Tet-repressible Regulation System in Mycobacteria for Application to Vaccine Design

Student: Nureen Mayat
Student No: MYTNUR001

SUBMITTED TO THE UNIVERSITY OF CAPE TOWN
In fulfilment of the requirements for the degree:
M.Sc (Medical Microbiology)

Faculty of Health Sciences
UNIVERSITY OF CAPE TOWN

Date of Submission: 03 July 2009
Supervisor: Dr H Stutz, Department of Virology, University of Cape Town

Contents

Chapter 2: Materials and Methods

Chapter 3:Optimisation of mycobacterial culture conditions

Chapter 4: Construction and selection of a revTetR controlled expression system

Chapter 5:Testing the revTetR regulatable expression system in *M. smegmatis*

Chapter 6:General Discussion and Conclusions

Abstract

Mycobacterium bovis bacille calmette Guerin (BCG), the current vaccine for tuberculosis has been used for the expression of heterologous genes i.e. rBCG. Since BCG has a long-standing safety profile and over 3 billion doses have bee administered worldwide it has been used as a vaccine vehicle to deliver immune-inducing antigens for various diseases. However, of major concern is the fact that many foreign antigens are only produced at very low levels and the recombinants are generally genetically unstable, resulting in weak immune responses. To overcome this challenge we proposed the use of a tetracycline-repressible expression system such that foreign gene expression is down regulated during the culturing process and up regulated once the vaccine is administered. This system has been used previously in mycobacteria for the generation of conditional knockouts but not in a vaccine context.

To test for repression and induction, a single, multicopy, episomal shuttle vector design containing the reporter gene GFP, driven by various strong P_{smyc} promoter-operator designs containing *tet* operator sites (pNM3.1, pNM3.2, pNM4.1 and pNM4.2), and the regulatory protein gene TetR r1.7 (revTetR) driven by the *hsp*60 promoter was used to control the expression of GFP in the model organism *Mycobacterium smegmatis*. The most promising shuttle-vector design proved to be pNM3.1 which contained two *tet*O2 sites situated on either side of the –35 promoter region. With this construct we successfully repressed GFP expression to ~ 40% with 100ng/ml of anhydrous tetracycline (ATc), an analogue of tetracycline and the co-repressor in this system. This was reversible, in that GFP expression was induced to ~68% in the same sample, in the absence of ATc. Repression and induction was observed to occur in a dose-dependent manner.

Acknowledgements

I would firstly like to thank my parents, whom have provided support and encouragement throughout my studies. Secondly my friends in Cape Town, my second family, particularly the sisters I never had; Laila and Nazia whom have helped me maintain my sanity throughout this project.

This project would not have been possible without Dr. Helen Stutz who went far beyond the call of duty in assisting, encouraging and guiding me. Additionally I would like to thank Prof. Anna-Lise Williamson for this opportunity, for her understanding and her input. I would also like to thank the other members and students of Prof. Anna-Lise's group, particularly Nicky Balfour who diverted many a catastrophe.

This study was in funded in part by the South African Aids Vaccine Initiative and the (SAAVI) and the National Research Foundation (NRF).

List of Abbreviations

2YT	2 x yeast tryptone
A	Adenine
ATc	Anhydrous tetracycline
Amp^R	Ampicillin resistance
BCG	Bacille Calmette-Guérin
bp	Base pair
BSA	Bovine serum albumin
C	Cytosine
CTL	Cytotoxic lymphocyte
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
ELISA	Enzyme-linked immunosorbent assay
G	Guanine
GFP	Green fluorescent protein
HIV	Human Immunodeficiency virus
HPLC	High purity liquid chromatography
HRP	Horseradish peroxidase
hsp	Heat shock protein
Kan^R	Kanamycin resistance
kb	Kilobase
kDa	KiloDalton
kV	Kilovolt
λ	Lambda
ln	Natural logarith
M	Molar
MIC	Minimum inhibitory concentration
M7H9	Middlebrook 7H9
M7H10	Middlebrook 7H10
mg	Milligram
ml	Millilitre
ng	Nanogram
nm	Nanometre
OADC	Oleic-acid albumin dextrose complex
ºC	Degree Celsiu
OD	Optical density
ORF	Open reading frame
ori	Origin of replication
oriE	*E. coli* origin of replication
oriM	Mycobacterial origin of replication
PBS	Phosphate buffered saline
PBS-T	Phosphate buffered saline with Tween
PCR	Polymerase chain reaction
Pr	Promoter
r	Recombinant
rpm	Revolutions per minute
RBS	Ribosome binding site
RFU	Relative fluorescence unit
RNA	Ribonucleic acid
RT	Room temperature
SIV	Simian Immunodeficiency virus
T	Thymidine
TB	Tuberculosis
Tc	Tetracycline
*tet*O1	Tetracycline operator sequence 1
tetO2	Tetracycline operator sequence 2
tt	Transcriptional terminator
µl	Microlitre
µM	Micromolar
UV	Ultraviolet

List of Figures

List of Tables

Chapter 1:Introduction

Chapter 1:Introduction

1.1. *Mycobacterium bovis* bacille Calmette-Guerin (BCG): A brief history

Tuberculosis (TB) is one of the world's major health problems and is responsible for approximately 1.8 million deaths annually[1]. *Mycobacterium tuberculosis,* the causative agent of TB was identified by Robert Koch in 1882 and was the first live vaccine developed against TB was BCG, which remains the only current TB vaccine to day. It was developed from a strain of *M. bovis* that was attenuated by serial passaging on solid media over a period of 13 years. Albert Calmette and Camille Guerin of the Pasteur Institute were credited with the development of the vaccine hence the name bacille Calmette-Guerin (BCG), with the first dose administered to a neonate in 1921. To date more than 3 billion doses have been administered worldwide[2], yet the vaccine is not effective in conferring protection against all the various presentations of the disease. BCG is protective against severe forms of childhood TB but does not reliably protect against the more prevalent form i.e. adult pulmonary TB[3]. This has lead to the extensive continued research and development of BCG with respect to improving vaccine efficacy. Additionally, due to the long-standing safety profile and extensive advances in mycobacterial genetics, BCG has been developed and employed for use as a vaccine vector for various other diseases in order to present or deliver foreign pathogen antigens to the host immune system.

1.2. BCG as a live recombinant vaccine vehicle

There are numerous advantages to developing BCG as a delivery vehicle for the expression of foreign antigens. BCG has a long-standing safety profile with more than 3 billion doses being administered to children with a low incidence of complications. It can be administered to neonates and may be administered via various routes including orally, percutaneously and intradermally[4].

A major advantage of using a live bacterial vector is that they continue to replicate following immunisation leading to continuous synthesis and presentation of antigens, thereby acting as an adjuvant by persistently stimulating the immune system[5]. Additionally, BCG is able to elicit both humoral and cell mediated immune responses, especially CD4 responses required for CD8 memory[6].

Furthermore BCG can accommodate the expression of fairly large regions of DNA, it is the most heat stable of all live vaccines, may be transported in the absence of a cold chain, is relatively inexpensive ($0.05 per dose) and easy to manufacture[7]. This is a vital consideration for developing countries with limited resources such as South Africa, which already possess the expertise for BCG production. Therefore BCG is recognised as a promising vaccine vector for the expression of a range of viral, bacterial and parasitic antigens.

1.2.1. Recombinant bacille Calmette-Guerin (rBCG) vaccines

As mentioned above, many heterologous proteins have been expressed in BCG for purposes of vaccine design. Stover *et al*., 1993[2] developed a rBCG strain that expressed the outer surface protein A (OspA) antigen of *Borrelia burgdorferi* (Lyme disease) as a membrane associated lipoprotein. BCG on it's own, as a result of it being an intracellular pathogen, induces a predominantly Th1 response that relies on cytotoxic T-cells to confer protection[8], but by expressing the antigen on the surface of rBCG , an-OspA specific protective humoral response was achieved against Lyme borreliosis in a mouse model. This group also achieved a protective antibody response against a *Clostridium tetani* challenge in mice with an rBCG strain expressing the non-toxinogenic Tetanus toxin fragment C (ToxC)[9]. In addition, rBCG expressing the pneumococcal surface protein A has elicited protective immune responses against a pneumococcal infection in a *Streptococcus pneumoniae* mouse challenge model[10].

rBCG vaccines containing parasitic antigens of; *Plasmodium falciparum*, *Leishmania major*, *Schistosoma mansoni* and *Toxoplasma gondii* have also been generated by

various teams that have been able to elicit appropriate protective immune responses against experimental infection in animal models[11-15].

rBCG vaccines expressing viral antigens including those from the viruses that cause Foot-and-mouth disease, measles, cervical cancer [human papilloma virus (HPV)], HIV/AIDS [human immunodeficiency virus (HIV-1)] and simian immunodeficiency virus (SIV) have been generated[16-18]. The majority of research involved in the expression of viral proteins has focused on HIV and SIV as a result of the significant impact of the HIV/AIDS pandemic. The expression of HIV-1 genes is of particular interest to us as it is one of the anti-HIV vaccine strategies being explored by our group with the goal of inducing protective anti-HIV CTL responses that are thought to play a role in the containment of HIV replication[19-21].

Since the initial report of the development of the *E. coli*-mycobaterial shuttle vectors in 1987[22] and the first report of the expression of a viral gene in rBCG by Matsuo *et al*., 1990[23] using the α antigen promoter of *M. kansaii* , many HIV and SIV antigens have been expressed in BCG and tested in various animal models including mice, guinea pigs and non-human primates.

Early studies such as the report by Kameoka *et al*., 1994[24] wherein they expressed the V3 region of the HIV-1 envelope protein fused to the *M. kansaii* α antigen (k- α) in BCG showed promising results. These included the induction V3 peptide specific cytotoxic lymphocytes (CTLs) upon inoculation of BALB/c mice[24]. Stover *et al*., 1991[25] expressed several HIV-1 antigens (Gag, RT, Polgp120 and gp41) in rBCG under the control of the heat shock promoters *hsp*60 or *hsp*70 on either an extra chromosomal or integrative vector. Aldovini and Young published an article in the same year wherein rBCG containing a multicopy plasmid expressing the HIV-1 *gag*, *pol* and *env* genes driven by the *hsp*70 promoter[26]. In both studies, both humoral and cell mediated HIV-1 specific responses were induced in mice immunised with these vaccines. Additionally, Winter and colleagues were able to express the HIV-1 *nef* gene under the control of the *gro*ES/EL1 operon promoter from *S. albus*. This vaccine was used to inoculate mice and *nef* specific immune responses were detected[27].

More recently, in cynomologous macaques, vaccinated with rBCG expressing the full-length gag SIV and challenge with SHIV (SIV-HIV synthetic virus) resulted in the reduction of viral load and plasma viremia[28]. Additionally, a recent study presented rBCG as a promising dual TB-HIV-1 vaccine platform in an attempt to prevent the vertical transmission of HIV-1 i.e. mother to child transmission (MTCT) of the virus via breastfeeding[29]. This is significant for TB endemic regions as South Africa where dual TB and HIV-1 infection is common.

1.2.2. Challenges for heterologous gene expression in rBCG with respect to vaccine design

These studies represent encouraging data to support the development of rBCG vaccines. However, in spite of the numerous advantages of BCG, there are several challenges. These include the fact that it is a slow growing organism with a doubling time of roughly twenty-four hours, it tends to aggregate in culture and is difficult to manipulate genetically[5, 30, 31]. Moreover, there exists BCG strain variation revealed by comparative genomics between early strains such as Japan, Sweden, Birkhaug and Russia, and late strains such as Pasteur, Danish, Glaxo and Prague[32, 33]. These variations have been attributed to gene modifications caused by selective pressures during passaging. These findings suggest that different strains result in variable immune responses and as such should be taken into consideration when selecting the vector strain.

Additionally, immune responses to the foreign antigen are usually masked by the vast milieu of bacterial antigens present. Thus, in order to obtain the desired response to the foreign antigen, it would have to be expressed at high levels to compete with the numerous immunogenic host antigens. However, the major challenge facing the successful development of rBCG vaccines lies in this fact that many foreign proteins can only be expressed at very low levels, and the recombinants are genetically unstable, resulting in weak immune responses to the foreign antigen[5, 30]. Thus, both these problems of low expression levels and genetic instability have to be addressed

in order to achieve a potent vaccine which can present an adequate amount of stable antigen to the host immune system to elicit the appropriate protective response.

1.2.2.1. Low expression levels and genetic instability associated with heterologous gene expression in rBCG

The literature reports that levels of heterologous proteins expressed in rBCG varies between 0.1% to 15% of total cellular protein[34]. The lower limit of this range is of concern as many foreign genes, typically viral genes can only be expressed at very low or even undetectable levels. Our group and others have found that in particular expressing viral genes in rBCG leads to genetic instability in which the gene encoding the immunogen may be modified or lost entirely which in turn affects the potency of the vaccine in terms of the magnitude of the immune response induced[35, 36].

The metabolic stress imposed on the host bacterium by the expression of a foreign gene has been defined by Glick, 1995[37] as 'the portion of the cells resources, either in the form of energy such as ATP or GTP, or raw materials such as amino acids that is required to maintain and express foreign DNA, as either RNA or protein in the cell'. It has been suggested that the metabolic burden is gene-specific and that the size of the plasmid can be an important variable[38, 39]. This increased metabolic burden can be severe especially when the foreign gene is under the control of a strong promoter which uses a vast amount of resources from the bacterium and can result in 'over expression toxicity' which may be lethal to the host bacterium[34].

Researchers commonly addressed the problems of low antigen expression by expressing the heterologous antigen from a multicopy episomal plasmid and/or by using stronger promoters. In order to overcome instability, researchers commonly integrated their gene into the chromosome. This is exemplified by Stover *et al*.,[25] who were unable to express HIV-1 gp120 from an episomal vector (n=5 copies per cell) but were able to from an integrative vector (n=1 copy per cell). However, neither the use of multicopy episomal or integrative constructs in combination with strong expression signal have been able to successfully solve both the problems of low viral antigen expression and instability. We now understand from extensive

research carried out in our group (Dr R. Chapman, publication in progress) that low expression and instability are not two separate issues but are intricately linked.

As mentioned above, the expression of viral genes can lead to a significant metabolic burden on the host bacterium due to the recruitment of a vast portion of the host cells metabolic machinery, which can be toxic to the host. The level of toxicity can depend on the specific gene itself, for example during the translation process wherein the presence of rare codons in the foreign gene potentially results in the sequestering of rare tRNAs away from mRNAs that encode essential host proteins[40]. Brinkmann *et al*., 1989[41] corroborated this hypothesis in their report that the plamsid-based expression of eukaryotic genes containing a large proportion of rare codons leads to plasmid instability, reduced protein synthesis and induction of a physiological stress response that upregulates cellular proteases that can degrade these foreign antigens[42, 43]. Langermann *et al*.,[44] experienced difficulty in expressing full-length PspA protein with its natural carboxy-terminal anchor domain from an episomal vector under the control of an *hsp60* promoter. Successful expression was only attained once this repetitive domain was removed suggesting its toxicity.

Inherent toxicity of a protein may be attributed to a post-translational event, for example, the expression of viral genes leads to the production of mis-folded proteins which results in the induction of expression of cellular proteases in order to degrade them[34, 45, 46]. Thus, the stress responses result in the degradation of viral antigens, hence the low level observed, but also leads to an added metabolic burden.

Metabolic burden subsequently affects the growth rate such that the higher the metabolic burden, the slower the growth rate of rBCG. Mutations occur randomly and continuously within a culture, however as soon as a mutation arises that gives that specific recombinant a growth rate advantage, it will out-compete the rest of the culture and become the dominant species. These dominant more 'fit' mutants contain either a heterologous gene that is modified or deleted completely which is termed genetic instability[34].

In bacterial culture, antibiotic selection is present, preventing the growth of plasmid-less bacteria. In the absence of antibiotic, as the case would be when the vaccine is

administered if the metabolic load is great this would increase the selection pressure for plasmid-less bacteria. Thus, genetic instability poses a major challenge when growing vaccine stocks if one considers that vaccine stocks are grown from a single colony off a plate, which contains approximately 10^6 cells, through starter cultures, taken through several sub-culturing processes, before finally producing a vaccine stock.

To obtain the high concentrations at which vaccines are usually administered i.e. about 10^7-10^{11} cells in a 100µl volume, cultures have to go through numerous passages and generations (more than 30 generations) providing more than ample time for a more "fit", less unstable mutant to take over a culture[47]. This highlights the importance of developing a genetically stable vaccine.

1.2.2.2. Overcoming challenges of low expression levels and genetic instability

In order to achieve high stable expression of antigens in rBCG it is necessary to minimise the imposed metabolic burden. A number of approaches can and have been used to find the balance between the expression of sufficient antigen whilst maintaining genetic stability.

Higher expression levels of recombinant antigen can be attained in some instances using strong promoters, optimised ribosome binding sites and high copy number vectors[48]. *E. coli*-mycobacterial shuttle vectors have been developed by the combination of the replication region (oriM) of the pAL5000 *M. fortuitum* replicon, allowing for plasmid replication in both *M. smegmatis* and BCG and an oriE allowing for replication in *E. coli* [49]. The use of this plasmid results in ~ 5 copies per cell versus the use of an integrative system which results in only 1 copy per cell. Making use of a plasmid with an even higher copy number, such as the mutant generated by our group that is capable of persisting at ~50 copies per cell may serve to further increase expression levels[48].

Although several different promoters have been used for foreign gene expression in mycobacteria, the most commonly used are the heat shock protein promoters, *hsp*60

and *hsp*70 that drive the expression of heat shock proteins in mycobacteria and are constitutively expressed when the cell experiences stress in its environment[50]. These promoters result in strong expression of recombinant antigens, however, the use of the *hsp*60 promoter was found to be disadvantageous in that it has been found to directly contribute towards genetic instability in a number of studies[51, 52]. More recently, the P_{smyc} promoter, isolated from *Mycobacterium smegmatis* by Ehrt *et al*., 2005[53] was found to be an even stronger promoter than *hsp*60 capable of potent expression that was equally strong in BCG Pasteur.

According to Dennehy and Williamson, 2005[34] 'The rate of translation of foreign genes may be influenced by their codon usage'. Mycobacterial genes have a high G+C content and as a result the translation of foreign genes that have a high A+T content may have compromised translation efficiency. In order to improve the efficiency of translation and hence improve expression of the foreign gene some groups have codon adapted their gene of interest to suit mycobacteria[54-56]. On a post-translational level, the mutation and the selection of a better-folded protein would result in a less toxic product and aid in attaining higher expression levels.

The addition of secretion signals to export the foreign antigen or tagging the foreign antigen for membrane display have also been used as means of minimising cellular toxicity and enhancing antigenicity. The *M. tuberculosis* 19kDa lipoprotein has proved to be the most widely used signal sequence for directing foreign antigen to the membrane[2, 10].

In vivo inducible promoters represent one way in which metabolic burden can be reduced during the culturing of vaccine stocks. The 18kDa promoter from *M. leprae* for example, exhibited weak activity during *in vitro* growth in *M. smegmatis* and BCG that was induced *in vivo* upon infection of macrophages[57, 58]. Additionally, the *M. tuberculosis mtr*A promoter is downregulated during *in vitro* growth and upregulated during *in vivo* growth, which would once again reduce the metabolic burden during culturing and potentially minimise genetic instability.

Another way in which to circumvent the effects of instability during vaccine culturing would be to temporally control the expression of the gene of interest i.e. to repress the expression of the gene during the culturing process and only induce its expression when required i.e. upon administration of the vaccine. This may be approached by using a regulated expression system, which may be controlled by the addition of an exogenous regulatory molecule.

1.3. Regulated expression systems

The regulation of gene expression evolved as a means to conserve energy. Bacteria utilise these systems to adapt to changes in their environment such as temperature fluctuations, water and nutrient availability and in response to toxic molecules originating from the environment or produced by their own metabolism. Regulatory proteins that modulate transcription, translation, or some other event in gene expression trigger rapid adaptive responses[59]. Regulated gene expression systems usually function at the level of transcription, allowing for transcription in the ON state and preventing it in the OFF state. The OFF state usually refers to basal levels of transcription, as there are no known systems where gene expression is completely shut down[60]. These natural systems have been manipulated for several purposes, including the identification of gene function either in its natural cellular environment or in specialised host organisms. In addition, these systems can be utilised to regulate the production of heterologous protein for the application to vaccine design.

The ideal regulated expression system that would control the level and the timing of the expression of the gene of interest should contain several key properties including high specificity, efficiency, dose dependency, reversibility and should allow for *in vivo* expression if it is to be applied to a vaccine setting. The most widely used and studied regulated expression system consists of the genes responsible for lactose utilisation in *E. coli.* Many of the regulatory systems utilised today are derived from *E. coli* but due to the species-specific nature of promoters these systems cannot be effectively utilised in mycobacteria i.e. BCG or the model organism *M. smegmatis*.[61]

1.3.1. Regulated expression systems in mycobacteria

The identification and elucidation of complex systems such as inducible expression systems in mycobacteria is scant due to the limited availability of mycobacterial research genetic tools due to the difficulties encountered in genetic manipulation and culturing of these bacteria in the laboratory. Historically, recombinant *E. coli* expression systems have been used to identify and study genes from less well-known organisms. Mycobacteria however, pose a challenge in that mycobacterial promoters lack recognition by *E. coli* transcription machinery thereby precluding the use of *E. coli* to express mycobacterial antigens. Only a few regulated expression systems have been identified and described for use in mycobacteria including the modification of inducible expression systems from other origins for application to mycobacteria. These are described below.

1.3.1.1. The Tra system

The Tra system consists of genes involved in plasmid transfer in *Streptomyces nigrifaciens*. It consists of a repressor molecule $TraR^{ts107}$ (temperature-sensitive derivative of TraR) that represses initiation of transcription at the P_{tra} promoter at temperatures of ≤ 28°C and induces at temperatures above this. This is not ideal for use in BCG or *M. smegmatis* since they grow optimally at 37°C. Furthermore, the promoter displays only intermediate expression levels in mycobacteria that, in the context of vaccine efficacy, would result in sub-optimal levels of antigen expression and hence immune responses[62]. Lim *et al*., 2000[63] showed a 15-fold reduction in promoter activity in comparison to the *M. bovis* BCG derived *hsp*60 promoter[63]. These features make the use of this genetic switch an unattractive choice for the development of an rBCG vaccine.

1.3.1.2. The Acetamidase inducible expression system

The acetamidase system was the first system used to regulate gene expression in mycobacteria[61]. Draper first alluded to the highly inducible nature of the acetamidase enzyme in *M. smegmatis,* strain NTC 8159, in 1967[64]. He identified the inducers as acetamide or butyramide, which are simple aliphatic amides. The induction of the

acetamidase 47-kDa enzyme enables the bacterium to utilise these substrates as the sole carbon source. It was evident that acetamidase expression was regulated by two systems, repression under uninduced conditions and induction in the presence of acetamide[65].

The acetamidase gene was cloned, sequenced and the protein purified by Mahenthiralingam *et al.,* in 1993[66]. In 1997, Parish *et al.*[67] identified the negative regulator of the system, AmiA and proposed a model for transcriptional induction and repression which was corroborated by Subbian and Narayanan in 2007 in addition to them defining several operator regions[68]. The acetamidase operon consists of the acetamidase gene (*amiE*) and the upstream open reading frames, *amiA*, *amiCD* and *amiS*. The gene, *amiA* encodes the negative regulator and genes *amiC* and *amiD*, encode positive regulators. *AmiS* on the other hand, encodes a component of a putative transporter. The arrangement of these genes and their promoters on the *M. smegmatis* chromosome are shown in Figure 1.1.

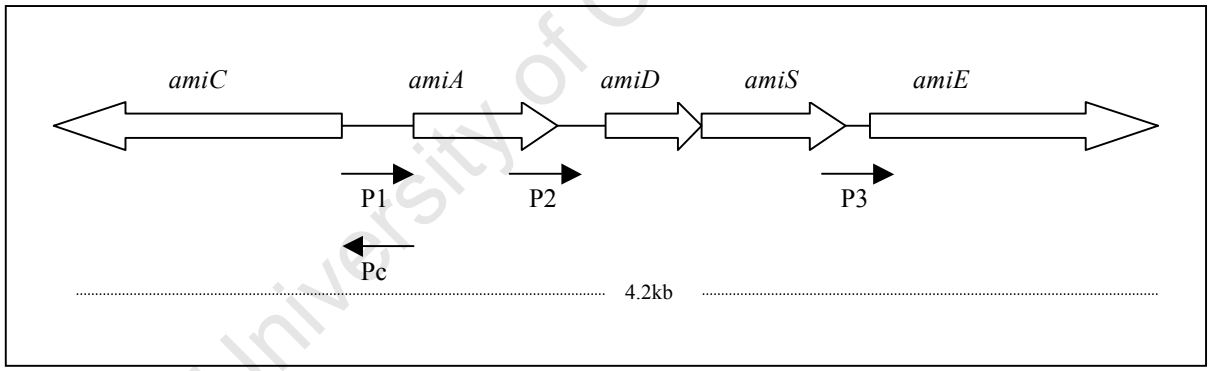


Figure 1.1. Arrangement of the acetamidase operon adapted from Parish *et al*., 2001[67]. The acetamidase operon consists of the acetamidase gene (*amiE*) and the upstream open reading frames (*amiA*, *amiCD* and *amiS*). *amiA* is the negative regulator and *amiC* and *amiD*, the positive regulators. *amiS* encodes a component of a putative transporter. In the absence of acetamide, AmiA binds to promoter P2 and prevents transcription of the three downstream genes (*amiD*, *ami*S and *ami*E). AmiA may also bind to promoter Pc and reduce AmiC expression. In the presence of acetamide, AmiA is no longer bound to the promoter P2 and transcription of the downstream genes occurs. It is postulated that the AmiC protein binds to both acetamide and AmiA thus preventing the repressor (AmiA) from interacting with the promoters.

In the absence of acetamide, AmiA, the negative regulator, binds to promoter P2 and prevents transcription of the three downstream genes. A basal level of transcription does occur from P2 leakage or via the P3 promoter. AmiA may also bind to promoter

Pc and reduce AmiC expression. In the presence of acetamide, AmiA is no longer bound to the promoter P2 and transcription of *amiD, amiS and amiE* occurs at a high level. It is likely that AmiC protein has an acetamide binding domain, binding to both acetamide and AmiA, thus preventing the repressors (AmiA) interaction with the promoters[67, 68]. Acetamidase is very highly expressed in its induced state by 100-fold and can account for up to 10% of the total protein production while only low levels of transcription occur during the repressed state.

Despite its highly inducible nature and its previous successful use in the over-expression of heterologous antigens in *M. smegmatis* from either *Mycobacterium leprae* or *M. tuberculosis,* this system is complex and not fully understood[69, 70]. The acetamidase operon is a large fragment (>4.2kb) making this system difficult to manipulate and since it would require the addition of acetamide for the induction of antigen expression, it is an unsuitable system for the application to vaccine design.

1.3.1.3. The Tetracycline inducible expression system

Tetracycline resistant determinants are found in abundance amongst Gram-negative bacteria of which the most commonly found resistance mechanisms are active export of the antibiotic via efflux pumps from the cell and ribosome protection of which the most extensively studied is the Tn*10* determinant, Tet(B) of *E. coli* [71, 72]. The Gram-negative promoters that drive the expression of these tetracycline efflux pumps are among the most tightly regulated bacterial promoters. Gram-positive bacteria have also been found to contain *tet* determinants, more recently, the TetR(Z) system of *Corynebacterium glutamicum* has been identified[73]. The Tn*10* determinant, Tet(B) of *E. coli* which is highly conserved and encodes a tightly regulated efflux system for the export of the antibiotic thus conferring resistance, is described in detail here to outline the principle of this regulated system. The system consists of two genes *tetA* and *tetR,* which are divergently oriented as shown in Figure 1.2. TetA is the resistance protein; a tetracycline metal-proton antiporter situated in the cytoplasmic membrane and is responsible for transporting tetracycline out of the cell. The intergenic region between *tetA* and *tetR* contains three overlapping promoters $P_{A,}$ $P_{R1,}$ P_{R2} and two distinguishable 19bp operator sequences (*tet*O1 and *tet*O2) separated by 11bp of sequence[71].

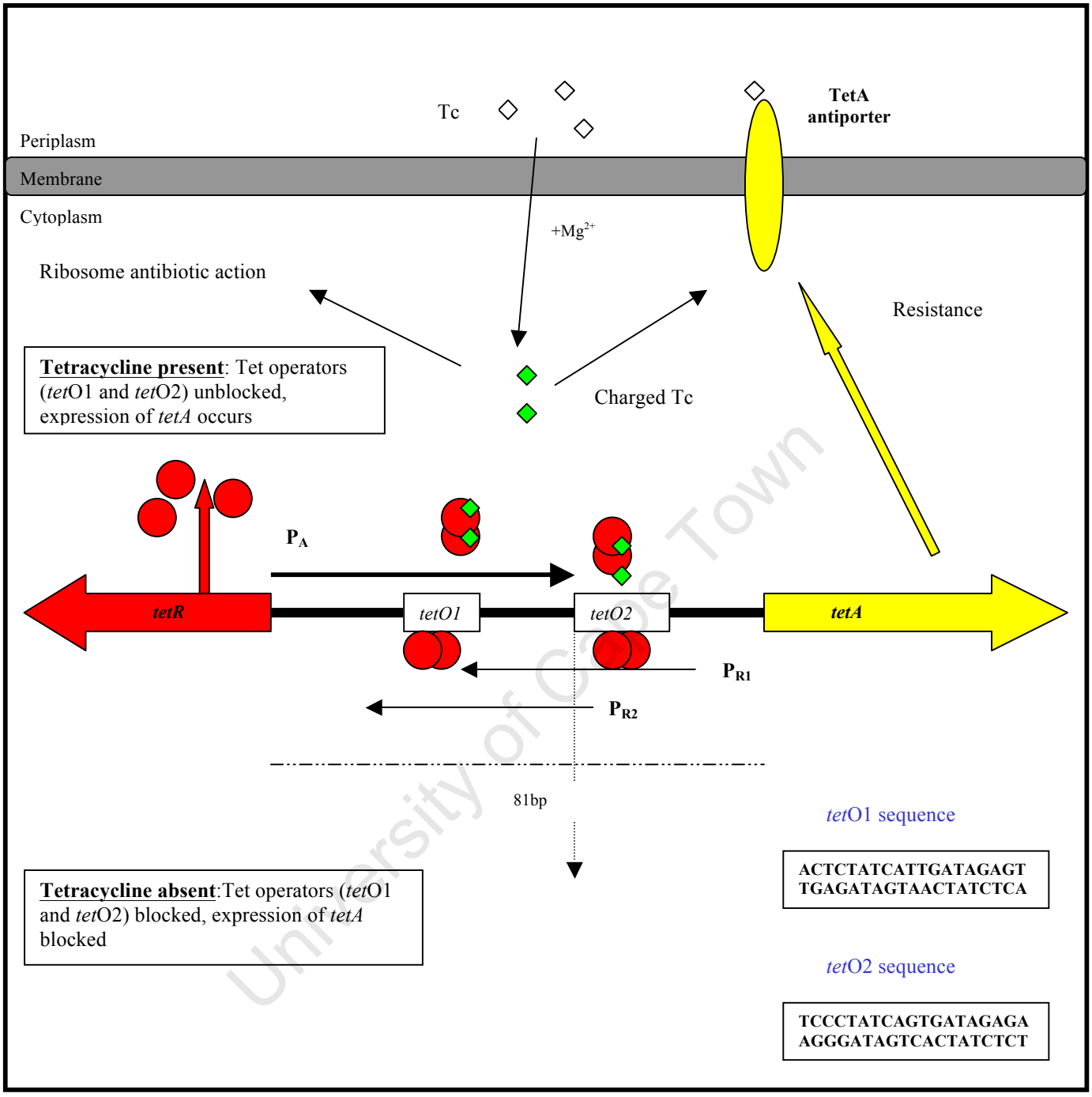


Figure 1.2. The genetic organisation and mechanism of the *E. coli* Tn*10* tetracycline resistance determinant. (Adapted from Hillen and Berens, 1994[71] and Rossi and Blau, 1998[74])
The system consists of two genes *tetA* and *tetR,* which are divergently oriented. TetA is the resistance protein; a tetracycline metal-proton antiporter (yellow oval) situated in the cytoplasmic membrane and is responsible for transporting tetracycline (white diamonds represent uncharged tetracycline molecules and green diamonds represent charged tetracycline molecules) out of the cell. The intergenic region between *tetA* and *tetR* contains overlapping promoters P_A, P_{R1}, P_{R2} and two distinguishable operator sequences (*tet*O1 and *tet*O2) separated by 11bp. In the absence of tetracycline, the *tet* operators are blocked and expression of TetA is blocked. In the presence of tetracycline, charged tetracycline molecules bind TetR (red circles) homodimers resulting in the dissociation from operator sites. Expression of TetA follows, resulting in the export of the antibiotic.

The repressor protein TetR consists of two distinct domains, a DNA- binding domain and a tetracycline-binding domain. In the absence of tetracycline (Tc), TetR represses the expression of both *tetA* and *tetR* by binding to both operator sites, with different affinities. The affinity of TetR for *tet*O2 is estimated to be 3-5 times higher than it is for *tet*O1[71]. Thus, occupation of *tet*O2 ensures repression of the expression of TetA, which is highly toxic to cells at high concentrations, while allowing the expression of TetR from the secondary promoter, $P_{R2,}$ when *tet*O1 becomes free, hence providing a tightly regulated security circuit and permitting low level read-through of TetR.

When Tc is present, it freely diffuses across biological membranes (i.e. without the aid of protein channels), chelates Mg^{2+} and binds to the repressor TetR with high affinity. This results in a conformational change within the DNA binding region of the protein causing it to dissociate from the DNA operator sites thus, inducing the expression of TetA. This system ensures that TetA is only expressed in the fully induced state when absolutely required, thus minimising toxicity[53, 59, 71, 75].

The precise sequence and position of these operators and the linker region between them is key for efficient and tight regulation of the system, and it has been suggested that the complex arrangement of these promoters and operators is conserved for tight regulation of TetA expression. The efficiency of this system is a result of the high affinity that Tc harbours for the TetR-*tet*O complex versus its affinity for the small (30S) ribosomal subunit i.e.$k_{ass} \approx 10^6\ M^{-1}$ versus $k_{ass} \sim 10^9\ M^{-1}$, respectively, resulting in the rapid induction of expression of TetA before Tc can exhibit its antibiotic activity. TetR shows high specificity for the *tet* operator sites above other sites on the chromosome ($k_{ass} \approx 10^{11}\ M^{-1}$) and the Tc-TetR complex has a reduced affinity for binding DNA thus, contributing to the specificity of this system[72].

Thus, the Tc regulated expression system is one of the simplest, best understood, most sensitive and extensively used systems for conditional gene expression. These regulatory elements Tc, TetR and operator sequences in conjunction with species-specific promoters have been used by researchers in various ways to control selected gene expression in a variety of different organisms including bacteria, yeast and

eukaryotes [71, 76]. An overview of the interchangeable components is presented below in Figure 1.3.

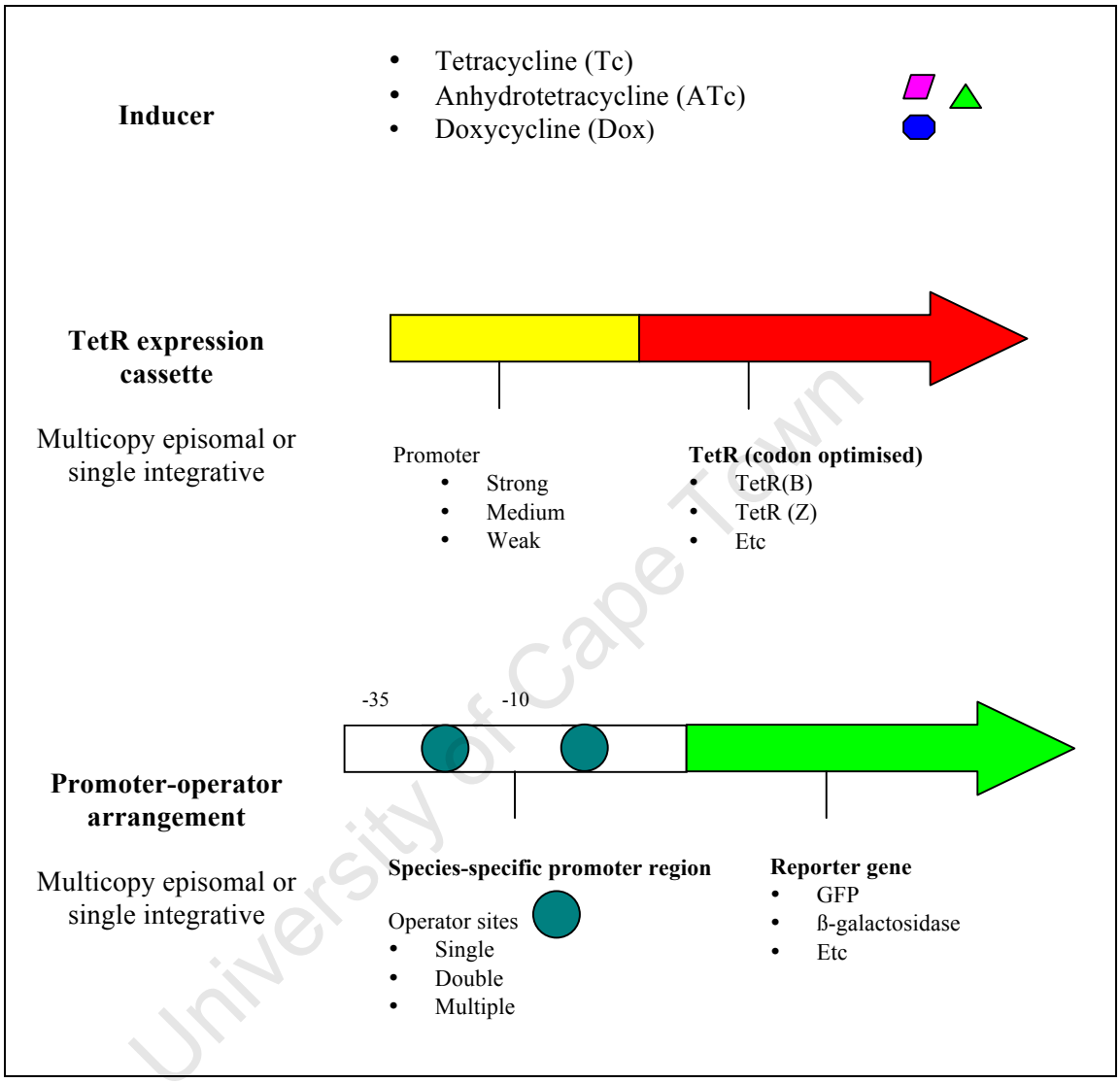


Figure 1.3. Generic schematic overview of the various combinations of components that have been used in Tc-regulated expression systems. Various analogues of Tc have been used. Additionally, selected promoters may drive TetR expression, either integratively or episomally. Promoter-operator arrangements with different combinations of operator sites and reporter genes represent additional components that may be selected.

In these systems a reporter gene or another gene of interest replaces the *tetA* gene and a variable number of operator sites are introduced into the appropriate promoter regions (See Fig. 1.3). Making use of various tetracycline analogoues has allowed for further customisation of this system. Anhydrous tetracycline (ATc) is an antibiotic derivative of tetracycline, which acts as a superior inducer since it binds to the Tet repressor more efficiently than tetracycline or doxycycline (analogue), is less toxic to cells and maintains high functional stability in cell culture. ATc is used at sub-inhibitory levels and is not metabolised by bacteria or higher organisms thus making it an ideal inducer for use in *in vivo* models[77, 78].

Also the selection of the appropriate type (*tet*O1 versus *tet*O2) and number (single, double or multiple) of operator sites may serve to further optimise the system, depending on what type of regulation is required. Kamionka *et al*., 2004[79] found that in *Bacillus subtilis* the insertion of two operator sites (*tet*O1) into a *spo*VG promoter functioned more efficiently in terms of repressing the *lacZ* reporter gene expression than one *tet* operator 1 site. Precisely 5.3-fold better repression was achieved with the additional operator site. This highlighted that the addition of an operator site serves to improve repression. Controlling the amount of TetR confers an additional level of control by placing the *tetR* gene under the control of promoters of various strengths. Additionally, codon usage adaptation of the *tetR* gene to replace codons that occur infrequently in the host cell leads to improving repression by enhancing the translation of the TetR protein[56].

The *E. coli* Tn*10* Tet (B) system has been successfully adapted for use in Gram-positive organisms, namely, *B. subtilis*, *M. smegmatis*, *M. bovis* BCG and *M. tuberculosis*, to generate conditional knockouts for the purpose of functional gene analysis [79-82]. A summary of Tc-regulated expression systems in mycobacteria and *B. subtilis* tested using reporter gene assays is presented in Table 1.1. Ehrt *et al*., 2005[53] isolated the P_{smyc} promoter from *Mycobacterium smegmatis* and selected it to drive the expression of their gene of interest since it was an even stronger promoter than *hsp*60, capable of potent expression and was reportedly equally strong in wild type BCG Pasteur. The isolated P_{smyc} promoter region (189bp) was found to be identical (100% sequence identity) to the upstream region of the putative *rspA* gene of *M. smegmatis* encoding the putative 30S ribosomal protein S1[53]. They developed this

system to generate conditional knockouts in mycobacteria, by replacing the gene of interest's promoter with the strong P_{smyc} promoter containing operator sites (2 x *tet*O2 sites on either side of the –35 region) and by expressing TetR chromosomally under the control of a P_{smyc} promoter (strong) or a P_{imyc} promoter (weak). In this study they showed efficient repression and a range of induction factors of either ß-galactosidase activity or GFP fluorescence with their arrangements. The results of which are summarised in table 1.1. In the same study, it was shown that gene induction is possible intra-cellularly using this system in *M. tuberculosis*, and that ATc is able to readily gain access to the bacterium that resides within a phagosome. This is significant as BCG is an intra-phagosomal bacterium. Guo *et al.,* 2007[82] used the same promoter-operator arrangement also reported efficient repression and significant induction of ß-galactosidase activity. More recently, codon optimised TetRs were generated that were episomally expressed (weak promoter) showed a significant improvement in repression using the same promoter-operator arrangement in which were tested in both *M. smegmatis* and BCG[56].

As mentioned previously, we ideally require a system that can down regulate expression of the foreign antigen during culturing until required upon vaccination thus, reducing the effect of genetic instability and low expression levels generally associated with rBCG vaccines. However, adapting this system for the regulated expression of selected genes would require tetracycline to be continually present for induction. Despite it being used at non-toxic levels, prolonged exposure to an antibiotic is not favourable and may lead to resistance accumulation. The fact that tetracycline would have to be administered to the vaccinees to 'switch on' antigen production would preclude such a vaccine design from approval for commercial development. However, more suitably a reverse mutant of the tetracycline repressor has been developed resulting in a reverse phenotype, which requires the addition of tetracycline to repress gene expression and the removal of tetracycline for induction of gene expression.

Table 1.1. Summary of tetracycline regulated systems in Gram-positive bacteria. This table summarises the results from various studies utilising tetracycline-regulated systems in mycobateria and *B. subtilis.*

Organism and reference	TetR expression	Promoter-operator arrangement	Reporter gene assay and results
M. smegmatis 2005[53]	TetR(B) under the control of a P_{smyc} Promoter (strong promoter)	P_{smyc} Promoter (strong promoter), 17 nucleotides upstream and downstream of –35 was replaced with *tet*O2 sites. (P_{myc1} *tet*O)	**GFP Fluorescence** Efficient repression 150- fold induction with 50ng/ml ATc Induction was high but incomplete and delayed.
M. smegmatis 2005[53]	TetR(B) under the control of P_{imyc} Promoter (weaker promoter)	P_{smyc} Promoter (strong promoter), 17 nucleotides upstream and downstream of –35 was replaced with *tet*O2 sites. (P_{myc1} *tet*O)	**GFP Fluorescence** Efficient repression Complete induction with 50ng/ml ATc. Maximal GFP induction was attained after 4 hours with 50ng/ml ATc 50ng/ml aTc and 100ng/ml ATc induced maximal GFP activity after 15 hours.25ng/ml, 250ng/ml and 500ng/ml l ed to reduced GFP activity.
M. smegmatis 2005[53]	TetR(B) under the contro of a P_{smyc} Promoter (strong promoter)	P_{smyc} Promoter (strong promoter), 17 nucleotides upstream and downstream of –35 was replaced with *tetO*2 sites. (P_{myc1} *tet*O)	**β-galactosidase** Efficient repression 2-orders of magnitude induction with 50ng/ml ATc after 15 hours
M. smegmatis 2005[53]	TetR(B) under the control of P_{imyc} Promoter (weaker promoter)	P_{smyc} Promoter (strong promoter), 17 nucleotides upstream and downstream of –35 was replaced with *tet*O2 sites. (P_{myc1} tetO)	**β-galactosidase** Efficient repression 2-orders of magnitude induction with 50ng/ml ATc after 15 hours
M. tuberculosis 2005[53]	TetR(B) under the control of P_{imyc} Promoter (weaker promoter)	P_{smyc} Promoter (strong promoter), 17 nucleotides upstream and downstream of –35 was replaced with *tet*O2 sites. (P_{myc1} tetO)	**β-galactosidase** Efficient repression 150 to 160-fold increase in β-galactosidase activity after 96h following the addition of 50 or 200ng/ml ATc. Dose dependent increase displayed.
M. smegmatis 2007[82]	TetR(B) under the control of a P_{smyc} Promoter (strong promoter)	P_{smyc} Promoter (strong promoter), 17 nucleotides upstream and downstream of –35 was replaced with *tet*O2 sites integrated into the chromosome. (P_{myc1} *tet*O)	**β-galactosidase** Efficient repression No induction with 30ng/ml ATc

Table 1.1. Continued: Summary of tetracycline regulated systems in Gram-positive bacteria. This table summarises the results from various studies reported in various studies utilising the tetracycline-regulated systems in mycobateria and *B. subtilis.*

Organism and reference	TetR expression	Promoter-operator arrangement	Reporter gene assay and results
M. smegmatis 2007[82]	TetR(B) under the control of P_{imyc} Promoter (weaker promoter)	P_{smyc} Promoter (strong promoter), 17 nucleotides upstream and downstream of –35 was replaced with *tet*O2 sites integrated into the chromosome. (P_{myc1} *tet*O)	**β-galactosidase** Efficient repression Induction with 30ng/ml Max regulation factor of 39 Dose responsive regulation with ATc (1-200g/ml used), 10ng/ml ATc was sufficient for maximum induction TetR protein below level of detection
B. subtilis 2005[79]	Plasmid containing TetR under the control of Pr-xylA (*B. subtilis*)	SpoVG promoter fused to lacZ on a non-replicative integrative plasmid containing 1 *tet*O1 site between the –10 and –35 sites.	**β-galactosidase** Repression at 1.6% β-galactosidase activity therefore did not apply this strain with 1 operator site further
B. subtilis 2005[79]	Plasmid containing TetR under the control of Pr-xylA (*B. subtilis*)	SpoVG promoter fused to lacZ on a non-replicative integrative plasmid containing 2 *tet*O1 sites between the –10 and –35 sites and downstream. (SpoVG-lacZ)	**β-galactosidase** Efficient repression (0.3% β-galactosidase activity) and 300-fold Induction
M. smegmatis **and** ***M. tuberculosis*** 2005[81]	Plasmid containing TetR from *S. aureus* under the control of the antigen 85A promoter (P_{Atet})	On the same plasmid but divergently orientated, GFP under the control of P_{tet} (contains two operator sites)	**GFP Fluorescence** Low levels of expression in un-induced state 20-40ng/ml of tetracycline used for *M. smegmatis* and 200ng/ml for *M. tuberculosis.* Maximum of ≈6 fold induction depending on the type of GFP gene used
M. smegmatis **and** ***M. tuberculosis*** **and** ***M. bovis*** **BCG** 2005[80]	Tet(Z) locus from *Corynebacterium glutamicum* consisting of 1 operator site between the divergently oriented P_{tet}-tetR and gene of interest	On the same plasmid but divergently orientated, luciferase under the control of P_{tet} (contains 1 operator site)	**Luciferase assay** 40-fold induction was observed after 2 hours of exposure to 20ng/ml tetracycline in *M. smegmatis*. 20-fold induction observed in *M. bovis* BCG and 10-fold induction in *M. tuberculosis.*

Table 1.1. Continued: Summary tetracycline regulated systems in Gram-positive bacteria. This table summarises the results from various studies reported in various studies utilising the tetracycline-regulated systems in mycobateria and *B. subtilis.*

Organism and reference	TetR expression	Promoter-operator arrangement	Reporter gene assay and results
M. smegmatis **and** ***M. bovis*** **BCG** 2009[56]	Codon optimised TetR(B)synthetic and TetR(BD)synthetic under the control of P_{imyc} Promoter (weaker promoter)	P_{smyc} Promoter (strong promoter), 17 nucleotides upstream and downstream of –35 was replaced with *tet*O2 sites integrated into the chromosome. (P_{myc1} *tet*O	**β-galactosidase** β-galactosidase activity repressed to 1% with TetR(B) derivatives and repressd to between 1-2% with TetR(BD) derivatives 300ng/ml ATc

1.3.1.3.1. The revTetR system

A reverse TetR system i.e. a revTetR system has been developed which reverse, requires the presence of tetracycline to shut off gene expression[77, 83, 84]. Several reverse mutants have been described that were created by mutagenesis of the TetR protein rendering tetracycline a co-repressor. Scholz *et al*., 2004[84] presented 100 revTetR mutants containing mutations mainly clustered at the interface of the DNA and inducer binding domains that were generated by performing DNA shuffling on *E. coli tetR.* Candidate reverse transactivators were subsequently selected on the basis of efficiency of repression and induction, represented as an induction factor (IF) using a ß-galactosidase assay. Induction factors represented the ratio of ß-galactosidase activity in the presence of ATc (repression) and without ATc (induction). The most efficient revTeR mutant obtained, resulted in an induction factor of 102 and contained three mutations within the alpha 1 helix of the DNA binding domain namely, E15A, L17G and L25V[82, 84]. This mutant was designated TetR r1.7. In contrast to the wild type TetR that originates from the *E. coli* transposon Tn*10* known as TetR(B), this reverse mutant (TetRr1.7) was generated from a chimaeric TetR(BD) protein in which the DNA binding region was from TetR(B) and the protein core was encoded by TetR(D).

The mutations in TetR r1.7 alter the conformational changes that occur such that in the presence of tetracycline, the Tc-revTetR complex binds operator sequences and shuts off gene expression. Upon removal of tetracycline, the revTetR protein

dissociates from the operator sites and the system is 'switched on'. Thus the functioning of the repressor is reversed[77, 83, 84].

The revTetR system, as with the native TetR system possesses the ideal characteristics required for a regulated expression system. It is specific for the co-repressor ATc and does not respond to endogenous activation or interfere with other cellular regulatory pathways. It is efficient in that the system exhibits low basal levels of gene expression during the uninduced state and rapidly achieves high levels of expression during the induced state, which can be efficiently reversed. It is dose dependent in that expression levels respond to the dose of the co-repressor. Additionally this system allows for *in vivo* expression, which is essential for therapeutic application.

1.3.1.3.2. Application of revTetR to mycobacteria

The revTetR (TetR r1.7) system has been used in the gram-positive *B. subtilis*[79] in order to generate conditional knockouts and in *M. smegmatis* to silence essential protein secretion[82]. The system used by Kamionka *et al*., 2005[79] in *B. subtilis* (summarised in table 1.2) makes use of two *tet*O1 sequence elements between the –10 and –35 regions of the *spo*VG promoter on an integrative plasmid. TetR r.1 was expressed of an episomal plasmid with the mutated Pr-*xyl*A promoter. With this system, they managed to repress β-galactosidase activity to a low 0.2% in the presence of 0.4μM ATc and attained 500-fold induction in its absence. Guo *et al.,* 2007[82] , using the same promoter-operator arrangement as for the TetR experiment (i.e. two *tet*O2 sites on either side of the–35 region of the P_{smyc} promoter) and with TetR r1.7 expression driven by the strong P_{smyc} promoter resulted in efficient repression in the presence of as little as 10ng/ml of ATc. They showed that this system responded quickly to non-toxic concentrations of the co-repressor (ATc) and allowed for regulation over a variety of co-repressor concentrations in a dose-dependent manner (10-200ng/ml ATc).

Recently Klotzsche *et al*., 2009[56] generated improved revTetRs by firstly adapting the codon usage of the *tetR(B)*gene to *M. tuberculosis* and secondly by mutating the codon adapted gene to the to include the three TetR r1.7 amino acid changes (E15A,

L17G and L25V) and introducing various TetR(D) residues. This allowed for an increase in revTetR expression without interfering with the DNA binding region of the repressor molecule and thus, an increase in repression. Table 1.2 summarises the use of the TetR r1.7 to control gene expression in mycobacteria.

Table 1.2. Summary of revTetR (TetR r1.7) regulated systems used in gram-positive bacteria. This table summarises the results from various studies reported in various studies utilising the revTetR-regulated systems in mycobateria and *B. subtilis.*

Organism and reference	TetR expression	Promoter-operator arrangement	Reporter gene assay and results
M. smegmatis 2007[82]	TetR r1.7 expressed episomally under control of P_{smyc} Promoter (strong promoter)	P_{smyc} Promoter (strong promoter), 17 nucleotides upstream and downstream of –35 was replaced with *tet*O2 sites integrated into the chromosome. (P_{myc1} *tet*O)	**β-galactosidase** Efficient repression in the presence of 30ng/ml ATc. Max regulation factor of 17-fold. Dose responsive regulation with ATc (1-200ng/ml used) 10ng/ml ATc was sufficient for maximum repression
M. smegmatis 2007[82]	TetR r1.7 expressed episomally under the control of P_{imyc} Promoter (weaker promoter)	P_{smyc} Promoter (strong promoter), 17 nucleotides upstream and downstream of –35 was replaced with *tet*O2 sites integrated into the chromosome. (P_{myc1} *tet*O)	**β-galactosidase** No repression TetR r1.7 protein below level of detection
B. subtilis 2004[79]	Plasmid containing revTetR (TetR r1.7) under the control of a mutated Pr-*xyl*A (*B. subtilis*)	*spo*VG promoter fused to lacZ on a non-replicative integrative plasmid containing 2 *tet*01 site between the –10 and –35 sites.	**β-galactosidase** Efficiently repressed to 0.2% β-galactosidase activity with 0.4μM ATc Induction Factor of 500
M. smegmatis **and *M. bovis* BCG** 2009[56]	Improved revTetRs designated #27 and #28 [codon adapted TetR(B) with E15A, L17G and L25V amino acid changes and several TetR(D) residues] chromosomally integrated under control of P_{smyc} Promoter (strong promoter	P_{smyc} Promoter (strong promoter), 17 nucleotides upstream and downstream of –35 was replaced with *tet*O2 sites integrated into the chromosome. (P_{myc1} *tet*O)	**β-galactosidase** Efficiently repressed β-galactosidase activity to <2% in *M. smegmatis* and #27 and #28 showed 7-fold and 10-fold reduction in % β-galactosidase activity in BCG, respectively

1. 4. Project rationale and design

The long-term goal of this project is to develop rBCG as a successful vaccine vector by overcoming the challenges of low antigen production and instability. The aim is to induce a strong and potent immune response by enhancing genetic stability and expression levels of foreign genes, particularly viral genes that pose a major challenge.

In this project we propose to achieve this by developing an episomal tetracycline repressible revTetR regulated expression system such that during *in vitro* growth of rBCG cultures in the presence of tetracycline, the tetracycline-revTetR complex will bind to promoter-operator regions and in doing so repress the expression of the selected antigen. Thus, minimising the metabolic load and/or toxicity associated with the expression of heterologous antigens and by doing so, maximising genetic stability until the rBCG culture has reached its required density. The removal of tetracycline by washing or by default upon vaccination of an animal will then result in the induction of gene expression to elevated levels, the desired requirement upon vaccination of a subject.

This system will be based on a single episomal shuttle vector design using the reverse mutant TetR (TetR r1.7) generated by Scholz *et al*., 2004[84]. This reverse system functions to repress the expression of the gene of interest with the co-repressor molecule, anhydrous tetracycline (ATc) binding to the regulatory molecule (revTetR protein), which in this 'active conformation' subsequently binds to *tet* operator sites within the promoter and blocks transcription of the gene of interest. By engineering selected strong mycobacterial promoters to contain *tet* operator sequences and expressing a codon optimised revTetR gene constitutively, one may switch off the expression of the target gene during culturing in the presence of ATc and induce expression upon it's removal, once the culture has reached the required density (See Figure 1.4).

In order to optimise this system the model organism *M. smegmatis* and the reporter gene, GFP will be used. However, the long-term goal would be to replace GFP with

our HIV vaccine antigens of choice. The application of this system to vaccine design is novel.

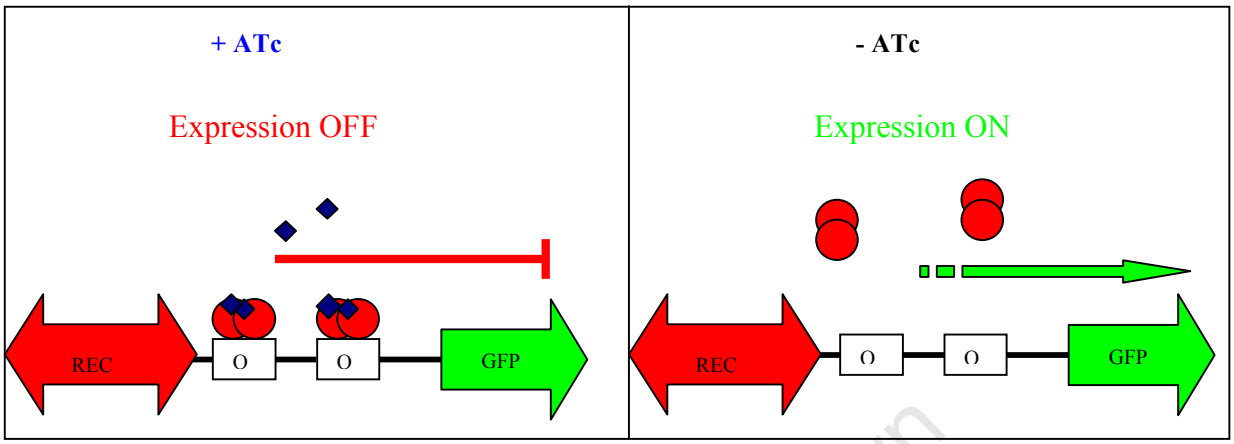


Figure 1.4. Schematic diagram of the proposed revTetR regulatory system. In the presence of ATc (blue diamonds), the expression of the gene of interest/reporter gene (GFP, green arrow) is blocked by the revTetR (red circles) homodimer-ATc complex binding to the *tet* operator sites (O), allowing low level transcription of the revTetR gene (red arrow, REC-revTetR expression cassette, in either orientation). Upon the removal of ATc, the revTetR homodimers dissociate from the operator site and expression of GFP occurs.

1.4.1. Specific aims

1. Characterisation of growth conditions for mycobacterial strains utilised in this study
    - Establishment of mycobacterial strain culturing conditions
    - Establishment of the range of different concentrations of chemical co-repressor (ATc) that may be used before the growth of the mycobacterial strains becomes inhibited
2. Construction of an episomal multicopy revTetR (TetR r1.7) regulated expression system
    - Selection of a suitable promoter to drive the expression of the foreign gene
    - Introduction of *tet* operator sites into the selected promoter region to which the regulatory protein TetRr1.7 binds in the presence of ATc, and the selection of designs which continue to support strong antigen expression

- Introduction of the revTetR cassette consisting of the codon optimised gene encoding the regulatory protein, TetR r1.7 under the constitutive control of a *hsp*60 promoter and strong ribosome binding site.

3. Assessment of the revTetR regulated system in the model organism *M. smegmatis* using the reporter gene, GFP and specifically testing for:
    - Repression of GFP expression in the presence of ATc
    - Induction of GFP expression upon the removal of ATc
    - Dose dependency of induction and repression

.

Chapter 2: Materials and Methods

Chapter 2: Materials and Methods

2.1. Mycobacterial strains and culture conditions

Three mycobacterial strains were used in this project, *M. smegmatis* mc^2155 (BioVac institute, Cape Town, S.A), *M. bovis* BCG Pasteur and a pantothenate auxotroph derived from *M. bovis* BCG Pasteur, BCG Pasteur Δ*panCD* mc^26000 (BCG panCD) obtained from Dr. William Jacobs Jr. (Albert Einstein College of Medicine, New York).

Mycobacteria are characterised by a complex cell wall containing a "network" of mycolic acids[85]. This poses a challenge when propagating mycobacteria in liquid culture, as cells tend to aggregate into macroscopic clumps that affect optical density readings (OD), colony forming unit (CFU) estimates and the uniform exposure of cells to nutrients and other metabolites added to the media. For this reason a detergent such as tween-80 or tyloxapol is routinely added to the growth media to ensure the homogenous dispersion of cells. We have chosen to use tyloxapol since it is better accepted in the formulation of vaccine stocks for the inoculation of animals.

General culturing conditions for all mycobacterial strains were as follows unless otherwise indicated. A 10% glycerol stock of mycobacterial cells stored at -80°C were removed and allowed to thaw on ice. A 0.5ml aliquot was used to inoculate 5ml of sterile Middlebrook 7H9 liquid media (M7H9) containing 0.01%w/v tyloxapol (TXL) detergent and supplemented with 10%v/v oleic-acid albumin dextrose complex (OADC) [BioLab]. These starter cultures were incubated in 50ml conical sterile tubes [Sterilin] standing at 37°C with loosened lids for approximately 2 days for *M. smegmatis* and 5 days for the BCG strains. The starter culture volumes were subsequently increased to 10ml with the same M7H9-TLX-OADC media and further incubated at 37°C until an OD_{600} of approximately 0.5 was obtained. The 10ml starters were then used to inoculate 100ml of the same M7H9-TLX-OADC media in 1L Schott bottles. BCG cultures were first incubated standing for 48hrs at 37°C with loosened lids before being placed on rollers until the required cell density was obtained.

Since the BCG Pasteur Δ*panCD* mc^{2}6000 strain carried a hygromycin marked deletion of the pan*CD* genes required for pantothenate biosynthesis, the propagation of this strain required the additional supplementation with and 0.05mg/ml Hygromycin B [Roche diagnostics] and a reference amount of 0.048mg/ml of pantothenic acid [Sigma-Aldrich] was indicated for use during culturing.

2.2. Pantothenate growth requirements for *M. bovis* BCG Pasteur Δ*panCD* mc^{2}6000

BCG panCD 10ml starter cultures were propagated as described in Section 2.1 in M7H9-TLX-OADC media approximately supplemented with 0.05mg/ml Hygromycin B and 0.048mg/ml of pantothenic acid. Once *s*tarter cultures had grown to an OD_{600} of approximately 0.5, equal volumes of 0.5ml of culture was used to inoculate 100ml of the same M7H9-TLX-OADC media containing various concentrations of pantothenate that varied from 0mg/ml to 0.12mg/ml. Cultures were further propagated as outlined above and OD_{600} readings were recorded every 24 hours on a spectrophotometer [Beckmann, DU-40] using cell density as an indicator of bacterial growth. For accuracy, cultures were diluted to read between OD_{600} 0.1 and 0.5 where necessary.

2.3. Minimum inhibitory concentration (MIC) of anhydrous tetracycline

The minimum inhibitory concentration (MIC) range of anhydrous tetracycline (ATc) [Acros Organics] was determined for the three mycobacterial strains *M. smegmatis*, BCG Pasteur and BCG panCD. ATc was obtained in powder form, reconstituted in high purity liquid chromatography (HPLC) [BDH] grade water to a stock concentration of 1mg/ml and stored as 1ml aliquots at -20°C for subsequent use. ATc is a light-sensitive molecule and therefore all procedures involving its preparation or subsequent handling were carried out under conditions of minimal light. Aliquots

were covered in foil during defrosting. Following the establishment of *M. smegmatis*, BCG Pasteur or BCG panCD 10ml starter cultures in the appropriate M7H9-TLX-OADC media as described in Section 2.1, identical volumes (1ml) of each of the three strains was inoculated into 100ml of the same media, containing various concentrations of ATc, ranging from 0 ng/ml to 800ng/ml for *M. smegmatis* and 0ng/ml to 5000ng/ml for BCG Pasteur and BCG panCD. Culture bottles were incubated in the dark on rollers and protected from light during handling. OD_{600} readings were recorded every 4 hours for *M. smegmatis* and every twenty-four hours for BCG.

2.4. *E. coli* culturing conditions

Escherichia coli DH5α was used as the cloning host. Recombinant *E. coli* clones were grown in 2YT broth (16g of tryptone, 10g of yeast extract and 5g NaCl per litre) supplemented with 1.6mg/ml kanamycin [Sigma-Aldrich] for plasmid selection and propagated at 37°C on shakers for aeration. Cells were plated onto 25ml 2YT agar medium solidified with 15g of agar per litre, supplemented with 1.6mg/ml kanamycin and incubated overnight at 37°C.

2.5. Preparation of *E. coli* competent cells

E. coli DH5α cells were made competent using a modification of the dimethyl sulphoxide (DMSO) method by Chung and Miller 1988 [86]. A single colony of *E. coli* picked off a freshly streaked 2YT agar plate was used to inoculate 5ml of 2YT broth. After overnight growth the culture was diluted 1/100 in 100ml of 2YT broth and propagated for approximately 3 hours to early log phase (OD_{600} 0.2-0.4). The cells were harvested by centrifugation at 5000 rpm for 5 min at 4°C. The pellet was resuspended in 1/10 volume (10ml) of ice-cold sterile TSB buffer (10% w/v PEG, 5% v/v DMSO, 10mM $MgCl_2$, 10mM $MgSO_4$). The cells were held on ice for ten minutes and either transformed immediately or stored in cold sterile glycerol to a final concentration of 10% v/v in 100µl aliquots at -80 °C.

2.6. Transformation of plasmid DNA into *E. coli* DH5α cells

E. coli DH5α cells were transformed using standard cold-shock transformation methodology[87]. Glycerol stocks of competent cells were thawed on ice. Plasmid DNA (0.1ng up to 100ng DNA) was immediately mixed into a 100µl aliquot of cells. The cells were then incubated on ice for a further 20min, followed by a heat-shock step at 37 °C for 5 min. Room temperature 2YT media (1ml) was added prior to incubation of the vials at 37 °C for 30-60 min to allow for expression of the antibiotic resistance marker. Tubes were gently mixed by inversion during this incubation step and 50 µl, 100 µl and 200 µl aliquots were plated onto the appropriate selective 2YT solid media. The remaining transformation mix was pelleted, resuspended in 100µl 2YT broth and plated as well.

2.7. General enzymatic manipulations

All restriction enzyme digests were performed according to the manufacturer's instructions [Roche Diagnostics] in a final volume of 20µl unless otherwise indicated. Ligation reactions were performed using a T4 DNA ligase kit [Roche Diagnostics] and were carried out overnight at 4°C at various vector: insert ratios (1:1, 1:3 and 1:5). Dephosphorylation reactions were performed with Shrimp Alkaline Phosphatase (SAP)[Roche Diagnostics] to remove 5'phosphate groups for the prevention of spontaneous religation of the plasmid backbone in question. A T4 DNA polymerase kit [Roche Diagnostics] was utilised for filling in 5' overhangs with unlabelled dNTPs. Plasmid DNA was purified from reaction mixes using a QIAquick Purification Kit [Qiagen] where required while desired DNA fragments were purified from agarose gels using the QIAquick Gel Extraction Kit [Qiagen]. Following any enzymatic manipulations, except ligations, DNA integrity and concentration was confirmed on a mini 10ml 1% agarose slide gel containing Ethidium Bromide (EtBr) that was electrophoresed at 10V for 10-15min. Restriction enzyme reactions were electrophoresed on 1% standard agarose gels (100-200ml) containing EtBr for 1.5hours at 90V using a 1Kb size marker [Promega]. All the

agarose gels were visualised over a UV box and digital images captured and edited with Uvipro [Uvitech Silver] equipment and software, respectively.

2.8. Design and synthesis of P_{smyc} promoter-operator constructs and the TetR r1.7 expression cassette

The promoter-operator constructs used in this study were designed by Dr. Helen Stutz and based on the promising results reported by Ehrt *et al*., 2005[53]. These authors isolated the P_{smyc} promoter from *Mycobacterium smegmatis* using a promoter trap library. It was chosen for use in our application as it was shown to be a particularly strong constitutive promoter supporting high expression levels in both *M. smegmatis* and BCG[53] and was only nominally downregulated by the insertion of two *tet* operator 1 (*tet*O1) sequences on either side of the –35 sequence.

Analysis of the 64bp published P_{smyc} promoter sequence against the incomplete *M. smegmatis* genome sequence using the Artemis software [Sanger Sequencing Institute] revealed that it was identical to the upstream region of the putative *rpsA* gene encoding the 30S ribosomal protein S1. Since there was a significant distance between the identified transcriptional start and translational start sequences (125bp) we sought to determine whether the lengthy leader sequence had a significant effect on the expression of the GFP reporter gene by designing two promoter elements, one which comprised the full-length natural P_{smyc} promoter (NM1) and a truncated version (NM2). Based on the outcome of this, duplicate insertions of the *tet* operator sequences *tet*O1 or *tet*O2 or a combination of the two were designed flanking the –35 region in order to assess the effect of the different sequence arrangements on the expression and regulation of GFP. These promoter and promoter-operator elements were synthesised by GeneArt® GmBh and supplied separately cloned into the manufacturer's backbone carrying an Ampicillin^R marker. Sequence features of the five different elements synthesised (NM1-NM5) are summarised in table 2.1.

The 208 amino acid mutant revTetR binding regulatory protein TetR r1.7[84], selected for use in this study, carries three amino acid substitutions (E15V, L17G and L25V)

in the DNA reading head such that it now binds DNA in the presence of ATc, and dissociates from DNA in it's absence. It was codon optimised for optimal expression in mycobacteria and incorporated into an 821bp expression cassette comprising the *hsp*60 promoter, a strong ribosome binding site (RBS D) and bordered by the *hsp*60 transcriptional terminator. This sequence cassette (designed by Dr. Helen Stutz) was synthesised by GeneArt® and supplied in a standard vector with an Ampicillin^{R} marker.

Table 2.1. Description of the P_{smyc} promoter-operator fragments synthesised. The 5' and 3' restriction sites utilised for purposes of cloning are underlined: *TCTAGA* represents one strand of the 5'-*Xba*1 site and *GAATTC* represents one strand of the 3'-*EcoR*1 site. Transcription start sites (+1) are boxed. Sequences in blue represent *tet* operator 2 sequences (*tet*O2) and those in red represent *tet* operator 1 sequence (*tet*O1). Bold and underlined 6bp sequences represent –35 and -10 regions.

Construct name	Promoter description	Length (bp)	Sequence
NM1	Full-length natural P_{smyc} promoter	194	5'-*TCTAGA*ATGTGCGTTCGCACGCACAGGCCCGGTG TGAGAAGGGTCTCTGCAGAGCGGGAGAACCCACCCG GGGTGGGCGAGT**TTGTCC**TGCGTGTGCTCGGTCGAG **TAGGCT**CTGGGAGTACCCGTGTGTACGACCAGCACG GCATACATCATTTCGACGCCGAGAGATTCGCCGCCC GAAATGAGCAC*GAATTC*-3'
NM2	Truncated P_{smyc} natural P_{smyc} promoter	90	5'-*TCTAGA*CTGCAGAGCGGGAGAACCCACCCGGGG TGGGCGAGT**TTGTCC**TGCGTGTGCTCGGTCGAG**TAG GCT**CTGGGAGTACCC*GAATTC*-3'
NM3	Full length P_{smyc} promoter containing two *tet*O2 sequences	199	5'-*TCTAGA*GGATCCTGTGCGTTCGCACGCACAGGC CCGGTGTGAGAAGGGTCTCTGCAGAGCGGGAGAAC**T CCCTATCAGTGATAGAGTTTGTCCTCCCTATCAGTG ATAGATAGGCT**CTGGGAGTACCCGTGTGTACGACCA GCACGGCATACATCATTTCGACGCCGAGAGATTCGC CGCCCGAAATGAGCAC*GAATTC*-3'
NM4	Full length P_{smyc} promoter containing two *tet*O1 sequences	200	5'-*TCTAGA*GGATCCTGTGCGTTCGCACGCACAGGC CCGGTGTGAGAAGGGTCTCTGCAGAGCGGGAGAAC**A CTCTATCATTGATAGAGTTTGTCCACTCTATCATTG ATAGAGTAGGCT**CTGGGAGTACCCGTGTGTACGACC AGCACGGCATACATCATTTCGACGCCGAGAGATTCG CCGCCCGAAATGAGCAC*GAATTC*-3'
NM5	Full length P_{smyc} promoter containing one *tet* O1 sequence and one *tet*O2 sequence	199	5'-*TCTAGA*GGATCCTGTGCGTTCGCACGCACAGGC CCGGTGTGAGAAGGGTCTCTGCAGAGCGGGAGAAC**A CTCTATCATTGATAGAGTTTGTCCTCCCTATCAGTG ATAGATAGGCT**CTGGGAGTACCCGTGTGTACGACCA GCACGGCATACATCATTTCGACGCCGAGAGATTCGC CGCCCGAAATGAGCAC*GAATTC*-3'

2.9. Construction, screening and confirmation of phase 1 and phase 2 *E. coli*-mycobacterial shuttle vectors

The episomal *E. coli*-mycobacterial shuttle plasmid, pRC100 was chosen as the parental backbone for the development of the inducible system. It was derived from the pAL5000-based plasmid pWB100[48] which consists of an oriM, an oriE, a kanamycin resistance gene. pRC100 contains a green fluorescent protein reporter gene (GFP) under the control of strong expression and translation initiation signals; the *M. tuberculosis hsp*60 promoter and ribosome binding site D (RBS D).

The inducible episomal system was developed in two stages referred to as Phase 1 and Phase 2. Phase 1 consisted of the construction of control and test constructs involving the deletion or replacement of the *hsp*60 promoter with our various promoter-operator elements (NM1-NM5) into the pRC100 backbone. Phase 2 involved the subsequent insertion of the TetR r1.7 expression cassette into selected Phase 1 constructs.

2.9.1. Phase 1 cloning: insertion of P_{smyc} promoter-operator elements into pRC100

The strategy for insertion of promoter-operator elements into the parental plasmid pRC100 is illustrated in Figure 2.1 and details of the cloning steps employed are outlined diagrammatically in Chapter 4.

The pRC100 vector backbone (Kan^R) was prepared by excising the existing *hsp*60 promoter region by double digestion with the restriction endonucleases *Xba*1 and *EcoR*1. The resulting fragments were separated by electrophoresis. The larger required 4845bp fragment was purified using a QIAquick gel extraction kit, and the recovered product quantitated on a mini gel. This vector backbone was further manipulated in two ways. One aliquot was subjected to filling in of the 5' overhangs with a T4 DNA polymerase kit and ligated to serve as a negative promoter-less control in all subsequent experiments. The remaining aliquot was subjected to a dephosphorylation reaction with SAP to prevent spontaneous religation of the

linearised backbone prior to use in the ligation reactions involving the promoter-operator elements NM1-NM5.

The size of the promoter-operator elements varied from 90bp to 218bp making it very difficult to visualise these small fragments on a gel and to purify them from the manufacturers backbone for subsequent cloning. Thus, the cloning strategy was designed using compatible restriction endonuclease sites to ensure directional cloning and a negative antibiotic selection step to preferentially select for the insertion of the small promoter-operator fragment over the GeneArt® ampicillin resistant backbone fragment.

The various inserts (NM1-NM5) were digested out of the manufacturer's backbone with the restriction endonucleases *Xba*1 and *EcoR*1 by first setting up two separate reactions with each of the enzymes (250ng of plasmid DNA per reaction), checking for complete digestion by mini gel electrophoresis, then pooling the two reactions and digesting for a further 1.5hrs to ensure complete digestion of the small inserts. Following a reaction clean up step, ligation mixes were set up to contain; linearised dephosphorylated vector backbone DNA (50ng), and the varying amounts of insert DNA (NM1-NM5) present as a mixture of GeneArt® backbone and insert, ligase and HPLC water to a final volume of 10μl. The concentration of GeneArt® backbone DNA and insert DNA added was calculated such that a vector: insert ratio of 1:1, 1:3 or 1:5 was achieved.

Resultant ligation mixes were transformed into *E. coli* cells DH5α using standard cold-shock transformation methodology[87]outlined above, and putative positive recombinants identified following duplicate patch plating of selected colonies onto two sets of 2YT agar plates; one set contained both kanamycin and ampicillin selection and the other contained kanamycin selection only. Those colonies that grew on the Kan^+ 2YT agar plates but not on the Kan^+ and Amp^+ 2YT agar plates were considered potential positive recombinants for further consideration since this implied that the smaller *Xba*1-*EcoR*1 promoter-operator fragment had been cloned in and not the larger GeneArt® backbone fragments carrying the Amp^R gene.

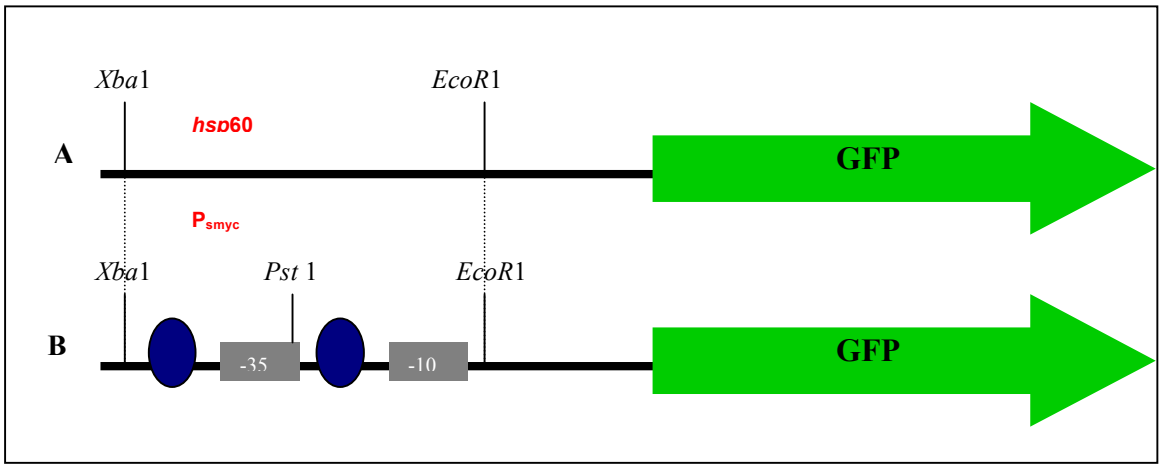


Figure 2.1. Strategy for the cloning of P_{smyc} promoter-operator constructs into pRC100. A) Represents the parental plasmid pRC100 and B) represents the derived P_{smyc} based constructs. The *hsp60* promoter was replaced via the *Xba*1 and *EcoR*1 sites with a full length P_{smyc} promoter (NM1), truncated P_{smyc} promoter (NM2) or a P_{smyc} promoter-operator fragment containing two *tet* operator 2 sites (NM3), two *tet* operator 1 sites (NM4) or a combination of both (NM5) situated on either side of the –35 promoter region. The blue circles represent the operator sites.

2.9.2. Phase 2 cloning: insertion of TetR r1.7 expression cassette into selected Phase 1 constructs

The second phase of developing this inducible system involved the introduction of the 821bp TetR r1.7 expression cassette, wherein the *tetR* r1.7 regulatory gene under the control of the *M. tuberculosis hsp60* promoter into selected Phase one constructs (see Figure 2.2). The TetR r1.7 cassette was excised out of the GeneArt® vector backbone by a double digest with *Xba*1 and *Spe*1 restriction enzymes, electrophoresed on a 1% agarose gel and the required 821bp fragment gel purified with a QIAquick gel extraction kit and quantitated on a mini slide gel. The selected Phase 1 constructs were linearised via *Xba*1, dephospohorylated with SAP and the DNA purified. Digestion with *Xba*1 generates a 5' overhang that is also compatible with the 5' overhang generated by *Spe*1 digestion. Ligation reactions were carried out overnight and resulting ligation mixes were transformed into *E. coli* DH5α cells. Expression mixes were plated on selective 2YT agar and potential positive recombinants screened for by restriction enzyme mapping of plasmid DNA that had been isolated by a small-scale plasmid DNA isolation process. It should be noted that the TetR r1.7 expression cassette could be cloned in either orientation and that the position of the reconstituted *Xba*1 site in conjunction with other restriction enzyme

sites could be used to determine the orientation of the insertion of the TetR r1.7 cassette.

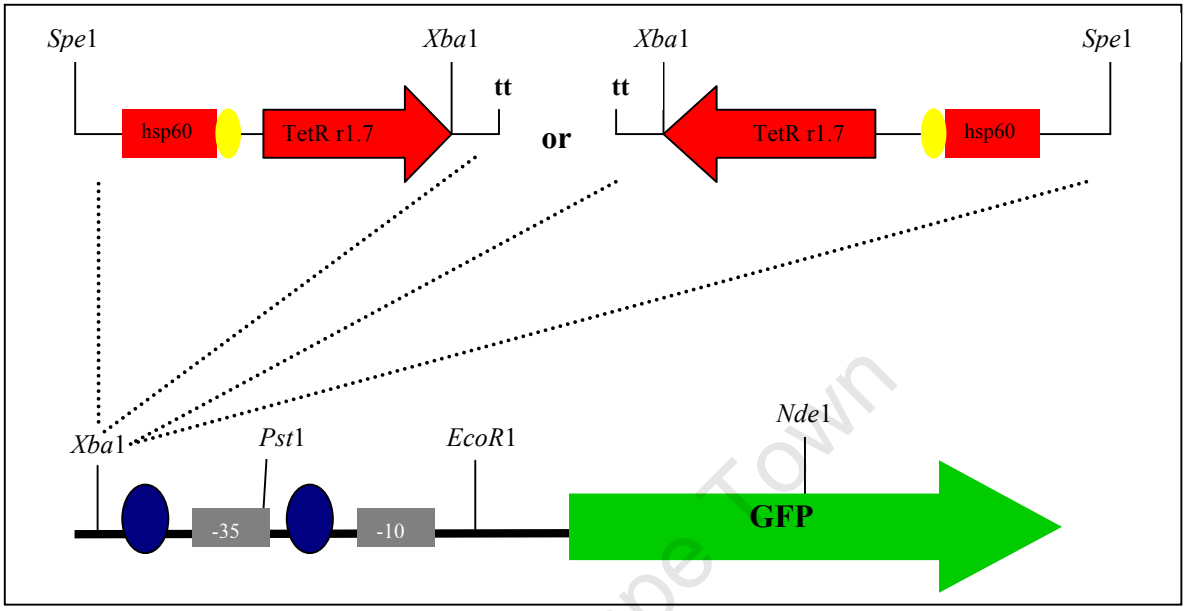


Figure 2.2. Strategy for the insertion of the TetR r1.7 cassette into selected Phase 1 promoter-operator plasmids. The TetR r1.7 cassette, comprising the *hsp60* promoter, ribosome binding site D (RBS D) {yellow ovals}, the *tetR r*1.7 regulatory gene and the *hsp60* transcriptional terminator (tt) was cloned into the *Xba1* site of selected Phase 1 constructs in either orientation. The blue circles represent operator sites.

2.9.3. Preliminary screening of phase 1 and phase 2 recombinants by restriction enzyme digestion of miniprep DNA

Transformants resulting from Phase 1 and Phase 2 cloning strategies were picked off 2YT agar plates and inoculated into 800µl of 2YT media containing the appropriate antibiotic in a microcentrifuge tube and cultures incubated overnight, standing at 37°C. Reference plates were prepared simultaneously by streaking each selected transformant onto a numbered grid on a 2YT agar plate, thus ensuring the maintenance of recombinant clones for further reference. Cultures were then subjected to a standard small-scale plasmid isolation procedure by alkaline lysis[88] and resuspended in 50µl of HPLC grade water Resulting Phase 1 and Phase 2 miniprep DNA was digested with a combination of enzymes according to the manufacturer's guidelines. The products were separated by electrophoresis visualised

over a UV box and the image captured with a digital camera. Samples that yielded the expected fragment sizes were identified.

2.9.4. Confirmation of positive recombinants by restriction enzyme mapping of Maxiprep DNA

Following initial screening, selected recombinants were picked off the reference plates prepared during the small-scale plasmid isolation process, cultured in 100ml 2YT media with selection on an orbital shaker at 37°C. Plasmid DNA (20-60μg) was extracted from 50ml of culture with an adapted Nucleobond AX PC-Kit [Machery-Nagel] procedure. For maximum plasmid recovery, subsequent to the elution step, 5ml of eluent was divided into five 2ml eppendorf™ tubes, followed by the addition of 700μl of isopropanol to each tube and centrifugation at 15000rpm at 4°C for 30 min. Supernatant was removed and each pellet resuspended in 90μl of HPLC water. Suspensions were pooled, 50μl of 3M Na-Acetate (1/10V) added, 1ml of 100% Ethanol added and tubes centrifuged at room temperature at 15000rpm for 10mins to precipitate the DNA. Supernatants were removed, pellets dried and washed with 2ml of 70% Ethanol. Tubes were then subjected to centrifugation at room temperature at 15000rpm for 10 min, followed by a second drying step. Pellets were finally resuspended in 30-40μl of HPLC water. Purified plasmid DNA was quantified using a Nanodrop Spectrophotometer [Inqapa Biotech] set to read DNA at 260nm.

Maxiprep DNA (500ng-1μg) was subjected to various enzymatic digestions. Phase 1 constructs were subjected to *Pst*1 and *Mlu*1 single enzyme reactions and *Xba*1+ *Mlu*1 double digests by standard methods. Phase 2 constructs were mapped by *EcoR*V single digests, *Mlu*1+ *Not*1 and *Xba*1+ *Nde*1 double digests. The products were electrophoresed on a 1% agarose gel alongside the 1kb marker, visualised over a UV box and images captured and edited.

2.9.5. Confirmation of positive recombinants by sequencing of Phase 1 and Phase 2 constructs

Recombinants were confirmed by sequencing across the pRC100 vector backbone-insert junctions. Primers used were designed with the aid of a Primer Design software package [Freeware, version1] and are listed in Table 2.2. Primers were synthesised by Integrated DNA Technologies [Whitehead Scientific].

Cycling reactions were performed using the Big Dye Terminator v3.1 cycle sequencing kit [Applied Biosystems] with 150-300ng of template DNA, 4μl of termination mix, 3.2pmol of each primer and 2.5 times dilution buffer made up to a final volume of 20μl with HPLC water. Cycling reactions were performed using the following conditions: 2 cycles of 96 °C for 30s, 1 cycle of 50 °C for 15s, 1 cycle of 60 °C for 4 min and 1 cycle of 72 °C for 7 min. Cycle sequencing reactions were performed in the GeneAmp® PCR System 9700 thermocycler [Applied Biosystems] and were resolved by the sequencing unit of the Department of Human Genetics, University of Cape Town. Sequences were viewed with the Chromas v2.3 software package [Technelysium] and edited and aligned against constructed theoretical sequences using DNAMAN bioinformatics software [Lynnon Biosoft, version 4.0.0.1].

Table 2.2. List of primers used for sequence confirmation of constructs. Primers were designed using Primer Design software package [Freeware, version1] and were synthesised by Integrated DNA Technologies [Whitehead Scientific].

Primer	Orientation	Sequence	Purpose
GFP_R	Reverse	5'-TTTGTGCCCATTAACATCACC-3'	Binds to GFP for sequencing across pRC100/promoter-operator junctions of Phase 1 constructs
26F	Forward	5'-TGGCGAACTCCGTTGTAGTG-3'	Binds vector for sequencing across TetRr1.7 / promoter-operator junctions and pRC100/ TetRr1.7 junctions
07-0320R	Reverse	5'-GATGTATGCCGTGCTGGTCG-3'	
FP1	Forward	5'-TGAATATCGTGGAGCTCACC-3'	Binds to vector, TetR r1.7 cassette and promoter-
FP2	Forward	5'-CCTGTACTGGCACGTGAAGA-3'	
FP3	Forward	5'-ATCTCGGCCGTGTCGCACTT-3'	

Primer	Orientation	Sequence	Purpose
FP4	Forward	5'-CTCCTCTGGTTGGTACTCTA-3'	operator elements for sequencing across entire promoter-operator region and TetRr1.7 cassette
RP1	Reverse	5'-CGCTTGTTCTTCACGTGCCA-3'	
RP2	Reverse	5'-AAGTGCGACACGGCCGAGAT-3'	
RP3	Reverse	5'-TCTAGAGTACCAACCAGAGG-3'	
RP4	Reverse	5'-CCGTATGTTGCATCACCTTC-3'	

2.10. Preparation of mycobacterial competent cells

The three mycobacterial strains used in this project were made competent by methods described by Parish and Stoker, 2001[89]. After starter cultures were established, inoculums were propagated in 100ml of M7H9-TLX-OADC media in the absence of antibiotic selection, except for BCG panCD, which required culturing in the presence of Hygromycin B, and pantothenic acid as described in Section 2.1 to an OD_{600} of 0.8-1 units. Cultures were then incubated on ice for 1.5 hours with slow rotation and harvested by centrifugation at 4^{o}C, at 5000rpm for 10min. Cells were kept cold on ice during all subsequent procedures to enhance competency. The pellets were then washed three times in pre-chilled 10% glycerol (prepared with HPLC grade water) and finally resuspended in 5ml pre-chilled 10% glycerol representing a 20–fold concentration. The competent cells were subsequently divided into 300µl aliquots and stored at –80^{o}C.

2.11. Electrotransformation of plasmid DNA into mycobacteria

Confirmed plasmid DNA constructs (500ng-1µg) resulting from Phase 1 and 2 were made up to a final volume of 5µl in HPLC water. Laboratory stocks of the reference plasmid DNA (pWB100) was used as a positive control and a no DNA sample served as a negative control for transformation. *M. smegmatis* mc^{2}155, BCG Pasteur and BCG panCD competent cells were thawed on ice and pelleted by centrifugation at room temperature, at 4000rpm for 1min. Supernatant was discarded and cells gently resuspended in 500µl ice-cold 10% glycerol and kept on ice. An aliquot of competent cells (40µl) was then mixed with a plasmid DNA sample and transferred to a pre-chilled 0.1mm cuvette [Biorad] and incubated on ice for a further minute.

The competent cell mixture was then electroporated using the Gene Pulser™ [BioRad] set at 1.8kV, 25 μF and1000Ω for efficient plasmid uptake. Time constants were noted and 1ml of M7H9-TLX-OADC expression medium supplemented with pantothenate and Hygromycin B for the BCG panCD samples was added immediately under sterile conditions. Transformation mixes were transferred into 20ml tubes [Sterilin] and left standing with loosened lids at 37°C for 3-4hrs for *M. smegmatis* and overnight for BCG. Cultures were plated at a range of dilutions on 35ml of Middlebrook 7H10 (M7H10) solid media supplemented with 10%v/v OADC, 10μg/ml kanamycin for plasmid selection as well as 0.048mg/ml pantothenate and 0.05mg/ml Hygromycin B for BCG panCD transformants. Plates were sealed in plastic bags to prevent dehydration and incubated at 37°C for 3 days for *M. smegmatis* and 3-5 weeks for rBCG.

Recombinant mycobacterial colonies were cultured in liquid media as describe in Section 2.1 except that colonies were first broken apart manually using sterile toothpicks and then dispersed in the 5ml starter culture by a gentle pipetting action. After significant growth was observed, the starters were swirled, allowed to settle and the top homogenous layer used to sub-inoculate a second 10ml starter in order to eliminate the presence of bacterial clumps in the cultures and incubated until the desired cell density was reached prior to scaling up to 100ml. Glycerol stocks were made for all the recombinant mycobacterial cultures and stored at -80°C for future reference.

2.12. Expression studies in mycobacteria

Three techniques were employed to assess comparative levels of GFP expression resulting from our different construct designs. Initially, GFP expression levels were compared by visualisation of plated transformants over a UV trans-illuminator. In order to quantitatively compare GFP expression levels between constructs two methods were employed. Initially, GFP capture ELISA was utilised but later fluorimetry became the assay of choice due to better reproducibility and shorter processing times.

2.12.1 Visualisation over a UV trans-illuminator

Plates containing similar numbers of colonies, plated at the same time and under identical conditions, were visualised over a Chromato-vue TM-20 Trans-illuminator [UVP] at 245nm (UV range) and photos taken at various exposures (-0.1s to +0.1s) using a digital camera [Nikon 4500]. The relative number of transformants, colony size and relative GFP fluorescent intensities were noted as a preliminary indication of GFP expression levels.

2.12.2 GFP quantification by Enzyme-Linked Immuno-absorbant Assay (ELISA)

GFP capture ELISAs were performed on proteins extracted from recombinant mycobacterial cultures.

2.12.2.1 Protein extractions from recombinant mycobacterial cultures and quantification

Starter cultures were grown from selected recombinant mycobacterial colonies as described in Section 2.11 and recombinant cultures were further propagated as outlined in Section 2.1 to an early logarithmic OD_{600} of between 0.4-0.5 units. A fraction (5ml) of the culture was stored as 1ml 15% glycerol stocks at -80°C for future reference. Samples of culture (25ml) were then centrifuged at 5000rpm in SS34 tubes at 4°C for 15min, and the pellets re-suspended in 5ml of sterile phosphate-buffered saline (PBS pH 7.4). Of this, 4ml were stored as 1ml aliquots at -80°C for future use. The remaining 1ml was transferred into FastRNA tubes [Bio 101] containing 0.5g of 0.1mm zirconia-silica beads [BioSpec] that had been sterilised by autoclaving and the cells lysed using a FastPrep machine [Bio 101]. The samples were processed using setting 6 for 30-second bursts, followed by incubation on an ice-slurry for 2min intervals. This regime was repeated three times. The samples were then subjected to centrifugation at 14000rpm at room temperature for 10 minutes. The supernatant was retained and centrifuged a second time to remove residual debris. The supernatant was retained and divided into three aliquots. Two of these were stored at -80°C for future use. The remaining aliquot was subjected to a standard Dc-Biorad protein quantification assay utilising Bovine Serum Albumin

(BSA) [Roche Diagnostics] to set up a standard curve. Samples of protein were stored as 1mg/ml and 0.5mg/ml aliquots at -80°C.

2.12.2.2. GFP quantification by GFP-capture ELISA

GFP present in the protein lysate extracted from recombinant mycobacterial cultures outlined above was quantitated in an ELISA system that used polyclonal anti-GFP pre-coated NUNC plates [Pierce], polyclonal anti-GFP HRP-conjugated antibody [Abcam] for detection, and recombinant purified GFP protein [Clontech] for the generation of a linear standard curve. The ELISA procedure had to be optimised with respect to the sample dilutions used to ensure that the ELISA readings for all the samples were within the range of the linear standard curve. In order to do this, a range of sample dilutions were tested; 1:50, 1:100, 1:200, 1:250 and 1:500. Furthermore, a range of secondary antibody concentrations were tested; 1:1000, 1:2000, 1:4000, 1:5000, 1:10 000 and 1:20 000. A combination of a 1:500 sample dilution and 1:20 000 secondary antibody dilution yielded the best results and was used routinely (data not shown). In addition, due to the finding that the mycobacterial proteins interfered with GFP detection, thus affecting ELISA readings, the rGFP standards and all the samples were diluted in 1:500 pWB100 lysate that had been made up in Phosphate-buffered saline, 0.5% Tween-20 (PBS-T) and 1% milk powder. By ensuring the concentration of mycobacterial proteins remained equivalent in all the samples and standards we minimised the effect of background interference.

Previously quantified protein samples were defrosted on ice and diluted 1:500 in PBS-T and 1% milk powder. A volume of 100µl of the samples and standards were loaded in triplicate in 96-well NUNC polyclonal anti-GFP pre-coated plates and incubated for 1hour at 37°C. Wells were washed four times with PBS-T using the Elx50 Autostrip washer [Bio-Tek Instruments] and incubated with 100µl of a 1:20 000 dilution of a goat polyclonal anti-GFP HRP-conjugated antibody for 1 hour at 37°C before being washed again as above. Wells were subsequently incubated with 100µl of substrate [TMB Microwell Peroxidase Substrate System, KPL] for 30mins at room temperature in the dark following the manufacturer's instructions. The

detection reaction was stopped using 100µl of 0.5M H_2SO_4 per well. Absorbency readings were taken using a PR2100 ELISA plate reader [Sanofi Diagnostic Pasteur] at 450nm and the readings for each sample averaged. The amount of GFP (ng of GFP/ng total protein) was determined from the standard curve. GFP capture ELISAs were performed at least 3 times.

2.12.3. Whole Cell GFP Fluorimetry in *M. smegmatis*

Relative GFP expression levels for recombinant *M. smegmatis* cultures were assessed by whole cell GFP fluorimetry, using an assay described by Carrol and Parish, 2004[81].

Recombinant cultures were propagated from 15% glycerol stocks and grown to an OD_{600} of 0.4-0.5 units. The cultures were centrifuged at 5000rpm for 10 min and the pellets were resuspended in Phosphate-buffered saline (PBS) to a calculated density of approximately 2 OD_{600} units/100µl, the density referenced by Carrol and Parish, 2004[81]. Samples were briefly vortexed prior to loading 100µl of sample per well into a 96-well white fluorimetry plate [NUNC] in triplicate. PBS was loaded as a negative control to ensure that it did not affect fluorescence readings. Culture harbouring the plasmid pWB100, which does not express GFP, was used to account for background interference caused by mycobacterial proteins. Fluorescence readings were taken using the Varian Cary Eclipse Fluorimeter set to read GFP fluorescence at an excitation wavelength of 485nm and emission wavelength of 515nm.

In order to obtain accurate fluorescence readings and to correct for any inaccuracies encountered during the concentration process, OD_{600} readings of the sample pellets resuspended in PBS were taken to determine the correction factor (2OD units/ OD_{600} reading). Fluorimetry results were calculated by taking an average of the triplicate readings, multiplying it by the correction factor, subtracting the average of the pWB100 readings, and then expressing GFP expression as relative fluorescence units (RFU).

2.12.3.1 Proof of concept experiment: GFP induction-repression studies in *M. smegmatis* with ATc

In order to test whether our model GFP system could be repressed in the presence of the co-repressor ATc and induced in it's absence, the following experiment was conducted as outlined in Figure 2.3. This experiment comprised two groups and was performed in two separate batches but each carried out identically. Recombinant *M. smegmatis* starter cultures containing plasmids pWB100, pNM3, pNM3.1, pNM3.2 (group 1) and pWB100, pNM4, pNM4.1 and pNM4.2 (group 2) were set up as described previously. These were subsequently increased to 10ml of appropriate M7H9-TLX-OADC media with antibiotic selection for a further 2days and then used to inoculate 120ml of the same media. These cultures were grown to an OD_{600} of approximately 0.5 units. A volume of 25ml of this culture was used to inoculate each of four1L Schott containing 5ml of fresh media and varying amounts of ATc such that the concentrations achieved were 0ng/ml, 25ng/ml, 50ng/ml or 100ng/ml in a 30ml final culture volume. This was considered time point zero. These cultures were placed on rollers at 37°C and 5ml of culture removed for fluorimetry assays as outlined above at various time points (0, 2, 8 and 12 hours) to determine repression. Induction was assessed after removal of ATc from cultures by centrifugation at 5000rpm and resuspension in the same media (without ATc). This washing step was performed twice. Culture samples were removed at 6 and 12 hours after the ATc had been removed.

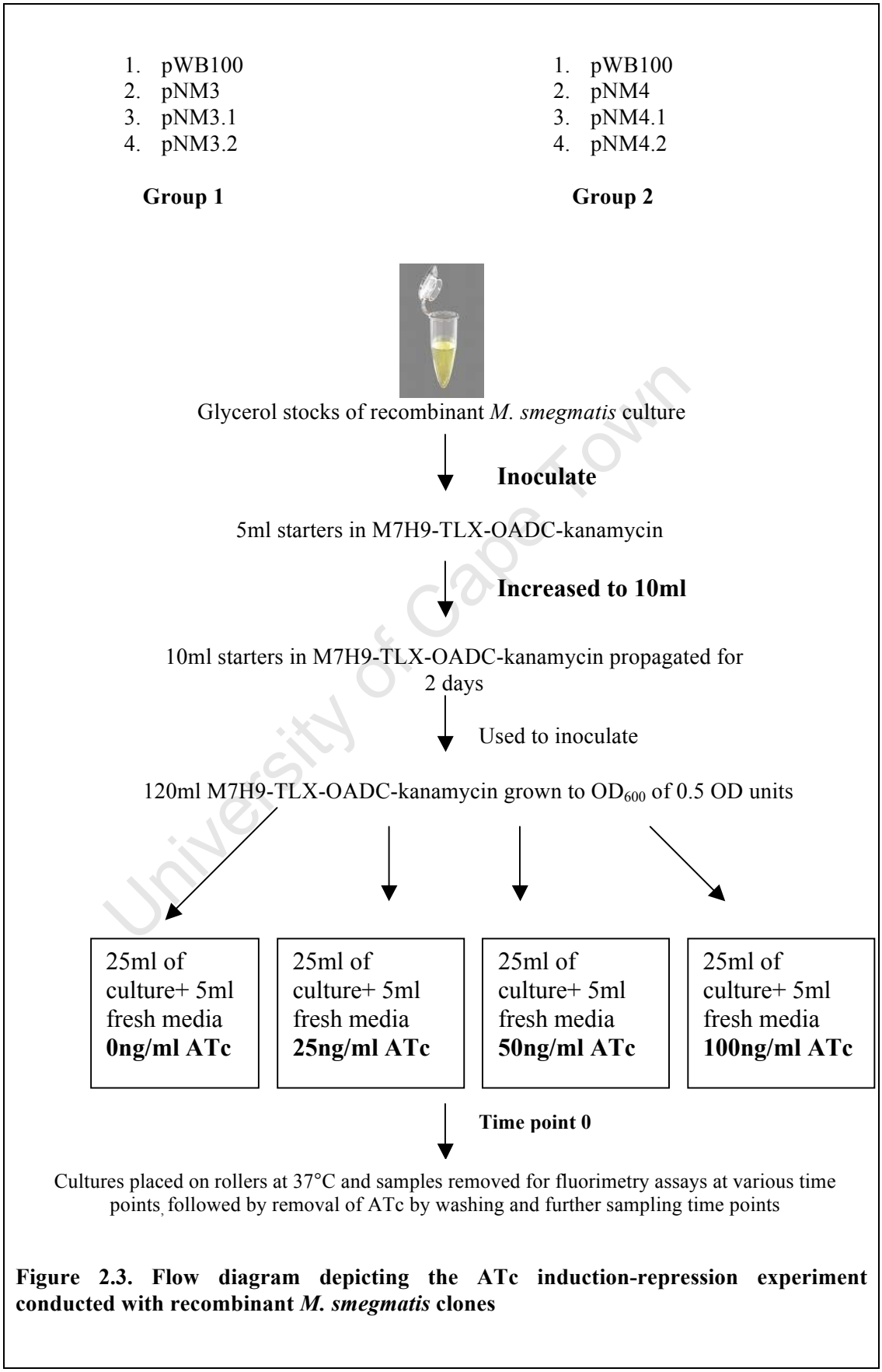


Figure 2.3. Flow diagram depicting the ATc induction-repression experiment conducted with recombinant *M. smegmatis* clones

2.13. Plasmid stability in the model organism *M. smegmatis*

Recombinant cultures stored as 15% glycerol stocks were diluted in 7H9 media (1:1000 and 1:10000) and plated on M7H10-OADC solid media with kanamycin selection. Following incubation for 3 days at 37°C colony size was noted and the percentage of GFP fluorescent colonies was calculated as an assessment of plasmid stability.

Chapter 3:Optimisation of mycobacterial culture conditions

Chapter 3:Optimisation of mycobacterial culture conditions

3.1. Introduction

To date BCG remains the only available vaccine for the treatment of tuberculosis (TB). Thus, as recommended by the World Health Organisation (WHO), it is routinely administered to newborns in TB endemic regions. BCG is a live attenuated vaccine that replicates in immunocompetent hosts before cell-mediated responses are mounted to halt its replication. It is however contraindicated for use in immunocompromised hosts, such as HIV positive infants, since recent evidence has suggested that BCG vaccination places them at risk of developing 'BCG disease' which is potentially fatal as a result of their compromised immune system[90, 91]. This has led to the WHO revising their recommendation and advising that BCG not be administered to HIV positive asymptomatic infants in TB endemic regions[92]. Thus, the need to develop safer and more efficacious TB vaccines has become more urgent. The use of an auxotrophic strain of BCG that is replication impaired provides an attractive alternative strategy as it would still provide the benefit of protection against severe pulmonary TB in infants, whilst minimising the risk of developing disseminated BCG in immunocompromised individuals, including these highly susceptible HIV positive infants. Such a replication deficient strain of BCG would also be an attractive vector for the development of a recombinant dual TB-HIV vaccine.

In order to optimise and standardise growth conditions to ensure consistent results, the following experiments were performed:

i. Optimising growth requirements of a pantothenic acid auxotrophic strain of BCG Pasteur i.e. *M. bovis* BCG Pasteur Δ*panCD* $mc^2$6000 (BCG panCD).

ii. Establishing the minimum inhibitory concentration (MIC) of anhydrous tetracycline (ATc) to the growth of BCG panCD, wild type BCG Pasteur and the model organism, *M. smegmatis*. ATc is a tetracycline analogue selected for use as the co-repressor molecule to control the expression of heterologous antigens in our vaccine vector systems.

3.2. *M. bovis* BCG Pasteur Δ*panCD* mc^{2}6000 (BCG panCD), a pantothenate auxotroph

BCG panCD (obtained from W.R Jacobs Jr.) is an auxotroph derived from wild type BCG Pasteur that is unable to synthesise the metabolite pantothenic acid (Vitamin B5). The strain was created by a specialized method of mycobacteriophage transduction[93], and contains a non-reverting deletion of two adjacent genes, *pan*C and *pan*D, that prevents the *de novo* biosynthesis of pantothenic acid. Pantothenic acid is necessary for the synthesis of coenzyme A and acyl carrier protein (ACP) that play a vital role in fatty-acid biosynthesis and metabolism, see Figure 3.1[94].

The first study of a mycobacterial pantothenate auxothroph was conducted with a *M. tuberculosis* mutant (mc^{2}6030) constructed by Sambandamurthy *et al* [93, 94]. Significantly, they observed that it possessed a superior safety profile in immunocompromised mice in comparison to the current BCG vaccine, *M. bovis* BCG Pasteur[95]. More recently, a pantothenic acid auxotroph recombinant BCG strain, rBCG (panCD)30 was generated in a similar manner by Tullius *et al*., 2008[96]. It was designed to episomally overexpress the *M. tuberculosis* 30-kDa major secretory protein r30, also known as antigen 85B. This strain demonstrated similar efficacy when compared to wild type BCG in the outbred guinea pig model of pulmonary TB, however it resulted in a superior safety profile in SCID mice. Sambandamurthy *et al* [93, 94] reported the supplementation of both liquid and solid phase growth with 24μg/ml of pantothenate for the *M. tuberculosis* (mc^{2}6030) strain [95] whilst Tullius *et al*., 2008[96] reported that 10-50μg/ml was sufficient to supplement growth of rBCG(panCD)30, after the initial starter innoculum was cultured in 50μg/ml pantothenate. Since there is no published information on the growth conditions for BCG panCD, our initial task was to establish the optimal concentration of pantothenate to supplement growth of this strain in order to approximate wild-type BCG growth and minimise stress imposed by the limitation of this essential growth nutrient.

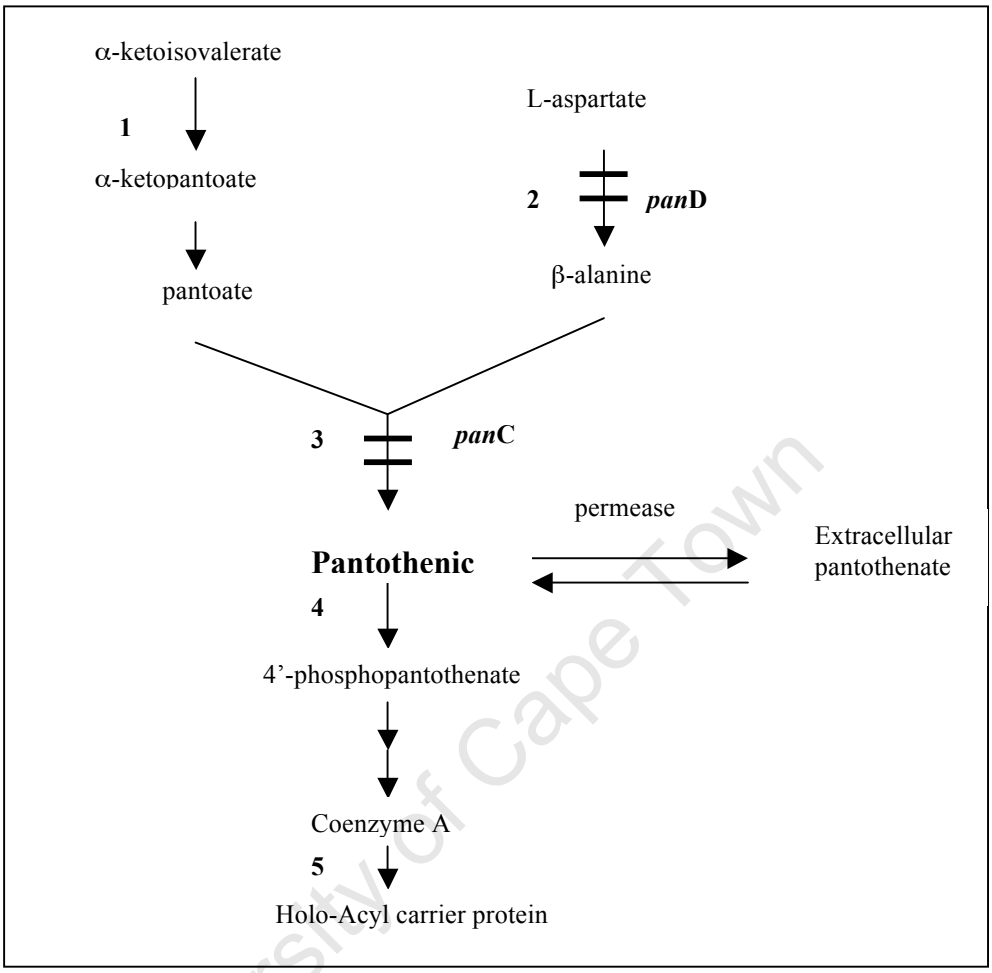


Figure 3.1. Biosynthetic pathway for the production of Pantothenic acid and Coenzyme A in *M. tuberculosis.* Bold numbers denotes the enzymes involved in this pathway. 1-*panB*, ketopantoate hydroxymethyl transferase. 2-*panD*, aspartate-1-decarboxylase. 3-*panC*, pantothenate synthetase. 4-*panK*, pantothenate kinase. 5-*acpS*, ACP synthase. Disruptions to the pathway are denoted by a = at positions 2 and 3.This diagram was reproduced from Sambandamurthy *et al*., 2002[94].

3.2.1. Determination of optimal pantothenate growth requirement for BCG panCD

Glycerol stocks were used to inoculate BCG panCD starter cultures as described in Section 2.1. Since this strain carries a hygromycin marked deletion of the panCD genes, the starter culture was supplemented with 0.05mg/ml hygromycin and grown in the presence of 24μg/ml of pantothenate until the required OD_{600} was reached. Identical aliquots of the starter culture were used to inoculate 100ml cultures

containing the various concentrations of pantothenate: 0µg/ml, 24µg/ml, 36µg/ml, 48µg/ml and 120µg/ml).

Growth of BCG panCD was notably enhanced in a dose-dependent manner up to a concentration of 48µg/ml (Figure 3.2). Once the pantothenate concentration exceeded this i.e. 120µg/ml, growth retardation was observed suggesting the existence of a threshold level for the enhancing effect of pantothenate. No growth was observed in the absence of pantothenate confirming its auxotrophic phenotype. Thus, 48ug/ml of pantothenate was selected as the optimal concentration for supplementation of growth of BCG panCD. In contrast, the addition of the various concentrations of pantothenate to our standard M7H9-TLX-OADC media had no effect on growth of the wild type Pasteur strain of BCG, with a maximum culture density at OD_{600} of 2.0 units being reached after 380 hours of incubation (data not shown).

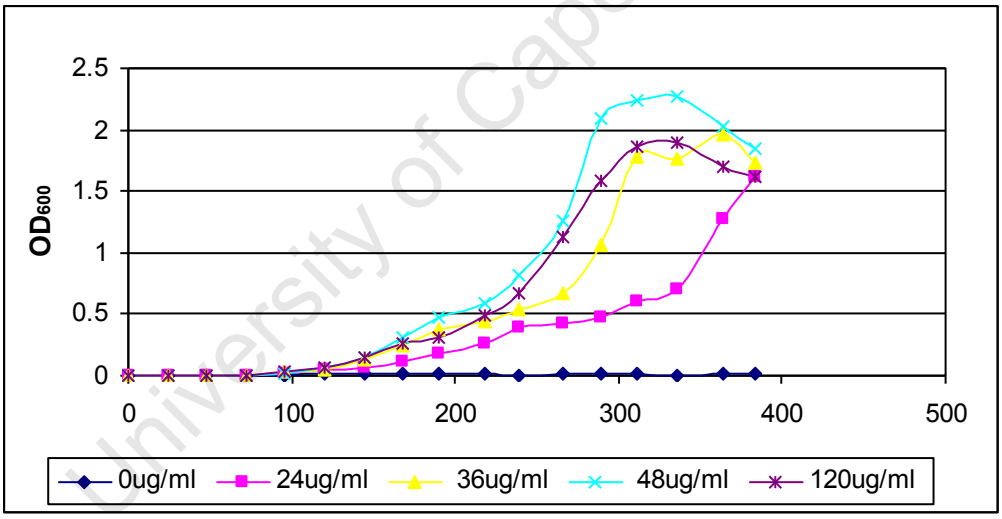


Figure 3.2. The effect of various concentrations of pantothenate on the growth of BCG panCD. Cultures containing the various amounts of pantothenate i.e. 0µg/ml, 24µg/ml, 36µg/ml, 48µg/ml and 120µg/ml were propagated under the appropriate conditions. OD_{600} readings were taken every 24 hours, using cell density as an indication of growth.

3.3 Determination of the minimum inhibitory concentration (MIC) of anhydrous tetracycline ATc for the growth of *M. smegmatis* mc^{2}155, *M. bovis* BCG Pasteur and BCG panCD.

As described in Section 1.3.1.3, tetracycline (Tc) is the natural inducer or repressor of the TetR and revTetR systems, respectively. Tetracycline exerts its antibiotic activity at the transcriptional level by binding to the small (30S) ribosomal subunit, thereby interfering with protein synthesis and retarding growth[72, 97]. However, since Tc possesses a considerably higher affinity for TetR and revTetR over the host 30S ribosomal subunit, in the order of 10^3-10^5 fold, it acts as a highly efficient inducer or repressor, respectively. Anhydrous tetracycline (ATc) is an antibiotic analogue of Tc that is commonly used in Tc regulatable systems as it binds the Tet repressor protein even more efficiently than Tc, is less toxic to mammalian cells, and retains high functional stability in culture, thus acting as a superior inducer or repressor[78].

In order to harness the full potential of this regulatory system, it was necessary to first establish what range of the co-repressor (ATc) may be used to regulate the revTetR system before inhibition of growth of the bacteria occurred i.e. it was necessary to establish the minimum inhibitory concentration. The MIC range was established for the three mycobacterial strains: BCG panCD, the wild type *M. bovis* BCG Pasteur from which the pantothenate auxotroph was derived and *M. smegmatis* mc^{2}155.

3.3.1 Minimun Inhibitory Concentration (MIC)

To accurately establish the MIC for all three strains, starter cultures were grown to a homogenous density at an OD_{600} of 0.5 units. Following the establishment of starter cultures, identical volumes were used to inoculate identically prepared M7H9-TLX-OADC media containing different concentrations of ATc, incubated at 37°C in the dark as outlined in Section 2.3 and the optical density (OD_{600}) of the cultures measured over time. No differences could be observed in the growth profiles of BCG panCD cultured in the presence of varying concentrations of ATc up to 1500ng/ml of

the compound (results not shown). This auxotrophic strain showed that in the presence of >1500ng/ml of ATc, growth was affected (Figure 3.3). However, growth was only significantly affected in the presence of 5000ng/ml of ATc, thus establishing an MIC for this strain. This is significantly lower than that seen for the wild type BCG Pasteur, since growth of wild type BCG Pasteur was not at all inhibited by the presence of up to 5000ng/ml of ATc (Figure 3.4). Thus, the MIC for wild type BCG Pasteur is at a much higher range than 5000ng/ml. The model organism *M. smegmatis*, is much more sensitive to ATc, with an MIC range between 400-600ng/ml (Figure 3.5).

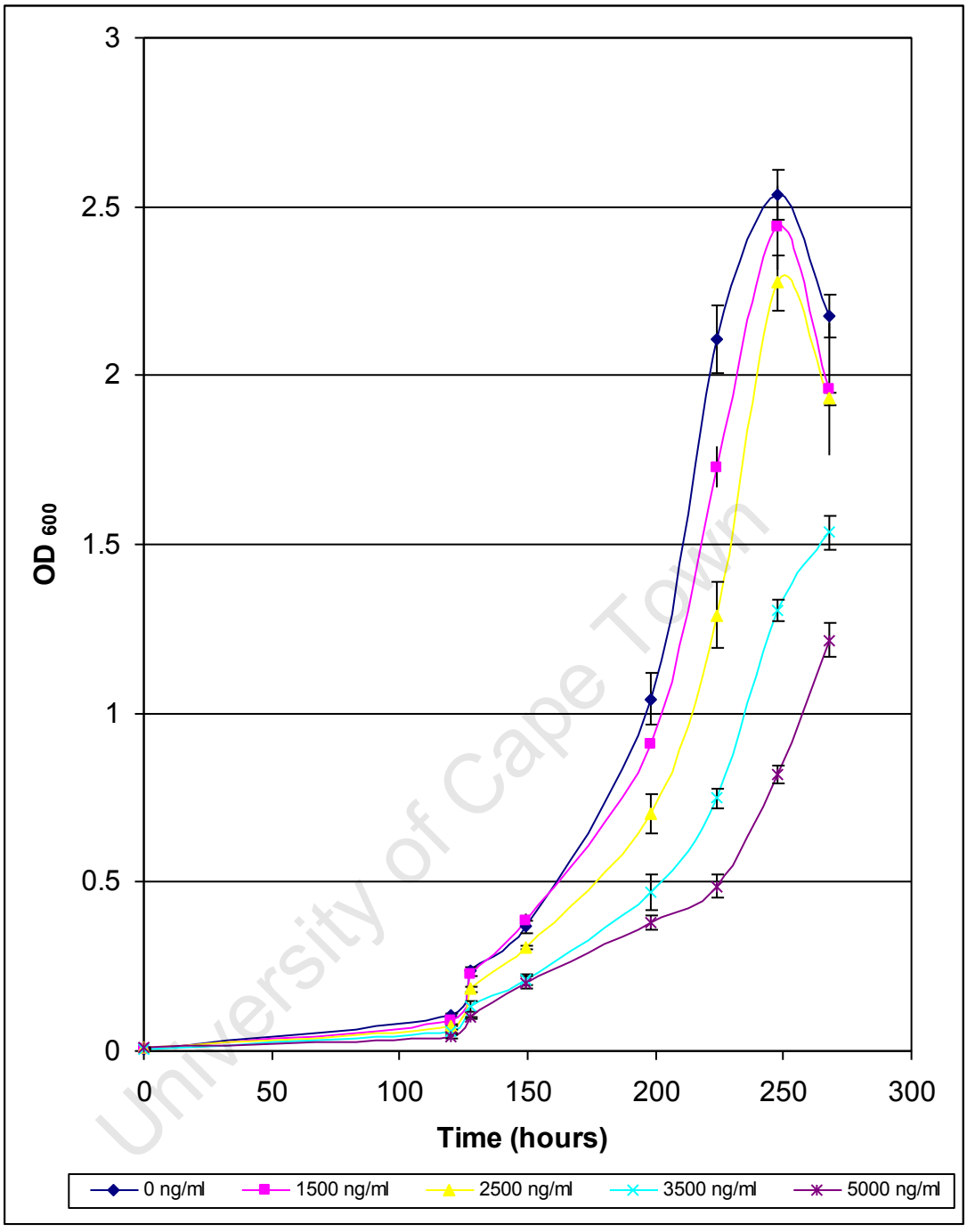


Figure 3.3. Growth curve of BCG panCD in the presence of varying concentrations of ATc to establish the MIC. Cultures were grown in the presence of various amounts of ATc, i.e. 0, 1500, 2500, 3500 and 5000ng/ml as listed in the legend below the graph. OD_{600} readings were taken every 24 hours, using cell density as an indication of growth. Data points represent the average of three independent readings and error bars are as indicated.

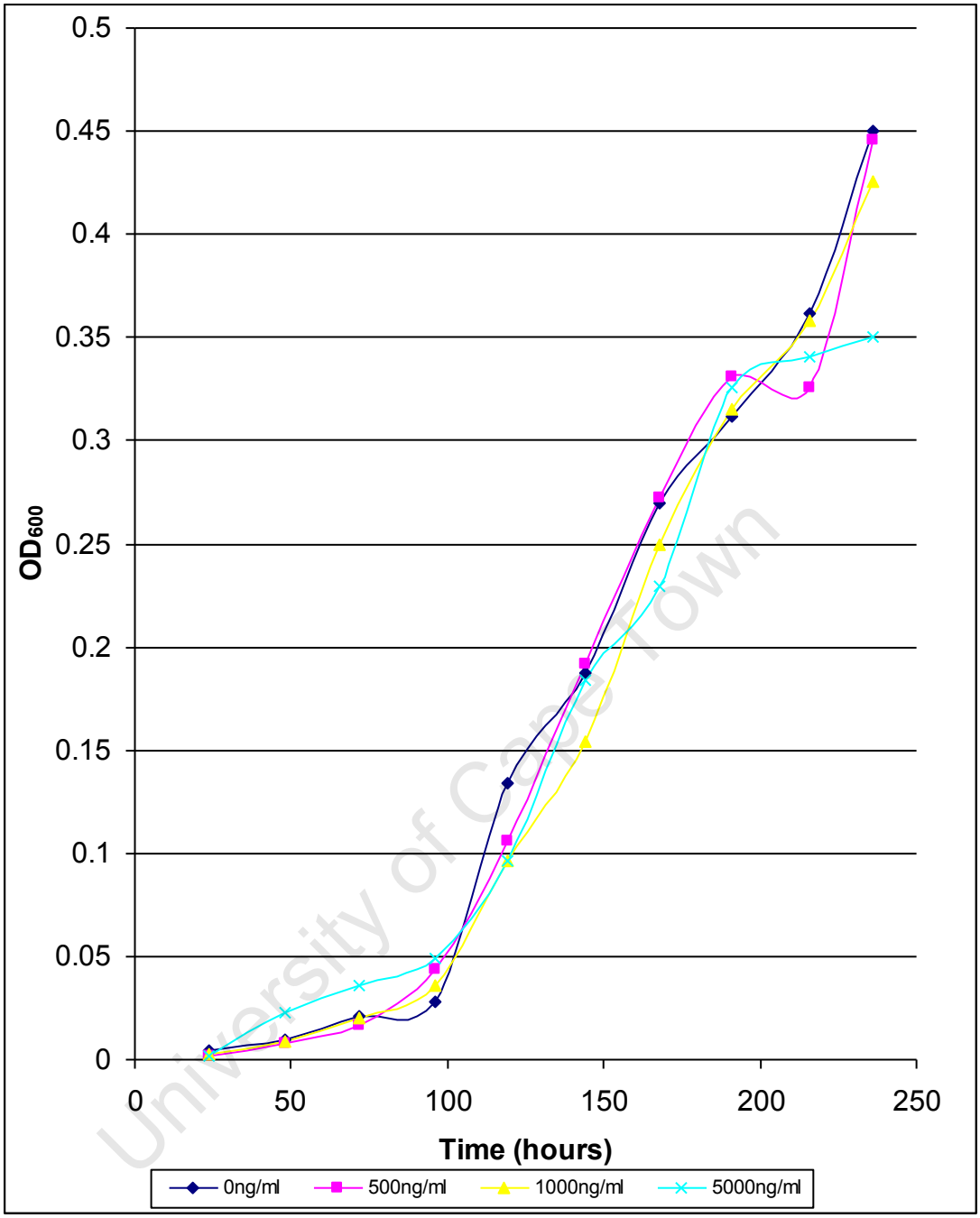


Figure 3.4. Growth curve of wildtype BCG Pasteur in the presence of varying concentrations of ATc to establish the MIC. Cultures were grown in the presence of various amounts of ATc, i.e. 0, 500, 1000 and 5000ng/ml as listed in the legend below the graph. OD_{600} readings were taken every 24 hours, using cell density as an indication of growth.

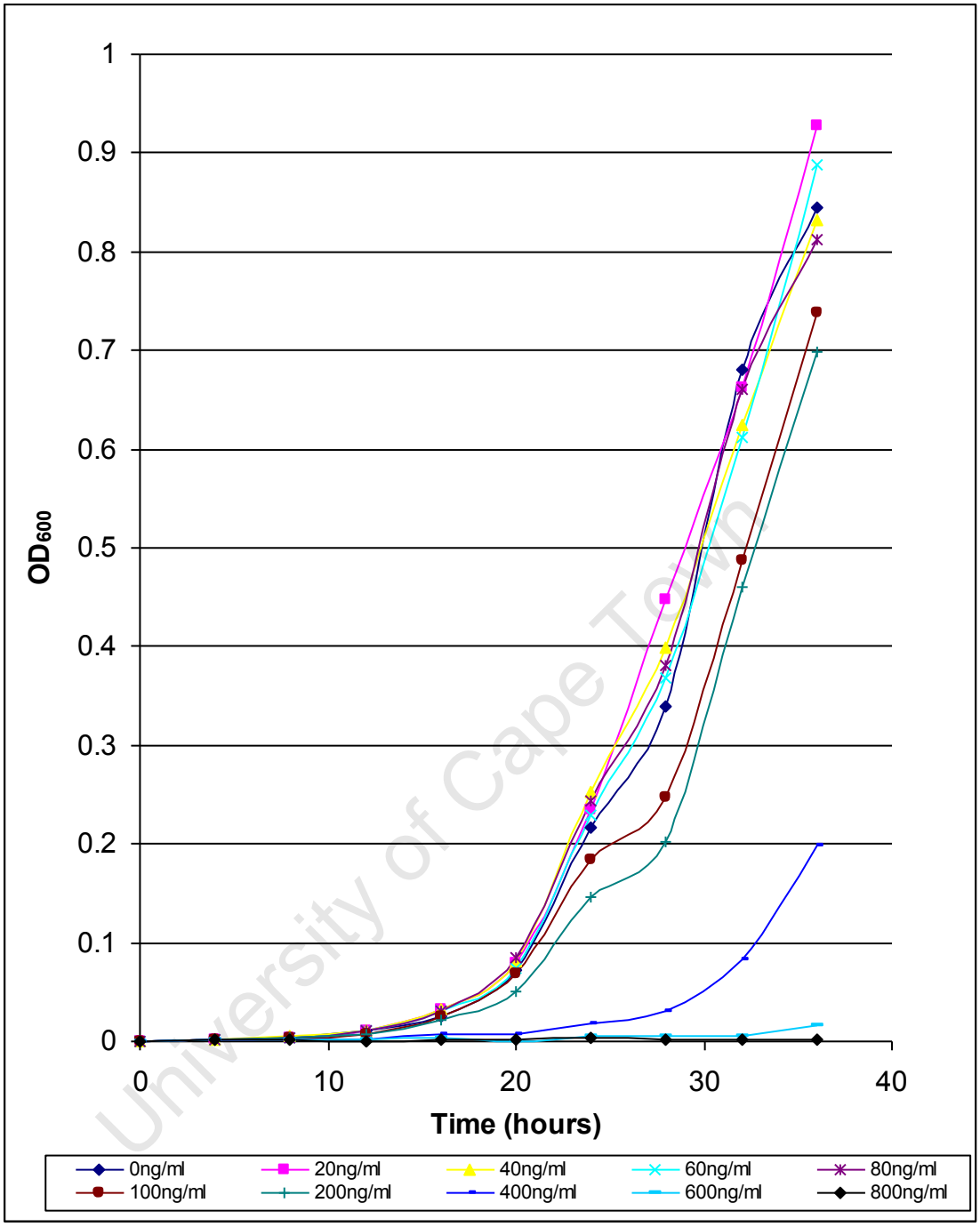


Figure 3.5. Growth curve of *M. smegmatis* mc^{2}155 in the presence of varying concentrations of ATc to establish the MIC. Cultures were grown in the presence of various amounts of ATc, i.e. 0, 20, 40, 60, 80, 100, 200, 400, 600 and 800ng/ml as listed in the legend below the graph. OD_{600} readings were taken every 24 hours, using cell density as an indication of growth.

3.4. Discussion

The addition of an adequate amount of exogenous pantothenate is essential for the optimal growth of the auxotrophic strain. Our results show that the addition of 48μg/ml of pantothenate represented the optimal concentration to supplement the growth of this auxotrophic strain, BCG panCD and eliminated any stress imposed by the limitation of an essential nutrient. This concentration represents twice the amount of pantothenate referenced for use in the *M. tuberculosis* $mc^2$6030 studies[95] and is 4.8 times the minimum concentration referenced for use in the rBCG(panCD)30 studies i.e. 10μg/ml[96]. While the growth rates observed for cultures containing either 48μg/ml or 120μg/ml of pantothenate appeared similar, the final optical density reached for the higher pantothenate concentration (120μg/ml) was lower, suggesting that excess pantothenate in fact became growth limiting and should be avoided.

Interestingly, we also noted that the rate of growth of the auxotroph exceeded that of the wild type strain (data not shown) as the concentration of pantothenate increased up to 48μg/ml. We hypothesise that this is due to this strain possessing a potentially "less complex" cell wall as a result of the altered metabolism of lipid biosynthesis[94] thus, allowing for easier access to nutrients in the media in comparison to wild type BCG.

The MIC's of ATc for the growth of BCG panCD and wild type BCG Pasteur have not been documented. Here we report that the MIC for the growth of BCG panCD was established at approximately 5000ng/ml of ATc. This was notably lower than that observed for wild type BCG Pasteur, which displayed no growth retardation up to a concentration of 5000ng/ml of ATc. This increased sensitivity of the auxotrophic strain to ATc could potentially be attributed to the aforementioned hypothesis of a "less-complex, lipid-poor" cell wall, thus allowing for a more rapid influx of ATc resulting in it exhibiting its antibiotic activity at a lower concentration in comparison to BCG Pasteur. Results for the MIC of ATc in *M. smegmatis* agree well with studies published by Ehrt *et al*., 2005[53] who reported an MIC of 500ng/ml of ATc, which is within the range we determined i.e. 400-600ng/ml of ATc. The literature sites a range of 0.03ng/ml-50ng/ml of ATc used for successful induction or repression

studies in *M. smegmatis* and wild type BCG[53, 80, 82, 98]. This is well below the MIC for both organisms, thus we do not expect to observe any toxic effects or growth retardation with similar ATc concentrations that we plan to use in this project.

BCG panCD not only represents a safer strain for use in immunocompromised individuals, but is also more sensitive to the presence of exogenous ATc making it more suitable for the development of a rBCG vaccine in which the expression of antigens are controlled by the presence or absence of ATc. However, since *M. bovis* BCG has a doubling time of twenty-four hours and requires six weeks for the formation of colonies on plates, *M. smegmatis* is routinely used as a model organism for studies involving mycobacteria[81] as it is non-pathogenic, only has a doubling time of four hours and requires only three days for the growth of transformants on solid media. Thus, for this project we focused our efforts on optimising the revTetR regulatory system in *M. smegmatis* with the long-term goal of applying it to the development of an effective rBCG panCD HIV vaccine.

Chapter 4:Construction and selection of a revTetR controlled expression system

Chapter 4:Construction and selection of a revTetR controlled expression system

4.1. Introduction

Our goal in this study was to develop a revTetR controlled expression system for application in mycobacteria such that expression of the selected gene of interest is repressed in the presence of ATc but induced to high levels upon removal of ATc. In order to achieve this, we chose to use a single episomal multicopy plasmid approach in which the DNA binding regulatory protein TetR r1.7 is constitutively expressed by a strong promoter, and GFP used as a model reporter gene under the control of an engineered mycobacterial promoter containing *tet* operator sites for regulation of GFP expression. The episomal revTetR system was developed using an *E. coli*-mycobacterial shuttle vector based on a vector backbone (pRC100) that had been previously optimised in our laboratory for stability in mycobacteria (manuscript in preparation). It is derived from the pAL5000 origin of replication and produces 3-5 copies per cell[99]. Thus, compared to integrative systems in which only 1 copy of the gene(s) of interest is present, a multicopy episomal system is expected to support higher levels of induced antigen expression, a notable requirement for the development of an effective vaccine strategy.

The episomal, multicopy, mycobacterial revTetR system will be developed in two phases:

Phase 1. Construction and selection of suitable promoters engineered to contain *tet* operator (*tet*O) sequences that support strong GFP expression.

Phase 2. Insertion of the TetR r1.7 expression cassette into selected Phase 1 plasmid constructs and selection of promising vaccine shuttle vector designs.

4.2. Phase 1: Construction and selection of promoter-operator constructs expressing GFP

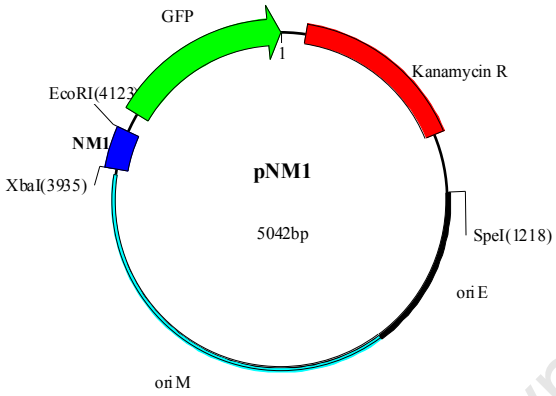
4.2.1. Selection of a strong mycobacterial promoter

The first step in developing this system involved the selection of a strong promoter to drive GFP expression in mycobacteria. The most commonly used promoters are *hsp*60 and *hsp*70, those of the 60kDa and 70kDa heat-shock proteins of *M. tuberculosis* respectively[100]. These promoters are naturally induced under stress conditions and capable of constitutively expressing heterologous genes to levels of up to 10% of the total mycobacterial protein content[2, 44, 101]. However, several groups have found that the expression of viral genes by the *hsp*60 promoter often resulted in genetic instability and thus it was decided against for the development of the promoter-operator control region. Instead, the P_{smyc} promoter, isolated from *Mycobacterium smegmatis* by Ehrt *et al*., 2005[53] was selected to engineer in the *tet* operator sites as it was an even stronger promoter than *hsp*60 capable of potent expression, and equally strong in wild type BCG Pasteur. Furthermore, the insertion of tandem *tet* operator 2 (*tet*O2) sequences flanking the –35 region, resulted in only a minimal decrease in promoter activity[53].

The isolated P_{smyc} promoter region (189bp) was found to be identical (100% sequence identity) to the upstream region of the putative *rspA* gene of *M. smegmatis* encoding the putative 30S ribosomal protein S1[53]. In order to first assess the importance of the extensive region between the reported transcriptional start site and the theoretical translation start site (125bp) on GFP expression levels, we designed and synthesised two promoter versions, a full-length 190bp version (NM1) and a truncated 90bp version (NM2). The design and synthesis of these promoter fragments are outlined in Section 2.8 and are described in Table 2.1. Both NM1 and NM2 were cloned into the selected expression vector, pRC100 in a directional manner upstream of the GFP reporter gene with restriction endonucleases *Xba*1 and *EcoR*1 thereby replacing the *hsp*60 promoter and generating the shuttle vectors pNM1 and pNM2, respectively. See Figure 4.3 for an illustration of the cloning strategy.

GFP from the jellyfish Aequorea Victoria is a frequently used reporter protein for monitoring gene expression in a variety of cells and organisms including *M. smegmatis* and BCG[58]. We have chosen to use GFP in this study as a simple and effective means to screen for the different strengths of the various promoter elements generated here that may have potential for future application in mycobacteria. The relative promoter strengths of the full-length and truncated P_{smyc} promoter contained in pNM1 and pNM2, respectively, were compared to a positive control, the parental plasmid pRC100 in which the GFP gene is expressed by the *hsp*60 promoter in conjunction with a strong ribosomal binding site (RBSD). The plasmid, pWB100, which only consisted of vector backbone and expressed no GFP served as a negative control. These four plasmids; pWB100, pRC100, PNM1 and pNM2 (Figure 4.1) were subsequently electroporated into *M. smegmatis,* BCG Pasteur and BCG panCD and promoter strengths compared by visualising GFP fluorescence intensities of the resulting colonies over a UV trans-illuminator.

The results for *M. smegmatis* and BCG Pasteur are presented in Figure 4.2, however a similar pattern of fluorescence was observed in all three strains. As expected, the negative control, r*M. smegmatis* [pWB100], showed no fluorescence while the positive control r*M. smegmatis* [pRC100], displayed strong GFP expression. However, r*M. smegmatis* [pNM1] and r*M. smegmatis* [pNM2] fluoresced more intensely than the positive control, confirming reports that the P_{smyc} promoter is stronger than *hsp*60. Although it may not be clear from the saturated images (Figure 4.2), when viewing the plates directly over UV light, it was consistently observed that the full-length P_{smyc} promoter (Fig 4.2 iii) was slightly stronger than the truncated promoter (Fig 4.2 iv) in all strains. Thus, all subsequent promoter-operator sequences were designed based on the full-length promoter region to attain maximum expression levels indicative of the fully induced state.



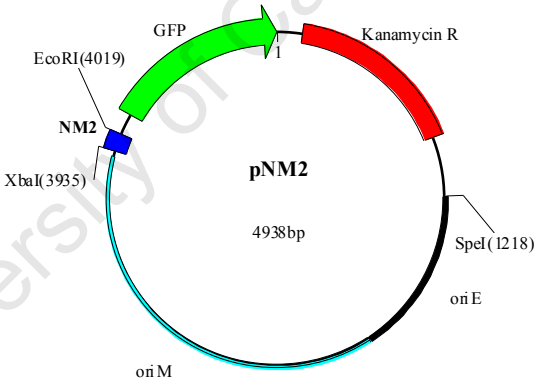


Figure 4.1. ***E. coli*** **–mycobacterial shuttle plasmids constructed for the selection of a strong mycobacterial promoter.** pNM1 and pNM2 were derived from pRC100 and contain either the full-length natural P_{smyc} promoter (NM1) or the truncated promoter (NM2), respectively.

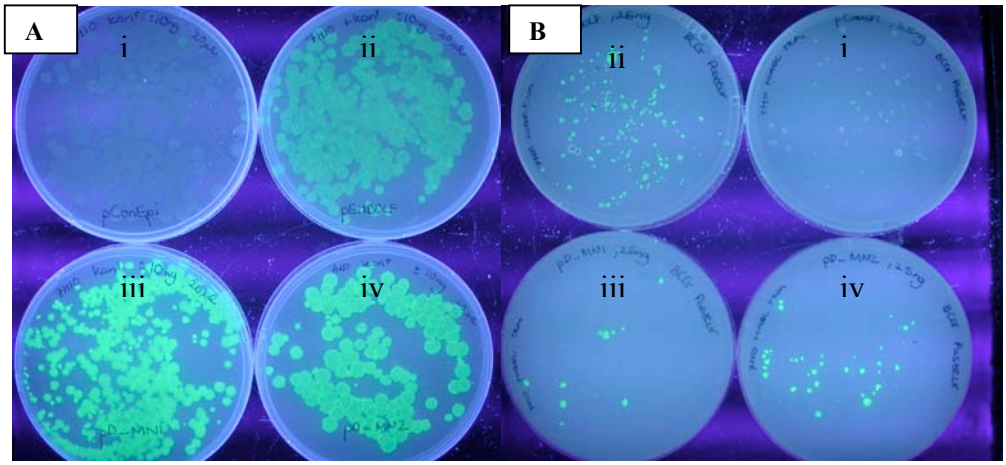


Figure 4.2. Fluorescent images of *M. smegmatis* and BCG Pasteur transformants harbouring the promoter test vectors. *M. smegmatis* (Panel A) or BCG Pasteur (Panel B) cells were transformed with the promoter test vectors: pWB100 (i), pRC100 (ii), pNM1 (iii) and pNM2 (iv) and plated on the appropriate selective media and incubated at 37°C for 3 days. Images were captured in the dark over a UV transilluminator.

4.2.2. Construction of Phase 1 promoter-operator constructs

The sequence, positioning and number of *tet* operator sites (*tet*O) inserted within a promoter can significantly affect the strength of a promoter and thus the efficiency of regulation of the inducible expression system[72]. Based on the promoter-operator design reported by Ehrt *et al*., 2005[53], three sequence elements were designed which contained the full-length P_{smyc} promoter sequence with either duplicate *tet* operator 1 or operator 2 sequences (*tet*O1 or *tet*O2) or a combination of them, inserted on either side of the –35 promoter position. Table 2.1 contains the annotated sequence of the three promoter-operator designs namely, NM3, NM4 and NM5. These fragments were synthesised by GeneArt® and provided in a plasmid containing an ampicillin resistance marker.

The three promoter-operator sequence elements (NM3-5) were excised out of the GeneArt® backbone and cloned into pRC100 in the same way as for NM1 and NM2, replacing the existing *hsp*60 promoter and generating plasmids pNM3-5. Removal of the promoter region from pRC100 generated a negative control, pNM6. Details of the cloning procedure are described in Section 2.9.1. An illustration of the generic cloning strategy employed to generate these Phase one constructs is presented in Figure 4.3, while Figure 4.4 represents the resulting plasmid maps for the constructs, pNM3-6.

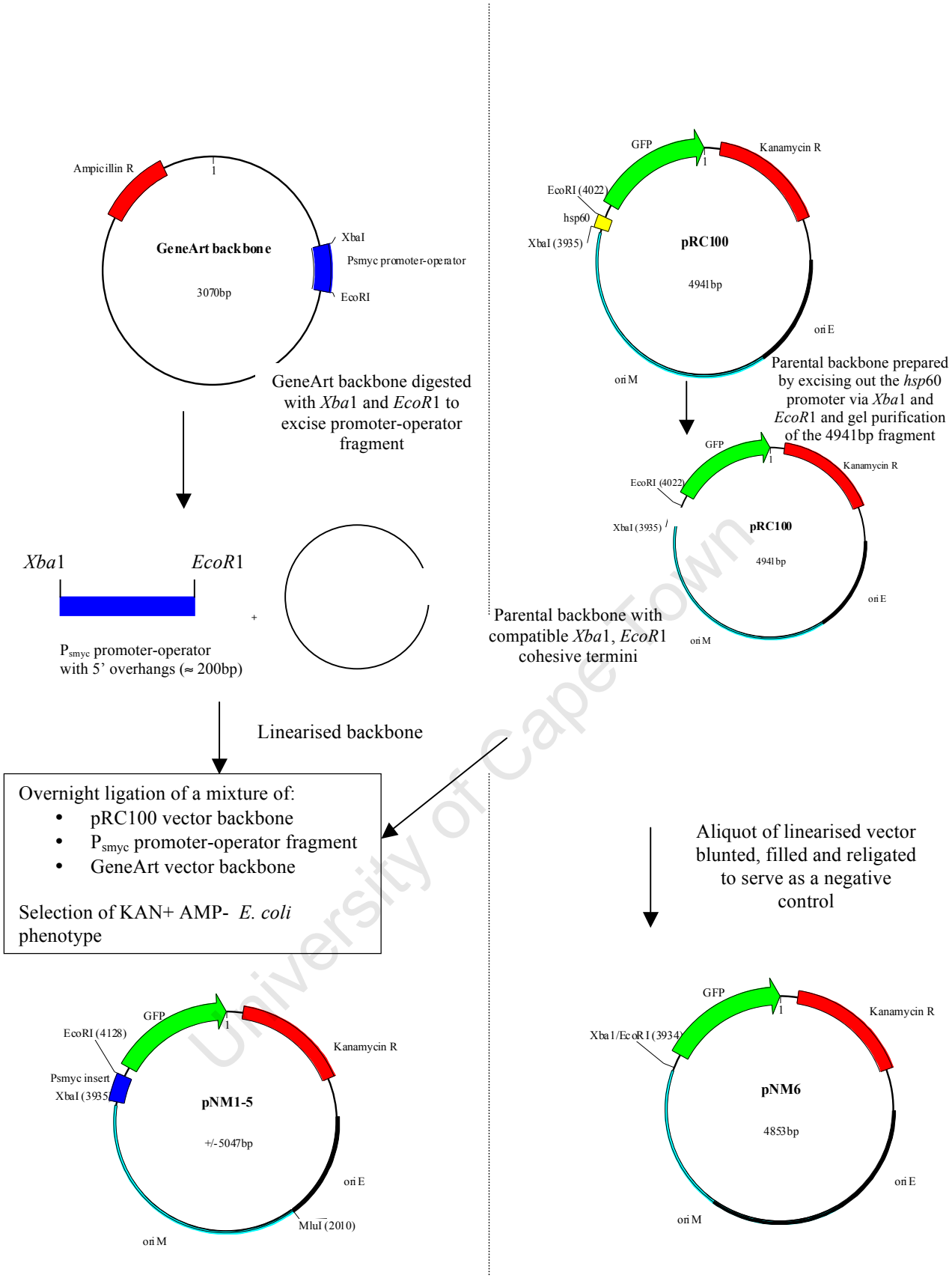
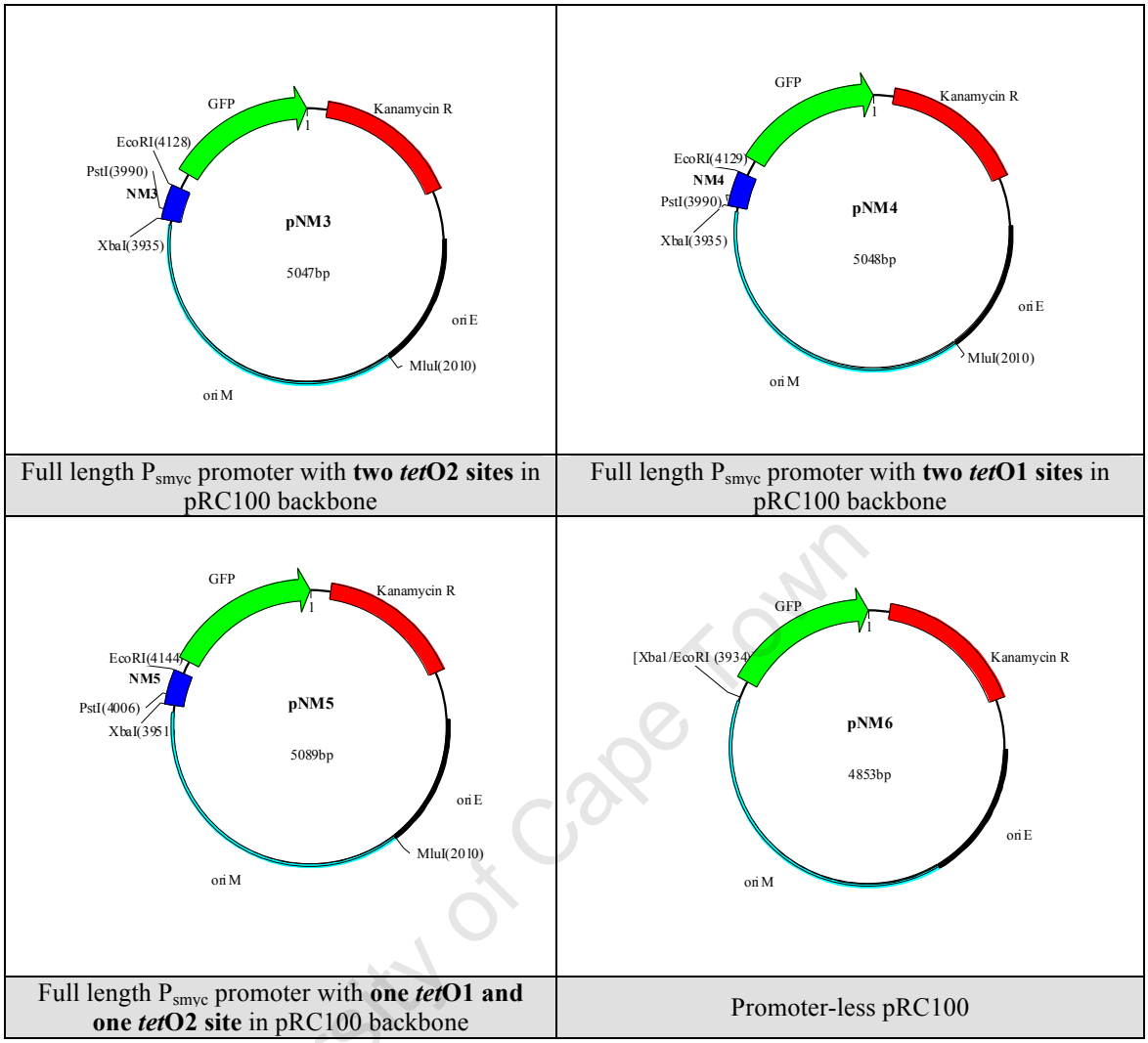


Figure 4.3. Illustration of the general cloning strategy for the generation of Phase 1 constructs



Full length P_{smyc} promoter with **two *tet*O2 sites** in pRC100 backbone	Full length P_{smyc} promoter with **two *tet*O1 sites** in pRC100 backbone
Full length P_{smyc} promoter with **one *tet*O1 and one *tet*O2 site** in pRC100 backbone	Promoter-less pRC100

Figure 4.4. Resulting Phase 1 promoter-operator shuttle vector constructs and promoter-less control.

4.2.3. Confirmation of Phase 1 plasmid constructs by restriction endonuclease mapping

Restriction enzyme mapping provides a convenient and quick method to for screen for potentially correct recombinants. After an initial screening of miniprep DNA isolated from *E. coli* clones transformed with the various Phase 1 construct ligation mixes, selected clones were maxiprepped and a combination of restriction enzyme digests performed to map the plasmid DNA and confirm constructs pNM1-6. A *Pst*1 digest provided a convenient reference for the presence of the P_{smyc} promoter-operator region as it represents a unique site within this region while *Mlu*1 represents a unique site within the parental backbone, pRC100. In addition, an *Mlu*1+ *Xba*1

double digest was performed to assess base pair size differences between the constructs. Expected band sizes and the resulting gel photos are presented in Table 4.1 and Figure 4.5, respectively.

Table 4.1. Expected fragment sizes (bp) resulting from restriction endonuclease mapping of the phase 1 plasmid constructs

	Restriction Endonucleases		
Plasmid	***Pst*1**	***Mlu*1**	***Mlu*1, *Xba*1**
pWB100	Uncut	4073bp	1925,2148
pRC100	Uncut	4941bp	1925,3016
pNM1	5042bp	5042bp	1925,3117
pNM2	4938bp	4938bp	1925,3013
pNM3	5047bp	5047bp	1925,3122
pNM4	5048bp	5048bp	1925,3123
pNM5	5047bp	5047bp	1925,3122
pNM6	Uncut	4858bp	1925,2933

Electrophoresis of the single enzyme digests resulted in the expected results. The *Pst*1 digests of pNM1-5 (Figure 4.5 A, lanes 5-9), confirmed the presence of the P_{smyc} promoter-operator element within the recombinant plasmids, as it linearised the plasmids (lane 7 however, displays a partial digestion of the plasmid DNA). The plasmids pWB100, pRC100 and pNM6 contained no P_{smyc} promoter-operator sequences thus no *Pst*1 sites, hence remained uncut as expected. The *Mlu*1 digests linearised all plasmids confirming the presence of the pRC100 backbone (Figure 4.5 B). A notable size difference can be seen between pWB100 (lane 2) and pRC100 (lane 3) that was derived from pWB100. This is attributed to the presence of the 868bp GFP gene under the control of an *hsp*60 promoter in pRC100.

Figure 4.6 illustrates that the double digestion of the Phase 1 constructs with *Xba*1 and *Mlu*1 resulted in fragment sizes as predicted in Table 4.1. A common 1925bp band for pWB100, pRC100 and pNM1-6 and a larger approximately 3Kb band was generated, as expected. It is not possible to distinguish the small size differences of the larger band between the constructs as they only differ by a few base pairs for the most part. The size difference between the larger band of pWB100 (lane 4), the empty vector, and all other plasmids is however evident. These results provided

additional confidence in the constitution of the constructs as *Xba*1 was one of the sites used for cloning and should be re-generated during the ligation process.

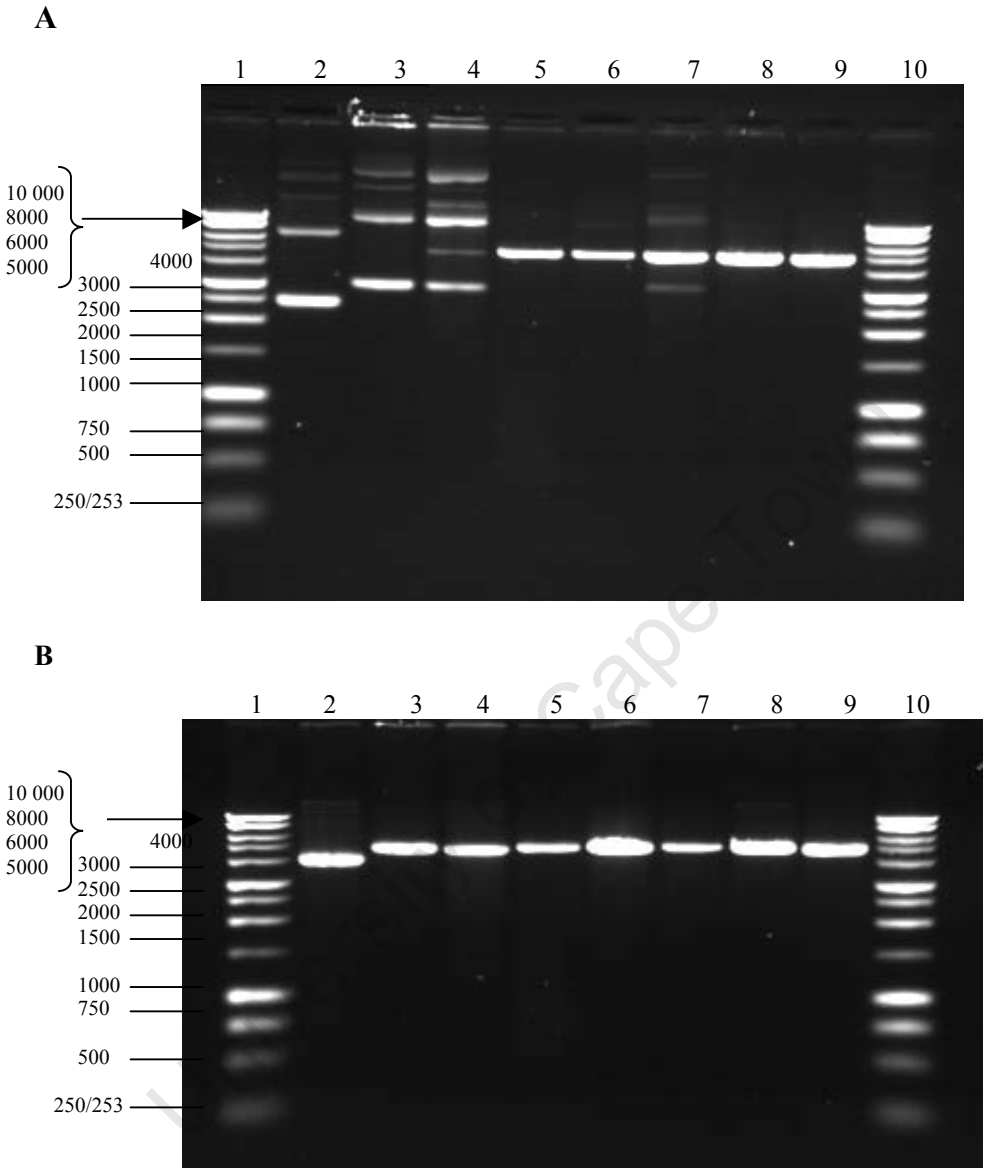


Figure 4.5. Gel electrophoresis of single restriction enzyme digests for Phase 1 plasmids. Phase 1 plasmids were either digested with *Pst*1 (A) or *Mlu*1 (B). Lane 1 and 10, 1Kb DNA ladder; Lane 2, pWB100; Lane 3, pRC100; Lane 4, pNM6; Lane 5, pNM3; Lane 6, pNM4; Lane 7, pNM5; Lane 8, pNM1; Lane 9, pNM2. Lanes 2-9 represents 500ng of digested plasmid DNA.

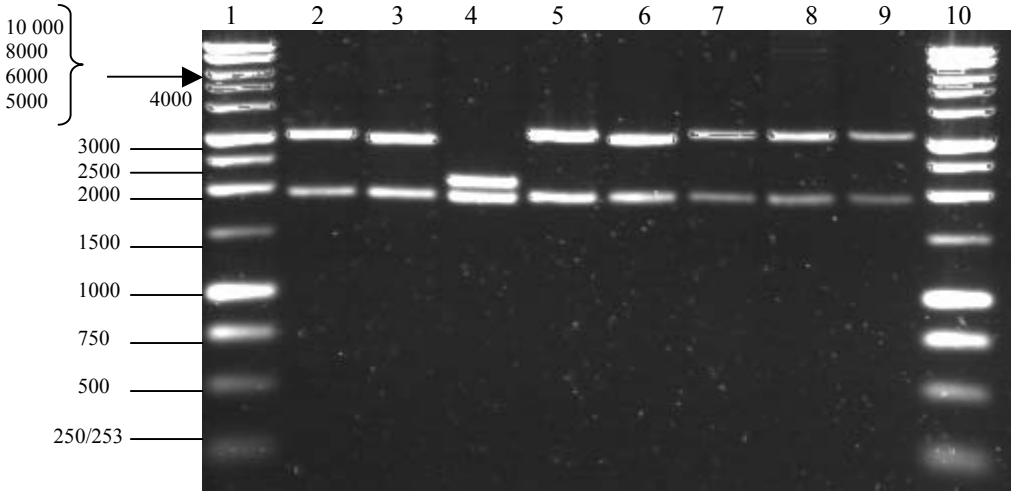


Figure 4.6. Gel electrophoresis of Phase 1 constructs simultaneously digested with *Xba*1 and *Mlu*1. Lane 1 and 10, 1 Kb DNA ladder; Lane 2, pRC100; Lane3, pNM6; Lane 4, pWB100; Lane 5, pNM1; Lane 6, pNM2; Lane 7, pNM3; lane 8, pNM4; Lane 9, pNM5.

4.2.4. Confirmation of Phase 1 constructs by sequencing

Sequencing reactions were performed for all Phase 1 constructs (pNM1-6) using a single reverse primer (GFP_R) that binds to the GFP gene at nucleotide positions 81-58, yielding a single strand of sequence. This allowed for the confirmation of the correct pRC100- P_{smyc} promoter-operator junctions and the re-creation of both *Xba*1 and *EcoR*1 sites, or the removal of *hsp*60 in the case of pNM6. Additionally, the sequencing reactions allowed for the verification of the presence of just a single insert within the vectors.

4.2.5. Selection of promoter-operator constructs that support strong GFP expression

To assess the effect of the different promoter-operator designs on GFP expression, and to select promoter-operator constructs that drive strong GFP expression, the phase 1 constructs were electroporated concurrently into both BCG strains (BCG Pasteur and BCG panCD) as well as the model organism *M. smegmatis*. Visualising the plated transformants over a UV trans-illuminator allowed for initial assessment of GFP expression levels relative to the controls. GFP capture ELISAs were performed subsequently for a more accurate comparison for the selection of those promoter-operator constructs that supported strong GFP expression to be further developed in Phase 2.

4.2.5.1 Comparison of GFP expression by visualising fluorescence over a UV trans-illuminator

GFP fluorescence was compared by visualisation of recombinant mycobacterial colonies over a UV trans-illuminator. This provided a preliminary guideline of relative expression levels between the various Phase 1 constructs. Photos of the r*M. smegmatis* plates can be seen in Figure 4.7 and photos of rBCG Pasteur presented in Figure 4.8. The results of this experiment are summarised in Table 4.2. An arbitrary scale was used to record relative fluorescence intensities where a value of 1 denotes very weak expression and 5 denotes maximum GFP expression.

Figure 4.7 panel A, shows that as expected *M. smegmatis* transformed with the positive control pRC100 displayed relatively good GFP fluorescence (panel A, i) whereas the negative control r*M. smegmatis*[pWB100] did not fluoresce at all (data not shown). However, both r*M. smegmatis*[pNM3] and r*M. smegmatis*[pNM4] fluoresced more brightly than the positive control r*M. smegmatis*[pRC100], with *M. smegmatis*[pNM4] displaying the highest level of fluorescence. r*M. smegmatis*[pNM5], on the other hand fluoresced visibly less than the positive control. Figure 4.7 panel B, further highlights that the brightest *M. smegmatis* recombinant resulted from the transformation with pNM1, the full-length natural P_{smyc} promoter without any operator sites, however there was not a big difference in the fluorescent intensities between r*M. smegmatis*[pNM1] and the slightly weaker r*M. smegmatis*[pNM4] and r*M. smegmatis*[pNM3], both also containing the full-length P_{smyc} promoter but with operator site inserted.

Similar trends of relative fluorescence intensities were also observed for recombinant BCG Pasteur (Figure 4.8) and recombinant BCG panCD transformants harbouring the same Phase 1 constructs. Results for BCG panCD were not presented here due to poor transformation efficiency. Colony sizes were noted and found to be inversely correlated to GFP expression levels, suggesting that the high level of GFP expression imposes a toxic or metabolic load on the bacteria.

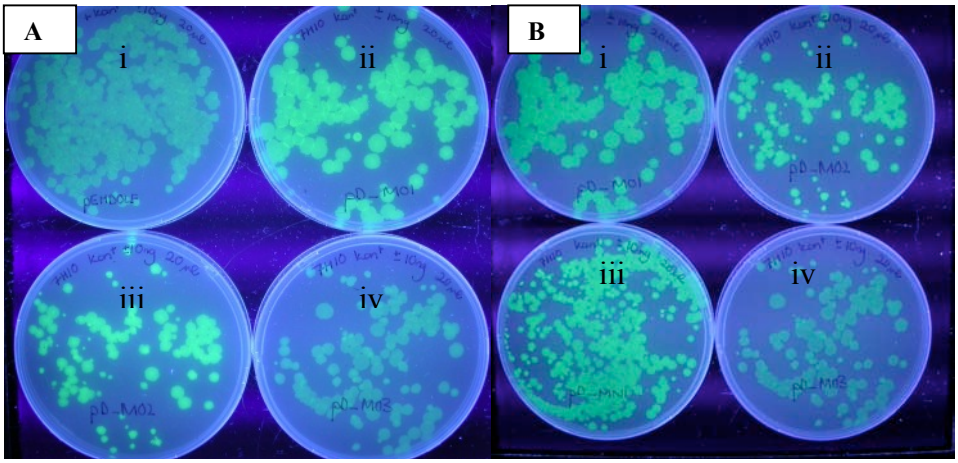


Figure 4.7. Photographic images of fluorescent *M. smegmatis* transformants harbouring the Phase 1 plasmids. *M. smegmatis* cells were electrotransformed with the various Phase 1 constructs and plated on appropriate selective media and visualised over UV following incubation at 37°C for 3 days. Panel A represents *M. smegmatis* harbouring the plasmids: pRC100 (i), pNM3 (ii), pNM4 (iii) and pNM5 (iv). Panel B represents *M. smegmatis* harbouring the plasmids: pNM3 (i), pNM4 (ii), pNM1 (iii) and pNM5 (iv)**.**

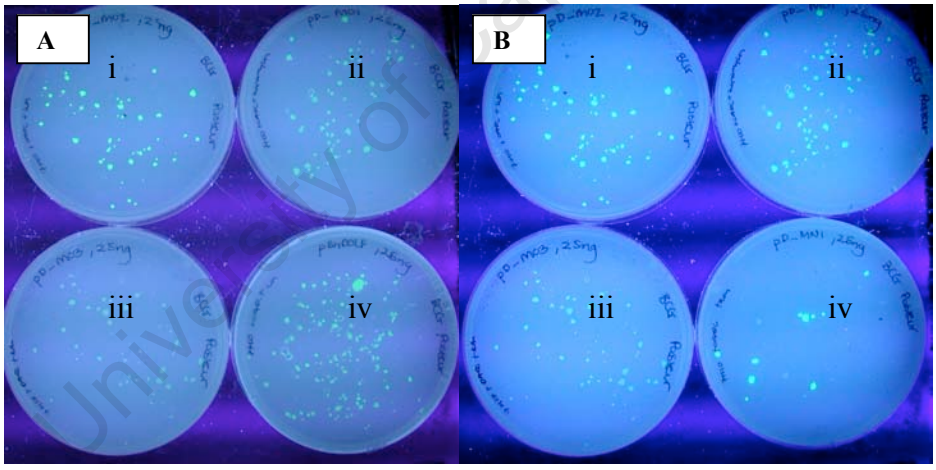


Figure 4.8. Photographic images of fluorescent wild type *M. bovis* BCG Pasteur transformants harbouring the Phase 1 plasmids. BCG Pasteur cells were electrotransformed with the various Phase 1constructs and plated on appropriate selective media and visualised over UV following incubation at 37°C for 3 days. Panel A represents BCG Pasteur harbouring the plasmids: pNM4 (i), pNM3 (ii), pNM5 (iii), pRC100 (iv). Panel B represents BCG Pasteur harbouring the plasmids: pNM4 (i), pNM3 (ii), pNM5 (iii), pNM1 (iv).

Table 4.2. Summary of the comparitive fluorescence trends observed for recombinant mycobacterial colonies harbouring the Phase 1 constructs. Relative fluorescence intensities were recorded as a score out of 5, with 5 representing the highest level of GFP expression observed.

Plasmid Name	Promoter Features	Relative Fluorescence in *M. smegmatis*, BCG Pasteur and BCG panCD
pRC100	*hsp*60 driven GFP	3.5
pWB100	Empty vector backbone	0
pNM6	Promoter-less pRC100	0
pNM1	Full-length (190bp) P_{smyc} promoter, no *tet* operator sites	5
pNM2	Truncated (90bp) P_{smyc} promoter, no *tet* operator sites	4.5
pNM3	Full-length P_{smyc} promoter with two *tet*O2 sequences situated on either side of -35	4.5
pNM4	Full-length P_{smyc} promoter with two *tet*O1 sequences situated on either side of -35	4.75
pNM5	Full-length P_{smyc} promoter with one *tet*O1 upstream of –35 and one *tet*O2 sequence downstream of –35.	3

4.2.5.2. Comparison of GFP expression levels in *M. smegmatis* by quantitative GFP -capture Enzyme-linked Immunoabsorbant Assay (ELISA).

The levels of GFP expressed by the various Phase 1 *M. smegmatis* recombinants were quantitated using polyclonal anti-GFP pre-coated plates to capture GFP present in bacterial lysate, and a polyclonal anti-GFP HRP-conjugated antibody for detection. Since this particular ELISA procedure had not been previously established in the laboratory, several optimisation experiments were performed in order to ascertain the appropriate sample and secondary antibody dilutions to be used. Furthermore, since the presence of a milieu of mycobacterial proteins affected the sensitivity of the ELISA assay, all rGFP standards were diluted in lysate prepared from *rM. smegmatis*[pWB100] culture such that there was an equivalent amount of mycobacterial protein lysate present in both the experimental samples and the rGFP standards in the assay. These optimisations along with the general procedure are outlined in Section 2.12.2.

This assay was used to quantify GFP expression in bacterial lysates prepared from r*M. smegmatis* cultures harbouring the Phase 1 constructs grown to an early logarithmic growth phase with OD_{600} readings of approximately 0.4-0.5 OD units. A summary of the results is presented in Figure 4.9 in a bar graph format as a percentage of total protein expression. As expected, sample *rM. smegmatis*[pWB100] (empty vector backbone) showed 0% GFP expression, however the promoter-less negative control sample, *rM. smegmatis*[pNM6], did in fact show minimal GFP expression of 0.1% . This suggests that some read-through is occurring from an upstream promoter-like region. The highest level of GFP expression measured was 1.42% obtained from cultures of *rM. smegmatis*[pNM1] in which GFP is expressed from the full-length natural P_{smyc} promoter. This was followed closely by *rM. smegmatis*[pNM2] (1.36%), *rM. smegmatis*[pNM3] (1.33%) and *rM. smegmatis*[pNM4] (1.25%), whereas *rM. smegmatis*[pNM5] produced significantly lower GFP(0.38%). These trends represent those seen previously on plates except that on the plates *rM. smegmatis*[pNM4] colonies appeared to produce more GFP (higher fluorescence intensity) than *rM. smegmatis*[pNM3].

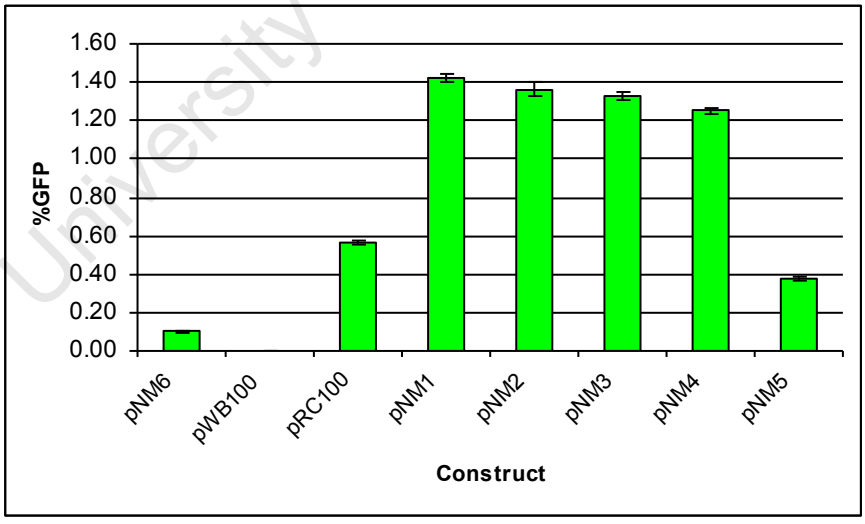


Figure 4.9. Comparison of GFP expression between r*M. smegmatis* cultures containing phase 1 constructs presented as a percentage of total protein produced. This bar graph represents a summary of the Phase 1 ELISA experiments. The amount of GFP produced was calculated by dividing the total amount of GFP expressed (ng) by the amount of total protein (ng), expressed as a percentage. The experiment was performed in triplicate and error bars are as included.

4.3. Development of Phase 2 of the revTetR regulatable expression system

4.3.1. Introduction

Having successfully constructed two shuttle vectors in which our model antigen GFP is strongly expressed by the P_{smyc} promoter engineered to contain two *tet*O2 sites (pNM3) and two *tet*O1 sites (pNM4) in Phase 1, Phase 2 involved the insertion of the revTetR expression cassette into these Phase 1 plasmids and the selection of promising gene arrangements for regulatory studies. Since pNM5 displayed such poor GFP expression it was not explored for further development.

In order to construct a revTetR regulatable system that could efficiently repress GFP expression in the presence of ATc and induce GFP expression upon its removal, we chose to base our system design on the revTet regulatory protein TetR r1.7[84]. The reverse TetR mutant, TetR r1.7, was selected on the basis of its highly efficient reverse regulation in *E. coli*, regulation that was found to be equal to that of the wild type TetR[84]. It contains three amino acid changes within the alpha one helix of the DNA binding domain namely, E15A L17G and L25V[84] and is the same mutant TetR that was used by Kamionka *et al*., 2005[102] to generate tetracycline dependent conditional gene knockouts in *B. subtilis* and the same mutant employed by Guo *et al*., 2007[82] to silence essential gene expression in *M. smegmatis*.

In our study design, a TetR r1.7 expression cassette was designed such that TetR r1.7 was placed under control of the *M. tuberculosis hsp*60 promoter for strong constitutive expression, a very strong ribososme binding site (RBSD) for maximum translation initiation and bordered by the *hsp*60 transcriptional terminator for efficient termination (see Section 2.8). In addition, to optimise expression and stability of TetR r1.7, the gene was codon optimised for use in mycobacteria (Dr. Helen Stutz, personal communication). This expression cassette was introduced into pNM3 and pNM4 for the development of a complete episomal multicopy revTetR system with the aim of continuously producing sufficient TetR r1.7 to efficiently repress GFP expression in the presence of ATc during *in vitro* culture.

4.3.2. Construction of the Phase 2 plasmids

The 821bp TetR r1.7 expression cassette was synthesised by GeneArt®, provided in a backbone bearing an ampicillin resistance marker and cloned into the Phase 1 plasmids pNM3 and pNM4 immediately upstream of the promoter-operator driven GFP. To achieve this, the TetR r1.7 cassette was excised out of the manufacturer's backbone by digestion with the flanking restriction enzyme sites *Xba*1 and *Spe*1 and the cassette gel purified. This generated compatible ends for cloning into the two Phase 1 vector backbones i.e. pNM3 and pNM4, that were simultaneously linearised with the restriction enzyme *Xba*1 and the products gel purified and dephosphorylated. Ligation reaction mixes were subsequently transformed into *E. coli* cells, with the 1:3 vector:insert ratio yielding the best results with respect to the number of recombinant colonies obtained. It should be noted that the experimental design purposely allowed for the insertion of the TetR r1.7 cassette in either orientation; orientation 1 in which the *tetR* r1.7 gene was present in the same orientation as GFP, or orientation 2 in which the *tetR* r1.7 gene and GFP gene are expressed divergently. The general cloning strategy for this phase is illustrated in Figure 4.10. Both orientations were screened and selected for as orientation may well affect the expression of both TetR r1.7 and GFP, which are closely situated, approximately 500bp apart on the plasmid.

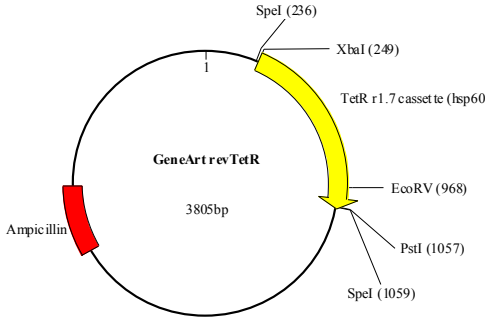
4.3.3.Confirmation of the Phase 2 constructs by restriction endonuclease mapping

Restriction enzyme mapping was used to screen for potentially correct recombinants. After an initial screening of miniprep DNA isolated from *E. coli* clones transformed with the various Phase 2 construct ligation mixes, selected clones were maxiprepped and a combination of restriction enzyme digests performed to map the plasmid DNA and confirm constructs pNM3.1 (pNM3 harbouring the TetR r1.7 cassette in orientation 1), pNM3.2 (pNM3 harbouring the TetR r1.7 cassette in orientation 2), pNM4.1 (pNM4 harbouring the TetR r1.7 cassette in orientation 1) and pNM4.2 (pNM4 harbouring the TetR r1.7 cassette in orientation 2). The predicted fragment sizes and resulting gel photos following electrophoresis are presented in Table 4.3 and Figures 4.12 and 4.13 respectively.

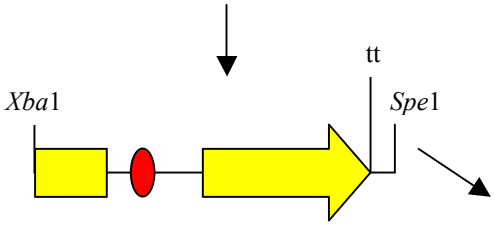
A single *EcoR*V digest was performed to verify the presence of the TetR r1.7 cassette as it represents a unique site within the cassette. Gel electrophoresis of the *EcoR*V restriction enzyme digests resulted in linear bands for all four of the Phase 2 constructs (pNM3.1, pNM3.2, pNM4.1 and pNM4.2) confirming the presence of the cassette (Lanes 3 and 4 of both Figures 4.12 and 4.13). As predicted, the digestion of the Phase 1 plamsids, pNM3 and pNM4 resulted in uncut plasmid (Lane 2 of both Figures 4.12 and 4.13).

An *Mlu*1 + *Not*1 double digest was performed to map the increase in size from Phase 1 vectors, pNM3 and pNM4 to new recombinant Phase 2 plasmids, pNM3.1 and pNM3.2 or pNM4.1 and pNM4.2 resulting from the insertion of the TetR r1.7 cassette. An *Xba*1+ *Nde*1 double digest was performed to ascertain the orientation of the cassette. These digests generated a common 1901bp smaller band derived from the vector backbone in all the Phase 1 and Phase 2 constructs, and a larger band that varied in size as predicted (Lanes 5-7 of both Fig 4.12 and 4.13).

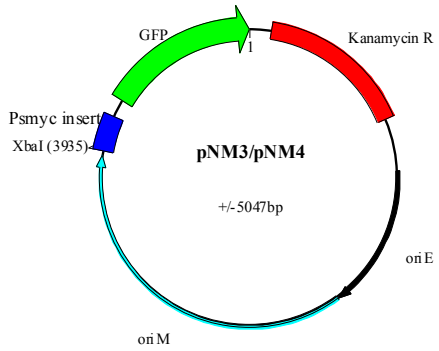
An *Xba*1 and *Nde*1 double digest was performed to determine the orientation of insertion of the cassette. Since the *Xba*1 site was not reconstituted upon ligation of the *Xba*1+*Spe*1 5' overhang sites; the constructs containing the TetR r1.7 cassette in orientation 1 (pNM3.1 and pNM4.1) resulted in a small approximately 500bp band and a larger approximately 5.3Kb band, as predicted (Lane 10 of both Figures 4.12 and 4.13). The constructs containing the TetR r1.7 cassette in orientation 2 (pNM3.2 and pNM4.2) however resulted in a smaller band of approximately 1.2Kb and a larger approximately 4.6Kb band, as predicted (Lane 9 of both Figures 4.12 and 4.13). In order to visualise the small ≈ 500bp band in lanes 8 (phase 1 controls) and 10 of both Figures 4.12 and 4.13, a large amount of DNA (1μg) was required for digestion thus resulting in the distorted overloaded larger band, making it difficult to accurately distinguish size differences.



The TetR r1.7 cassette was excised out of the GeneArt backbone using the restriction enzymes, *Xba*1 and *Spe*1 and the resulting 816bp product was subsequently gel purified



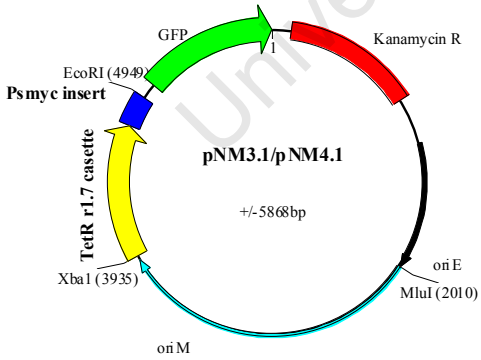
TetR r1.7 gene under the control of an *hsp*60 promoter with RBSD and transcriptional terminator



Phase 1 Shuttle vectors (pNM3 and pNM4) were linearised with *Xba*1 and the resulting 4853bp backbone gel purified, dephosphorylated and reaction purified.

Overnight ligation at vector:insert ratios of 1:1, 1:3, 1:5

Orientation 1



Orientation 2

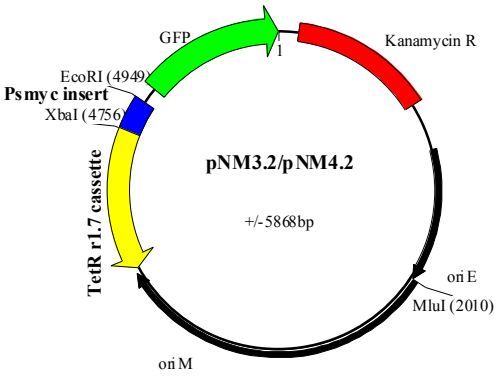


Figure 4.10. Illustration of generic cloning strategy employed to generate Phase 2 shuttle vectors

Table 4.3. Expected fragment sizes (bp) resulting from restriction endonuclease digestion of the Phase 2 plasmid constructs

Plasmid	Restriction Endonuclease digestions		
	***EcoR*V**	***Mlu*1 + *Not*1**	***Xba*1 +*Nde*1**
pNM3	uncut	3466bp, 1901bp	4561bp, 511bp
pNM3.1	5857bp	3956bp, 1901bp	5346bp, 511bp
pNM3.2	5857bp	3956bp, 1901bp	4597bp, 1260bp
pNM4	uncut	3447bp, 1901bp	4562bp, 511bp
pNM4.1	5858bp	3957bp, 1901bp	5346bp, 512bp
pNM4.2	5858bp	3957bp, 1901bp	4597bp, 1261bp

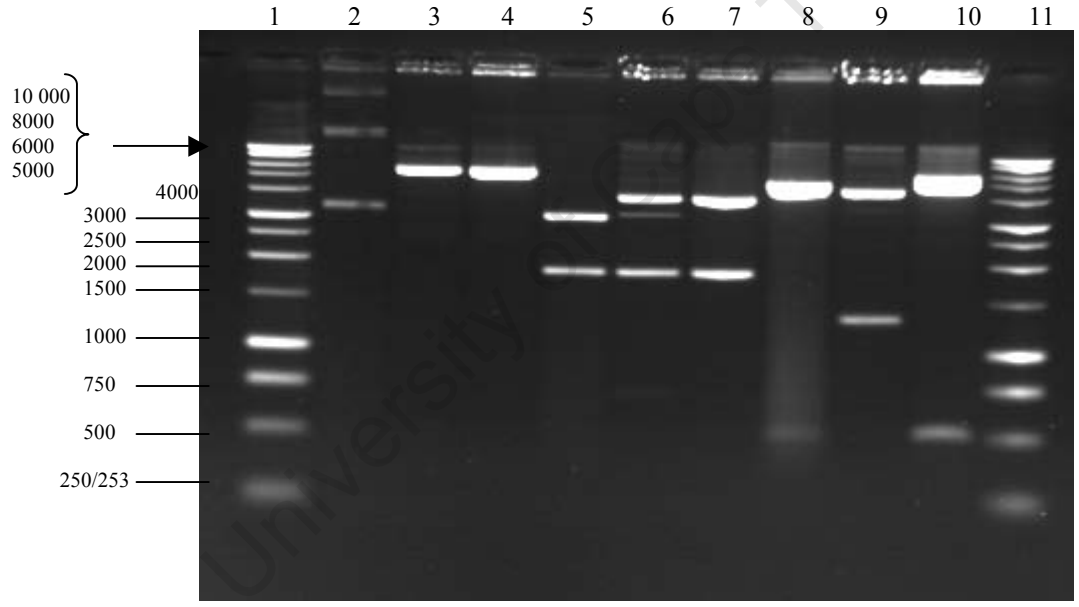


Figure 4.11. Gel photo of electrophoresed restriction enzyme digests for pNM3, pNM3.1 and pNM3.2. Lanes 1 and 11 were loaded with a 1Kb DNA ladder. Lanes 2-4, *EcoR*V digests of 250ng of pNM3, pNM3.2 and pNM3.1, respectively. Lanes 5-7, *Mlu*1+*Not*1 digest of 250ng of pNM3, pNM3.2 and pNM3.1, respectively. Lanes 8-10, Xba1 + *Nde*1 digest of pNM3 (1μg), pNM3.2 (250ng), and pNM3.1 (1μg) respectively.

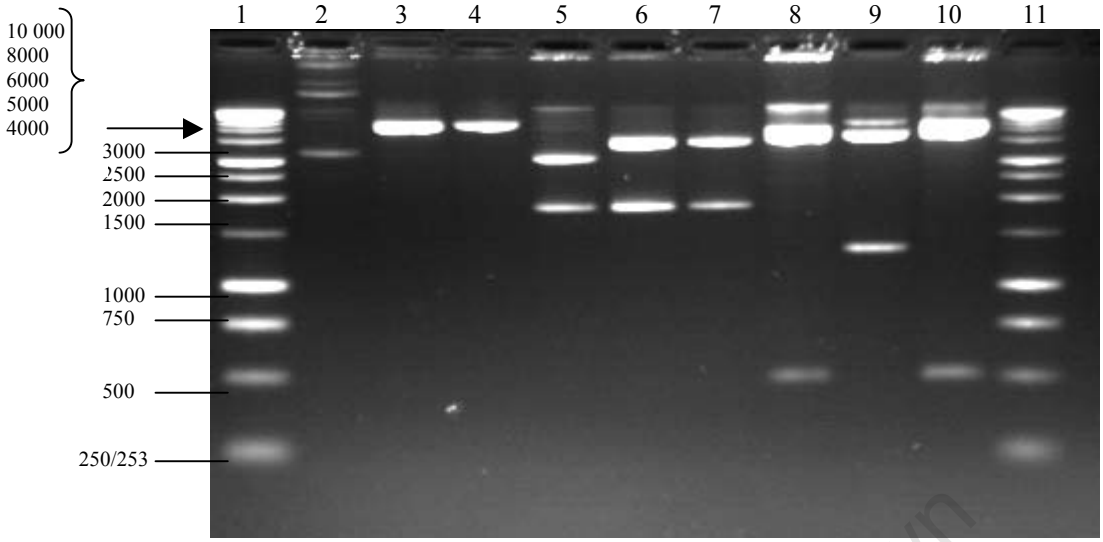


Figure 4.12. Gel photo of electrophoresed restriction enzyme digests for pNM4, pNM4.1 and pNM4.2. Lanes 1 and 11 were loaded with a 1Kb DNA ladder. Lanes 2-4, *EcoR*V digests of 250ng of pNM4, pNM4.2 and pNM4.1, respectively. Lanes 5-7, *Mlu*1+*Not*1 digest of 250ng of pNM4, pNM4.2 and pNM4.1, respectively. Lanes 8-10, *Xba*1 + *Nde*1 digest of pNM4 (1µg), pNM4.2 (250ng), and pNM4.1 (1µg) respectively.

4.3.4. Confirmation of Phase 2 constructs by sequencing

Sequencing reactions were performed firstly, in order to confirm the cloning junctions (pRC100/TetR r1.7 cassette junction and the TetR r1.7 cassette/ P_{smyc} promoter-operator junction) that would also confirm the orientation of the TetRr1.7 cassette and secondly, to confirm the integrity of the TetR r1.7 gene cassette supplied by GeneArt®. In order to confirm the cloning junctions of the TetR r1.7 cassette, sequencing reactions were performed with primers 26F (forward), which binds to the OriM region of pNM3 and pNM4 and 07-0320R (reverse), which binds a common region within the P_{smyc} promoter of both pNM3, and pNM4. This was done for all the Phase 2 plasmids and confirmed that for the orientation 1 constructs, pNM3.1 and pNM4.1; the *Xba*1 site was reconstituted adjacent to the P_{smyc} promoter-operator sequence whilst the remnant of the *Xba*1-*Spe*1 religated site occurred proximal to the OriM vector sequence and vice versa for the orientation 2 constructs i.e. pNM3.2 and pNM4.2.

A single plasmid, pNM3.1 was chosen to fully sequence across the TetR r1.7 cassette region to ensure the entire cassette was intact. A series of four forward and four

reverse primers were designed for this purpose (See Table 2.1 for list of primers used), spanning the entire 821bp cassette. Sequencing output showed 100% sequence identity to the sequence designed and requested from the manufacturer, which verified the presence of an intact TetR r1.7 cassette. The entire annotated sequence of pNM3.1 is shown in Appendix A. The confirmed plasmid maps are illustrated in Figure 4.13.

4.3.5. Selection of Phase 2 constructs that support strong GFP expression

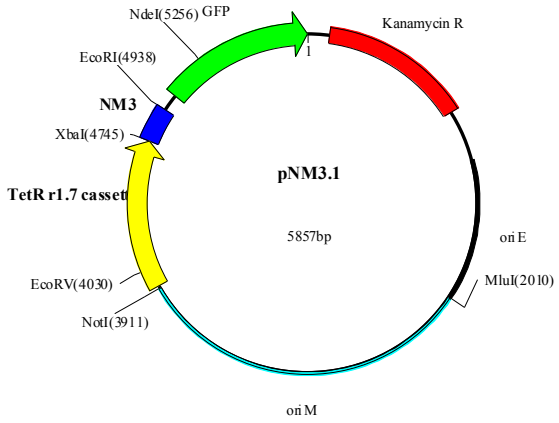
The presence of TetR r1.7 constitutively expressed at elevated levels by the *hsp*60 promoter and it's orientation in the shuttle-vector with respect to GFP may well influence the level of GFP expression. Thus, promising Phase 2 constructs were selected on the basis of continued strong GFP expression.

4.3.5.1. Comparison of GFP expression by visualising fluorescence over a UV trans-illuminator

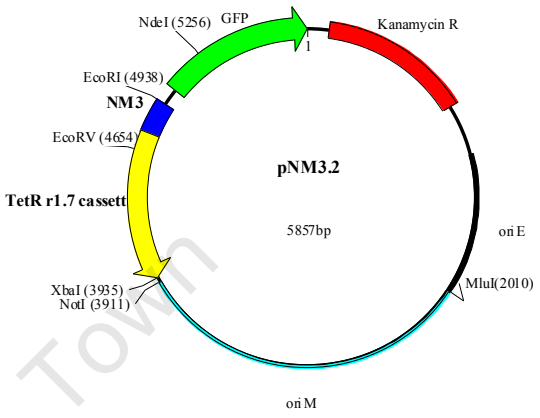
M. smegmatis and BCG panCD cells were electrotransformed with the control plasmid DNA; pWB100 and pRC100, the selected Phase 1 plasmid DNA; pNM3 and pNM4 and the Phase 2 plasmid DNA; pNM3.1, pNM3.2, pNM4.1 and pNM4.2, and the GFP fluorescent intensities of the resulting transformants viewed over a UV trans-illuminator. The images of recombinant *M. smegmatis* colonies on plates are presented in Figure 4.15 in Panels A-C. It should be noted however, that due to saturation of the images presented, small differences in the fluorescent intensities might not be as apparent as in direct observation of the plates. Hence, the results for this experiment are summarised in Table 4.4.

A number of salient observations were made. While all the recombinant colonies plated were fluorescent, colonies containing the TetR r1.7 expression cassette showed a noticeable reduction in GFP fluorescence and were slightly smaller in comparison to their Phase 1 counterparts. However, consistent with observations for Phase 1 plated recombinants, colonies containing pNM4-derived constructs (r*M. smegmatis*[pNM4.1] and r*M. smegmatis*[pNM4.2] containing two *tet*O1 sites

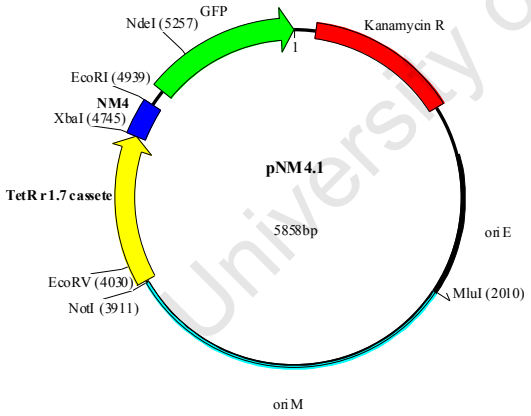
engineered into the P_{smyc} promoter), showed greater GFP expression than pNM3-derived constructs (r*M. smegmatis* [pNM3.1] and r*M. smegmatis* [pNM3.2], containing two *tet*O2 sites engineered into the P_{smyc} promoter).



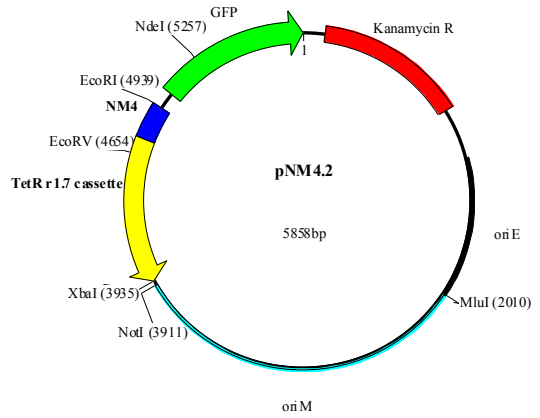
Full length P_{smyc} promoter with **two *tet*O2 sites** in pRC100 backbone and TetR r1.7 cassette in the same orientation as GFP (orientation 1)



Full length P_{smyc} promoter with **two *tet*O2 sites** in pRC100 backbone and TetR r1.7 cassette in the opposite orientation to GFP (orientation 2)



Full length P_{smyc} promoter with **two *tet*O1 sites** in pRC100 backbone and TetR r1.7 cassette in the same orientation as GFP (orientation 1)



Full length P_{smyc} promoter with **two *tet*O1sites** in pRC100 backbone and TetR r1.7 cassette in the opposite orientation to GFP (orientation 2)

Figure 4.13. Plasmid maps for Phase 2 promoter-operator shuttle vector constructs containing the TetR r1.7 cassette.

In addition, the constructs containing the TetR r1.7 cassette in the opposite orientation to GFP i.e. orientation 2 (r*M. smegmatis*[pNM3.2] and r*M. smegmatis* [pNM4.2]), displayed higher expression than their orientation 1 counterparts (r*M. smegmatis*[pNM3.1] and r*M. smegmatis*[pNM4.1]). Thus the highest Phase 2 GFP expresser was r*M. smegmatis* [pNM4.2]. Taken together, these observations indicate that the presence of the TetR r1.7 expression cassette not only affects cell "fitness', causing them to grow at a slower rate but also affects the level of GFP expressed which is further dependent on the orientation of insertion.

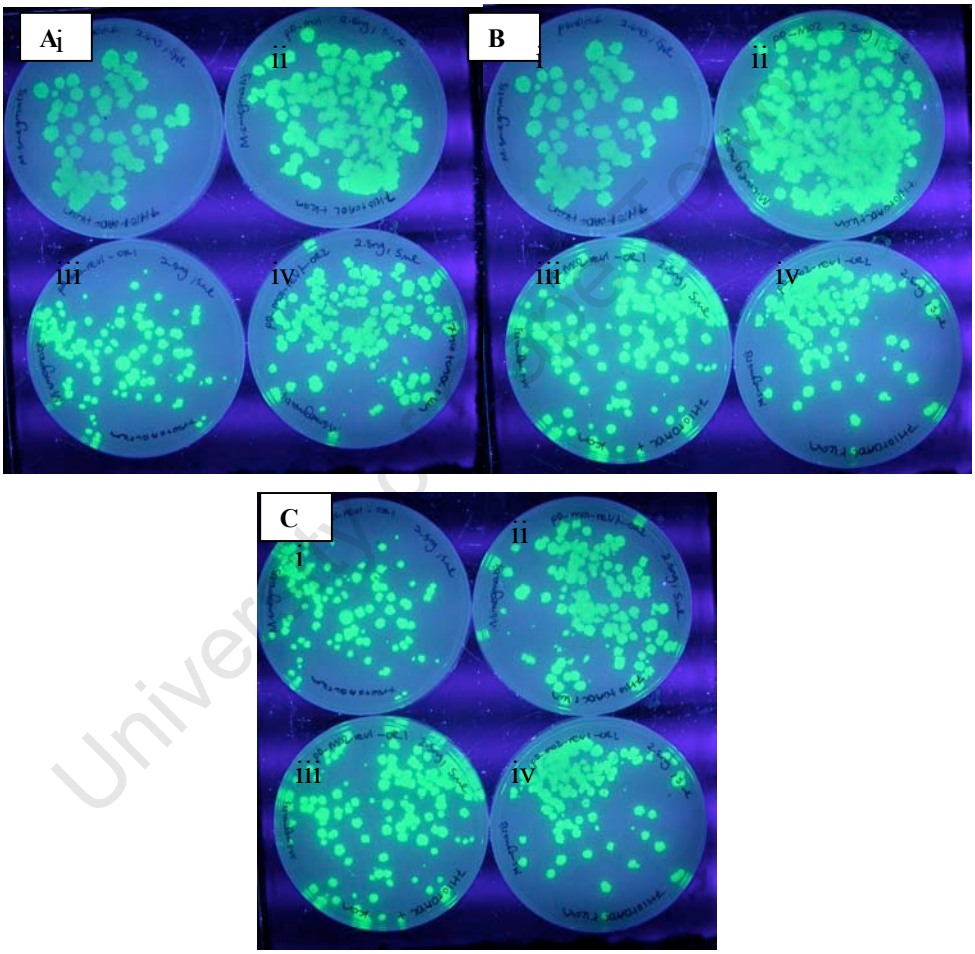


Figure 4.14. Photographic images of fluorescent *M. smegmatis* transformants harbouring the Phase 2 plasmids. *M. smegmatis* cells were electrotransformed with the various Phase 1 and two constructs and plated on appropriate selective media and incubated for 3 days at 37°C. Panel A represents *M. smegmatis* harbouring the plasmids, pRC100 (i), pNM3 (ii), pNM3.1 (iii) and pNM3.2 (iv). Panel B represents *M. smegmatis* harbouring the plasmids, pRC100 (i), pNM4 (ii), pNM4.1 (iii), and pNM4.2 (iv). Panel C represents *M. smegmatis* harbouring the plasmids, pNM3.1 (i), pNM3.2 (ii), pM4.1 (iii) and pNM4.2 (iv).

A similar trend was seen for BCG panCD cells that were also transformed with the Phase 2 plasmids, except that the higher expressers resulted in significantly smaller colony sizes and a high frequency of large non-fluorescent colonies (data not shown).

Table 4.4. Summary of the comparative fluorescent trends observed for recombinant mycobacteria harbouring Phase 2 constructs. Relative fluorescence intensities were recorded as a score out of 5, with 5 representing the highest level of GFP expression observed.

Plasmid Name	Promoter-operator-TetR r1.7 description	Relative Fluorescence in r*M. smegmatis* and rBCG panCD
pWB100	Empty vector	0
pRC100	*hsp*60 driven GFP	3.5
pNM3	Full-length P_{smyc} promoter two *tet*O2 sequences situated on either side of -35 region.	4.5
pNM4	Full-length P_{smyc} promoter and two *tet*O1 sequences situated on either side of -35 region.	4.75
pNM3.1	pNM3 with the TetR r1.7 cassette in the same orientation as GFP	3.75
pNM3.2	pNM3 with the TetR r1.7 cassette in the opposite orientation to GFP	4
pNM4.1	pNM4 with the TetR r1.7 cassette in the same orientation as GFP	4.25
pNM4.2	pNM4 with the TetR r1.7 cassette in the opposite orientation to GFP	4.5

4.3.5.2. Comparison of GFP expression in r*M. smegmatis* cultures by quantitative GFP-capture Enzyme-linked Immunoabsorbant Assay (ELISA)

In order to quantitate the difference in GFP expression between the various Phase 2 constructs, recombinant *M. smegmatis* cultures were grown in the appropriate M7H9-TLX-OADC media to an OD_{600} of approximately 0.8 units. This represents an approximately 2-fold higher density than that used to harvest the Phase 1 proteins. Proteins were prepared from bacterial lysate and ELISAs performed as described for Phase 1 constructs as outlined in Section 2.12.2. A summary of the results is presented here in Figure 4.15 in a bar graph format as a percentage of total protein expression.

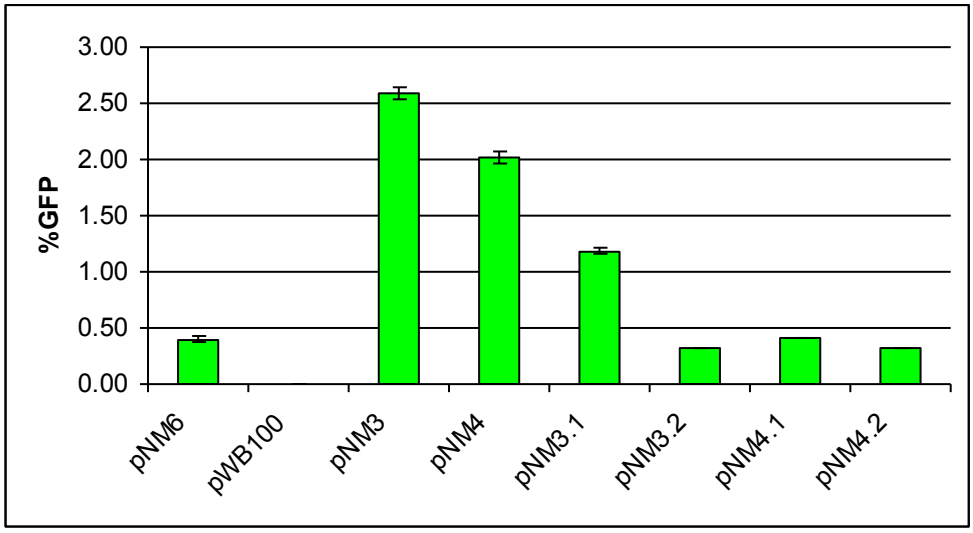


Figure 4.15. Comparison of the percentage of GFP expressed by recombinant *M. smegmatis* cultures containing Phase 2 constructs as determined by ELISA. The amount of GFP produced was calculated by dividing the total amount of GFP expressed (ng) by the amount of total protein (ng), and expressed as a percentage. The experiment was performed in triplicate and error bars are as included.

As expected, the r*M. smegmatis*[pWB100] negative control sample showed no GFP expression whilst, as for the previous Phase 1 ELISA, the r*M. smegmatis*[pNM6] promoter-less control displayed a low background level of GFP expression (0.4%), an indication that read-through is occurring from an upstream promoter-like region. Consistent with observations from the Phase 1 ELISA results (Figure 4.9), the r*M. smegmatis*[pNM3] sample produced more GFP (2.6%) than r*M. smegmatis*[pNM4] sample (2%). Once again, this is contrary to what was observed from the plated r*M. smegmatis* colonies. Furthermore, the 'weakest' Phase 2 recombinant on plates, r*M. smegmatis*[pNM3.1], produced more GFP than r*M. smegmatis*[pNM3.2], r*M. smegmatis*[pNM4.1] or r*M. smegmatis*[pNM4.2], which were all at low levels comparable to the read through levels seen for r*M. smegmatis*[pNM6]. Again, these results were very surprising and contrary to what was observed from the plated r*M. smegmatis* colonies. However, even at these low levels it was evident that the r*M. smegmatis* samples harbouring pNM3.2 and pNM4.2 (containing the TetR r1.7 cassette in the opposite orientation to GFP) expressed less GFP than the r*M. smegmatis* samples harbouring their orientation 1 counterparts i.e. pNM3.1 and pNM4.1 (containing the TetR r1.7 cassette in the same orientation as GFP).

Overall GFP yields for the r*M. smegmatis*[pNM3] and r*M. smegmatis*[pNM4] samples were approximately two-fold higher than in the previous Phase 1 ELISA. This could be attributed to differences in promoter activity at the different growth stages since cultures containing Phase 2 constructs were harvested at a later OD_{600} than the Phase 1 containing cultures i.e. an OD_{600} of 0.8 units versus an OD_{600} of 0.4-0.5 units, respectively.

4.4. Discussion

Phase 1, the selection of a strong mycobacterial promoter engineered to contain a combination of two 19bp *tet* operator sequences, yet which still support strong GFP expression, was successfully achieved. The design was based on the study reported by Ehrt *et al*., 2005[53], in which they identified and isolated the strong P_{smyc} promoter from *M. smegmatis* and demonstrated successful TetR (wild type tetracycline regulatory protein) regulation of expression of their GFP reporter gene and essential mycobacterial gene, *ftsZ* following the insertion of two *tet*O2 elements flanking the putative –35 region. However, since sequence analysis of the limited published region of the P_{smyc} promoter revealed the presence of an extensive region of 125bp between the identified transcriptional start site and the start of the coding sequence for the putative 30S ribosomal protein S1, we sought to determine the influence the presence or absence of this region would have on the strength of GFP expression, as our aim was to design the strongest possible TetR r1.7 regulatable promoter. It is thought that the presence of extensive transcripts preceding a gene may lead to enhance the stability of mRNA transcripts and thus enhanced gene expression. Thus, it serves as a means of post-transcriptional regulation of gene expression. A study published by Hillmann *et al*., 2007[103] elucidated a 200bp region upstream of the *msp*A gene (which encodes a major porin protein MspA) that increased *lacZ* expression by 12-fold.

To address this we designed two versions of the P_{smyc} promoter; a full-length 190bp version (NM1) that contained the complete region upstream of the putative *rsp*A gene to the transcriptional start site (70bp), and a 90bp truncated version (NM2) that contained only 5bp after the transcriptional start site (not including the engineered

*Xba*1 cloning site). Although there was not a dramatic difference in the intensity of GFP fluorescence when the resulting shuttle-vectors pNM1 and pNM2 were transformed into any of the three mycobacterial strains (*M. smegmatis*, BCG Pasteur and BCG panCD), the truncated version (pNM2) resulted in slightly less GFP fluorescence when recombinant colonies were visualised over a UV trans-illuminator. It was thus concluded that the full-length pre-gene transcript played a role in enhancing or stabilising GFP expression and was henceforth included in all further promoter-operator designs. These fluorescent studies further supported the reports that the P_{smyc} promoter was indeed significantly stronger than the *M. tuberculosis hsp*60 promoter present in the parental shuttle-vector pRC100, in both *M. smegmatis* and BCG[53]. The *hsp*60 promoter is commonly used as a gold standard of a strong mycobacterial promoter[57, 104].

In the natural tetracycline regulated system the affinity of TetR for *tet*O2 is 3-5 times stronger than its affinity for *tet*O1 sequence elements[71]. If the reverse mutant TetR r1.7 regulatory protein shares these characteristics, then it may be expected that in the presence of ATc, GFP expression will be more strongly down regulated (switched off) when *tet*O2 sequence elements are present compared to when *tet*O1 sequence elements are present in the promoter. However, this would also be expected to result in more poorly induced GFP expression upon removal of ATc for *tet*O2 versus *tet*O1. Since it is difficult to predict how well a theoretical system design will work in practice, we tested three different promoter-operator designs in our multicopy episomal system; one which contained two identical *tet*O2 sequence elements on either side of the –35 region of the P_{smyc} promoter (pNM3) as reported by Ehrt *et al*., 2005[53, 56], a second which contained two identical *tet*O1 sequence elements on either side of the –35 region of the P_{smyc} promoter (pNM4), and a third which contained one *tet*O1 sequence element and one *tet*O2 sequence element on either side of the –35 region of the P_{smyc} promoter (pNM5).

Fluorescent intensities of all these Phase 1 constructs and controls were compared by both visualisations of transformed colonies over a UV trans-illuminator and by quantitative GFP capture ELISAs of prepared cell lysates. It should be noted that in the designing of the promoter-operator fragments foreign operator sequences replaced potentially imperative sequences surrounding the -35 region of the P_{smyc}

promoter. Thus, some effect on GFP expression levels was expected. In both assay systems it was evident that the insertion of either identical *tet*O2 sequence elements (pNM3) or identical *tet*O1 sequence elements (pNM4) resulted in only a slight decrease in GFP expression when compared to pNM1 (full-length natural promoter construct without operator sites). r*M. smegmatis*[pNM3] colonies scored a fluorescent intensity of 4.5 whilst r*M. smegmatis*[pNM4] colonies scored 4.75 in comparison to r*M. smegmatis*[pNM1] colonies that scored 5. In the ELISAs r*M. smegmatis*[pNM3] and r*M. smegmatis*[pNM4] samples appeared to express very similar levels of GFP expression i.e. 1.33% and 1.25%, respectively, and only marginally less than r*M. smegmatis*[pNM1] which produced 1.42% GFP. Both r*M. smegmatis*[pNM3] and r*M. smegmatis*[pNM4] however supported greater GFP expression than the *hsp*60 driven parental construct sample i.e. r*M. smegmatis*[pRC100] which produced 0.56% GFP. However, the insertion of one *tet*O1 sequence element and one *tet*O2 sequence element to form the chimaeric promoter-operator sequence NM5 clearly disrupted efficient promoter functioning resulting in dramatically reduced GFP expression with r*M. smegmatis*[pNM5] colonies scoring a low 3 in the plating assay and producing just 0.38% GFP in the ELISA, well below that seen for the parental plasmid in r*M. smegmatis*[pRC100]. Thus, the high GFP expressing shuttle-vectors pNM3 and pNM4 were considered suitable candidates for further development of the reverse tetracycline inducible expression system i.e. for the insertion of the TetR r1.7 cassette.

High expression levels of heterologous antigens can lead to toxic and or a metabolic burden on bacterial cells[25, 44, 46]. This typically leads to slower growth rates and results in genetic instability of the gene being expressed since as soon as a mutant arises that has a growth rate advantage it will out compete the rest of the colony or culture and take over the colony or culture in question. Thus, the larger the toxic or metabolic burden, the slower the growth rate and the higher the rate of instability. This is reflected in relative colony sizes wherein; the larger the colonies, the more 'healthy' the recombinants and the smaller the colonies, the more stressed or 'unhealthy' the recombinants are. It was evident that colonies containing the Phase 1 constructs that supported high GFP expression (pNM1, pNM2, pNM3 and pNM4) grew at a slower rate than colonies containing pRC100, resulting in slightly smaller colonies on the transformation plates. This is attributed to the metabolic burden

placed on the cell by having to produce higher levels of the foreign protein in this case, GFP.

With the introduction of the TetR r1.7 cassette into pNM3 and pNM4, a further reduction in colony size was observed for all the Phase 2 plasmids; pNM3.1, pNM3.2, pNM4.1 and pNM4.2 upon transformation of all three mycobacterial strains. Furthermore, a number of large non-fluorescent colonies were observed for the rBCG Pasteur and rBCG panCD plated transformants but not for r*M. smegmatis* transformants. This suggests that, despite the fact that TetR r1.7 is of bacterial origin and that the gene had been codon optimised for expression in mycobacteria, the high constitutive expression from the naturally stress induced *hsp*60 promoter[50], has increased the metabolic load resulting in the appearance of faster growing non-fluorescent mutants.

Furthermore, insertion of the TetR r1.7 expression cassette lead to an overall decrease in GFP expression compared to their cognate Phase 1 constructs. Even so, consistent with trends seen for the Phase 1 constructs, mycobacterial colonies harbouring pNM4.1 and pNM4.2 resulted in stronger fluorescence than those harbouring pNM3.1 and pNM3.2. In addition, the orientation of the insertion strongly influenced the level of GFP expression observed in mycobacterial transformants exposed to UV illumination. In both instances those recombinant mycobacterial colonies containing constructs with the TetR r1.7 cassette divergently oriented to GFP i.e. in orientation 2 (pNM3.2 and pNM4.2) were 'brighter' i.e. displayed greater expression of GFP than the orientation 1 versions (pNM3.1 and pNM4.1). This suggests that insertion of the TetR r1.7 expression cassette affects GFP expression in an orientation dependent manner, perhaps due to the secondary mRNA structure in the region of the promoter as a result of close proximity of the GFP gene and TetR r1.7 expression cassette.

Unexpectedly, we observed an inverse-type relationship between *M. smegmatis* recombinants that displayed high GFP expression on plates (strong fluorescence) and GFP expression in the ELISA, which were lower than expected. In the ELISAs performed during Phase 2, r*M. smegmatis*[pNM4] samples showed 0.6% less GFP expression than the r*M. smegmatis*[pNM3] samples in contrast to what was observed

from the plated transformants over UV where pNM4 was deemed the stronger expresser. Furthermore, the r*M. smegmatis* samples harbouring the plasmids, pNM4.1, pNM4.2 and pNM3.2 showed such diminished levels of GFP production that were comparable to the promoter-less control sample r*M. smegmatis*[pNM6] which displayed a low level (0.4%) of read through GFP expression. However, the Phase 2 sample r*M. smegmatis*[pNM3.1] produced 1.2% GFP expression that was significantly greater than the promoter-less control. Notably, this construct appeared to produce the least GFP from all the Phase 2 plated transformants.

We propose that these results reflect different rates of instability between the constructs, as poor levels of GFP expression seen in the ELISA correlated with strong GFP fluorescent intensity and small colony size. It should be noted that the cultures utilised in these ELISAs would have gone through several sub-culturing processes (average of 3 sub-cultures) resulting in approximately 15 additional generations that could allow a more 'fit' GFP mutant to dominate the culture. These results suggest that the entire plasmid has not been lost as recombinant mycobacterial colonies grew with antibiotic selection, but that the GFP gene had either been deleted or somewhat modified, preventing its expression.

This instability is the very problem that this regulatory system wishes to address by employing the reverse tetracycline regulated expression system to repress gene expression during *in vitro* culturing in the presence of ATc, thus reducing the metabolic load encountered by the mycobacteria and promoting genetic stability. Thus, all four Phase 2 constructs (pNM3.1, pNM 3.2, pNM4.1 and pNM 4.2) were used to further test the efficacy of the regulated expression system in the presence of ATc.

Chapter 5:Testing the revTetR regulatable expression system in *M. smegmatis*

Chapter 5:Testing of the revTetR regulated expression system in *M. smegmatis*

5.1. Introduction

The revTetR regulated system has previously been engineered for use in mycobacteria for purposes of conditional gene silencing but to date there has been nothing published about the use of the revTetR system to conditionally regulate the expression of foreign antigens in mycobacteria in a vaccine context[56, 82, 105]. Ideally, we would require repression of foreign gene expression by ATc during the culturing processes to minimise the toxic and metabolic load to ensure the genetic integrity of the vaccine, and would induce expression upon delivery of the vaccine by the removal of ATc. In order to try and achieve this goal, multiple versions of an episomal revTetR regulatable system were successfully constructed, resulting in the four candidate constructs, pNM3.1, pNM3.2, pNM4.1 and pNM4.2 expressing the model antigen GFP.

This chapter explores how efficiently the co-repressor molecule ATc, could control GFP expression in the different construct designs i.e.:

- Could GFP expression be repressed by ATc in a dose-dependent manner and if so to assess the time required to achieve this.
- Could GFP expression be induced upon removal of ATc and if so to assess the time required to achieve this.

5.2. Optimisation of the GFP fluorescence assay

In order to conduct these experiments, we ideally required a simpler and quicker assay than the more laborious and time-consuming ELISAs. GFP flourimetry assays of whole cells as reported by Carrol and Parish, 2004[98] provided a quick, reliable and reproducible method. This procedure essentially involved a comparison of fluorescence emitted by whole live cells harvested from culture at the same growth

phase and resuspended to the same cell density. In this paper, harvested mycobacterial culture samples expressing GFP were diluted with phosphate-buffered saline (PBS) to an OD_{600} of 2 units prior to taking readings on the fluorimeter.

Before implementing these methods to our experiments we wished to first establish the relationship between culture density (OD_{600}) and GFP fluorescence in order to familiarise ourselves with the assay and ensure that we use the optimal condition of cell density in our experiments. Thus, a preliminary fluorimetry experiment was performed using the construct, pNM3. An *M. smegmatis* culture harbouring pNM3 was propagated from a glycerol stock to an OD_{600} of 0.4 units in 25ml of the appropriate M7H9-TLX-OADC media with kanamycin selection in the absence of ATc. The r*M. smegmatis*[pNM3] culture (25ml) was concentrated into 1ml of PBS resulting in an OD_{600} of approximately 11.025 units that was subsequently diluted in order to achieve an OD_{600} of 1, 2 and 3 units. These test samples were loaded into white fluorimetry plates in triplicate and fluorescence readings recorded.

As can be seen in Figure 5.1, under these specific culture preparation conditions, there exists a linear relationship between culture density and GFP fluorescence readings, between an OD_{600} of 1 to 11 units. The correlation coefficient was greater than 0.9 ($R^2>0.9$) signifying a statistcally significant positive correlation between culture density and relative fluorescence readings. This experiment thus provided us with the confidence to use an OD_{600} of 2 units for subsequent fluorimetry experiments as reported by Carroll and Parish in 2004[98].

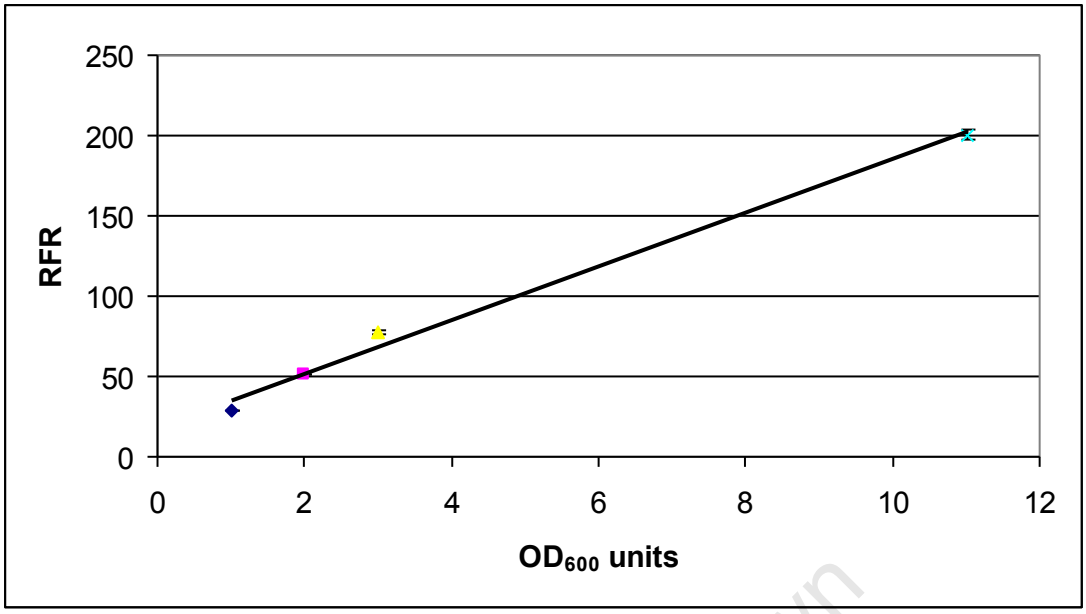


Figure 5.1. Optimisation of the whole cell fluorimetry assay to determine the relationship between *rM. smegmatis* culture densities and GFP fluorescence. Fluorescence readings were taken for the r*M. smegmatis*[pNM3] experimental sample at an OD_{600} of 1, 2, 3 and 11.025 units (x-axis) to explore the correlation between culture density and GFP production. The relative fluorescence readings (RFR, y-axis) represent values recorded directly from the fluorimeter. Error bars are as included.

5.3. Preliminary testing of the revTetR system in *M. smegmatis*

Having confirmed the assay conditions, a small preliminary pilot experiment was designed to assess whether:

- GFP expression could be repressed at all in the presence of the co-repressor molecule ATc
- GFP expression could in fact be induced upon removal of the co-repressor molecule ATc

This pilot experiment was performed using a concentration of 100ng/ml of ATc. r*M. smegmatis* cultures containing the constructs pWB100 (served as the negative control instead of pNM6 as basal read-through was observed in ELISAs), pNM3 (control) and pNM3.1 (experimental) were propagated in 200ml of the appropriately supplemented M7H9-TLX-OADC media, without ATc until an OD_{600} of 2 units was obtained. Each of these 'starter' cultures was then equally subdivided and 40ml used to inoculate bottles containing 30ml of fresh media with either no ATc (0ng/ml) or

supplemented such that the final concentration of ATc was 100ng/ml. After 12 hours of incubation following the subculturing step, 20ml samples were removed from the culture bottles and prepared for fluorimetry as described in section 2.12.3. The remaining 50ml of each culture was then washed to remove the ATc from those that originally contained it and the cultures re-incubated. Approximately 6 hours after the removal of ATc, the remaining culture was prepared for the fluorimetry assays as before.

Results for this experiment are summarised in bar graphs in Figure 5.2. r*M. smegmatis*[pWB100] resulted in fluorimeter output readings of approximately 1.80 that could be attributed to fluorescence of mycobacterial proteins in the cultures and was used to account for background fluorescence readings and subtracted from the experimental sample readings. As can be seen from Figure 5.2 (A), following culturing in the presence of 100ng/ml of ATc, approximately 3.2-fold repression of expression was seen for r*M. smegmatis*[pNM3.1] when compared to it's growth in the absence of ATc. However, unexpectedly r*M. smegmatis*[pNM3.1] displayed higher relative fluorescent units (RFU) than r*M. smegmatis*[pNM3] in the absence of ATc. This is contrary to what was observed from the plated transformants, which showed that the insertion of the TetR r1.7 cassette reduced GFP expression. We speculate that this result reflects differences in stabilities of the constructs such that when the starter cultures reached an OD_{600} of 2 units, the higher expressing rM. *smegmatis*[pNM3] culture was more unstable than the r*M. smegmatis*[pNM3.1] culture which contained a higher proportion of GFP expressing cells. Furthermore, we observed that r*M. smegmatis*[pNM3] cultured in the presence of 100ng/ml ATc showed a slight (1.37-fold) increase in GFP expression with respect to r*M. smegmatis*[pNM3] cultured without any ATc. ATc was not expected to affect GFP expression in constructs such as pNM3 that do not contain a TetR r1.7 cassette.

Figure 5.2 (B) depicts the relative fluorescence intensities measured after washing of the cells to remove any ATc present. Notably, all the RFU readings were significantly depressed after washing and further incubation compared to before washing. Once again the same trends were observed with respect to the higher RFUs recorded for the r*M. smegmatis*[pNM3.1] sample versus the r*M. smegmatis*[pNM3] sample originally cultured in the absence of ATc, as well as the increase in

fluorescence observed for r*M. smegmatis*[pNM3] originally cultured in the presence of 100ng/ml of ATc. However, 6 hours post removal of ATc, a significant reduction in the difference of the relative fluorescence intensities was measured between the r*M. smegmatis*[pNM3.1] cultures, which was reduced to just 1.37-fold. This reduction could be attributed to the induction of GFP expression in the absence of ATc.

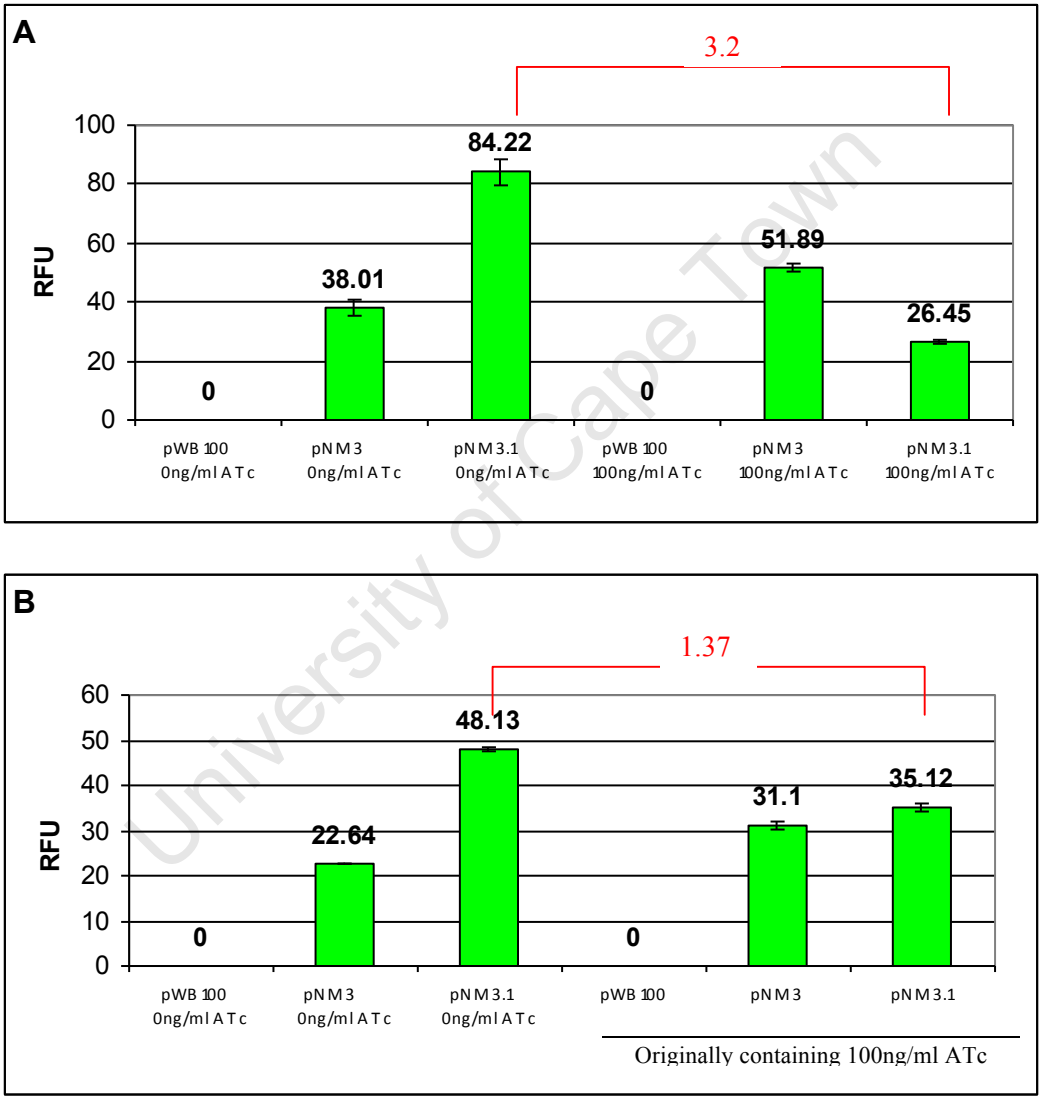


Figure 5.2. Pilot study to assess the repression and induction potential of the selected *rM. smegmatis*[pNM3.1] culture expressing GFP in the presence and absence of ATc, respectively. A) Fluorimetry results for r*M. smegmatis*[pNM3] and r*M. smegmatis*[pNM3.1] cultures after 12 hours of growth in the absence or presence of 100ng/ml of ATc. B) Fluorimetry results of the same r*M. smegmatis* cultures 6 hours after washing to remove any ATc if present. GFP expression was measured using the whole cell fluorimetry assay and expressed as relative fluorescence units (RFU) after subtracting the background readings of *rM. smegmatis*[pWB100]. Error bars are as included.

5.4. ATc-dependent repression and induction of GFP expression in r*M. smegmatis* cultures (proof of concept)

Having established that the addition of ATc was able to repress GFP expression in r*M. smegmatis*[pNM3.1], and that its removal appeared to lead to the de-repression (induction) of GFP expression, the following proof of concept experiment was designed to assess the time required to achieve maximum repression and how this related to the dose of ATc used, as well as the time course for induction to occur once ATc was removed.

All four Phase 2 constructs were tested in this proof of concept experiment described in detail in section 2.12.3.1. In order to accommodate all the experimental constructs and their controls, the experiment was performed identically in two batches. Group 1 comprised r*M. smegmatis* cultures containing the constructs, pWB100 (negative control), pNM3 (control), pNM3.1 and pNM3.2 and Group 2 consisted of r*M. smegmatis* cultures containing the constructs pWB100, pNM4 (control), pNM4.1 and pNM4.2. Starter cultures were established from glycerol stocks and the same volume used to inoculate fresh M7H9-TLX-OADC media containing kanamycin and either no ATc (0ng/ml) or three different concentrations of ATc i.e. 25ng/ml, 50ng/ml and 100ng/ml. All cultures were incubated appropriately and 5ml of culture was removed from each of the sample bottles following 2, 8 and 12 hours of incubation. At each time point the 5ml samples were prepared as described in Section 2.12.3, fluorimetry readings taken in triplicate and assessed. ATc was removed after 12 hours of incubation by washing the cells in M7H9-TLX-OADC media twice and gently resuspending them in the same media but without any ATc. The 'washed' cultures were further incubated and 5ml of sample removed at 6 and 12 hours post washing. The removed samples were prepared as before, fluorimetry readings taken in triplicate and compared. The results for this experiment are tabulated in Table 5.1 and summarized in Figure 5.3. The r*M. smegmatis*[pWB100] samples, which served as the negative control, consistently displayed low background fluorescence readings in the presence or absence of ATc (average 1.80 RFR). Thus, all fluorimetry readings were expressed as RFU's in which this background fluorescence has been subtracted.

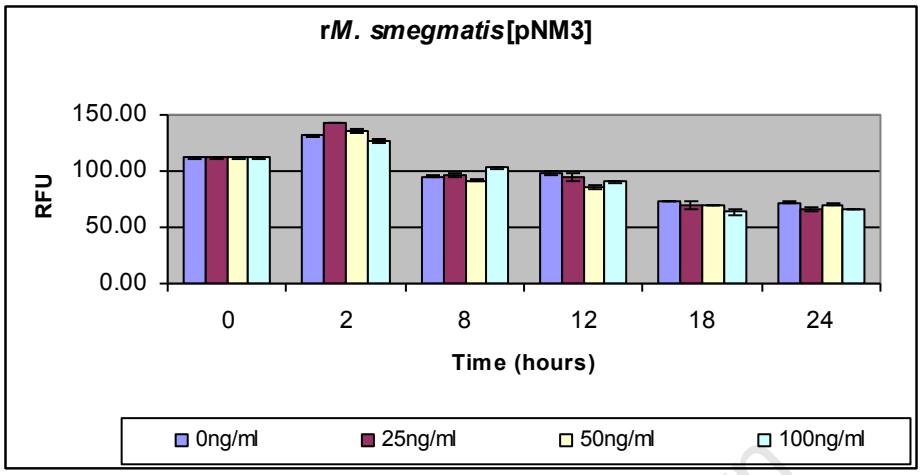
Table 5.1. Summary of the r*M. smegmatis* ATc regulated GFP expression proof of concept study by whole cell fluorimetry. Fluorimetry readings were taken at various time points before and after the washing step. The RFU output represents an average of three readings from which the background RFU readings (r*M. smegmatis*[pWB100] 0-100ng/ml ATc) were subtracted and normalised for discrepancies in culture density.

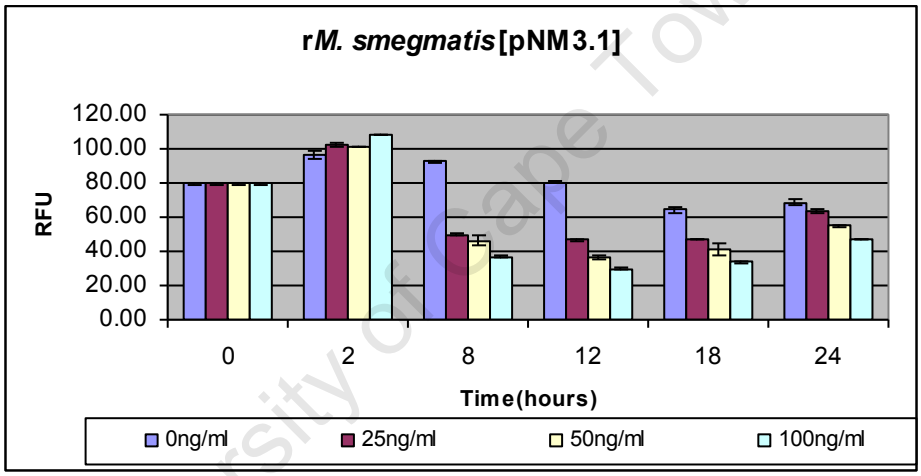
r*M. smegmatis* sample	ATc concentration (ng/ml)	RFU at 0 hours	RFU at 2 hours	RFU at 8 hours	RFU at 12 hours	Washing Step	RFU at 6 hours	RFU at 12 hours
pNM3 (control)	0	111.67	131.35	95.52	98.19		72.94	72.00
	25	111.67	142.18	96.17	95.26		69.85	66.31
	50	111.67	136.05	91.84	85.94		69.57	69.82
	100	111.67	126.59	103.53	90.51		63.4	66.00
pNM3.1	0	79.75	96.49	92.68	80.21		64.20	68.79
	25	**79.75**	**102.24**	**49.58**	**46.49**		**46.67**	**63.17**
	50	**79.75**	**101.36**	**46.37**	**36.65**		**40.62**	**55.21**
	100	**79.75**	**108.63**	**36.78**	**29.85**		**33.76**	**46.92**
pNM3.2	0	3.94	6.30	4.52	3.46		3.33	4.02
	25	3.94	6.54	4.02	1.29		2.56	2.30
	50	3.94	7.66	3.50	0.47		2.49	2.39
	100	3.94	5.95	2.12	0.27		0.91	1.31
pNM4 (control)	0	115.87	121.81	123.92	85.65		59.13	58.69
	25	115.87	142.23	135.30	92.30		64.25	64.71
	50	115.87	151.90	154.87	95.80		59.58	59.53
	100	115.87	134.78	132.65	90.29		56.84	69.85
pNM4.1	0	17.63	22.94	20.70	13.37		8.92	8.86
	25	17.63	21.83	16.24	10.48		6.30	7.03
	50	17.63	18.73	14.98	9.25		5.55	7.45
	100	17.63	18.34	13.07	9.89		8.19	7.78
pNM4.2	0	8.44	11.53	10.20	8.84		4.20	5.09
	25	8.44	11.30	10.85	6.38		4.74	4.82
	50	8.44	9.64	9.95	7.39		4.46	5.24
	100	8.44	8.88	6.80	6.99		3.11	4.84

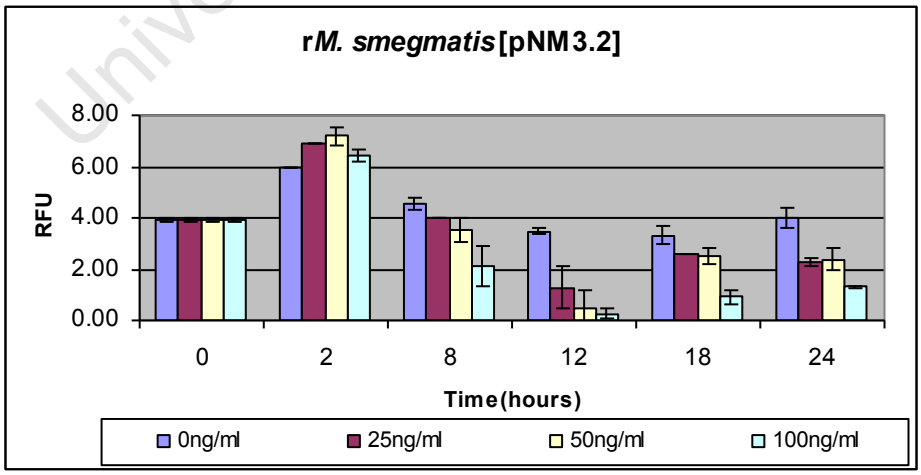
Both the control samples r*M. smegmatis*[pNM3] and r*M. smegmatis*[pNM4] resulted in high initial levels of GFP fluorescence. We observed that these levels increased further for all samples at 2 hours (and 8 hours for r*M. smegmatis*[pNM4] in the presence of 0ng/ml and 50ng/ml) after the initial inoculations, but that all the readings then decreased with time by 12 hours and even further after the washing step. As expected, no evidence of GFP repression in the presence of increasing concentrations of ATc or induction after its removal was observed, since these constructs do not contain the TetR r1.7 cassette for ATc regulation.

For all four test cultures r*M. smegmatis*[pNM3.1], r*M. smegmatis*[pNM3.2], r*M. smegmatis*[pNM4.1] and r*M. smegmatis*[pNM4.2], we again observed a similar trend of an overall increase in RFU at 2 hours post sub-inoculation of the experimental culture flasks, and a general diminishing of fluorescence over time. However, in contrast to r*M. smegmatis*[pNM3.1], the r*M. smegmatis* cultures harbouring shuttle vector constructs pNM3.2, pNM4.1 and pNM4.2, all displayed dramatically reduced RFU values for all samples in the presence or absence of ATc throughout the time course of the experiment. This is clearly depicted in Figure 5.4 in which the 0 hour time point was selected to illustrate this difference. Once again, this was contradictory to observations from transformed r*M. smegmatis* colonies viewed on plates over a UV trans-illuminator where these constructs appeared to support stronger GFP expression than r*M. smegmatis*[pNM3.1]. However, even though the readings were low and the relative differences in RFU were not large and somewhat erratic, a distinct trend could be seen between the levels of GFP repression and the concentration of ATc used over the 12-hour assay period. Furthermore, these results suggest possible induction of GFP expression again after washing to remove the ATc (in culture samples r*M. smegmatis*[pNM4.1] and r*M. smegmatis*[pNM4.2] specifically), since the RFU levels recorded for the samples that were originally incubated in the presence of 25 to 100ng/ml of ATc achieved the same level of RFU 12 hours post washing (i.e. the 24 hr time point) as the control sample that was cultured in the absence of ATc. However, the accuracy of these readings may be questionable at such low levels.

A







B

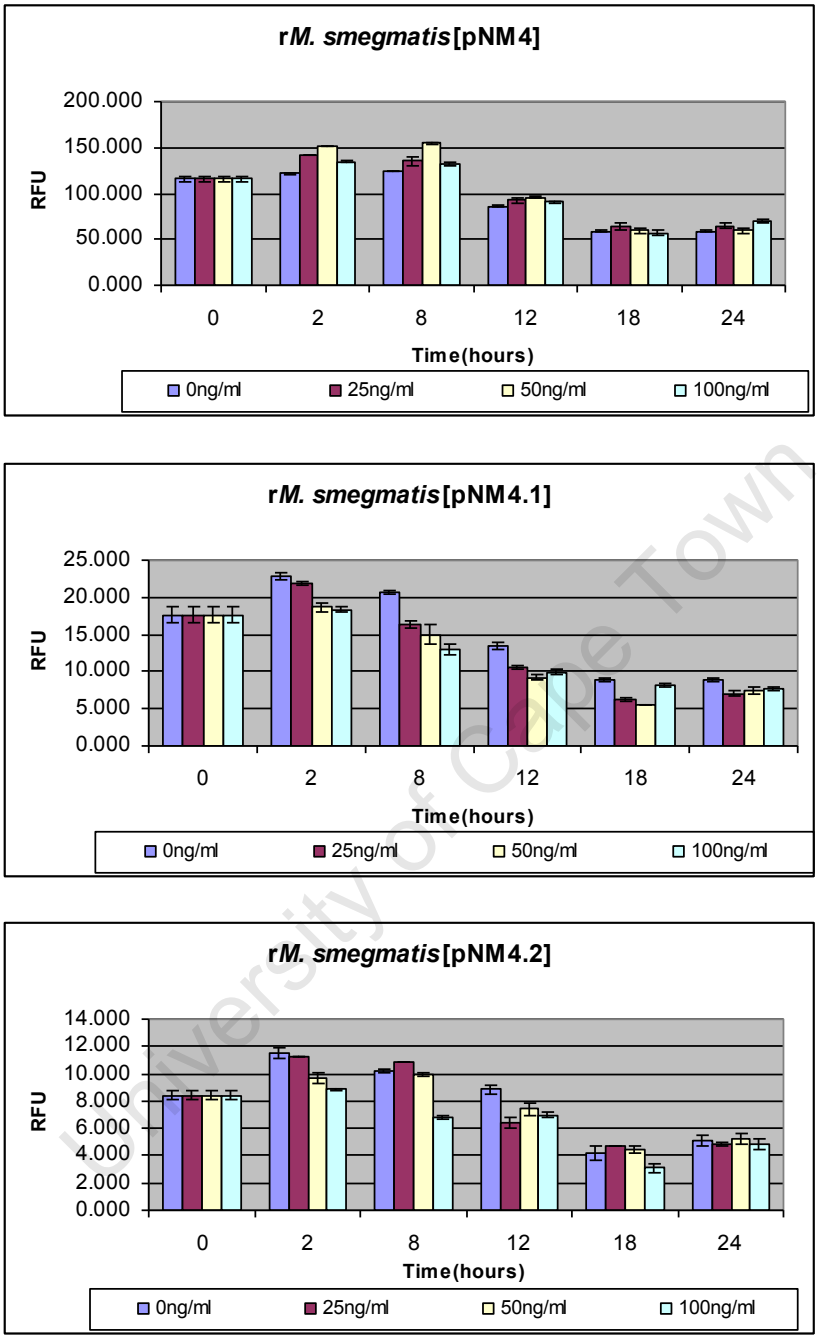


Figure 5.3. Relative GFP fluorescence between r*M. smegmatis* cultures containing various constructs at various time points measured by the whole cell fluorimetry assay. A) RFU readings for the Group 1, r*M. smegmatis* cultures containing the constructs; pNM3, pNM3.1 and PNM3.2. B) RFU readings for group 2, r*M. smegmatis* cultures containing the constructs; pNM4, pNM4.1 and pNM4.2. The ATc concentration for each culture is indicated on the x-axis 0, 25, 50 or 100ng/ml. The RFU output (y-axis) represents an average of three readings from which the background RFU readings (r*M. smegmatis*[pWB100] 0-100ng/ml ATc) were subtracted and normalised for discrepancies in culture density. Outlier values have been dismissed where appropriate and error bars are as indicated.

By comparison, r*M. smegmatis*[pNM3.1] showed high levels of RFU's in the absence of ATc at the start of the experiment (time 0 hours) which increased further 2 hours post inoculation of the fresh media as noted for all the other r*M. smegmatis* samples as well. Furthermore, as presented in the bar graph in Figure 5.3A and the line graph in Figure 5.5, significant repression of GFP expression was observed in the presence of ATc which was dose and time dependent i.e. the higher the ATc concentration and the longer the incubation time the greater the repression. Furthermore, following washing of the cultures to remove any ATc present, repression of GFP expression was reversed and GFP expression was induced in all samples as a function of time. The induction levels achieved were dependent on the original concentration of ATc present in the cultures.

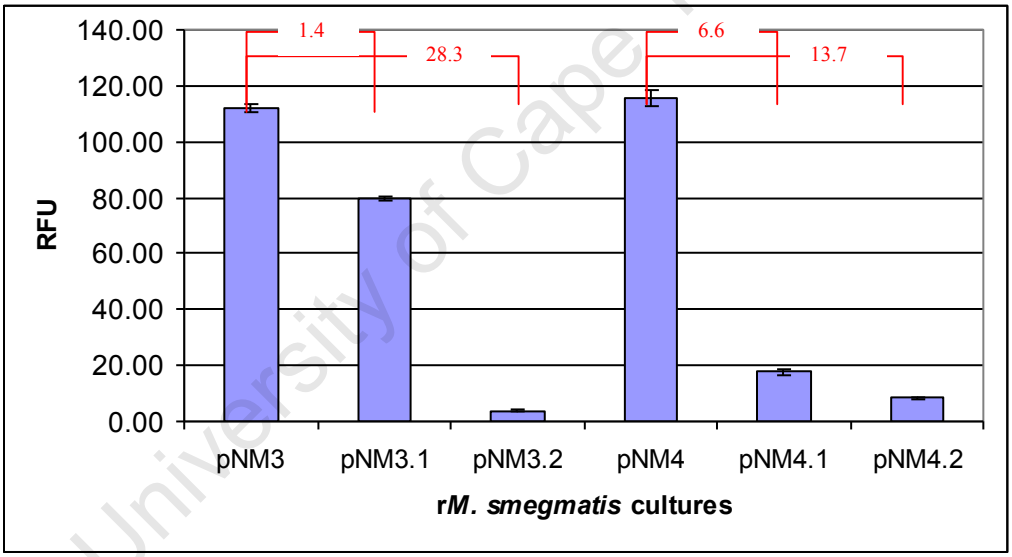


Figure 5.4. Relative GFP fluorescence between r*M. smegmatis* cultures containing various constructs at the 0-hour time point, before the addition of ATc measured by the whole cell fluorimetry assay. The RFU output represents an average of three readings from which the background RFU readings (r*M. smegmatis*[pWB100] 0-100ng/ml ATc) were subtracted and normalised for discrepancies in culture density. Outlier values have been dismissed where appropriate and error bars are as indicated.

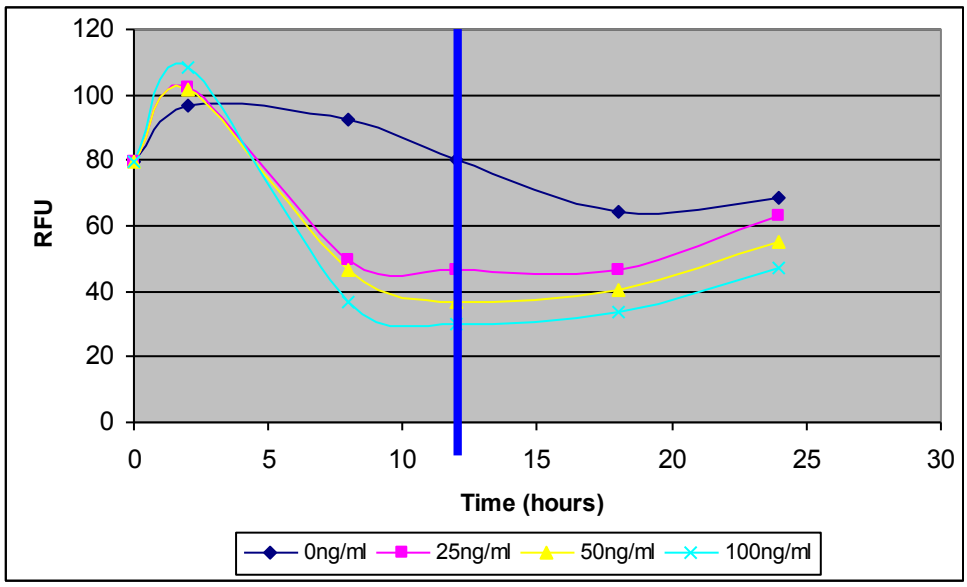


Figure 5.5. GFP expression (RFU) versus culturing time for the r*M. smegmatis*[pNM3.1] cultures as determined by the whole cell fluorimetry assay. This graph represents the fluorimetry data (RFU readings) at each of the time points i.e. 0, 2, 8, 12 hour time points before washing and the 6 and 12 hour time points after washing step as indicated by the blue line.

These results are more clearly illustrated in Figure 5.6 where the GFP expression levels of the r*M. smegmatis*[pNM3.1] culture containing no ATc was assigned 100% GFP production, and the GFP expression of the cultures containing the various amounts of ATc was expressed as a percentage of this. We noted that most of the repression occurred during the first 8 hours for all samples irrespective of ATc concentration, with GFP expression being reduced to between 40-53% at the 8-hour time point. Only marginal repression occurred after this. After 12 hours growth in the presence of 100ng/ml ATc, GFP expression had been repressed to 37% in relation to the control sample (0ng/ml ATc). This represented a 3- fold (60%) repression and was the maximum repression that could be achieved in this experiment in the presence of 100ng/ml of ATc. After 24 hours i.e. 12 hours following the removal of ATc, the sample that originally contained 25ng/ml ATc regained 92% of its maximum GFP expression while the sample that originally contained 100ng/ml of ATc, regained 68% GFP expression.

It should be noted that induction of GFP expression by the removal of ATc was slower than the repression of GFP expression by the addition of ATc, and even though we managed to almost obtain 100% induction of GFP expression (92%) in the sample containing only 25ng/ml GFP, full induction (100%) was not attained

even at 12 hours after the removal of ATc. Thus, the stronger the repression, the longer the period of induction required.

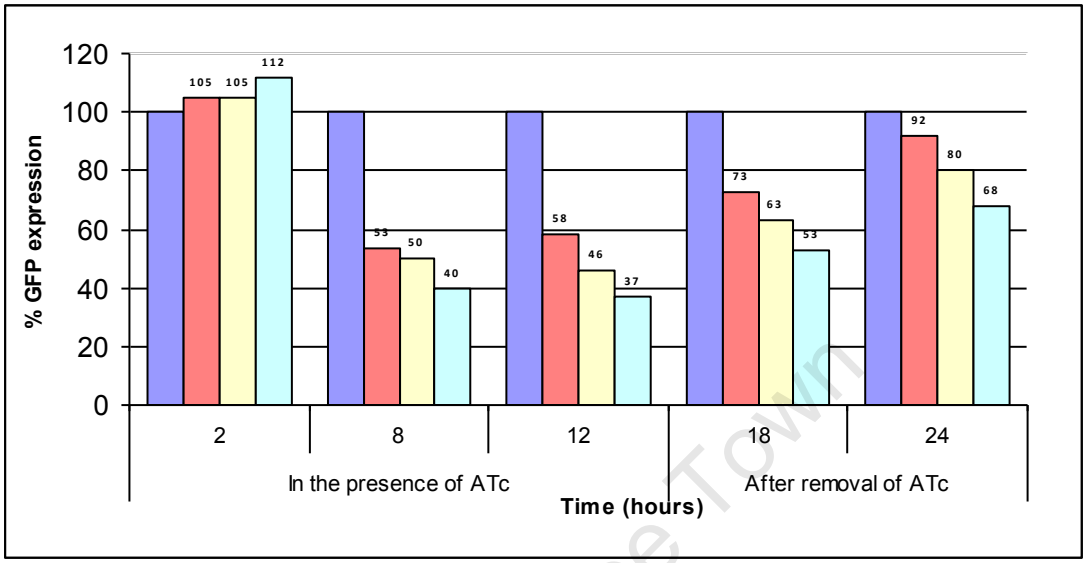


Figure 5.6. ATc-dependent repression of GFP expression in r*M. smegmatis*[pNM3.1] culture as determined by the whole cell fluorimetry assay. GFP expression was recorded as a percentage of the GFP expression recorded for r*M. smegmatis*[pNM3.1] without any ATc (0ng/ml-blue bar) which was taken to represent 100% GFP expression. Three ATc concentrations (25ng/ml-pink bar, 50ng/ml-pale pellow bar and 100ng/ml-pale green bar) were used.

5.5 Stability of recombinant plasmids

Caveats of instability were alluded to early in Chapter 4 wherein contradictory GFP expression levels were observed between different experiments i.e. between the plates containing freshly transformed r*M. smegmatis* colonies viewed over a UV trans-illuminator and ELISAs. A general trend was seen on the plates, where the higher the relative expression levels the more non-fluorescent colonies were observed and the smaller or less healthy the fluorescent colonies appeared. Generally, those constructs that displayed elevated GFP expression levels on plates showed depressed GFP expression levels in the ELISAs, indicating that an inverse relationship existed between expression levels and genetic stability. These contradictory results were also observed in the fluorimetry assays where the predicted higher expressers r*M. smegmatis*[pNM3.2], r*M. smegmatis*[pNM4] and r*M. smegmatis*[pNM4.2] exhibited negligible levels of GFP expression. As mentioned previously, the r*M. smegmatis* samples used for the ELISAs and fluorimetry experiments would have gone through several additional culture generations before being used in the assays allowing more time for more "fit" non-fluorescent mutants to dominate the culture.

Therefore, since GFP fluorescence serves as such a convenient marker of antigen expression we sought to determine the role genetic instability might play in these unexpected results by conducting the following fluorescent versus non-fluorescent colony counting experiment. Recombinant *M. smegmatis* cultures that were stored as glycerol stocks (the same as those used in fluorimetry experiments to initiate starter cultures) were plated at two dilutions (1/1000 and 1/10 000). The plates were incubated for 2-3 days and the number of fluorescent colonies versus non-fluorescent colonies compared over a UV trans-illuminator as an indication of plasmid integrity. The results are summarised in Table 5.2.

Table 5.2. Summary of the stability of the recombinant plasmids by visualising plated transformants over a UV trans-illuminator. Relative fluorescent colony size (RFCS) was described using an arbitrary scoring system of between 1 and 5 and relative colony fluorescent intensity (RFI) was described using an arbitrary scoring system of between 1 and 5 as used in Chapter 4.

Construct	Colony description	Plating dilution	Total No. of colonies	No. of fluorescent colonies	% Fluorescent colonies
pWB100	RFCS=5 RFI=0	1: 1000	480	0	0
		1:10000	108	0	0
pNM3	RFCS=4 RFI=4.5	1: 1000	TNTC	All	100%
		1:10000	152	All	100%
pNM4	RFCS=3.75 RFI=4.75	1: 1000	TNTC	All	100%
		1:10000	364	All	100%
pNM3.1	RFCS=3.75 RFI=3.75	1: 1000	TNTC	All	100%
		1:10000	120	116	97%
pNM3.2	RFCS= 3.5 RFI=4	1: 1000	TNTC	Most	+-80%
		1:10000	320	33	10%
pNM4.1	RFCS=3.25 RFI=4.25	1: 1000	TNTC	Most	+-80%
		1:10000	480	240	50%
pNM4.2	RFCS=2.75 RFI=4.5	1: 1000	336	2	0.6%
		1:10000	NO colonies	NO colonies	N/a

Although the colony numbers were not ideal, this experiment provided an insightful trend with respect to fluorescence intensity, colony size and instability. From the above table, it can be seen that the Phase 1 constructs, pNM3 and pNM4 appeared 100% stable at the stage the glycerol stocks were made. However, the Phase 2 constructs (pNM3.1, pNM3.2, pNM4.1 and pNM4.2) did display enhanced instability as seen from the high incidence of non-fluorescent colonies and small colony size of those that did fluoresce. Although pNM3.1 did display a very low percentage of instability, it was the most stable of all the Phase 2 constructs at this stage. Additionally, those constructs containing TetR r1.7 in the opposite direction to GFP (orientation 2) seemed even more unstable than those containing TetR r1.7 in the same direction as GFP (orientation 1).

5.6. Discussion

We have successfully developed a mycobacterial episomal revTetR regulatable shuttle vector, pNM3.1, in which the reporter GFP can be repressed by ATc in a dose dependent manner, and induced upon its removal. In both the pilot and proof of

concept study, approximately a 3-fold repression (~60%) of GFP expression was achieved for *rM. smegmatis*[pNM3.1] in the presence of 100ng/ml ATc after 12 hours of culturing. Since the level of repression achieved with the various concentrations of ATc (25ng/ml, 50ng/l and 100ng/ml) were very similar at 8 and 12 hours of culturing, this suggests that the maximum amount of repression achievable for the various ATc concentrations occurred by 8 hours. Upon removal of ATc induction occurred in a concentration dependent manner, such that the culture that contained only 25ng/ml of ATc achieved 92% induction of GFP expression after 12 hours, whereas the culture that contained 100ng/ml ATc was only induced to 68% GFP activity. Although complete induction was not achieved during the 12 hour assay procedure, it could probably have been observed at a later time point, suggesting that induction by washing to remove ATc was not as efficient as repression by the addition of ATc to the media.

Once again, as seen in the ELISA assays, only very low levels of GFP expression were recorded in the whole cell fluorimetry assays for r*M. smegmatis*[pNM3.2], r*M. smegmatis*[pNM4.1] and r*M. smegmatis*[pNM4.2] contradicting the high GFP expression levels observed following plating of these electrotransformed colonies. Since we proposed that these contradictory results could be due to instability of the shuttle vector constructs during culture, and since the recombinant *M. smegmatis* cultures used for both the ELISA and the fluorimetry experiments were started from the same batch of glycerol stocks prepared from a single selected recombinant colony grown in the absence of ATc, we sought to determine the stability status of these glycerol stocks used to start these cultures from.

Plating of these glycerol stocks confirmed that r*M. smegmatis*[pNM4.2] was the strongest fluorescer followed by r*M. smegmatis*[pNM4.1], r*M. smegmatis*[pNM3.2] and lastly the weakest of the Phase two recombinants r*M. smegmatis*[pNM3.1], and that the stronger the fluorescence the smaller the colony sizes. Furthermore, this experiment also revealed a trend that the higher the GFP expression levels and smaller or less healthy the fluorescent colonies, the larger the percentage of non-fluorescent colonies present, such that plates of r*M. smegmatis*[pNM4.2] colonies contained a very low percentage of fluorescent colonies (0.6% fluorescent colonies), while r*M. smegmatis*[pNM3.1] contained only a very few non-fluorescent colonies

(97-100% fluorescent colonies). The loss of fluorescence of colonies plated on selective media indicated that, while the kanamycin gene was intact, plasmid integrity was compromised such that these non-fluorescent colonies either contained a deleted or truncated non-functional GFP gene. Thus, this plasmid stability experiment served to corroborate the theory that genetic instability was indeed the cause of such low levels of GFP expression seen in the ELISA and fluorimetry assays for the "stronger expressers". Since this instability was already evident in the glycerol stocks, and since several additional culture generations would have followed prior to the assays being performed, a much faster growing 'fitter' mutant, in which the GFP gene had been lost or altered, could have dominated these cultures resulting in the unexpectedly low levels of GFP expression seen in the absence of ATc. Since the appearance of a mutant is a random event, different rates of instability may also explain the discrepancy in GFP levels observed between r*M. smegmatis*[pNM3] and its phase two counterpart r*M. smegmatis*[pNM3.1] in the absence of ATc in the pilot study, as well as the increased GFP expression observed for r*M. smegmatis*[pNM3] in the presence of 100ng/ml ATc as compared to no ATc present. From the MIC studies it was shown that 100ng/ml ATc did not affect the growth rate of *M. smegmatis*, although it should be noted that these experiments were not performed on recombinant cultures expressing GFP.

Chapter 6:General Discussion and Conclusions

The development of a successful episomal revTetR controlled expression system for application to vaccine design will depend on optimising a number of different factors. Firstly, in the induced state (absence of ATc) high levels of antigen expression would be desirable in order to promote a strong immune response. We chose to design our antigen expression cassette using the P_{smyc} promoter from *M. smegmatis* since it was found to drive even stronger expression than the *hsp*60 promoter and was found to be equally strong in BCG[61]. We confirmed that stronger expression could be achieved if the full-length transcript region was included in the promoter designs i.e. pNM1 (full-length) versus pNM2 (truncated) and thus based the subsequent promoter designs on the full-length sequence. Since alterations of sequences within the promoter region can dramatically affect promoter activity, it was also imperative that the introduction of the *tet* operator sequences did not decrease promoter activity to a large extent. We showed that addition of either two 19bp *tet*O1 (pNM4) or *tet*O2 (pNM3) operator sequence elements on either side of the putative -35 region lead to only a marginal decrease in GFP expression compared to the control pNM1 (full-length without any operator sites) as monitored by the intensity of fluorescence of r*M. smegmatis* transformants. However, in contrast the construct pNM5 in which chimeric operator sites were present (*tet*O1 and *tet*O2) resulted in GFP levels lower than all the other constructs containing a P_{smyc} promoter including pNM2 (truncated), and thus lead to its exclusion from further development of Phase 2 constructs.

The success of the system will also depend on the efficient binding of the repressor protein TetR r1.7 to the operator sites placed within the modified P_{smyc} promoter in the presence of ATc, thereby repressing expression of the selected antigen, as well as the efficient release from the operator sites upon removal of ATc. This is likely to depend on the amount of TetR r1.7 produced, its stability or half-life, its affinity for the specific operator site in the presence of ATc and the concentration of ATc used. Thus, in order to achieve high levels of stable protein, we designed an expression cassette constitutively expressing a codon optimised TetR r1.7 protein from the

*hsp*60 promoter and using a strong RBS terminated by the *hsp*60 transcriptional terminator.

Thus, we have constructed and explored the potential of four different episomal shuttle vector designs to repress the expression of the model antigen GFP in the presence of ATc and induce its expression in the absence of ATc, within the model organism *M. smegmatis*. The four designs (pNM3.1, pNM3.2, pNM4.1 and pNM4.2) included the full-length P_{smyc} promoter containing either the *tet*O1 or *tet*O2 operators inserted and the TetR r1.7 expression cassette in either the forward or reverse orientation with respect to the GFP gene on the episomal plasmid.

We observed from the fluorescent intensities of transformed colonies that the level of GFP expression was not only affected by the strength of the modified P_{smyc} promoter design, in which the insertion of two *tet*O1 sites (pNM4) resulted in slightly lower fluorescence than two *tet*O2 sites (pNM3), but more significantly by the orientation of insertion of the TetR r1.7 expression cassette. r*M. smegmatis* colonies that contained the cassette in the opposite orientation (pNM3.2 and pNM4.2) to GFP fluoresced noticeably stronger than those that contained the cassette in the same orientation (pNM3.1 and pNM4.1). In orientation 1 (same orientation as GFP), the expression of TetR r.1.7 somehow affected the expression of GFP. This could be at the level of transcription or translation since the TetR r1.7 and the GFP expression cassettes lie adjacent to each other in close proximity. It is possible that during transcription or translation initiation of TetR r.17 that transcription initiation from the nearby P_{smyc} promoter region is negatively affected thus, resulting in lower GFP expression levels. This serves to emphasise the importance of testing theoretical designs in experiments.

Furthermore we observed that the stronger the Phase 2 constructs fluoresced, the smaller the colonies were such that pNM4.2 was the strongest expresser and formed the smallest colonies, followed by pNM4.1, pNM3.2 and pNM3.1. This is indicative of the higher the level of antigen expression the slower the growth rate of the recombinant due to a toxic or metabolic burden on the cell. Mutants arise spontaneously, but unless they have a growth rate advantage will go undetected. But as soon as a mutant arises that can grow faster than the rest of the recombinant

population, it will out compete other species in the culture or colony and take over such that this will then become the dominant mutant species. Thus, it follows that the slower the growth rate of the recombinant, the higher the rate of instability. This was indeed confirmed by plating of the glycerol stocks of r*M. smegmatis* that had been grown up as seed stocks in the absence of ATc. Using the equation: ln OD_{600} (final) – OD_{600} (initial) divided by ln2, we estimate that the stored starter cultures would have gone through approximately 8-10 generations from the time the colony was picked off the plate through a sub-inoculation step to remove colony debris to the preparation of the final starter culture. We found that as for the transformed colonies, the higher the level of GFP expression the smaller the colonies but also the higher the % of faster growing large non fluorescent colonies present on the plates, especially for the plated glycerol stocks of r*M. smegmatis*[pNM4.2], [pNM4.1] and [pNM3.2]. This explains the inverse relationship observed between the higher GFP expressing smaller transformants, and the unexpectedly low levels of GFP expression measured in the ELISA and fluorimetry assays for r*M. smegmatis* [pNM4.2], [pNM4.1] and [pNM3.2]. Since it is estimated that the starter cultures would have undergone a further 5 or more generations before the ELISA or fluorimetry assays were performed, this would have allowed further opportunity for faster growing non-fluorescent mutants to dominate the culture.

However, the Phase 2 fluorescer r*M. smegmatis* [pNM3.1] which represents the weakest of the four Phase 2 constructs on plates, yielded the highest levels of GFP expression in the proof of concept fluorimetry assay of ~100RFUs. This represented ~20RFUs less than that of the r*M. smegmatis* [pNM3] control. This sample (r*M. smegmatis* [pNM3.1]) only resulted in between 0 to 3% of non-fluorescent colonies when the glycerol stocks were plated i.e. it appeared to be the most stable of the four Phase 2 constructs. Furthermore, we were able to achieve a maximum of 60% repression with this system in the presence of 100ng/ml ATc after 12 hours of culturing. These results are very encouraging considering that this assumes that there is no instability occurring in the culture that does not contain ATc. If on the other hand instability has occurred during culturing in the absence of ATc, then our repression ratio would be even better than reported. In conclusion we have successfully developed a single episomal shuttle vector, pNM3.1, that in the model organism *M. smegmatis*, can effectively down regulate GFP expression in the

presence of ATc and upregulate GFP expression upon the removal of ATc in a dose-dependent manner.

Using the same TetR r1.7 reverse mutant, Guo *et al*., 2007[82] reported a maximum of 17-fold repression of β-galactosidase activity with a single copy integrated promoter-operator-reporter construct (containing two *tet*O2 sites flanking the –35 P_{smyc} promoter region) designed for gene silencing in *M. smegmatis*. Similarly, Kamionka *et al*., 2005[79] reported a 99.8% repression of β-galactosidase activity with their single-copy integrated promoter-operator-reporter construct (containing two *tet*O1 sites flanking the -10 promoter region), and a maximum induction factor of 500. In both these systems the TetR r1.7 regulator is expressed from a multicopy episomal plasmid, whereas the promoter-operator-reporter construct is a single copy integrated into the chromosome. In these studies the aims of researchers was to repress gene expression for the generation of conditional knockouts for purposes of functional gene analysis. The balance between elevated levels of gene expression and genetic stability is our challenge. Our design based rather on an episomal system was chosen since these plasmids are maintained at approximately five copies per cell, and so the level of antigen expression in the induced state will be higher than a single copy of the gene integrated into the chromosome. Secondly, we have established from other vaccine designs that we do not require complete shut off to ensure stability. Supporting data from a study using the *mtr*A promoter from *M. tuberculosis*, which results in low levels of antigen expression during *in vitro* growth but which is induced upon infection of macrophages (unpublished data generated from our laboratory), shows that BCG can tolerate a certain amount of antigen expression without resulting in instability over 30 culture generations. This represents a theoretical value calculated to generate sufficient vaccine to vaccinate a population.

The question we need to answer is whether the repression we have achieved with our newly developed shuttle vector pNM3.1 system is enough to maintain stability over a number of culture generations in the presence of ATc. Additionally, the concentration of ATc can be further explored, especially the 50-100ng/ml range. The ideal concentration of the co-repressor, ATc should be as low as possible to make its

removal from the culture easier and thus induction quicker but high enough to ensure stability. This will be the focus of ongoing research in our lab.

The reporter system used is also important. Whilst GFP represents a convenient reporter, it is very stable with a long half-life.Thus, it would be better to use a less stable GFP reporter gene containing tags targeting them for degradation by cellular proteases, as did Carroll *et al*., 2005[98] to more accurately measure real time levels of protein expression.

This study highlights the importance of considering instability of heterologously expressed proteins in mycobacteria and the critical importance of developing a stable expression system in BCG. This is very often underestimated and overlooked by researchers trying to develop BCG as a vaccine vehicle. Future research will be focussed on further improving our approach. In order to minimise instability, ATc will be included in all the transformation expression mixes, recombinant transformants will be plated on media containing ATc and starter cultures grown in the presence of ATc to ensure maximum stability. Stability will be continuously monitored by plating samples of the recombinant cultures, visualising colony sizes and gauging the loss of fluorescence intensity with respect to the appropriate controls. Additionally plasmid DNA will be extracted, transformed into *E. coli* and plasmid mapping performed with the isolated DNA.

However, since the metabolic burden observed is a cumulative effect of both high GFP expression and constitutive TetR r1.7 expression it is possible that, even if the presence of ATc from the start enhances stability of all the shuttle vectors tested, shuttle vector pNM3.1 rather than the higher expressers pNM3.2, pNM4.1 or pNM4.2 will still be more desirable as it represents a compromise between high levels of heterologous gene expression and genetic stability in the absence of ATc. It would not be desirable for the vaccine shuttle vector to be mutated or lost upon or soon after vaccination of the host before sufficient antigen is produced and before an adequate immune response can be mounted. Even though TetR r1.7 is of bacterial origin and codon optimised for expression in mycobacteria, the decrease in colony sizes observed upon introduction of the TetR r1.7 cassette into the various shuttle vectors suggests its expression leads to increased metabolic load. Klotzsche *et al*.,

2009[56] presented the use of improved revTetRs, managing to repress β-galactosidase activity to <2% which would serve to improve repression in our system as well. Additionally, since the *hsp*60 promoter is itself a stress induced promoter and has been associated with instability[51, 52], we could test out other more suitable promoters that are of medium strength as well.

This system that we developed with the model antigen GFP, in the model organism *M. smegmatis* needs to be tested in the vaccine vector BCG panCD. A number of features of BCG panCD make it an attractive vector for the development of rBCG HIV vaccines. Due to the limited availability of pantothenate in macrophages (*in vivo*) the ability for this strain to replicate within the host or vaccinee is severely impaired making it a much safer strain for use in immunocompromised individuals than the current BCG strains. Furthermore, data generated in our laboratory suggest that BCG panCD expressing HIV Gag fused to signal sequences induces a better CD8+ T cell response to foreign proteins than BCG Pasteur. It has been hypothesised that the compromised ability of this auxotrophic strain to produce fatty acids has resulted in a less-complex more permeable cell wall allowing for better antigen export, cross presentation and thus immune responses. It has also been suggested that these enhanced immune responses could be attributed to the fact that BCG panCD is unable to manufacture fatty acids or polyketides that play a role in host immune suppression during mycobacterial invasion (personal communication with Dr. W.R Jacobs). In both cases BCG panCD represents a better strain for the development of a vaccine vector in which the expression of heterologous antigens are controlled by the presence or absence of exogenous ATc. It should be recounted that the MIC of ATc to BCG panCD was significantly lower than that of BCG Pasteur, which could be justified by the theory that the panCD strain possesses a "less-complex" cell wall resulting in ATc being able to cross the cell wall quicker, yet another advantage of using this strain for tetracycline controlled expression.

This project provides encouraging preliminary results that will provide a guideline for further work in expressing stable foreign genes in BCG panCD and finally the GFP gene will be replaced with our gene of interest, an HIV gene/s.

Reference List

1. WHO report 2009 Global Tuberculosis control. 12-15. 2009. 6-9-0090. Ref Type: Report

2. Stover,C.K. *et al.* Protective immunity elicited by Recombinant Bacille Calmette-Guerin expressing outer surface protein A (ospA) lipoprotein: A candidate Lyme disease vaccine. *Journal of Experimental Medicine* **178**, 197-209 (1993).

3. Gheorghiu,M., Lagrange,P.H., & Fillastre,C. The stabiltiy and immunogenicity of a dispersed-grown freeze-dried Pasteur BCG vaccine. *J Biol Stand* **16**, 15-26 (1988).

4. Kawahara,M. *et al.* Combined intrarectal/intradermal inoculation of rBCG induces enhanced immune responses against the inserted HIV-1 V3 antigen. *Vaccine* **21**, 158-166 (2002).

5. Ohara,N. & Yamada,T. Recombinant BCG vaccines. *Vaccine* **19**, 4089-4098 (2001).

6. Wada,N. *et al.* Long-lasting immune response induced by recombinant bacillus Calmette-Guerin (BCG) secretion system. *ScadanavianJ ournal of Immunology* **43**, 202-209 (1996).

7. Kauffman,S.H.E. Envisioning future strategies for vaccination against tuberculosis. *Nature Reviews Immunology* **6**, 699-704 (2006).

8. Lazarevic,V. & Flynn,J. CD8+ T cells in TB. *Am J Respir Crit Care Med* **166**, 1116-1121 (2002).

9. Cassatt,D.R. *et al.* Protection of mice against tetanus vaccine based on rBCG. *Vaccines* **93**, 385-389 (2000).

10. Langermann,S. *et al.* Protective Humoral response against pneumococcal infection in mice elicited by rBCG vaccines expressing pneumococcal surface protein A. *Journal of Experimental Medicine* **180**, 2277-2286 (1994).

11. Matsumotomo,S., Yukitake,H., Kanbara,H., & Yamada,T. Recombinant *Mycobacterium bovis* bacillus Calmette-Guerin secreting merozoite surface protein 1 (MSP1) induces protetion against rodent malaria parasite infection depending on MSP1-stimuated IFNg and parasite specific antibodies. *Journal of Experimental Medicine* **188**, 845-854 (1998).

12. Abdelhak,S., Louzir,H., & Timm,J. Recombinant BCG expressing the leishmania surface anitgen Gp63 induces protective immunity agianst *Leishmania major* infection in BALB/c mice. *Microbiology* **141**, 1585-1592 (1995).

13. Streit,J.A., Recker,T.J., Donelson,J.E., & Wilson,M.E. rBCG expressing LCR1 of *Leishmania chagasi* induces protective immunity in susceptible mice. *Experimental Parasitology* **94**, 33-41 (2000).

14. Kremer,L., Riveau,G., Baulard,A., Capron,A., & Locht,C. Neutralising antibody responses elicited in mice immunised with rBCG producing the

Schistosoma mansoni glutathione S-transferase. *Journal of Immunology* **156**, 4309-4317 (1996).

15. Supply,P., Sutton,P., Coughlan,S.N., & et al Immunogenicity of rBCG producing GRA1 antigen from *Toxoplasma gondii*. *Vaccine* **17**, 705-714 (1999).

16. Jabbar,I.A. *et al.* Immune responses induced by BCG recombinant for human papillomavirus L1 and E7 proteins. *Vaccine* **18**, 2444-2453 (2000).

17. Leung,N.J. *et al.* The Kinetics of Specific Immune Responses in Rhesus Mokeys Inoculated with Live Recombinant BCG Expressing SIV Gag,Pol, Env and Nef Proteins. *Virology* **268**, 94-103 (2000).

18. Zhu,Y.D. *et al.* rBCG expressing the measles virus nucleo-protein protects infant rhesus macaques from measles virus pneumonia. *Journal of infectious disease* **176**, 1445-1453 (1997).

19. Goulder P. & Watkins D. HIV and SIV CTL escape: Implications for vaccine design. *Nature* **4**, 630-640 (2004).

20. Kuroda,M.J. *et al.* Emergence of CTL coincides with clearance of virus during primary simian immunodeficiency virus infection in rhesus monkeys. *Journal of Immunology* **162**, 5127-5133 (1999).

21. Ogg,G.S., Jin,X., Bonhoeffer,S., & et al Quantitation of HIV-1 specific cytotoxic T lymphocytes and plasma load of viral RNA. *Science* **279**, 2103-2106 (1998).

22. Jacobs,W.R., Tuckman,M., & Bloom,B.R. Introduction of foreign DNA into mycobacteria using a shutle plasmid. *Nature* **327**, 532-535 (1987).

23. Matsuo,K., Yamaguchi,R., & Yamazaki,A. Establishment of a foreign antigen secretion system in mycobacteria. *Infect Immun* **58**, 4049-4054 (1990).

24. Kameoka,M., Nishino,Y., Matsuo,K., & et al Cytotoxic T lymphocyte response in mice induced by a rBCG vaccination which produces an extracellular α antigen that fused with the HIV-1 envelope immunodominant domain in the V3 loop. *Vaccine* **12**, 153-158 (1994).

25. Stover,C.K. *et al.* New use of BCG for recombinant BCG vaccines. *Nature* **351**, 456-460 (1991).

26. Aldovini,A. & Young,R.A. Humoral and cell mediated immune responses to live recombinant BCG-HIV vaccines. *Nature* **351**, 479-482 (1991).

27. Winter,N., Lagranderie,M., Rauzier,J., & et al Expression of heterologous genes in BCG: induction of acellular response against HIV-1 nef protein. *Gene* **109**, 47-54 (1991).

28. Ami,Y. *et al.* Priming-Boosting vaccination with Recombinant *Mycobacterium bovis* Bacillus Calmette-Guerin and a Non-replicating Vaccinia Virus recombinant Leads to Long-Lasting and Effective Immunity. *Journal of Virology* **79**, 12871-12879 (2005).

29. Im,E.-J. *et al.* Vaccine Platform for Prevention of Tuberculosis and Mother-to-Child Transmisssion of Human Immunodeficiency Virus Type 1 through Breastfeeding. *Journal of Virology* **81**, 9408-9418 (2007).

30. Dennehy M., Bourn W., Steele D., & Williamson A.-L. Evaluation of recombinant BCG expressing rotavirus VP6 as an anti-rotavirus vaccine. *Vaccine* **25**, 3646-3657 (2007).

31. Connell,N.D. Mycobacterium: Isolation, Maintenance, Transformation and Mutant Selection in *Methods in Cell Biology* 107 (Academic Press, 1994).

32. Brosch,R. *et al.* Genome plasticity of BCG and impact on vaccine efficacy. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 5596-5601 (2007).

33. Behr,M. *et al.* Comaparative Genomics of BCG Vaccines by Whole-Genome DNA Microarray. 248, 1520-1523. 1999. Ref Type: Report

34. Dennehy M. & Williamson A.-L. Factors influencing the immune response to foreign antigen expressed in recombinant BCG vaccines. *Vaccine* **23**, 1209-1224 (2005).

35. Lim,E.M. *et al.* Recombinant Mycobacterium bovis BCG producing the N-terminal half of SIVmac251 Env antigen induces neutralizing antibodies and cytotoxic T lymphocyte responses in mice and guinea pigs. *AIDS Res. Hum. Retroviruses* **13**, 1573-1581 (1997).

36. Fuerst,M., de la Cruz,V.F., Bansal,G.P., & Stover,C.K. Development and analysis of recombinant BCG vector systems. *AIDS Research and Human Retroviruses* **8**, 1451-1455 (1992).

37. Glick,B.R. Metabolic load and heterologous gene expression. *Biotechnol. Adv* **13**, 247-261 (1995).

38. Smith,M.A., Bidochka,M., & J. Bacterial Fitness and plasmid loss: The importance of culture conditions and plasmid size. *Canadian Journal of Microbiology* **44**, 351-355 (1998).

39. Corchero,J.L. & Villaverde,A. Plasmid maintenance in *E. coli* recombinant cultures is dramatically and specifically influenced by features of the encoded protein. *Biotechnology Bioengineering* **58**, 625-632 (1998).

40. Zahn,K. Overexpression of an mRNA dependent on rare codons inhibits protein synthesis and cell growth. *Journal of Bacteriology* **178**, 2926-2933 (1996).

41. Brinkmann,U. High-level expression of recombinant genes in *Escherichia coli* is dependent on the availability of the dnaY gene product. *Gene* **85**, 109-114 (1989).

42. Harcum,S.W. & Bentley,W.E. Heat shock and stringent responses have overlapping protease activity in recombinant *Escherichia coli*: implications for heterologous protein yield. *Appl. Microbiol. Biotechnol.* **80**, 23-27 (1999).

43. Cserjan-Puschmann,M. & et al Metabolic approaches for the optimisation of recombinant fermentation processes. *Appl. Microbiol. Biotechnol.* **53**, 43-50 (1999).

44. Langermann,S. *et al.* Protective Humoral response against Pneumococcal infection in mice elicited by recombinant Bacille Calmette-Guerin Vaccines expressing Pneumococcal surface protein A. *Journal of Experimental Medicine* **180**, 2277-86 (1994).

45. Galen,J. & Levine,M. Can a "flawless" live vector vaccine strain be engineered? *TRENDS in Microbiology* **9**, 372-376 (2001).

46. Bourn W. Candidate HIV-1 Vaccines based on Recombinant Bacterial Delivery systems, Unpublished. 2005.
Ref Type: Report

47. Labidi,A., Estes,R., David,H., & Bollon,A. Mycobacterium Recombinant Vaccines. *La Tunisie Medicale* **79**, 65-81 (2001).

48. Bourn W. *et al.* Creation and characterisation of a high-copy-number version of the pAL5000 mycobacterial replicon. *Tuberculosis* **87**, 481-488 (2007).

49. Labidi,A., Mardis,E., Roe,B.A., & Wallace,F.J. Cloning and DNA sequence of the *Mycobacterium fortuitum* var *fortuitum* plasmid pAL5000. *Plasmid* **27**, 130-140 (1992).

50. Young,D.B., Lathigra,R., Hendrix,R., Sweetster,D., & Young,R.A. Stress proteins are immune targets in leprosy ad tuberculosis. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 4267-4270 (1988).

51. Zarouni,A.L. & Dale,J.W. Expression of foreign genes in Mycobacterium bovis BCG starins using different promoters reveals instabiltiy of hsp60 promoter for expression of foreign genes in BCG starins. *Tuberculosis* **82**, 283-291 (2002).

52. Haeseleer,F., Pollet,J.F., Haumont,M., Bollen,A., & Jacobs,P. Stable and integration and expression of the plasmodium falciparum circum-sporozoite protein coding sequence in mycobacteria. *Mol Biochem Parasitol* **57**, 117-126 (1993).

53. Ehrt,S. *et al.* Controlling gene expression in mycobacteria with anhydrotetracycline and Tet repressor. *Nucleic Acids Research* **33**, (2005).

54. Kanekiyo,M. *et al.* Mycobacterial Codon Optimisation Enhances Antigen Expression and Virus-Specific Immune responses in Recombinant Mycobacterium bovis Bacille Calmette-Guerin Expressing Human Immunodeficiency Virus Type 1 Gag. *Journal of Virology* **79**, 8716-8723 (2005).

55. Connell,N.D., Medina-Acosta,E., Mccaster,W.R., & Bloom,B.R. Effective Immunisation against cutaneous Leishmaniasis with recombinant bacille Calmette-Guerin expressing the Leishmania surface proteinase gp63. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 11473-11477 (1993).

56. Klotzsche,M., Ehrt,S., & Schnappinger,D. Improved tetracycline repressors for gene silencing in mycobacteria. *Nucleic Acids Res.* **37**, 1778-1788 (2009).

57. Dellagostin,O.A., Esposito,G., Eales,L.J., Dale,J.W., & McFadden,J. Activity of mycobacterial promoters during intracellular and extracellular growth. *Microbiology* **141**, 1785-1792 (1995).

58. Dhandayuthapani,S., Via,L.E., Thomas,C.A., & Horrowitz,C.A. Green fluorescent protein as a marker for gene expression and cell biology of mycobacterial interactions with macrophages. *Molecular Microbiology* **15**, 901-912 (1995).

59. Ramos,J.L. *et al.* The TetR family of Transcriptional Repressors. *Microbiology and Molecular Biology Reviews* **69**, 326-356 (2005).

60. Freifelder,D. Regulation of the Activity of Genes and of Gene Products in Prokaryotes in *Molecular Biology: A comprehensive introduction to Prokaryotes and Eukaryotes* 540-595 (Jones and Bartlett, 1983).

61. Ehrt,S. & Schnappinger,D. Controlling gene expression in mycobacteria. *Future Microbiology* **1**, 177-184 (2006).

62. Kataoka,M. *et al.* Development of a temperature-inducible expression system for Streptomyces spp. *J. Bacteriol.* **178**, 5540-5542 (1996).

63. Lim,A., Boon,C., & Dick,T. Inducibility of the Streptomyces traRts107-Ptra expression cassette in Mycobacterium smegmatis. *Biological Chemistry* **381**, 517-519 (2000).

64. Draper,P. The aliphatic acylamide aminohydrolase of Mycobacterium smegmatis:its inducible nature and relation to acyl-transfer to hydroxylamine. *Journal of General Microbiology* **46**, 111-123 (1967).

65. Parish,T., Mahenthiralingam,E., Draper,P., Davis,E.O., & Colston,M.J. Regulation of the inducible acetamidase gene of Mycobacterium smegmatis. *Microbiology-Uk* **143**, 2267-2276 (1997).

66. Mahenthiralingam,E., Draper,P., Davis,E.O., & Colston,M.J. Cloning and Sequencing of the Gene Which Encodes the Highly Inducible Acetamidase of Mycobacterium-Smegmatis. *Journal of General Microbiology* **139**, 575-583 (1993).

67. Parish,T., Turner,J., & Stoker,N.G. amiA is a negative regulator of acetmidase expression in *Mycobacterium Smegmatis*. *Bmc Biotechnology* **1**, (2001).

68. Subbian,S. & Narayanan,S. Identification and characterisation of the regulatory elements of the inducible acetamidase operon from Mycobacterium smegmatis. *Canadian Journal of Microbiology* **53**, 599-606 (2007).

69. Daugelat,S., Kowall,J., Mattow,J., & et al The RD1 proteins of Mycobacterium tuberculosis: expression in M. smegmatis and biochemical chracterisation. *Microbes Infect.* **5**, 1082-1095 (2003).

70. Triccas,J.A., Parish,T., Britton,W.J., & Gicquel,B. An inducible expression system permitting the efficient purification of a recombinant antigen from *M. smegmatis*. *FEMS Microbiol. Lett.* **167**, 151-1156 (1998).

71. Hillen,W. & Berens,C. Mechanisms underlying expression of TN*10* encoded Tetracyline resistance. *Annual Reviews in Microbiology* **48**, 345-369 (1994).

72. Saenger,W., Orth,P., Kisker,C., Hillen,W., & Hinrichs,W. The Tetracycline Repressor-A Paradigm for a Biological Switch. *Angew. Chem. Int. Ed Engl.* **39**, 2042-2052 (2000).

73. Tauch,A., Puhler,A., Kalinowski,J., & Thierbach,G. TetZ, a new tetracycline resistance determinant discovered in Gram-positive bacteria, shows high homology to Gram-negative regulated efflux systems. *Plasmid* **44**, 285-291 (2000).

74. Rossi,F.M.V. & Blau,H.M. Recent advances in inducible gene expression systems. *Current Opinion in Biotechnology* **9**, 451-456 (1998).

75. Berens,C. & Hillen,W. GENE REGULATION BY TETRACYCLINES in *Genetic engineering* (ed. Kluwer Academic/Plenum Publishers) 255-277 2004).

76. Berens,C. & Hillen,W. Gene regulation by tetracyclines. *European Journal of Biochemistry* **270**, 3109-3121 (2003).

77. Henssler,E.M., Scholz,O., Lochner,S., Gmeiner,P., & Hillen,W. Structure-based design of Tet repressor to optimize a new inducer specificity. *Biochemistry* **43**, 9512-9518 (2004).

78. Gossen,M. & Bujard,H. Anhydrotetracycline, a novel effector for tetracycline controlled gene expression systems in eukaryotic cells. *Nucleic Acids Res.* **21**, 4411-4412 (1993).

79. Kamionka,A., Bertram,R., & Hillen,W. Tetracycline-dependent Conditional Gene Knockout in *Bacillus subtilis*. *Applied and Environmental Microbiology* **71**, 728-733 (2004).

80. Blokpoel,M.C.J. *et al.* Tetracycline-inducible gene regulation in mycobacteria. *Nucleic Acids Res.* **33**, (2005).

81. Carroll,P., Muttucumaru,D.G.N., & Parish,T. Use of a Tetracycline-Inducible System for Conditional Expression in *Mycobacterium tuberculosis* and *Mycobacterium Smegmatis*. *Applied and Environmental Microbiology* **71**, 3077-3048 (2004).

82. Guo,X.V. *et al.* Silencing essential protein secretion in Mycobacterium smegmatis by using tetracycline repressors. *Journal of Bacteriology* **189**, 4614-4623 (2007).

83. Schnappinger,D., Schubert,P., Pfleiderer,K., & Hillen,W. Determinants of protein-protein recognition by four helix bundles: changing the dimerization specificity of Tet repressor. *Embo Journal* **17**, 535-543 (1998).

84. Scholz,O. *et al.* Activity reversal of Tet repressor caused by single amino acid exchanges. *Mol. Microbiol.* **53**, 777-789 (2004).

85. Schlegel,H.G. The grouping of prokaryotes in *General Microbiology* (ed. Schlegel,H.G.) 103-142 (Cambridge University Press, Cambridge, 1996).

86. Chung,C.T. & Miller,R.H. A rapid and convenient method for the preparation and storage of competent bacterial cells. *Nucleic Acids Res.* **16**, 3580 (1988).

87. Sambrook,J. & Russell,D.W. *Molecular cloning:A laboratory Manual, 3rd edition*(Cold Spring Harbour Laboratory Press,2001).

88. Ish-Horowicz,D. & Burke,J.F. Rapid and efficient cosmid cloning. *Nucleic Acids Res.* **9**, 2989-2998 (1981).

89. Parish,T. & Stoker,N.G. *Mycobacterium tuberculosis Protocols*(Humana Press,2001).

90. Hesseling,A.C. *et al.* Bacille Calmette-Guerin Vaccine -induced disease in HIV-infected and HIV uninfected individuals. *Clinical Infectious diseases* **42**, 548-558 (2006).

91. Reza,M.S., Khotai,G., Mamishi,S., Kheirkhah,M., & Parvaneh,N. Disseminated Bacillus Calmette-Guerin infection after BCG vaccination. *J Trop Paediatr.* **54**, 413-416 (2008).

92. World Health Organisation Revised BCG vaccination guidelines for infants at risk of HIV infection. *Wkly. Epidemiol. Rec* **82**, 193-196 (2007).

93. Bardarov,S. *et al.* Specialised transduction: an efficient method for generating marked and unmarked targeted gene disruptions in *Mycobacterium tuberculosis*, *M. bovis* BCG and *M. smegmatis*. *Microbiology-Uk* **148**, 3007-3017 (2002).

94. Sambandamurthy,V.K. *et al.* A pantothenate auxotroph of *Mycobacterium tuberculosis* is highly attenuated and protects mice against tuberculosis. *Nature* **8**, 1171-1174 (2002).

95. Jacobs,W.R. *et al. Mycobacterium tuberculosis RD1 panCD*: A safe and limited replicating mutant strain that protects immunocompetent and immunocompromised mice against experimental tuberculosis. *Vaccine*(2006).

96. Tullius,M.V., Harth,G., Maslesa-Galic,S., Dillon,B.J., & Horwitz,M.A. A replication-limited recombinant Mycobacterium bovis BCG vaccine against Tuberculosis designed for Human Immunodeficiency Virus-positive persons is safer and more efficacious than BCG. *Infect Immun* **76**, 5200-5214 (2008).

97. Gossen,M. & Bujard,H. Anhydrotetracycline, a novel effector for tetracycline controlled gene expression systems in eukaryotic cells. *Nucleic Acids Res.* **21**, 4411-4412 (1993).

98. Carroll,P., Muttucumaru,D.G.N., & Parish,T. Use of a Tetracycline-Inducible System for Conditional Expression in *Mycobacterium tuberculosis* and

Mycobacterium Smegmatis. *Applied and Environmental Microbiology* **71**, 3077-3048 (2005).

99. Ranes,M.G., Rauzier,J., Lagranderie,M., Gheorghiu,M., & Gicquel,B. Functional analysis of pAL5000, a plasmid from *Mycobacterium fortuitum* : construction of a"mini" mycobacterium-*Escheria coli* shuttle vector. *Journal of Bacteriology* **172**, 2793-2797 (1990).

100. Bukau,B. & Horwich,A.L. The Hsp60 and Hsp70 Chaperone machines. *Cell* **92**, 351-366 (1998).

101. Edelman,R. *et al.* safety and immunogenicity of rBCG expressing Borrelia burgdorferi outer surface protein A (OspA) lipoprotein in adult volunteers: a candidate for Lyme disease Vaccine. *Vaccine* **17**, 904-914 (2009).

102. Kamionka,A., Bertram,R., & Hillen,W. Tetracycline-dependent Conditional Gene Knockout in *Bacillus subtilis*. *Applied and Environmental Microbiology* **71**, 728-733 (2005).

103. HIllmann,D., Eschenbacher,I., Thiel,A., & Niederweis,M. Expression of the Major porin Gene mspA Is regulated in *M. smegmatis*. *Journal of Bacteriology* **189**, 958-967 (2007).

104. Dhandayuthapani,S. *et al.* Green Fluorescent protein as a marker for gene expression and cell biology of mycobacterial interactions with macrophages. *Molecular Microbiology* **17**, 901-912 (1995).

105. Gandotra,S., Schnappinger,D., Monteleone,M., Hillen,W., & Ehrt,S. In vivo gene silencing identifies the mycobacterium tuberculosis proteosome as essential for the bacteria to persist in mice. *Nature Medicine* **13**, 1515-1520 (2007).

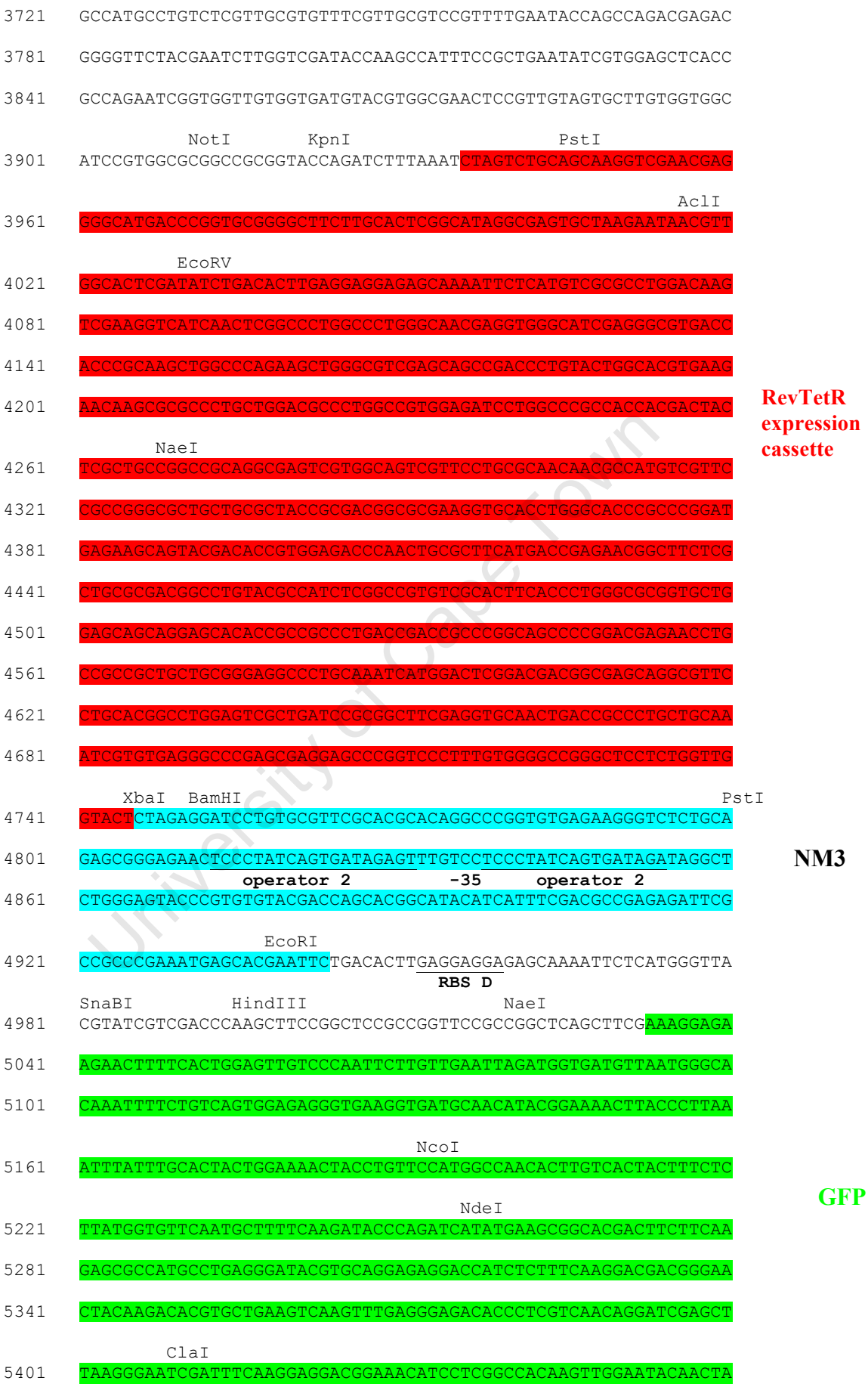
Appendix A

pNM3.1 Annotated sequence

Total base pairs: 5857bp

```
1       GCTAGCAACAAAGCGACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAA

61      ATATATCATCATGAACAATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTA

                                      XhoI
121     TGAGCCATATTCAACGGGAAACGTCTTGCTCGAGGCCGCGATTAAATTCCAACATGGATG

181     CTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAATCT

241     ATCGCTTGTATGGGAAGCCCCATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCG

301     TTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCTC

361     TTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGA

421     TCCCCGGGAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATATTG

481     TTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCCTT

541     TTAACAGCGATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTTGG

601     TTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAG

661     AAATGCATAATCTTTTGCCATTCTCACCGGATTCAGTCGTCACTCATGGTGATTTCTCAC

721     TTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCG

781     GAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTC

841     CTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAAT

901     TGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAATTGGTTGTAAC

961     ACTGGCAGAGCATTACGCTGACTTGACGGGACGGCGGCTTTGTTGAATAAATCGAACTTT

1021    TGCTGAGTTGAAGGATCAGATCACGCATCTTCCCGACAACGCAGACCGTTCCGTGGCAAA

1081    GCAAAAGTTCAAAATCACCAACTGGTCCACCTACAACAAAGCTCTCATCAACCGTGGCTC

1141    CCTCACTTTCTGGCTGGATGATGGGGCGATTCAGGCCTGGTATGAGTCAGCAACACCTTC

                        SpeI
1201    TTCACGAGGCAGACCTCACTAGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGG

1261    ATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACC

1321    GCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAC

1381    TGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCA

1441    CCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGT

1501    GGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACC

1561    GGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCG

1621    AACGACCTACACCGAACTGAGATACCTACAGCGTGAGCATTGAGAAAGCGCCACGCTTCC

1681    CGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCAC
```

```
1741  GAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCT

1801  CTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGC

1861  CAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTT

1921  TCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATAC

                                      MluI
1981  CGCTCGCCGCAGCCGAACGACCGAGCGCAACGCGTGAGCCCACCAGCTCCGTAAGTTCGG

2041  GTGCTGTGTGGCTCGTACCCGCGCATTCAGGCGGCAGGGGGTCTAACGGGTCTAAGGCGG

2101  CGTGTACGGCCGCCACAGCGGCTCTTAGCGGCCCGGAAACGTCCTCGAAACGACGCATGT

2161  GTTCCTCCTGGTTGGTACAGGTGGTTGGGGGTGCTCGGCTGTCGCTGGTGTTTCATCATC

2221  AGGGCTCGACGGGAGAGCGGGGGAGTGTGCAGTTGTGGGGTGGCCCCTCAGCGAAATATC

2281  TGACTTGGAGCTCGTGTCGGACCATACACCGGTGATTAATCGTGGTTTATTATCAAGCGT

2341  GAGCCACGTCGCCGACGAATTTGAGCAGCTCTGGCTGCCGTACTGGTCCCTGGCAAGCGA

        XhoI
2401  CGATCTGCTCGAGGGGATCTACCGCCAAAGCCGCGCGTCGGCCCTAGGCCGCCGGTACAT

2461  CGAGGCGAACCCAACAGCGCTGGCAAACCTGCTGGTCGTGGACGTAGACCATCCAGACGC

2521  AGCGCTCCGAGCGCTCAGCGCCCGGGGGTCCCATCCGCTGCCCAACGCGATCGTGGGCAA

2581  TCGCGCCAACGGCCACGCACACGCAGTGTGGGCACTCAACGCCCCTGTTCCACGCACCGA

2641  ATACGCGCGGCGTAAGCCGCTCGCATACATGGCGGCGTGCGCCGAAGGCCTTCGGCGCGC

2701  CGTCGATGGCGACCGCAGTTACTCAGGCCTCATGACCAAAAACCCCGGCCACATCGCCTG

2761  GGAAACGGAATGGCTCCACTCAGATCTCTACACACTCAGCCACATCGAGGCCGAGCTCGG

2821  CGCGAACATGCCACCGCCGCGCTGGCGTCAGCAGACCACGTACAAAGCGGCTCCGACGCC

2881  GCTAGGGCGGAATTGCGCACTGTTCGATTCCGTCAGGTTGTGGGCCTATCTTCCCGCCCT

2941  CATGCGGATCTACCTGCCGACCCGGAACGTGGACGGACTCGGCCGCGCGATCTATGCCGA

3001  GTGCCACGCGCGAAACGCCGAATTTCCGTGCAACGACGTGTGTCCCGGACCGCTACCGGA

3061  CAGCGAGGTCCGCGCCATCGCCAACAGCATTTGGCGTTGGATCACAACCAAGTCGCGCAT

3121  TTGGGCGGACGGGATCGTGGTCTACGAGGCCACACTCAGTGCGCGCCATGCGGCCATCTC

3181  GCGGAAGGGCGCAGCAGCGCGCACGGCGGCGAGCACAGTTGCGCGGCGCGCAAAGTCCGC

        NcoI
3241  GTCAGCCATGGAGGCATTGCTATGAGCGACGGCTACAGCGACGGCTACAGCGACGGCTAC

3301  AACTGGCAGCCGACTGTCCGCAAAAAGCGGCGCGTGACCGCCGCCGAAGGCGCTCGAATC

3361  ACCGGACTATCCGAACGCCACGTCGTCCGGCTCGTGGCGCAGGAACGCAGCGAGTGGTTC

3421  GCCGAGCAGGCTGCACGCCGCGAACGCATCCGCGCCTATCACGACGACGAGGGCCACTCT

3481  TGGCCGCAAACGGCCAAACATTTCGGGCTGCATCTGGACACCGTTAAGCGACTCGGCTAT

3541  CGGGCGAGGAAAGAGCGTGCGGCAGAACAGGAAGCGGCTCAAAAGGCCCACAACGAAGCC

3601  GACAATCCACCGCTGTTCTAACGCAATTGGGGAGCGGGTGTCGCGGGGGTTCCGTGGGGG

3661  GTTCCGTTGCAACGGGTCGGACAGGTAAAAGTCCTGGTAGACGCTAGTTTTCTGGTTTGG
```



```
5461   CAACTCCCACAACGTATACATCACGGCAGACAAACAAAAGAATGGAATCAAAGCTAACTT

5521   CAAAATTAGACACAACATTGAAGATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAA

5581   TACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCCACACAATC

5641   TGCCCTTTCGAAAGATCCCAACGAAAAGAGAGACCACATGGTCCTTCTTGAGTTTGTAAC

          PvuII                                          HpaI
5701   AGCTGCTGGGATTACACATGGCATGGATGAACTATACAAATGAGTTAACTAGCGTACGAT

5761   CGACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTA

5821   TCTGTTGTTTGTCCGGCCATCATGGCCGCGGTGATCA
```

Table listing suggested corrections and responses			**Nureen Mayat**
Page number	**Examiner**	**Minor correction /typo/comment**	**Response**
N/a	Joan Joseph	Distribution of chapters as follows: intro, materials and methods, results then conclusion	After consideration, changes were not appropriate
Contents	Bhavna Gordhan	Contents, spacing between Chapter and 4 is incorrect	Corrected
List of figures	Bhavna Gordhan	Superfluous "figure 1.2"	Deleted
3	Bhavna Gordhan	Section 1.2.1, rBCG give full nomenclature	Corrected
4	Bhavna Gordhan	Line 6 from bottom, full stop should be included after ref 26	Corrected
4	Joan Joseph	Paragraph 3, delete "many variations of HIV and SIV genes" insert "many HIV and SIV antigens have been expressed in BCG"	Corrected
5	Bhavna Gordhan	Line 2, insert "in" after "resulted"	Corrected
6	Bhavna Gordhan	Line 4 from bottom, correct spacing before "however"	Corrected
7	Bhavna Gordhan	Line 5 from bottom, spelling error correct to "Mutants"	Corrected
8	Joan Joseph	Line 7 from bottom, sentence misleading change	Changed to "in both *M. smegmatis* and BCG and an oriE allowing for replication in E. *coli* [49]"
12	Bhavna Gordhan	Fig 1.1, italicise gene names	Corrected
13	Bhavna Gordhan	Line 10, remove extraneous full stop and un-italicise refs 69 and 70	Corrected
13	Bhavna Gordhan	Line 5 from bottom, change Figure "2" to "1.2"	Corrected
19-21	Joan Joseph	Table 1.1, ATC change to ATc	Corrected
22	Joan Joseph	Name of promoters Pr-xylA and spoVG –should give original name of protein they drive expression of	Have referenced the original paper
25	Joan Joseph	Point 1, delete" establishment of diff conc: insert "Establishment of the range of different concentrations of chemical co-repressor (ATc)"	Corrected
30	Joan Joseph	Line 3, the concentration and supplier of pantothenic acid omitted	Disregarded, not necessary
31	Bhavna Gordhan	Line 4, insert "of the" after "each"	Corrected
36	Bhavna Gordhan	Line 5 and 6 from bottom, correct spacing	Corrected

Page number	Examiner	Minor correction /typo/comment	Response
38	Bhavna Gordhan	Line 3, insert “incubated” after “cultures”	Corrected
40-41	Bhavna Gordhan	Table 2.2, table irregular spacing	Corrected
47	Bhavna Gordhan	Figure 2.3 legend, ATC change to ATc	Corrected
49	Joan Joseph	Replace “3.2” with correct numbering “3.4”	Corrected
50	Joan Joseph	Last line, paragraph 1, change “dual BCG-HIV” to “dual TB-HIV”	Corrected
51-52	Bhavna Gordhan	Correct spelling of author to “Sambandamurthy”	Corrected
54	Bhavna Gordhan	Last line, insert “of” after presence	Corrected
55	Joan Joseph	Huge gap between text and next figure page	Difficult to change without significantly affecting the readability and format
59	Joan Joseph	Title change numbering to 3.4	Corrected
59	Bhavna Gordhan	Paragraph 1, says growth of auxotrophic strain was optimal with 48ug/ml and notably lower with 120ug/ml (growth limiting).	Changed “notably lower” to just “lower”
61	Joan Joseph	Change numbering from 4.3 to 4.4 and 4.5 to 4.4 in contents and change in text	Corrected
63	Joan Joseph	Line 6, change “our group” to “several groups”	Corrected
66	Both examiners	Figure 4.2, correctly label panels Bi and Bii	Corrected
66	Joan Joseph	Figure 4.3 suggest move to materials and methods	Figure is appropriately positioned
71	Bhavna Gordhan	Remove vertical lines in margin	Removed
73	Bhavna Gordhan	Remove vertical lines in margin	Removed
73	Joan Joseph	Figure 4.8, figure legend change to BCG Pasteur	Corrected
75	Bhavna Gordhan	Font size seems smaller	Corrected
70 and 80	Bhavna Gordhan	Should have ideally run uncut along with cut as well as partial digestion to acknowledged in text Fig 4.5A, lane 7	Suggestion appreciated and partial digestion acknowledged
82	Joan Joseph	Line 6,Fig “4.14” replace with“4.13”	Corrected

Page number	Examiner	Minor correction /typo/comment	Response
85	Joan Joseph	Last line, Figure 4.14 change to 4.15	Corrected
94	Joan Joseph	Line 2, change sentence to " harvested mycobacterial culture samples…"	Corrected
99	Bhavna Gordhan	Line 4, change 0ng.ml to 0ng/ml	Corrected
101-102	Joan Joseph	Suggest separate Figure 5.3 into separate figures	After consideration decided it would be too repetitive
103	Joan Joseph	Figure legend continued on next page	Corrected
106	Joan Joseph	Last line, change plastid to plasmid	Corrected
107	Bhavna Gordhan	Suggestion of alternate experiment that would yield better results for table 5.2	Suggestion for future experiments appreciated
108	Joan Joseph	Line 2 paragraph 2, recorder change to recorded	Corrected
110	Bhavna Gordhan	Line 6, "to found" change to "found to"	Corrected
116	Bhavna Gordhan	Change font of references to match rest of document and correct all spelling errors within	Reference list in a smaller font for purposes of economy and corrected all spelling errors
Supervisor		**Dr. Helen Stutz** Signature:	