and TB Database (TBDB).

Cloning strategies

For phenotypical studies, the three candidates were first cloned (STRATEGY 1) into plasmids pMV306 integrating vector and transformed into *M. smegmatis* wild-type cells.

Generated PCR products on the plasmid contained the sRNA coding sequence as well as approximately 50-100bp on either side. An assessment was then done to investigate the effect this expression would have on the growth rate of *M. smegmatis* and also on its drug resistance profile against different drugs (KAN and MOX).

In the second cloning strategy (STRATEGY 2), a plasmid containing a controllable promoter, and the ability to integrate at the *att*B site of *M. smegmatis* were first constructed. PCR products of the sRNAs were then cloned behind the controllable promoter and constructs transformed into the target bacterium. The newly constructed plasmid contained (1) a Tet operator, (2) the pmyc1tetO fragment with a promoter to regulate the Tet ON/Off system in mycobacteria, and which was obtained from pSE100 expression plasmid, and (3) a fragment encoding for an integrase from pMV306 integrating plasmid.

PCR amplification of sRNAs for use in cloning experiments

PCR products of the three candidates, mpr6, mcr3 and ASpks (**Figure 3.1**), were generated for use in downstream cloning experiments. It is important to note that the sRNA fragments contain

surrounding nucleotides so as to ensure no cleavage of the actual sRNA occurred when carrying out restriction digestion reactions. Secondly, since we have no knowledge of potential cleavage and post-transcriptional processing of these candidates, we decided to allow extra bases in case there was post-transcriptional processing.

All primers for this study were designed using the Primer 3 program [174] and using the whole genome sequence of the reference strain H37Rv as the template for primer design. Restriction sites for *Eco*RV and *Hind*III restrictions sites were added on the ends of sequences to facilitate restriction endonuclease digestion and cloning (**Table 3.2**).

Table 3.2 Primers used to amplify sRNA candidates

Primer		**Sequence (5'-3')**
Asks (full)	Forward	GGAATTCGATATCAAG**CTTGATCACCCCGGCCAC**
	Reverse	TAGCCTAAGCTTGAT**GGGGGATCCGATTGAGGC**
Mpr6	Forward	GGAATTCGATATCAAG**AGCGCACCCCGAACATGG**
	reverse	TAGCCTAAGCTTGAT**CGATGGCCTCGATTGACAGG**
Mcr3	Forward	GGAATTCGATATCAAG**CCGAGGTCCACGATGTATTCC**
	Reverse	TAGCCTAAGCTTGAT**TTTGGAAAAGGGAGCCAAAGG**

***Restriction sites for *Eco*RV (gatatc) and *Hind*III (aagctt) are underlined.**

PCR amplification reaction was set up using the Thermo Scientific Long PCR Enzyme Mix. This mix contains a thermostable Taq high fidelity DNA polymerase with proof reading activity. The Long PCR Enzyme Mix generates amplicons which are 3'-dA tailed. An *M. tb* crude boiled DNA template was used for amplification of the sRNA regions using the above mentioned modified primers. The amplification reaction was set up in 25µl reactions which consisted of 2.5µl Long Buffer, 2.5µl of 2.5nM dNTPs, 2.5µl forward primer, 2.5µl reverse primer, 11.75µl ddH_2O, 3µl DNA template and 0.25µl of the Long PCR Taq high fidelity polymerase.

Thermocycler PCR conditions for a one step-Long PCR were set up as described in Figure 3.1.

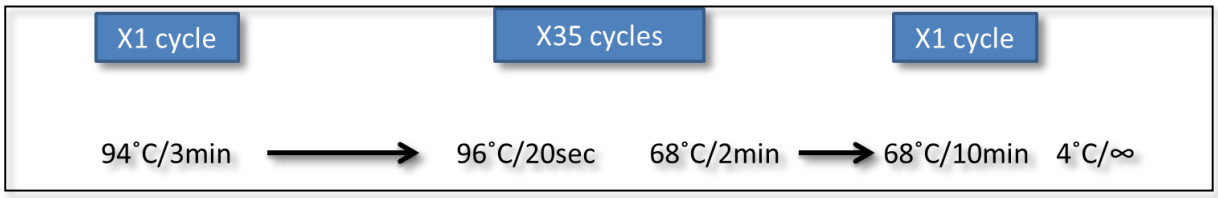

Figure 3.1 Thermo cycling reaction conditions for the Thermo Scientific long PCR used to amplify the sRNA using primers with restriction sites. DNA template denaturation was done at 94°C for 3 minutes, extension/elongation step was performed at 68°C for 35 two minute cycles, final elongation step was for 10 minutes and a final hold step was added at a temperature of 4°C.

Products were analysed with agarose gel electrophoreses to verify the size. Electrophoresis of the PCR products was performed on a 1% agarose gel with ethidium bromide (4µl of 10mg/ml per 250ml agarose gel), and visualised under UV light at 240nm. Thereafter it was sequenced using Sanger sequencing at Central Analytical Facilities (CAF) Stellenbosch University to verify the sequence and homology to published sequences (H37Rv as reference genome). When PCR products were used for cloning, bands were visualised with long range UV light, excised from gels and purified using the Zymoclean DNA gel recover kit (Zymoresearch, California, USA).

Cloning Strategy 1: Expression of sRNAs using pMV306 plasmid

Construction of pMV306 mcr3/mpr6/ASpks plasmid

Multiple PCR reactions were set up to ensure that excess product is available for purification, restriction reactions and final purification steps. PCR products were purified with the Zymoclean DNA gel recover kit before restriction enzyme digestion. The restriction digestion reactions contained 16µl of the cleaned sRNA PCR product (~1.5µg), 2µl of 10X reaction buffer and 1µl of *Eco*RV and *Hind*III restriction enzymes. The mixture was incubated at 37˚C for 16 hours to further increase the chance of digestion as a large amount of PCR product was being used. Upon completion of digestion the mixture was cleaned up using the Zymoclean DNA gel recover kit.

A digestion reaction to linearize pMV306 was set up as follow: 1µl of each restriction (10U each) enzyme *Hind*III and *EcoRV*, 2µl of 10 X reaction buffer and 16µl of the pMV306 plasmid. The digested plasmid was run on a gel and purified using the Zymoclean DNA gel recovery kit.

Ligations reactions

The linearized pMV306 fragment was ligated to each of the sRNAs (mpr6, mcr3 and ASpks) PCR product in the following reaction: 1µl pMV306 plasmid (approximately 100ng), 7µl purified PCR product (50-100ng), 1µl H_2O, 10µl of 2X ligation buffer and 1µl of T4 DNA ligase. The ligation mixture was incubated at 22˚C for 20 minutes prior to transformation.

Transformation of prepared new constructs

The ligation mixtures were used to transform the chemical competent *E.coli* cells by mixing 3µl of respective ligation mixtures and 200µl of bacterial cells. Transformants were selected by plating on LB agar plates containing 100µg/ml of the antibiotic Hygromycin B and incubated at 37ºC over-night (~16 hours).

Single colonies were picked and inoculated into 5ml LB broth (plus 5µl Hygromycin B) and grown at 37˚C on a shaker for ~16 hours. The presence of inserts was confirmed with colony PCR. A schematic diagram for the construction of pMV306-sRNA plasmid is shown in Figure 3.2.

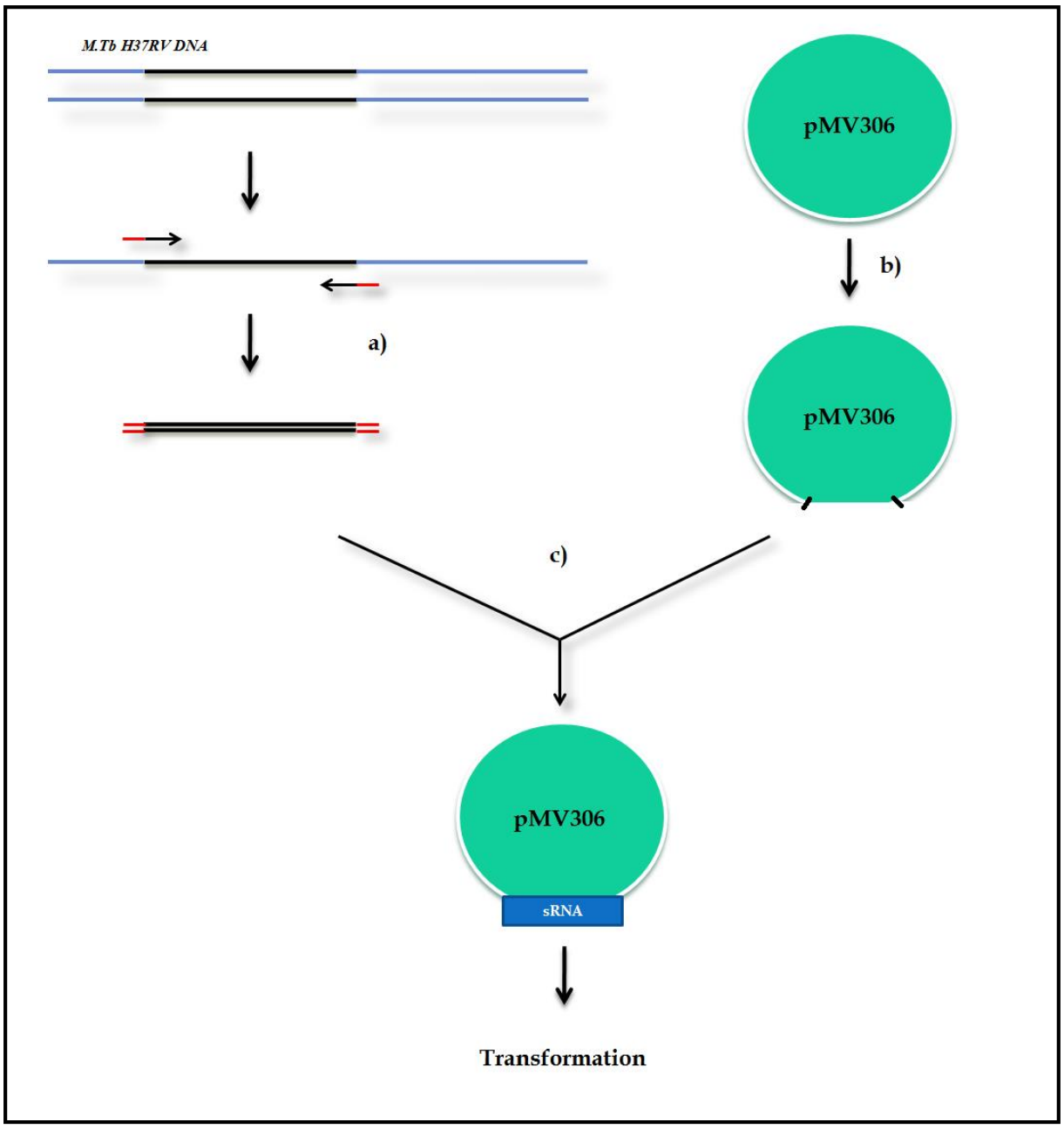

Figure 1.2 A schematic presentation of how the pMV306-sRNA plasmids were made. a) PCR with modified primers followed by digestion of the PCR product with *EcoRV* and *Hind*III restriction enzyme, b) pMV306 plasmid linearized by digestion with *EcoRV* and *Hind*III restriction enzyme, c) ligation of each sRNA to linearized pMV306 plasmid followed by transformation into *E.coli* cells

Transformation of *M. smegmatis*

The pMV306 + insert plasmid constructs for each sRNA were used to transform electro-competent *M. smegmatis* cells. Approximately 3µl (300ng) of the plasmid (pMV306-mcr3/mpr6/ASpks/wild type WT) was used to transform 400µl of electro-competent *M. smegmatis* cells previously described [167,168]. Cells with no added DNA were used as negative

control and cells with pMV306 wild type plasmid served as positive controls. The plasmid/ice-cold cell mixture was pipetted into a cold electroporation cuvette and transformed using the electroporation conditions of 2.5kV, 25µF, 1000Ω in the Bio-Rad gene Pulser II electroporator. Cells were recovered by immediately adding 600µl of 7H9-ADC media into the cuvette and transferring cells into a micro centrifuge tube. Cells were incubated at 37°C for 3hrs or overnight before plating on LB agar with Hyg as selection agent. All transformants, including those cells transformed with and without the control plasmid, were serially diluted (10^{-3} and 10^{-6}) and 100µl aliquots plated on LB plus Hyg plates. Plates were incubated at 37°C for approximately three days.

Investigation of the effect of sRNA expression using pMV306 plasmid on the growth of *M. smegmatis*

To investigate whether sRNAs have an effect on the growth rate of mycobacterial species, colonies from the above transformations for each sRNA were picked and inoculated into 2ml of 7H9-ADC broth and vigorously vortexed, prior to incubation at 37˚C with shaking. The cultures where grown to OD_{600nm} ~1. Thereafter each culture was sub-cultured into 100ml LB broth and the volume of cultures used to inoculate new media for sub culturing was calculated to obtain a starter OD_{600nm} of 0.01. The new cultures were then incubated at 37˚C in a shaker incubator. Note that the cultures where grown in three biological replicates for each sRNA. ODs were taken every 3 hours in a spectrophotometer (Ultrospec 4051, LKB Biochrom). Once an OD_{600nm} of ~1 was obtained subsequent cultures were diluted before an OD was taken (500µl of culture plus 500µl of 7H9 broth). The resulting values were multiplied by 2 to obtain the actual ODs of the cultures. The mean values and standard errors were calculated from the three independent biological experiments and used to plot growth curves.

Characterisation of drug resistance phenotype of *M. smegmatis* cells over-expressing the sRNAs mcr3, mpr6 and ASpks

To investigate whether sRNAs had an effect on altering drug resistant phenotypes through their proposed involvement in bacterial gene regulation, a drug induction experiment was designed. We hypothesised that expressing extra copies of sRNAs can induce a change in drug resistant phenotype from resistance to a sensitive phenotype or even just a lower the MIC. We exposed mutant *M. smegmatis* cells to three drugs (KAN, INH and MOX), all of which are currently used in various levels of TB therapy. The mutants were exposed to the drugs at MIC levels. Two experiment designs were followed:

Experiment 1: Mutant *M. smegmatis* cells transformed with pMV306 plasmids with sRNA insert were grown in 100ml 7H9-ADC broth plus 100µl of the antibiotic Hygromycin to an OD_{600nm} of ~0.5. Thereafter the 100ml cultures were split into 20ml cultures, appropriate drugs added at sub-lethal (MIC) concentrations and incubated for 6 hours at 37°C in a shaker incubator. To monitor any growth (indicating an altered resistance phenotype), aliquots of 1ml each were taken from the cultures, serial dilutions made (10^{-0}-10^{-5}). 10µl aliquots were spotted onto antibiotic-free LB agar plates (**Figure3.3**) and incubated for 3 days. Figure 3.3 shows a schematic representation of the experiment. Experiments were repeated to obtain three biological replicates of each drug treatment.

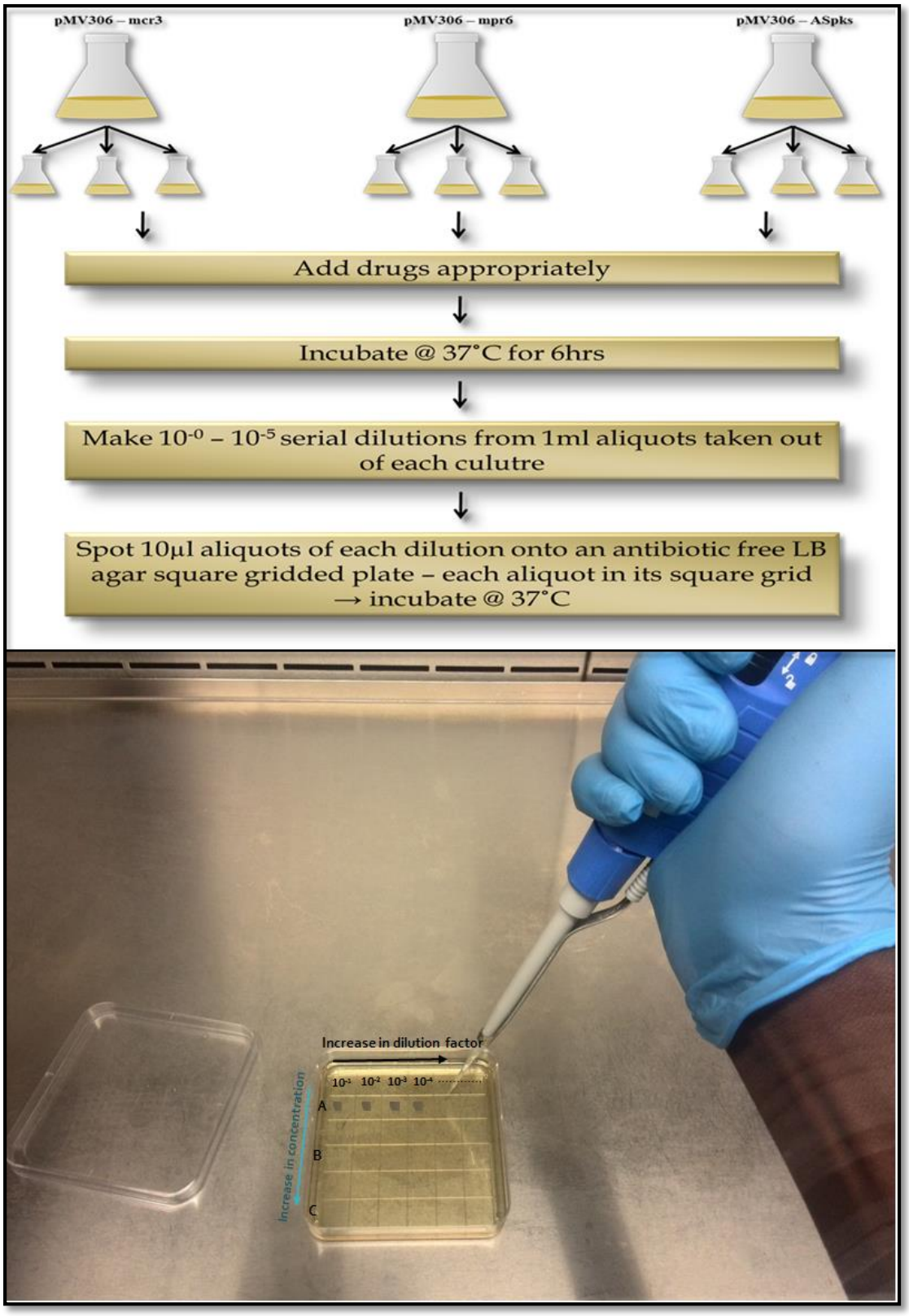

Figure 3.3 Drug induction experiment 1 schematic representation. Aliquots of 1ml each were taken from the cultures, serial dilutions made (10^{-0}-10^{-5}). 10µl aliquots were spotted onto antibiotic-free LB agar plates

Experiment 2: Mutant *M. smegmatis* cells transformed with pMV306 plasmids with sRNA insert were grown in 100ml 7H9-ADC broth plus 100µl of the antibiotic Hygromycin B to an

OD_{600nm} of ~0.5. Thereafter the 100 cultures were split into 20ml cultures in three separate replicates, appropriate drugs added at sub-lethal concentrations and incubated for 6 hours at 37°C in a shaker incubator. Thereafter aliquots of the cultures were taken out and sub-cultured into fresh antibiotic free 7H9-ADC (20ml) plus hygromycin to make a starter culture of OD_{600nm} ~0.01. The culture was incubated at 37°C on a shaker for 3 days. OD_{600nm} readings were taken every 6 hours for the first 36hrs to observe any growth, thereafter readings were taken every 24 hours. **Figure 3.4** shows a schematic representation of the experiment.

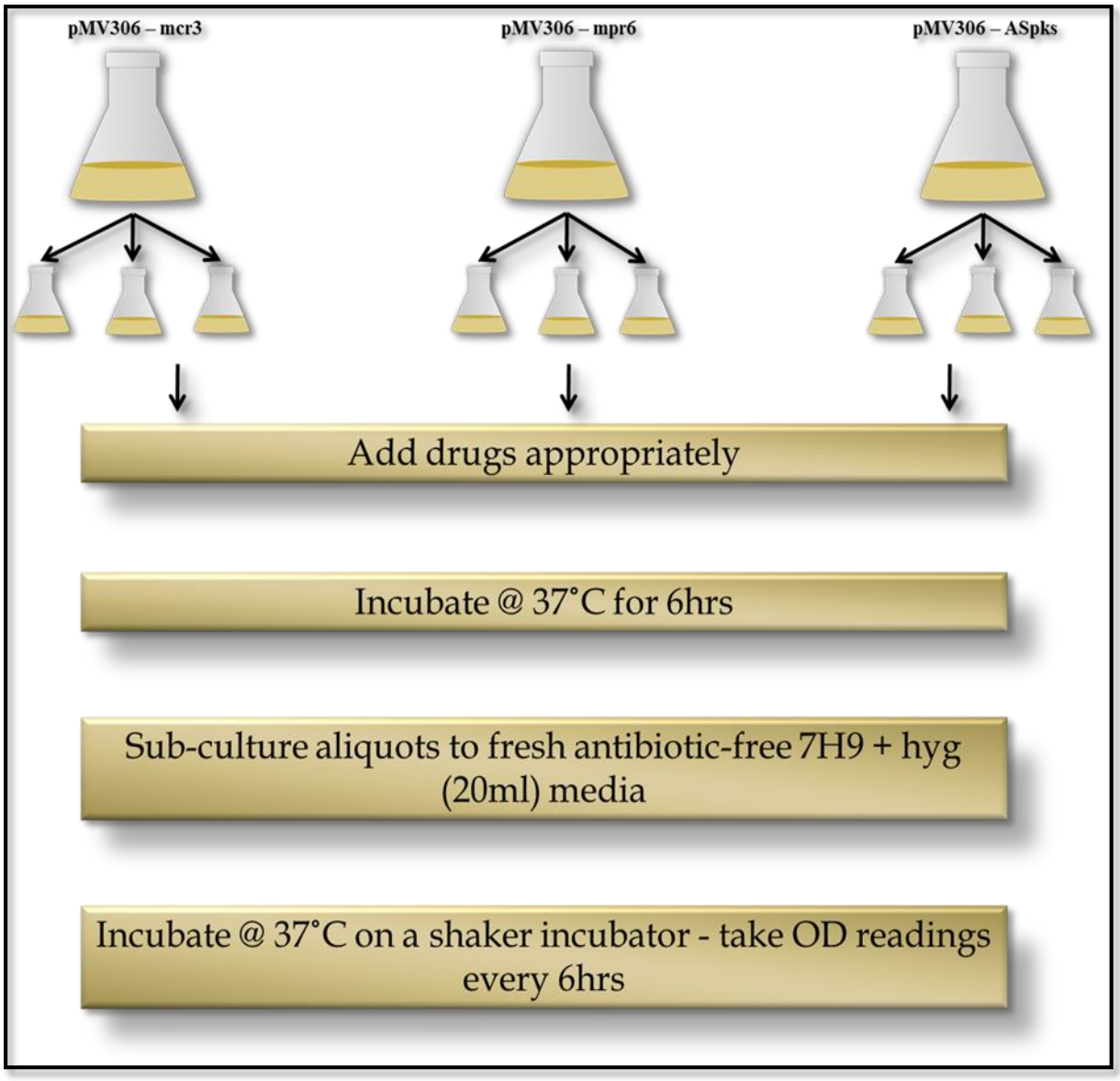

Figure 3.4 Drug induction experiment 2 schematic representation. Cultures were incubated at 37°C on a shaker for 3 days. OD_{600nm} readings were taken every 6 hours for the first 36hrs to observe any growth, thereafter readings were taken every 24 hours.

Strategy 2: Controlled expression of small RNA in *M. smegmatis* using a modified plasmid containing the Tet ON/OFF promoter expression system

Many processes in mycobacteria are poorly understood because they are essential for mycobacterial growth hence making it a challenge to make knockout mutants or induce mutations. The use of controllable promoters on plasmids provides opportunity to regulate the level of gene expression in a bacterium thereby providing a way to study such bacterial processes. The tetracycline controllable promoter (Tet on and Off system) in plasmid pSE100 provides such control in mycobacteria [175]. The Tet on and off system provides a means of regulating the expression of essential genes creating conditional/controllable mutants for these essential genes. The inducible gene expression system described here makes use of a tetracycline repressor (TetR) that can be repressed or induced by tetracyclines. TetR proteins regulate the expression of tetracycline-exporting proteins. TetR proteins are tightly bound to Tet-operators (TetO) in the *tetA* promoter in the absence of tetracycline thereby suppressing the transcription of *tetA* which encodes for a tetracycline exporter. Once tetracycline is introduced it binds to TetR resulting in a conformational change leading to the dissociation of TetR from TetO. This dissociation paves way for the transcription of the tetracycline exporter genes. Guo and colleagues constructed a mycobacterial conditional expression system based on the above mechanism by placing TetR responsive operators (TetOs) in regions of a strong mycobacterial promoter not essential for promoter activity [175]. This resulted in a Pmyc1tetO promoter that can be repressed by TetR in *M. tb* and *M. smegmatis*. They also showed that the use of different anhydrotetracycline concentrations can also control (inducer concentration-dependent) the level of expression [175].

However, this vector integrates into mycobacteria at the site of the homologous DNA which is present on the plasmid and in the mycobacterial genome. This results in promoter substitution of a native mycobacterial promoter with the tetracycline controllable promoter. As sRNAs frequently are situated in promoter regions of genes or inside genes, we did not want to interrupt

native genes or promoters, but simply have the sRNAs cloned after the TET on/off promoter, integrated at the *att*B site of mycobacteria. Therefore we aimed to create a plasmid containing a pSE100 fragment with the TET on/off promoter, combined with the fragment from pMV306, which contains the *att*P site necessary for integration at *att*B in mycobacteria. The sRNA could then be cloned directly after the controllable promoter, but integration occurs at a site distant from the homologous genomic DNA target, i.e. at the *att*B site. This ensures that a gene is not interrupted, which is important in this experiment, as we did not want to do promoter replacement or interrupt genes.

Construction and verification of plasmid pTKL

The plasmid pTKL was constructed by combining sections of two different plasmids, pSE100 and pMV306. pSE100 is an expression vector, derivative of a shuttle vector pMS2 and contains a Tet operator pmyc1tetO with a mycobacterial promoter [175]. It also contains an *E.coli* (pMB1) and a Mycobacteria (pAL500) origin of replication and a multiple cloning site with two transcriptional terminators flanking it. PSE100 in *E.coli* is high copy plasmid and low copy in mycobacteria. It carries a Hygromycin cassette conferring bacterial resistance to hygromycin (100mg/ml - *E.coli*, 50mg/ml – mycobacteria). **Figure 3.5** shows the pMV306 and pSE100 plasmid maps.

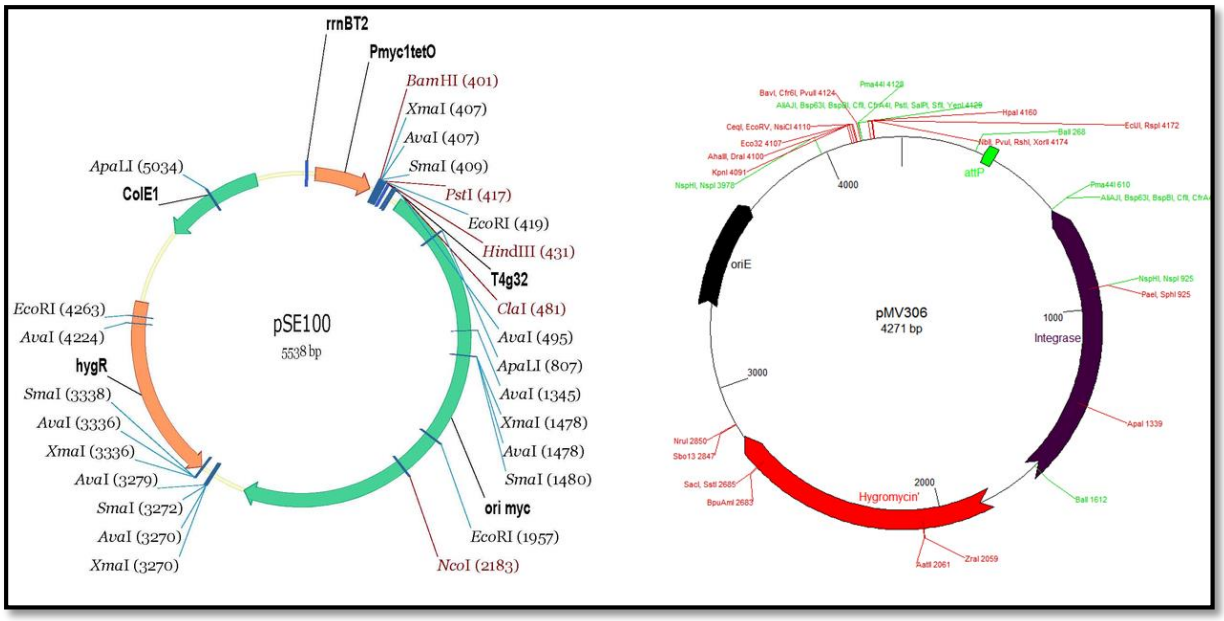

Figure 3.5 pSE100 (5538bp) and pMV306 (4271bp) plasmid maps. A fragment with a Tet operator pmyc1tetO, a mycobacterial promoter, a Hygromycin cassette conferring bacterial resistance to hygromycin and origin of replication from the pSE100 plasmid was merged to a fragment from pMV306 containing an *M. tb* integrase sequence.

For the purposes of our experiment we excised a fragment containing the hygromycin B and Tet operator cassette from pSE100 using the restriction enzymes *Cla*I (AT^CGAT) and *Msc*I (TGG^CCA). The reaction conditions were as follows; 14µl H_2O, 3µl (~400ng/µl) PSE100 plasmid, 2µl 2X Tango buffer (Thermo scientific) and 1µl of each enzyme *Cla*I and *Msc*I (Thermo Scientific). The mixture (final volume was 20µl) was incubated at 37˚C for 2 hours. The restriction digest was run on a 2% agarose gel for 2 hours at 140Volts, after which the desired band (3174bp) was cut out and purified using the Zymoclean DNA gel recover kit (Zymoresearch).

We excised the target fragment containing the integrase sequence of the pMV306 using MscI restriction enzyme, thereby cutting on both sides of the integrase fragment in PMV306 plasmid. The restriction digest mixture was run on a 2% agarose gel at 120 volts to obtain the desired band size of 1344bp. It was cleaned thereafter using the Zymoclean DNA gel recover kit (Zymoresearch, California).

To combine the two fragments we followed the protocol below and illustrated in **Figure 3.6**. pSE100 fragment was first dephosphorylated by adding 7µl H_2O, 10µl pSE100 fragment, 2µl Phosphatase Buffer and 1µl Phosphatase. Incubate at 37˚C for 10min, inactivate enzyme at 75˚C for 2min and chill on ice for 2min. Ligation of the pMV306 fragment to the dephosphorylated (dephos) pSE100 fragment was done by adding 4µl of the dephos pSE100, 5µl pMV306 fragment, 10µl 2X ligation buffer (Thermo Scientific) and 1µl T_4 DNA ligase (Thermo Scientific) and incubation of the mixture at 22˚C or room temperature (RTM) for 20min. A Clean up reaction then followed using the Zymoclean Gel clean-up kit (contains a DNA binding buffer that can be used to clean up DNA from contaminating reagents. Generation of blunt ends of the pSE100 end was done using the restriction digestion enzyme *Cla*I to cut the plasmid. Initial ligations with compatible ends failed. Thus blunt ended fragments were generated for further ligations: Mix 8µl pSE100-pMV306 fragment, 10µl 2X reaction buffer (Thermo Scientific) and 1µl Blunting enzyme (Thermo Scientific). Incubate at 70˚C for 5min and chill on ice for 2min. The final Ligation (circularizing the plasmid) and transformation reaction was done by adding 1µl T_4 DNA ligase to the reaction mixture above and incubating at 22˚C or room temperature (RTM) for 20min. The whole mixture (20µl) was then used to transform 200 µl of chemical competent *E.coli* cells which were then plated on LB agar hygromycin plates and incubate over night at 37˚C. Refer to appendix 1 for the above protocol in point form.

Plasmids were isolated from selected transformants which were cultured overnight, in 5ml LB media. The size and orientation of the two fragments in the new plasmid, named pTKL were verified by doing a restriction digest reaction with the restriction enzyme *Bam*HI.

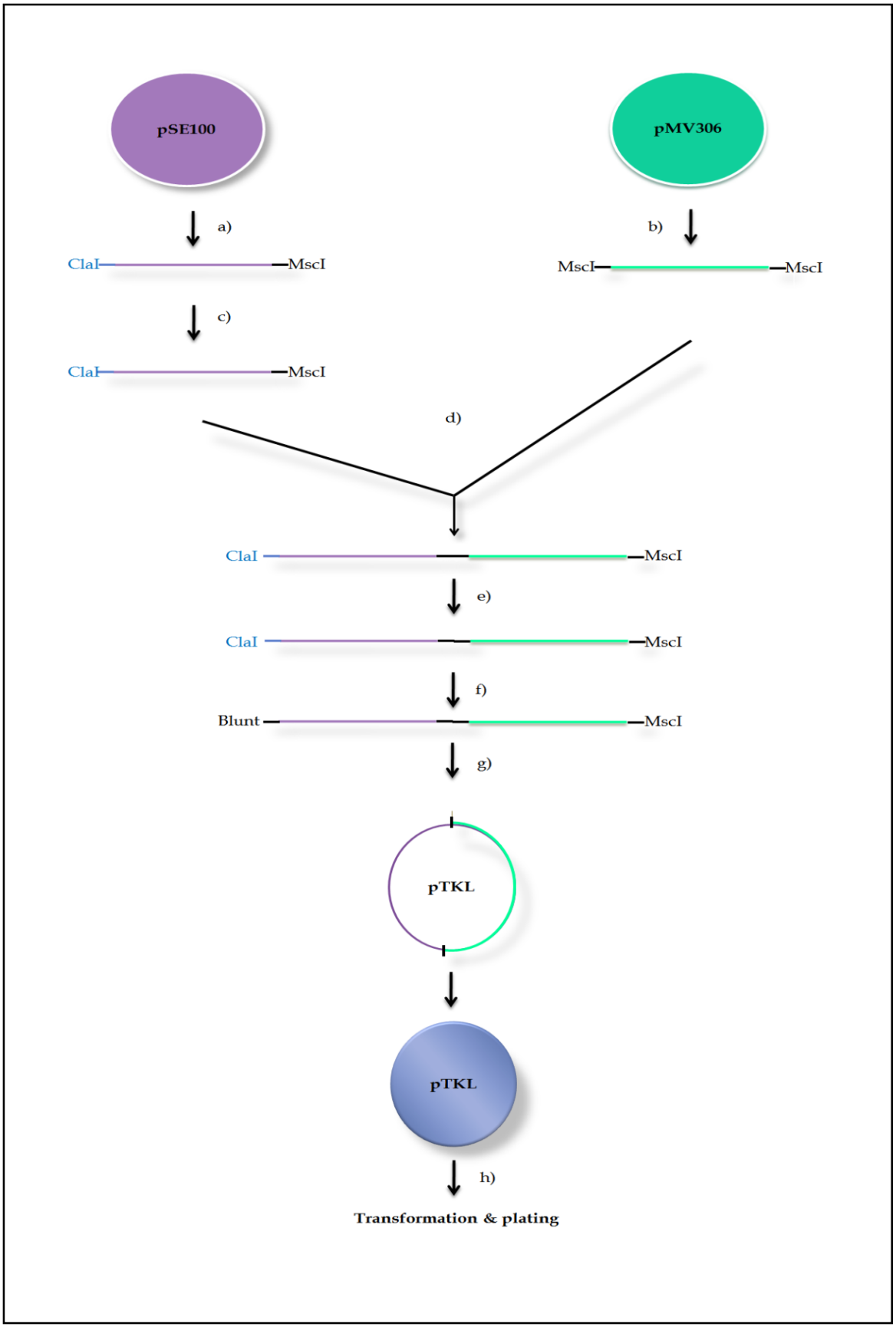

Figure 3.6 Construction of plasmid pTKL. a) Digestion of the pSE100 with restriction enzymes *Cla*I and *Msc*I to obtain a pSE-fragment containing the *PmyctetO* gene and the Hygromycin antibiotic resistance cassette, b) Digestion of the pMV306 plasmid with the restriction enzyme *Msc*I to obtain the pMV-fragment containing the integrase cassette, c) dephosphorylation of the pSE-fragment, d) ligation of the two fragments, e) clean-up with the Zymoclean DNA recovery kit to remove reagents that can inhibit downstream steps, f) blunting of the pSE-fragment cut with *Cla*I followed by g) ligation of the blunt end to the pMV-fragment resulting in a circular plasmid. Finally h) the plasmid is then transformed into *E.coli* to obtain more copies of the plasmid.

Construction of pTKL-derived constructs containing sRNA genes

After obtaining the desired plasmid, 16µl of the plasmid pTKL was linearized by digestion with 1µl of *EcoRV* and 1µl of *Hind*III restriction enzyme for compatibility and easy ligation with the inserts.

The digested plasmid was visualised on a 1% agarose gel and cut out under white light to avoid prolonged UV exposure which can induce mutations. The cut out fragments were cleaned up with the Zymoclean Gel DNA Recovery Kit recovering approximately 90ng/µl of cut plasmid.

For ligation of digested PCR sRNA insert (prior cleaned by the Zymoclean DNA recovery Kit as in over-expression experiments above) to linearized pTKL plasmid the reaction consisted of 1µl H_2O, 7µl of the digested insert, 1µl linearized pTKL plasmid, 10µl 2X reaction buffer and 1µl of T_4 DNA ligase (illustrated in **Figure 3.7**). Incubation was for 30 minutes at 22˚C. The ligation controls included a wild type pTKL plasmid digested with *Hind*III restriction enzyme which, after addition of ligation reaction reagents, would re-ligate and form colonies on plates with hygromycin confirming ligation efficiency. To confirm that the linearized pTKL plasmid used in ligation of the sRNA inserts did not self-ligate a ligation reaction was set up consisting of 5µl of H_2O, 4µl of the linearized pTKL plasmid (using *EcoR*V and *Hind*III), and 10µl 10X ligation/reaction buffer and 1µl T_4 DNA ligase. No colonies were expected on the plates with hygromycin, additionally confirming that whatever is seen on the experiment plates was pTKL plasmid plus insert**.** All the ligation reaction mixtures for experiment, positive and negative controls were used to transform chemical competent *E.coli* cells, plated on LB hygromycin plates and incubated at 37°C overnight (~16hours). Colonies were picked, inoculated into 5ml of LB plus 5µl of hygromycin and grown at 37°C on a shaker overnight.

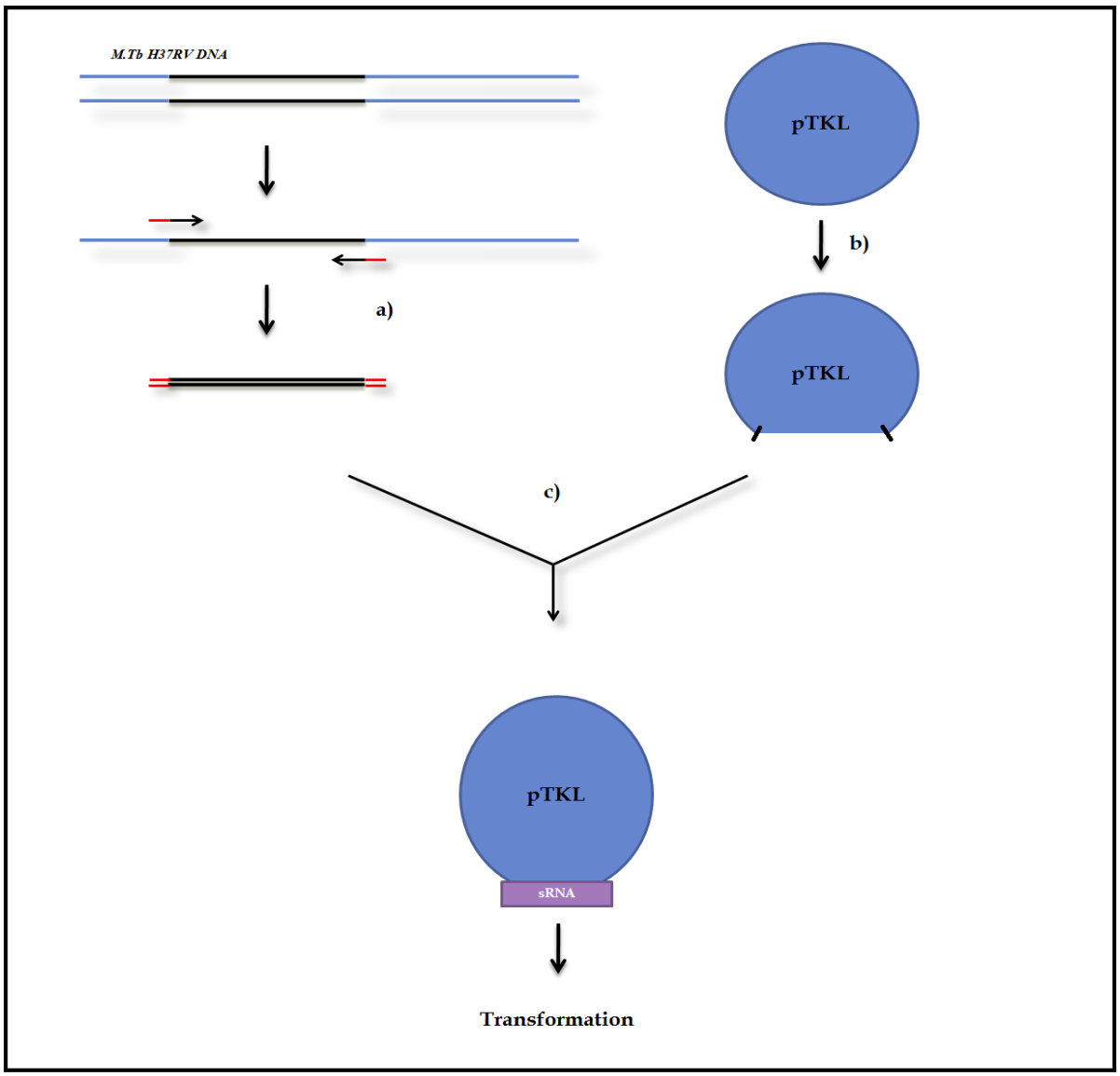

Figure 3.7 Schematic representation of how the pTKL-sRNA plasmids were made. a) PCR of with modified primers followed by digestion of the PCR product with *EcoRV* and HindIII restriction enzyme, b) pTKL plasmid linearized by digestion with *EcoRV* and HindIII restriction enzyme, c) ligation of each sRNA to linearized pTKL plasmid followed by transformation into *E.coli* cells.

Part of the culture (2µl) was then used, in a PCR reaction, to confirm the presence of inserts by performing a Thermo Scientific Long PCR reaction, set up as described above and with the same sRNA primers. The amplification products were run on a 1% gel and visualised under UV. The remaining portion of the 5ml culture, with positive confirmation of inserts by PCR, was then used to extract plasmid for downstream experiments (yielded ~350ng of plasmid). To further confirm the presence of inserts and orientation a restriction digestion reaction was set up for the extracted plasmid as follows; 1µl *Hinc*II restriction enzyme, 14µl H_2O, 2µl 10X

restriction buffer and 3µl of each plasmid. The mixture was then incubated at 37°C for 2hrs (or overnight), after which it was run on a 1% agarose gel and visualised under UV light.

Characterisation of growth and drug resistance phenotypes of *M. smegmatis* expressing sRNA genes mcr3, mpr6 and ASpks, using the Tet on and off system

To characterise the effect sRNAs on *Mycobacterium* physiology, the above constructed pTKL plasmids with inserts were used to transform *M. smegmatis* cells using the method described above (to transform pMV306 + inserts into *M. smegmatis*). The resulting transformed cells were cultured in 100ml 7H9 media and plated on agar plates with hygromycin. Repeated transformations failed consistently, thus this strategy was abandoned and no further experiments could be conducted.

CHAPTER 4

ESTABLISHING A STANDARD OPERATING PROCEDURE FOR sRNA EXTRACTION FROM M. TUBERCULOSIS USING M. SMEGMATIS

4.1 INTRODUCTION

The conventional method (chloroform/phenol method) for RNA extractions in bacterial RNA contains no enrichment for sRNAs. In order to obtain sRNAs initially one would have to extract total RNA followed by a gel electrophoresis of the extracted RNA. After the electrophoresis the bands of size smaller than 500bp (presumed to be sRNA bands) will be cut out and purified to obtain a sRNA enriched sample. Due to the rigidity of the *Mycobacterium* cell wall it is difficult to extract a high yield of quality RNA**.** In addition to this problem, sRNAs constitute a small proportion of the total RNA and purification from a gel piece results in a significant loss of the sRNA. Also, for accurate and reproducible gene expression analysis, it is imperative to eliminate the genomic DNA, therefore a DNase step follows, resulting in further loss of the sRNAs**.** It is important to note that the DNase treatment step can also be performed right after total RNA extraction prior to gel electrophoresis but still the loss of RNA is inevitable. The whole process from total RNA extraction, gel electrophoresis and clean-up to DNase treatment is time consuming (4-5 days) and increases the chances of loss of yield and contamination. This chapter describes a method in which one can quickly extract and enrich for sRNA, including DNase treatment, in fewer steps and a shorter time frame of approximately 4 hours.

4.2 SMALL RNA ENRICHMENT PROTOCOL

We attempted to design a protocol specifically for isolating sRNAs from total RNA without using purification of the sRNA bands from agarose gel slices. The method entailed combining the protocols of the FastRNA PRO™ BLUE KIT (Catalog number: 6025-050 MP Biomedicals, USA)[176], the RNeasy Mini™ Kit (catalog number: 74104, Qiagen, USA)[177] and the RNeasy MiniElute™ kit (catalog number: 74204, Qiagen, USA) [178] (detailed below). The first kit was used to isolate total RNA, the second to purify and DNase treat RNA and separate the sRNAs from the bigger RNA transcripts, whilst the last kit was used to concentrate the sRNAs and elute it in a smaller volume of elution buffer or water.

4.3 CULTURING OF STRAINS FOR sRNA ISOLATION

M. smegmatis strains were cultured, with aeration, in 200ml 7H9 media to mid-log phase, 20ml of the culture was then distributed into individual flasks. Prior to extraction, RNAprotect Reagent from Qiagen (catalogue number: 76526) was added to cultures followed by 5min incubation at room temperature. This agent is reported to avoid changes in bacterial expression profiles [179]. The cells from these 20ml culture were harvested by centrifugation at 4000rpm (Eppendorf Centrifuge 5810 R). RNA and sRNA extractions were done as described in the protocol below. sRNAs were eluted in 20ul of RNase-free water and stored at -80ºC. RNA was quantitated using the Nanodrop 2000c machine (Nano Drop technologies. Inc. USA).

4.4 PROTOCOL:

The cell lysis and first extraction step is done using the FastRNA Blue Kit:

1. After centrifugation of cultures pellet cells, disrupt cells using the FastRNA blue kit protocol
   - Follow the Fast RNA Blue Kit protocol, including the the chloroform extraction step;

The separation of different RNA species is done using the RNEasy Kit:

2. Transfer the upper phase, obtained from the chloroform step, to a new reaction tube. Add 350ul 70% ethanol, vortex vigorously;
3. Transfer sample into an RNeasy Mini spin column placed in a 2ml collection tube.
4. Centifuge at ≥8000xg (≥10,000rpm) for 15s at room temperature, COLLECT THE FLOW THROUGH (which contains the sRNAs); If you want to isolate the larger RNAs from a culture, keep this column and follow the RNEasy protocol to isolate
5. Add 450ul of 100% ethanol to flow through liquid, vortex to mix;

The enrichment of sRNA steps are now performed on the MiniElute columns:

6. Pipette700ul of this mixture into the RNeasy MinElute column, which is placed in a 2ml collection tube, then centrifuge at ≥8000xg for 15s at rtm
7. Discard flow through. NB repeat the step until the whole sample has been pipetted into the spin column
8. Do an on-column DNase treatment*

***On-column DNase treatment was done according to the RNeasy protocol (QIAGEN):**

8.1. Pipette 350µl Buffer RWT into the RNeasy Mini Spin Column/ RNeasy MinElute spin column and centrifuge for 15 seconds at ≥8000 x *g* (≥10,000 rpm). Discard the flow-through.

8.2. Add 10µl DNase I stock solution to 70µl Buffer RDD. Mix by gently inverting the tube. Do not vortex, as DNase is sensitive to physical shearing and easily inactivated.

8.3. Pipette the DNase I incubation mix (80µl) directly onto the RNeasy Mini Spin Column membrane/ or RNeasy MinElute spin column membrane and place on the bench top at 20–30°C for 1 hour.

Note: Make sure to pipet the DNase I incubation mix directly onto the RNeasy membrane. DNase digestion will be incomplete if part of the mix sticks to the walls or the O-ring of the RNeasy Mini Spin Column/ RNeasy MinElute spin column.

8.4. Pipette 350µl Buffer RWT into the RNeasy MinElute spin column and centrifuge for 15s at ≥8000 x *g* (≥10,000 rpm). Discard the flow-through.

Continue with extractions and elution for sRNA at step 9:

9. Pipette 500ul buffer RPE* into RNeasy MinElute Column. Centrifuge at ≥8000xg for 15s, discard flow through

10. Add 500u1 of 80% ethanol to the RNeasy MinElute column. Centrifuge at 8000 x g for 2min to dry column. Discard flow through.

    NB. After centrifugation, remove column from collection tube. Be careful not to touch the flow through

11. Place column into a 2ml collection tube. Open lid and centrifuge at ≥8000xg for 5min
12. Place column into a 1.5ml collection tube, add 12ul RNase free water onto the spin column membrane. Close lid and centrifuge at ≥ 8000xg for 1min to elute sRNA enriched fraction.
    - Repeat 12
13. Store RNA and perform downstream experiments as desired.

NOTE: On-column DNase treatment to obtain mRNA can also be done on the RNeasy column, using the same protocol as above, except that it is done on the RNeasy columns.

4.5 ANALYSIS OF THE QUALITY OF ISOLATED sRNAs VERIFYING THAT ISOLATED sRNAs IS DNA-FREE

To check for remaining DNA, a standard PCR, using the Hotstar Taq kit from Qiagen [180], was done with the following primers; 6247 SF (GTTCTCGGGACGGAACTTG) and 6247 SR1 (ACCGATCTGGTGAAGGACAC) that amplify a fragment of the *msmeg6247* gene, encoding for a hypothetical protein of *M. smegmatis* [174].

Analysis of RNA quality using denaturing gel electrophoreses

Denaturing-RNA gel electrophoresis was done according to a previously described protocol, to visualise the 3 bacterial mRNA species and the sRNAs [170].

Testing the extraction protocol in *M. tuberculosis*

In this section of the study, we designed an experiment to isolate sRNA transcriptome from *M. tb* that were cultured under a drug-free and drug induced stress environment. The objective of

this experiment was to observe if good quality sRNAs can be isolated from stressed cells and is useful for downstream reactions such as quantitative reverse transcription PCR (qPCR).

In this experiment we chose to look at two sRNAs, mcr3 and ASpks, based on their association with genes previously published to be involved in drug resistant mechanisms. The mcr3 sRNA is located upstream (promoter region) of the *M. tb rrs* gene, known to be involved in aminoglycoside synthesis [147]. The *rrs* gene is widely associated with resistance to the anti-TB drug Kanamycin [181]. The ASpks sRNA is coded within the *pks12* gene which has been shown to cause intrinsic resistance to a range of antibiotics in the *Mycobacterium avium* complex [182], but mutations in this transcript were also associated with drug resistant isolates in previous studies. Kanamycin was chosen as the test drug for the mcr3 sRNA and Moxifloxacin was chosen for the ASpks sRNA. Moxifloxacin was chosen as currently it is a promising drug with a very high bactericidal activity as a combination drug with first-line drugs and is associated with a decrease in the treatment length of drug susceptible TB and MDR TB [183]. One of the factors that lead to development of resistance is that patients default treatment creating an environment for the bacteria to develop resistance conferring mutations, and this is attributed by the intolerable length of the treatment (currently 6 months for drug susceptible TB). Addition of MOX to the current regimen could improve the efficacy of TB therapy and hence it is important to investigate potential drug resistance mechanisms before its use.

4.6 ISOLATION AND MEASUREMENT OF sRNA LEVELS IN *M. TB* EXPOSED TO SUB-LETHAL CONCENTRATIONS OF KANAMYCIN OR MOXIFLOXACIN

Six cultures of *M. tuberculosis* (50ml of 7H9-ADC broth each), were prepared with starting inoculums to obtain an optical density of approximately0.05. When cultures, grown at 37^{o}C, reached an optical density of 0.4 to 0.5, two biological replicates each were combined in order to obtain a big enough volume for induction experiments. This volume of 100 ml was split into separate culture flasks of 20ml each (**Figure 4.1**), and antibiotics added as indicated in **Table 4.1**. One day after the addition of the antibiotics, cells were pelleted, resuspended in

1ml of RNAProtect as previously described, and RNA extractions proceeded as described for *M. smegmatis* above.

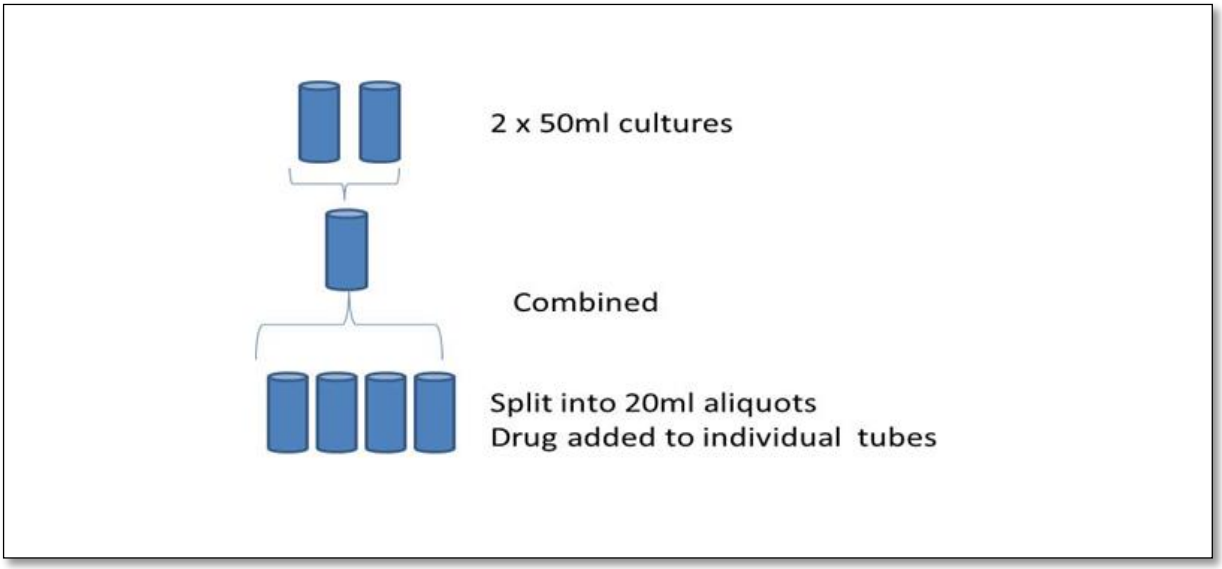

Figure 4.1 Experimental setup for s RNA extraction from *M. tb* cultures. In order to increase starter culture two 50ml cultures were combined before the drug induction step.

Pooling of two cultures was done to increase the amount of starter material, as culturing in the BSL3 laboratory is limited to 50ml volumes of *M. tb* per culture tube, and we isolated RNA from early log phase cells.

Table 4.1 Exposure conditions for *M. tuberculosis* cultures.

Tube 1	**Extraction time**	**Moxifloxacin final concentration (µg/ml)**
1	Day 0, control	0
2	After 24hrs, control	0
3	After 24hrs	0.01
4	After 24hrs	0.02
Tube		**Kanamycin final concentration (µg/ml)**
1	Day 0 control	0
2	After 24hrs, control	0
3	After 24hrs	1
4	After 24hrs	2

4.7 qPCR REACTIONS TO ANALYSE IF cDNA CAN BE OBTAINED FROM ISOLATED sRNAs

cDNA Synthesis and PCR reactions

In preparation for the qPCR, new primers were designed, in order to generate PCR fragments of approximately 100bps in length (**Table 4.2**).

Table 4.2 sRNA primers used for qPCR.

Primer		Sequence (5'-3')	TM °C	Length (bp)
Mcr3	Forward	CCATTGCCGGATTTGTATTAG	59.33	102
	Reverse	GTGTGTTTGGTGGTTTCACA	57.81	
Asks	Forward	CTGCGTATGACCCATATTCGG	59.02	108
	Reverse	ATCCGATTGAGGCTCAGG	60.72	

cDNA synthesis was performed on the extracted RNA samples, using the iScript cDNA synthesis kit (catalogue number: 1708890 Bio-Rad laboratories, Inc.), including the random hexamers. Reactions contained 50ng of purified RNA, 4µl 5x iScript reaction mix, 1µl iScript reverse transcriptase and made up with Nuclease-free water to a final volume of 20µl. The reaction mix was incubated for 5 minutes at 25°C, 30 minutes at 42°C and finally 5 minutes at 85°C to inactivate the iScript reverse transcriptase. For the second step PCR: A master mix was made and 18µl aliquots pipetted per well, in a 96-well plate. Each aliquot consisted of 10µl of iTaq Universal SyBr green supermix, 0.2µl of the forward primer, 0.2µl of the reverse primer and 7.6µl of RNase-free H_2O. cDNA, 2µl each, were added per well.

The 96 well ABgene plate (Thermo Fisher Scientific, Inc., Waltham, MA, USA) was centrifuged (3rpm, 10 seconds, room temperature) before PCR was performed on the Bio Rad

CFX96™ Real-Time System (Bio Rad C1000™ Thermal Cycler) instrument (Bio Rad Laboratories, Inc., Hercules, California, USA). The protocol was as follows: Activation at 95°C for 15 minutes, for 40 cycles- Denaturation at 95°C for 15 seconds, Annealing at 60°C for 30 seconds and Elongation at 72°C for 30 seconds. Lastly a melting curve program (95°C for 0 seconds, 60°C for 15 seconds, 90°C for 0 seconds with a heating rate of 0.1°C/s) and Cooling down program (40°C for 10 seconds) were executed.

Quantitative analysis

Standard curves were initially constructed using the cDNA derived from the sRNAs. Subsequent standard curves were generated using serial dilutions made from *M. tb* genomic DNA. The standard curves were constructed by first making serial dilutions (undiluted DNA at concentration of 50ng/µl, 1/10, 1/100, 1/1000, 1/10000, 1/100000, 1/1000000) of a genomic DNA sample. The *M. tb* gene, *sigA*, usually used as reference gene, was used as a control to verify that cDNA was successfully made, albeit from the rRNA. A linear standard curve was plotted using the Bio-Rad CFX96^{TM} Real-Time System (Bio-Rad Laboratories, Inc., USA). The relative quantification was not carried out as a result of limited availability of reference genes optimised for sRNAs.

CHAPTER 5: RESULTS

5.1 BIOINFORMATIC ANALYSIS OF MUTATIONS IN *M. tuberculosis* sRNA GENES AND INTERGENIC REGIONS

This study investigated sRNAs which were previously associated with drug-resistant *M. tb* drug resistance. First we verified if mutations in the same species could be identified in a genome bank of clinical samples from our local population. We further selected a sub-set for further analyses. These candidates, mpr6, mcr3 and ASpks were selected based on the following:

1. mpr6 is associated with *sig*E that is part of the *sigH* regulon which is known to be differentially regulated in the presence of INH [184,137]. This sRNA is transcriptionally associated with *sig*E (see genomic context, **Figure 3.1**), thus potentially is also differentially regulated by the presence of the drug.
2. mcr3 is associated with *rrs*, an essential gene and the main gene associated with kanamycin resistance when *rrs* is mutated [185] (**Figure 3.1**). This sRNA is part of the *rrs*, *rrl* and *rrf* tri-cistronic operon (**Figure 3.1**).
3. The ASpks sRNA was recently associated with drug resistance, when bioinformatics analysis revealed that mutations in specific mRNAs clustered uniquely with various drug resistant *M. tb* clinical isolates. This sRNA is found inside the *pks*12 gene, which is significantly associated with drug resistance in *M. avium* [182]. The precise role of *pks12* in drug resistance is not clear though (**Figure 3.1**).

Previous studies have identified regions in the *M. tb* genome that, in addition to already known genes, are associated with drug resistance. In this study we analysed only those previously identified regions coding for known and suspected sRNAs. This data set therefore looked at 36 previously described target regions. Of these 36 targets, SNPs were identified in 29 targets in our genomic bank (**Figure 5.1**). The largest proportion of isolates that contained a base change in a sRNA contained it in a region inside locus Rv2048c. Rv2048C encodes for polyketide synthase 12. Polyketide synthases are increasingly identified in drug resistant *M. tb*, although

the mechanism whereby it influences drug sensitivity is not clear [150]. All these isolates were classified as drug resistant based on drug resistance tests done in our department.

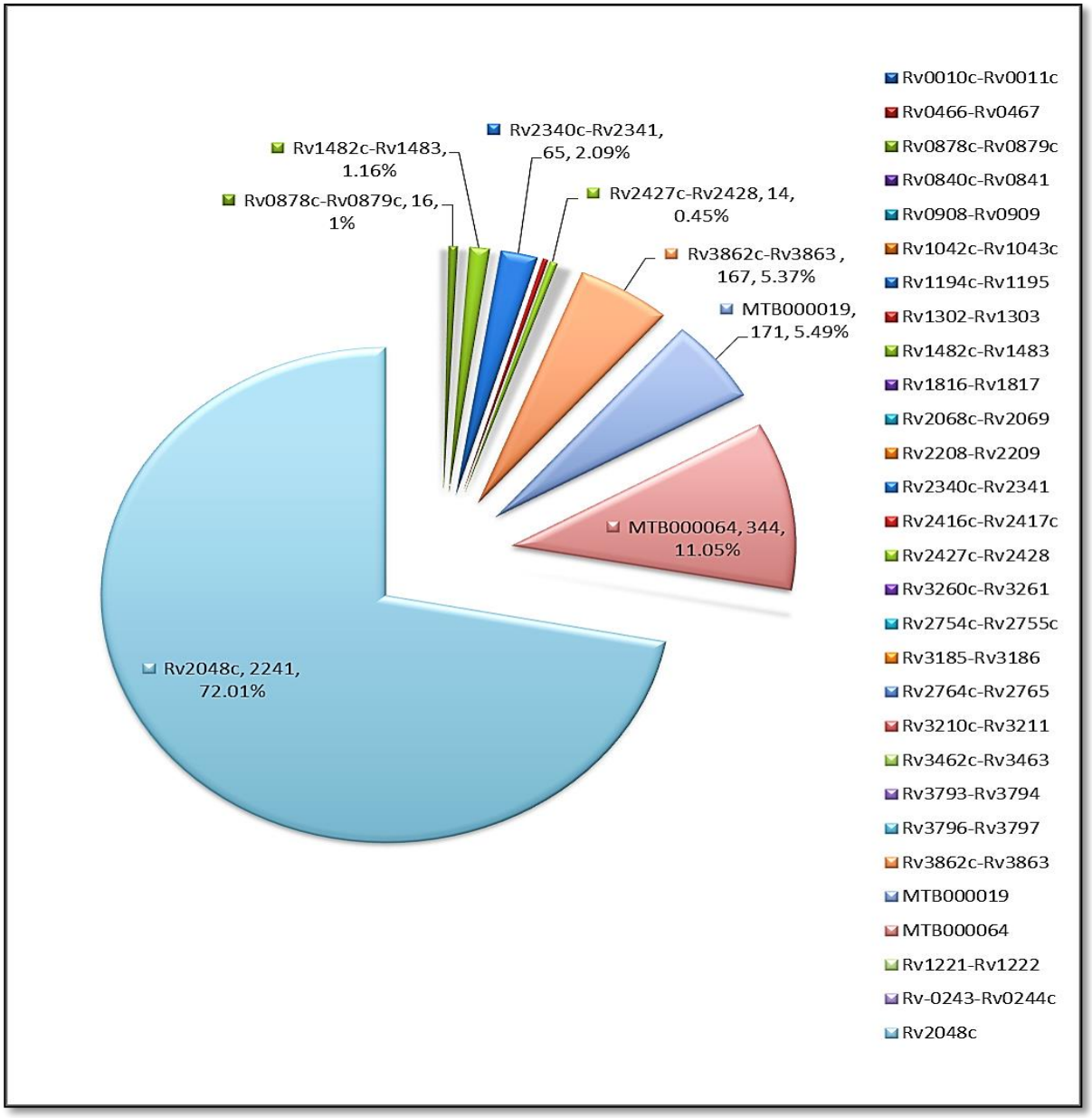

Figure 5.1. The percentage of analysed isolates which contain one or more SNP in each sRNA or putative sRNA region.

The following observations were made:

1. Isolates with SNPs in the different sRNA regions, were drug resistant. The drug resistant phenotypes ranged from Rif/INH mono-resistant isolates to XDR isolates (**Table 5.1**).

2. Further analysis of the isolates with SNPs and whose genotype was available on the database revealed the predominance of the Beijing genotype (69%) (**Figure 5.2)**. However, this result might be co-incidental as the majority of isolates in the Western Cape region are from this family. This overrepresentation might disappear if more isolates from other families are available for analysis.
3. The sRNA ASpks had the most number of SNPs, (**Tables 5.1** and **5.2)** in the analyzed isolates. This includes strains containing more than one SNP in this candidate. The mcr3 candidate is also featuring in many of the drug resistant strains.

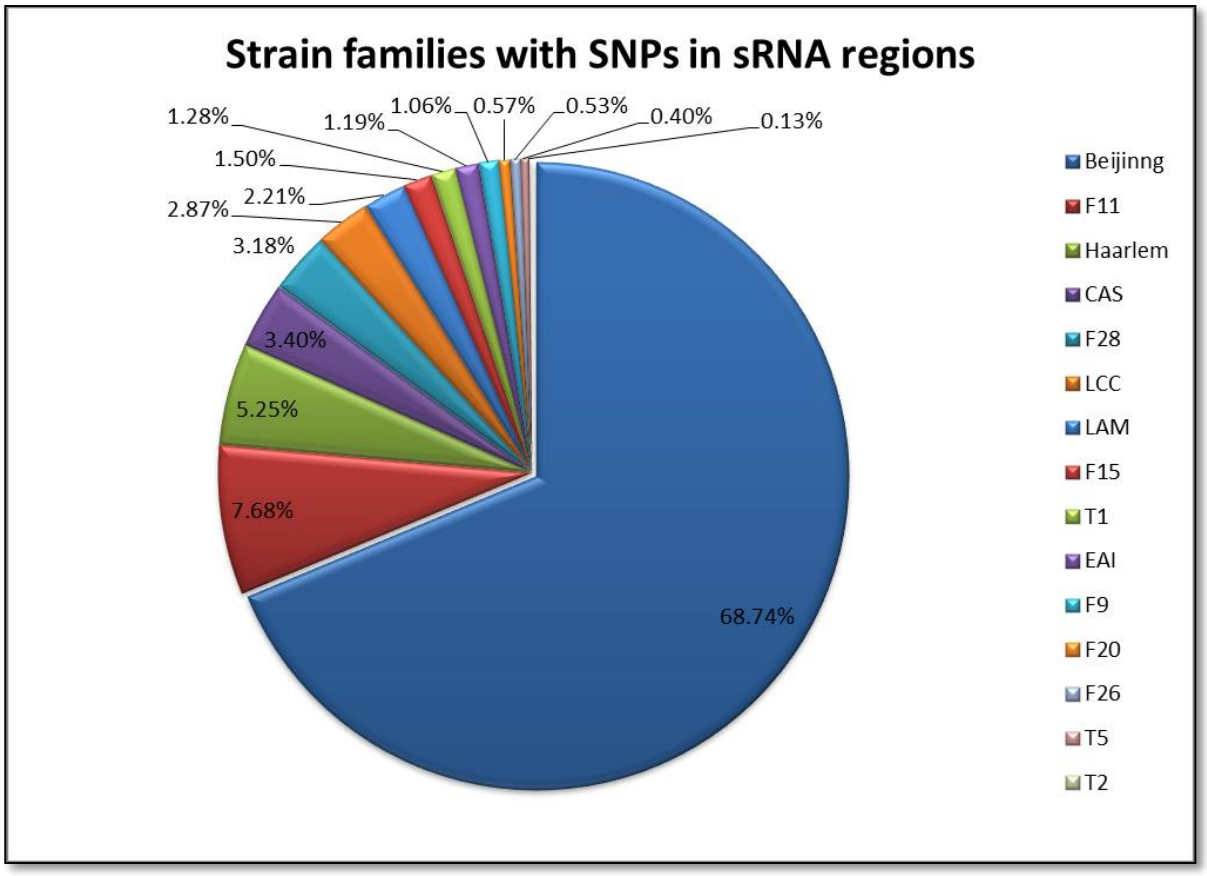

Figure 5.2 The percentage distribution of the M. tb strain families with SNPs in sRNAs.

Table 5.1 Association of mutations in sRNAs or intergenic regions of *M. tuberculosis* with genotype and drug resistance patterns. RIF – Rifampicin, INH – isoniazid.

Region/sRNA	Intergenic/ sRNA	Drug phenotypes (sensitive/resistance)	Families
Rv0010c-Rv0011c	INTERGENIC	RIF Mono, Pre XDR, MDR, N/A	Beijing X2, F9, EAI
Rv0466-Rv0467	INTERGENIC	RIF Mono (2), MDR, Pre XDR	Beijing X4, LCC
Rv0744c-Rv0745	INTERGENIC	N/A	N/A
Rv0878c-Rv0879c	INTERGENIC	RIF Mono (2), INH Mono (3), MDR (4)	Beijing, Haarlem, LCC X 12, N/A X2
Rv0840c-Rv0841	INTERGENIC	Rif Mono, INH mono, N/A	Beijing, F28 X2
Rv0908-Rv0909	INTERGENIC	N/A	N/A
Rv0920c-Rv0921	INTERGENIC	N/A	N/A
Rv1042c-Rv1043c	INTERGENIC	N/A	F28 X2
Rv1194c-Rv1195	INTERGENIC	N/A	Beijing X2, F11, Haarlem
Rv1080c-Rv1081c	INTERGENIC	N/A	N/A
Rv1302-Rv1303	INTERGENIC	N/A	Beijing
Rv1347c-Rv1348	INTERGENIC	N/A	N/A
Rv1482c-Rv1483	INTERGENIC	RIF Mono (2), INH Mono (3), MDR (6), POLY (2), Pre XDR (9), XDR (11), BLANKS (2)	
Rv1816-Rv1817	INTERGENIC	POLY, N/A	Beijing X2, LCC
Rv1900c-Rv1901	INTERGENIC	N/A	N/A
Rv2068c-Rv2069	INTERGENIC	N/A	F11
Rv2208-Rv2209	INTERGENIC	INH Mono	Beijing
Rv2340c-Rv2341	INTERGENIC	RIF Mono (12), INH Mono (4), POLY (2), MDR (9), Pre XDR (9), XDR (13), BLANKS (15)	13409 X2, Beijing X54, CAS, EAI, F11, F20, Haarlem, LCC
Rv2416c-Rv2417c	INTERGENIC	MDR, Pre XDR (3), XDR (2), BLANKS (4)	Beijing X5, F11, F26, F28, BLANKS X2
Rv2427c-Rv2428	INTERGENIC	INH Mono (2), MDR (5), BLANKS (7)	Beijing X4, CAS, EA1, F11, F20, Haarlem, LCC X3, BLANKS X2
Rv3260c-Rv3261	INTERGENIC	RIF Mono (2), INH Mono, MDR, Pre XDR	Beijing X3, F11 X2, MIX, BLANKS
Rv2754c-Rv2755c	INTERGENIC	MDR, N/A	Beijing X2, F28, LCC X2
Rv3185-Rv3186	INTERGENIC	N/A	Beijing
Rv2764c-Rv2765	INTERGENIC	N/A	F9
Rv3210c-Rv3211	INTERGENIC	N/A	Beijing
Rv3462c-Rv3463	INTERGENIC	N/A	Beijing
Rv2733c-Rv2734	INTERGENIC	N/A	N/A
Rv3793-Rv3794	INTERGENIC	MDR (2), Pre XDR	LCC X2, F11
Rv3796-Rv3797	INTERGENIC	MDR	Haarlem
Rv3862c-Rv3863	INTERGENIC	RIF Mono (25), INH Mono(15), MDR (25), POLY (5), Pre XDR (17), XDR (20), BLANKS (58)	13409, Beijing X87, CAS, EAI X2, F9 X3, F11 X23, F14, F15, F20, F26, F28 X7, Haarlem X4, LCC X22, BLANKS X13
MTB000019	sRNA-mcr3	RIF Mono (26), INH Mono (16), MDR (23), POLY (6), Pre XDR (20), XDR (14), BLANKS (66)	Beijing X78, CAS, EAI, F9 X3, F11 X26, F14, F15 X2, F20, F26, F28 X8, Harlem X4, LCC X22, mix, blanks X21, 13409
MTB000064	sRNA-mcr3	RIF Mono (26), INH Mono (15), MDR (35), POLY (4), Pre XDR (41), XDR (100), BLANKS (53), XDR (14), BLANKS X66	Beijing X172, CAS X8, EAI, F9 X2, F11 X18, F15 X2, F20, F26, F28 X7, Haarlem X15, LAM X7, LCC X25,MIX, Quebec/s, T1 X7, T2, T5 X2, BLANKS X72
Rv1221-Rv1222	sRNA-mpr6	N/A	Beijing, F11
Rv-0243-Rv0244c	sRNA-F6	MDR	N/A
Rv3651-Rv3652	sRNA-mpr18	N/A	N/A
Rv2048c	sRNA-ASpks	RIF Mono (175), INH Mono (78), MDR (224), POLY (26), Pre XDR (254), XDR (629), BLANKS (418)	Beijing X1134, CAS X65, EAI X20, F9 X14, F11 X98, F15 X29, F20 X8, F26 X8, F28 X44, Haarlem X95, LAM X43, LCC X122, MIX X6, Quebec/S X6, T1 X22, T2 X2, T5 X7, BLANKS X518

Table 5.2 SNPs identified in the three sRNA candidates used for wet bench experiments.

sRNA	BASE CHANGE	Frequency (Numbers)	Genomic loci	Notes
rrs	T-C	2	1471644	
	C-T	169	1471659	
mpr6	G-T	2	1365327	
ASpks	**C-A**	**143**	**2295685**	
	C-T	11	2295917	
	G-C	**306**	**2296042**	
	A-G	34	2296181	
	C-A	9	2296406	
	C-G	10	2296876	
	G-A	5	2297766	
	T-C	5	2297974	
	G-A	**317**	**2297976**	
	G-T	**88**	**2298194**	
	A-G	**302**	**2300237**	
	A-T	**279**	**2300546**	
	T-G	**283**	**2300552**	
	A-G	**282**	**2300555**	
	C-T	3	2300674	3 isolates had all 4 SNPs, each SNP 2 bases from the other
	T-A	3	2300676	
	T-C	3	2300678	
	G-A	3	2300680	
	T-C	5	2301089	
	G-A	75	2302033	
	C-A	10	2304068	
	A-G	9	2306306	
	T-C	10	2306472	
	T-G	3	2306630	all 3 SNPs were found in 3 isolates
	A-G	3	2306632	
	C-G	3	2306633	

5.2 ANALYSIS OF *M. smegmatis* GROWTH AND DRUG SENSITIVITY

Initial experiments were done to verify the growth pattern of the test strain under the conditions used in our assays. From this data (**Figure 5.3**), it is clear that a mid-log phase is reached between 9 and 15 hours, whereafter the growth decelerate as stationary phase ensued. From this data, we decided to use an optical density of 1.5 to represent mid-log phase *M. smegmatis* growth which achieved at approximately 12 hrs.

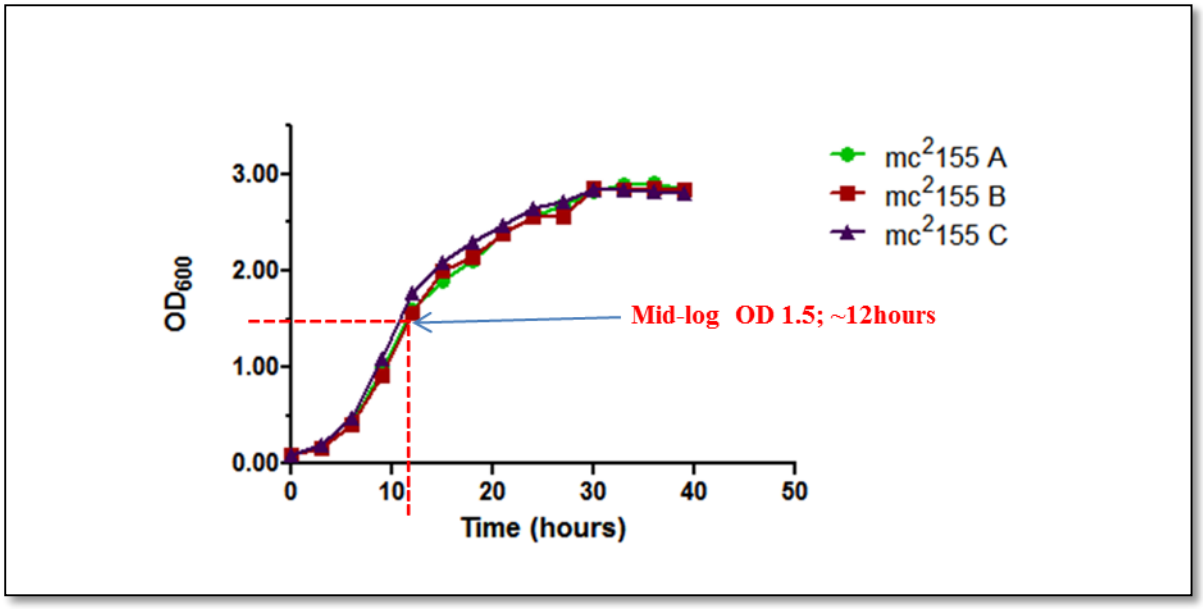

Figure 5.3 *M. smegmatis* growth curve in 7H9 media. A, B, C represents each replicate.

M. smegmatis minimum inhibitory concentration (MIC) determination

The MICs for the anti-TB drugs INH, KAN, MOX and OFL in *M. smegmatis* were determined. *M. smegmatis* cells (100ml for each drug) were grown to mid-log phase and split into five 20ml cultures. For each drug to be investigated different concentrations were added to the 20ml cultures as indicated on (**Table 5.2**) and incubated at 37°C for 6hrs. As described above, 1ml was then taken from these cultures, serial dilutions made and 10µl aliquots spotted onto antibiotic free LB agar plates. Plates were incubated for 36hrs. MICs were translated as the concentration at which very little to no growth was observed (**Figure 5.2**) and these values recorded in (**Table 5.2**).

Sub-lethal concentrations, as shown in Table 5.2, were used in further experiments.

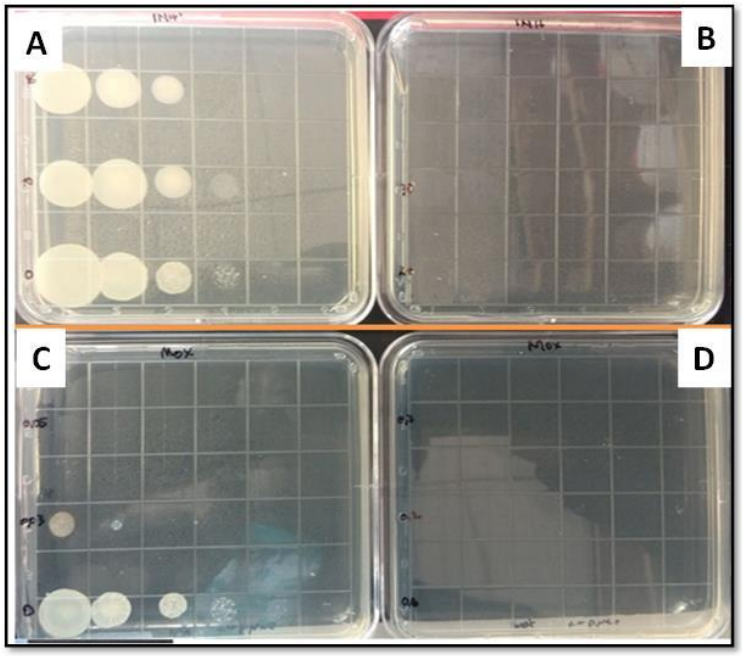

Figure 5.4 Determination of drug sensitivity of *M. smegmatis*. For this assay, serial dilutions were spotted on plate, with lowest to highest dilutions from left to right of plates. Represented in figure are results from some of the experiments, using INH and Moxifloxacin to illustrate how it was done. A: Bottom row, no drug; middle row, 5µg/ml INH; top row, 10µg/ml. B: Bottom row, 20µg/ml; middle row, 30µg/ml INH. C: .Bottom row, no drug; middle row, 0.03µg/ml MOX; top row, 0.05µg/ml MOX. D: Bottom row, 0.2µg/ml; middle row, 0.4µg/ml MOX; top row, 0.6µg/ml MOX. The same experiment was done also with Kanamycin and Ofloxacin.

Table 5.4 MICs for the different drugs and the concentrations to be used in qPCR exposure experiments.

Drug	MIC (µg/ml)	Concentration (µg/ml) determined as sub lethal
Ofloxacin	0.75	0.5
Kanamycin	0.5	0.1
Moxifloxacin	0.03	0.01
Isoniazid	24	10

5.3 EXPRESSION OF *M. TB* sRNAs IN *M. SMEGMATIS* AND THEIR EFFECT ON GROWTH

Three strategies were followed to assess expression of *M. tb* sRNA on the phenotype of *M. smegmatis*. Firstly, direct association with the transcript of a drug-resistance associated gene

was used to select the mcr3 sRNA, which is co-transcribed with the *rrs* gene, which is involved in Kanamycin resistance [186].

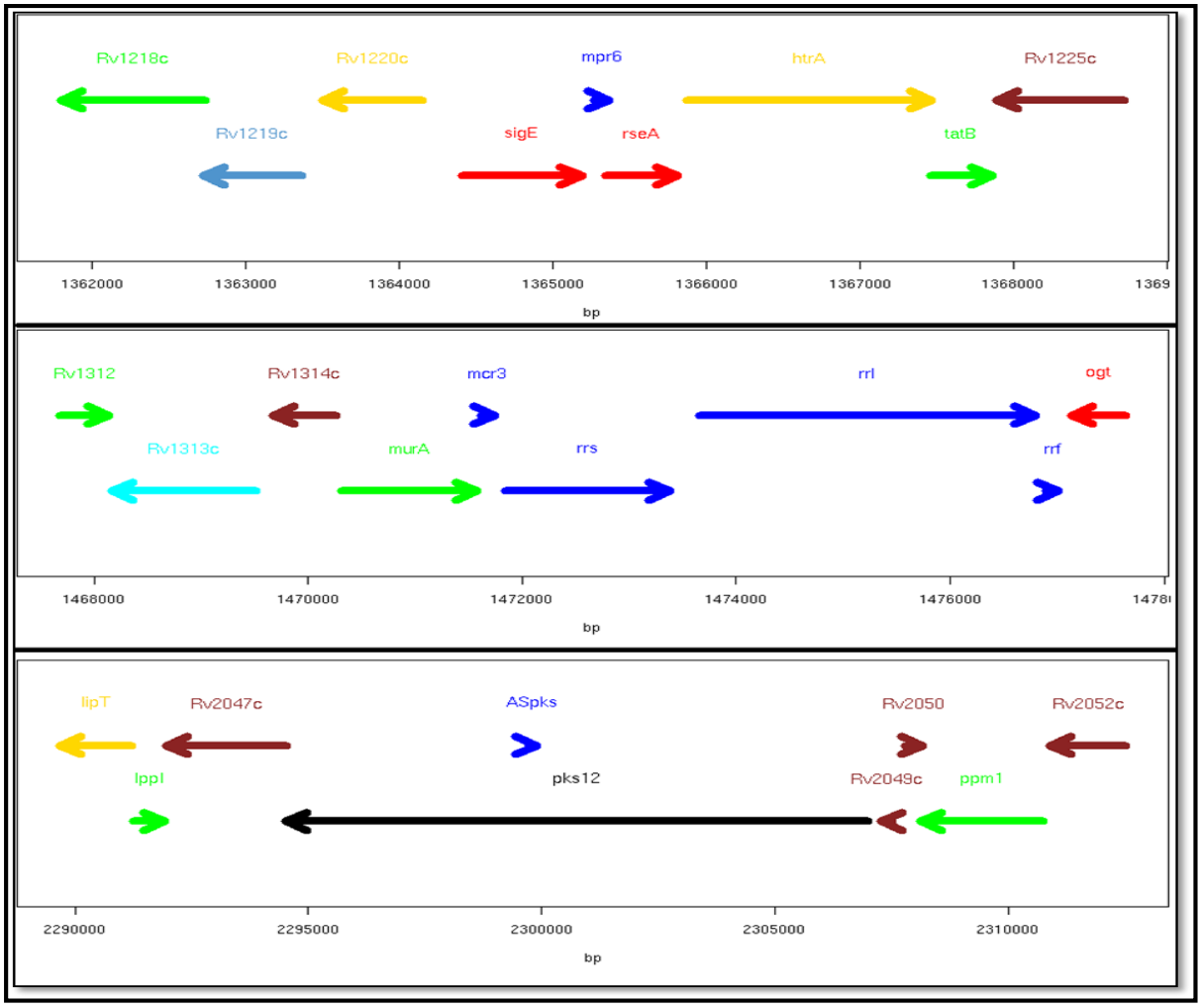

Figure 5.5 Genomic context of the sRNA mpr6, mcr3 and ASpks. The *rrs*, or 16S RNA gene is associated with kanamycin resistance; the *pks12*, gene encodes a membrane associated protein, and in which mutations are frequently found in drug resistant isolates [150,192]. Image was adapted from Tuberculist [191]

An indirect association of a small RNA and drug resistance genes was used to select the mpr6 sRNA, which is suspected to play a role in the regulation of *sigE*, which is known to affect *katG* (involved in INH resistance) transcription [187–190,152]. This sRNA is situated in the intergenic region between the sigma factor s*igE* and its anti-sigma factor *RseA*. We hypothesized that mpr6 works by facilitating the interaction of the two regulatory factors. Sigma factors play a role in regulating *M. tb* response to external stress. *SigE* is particularly important in regulating response to intracellular stress by interacting with genes involved in maintaining cell wall integrity. Furthermore, INH targets cell wall mycolic acid formation, the

M. tb sigE transcript as part of the *sigA* regulon was previously shown to increase after exposure to INH. [137,184]. The last strategy was to select a candidate identified in previous publications as a potential role-player in the development of drug resistance. For this reason, the ASpks sRNA was selected. This sRNA code is found within the *pks12* gene, which was previously shown to induce intrinsic multiple drug resistance in mycobacterial species, although the mechanism is unclear [149,147,191,143,148]. The genomic loci of the sRNA is shown in figure 5.5.

Two plasmid systems were designed to investigate the effect of sRNAs on the physiology of mycobacteria. In these two experiments we aimed at characterising the effect of sRNAs on *Mycobacterium* physiology particularly looking at growth and drug resistance. DNA from a wild-type *M. tuberculosis* sample was successfully used to amplify and clone the sRNA regions of interest (mcr3, mpr6 and ASpks). **Figure 5.6** shows the gel picture showing positive amplification of the desired regions in *E.coli* transformants of pMV306 constructs.

Construction of pMV306 and derivative plasmids containing the sRNAs

In the first strategy we designed a plasmid that allowed us to express each sRNA in *M. smegmatis*, a *Mycobacterium* model. An assessment was done to investigate the effect this over-expression would have on the growth rate of *M. smegmatis* and also on its drug resistance profile against different drugs (KAN, MOX, OFL and INH).

The amplified sRNAs with attached restriction sites were digested with the respective restriction enzymes (*EcoRV* and *Hind*III), the same enzymes used to digest pMV306 plasmid for cloning (**Figure 5.7**).

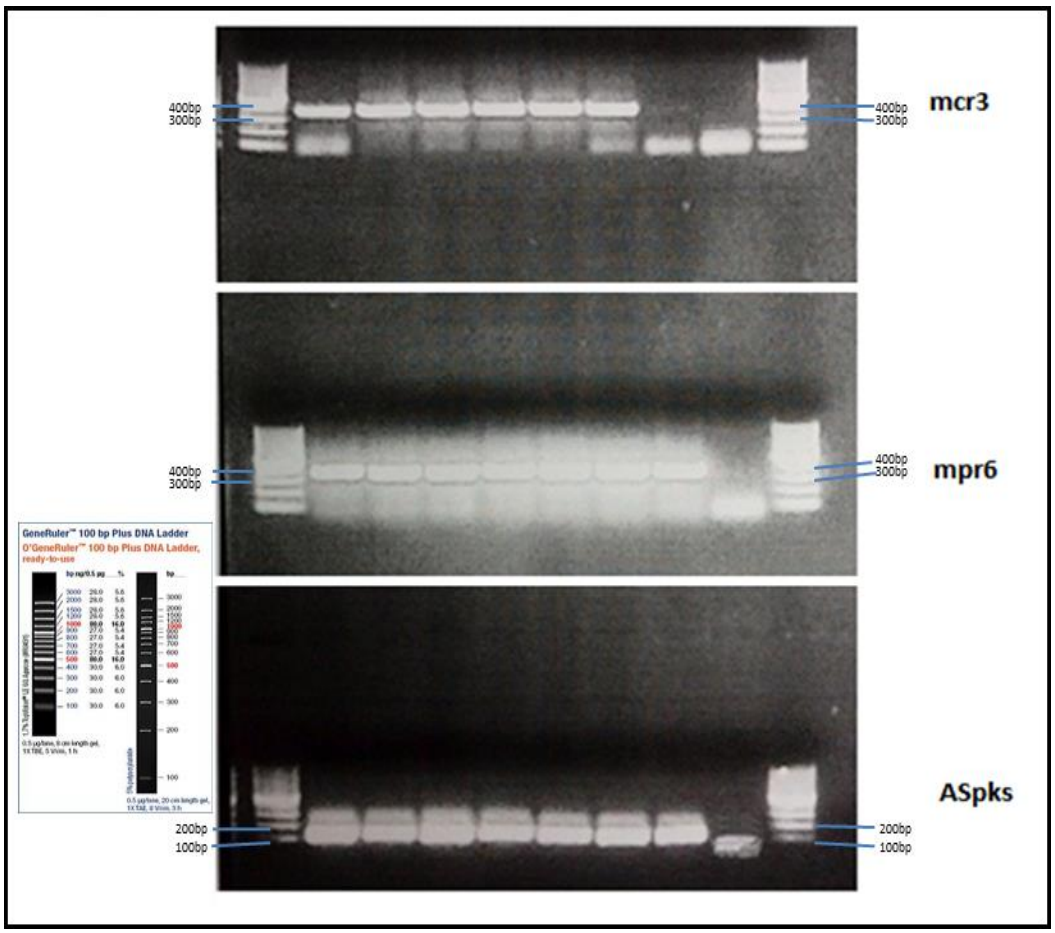

Figure 5.6 sRNA PCR amplification of target sRNAs. Multiple PCRs were performed to ensure enough yields after downstream purification steps. Expected sizes: mcr3 – 367bp, mpr6 - 351bp and ASpks – 142bp. Lane 9 is a control PCR with no template. We used the Gene Ruler 100bp DNA ladder (Catalogue number: SM0241, Thermo Fisher Scientific Inc. USA)

The digested sRNA PCR amplicons were successfully ligated to the linearized pMV306 plasmid and used to transform *E.coli.*

Colonies were picked and inoculated in liquid cultures. Thereafter plasmid extraction was done and a PCR done to confirm successful ligation of insert to plasmid (**Figure 5.6**), and also confirmed with restriction digestion reactions.

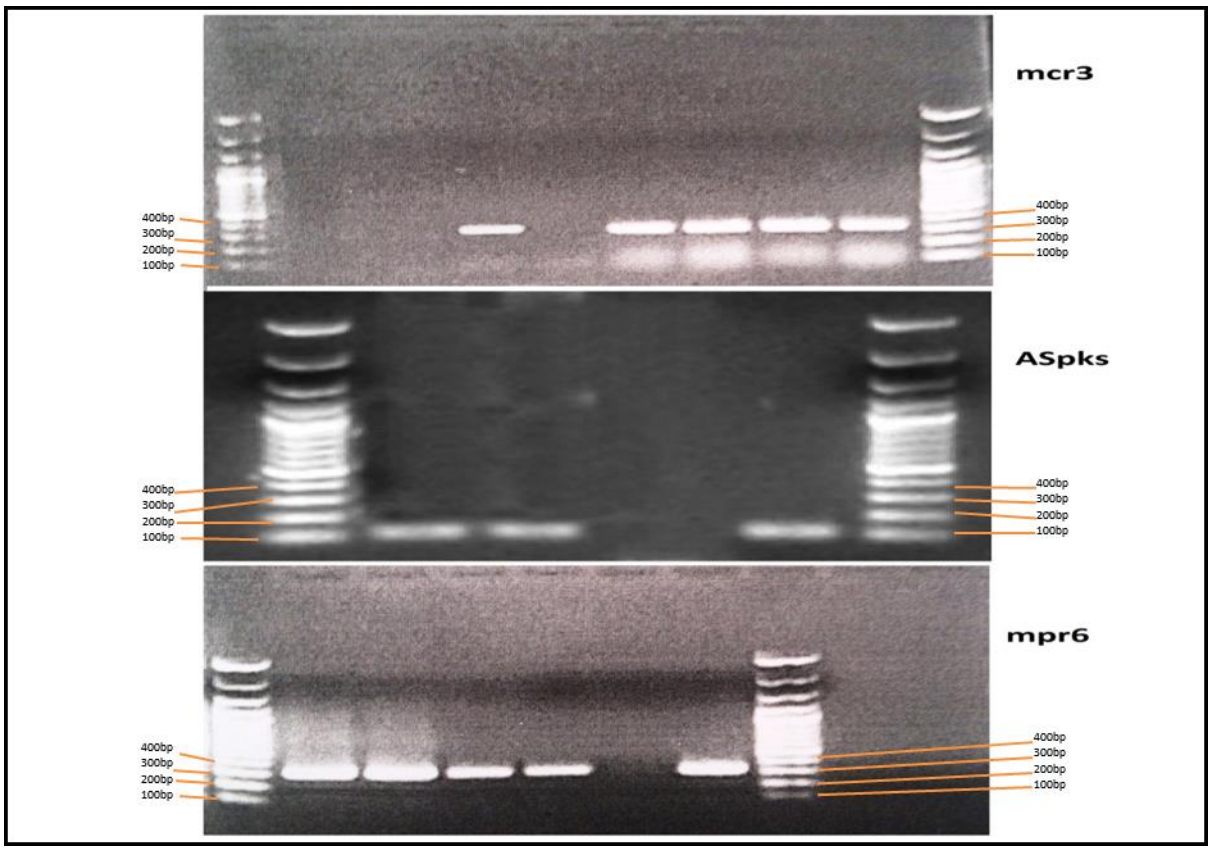

Figure 5.7PCR to confirm successful ligation of insert to plasmids pMV306. Lanes before the size marker on right contained positive control PCRs, where genomic DNA was used as template.

Restriction analysis was used to guide the construction of maps as depicted in Figure 5.8. Plasmid maps were constructed using the software developed by Magnus Manske, GENtle V1.9.4 [193].

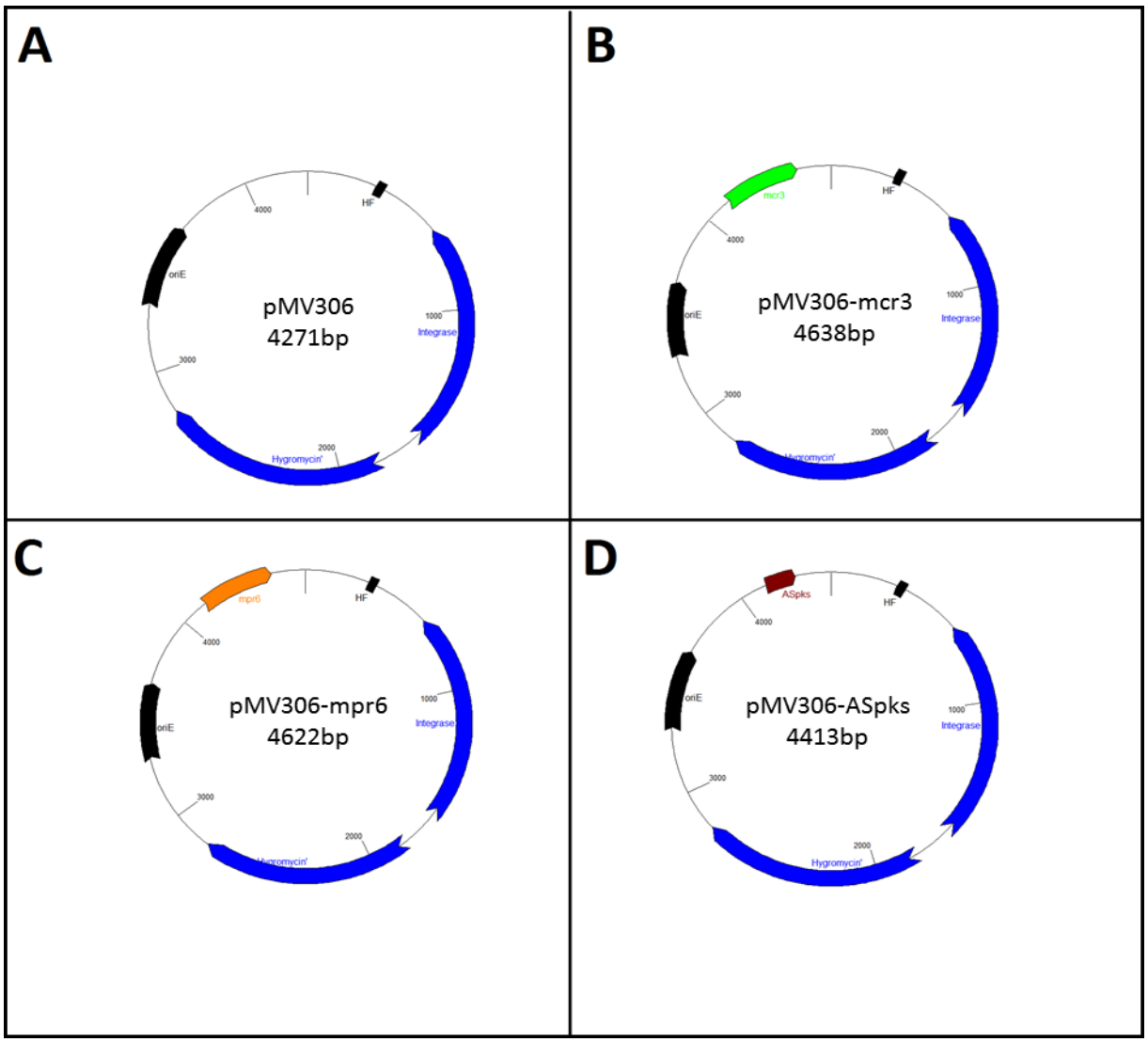

Figure 5.8 Schematic representation of pMV306-sRNA plasmid generated using GENtle, A) pMV306, B) pMV306-mcr3, C) pMV306-mpr6, D) pMV306-ASpks.

Electro competent *M. smegmatis* cells were checked for contamination using ZN-stains, showing we used pure cultures for downstream experiments (**Figure 5.9**).

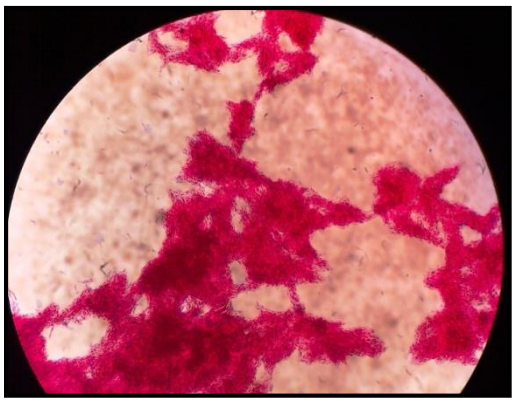

Figure 5.9 Ziehl Neelsen stain test on *M. smegmatis* electro-competent cells.

Growth assessment of *M. smegmatis* after over-expression of sRNAs

A growth curve experiment was performed to assess if the selected sRNAs had an effect on *Mycobacterium* growth rate. *M. smegmatis* mutant cells transformed with pMV306 plasmid containing a sRNA insert were grown in liquid media and growth rate measured by taking ODs at selected time intervals (3 hours) (**Table 5.5**). These ODs where used to plot growth curves for each sRNA (**Figure 5.11)**. In order to assess if the sRNAs had an effect on growth in *M. smegmatis*, controls included *M. smegmatis* transformed with only pMV306 plasmid in order to verify that any change noticed is brought about by the sRNA under investigation. To verify that the plasmid itself does not interfere, a negative control, wild-type *M. smegmatis* culture without plasmid, was also included.

These differences were only seen in the strains with plasmid constructs containing sRNA genes, and not in strains transformed with the control plasmid.

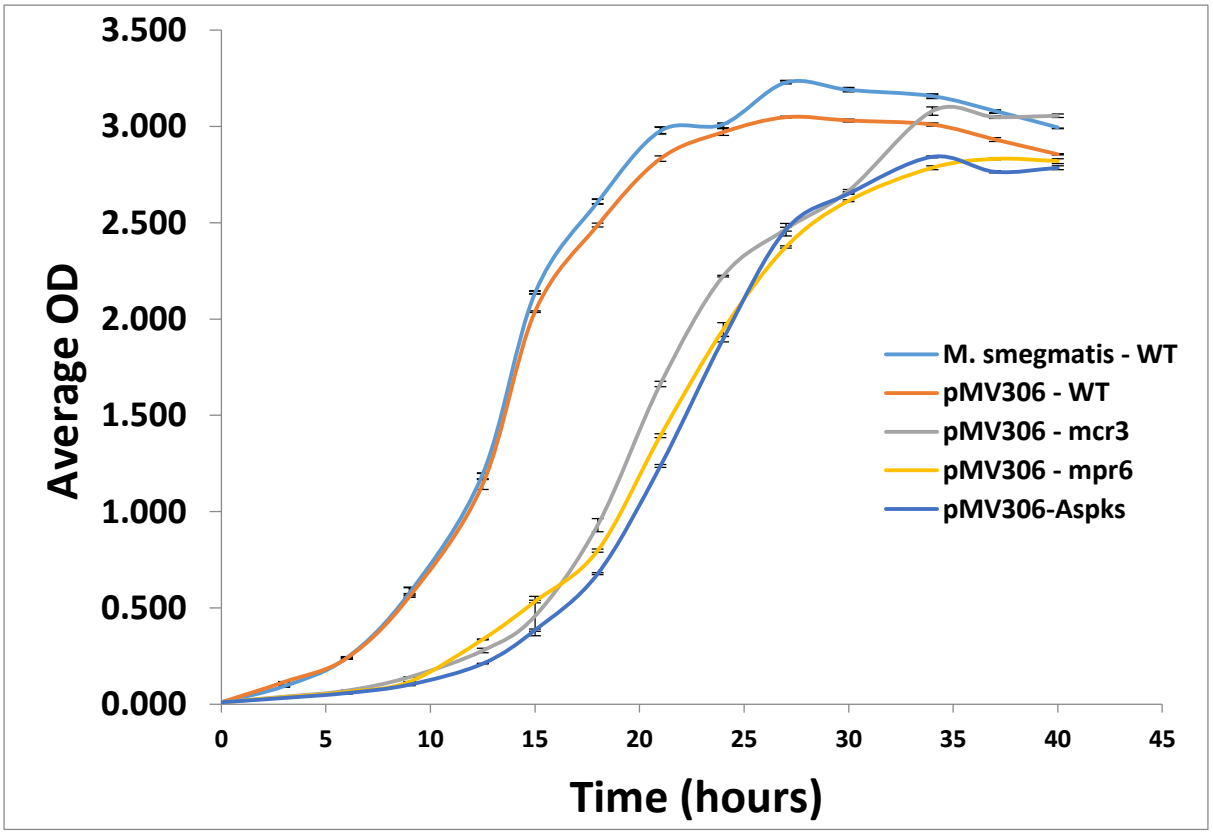

Figure 5.11 *M. smegmatis* growth curve showing the growth rate of sRNA transformed cells in comparison to wild-type cells.

It is clear that the presence of the extra copies of the RNA genes results in later onset of growth, with the ODs of control cultures almost tenfold more within three hours, whereas cultures with the sRNA copies only reached mid-log after approximately 22 hours. Control cultures also reached mid-log phase at about 12.5 hours of growth, whilst the cultures with the sRNA-expressing plasmids exhibited an OD of maximum 0.5. The onset of stationary phase was also much later in the latter cultures (25hrs compared to 15 hours).

Interestingly, this altered growth of *M. smegmatis* in the presence of expressed *M. tb* sRNA genes had been described before by Arnvig and co-workers [66]. Although this was observed with other sRNAs and not these described here, it would suggest that optimal conditions and plasmid constructs to assess sRNA function is still not clear, and that tight control of expression is needed to observe a real effect in mycobacteria.

Characterisation of drug resistance phenotypes of *M. smegmatis* cells over-expressing the sRNAs mcr3, mpr6 and ASpks

Small RNAs have been hypothetically associated with regulatory effects on the genes they target. However the mechanism by which these sRNAs function is not known. Here we set out to investigate what effect these candidate sRNAs had on the drug resistance profiles associated with the genes co-transcribed with or adjacent to them. In the first experiment, described above, no growth was seen after exposing the mutant *M. smegmatis* cells to sub-lethal (MIC) concentrations of the respective drugs up to day five (**Figure 5.12**). Similar results were obtained in experiment 2 involving re-growing drug exposed *M. smegmatis* cells in liquid cultures. These results suggest that the presence of the extra copies of a sRNA is detrimental to the bacterium, with regards to adaptation to antibiotic induced stress.

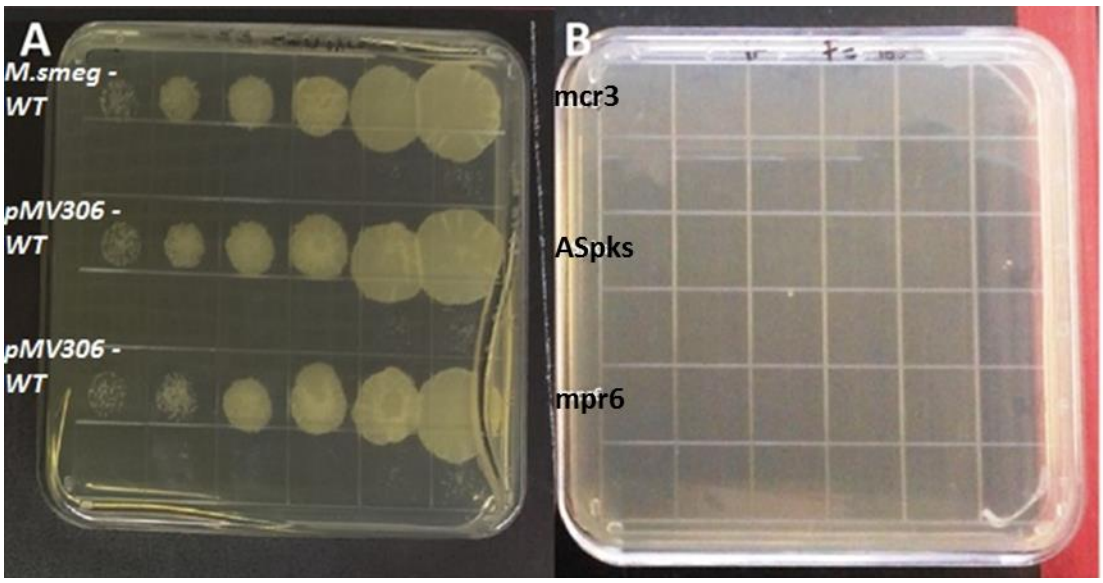

Figure 5.12 Exposure of mutant *M. smegmatis* cells to sub-lethal (MIC) concentrations. No growth was seen after exposing the mutant *M. smegmatis* cells to sub-lethal (MIC) concentrations of the respective drugs up to day five (B). A) Shows the positive control with growth.

As these were the wild type sRNAs, one is tempted to speculate that this observed negative effect, if true, would be absent if base changes are present in these sRNAs.

Controlled expression of small RNA in *M. smegmatis* using a modified plasmid containing the Tet ON/OFF promoter expression system

In an attempt to express the sRNA genes under the influence of a controllable promoter, a plasmid was created using the controllable Tet ON/OFF promoter system. However, to prevent integration at, and thus interruption of targeted chromosomal areas with functional genes, we used the fragment of the pSE100 plasmid [175], containing only the controllable promoter and an antibiotic marker. This fragment was ligated to the fragment from pMV306, which contains the integration cassette from the parent plasmid as well as the *att*P site. This new construct enabled the integration of the plasmid at the bacterial *attB* site, as opposed to homologous recombination at the mycobacterial homologous DNA found on the plasmid and the bacterial genome.

The pSE100 plasmid cut with the restriction enzymes *Cla*I and *Msc*I yielded two bands and pMV306 cut with *Msc*I also yielded two band and the desired bands, 3174bp and 1344bp respectively (**Figure 5.13**).

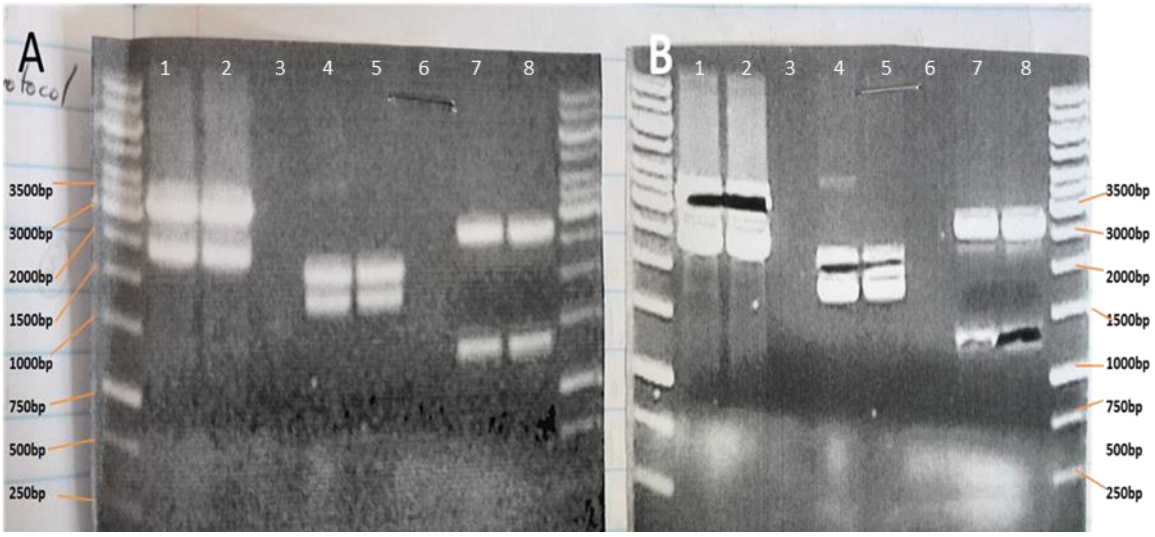

Figure 5.13 pSE100 plasmid cut with the restriction enzymes *Cla*I and *Msc*I. A). Lane 1 and 2 shows the restriction map for pSE100 plasmid cut by restriction enzymes *ClaI* and *Msc*I , lane 4 and 5 is the positive control (pMV306 cut by *SalI*) and lane 7 and 8 is the pMV306 plasmid cut by *Msc*1 for the Integrase fragment. Lane 2 in both gels:top band represent the 3174bp fragment from plasmid pSE100; Lane 7 and 8:The 1344bp fragment from plasmid pMV306 was excised as shown in B. Lane 3 and 6 are negative controls.

The pMV306 fragment could ligate in one of two ways, resulting either in plasmids with *Bam*HI restriction fragments of 3383 and 1135bp (**Figure 5.14,** lane 2), or 4149 and 359 bps (**Figure 5.14**, lanes 5 and 7). However the plasmids in lane 5 and lane 7 were the desired ligates and therefore were chosen for downstream experiment (**Figure**s **5.14**). Figure 5.15 shows the desired ligates in two orientations**.**

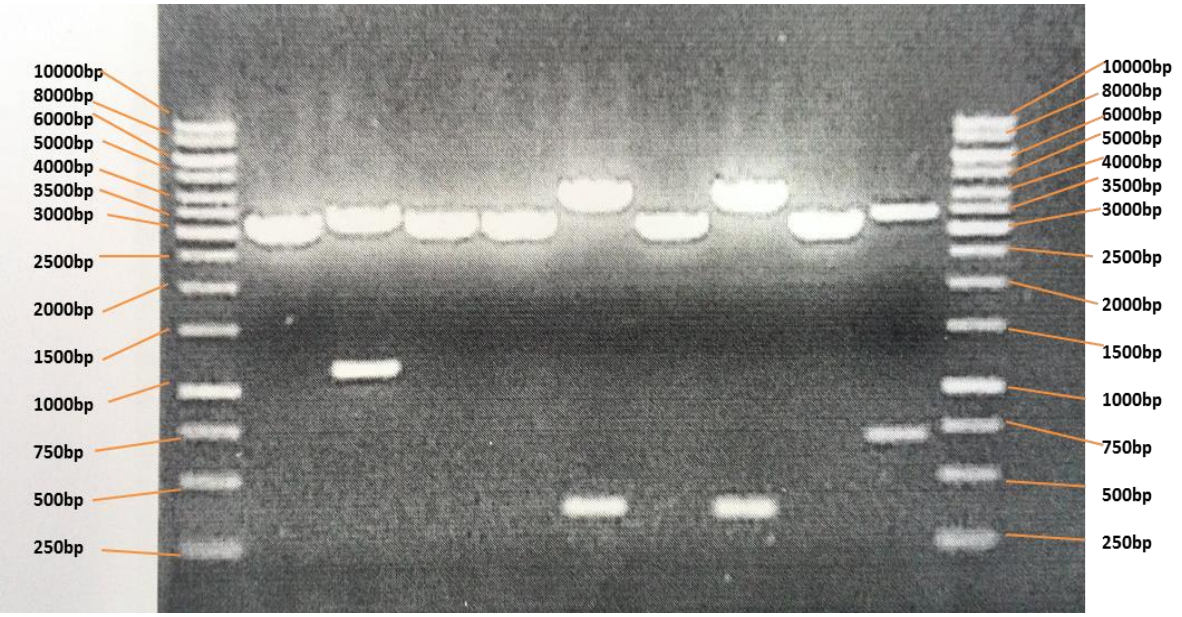

Figure 5.14 Restriction enzyme (*Bam*HI) analysis of plasmid pTKL. Lane 3, 6, 8: plasmid pTKL digested with *Bam*HI. Lanes 1 and 11, DNA marker, Lanes 2,4,5,7,9 and 10, Plasmids without insert or incorrect sized inserts (the latter potentially from multiple ligations of the insert in a single clone). The plasmid in lane 2 shows a different ligation orientation of the two fragments.

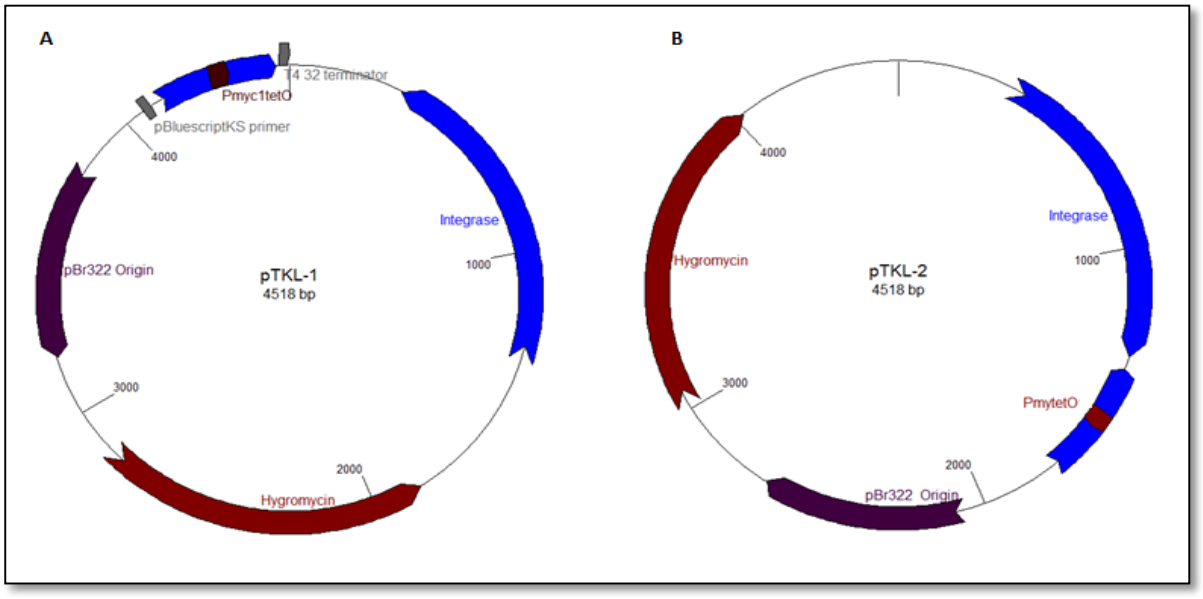

Figure 5.15 the plasmid maps (A and B) showing the newly constructed pTKL plasmid in both orientations. The pTKL plasmid shown in A was then used for all downstream experiments.

The pTKL plasmid was successfully linearized by digesting it with the restriction enzymes *EcoRV* and *Hind*III for the purposes of constructing pTKL plasmids containing the sRNAs (**Figure 5.16**)

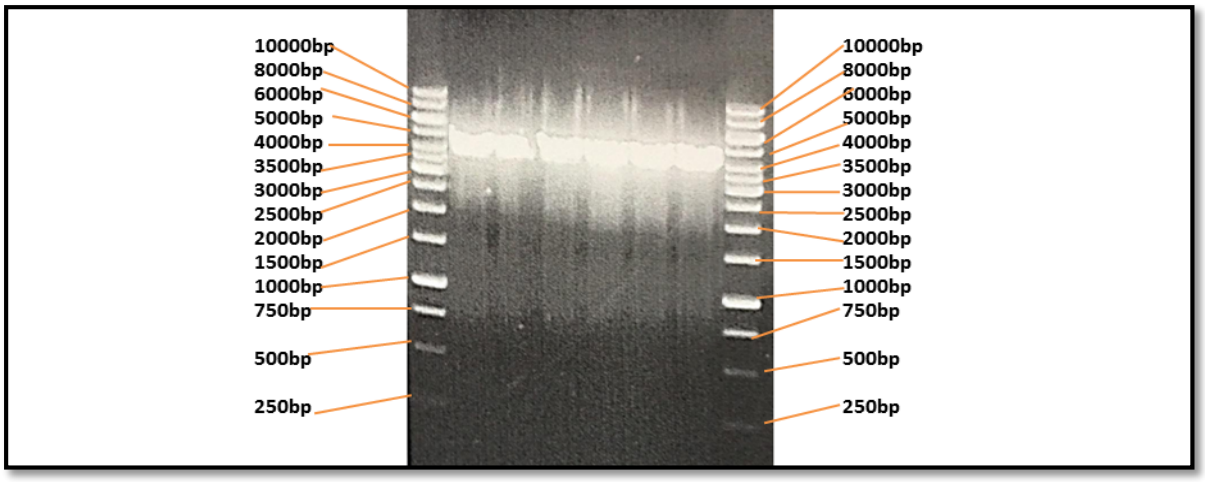

Figure 5.16 plasmid pTKL digested with *Eco*RV and *Hind*III restriction enzymes that cut the pTKL plasmid at sites one base pair from each other giving a band similar to a linearized pTKL plasmid of 4518bp length.

The respective sRNAs were ligated to the plasmids and controls added to confirm ligation efficiency and that no self-ligation occurred. The controls included a wild type pTKL plasmid digested with *Hind*III restriction enzyme which after addition of ligation reaction reagents would re-ligate and form colonies on Hygromycin plates confirming ligation efficiency (**Figure 5.17**). To confirm that the linearized pTKL plasmid used in ligation of the sRNA inserts did not self-ligate, a ligation reaction was set up for the linearized pTKL plasmid (using *Eco*RV and *Hind*III restriction enzymes). No colonies were expected on the Hygromycin plates additionally confirming that whatever is seen on the experiment plates was pTKL plasmid plus insert (**Figure 5.17**).

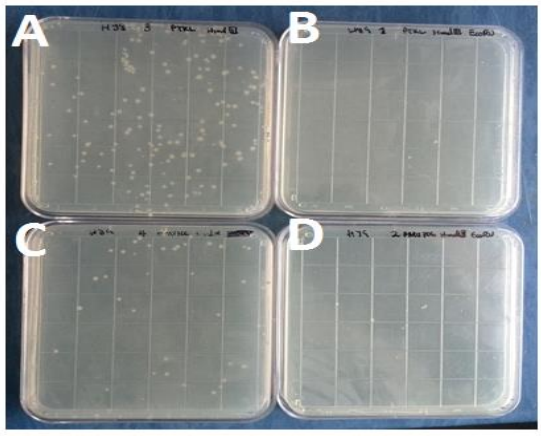

Figure 5.17 Self-ligation controls showing that there was no self-ligation that occurred in the formation of both pTKL-sRNA plasmids and pMV306-sRNA plasmids. A) pTKL plasmid linearized with *Hind*III restriction enzyme, B) pTKL plasmid linearized with both *Eco*RV and *Hind*III enzymes, C) pMV306 plasmid linearized with *Hind*III restriction enzyme, D) pMV306 plasmid linearized with both *Eco*RV and *Hind*III enzymes.

After confirming that the plasmids did not self-ligate we verified that the plasmids indeed ligated with the sRNAs and in order to do that we set up two experiments. The first experiment was PCR amplification with the respective sRNA primers. A positive amplification, as shown in Figure 5.18, indicated positive ligation of plasmid to sRNA fragment. The second verification experiment not only verified the presence of sRNA inserts but also the orientation to which they ligated to the plasmid. Figure 5.19 shows a diagrammatic presentation of the pTKL-sRNA plasmids as derived from the GENtle software [193].

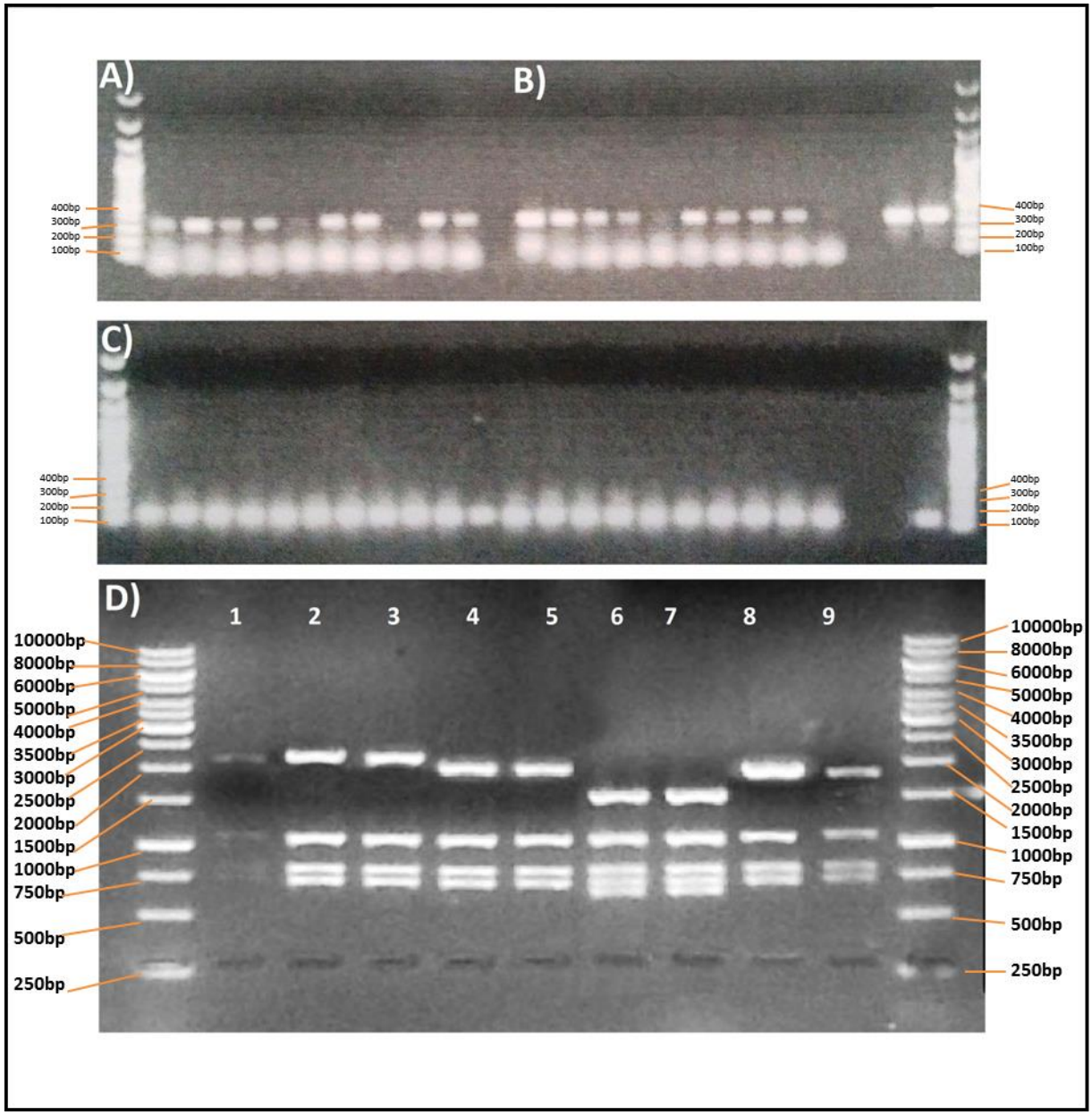

Figure 5.18 PCR and gel analysis to verify the presence of sRNA inserts in the pTKL plasmids. Positive PCR amplification of: A) mcr3 sRNA in pTKL-mcr3 plasmid, B) mpr6 sRNA in pTKL-mpr6 plasmid and C) ASpks sRNA in pTKL-ASpks plasmid. D) Restriction map obtained by digesting the pTKL-sRNA plasmids and controls: lane 1-3 is

the restriction map for pTKL-mcr3 plasmid, lane 4 and 5 is pTKL-ASpks plasmid, lane 6 and 7 – pTKL-mp6, lane 8 – wild-type pTKL plasmid and lane 9 – linearized pTKL plasmid.

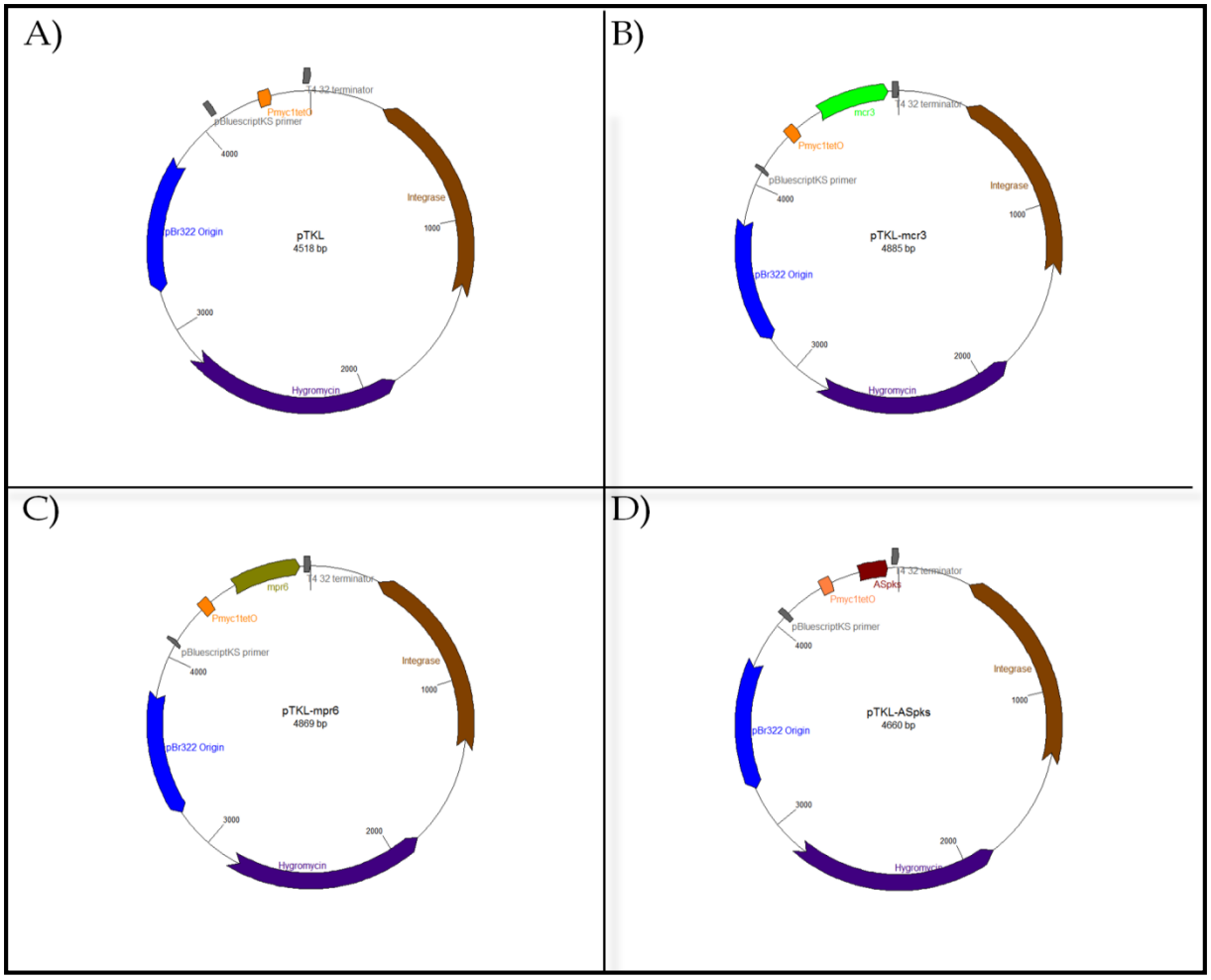

Figure 5.19 schematic representation of pTKL-sRNA plasmid generated using GENtle, A) pTKL, B) pTKL-mcr3, C) pTKL-mpr6, D) pTKL-ASpks.

Constructs containing sRNA gene candidates were successfully created and the presence of inserts was confirmed with colony PCR and restriction enzyme analysis. Unfortunately, attempts to transform *M. smegmatis* cells with pTKL and its derivatives, using the same method as described above to transform pMV306 constructs into *M. smegmatis*, failed repeatedly. No growth on plates pTKL + insert plasmids were obtained, whereas growth was seen on positive controls (cells transformed with control plasmids). In an attempt to troubleshoot several alterations to the original protocol for transformations were done, but to no avail (**Figure 5.20)**.

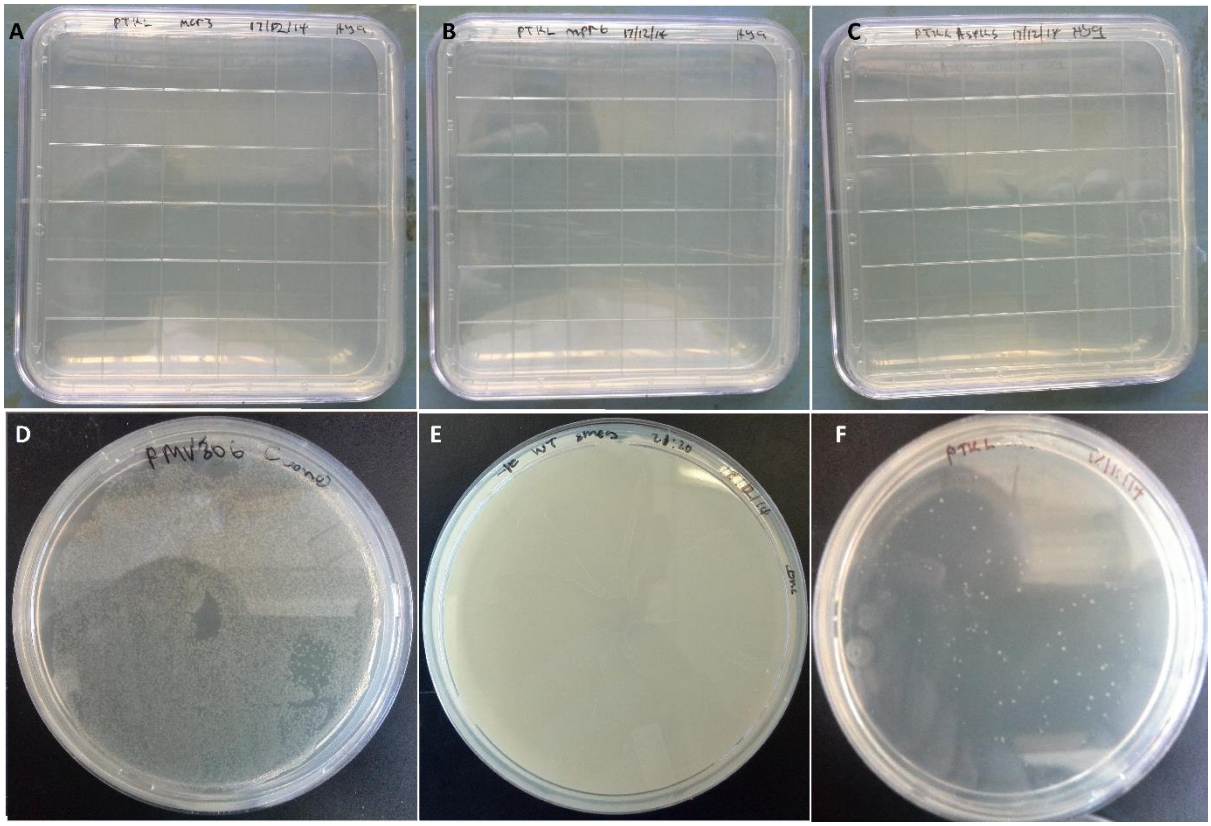

Figure 5.20 Growth of *M. smegmatis.* Growth not seen on plates containing and pTKL + sRNA plasmids (A = Mcr3, B = Mpr6, C = ASpks), but growth was seen on positive controls (D = wild type pMV306, E = wild type *M. smegmatis*, F = wild type pTKL). Square plates were used to ensure isolation of discrete colonies.

Although the plasmids were constructed successfully, their presence somehow was lethal to the bacterium. This strategy with pTKL constructs was thus abandoned after few attempts. Due to lack of time, no further experiments were done to find conditions where transformed cells could be isolated. Future work could include optimisation of culture protocols, perhaps in the presence and absence of promoter induction/repression with anhydrous tetracycline

5.4 EXPRESSION OF MCR3 AND ASPKS sRNAs IN *M. TB* EXPOSED TO SUB-LETHAL CONCENTRATIONS OF KANAMYCIN OR OFLOXACIN

The objective of this experiment was to observe if the sRNA could be used for quantitative PCR (qPCR). For this experiment we chose two sRNAs, mcr3 and ASpks, based on their association with genes previously published to be involved in drug resistant mechanisms. The mcr3 sRNA is located upstream (promoter region) of the *M. tb rrs* gene, known to be involved

in aminoglycoside synthesis [147]. Kanamycin was chosen as the test drug for the mcr3 sRNA and Moxifloxacin was chosen for the ASpks sRNA.

RNA was successfully extracted using the above modified protocol (Chapter 4) even though the quantity was low it was clean and intact, from both bacterial strains. The quality and quantity of the sRNAs was assessed using the Nanodrop 2000c (Nano Drop technologies, Inc., USA) to ensure the absence of contaminating reagents. However, when used in proper experiments, assessment of purity and quality should be done with better instrumentation.

Table 5.5 shows the sRNA and the bigger RNA concentrations obtained in the drug exposure experiment for each drug (KAN- mcr3 and MOX-ASpks).

Table 5.5 RNA concentration from the in vitro drug exposure experiments Concentration as determined by the Nanodrop 2000c software.

Sample I.D	KANAMYCIN CONC ng/µl		Sample ID	MOXIFLOXACIN CONC ng/µl	
	sRNA	bigger RNA		sRNA	bigger RNA
A1	23.7	22.0	**A0**	7.0	20.0
A2	18.8	19.6	**A0.015**	16.5	47.7
A3	14.7	35.0	**A0.03**	38.1	28.7
A4	77.1	25.4	**A1**	13.4	28.5
B1	49.6	16.2	**B0**	98.6	22.8
B2	6.8	17.9	**B0.015**	37.6	35.8
B3	62.3	27.3	**B0.03**	8.0	46.8
B4	66.1	35.1	**B1**	27.1	32.7
C1	9.7	11.5	**C0**	27.7	17.5
C2	8.9	19.2	**C0.015**	33.5	24.1
C3	18.4	26.1	**C0.03**	36.6	44.1
C4	16.4	30.5	**C1**	21.5	42.2

RNA concentrations appear to be low, but that can be optimised by adapting elution volumes when eluting nucleic acid from columns. Secondly, depending on your downstream application, the amount of input material should be optimised as well.

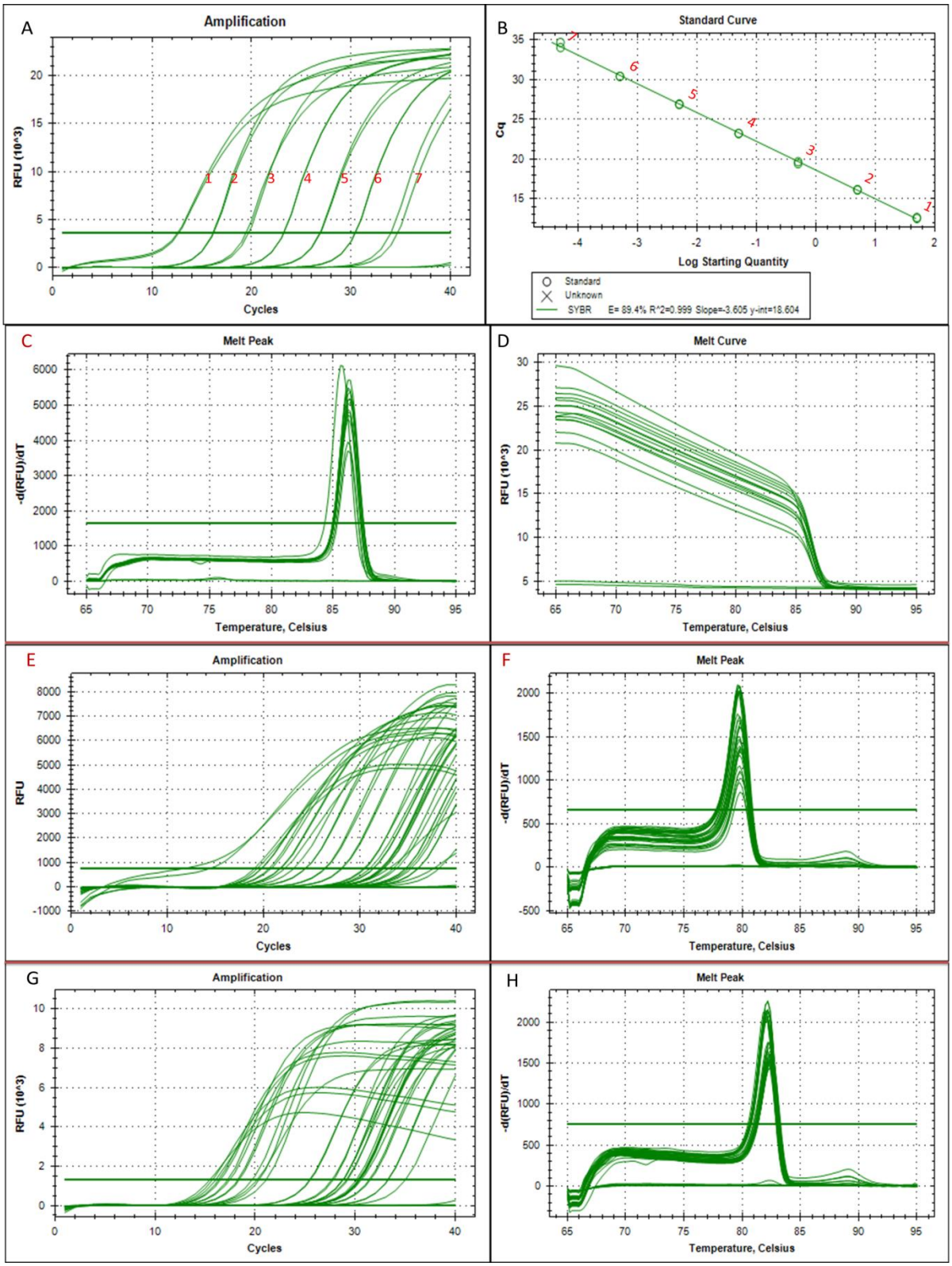

Figure 5.21 A representation of amplification curves (A-standard curve, E-mcr3 and G-ASpks). B) Shows the dilution factors and standard curve, using gDNA and sigA, constructed from the respective amplification curves A) by the Bio-Rad CFX96™ Real-Time System. Curves E and F represent mcr3 sRNA and G and H represent ASpks sRNA.

It is also noteworthy that biological replicates did not yield more or less the same yield, indicating that more optimisation is needed and perhaps more experience. Melting peak analysis for the qPCR products as analysed on the Bio-Rad equipment showed no primer dimers and each gene had a unique melting peak (**Figure 5.21**). The standard curve generated had a best-fit linear regression line with a slope value of 3.6, an efficiency of 89.4% and a goodness-of-fit value of $R^{\wedge 2}$ 0.999.

We did not quantitate transcription levels of the candidates here, but rather used this exercise to demonstrate that good quality RNA was isolated and that it could be used to make cDNA. This method thus, with further optimisation could be used to prepare sRNAs for downstream technologies.

CHAPTER 6: DISCUSSION

Drug resistance in TB is a growing problem, with "totally drug resistant" *M. tb* strains reported worldwide [194–198]. The biology of how resistance is initiated and regulated in *M. tb* is still poorly understood; therefore a lot still needs to be done to elucidate this phenomenon of drug resistance. There has been a recent increase in studies focused on identifying sRNAs in Mycobacterial species (and other bacteria), but so little has been done to attach functionality to these sRNAs. Putative associations of drug resistance and sRNA genes are appearing in the literature, including in *M. tuberculosis*. Many drug-resistance associated gene functions in clinically significant bacteria are not well comprehended and perhaps sRNA putative regulatory functions, if identified, will help to diminish that knowledge gap.

We have therefore attempted, for the first time, to investigate the physiological effect some sRNAs play in *Mycobacterial* species. In order to do this, we have combined basic bioinformatics analysis processes with wet bench experiments to identify sRNAs with potential roles in drug resistance mechanisms.

Firstly, we used an in-house Python Script developed in our department (by Ruben van der Merwe - unpublished data) to screen for SNPs in regions with previously described sRNAs. We aimed to discern if suspected SNPs were present, which strains they were found in, and if it they were found solely in drug resistant isolates. This analysis revealed that most isolates with SNPs in candidate sRNAs were from the Beijing genotype, which was dominant as compared to other genotypes. This genotype is the most successful genotype, originally identified in Asia [199], but has successfully established itself as the most frequent isolate found amongst TB-infections worldwide, including South Africa. The observed predominance in our study set thus could therefore be an artefact of its predominance in infections in this country. The Beijing family indeed is wide spread and owing to its genetic homogeneity several studies have hypothesised that it possesses a selective advantage over other strains. A study in 2005 showed that the percentage of the Beijing genotype increased with drug resistance as compared

to other strains contributing to 75% of the drug resistant strains as per their study in central Asia [200]. In the Eastern Cape province of South Africa, a study, done by Marisa Klopper et al, in our group showed that MDR-TB strains were diverse but pre-XDR and XDR-TB patients showed a genetic background restricted to the Beijing genotype [198]

. This adds more weight to the above hypothesis and also adds to hypothesise that Beijing strains more readily acquire drug resistance associated mutations than other strains and are capable of spreading extensively regardless of fitness costs associated with mutations. However a study in Sweden in 2010 contradicted this hypothesis as they showed that in their population there were no extensive outbreaks of this strain as compared to other strains [201]. Based on previous studies it is clear that the strain family Beijing has a potential selective advantage over other strains and from our findings sRNAs might have a role in explaining this selective advantage.

Of the 36 regions previously associated with drug resistance and/or drug resistance genes, SNPs were we identified found in 29 of these SNPs in regions in our genome bank, all in consisting of drug resistant isolates. How and if these SNPs contribute to drug resistance is still unclear. These SNPs could play a role in drug resistance given that drug resistance arises from mutations in genes associated with drug action and therefore warranty further investigations. The largest frequency of SNPs between isolates were found in the sRNAs namely; mcr3 - in the promoter region of a known drug resistant gene *rrs* and ASpks a sRNA confined in the *pks12* gene were found in most of the drug resistant isolates in our clinical isolate genome bank. However, larger data sets need to be screened for statistical associations to be made. However, since these SNPs were only found in resistant isolates, it does suggest an association with resistance, as previously suggested. The frequency of these SNPs could translate to function; future studies would therefore include investigating whether these SNPs have an effect on gene/sRNA function using wet bench techniques. Within the *rrs* gene, which is co-transcribed with *mcr*3, are SNPs shown to be associated with drug resistance. It is thus possible

that mcr3 could play a role in aminoglycoside resistance. Although base changes in non-coding regions does not result in changes in encoded products, such as found with non-synonymous SNPs in genes, it is becoming increasingly clear that even synonymous SNPs might have functional consequences in organisms. Could these include alterations in the way drug-resistance genes are regulated? It is with this in mind, and the need to fill this knowledge gap, that we selected three candidates for functional characterisations. These candidate sRNAs were selected based on their associated with known drug resistant genes and or pathways.

In a bid to investigate the effect of sRNAs on the physiology of *M. tb*, we over-expressed candidate sRNAs (mcr3, ASpks and mpr6) using the expression vector pMV306 in *M. smegmatis* and observed its effect on growth and drug resistance profile. Although we did not check for expression of the cloned candidates in *M. smegmatis*, the difference in growth indicates that the presence of the extra copies had a detrimental phenotypical effect on the bacterium. There was an approximately 5 fold decrease in growth rate when comparing the sRNA-transformed cell mutants with the wild type *M. smegmatis* strain or the strain transformed with the control plasmid. These results show that these sRNAs have an effect on the physiology of mycobacteria. For mcr3 and ASpks it can be speculated that their targets would be the genes with which they are co-transcribed with, namely *rrs* and *pks12* respectively. As for mpr6, its effect on physiology might be more significant as its genomic locus is between two interacting sigma factors which are important regulators of various characteristics of bacteria.

Unfortunately, we failed to utilize the plasmid pTKL, where the sRNA targets were cloned directly after a controllable promoter. For these plasmids to be useful, it appears that one has to find conditions where even its transformation into mycobacteria is tightly controlled by conditions that enable growth and prevents the lethal phenotype observed in this study. Thus controlled expression of the sRNAs in *M. smegmatis* using a modified plasmid containing the

Tet ON/OFF promoter expression system still has to be optimized. Controlled expression vectors give you the flexibility of controlling expression levels particularly for toxic candidates or when a gene is only required to be expressed at particular time points and/in response to particular stimuli. This is effective in the analysis of gene function and hence why it is important to make this optimisation.

In terms of sRNA effect on drug resistance: due to time constraints we only investigated whether expressing the sRNAs would induce sensitivity of *M. smegmatis* to anti-TB drugs. No growth was observed in both liquid and solid media suggesting that over expressing the sRNAs potentially affect/induce sensitivity to drugs in in vitro cultures. However, such studies need to be done under various conditions that more closely resemble in vivo conditions and not laboratory in vitro conditions. More still needs to be done before concluding the effect of sRNAs on drug resistance.

Lastly, we devised a shorter protocol to isolate various RNA species from mycobacteria. This protocol yielded intact RNA, both bigger transcripts and the sRNAs, by using a combination of columns to separate the different RNAs and concentrate the sRNAs. This RNA was DNA free and cDNA was successfully made from it. Unfortunately low yields were obtained. Therefore this approach is still limited by the amount of culture that maximally is allowed per flask in our biosafety laboratory. In order to yield increased concentrations of sRNAs, one can look at combining extractions from cells cultured in individual flasks under identical conditions, and extracted at exact identical time-points. Due to low yields, the sRNAs could not be used to construct standard curves. An alternative for standard curves thus could be to use genomic DNA instead of the cDNA.

Another problem that was encountered when attempting to quantitate sRNA expression with quantitative qPCR is the fact that no sRNA reference “genes” have been established yet. Other researchers published results using 16S mRNA or other mRNA species as reference genes. No

doubt though, that experimental error can occur as all methods separate the isolation of sRNAs and mRNA steps. Unfortunately one cannot spike sRNAs with RNA to generate internal standards, as sRNA species are mostly associated with mRNA transcripts, either before or after genes or in the middle of transcripts.

More sensitive and accurate methods of quantification such as the Bio-Analyzer should be used instead of the Nanodrop only.

Although the level of the sRNAs were too low to generate a proper standard curve with cDNA made from sRNAs, there was enough template to generate a signal for each species we assessed. Instead of using the sRNA-derived cDNA, we decided to use genomic DNA to prepare a standard curve for DNA extractions as the amount of cDNA did not yield enough signal with the higher dilutions, thus yield is a problem. Various dilution ranges were tested to find a set of concentrations that would accommodate Ct values obtained from sRNA derived-cDNA.

CHAPTER 7: CONCLUSION

With the growing problem of drug resistant TB it is paramount that new therapies evolve to combat new *M. tb* strains. From the discovery of the first TB drugs to date there has been a few new drugs added to TB treatment regimens. Partly this is a result of the lack of a comprehensive understanding of how *M. tb* metabolic functions occur.

Small RNA studies could shed light as to how regulatory mechanisms in mycobacteria occur filling a big part in the current knowledge gap. They could provide new drug targets turning drug resistant strains into sensitive strains and help treat patients regarded as therapeutically destitute. The field of sRNAs is new and hence more still needs to be done such as optimising protocols for some of the most used techniques such as qPCR and cloning.

This study has confirmed the presence of previous identified SNPs in sRNAs of drug resistant isolates. We also have shown that the sRNAs investigated here have an effect on the physiology of mycobacteria, both in the presence and absence of antibiotics, although more in-depth studies need to be performed to confirm these observations. Most importantly, we designed a protocol to isolate sRNAs successfully using commercial kits. Future studies include further optimisation of this protocols for use with sRNAs in order to increase yield.

The limitations of the study included the lack of sRNA reference genes for qPCR; because sRNAs are associated with mRNA transcripts and given that their target mRNAs are still unknown current reference genes are not suitable for qPCR. The lack of understanding of how sRNAs are transcribed and post transcriptionally modified made it difficult to design proper cloning strategies. Therefore in future actually investigating to see if a particular sRNA is transcribed instead of assuming it is based on a change of phenotype would be paramount in elucidating function of sRNAs. Generally sRNAs constitute a small percentage of the total RNA therefore proper extraction methods are required that utilise less harsh chemicals like phenol and have very few steps at which RNA can be lost or contaminated. The method we designed is promising in the sense that it eliminates unnecessary steps and chemicals that lead

to the loss of RNA although it still needs to be improved in terms of consistency as one can notice the wide ranges of the RNA concentration. Although the quantity was low the quality was good, one can work around the low quantity issue by increasing the amount of starter culture. And because of this low quantity of sRNAs a standard curve using the sRNAs could not be created but we showed that one can use genomic DNA to create a standard curve. It is unfortunate that although we were successful in creating a controllable plasmid (pTKL) with the derivative sRNAs we could not successfully transform *M. smegmatis*; in our future studies we aim to optimise transformation by changing the reaction conditions etc.

Overall, future studies will be to understand the role of sRNAs in mycobacterial virulence and drug resistance. Also we would like to optimise major molecular biology processes like cloning, qPCR etc., and also identify more sRNAs that have any association with drug resistant pathways.

APPENDICES

APPENDIX 1 - Protocols

1. Plasmid pTKL construction

Protocol A

1. Dephosphorylation of pSE100 fragment (company name)

- 7µl H_2O, 10µl pSE100 fragment, 2µl Phosphatase Buffer and 1µl Phosphatase
- Incubate at 37˚C for 10min, inactivate enzyme at 75˚C for 2min and chill on ice for 2min

2. Ligation of the pMV306 fragment to the dephosphorylated (dephos) pSE100 fragment

- 4µl of the dephos pSE100, 5µl pMV306 fragment, 10µl 2X ligation buffer (Thermo Scientific) and 1µl T_4 DNA ligase (Thermo Scientific).
- Incubate at 22˚C or room temperature (RTM) for 20min.

3. Clean up

- Clean up the reaction from 2 with the Zymoclean Gel clean-up kit (contains a DNA binding buffer that can be used to clean up DNA from contaminating reagents.

4. Blunting of the pSE100 end cut with *Cla*I

- 8µl pSE100-pMV306 fragment, 10µl 2X reaction buffer (Thermo Scientific) and 1µl Blunting enzyme (Thermo Scientific).
- Incubate at 70˚C for 5min and chill on ice for 2min.

5. Final Ligation (circularizing the plasmid) and transformation

- Add 1µl T_4 DNA ligase to the reaction mixture from (4).
- Incubate at 22˚C or room temperature (RTM) for 20min.

- Use the whole mixture (20µl) to transform 200 µl of chemical competent *E.coli* cells (protocol in appendix--).
- Plate on LB agar hygromycin plates and incubate over night at 37˚C

Protocol B

1. Blunting of pSE100 fragment

- 9µl pMV306 fragment, 10µl 2X reaction buffer (Thermo Scientific) and 1µl Blunting enzyme (Thermo Scientific).
- Incubate at 70˚C for 5min and chill on ice for 2min.

2. Clean up

- Clean up the reaction from 1 with the Zymoclean Gel clean-up kit (contains a DNA binding buffer that can be used to clean up DNA from contaminating reagents.

3. Dephosphorylation of pSE100 blunted fragment (company name)

- 17µl dephos pSE100 fragment, 2µl Phosphatase Buffer and 1µl Phosphatase
- Incubate at 37˚C for 10min, inactivate enzyme at 75˚C for 2min and chill on ice for 2min

4. Ligation of the pMV306 fragment to the dephosphorylated (dephos) pSE100 fragment

- Add 1µl T_4 DNA ligase to the reaction mixture from (4).
- Incubate at 22˚C or room temperature (RTM) for 20min.

Use the whole mixture (20µl) to transform 200 µl of chemical competent *E.coli* cells (protocol in appendix--).

BIBLIOGRAPHY

1. Global tuberculosis control - epidemiology, strategy, financing, WHO. WHO | Global tuberculosis control - epidemiology, strategy, financing. *WHO* (2009). at <http://www.who.int/tb/publications/global_report/2009/en/index.html>
2. TB Online. TB Online - South Africa: Infectious XDR-TB patients being sent home. (2014). at <http://www.tbonline.info/posts/2014/4/7/south-africa-infectious-xdr-tb-patients-being-sent/>
3. CDC | TB | Drug-Resistant TB. CDC | TB | Drug-Resistant TB. (2015). at <http://www.cdc.gov/tb/topic/drtb/>
4. Storz, G., Vogel, J. & Wassarman, K. M. Regulation by small RNAs in bacteria: expanding frontiers. *Mol. Cell* **43,** 880–891 (2011).
5. Gottesman, S. Micros for microbes: non-coding regulatory RNAs in bacteria. *Trends Genet. TIG* **21,** 399–404 (2005).
6. Lee, R. C., Feinbaum, R. L. & Ambros, V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75,** 843–854 (1993).
7. Ambros, V. The functions of animal microRNAs. *Nature* **431,** 350–355 (2004).
8. Chen, K. & Rajewsky, N. The evolution of gene regulation by transcription factors and microRNAs. *Nat. Rev. Genet.* **8,** 93–103 (2007).
9. Tjaden, B. *et al.* Target prediction for small, noncoding RNAs in bacteria. *Nucleic Acids Res.* **34,** 2791–2802 (2006).
10. Wassarman, K. M., Repoila, F., Rosenow, C., Storz, G. & Gottesman, S. Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev.* **15,** 1637–1651 (2001).

11. Gottesman, S. The small RNA regulators of Escherichia coli: roles and mechanisms*. *Annu. Rev. Microbiol.* **58,** 303–328 (2004).

12. Stougaard, P., Molin, S. & Nordström, K. RNAs involved in copy-number control and incompatibility of plasmid R1. *Proc. Natl. Acad. Sci. U. S. A.* **78,** 6008–6012 (1981).

13. Tomizawa, J., Itoh, T., Selzer, G. & Som, T. Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA. *Proc. Natl. Acad. Sci. U. S. A.* **78,** 1421–1425 (1981).

14. Simons, R. W. & Kleckner, N. Translational control of IS10 transposition. *Cell* **34,** 683–691 (1983).

15. Mizuno, T., Chou, M. Y. & Inouye, M. A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). *Proc. Natl. Acad. Sci.* **81,** 1966–1970 (1984).

16. Hershberg, R., Altuvia, S. & Margalit, H. A survey of small RNA-encoding genes in Escherichia coli. *Nucleic Acids Res.* **31,** 1813–1820 (2003).

17. Romby, P., Vandenesch, F. & Wagner, E. G. H. The role of RNAs in the regulation of virulence-gene expression. *Curr. Opin. Microbiol.* **9,** 229–236 (2006).

18. Altier, C., Suyemoto, M., Ruiz, A. I., Burnham, K. D. & Maurer, R. Characterization of two novel regulatory genes affecting Salmonella invasion gene expression. *Mol. Microbiol.* **35,** 635–646 (2000).

19. Figueroa-Bossi, N., Valentini, M., Malleret, L., Fiorini, F. & Bossi, L. Caught at its own game: regulatory small RNA inactivated by an inducible transcript mimicking its target. *Genes Dev.* **23,** 2004–2015 (2009).

20. Iyer, V. R., Sharma, R., Pathania, R. & Navani, N. K. Small RNAs of Pathogenic Bacteria: Not Small Enough to be Overlooked for Therapeutics. *Mol. Cell. Pharmacol.* **4,** 17–30 (2012).

21. Lenz, D. H. *et al.* The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in Vibrio harveyi and Vibrio cholerae. *Cell* **118,** 69–82 (2004).

22. Lenz, D. H., Miller, M. B., Zhu, J., Kulkarni, R. V. & Bassler, B. L. CsrA and three redundant small RNAs regulate quorum sensing in Vibrio cholerae. *Mol. Microbiol.* **58,** 1186–1202 (2005).

23. Mraheil, M. A. *et al.* The intracellular sRNA transcriptome of Listeria monocytogenes during growth in macrophages. *Nucleic Acids Res.* **39,** 4235–4248 (2011).

24. Schilling, D., Findeiss, S., Richter, A. S., Taylor, J. A. & Gerischer, U. The small RNA Aar in Acinetobacter baylyi: a putative regulator of amino acid metabolism. *Arch. Microbiol.* **192,** 691–702 (2010).

25. Wilderman, P. J. *et al.* Identification of tandem duplicate regulatory small RNAs in Pseudomonas aeruginosa involved in iron homeostasis. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 9792–9797 (2004).

26. Altuvia, S. Identification of bacterial small non-coding RNAs: experimental approaches. *Curr. Opin. Microbiol.* **10,** 257–261 (2007).

27. Billoud, B., Kontic, M. & Viari, A. Palingol: a declarative programming language to describe nucleic acids' secondary structures and to scan sequence database. *Nucleic Acids Res.* **24,** 1395–1403 (1996).

28. Chen, S. *et al.* A bioinformatics based approach to discover small RNA genes in the Escherichia coli genome. *Biosystems* **65,** 157–177 (2002).

29. Delihas, N. Annotation and evolutionary relationships of a small regulatory RNA gene micF and its target ompF in Yersinia species. *BMC Microbiol.* **3,** 13 (2003).

30. Carter, R. J., Dubchak, I. & Holbrook, S. R. A computational approach to identify genes for functional RNAs in genomic sequences. *Nucleic Acids Res.* **29,** 3928–3938 (2001).

31. Salari, R. *et al.* smyRNA: a novel Ab initio ncRNA gene finder. *PloS One* **4,** e5433 (2009).

32. Salzberg, S. L., Delcher, A. L., Kasif, S. & White, O. Microbial gene identification using interpolated Markov models. *Nucleic Acids Res.* **26,** 544–548 (1998).

33. Delcher, A. L., Harmon, D., Kasif, S., White, O. & Salzberg, S. L. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.* **27,** 4636–4641 (1999).

34. Delcher, A. L., Bratke, K. A., Powers, E. C. & Salzberg, S. L. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinforma. Oxf. Engl.* **23,** 673–679 (2007).

35. Lukashin, A. V. & Borodovsky, M. GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res.* **26,** 1107–1115 (1998).

36. Besemer, J. & Borodovsky, M. Heuristic approach to deriving models for gene finding. *Nucleic Acids Res.* **27,** 3911–3920 (1999).

37. Besemer, J., Lomsadze, A. & Borodovsky, M. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res.* **29,** 2607–2618 (2001).

38. Lee, Y. *et al.* The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425,** 415–419 (2003).

39. Han, J. *et al.* Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* **125,** 887–901 (2006).

40. Kim, V. N., Han, J. & Siomi, M. C. Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell Biol.* **10,** 126–139 (2009).

41. Chekulaeva, M. & Filipowicz, W. Mechanisms of miRNA-mediated post-transcriptional regulation in animal cells. *Curr. Opin. Cell Biol.* **21,** 452–460 (2009).

42. Arnvig, K. B. *et al.* Sequence-based analysis uncovers an abundance of non-coding RNA in the total transcriptome of Mycobacterium tuberculosis. *PLoS Pathog.* **7,** e1002342 (2011).

43. Arnvig, K. & Young, D. Non-coding RNA and its potential role in Mycobacterium tuberculosis pathogenesis. *RNA Biol.* **9,** 427–436 (2012).

44. Waters, L. S. & Storz, G. Regulatory RNAs in bacteria. *Cell* **136,** 615–628 (2009).

45. Brantl, S. Regulatory mechanisms employed by cis-encoded antisense RNAs. *Curr. Opin. Microbiol.* **10,** 102–109 (2007).

46. Schumacher, M. A., Pearson, R. F., Møller, T., Valentin-Hansen, P. & Brennan, R. G. Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. *EMBO J.* **21,** 3546–3556 (2002).

47. Picard, F., Dressaire, C., Girbal, L. & Cocaign-Bousquet, M. Examination of post-transcriptional regulations in prokaryotes by integrative biology. *C. R. Biol.* **332,** 958–973 (2009).

48. Sobrero, P. & Valverde, C. The bacterial protein Hfq: much more than a mere RNA-binding factor. *Crit. Rev. Microbiol.* **38,** 276–299 (2012).

49. Arluison, V. *et al.* Spectroscopic observation of RNA chaperone activities of Hfq in post-transcriptional regulation by a small non-coding RNA. *Nucleic Acids Res.* **35,** 999–1006 (2007).

50. Li, S.-K. *et al.* Identification of small RNAs in Mycobacterium smegmatis using heterologous Hfq. *RNA* **19,** 74–84 (2013).

51. Storz, G., Vogel, J. & Wassarman, K. M. Regulation by small RNAs in bacteria: expanding frontiers. *Mol. Cell* **43,** 880–891 (2011).

52. Man, S. *et al.* Artificial trans-encoded small non-coding RNAs specifically silence the selected gene expression in bacteria. *Nucleic Acids Res.* (2011). doi:10.1093/nar/gkr034

53. Folichon, M. The poly(A) binding protein Hfq protects RNA from RNase E and exoribonucleolytic degradation. *Nucleic Acids Res.* **31,** 7302–7310 (2003).

54. Henderson, C. A. *et al.* Hfq binding changes the structure of Escherichia coli small noncoding RNAs OxyS and RprA, which are involved in the riboregulation of rpoS. *RNA* **19,** 1089–1104 (2013).

55. Arluison, V. *et al.* Three-dimensional Structures of Fibrillar Sm Proteins: Hfq and Other Sm-like Proteins. *J. Mol. Biol.* **356,** 86–96 (2006).

56. Geissmann, T. A. & Touati, D. Hfq, a new chaperoning role: binding to messenger RNA determines access for small RNA regulator. *EMBO J.* **23,** 396–405 (2004).

57. Mohanty, B. K., Maples, V. F. & Kushner, S. R. The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in Escherichia coli. *Mol. Microbiol.* **54,** 905–920 (2004).

58. Moll, I., Afonyushkin, T., Vytvytska, O., Kaberdin, V. R. & Bläsi, U. Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. *RNA N. Y. N* **9,** 1308–1314 (2003).

59. Jousselin, A., Metzinger, L. & Felden, B. On the facultative requirement of the bacterial RNA chaperone, Hfq. *Trends Microbiol.* **17,** 399–405 (2009).

60. Nielsen, J. S. *et al.* Defining a role for Hfq in Gram-positive bacteria: evidence for Hfq-dependent antisense regulation in Listeria monocytogenes. *Nucleic Acids Res.* **38,** 907–919 (2010).

61. Sonnleitner, E. *et al.* Reduced virulence of a hfq mutant of Pseudomonas aeruginosa O1. *Microb. Pathog.* **35,** 217–228 (2003).

62. Chiang, M.-K., Lu, M.-C., Liu, L.-C., Lin, C.-T. & Lai, Y.-C. Impact of Hfq on Global Gene Expression and Virulence in Klebsiella pneumoniae. *PLoS ONE* **6,** e22248 (2011).

63. Guisbert, E., Rhodius, V. A., Ahuja, N., Witkin, E. & Gross, C. A. Hfq Modulates the σE-Mediated Envelope Stress Response and the σ32-Mediated Cytoplasmic Stress Response in Escherichia coli. *J. Bacteriol.* **189,** 1963–1973 (2007).

64. Blumenthal, T. & Carmichael, G. G. RNA replication: function and structure of Qbeta-replicase. *Annu. Rev. Biochem.* **48,** 525–548 (1979).

65. Kidmose, R. T., Vasiliev, N. N., Chetverin, A. B., Andersen, G. R. & Knudsen, C. R. Structure of the Qβ replicase, an RNA-dependent RNA polymerase consisting of viral and host proteins. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 10884–10889 (2010).

66. Arnvig, K. B. & Young, D. B. Identification of small RNAs in Mycobacterium tuberculosis. *Mol. Microbiol.* **73,** 397–408 (2009).

67. Thomason, M. K. & Storz, G. Bacterial antisense RNAs: How many are there and what are they doing? *Annu. Rev. Genet.* **44,** 167–188 (2010).

68. Faubladier, M. & Bouché, J. P. Division inhibition gene dicF of Escherichia coli reveals a widespread group of prophage sequences in bacterial genomes. *J. Bacteriol.* **176,** 1150–1156 (1994).

69. Cayrol, B. *et al.* Auto-assembly of E. coli DsrA small noncoding RNA: Molecular characteristics and functional consequences. *RNA Biol.* **6,** 434–445 (2009).

70. Urban, J. H. & Vogel, J. Translational control and target recognition by Escherichia coli small RNAs in vivo. *Nucleic Acids Res.* **35,** 1018–1037 (2007).

71. Lease, R. A. & Belfort, M. A trans-acting RNA as a control switch in Escherichia coli: DsrA modulates function by forming alternative structures. *Proc. Natl. Acad. Sci. U. S. A.* **97,** 9919–9924 (2000).

72. Majdalani, N., Cunning, C., Sledjeski, D., Elliott, T. & Gottesman, S. DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 12462–12467 (1998).

73. Sledjeski, D. D., Whitman, C. & Zhang, A. Hfq is necessary for regulation by the untranslated RNA DsrA. *J. Bacteriol.* **183,** 1997–2005 (2001).

74. Poole, R. K., Gibson, F. & Wu, G. The cydD gene product, component of a heterodimeric ABC transporter, is required for assembly of periplasmic cytochrome c and of cytochrome bd in Escherichia coli. *FEMS Microbiol. Lett.* **117,** 217–223 (1994).

75. Boysen, A., Møller-Jensen, J., Kallipolitis, B., Valentin-Hansen, P. & Overgaard, M. Translational regulation of gene expression by an anaerobically induced small non-coding RNA in Escherichia coli. *J. Biol. Chem.* **285,** 10690–10702 (2010).

76. Opdyke, J. A., Kang, J.-G. & Storz, G. GadY, a small-RNA regulator of acid response genes in Escherichia coli. *J. Bacteriol.* **186,** 6698–6705 (2004).

77. Urbanowski, M. L., Stauffer, L. T. & Stauffer, G. V. The gcvB gene encodes a small untranslated RNA involved in expression of the dipeptide and oligopeptide transport systems in Escherichia coli. *Mol. Microbiol.* **37,** 856–868 (2000).

78. Pulvermacher, S. C., Stauffer, L. T. & Stauffer, G. V. Role of the Escherichia coli Hfq protein in GcvB regulation of oppA and dppA mRNAs. *Microbiol. Read. Engl.* **155,** 115–123 (2009).

79. Sharma, C. M., Darfeuille, F., Plantinga, T. H. & Vogel, J. A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. *Genes Dev.* **21,** 2804–2817 (2007).

80. Pulvermacher, S. C., Stauffer, L. T. & Stauffer, G. V. The small RNA GcvB regulates sstT mRNA expression in Escherichia coli. *J. Bacteriol.* **191,** 238–248 (2009).

81. Pulvermacher, S. C., Stauffer, L. T. & Stauffer, G. V. Role of the sRNA GcvB in regulation of cycA in Escherichia coli. *Microbiol. Read. Engl.* **155,** 106–114 (2009).

82. Padalon-Brauch, G. *et al.* Small RNAs encoded within genetic islands of Salmonella typhimurium show host-induced expression and role in virulence. *Nucleic Acids Res.* **36,** 1913–1927 (2008).

83. Rasmussen, A. A. *et al.* Regulation of ompA mRNA stability: the role of a small regulatory RNA in growth phase-dependent control. *Mol. Microbiol.* **58,** 1421–1429 (2005).

84. Udekwu, K. I. *et al.* Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA. *Genes Dev.* **19,** 2355–2366 (2005).

85. Chen, S., Zhang, A., Blyn, L. B. & Storz, G. MicC, a second small-RNA regulator of Omp protein expression in Escherichia coli. *J. Bacteriol.* **186,** 6689–6697 (2004).

86. Delihas, N. & Forst, S. MicF: an antisense RNA gene involved in response of Escherichia coli to global stress factors. *J. Mol. Biol.* **313,** 1–12 (2001).

87. Guillier, M. & Gottesman, S. Remodelling of the Escherichia coli outer membrane by two small regulatory RNAs. *Mol. Microbiol.* **59,** 231–247 (2006).

88. Guillier, M. & Gottesman, S. The 5' end of two redundant sRNAs is involved in the regulation of multiple targets, including their own regulator. *Nucleic Acids Res.* **36,** 6781–6794 (2008).

89. Holmqvist, E. *et al.* Two antisense RNAs target the transcriptional regulator CsgD to inhibit curli synthesis. *EMBO J.* **29,** 1840–1850 (2010).

90. Altuvia, S., Zhang, A., Argaman, L., Tiwari, A. & Storz, G. The Escherichia coli OxyS regulatory RNA represses fhlA translation by blocking ribosome binding. *EMBO J.* **17,** 6069–6075 (1998).

91. Svenningsen, S. L., Tu, K. C. & Bassler, B. L. Gene dosage compensation calibrates four regulatory RNAs to control Vibrio cholerae quorum sensing. *EMBO J.* **28,** 429–439 (2009).

92. Tu, K. C., Waters, C. M., Svenningsen, S. L. & Bassler, B. L. A small-RNA-mediated negative feedback loop controls quorum-sensing dynamics in Vibrio harveyi. *Mol. Microbiol.* **70,** 896–907 (2008).

93. Updegrove, T., Wilf, N., Sun, X. & Wartell, R. M. Effect of Hfq on RprA-rpoS mRNA pairing: Hfq-RNA binding and the influence of the 5' rpoS mRNA leader region. *Biochemistry (Mosc.)* **47,** 11184–11195 (2008).

94. Jones, A. M., Goodwill, A. & Elliott, T. Limited role for the DsrA and RprA regulatory RNAs in rpoS regulation in Salmonella enterica. *J. Bacteriol.* **188,** 5077–5088 (2006).

95. Mika, F. *et al.* Targeting of csgD by the small regulatory RNA RprA links stationary phase, biofilm formation and cell envelope stress in Escherichia coli. *Mol. Microbiol.* **84,** 51–65 (2012).

96. Papenfort, K. *et al.* SigmaE-dependent small RNAs of Salmonella respond to membrane stress by accelerating global omp mRNA decay. *Mol. Microbiol.* **62,** 1674–1688 (2006).

97. Antal, M., Bordeau, V., Douchin, V. & Felden, B. A small bacterial RNA regulates a putative ABC transporter. *J. Biol. Chem.* **280,** 7901–7908 (2005).

98. Argaman, L. *et al.* Novel small RNA-encoding genes in the intergenic regions of Escherichia coli. *Curr. Biol. CB* **11,** 941–950 (2001).

99. Papenfort, K. *et al.* Systematic deletion of Salmonella small RNA genes identifies CyaR, a conserved CRP-dependent riboregulator of OmpX synthesis. *Mol. Microbiol.* **68,** 890–906 (2008).

100. Massé, E. & Gottesman, S. A small RNA regulates the expression of genes involved in iron metabolism in Escherichia coli. *Proc. Natl. Acad. Sci.* **99,** 4620–4625 (2002).

101. Vanderpool, C. K. & Gottesman, S. The novel transcription factor SgrR coordinates the response to glucose-phosphate stress. *J. Bacteriol.* **189,** 2238–2248 (2007).

102. Rice, J. B. & Vanderpool, C. K. The small RNA SgrS controls sugar-phosphate accumulation by regulating multiple PTS genes. *Nucleic Acids Res.* **39,** 3806–3819 (2011).

103. Kawamoto, H., Koide, Y., Morita, T. & Aiba, H. Base-pairing requirement for RNA silencing by a bacterial small RNA and acceleration of duplex formation by Hfq. *Mol. Microbiol.* **61,** 1013–1022 (2006).

104. Maki, K., Morita, T., Otaka, H. & Aiba, H. A minimal base-pairing region of a bacterial small RNA SgrS required for translational repression of ptsG mRNA. *Mol. Microbiol.* **76,** 782–792 (2010).

105. Møller, T., Franch, T., Udesen, C., Gerdes, K. & Valentin-Hansen, P. Spot 42 RNA mediates discoordinate expression of the E. coli galactose operon. *Genes Dev.* **16,** 1696–1706 (2002).

106. Polayes, D. A., Rice, P. W. & Dahlberg, J. E. DNA polymerase I activity in Escherichia coli is influenced by spot 42 RNA. *J. Bacteriol.* **170,** 2083–2088 (1988).

107. Papenfort, K. *et al.* Specific and pleiotropic patterns of mRNA regulation by ArcZ, a conserved, Hfq-dependent small RNA. *Mol. Microbiol.* **74,** 139–158 (2009).

108. Mandin, P. & Gottesman, S. Integrating anaerobic/aerobic sensing and the general stress response through the ArcZ small RNA. *EMBO J.* **29,** 3094–3107 (2010).

109. Kalamorz, F., Reichenbach, B., März, W., Rak, B. & Görke, B. Feedback control of glucosamine-6-phosphate synthase GlmS expression depends on the small RNA GlmZ and involves the novel protein YhbJ in Escherichia coli. *Mol. Microbiol.* **65,** 1518–1533 (2007).

110. Urban, J. H., Papenfort, K., Thomsen, J., Schmitz, R. A. & Vogel, J. A conserved small RNA promotes discoordinate expression of the glmUS operon mRNA to activate GlmS synthesis. *J. Mol. Biol.* **373,** 521–528 (2007).

111. Reichenbach, B., Maes, A., Kalamorz, F., Hajnsdorf, E. & Görke, B. The small RNA GlmY acts upstream of the sRNA GlmZ in the activation of glmS expression and is subject to regulation by polyadenylation in Escherichia coli. *Nucleic Acids Res.* **36,** 2570–2580 (2008).

112. Urban, J. H. & Vogel, J. Two seemingly homologous noncoding RNAs act hierarchically to activate glmS mRNA translation. *PLoS Biol.* **6,** e64 (2008).

113. Rasmussen, A. A. *et al.* A conserved small RNA promotes silencing of the outer membrane protein YbfM. *Mol. Microbiol.* **72,** 566–577 (2009).

114. Mandin, P. & Gottesman, S. A genetic approach for finding small RNAs regulators of genes of interest identifies RybC as regulating the DpiA/DpiB two-component system. *Mol. Microbiol.* **72,** 551–565 (2009).

115. Vogel, J. *et al.* RNomics in Escherichia coli detects new sRNA species and indicates parallel transcriptional output in bacteria. *Nucleic Acids Res.* **31,** 6435–6443 (2003).

116. Jain, C. Identification and Characterization of Growth Suppressors of Escherichia coli Strains Lacking Phosphorolytic Ribonucleases. *J. Bacteriol.* **191,** 5622–5627 (2009).

117. Akey, D. L. & Berger, J. M. Structure of the nuclease domain of ribonuclease III from M. tuberculosis at 2.1 ? *Protein Sci. Publ. Protein Soc.* **14,** 2744–2750 (2005).

118. Kovacs, L., Csanadi, A., Megyeri, K., Kaberdin, V. R. & Miczak, A. Mycobacterial RNase E-Associated Proteins. *Microbiol. Immunol.* **49,** 1003–1007 (2005).

119. Rustad, T. R. *et al.* Global analysis of mRNA stability in Mycobacterium tuberculosis. *Nucleic Acids Res.* gks1019 (2012). doi:10.1093/nar/gks1019

120. Bandyra, K. J. *et al.* The Seed Region of a Small RNA Drives the Controlled Destruction of the Target mRNA by the Endoribonuclease RNase E. *Mol. Cell* **47,** 943–953 (2012).

121. Csanadi, A. *et al.* Towards new antituberculotic targets: biochemical characterisation of mycobacterial RNase E/G. *Eur. Respir. Rev.* **17,** 76–77 (2008).

122. Majdalani, N., Vanderpool, C. K. & Gottesman, S. Bacterial small RNA regulators. *Crit. Rev. Biochem. Mol. Biol.* **40,** 93–113 (2005).

123. Waters, L. S. & Storz, G. Regulatory RNAs in bacteria. *Cell* **136,** 615–628 (2009).

124. Brown, L. & Elliott, T. Mutations that increase expression of the rpoS gene and decrease its dependence on hfq function in Salmonella typhimurium. *J. Bacteriol.* **179,** 656–662 (1997).

125. Zhang, A., Wassarman, K. M., Ortega, J., Steven, A. C. & Storz, G. The Sm-like Hfq Protein Increases OxyS RNA Interaction with Target mRNAs. *Mol. Cell* **9,** 11–22 (2002).

126. McHugh, J. P. *et al.* Global iron-dependent gene regulation in Escherichia coli. A new mechanism for iron homeostasis. *J. Biol. Chem.* **278,** 29478–29486 (2003).

127. Massé, E., Escorcia, F. E. & Gottesman, S. Coupled degradation of a small regulatory RNA and its mRNA targets in Escherichia coli. *Genes Dev.* **17,** 2374–2383 (2003).

128. Calvori, C., Frontali, L., Leoni, L. & Tecce, G. Effect of rifamycin on protein synthesis. *Nature* **207,** 417–418 (1965).

129. Bohn, C., Rigoulay, C. & Bouloc, P. No detectable effect of RNA-binding protein Hfq absence in Staphylococcus aureus. *BMC Microbiol.* **7,** 10 (2007).

130. Novick, R. P. *et al.* Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. *EMBO J.* **12,** 3967–3975 (1993).

131. Bhakdi, S. & Tranum-Jensen, J. Alpha-toxin of Staphylococcus aureus. *Microbiol. Rev.* **55,** 733–751 (1991).

132. Graille, M. *et al.* Crystal structure of a Staphylococcus aureus protein A domain complexed with the Fab fragment of a human IgM antibody: Structural basis for recognition of B-cell receptors and superantigen activity. *Proc. Natl. Acad. Sci.* **97,** 5399–5404 (2000).

133. Huntzinger, E. *et al.* Staphylococcus aureus RNAIII and the endoribonuclease III coordinately regulate spa gene expression. *EMBO J.* **24,** 824–835 (2005).

134. Geisinger, E., Adhikari, R. P., Jin, R., Ross, H. F. & Novick, R. P. Inhibition of rot translation by RNAIII, a key feature of agr function. *Mol. Microbiol.* **61,** 1038–1048 (2006).

135. Majdalani, N., Vanderpool, C. K. & Gottesman, S. Bacterial small RNA regulators. *Crit. Rev. Biochem. Mol. Biol.* **40,** 93–113 (2005).

136. Ramos, C. G., da Costa, P. J. P., Döring, G. & Leitão, J. H. The Novel Cis-Encoded Small RNA h2cR Is a Negative Regulator of hfq2 in Burkholderia cenocepacia. *PLoS ONE* **7,** e47896 (2012).

137. DiChiara, J. M. *et al.* Multiple small RNAs identified in Mycobacterium bovis BCG are also expressed in Mycobacterium tuberculosis and Mycobacterium smegmatis. *Nucleic Acids Res.* **38,** 4067–4078 (2010).

138. Miotto, P. *et al.* Genome-Wide Discovery of Small RNAs in Mycobacterium tuberculosis. *PLoS ONE* **7,** e51950 (2012).

139. Pellin, D., Miotto, P., Ambrosi, A., Cirillo, D. M. & Di Serio, C. A Genome-Wide Identification Analysis of Small Regulatory RNAs in Mycobacterium tuberculosis by RNA-Seq and Conservation Analysis. *PLoS ONE* **7,** e32723 (2012).

140. Tsai, C.-H. *et al.* Identification of Novel sRNAs in Mycobacterial Species. *PLoS ONE* **8,** e79411 (2013).

141. Gandhi, N. R. *et al.* Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa. *Lancet* **368,** 1575–1580 (2006).

142. Gandhi, N. R. *et al.* Risk factors for mortality among MDR- and XDR-TB patients in a high HIV prevalence setting. *Int. J. Tuberc. Lung Dis. Off. J. Int. Union Tuberc. Lung Dis.* **16,** 90–97 (2012).

143. Zhang, H. *et al.* Genome sequencing of 161 Mycobacterium tuberculosis isolates from China identifies genes and intergenic regions associated with drug resistance. *Nat. Genet.* **45,** 1255–1260 (2013).

144. Yu, J. & Schneiders, T. Tigecycline challenge triggers sRNA production in Salmonella enterica serovar Typhimurium. *BMC Microbiol.* **12,** 1–13 (2012).

145. Howden, B. P. *et al.* Analysis of the small RNA transcriptional response in multidrug-resistant Staphylococcus aureus after antimicrobial exposure. *Antimicrob. Agents Chemother.* **57,** 3864–3874 (2013).

146. Eyraud, A., Tattevin, P., Chabelskaya, S. & Felden, B. A small RNA controls a protein regulator involved in antibiotic resistance in Staphylococcus aureus. *Nucleic Acids Res.* gku149 (2014). doi:10.1093/nar/gku149

147. Johnson, R. *et al.* Drug-resistant tuberculosis epidemic in the Western Cape driven by a virulent Beijing genotype strain. *Int. J. Tuberc. Lung Dis. Off. J. Int. Union Tuberc. Lung Dis.* **14,** 119–121 (2010).

148. Farhat, M. R. *et al.* Genomic analysis identifies targets of convergent positive selection in drug-resistant Mycobacterium tuberculosis. *Nat. Genet.* **45,** 1183–1189 (2013).

149. Philalay, J. S., Palermo, C. O., Hauge, K. A., Rustad, T. R. & Cangelosi, G. A. Genes Required for Intrinsic Multidrug Resistance in Mycobacterium avium. *Antimicrob. Agents Chemother.* **48,** 3412–3418 (2004).

150. Matsunaga, I., Maeda, S., Meda, S., Nakata, N. & Fujiwara, N. The polyketide synthase-associated multidrug tolerance in Mycobacterium intracellulare clinical isolates. *Chemotherapy* **58,** 341–348 (2012).

151. Takayama, K. & Kilburn, J. O. Inhibition of synthesis of arabinogalactan by ethambutol in Mycobacterium smegmatis. *Antimicrob. Agents Chemother.* **33,** 1493–1499 (1989).

152. Ando, H. *et al.* Downregulation of katG expression is associated with isoniazid resistance in Mycobacterium tuberculosis. *Mol. Microbiol.* **79,** 1615–1628 (2011).

153. Kim, Y. & Wood, T. K. Toxins Hha and CspD and small RNA regulator Hfq are involved in persister cell formation through MqsR in Escherichia coli. *Biochem. Biophys. Res. Commun.* **391,** 209–213 (2010).

154. Köten, B. *et al.* RNase 7 contributes to the cutaneous defense against Enterococcus faecium. *PloS One* **4,** e6424 (2009).

155. Singh, A. & Batra, J. K. Role of unique basic residues in cytotoxic, antibacterial and antiparasitic activities of human eosinophil cationic protein. *Biol. Chem.* **392,** 337–346 (2011).

156. Pulido, D. *et al.* Antimicrobial action and cell agglutination by the eosinophil cationic protein are modulated by the cell wall lipopolysaccharide structure. *Antimicrob. Agents Chemother.* **56,** 2378–2385 (2012).

157. Boix, E. *et al.* Structural determinants of the eosinophil cationic protein antimicrobial activity. *Biol. Chem.* **393,** 801–815 (2012).

158. Pulido, D., Torrent, M., Andreu, D., Nogues, M. V. & Boix, E. Two Human Host Defense Ribonucleases against Mycobacteria, the Eosinophil Cationic Protein (RNase 3) and RNase 7. *Antimicrob. Agents Chemother.* **57,** 3797–3805 (2013).

159. Mraheil, M. A. *et al.* Comparative genome-wide analysis of small RNAs of major Gram-positive pathogens: from identification to application: sRNAs of Gram-positive pathogens used as drug targets and for microbial diagnostics. *Microb. Biotechnol.* **3,** 658–676 (2010).

160. Kulyté, A., Nekhotiaeva, N., Awasthi, S. K. & Good, L. Inhibition of Mycobacterium smegmatis gene expression and growth using antisense peptide nucleic acids. *J. Mol. Microbiol. Biotechnol.* **9,** 101–109 (2005).

161. Park, H., Bak, G., Kim, S. C. & Lee, Y. Exploring sRNA-mediated gene silencing mechanisms using artificial small RNAs derived from a natural RNA scaffold in Escherichia coli. *Nucleic Acids Res.* **41,** 3787–3804 (2013).

162. Kumar, A., Chandolia, A., Chaudhry, U., Brahmachari, V. & Bose, M. Comparison of mammalian cell entry operons of mycobacteria: in silico analysis and expression profiling. *FEMS Immunol. Med. Microbiol.* **43,** 185–195 (2005).

163. WHO | Tuberculosis. WHO | Tuberculosis. *WHO* (2013). at <http://www.who.int/mediacentre/factsheets/fs104/en/>

164. Pierre-Audigier, C. *et al.* Fatal disseminated Mycobacterium smegmatis infection in a child with inherited interferon gamma receptor deficiency. *Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am.* **24,** 982–984 (1997).

165. Dagert, M. & Ehrlich, S. D. Prolonged incubation in calcium chloride improves the competence of Escherichia coli cells. *Gene* **6,** 23–28 (1979).

166. Ellis, R. C. & Zabrowarny, L. A. Safer staining method for acid fast bacilli. *J. Clin. Pathol.* **46,** 559–560 (1993).

167. Jacobs, W. R., Tuckman, M. & Bloom, B. R. Introduction of foreign DNA into mycobacteria using a shuttle phasmid. *Nature* **327,** 532–535 (1987).

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

169. Broussard, G. W. Electrocompetent Mycobacterium Cells. *Gregory W. Broussard, Ph.D.* (2009). at <http://gregorybroussard.com/2009/04/21/electrocompetent-mycobacterium-cells/>

170. Molecular cloning : a laboratory manual / Joseph Sambrook, David W. Russell. - Version details - Trove. at <http://trove.nla.gov.au/work/13615226?q&versionId=45739910>

171. Smabrook, J., Fritsch, E. . & Maniatis, T. Molecular cloning : a laboratory manual (2rd. ed). Plain View, N.Y.: Cold Spring Harbor Laboratory Press. (1989). at <http://trove.nla.gov.au/work/13615226?q&versionId=45739910>

172. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinforma. Oxf. Engl.* **26,** 589–595 (2010).

173. Ponstingl, H. SMALT - Wellcome Trust Sanger Institute. at <https://www.sanger.ac.uk/resources/software/smalt/#t_1>

174. Rozen, S. & Skaletsky, H. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol. Clifton NJ* **132,** 365–386 (2000).

175. Guo, X. V. *et al.* Silencing Essential Protein Secretion in Mycobacterium smegmatis by Using Tetracycline Repressors. *J. Bacteriol.* **189,** 4614–4623 (2007).

176. Mirouze, N., Parashar, V., Baker, M. D., Dubnau, D. A. & Neiditch, M. B. An Atypical Phr Peptide Regulates the Developmental Switch Protein RapH ▿. *J. Bacteriol.* **193,** 6197–6206 (2011).

177. QIAGEN, Germany. QIAGEN - miRNeasy Mini Kit. at <http://www.qiagen.com/products/catalog/sample-technologies/rna-sample-technologies/rna-cleanup/rneasy-minelute-cleanup-kit>

178. QIAGEN, Germany. QIAGEN - RNeasy MinuElute Kit. at <http://www.qiagen.com/products/catalog/sample-technologies/rna-sample-technologies/total-rna/rneasy-mini-kit#resources>

179. Koul, A. *et al.* Delayed bactericidal response of Mycobacterium tuberculosis to bedaquiline involves remodelling of bacterial metabolism. *Nat. Commun.* **5,** 3369 (2014).

180. QIAGEN, Germany. HotStarTaq DNA Polymerase. at <http://www.qiagen.com/products/catalog/assay-technologies/end-point-pcr-and-rt-pcr-reagents/hotstartaq-dna-polymerase/>

181. Sirgel, F. A. *et al.* gyrA mutations and phenotypic susceptibility levels to ofloxacin and moxifloxacin in clinical isolates of Mycobacterium tuberculosis. *J. Antimicrob. Chemother.* **67,** 1088–1093 (2012).

182. Philalay, J. S., Palermo, C. O., Hauge, K. A., Rustad, T. R. & Cangelosi, G. A. Genes required for intrinsic multidrug resistance in Mycobacterium avium. *Antimicrob. Agents Chemother*. **48,** 3412–3418 (2004).

183. Dawson, R. *et al.* Efficiency and safety of the combination of moxifloxacin, pretomanid (PA-824), and pyrazinamide during the first 8 weeks of antituberculosis treatment: a phase 2b, open-label, partly randomised trial in patients with drug-susceptible or drug-resistant pulmonary tuberculosis. *The Lancet* **385,** 1738–1747 (2015).

184. Raman, S. *et al.* The Alternative Sigma Factor SigH Regulates Major Components of Oxidative and Heat Stress Responses in Mycobacterium tuberculosis. *J. Bacteriol.* **183,** 6119–6125 (2001).

185. Suzuki, Y. *et al.* Detection of kanamycin-resistant Mycobacterium tuberculosis by identifying mutations in the 16S rRNA gene. *J. Clin. Microbiol.* **36,** 1220–1225 (1998).

186. Sirgel, F. A. *et al.* Mutations in the rrs A1401G gene and phenotypic resistance to amikacin and capreomycin in Mycobacterium tuberculosis. *Microb. Drug Resist. Larchmt. N* **18,** 193–197 (2012).

187. Marttila, H. J. *et al.* A Ser315Thr Substitution in KatG Is Predominant in Genetically Heterogeneous Multidrug-Resistant Mycobacterium tuberculosis Isolates Originating from the St. Petersburg Area in Russia. *Antimicrob. Agents Chemother.* **42,** 2443–2445 (1998).

188. van Soolingen, D. *et al.* Mutations at amino acid position 315 of the katG gene are associated with high-level resistance to isoniazid, other drug resistance, and successful transmission of Mycobacterium tuberculosis in the Netherlands. *J. Infect. Dis.* **182,** 1788–1790 (2000).

189. Tracevska, T., Jansone, I., Broka, L., Marga, O. & Baumanis, V. Mutations in the rpoB and katG Genes Leading to Drug Resistance in Mycobacterium tuberculosis in Latvia. *J. Clin. Microbiol.* **40,** 3789–3792 (2002).

190. Yao, C. *et al.* Detection of rpoB, katG and inhA gene mutations in Mycobacterium tuberculosis clinical isolates from Chongqing as determined by microarray. *Clin. Microbiol. Infect. Off. Publ. Eur. Soc. Clin. Microbiol. Infect. Dis.* **16,** 1639–1643 (2010).

191. Lew, J. M., Kapopoulou, A., Jones, L. M. & Cole, S. T. TubercuList--10 years after. *Tuberc. Edinb. Scotl.* **91,** 1–7 (2011).

192. Cangelosi, G. A. *et al.* The Two-Component Regulatory System mtrAB Is Required for Morphotypic Multidrug Resistance in Mycobacterium avium. *Antimicrob. Agents Chemother.* **50,** 461–468 (2006).

193. Manske, M. GENtle V1.9.4. (2004). at <http://activities.aliexpress.com/bestsellers.php?aff_platform=aaf&sk=zzVR7UnY3%3A&cpt=1429007454537>

194. Velayati, A. A. *et al.* Emergence of New Forms of Totally Drug-Resistant Tuberculosis Bacilli Super Extensively Drug-Resistant Tuberculosis or Totally Drug-Resistant Strains in Iran. *Chest* **136,** 420–425 (2009).

195. Udwadia, Z. F., Amale, R. A., Ajbani, K. K. & Rodrigues, C. Totally Drug-Resistant Tuberculosis in India. *Clin. Infect. Dis.* **54,** 579–581 (2011).

196. Kozińska, M. *et al.* [MDR, pre-XDR and XDR drug-resistant tuberculosis in Poland in 2000-2009]. *Pneumonol. Alergol. Pol. Organ Pol. Tow. Ftyzjopneumonologicznego Pol. Tow. Alergol. Inst. Gruźlicy Chorób Płuc* **79,** 278–287 (2011).

197. Mudur, G. Indian health ministry challenges report of totally drug resistant tuberculosis. *BMJ* **344,** e702 (2012).

198. Klopper, M. *et al.* Emergence and Spread of Extensively and Totally Drug-Resistant Tuberculosis, South Africa. *Emerg. Infect. Dis.* **19,** 449–455 (2013).

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

200. Cox, H. S. *et al.* The Beijing genotype and drug resistant tuberculosis in the Aral Sea region of Central Asia. *Respir. Res.* **6,** 134 (2005).

201. Ghebremichael, S. *et al.* Drug Resistant Mycobacterium tuberculosis of the Beijing Genotype Does Not Spread in Sweden. *PLoS ONE* **5,** e10893 (2010).