Towards an understanding of the regulation of the oxidative stress responses of *Mycobacterium tuberculosis*.

Dominic Hoar

A thesis submitted in partial fulfilment

of the requirements of University College London

for the degree of Doctor of Philosophy

Division of Mycobacterial Research

National Institute for Medical Research

Mill Hill

London

September 2005

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Acknowledgements

First and foremost I would like to acknowledge the huge amount of support and assistance given to me by my supervisor, Dr. Roger Buxton, without whom none of this work would have been possible. I would also like to acknowledge the early guidance given to me by the late Dr. Jo Colston, and the advice and instruction of my second supervisors, Drs. Elaine Davis and Jonathan Millar. Dr. Davis in particular has often lent a critical eye when results seemed to be taking on a life of their own.

I would like to thank the Director of the NIMR, Sir John Skehel, for allowing me the opportunity to work at the Institute, and the Medical Research Council for providing my studentship and research funding. I would also like to thank the Society for General Microbiology for funding that allowed me to travel to meetings and work with our collaborators at HPA Porton Down, and University College London for assisting with expenses that allowed me to attend the 2005 Keystone meeting in Whistler, Canada.

I would also like to acknowledge our collaborators who have lent their technical expertise, time, laboratories, material and patience to enhance and further the work contained in this thesis. So, in no particular order, I would like to thank: Dr. Jason Hinds at St. George's Hospital Medical School, London, for microarray provision; Dr. José Saldanha at the NIMR for modelling the PAS domain of SenX3; Dr. Tanya Parish and Gretta Roberts at the Royal London Hospital for the gift of the Tame15 mutant; Joanna Bacon, Ian Beech, Kim Hatch and Jon Allnut at HPA Porton Down for their accommodation and ongoing chemostat work on the *senX3* null mutant.

I would very much like to thank all who provided vast amounts of technical advice and assistance, without which I would have been completely lost from day one.

Debbie Hunt for teaching me how to clone and how to organise my fridge; John Brennan, Vangelis Stavropolous and Steve Coade for their assistance with the infection work; Drs. Ed McGovan, Elena Curti, Tine Arnvig and Salvatore Adinolfi for their advice and help with the protein work; Bob Butler and Pat Brooks for their patience in supervising me in containment. I would also like to thank Dr. Steven Howell for his assistance with the mass spectrometry.

Last but by no means least, I would like to thank Alix Blockley, Bosco Chan, Roxana Curtis and Drs. Rachael Whalan, Calvin Boon, Jos Downey, Lucinda Rand and Robert Cox for their ideas and conversation, some of whom still owe me cups of tea ...

Abstract

Knowledge of the mechanisms by which *Mycobacterium tuberculosis* regulates its defence against oxidative stress will aid in the development of drugs and vaccines to combat the extensive morbidity and mortality caused by tuberculosis.

In this thesis I have investigated the SenX3-RegX3 two component signal transduction system, one of 12 such systems in *M. tuberculosis*, which is essential for virulence in a mouse model and is similar to systems in other organisms that are involved in sensing oxygen. By complementing the attenuated growth of the *senX3* and *regX3* null mutants in mice I have shown that this phenotype is truly due to the targeted mutations in these strains. *In vitro* assays of oxygen stress have shown the *senX3* and *regX3* null mutants to be less sensitive than the wild-type to superoxide and organic hydroperoxide stresses.

Microarray analysis of mutants in the SenX3-RegX3 system after growth in microaerobic conditions suggested some functions for RegX3; but there was little overlap between strains. However, microarray comparison of a strain overproducing *regX3* transcript demonstrated control of genes such as *ahpC* and *cydB*, indicating that this system may indeed play a role in the resistance of *M. tuberculosis* to oxidative stress.

RegX3 was expressed and purified in *E. coli* and used to examine the interaction of this protein to different promoters, including that of *senX3*.

Two further genes, which are probable one component regulators of oxidative stress in *M. tuberculosis*, Rv0465c and Rv1049, as judged by similarity to proteins in other organisms, were analysed through the isolation of null mutations. *In vitro* analysis suggests that Rv1049 is necessary for resistance to organic hydroperoxide stress,

although neither the Rv0465c nor the Rv1049 null mutants were attenuated in macrophage or murine intravenous injection models of infection.

List of Figures

List of Tables

Abbreviations

ΔRv0465c	Rv0465c null mutant of *M. tuberculosis*
ΔRv1049	Rv1049 null mutant of *M. tuberculosis*
Δ*regX3*	*regX3* null mutant of *M. tuberculosis*
Δ*senX3*	*senX3* null mutant of *M. tuberculosis*
BCG	Bacille Calmette-Guérin
bp	base pairs
bv	bed volume
BSA	bovine serum albumin
B. subtilis	*Bacillus subtilis*
cDNA	copy DNA
cfu	colony forming units
Crp	catabolite repressor protein
Cy	cyanine dye
dATP	2'-deoxyadenosine 5'triphosphate
dCTP	2'-deoxycytosine 5'triphosphate
dGTP	2'-deoxyguanosine 5'triphosphate
dH_2O	distilled water
DMSO	dimethyl sulphoxide
DNA	deoxyribonuclaeic acid
DNase	deoxyribonuclease
dNTP	2'-deoxynucleoside 5'triphosphate
dsDNA	double stranded DNA

DTT	dithiothreitol
dTTP	2'-deoxythymidine 5'triphosphate
E. coli	*Eschericia coli*
EDTA	ethylenediaminetetraacetic acid
F	Farads
g	grams
GSE	guanidinium chloride, N-lauroylsarcosine and EDTA solution
H_2O_2	hydrogen peroxide
HCl	hydrochloric acid
HIV	human immunodeficiency virus
HSR	head-space ratio; media to air ratio inside a sealed container
Hyg^r	hygromycin resistance
IPTG	isopropyl-β-D-thiogalactopyranoside
IS	insertion element
Kan^r	kanamycin resistance
kb	kilobase
k-	kilo-
l	litre
L-agar	Luria-Bertani agar
L-broth	Luria-Bertani broth
log	logarithm
M	molar
μ-	micro-

m-	milli-
MDR	multi-drug resistant
mol	mole
mRNA	messenger RNA
M. leprae	*Mycobacterium leprae*
M. tuberculosis	*Mycobacterium tuberculosis*
n-	nano-
NaAc	sodium acetate
NaCl	sodium chloride
NaOH	sodium hydroxide
p-	pico
$O2^{\cdot-}$	superoxide radical
OD_{600}	optical density at 600nm wavelength
$^{\cdot}OH$	hydroxyl radical
ORF	open reading frame
PBS	phosphate buffered saline
PCR	polymerase chain reaction
rpm	revolutions per minute
RNA	ribonucleic acid
RNase	ribonuclease
RNI	reactive nitrogen intermediates
ROI	reactive oxygen intermediates
SDS	sodium dodecyl sulphate

SSC	salt and sodium citrate solution
SET	sucrose and Tris-chloride solution
Suc^r	sucrose resistance
TB	tuberculosis
TBE	Tris (hydroxymethyl) aminomethane-boric acid-ethylenediaminetetraacetic acid buffer
TE	Tris (hydroxymethyl) aminomethane-ethylenediaminetetraacetic acid buffer
T_m	melting temperature
UV	ultra-violet light
V	volts
v/v	volume to volume ratio
WHO	World Health Organisation
w/v	weight to volume ratio
wt	wild-type
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

1 Introduction

1.1 Tuberculosis: disease and developments

Mycobacterium tuberculosis is the causative agent of tuberculosis, a disease which infects around one third of the world's population and is the cause of death of between two and three million people every year (Kochi 1991). Although tuberculosis can infect any part of the body, it typically attacks the lungs and throat. Whereas most foreign organisms that enter the body by this route are trapped and killed by the innate immune system, *M. tuberculosis* is able to survive and will usually grow until it is contained within a granulatomous lesion. The granuloma is an immune defence that is thought to control the replication of *M. tuberculosis* through limiting the oxygen and nutrients available to the bacterium. Despite this, the bacteria are able to survive in the host for long periods of time in a state of metabolic dormancy. If the host later becomes immunocompromised through, for example, malnutrition or immunodeficiency caused by HIV (Hopewell 1992), then he or she is no longer able to contain the growth of the *M. tuberculosis* and the lesion reactivates. Death typically occurs when the lesions caused by the mycobacteria erode the bronchial lining and cause the lungs to fill with blood.

There have been a number of trends in the incidence of tuberculosis in the last decade. Firstly, the high levels of *M. tuberculosis*-infected individuals in developing countries and the increasing presence of HIV in these countries means that there has been a growing trend for co-infection with both *M. tuberculosis* and HIV. It seems that these two diseases are synergistic, with death occurring within a matter of months in co-infected individuals. Secondly, more virulent strains of *M. tuberculosis* have been appearing in the Far East, such as the W, or Beijing, strain (Bifani *et al.* 2002). These isolates are characterised by a far higher infectivity index than other strains of *M.*

tuberculosis. Infection by the Beijing strains also proceeds more rapidly, and the bacteria are more likely to be multiply drug resistant – bringing us to the third trend. Drug resistance in *M. tuberculosis* has been a problem since the first antibiotic was invented to treat the disease, however due to poor adherence to treatment regimens and a dearth of new antibiotics the incidence of multiply drug resistant (MDR) tuberculosis is increasing. Combined with the increase in world travel and an often complacent attitude to the risks of tuberculosis infection in developed countries, this means that there has been a growing number of *M. tuberculosis* outbreaks in cities in developed countries where first-line antibiotics such as isoniazid have not been effective.

In recent years the World Health Organisation has had success at controlling tuberculosis infections with the DOTS (directly-observed therapy, short-course) treatment strategy (WHO, 2004), which involves the simultaneous administration of four antibiotics to which the isolate has proved to be sensitive. It is hoped that investigation of the molecular interactions within *M. tuberculosis* may allow the design of new antimicrobials to signalling or metabolic pathways of the bacterium, perhaps even to target *M. tuberculosis* during the latent stages of *M. tuberculosis* infection.

1.1.1 Vaccine use in the prophylactic treatment of *M. tuberculosis*

The Bacille Chalmette-Guerin (BCG) vaccine has been used prophylactically in the UK for several generations. It is a live attenuated strain of *M. bovis* that was grown on potato agar for an extended period of time until it lost its ability to create a progressive infection. See Oettinger *et al.* (1999) for a review. After its initial isolation in the 1920s the strain was distributed to a number of different laboratories around the world. The

BCG vaccine is highly variable in efficacy as a prophylactic vaccination against adult tuberculosis, where it is typically given during the teenage years. Although used widely in the UK, it has never been shown to be effective in adults in the developing world (Fine 1988; Roche *et al.* 1995; Fine 1996), and was of such limited efficacy in the US that it was decided to rely on selective application of drugs alone as a prophylactic measure in an outbreak. This also removed problems connected with the antigenic masking effect of the BCG vaccine since until very recently been difficult to distinguish vaccinated from infected individuals by serology due to the similarity between the BCG and *M. tuberculosis* bacteria. More recent genotyping of the BCG strains that were originally distributed to laboratories has shown that the original strain acquired a large deletion encompassing a number of genes now thought to encode parts of a secretory system involved in the virulence of *M. tuberculosis*. The distributed isolates, however, acquired a range of other large deletions and these may be in part responsible for the variable efficacy of the vaccine; see Behr (2002) for a review on this subject.

Other factors influencing the efficacy of the BCG vaccine are the underlying nutritional status of the population, and pre-existing immunity acquired from exposure to environmental mycobacteria. Despite the vaccine's variable protective efficacy in adults, it has been relatively successful as a neonatal vaccination protecting against childhood forms of the disease (Clarke & Rudd 1992). Furthermore, the antigenicity and widely accepted safety record make it an ideal candidate as a vector for the delivery of foreign, i.e., non-mycobacterial, antigens to protect against other diseases requiring strong intracellular immune responses.

Of greatest prophylactic use would be a live attenuated strain of *M. tuberculosis* to use as a vaccine that had been minimally modified in order to maintain its wild-type immunological profile. As more is understood about the dormant stages of *M. tuberculosis*, it may also be possible to design vaccines that are effective after infection, able to produce sterilising immunity in those with latent tuberculosis. Progress towards a modern vaccine for tuberculosis, as well as targeted antimicrobials, will rely on the elucidation of the genetic and molecular interactions of *M. tuberculosis* at all stages of the infection process.

1.1.2 Drug treatment for tuberculosis

The current treatment for tuberculosis is the widely-adopted WHO DOTS (directly-observed treatment, short-term) strategy. This involved the supervised treatment of patients with isoniazid, rifampicin, pyrazinamide and ethambutol for two months, followed by treatment with isoniazid and rifampicin alone for four further months. DOTS plus is used for cases of multi-drug resistant (MDR) tuberculosis, where second-line antimicrobials are included in the treatment regimen, and the period of the treatment is extended to 24 months.

There are several problems with the current use of antimicrobials to treat TB. It is difficult to find drugs that will penetrate caseous granulomas to come into contact with the bacteria. It is also difficult to find antimicrobials that have a sterilising ability, and this is of particular relevance to HIV-TB coinfected individuals whose immune systems will have trouble supporting a bacteriostatic antimicrobial attack on the bacilli. Furthermore, the majority of antimicrobials will have no effect on *M. tuberculosis* if it

has entered a latent or nonreplicating state, meaning that it is very difficult to treat patients with latent tuberculosis. Please see Zhang (2005) for a review.

1.1.2.1 Current anti-tuberculosis drugs, their mechanisms of action and their targets.

The key antimicrobials for tuberculosis therapy predominantly act as inhibitors of cell wall synthesis, and inhibitors of nucleic acid synthesis. They rely on interaction with components of the bacterial cell preferentially over components of the host cells, and an ability to cause bacteriostasis or bacteriocidal activity in the bacterial population. In the case of bacteriostatic drugs, or if bactericidal activity has not been total, then the administration of these drugs attempts to control bacteria pathogenesis and replication to the extent where the immune system can 'mop up' the remaining bacteria.

1.1.2.2 Inhibitors of cell wall synthesis

Isoniazid, for example, is a prodrug that is activated by the *katG* gene product to produce reactive oxygen intermediates. One product is the isonicotinic acyl radical, which inhibits enoyl ACP reductase (encoded by *inhA*), an enzyme involved in mycolic acid biosynthesis. Resistance in clinical isolates is usually due to the inactivation of *katG*, preventing activation of the prodrug. Ethambutol, as another example of cell wall synthesis inhibitors, prevents the polymerisation of arabinogalactan and lipoarabinomannan, probably by inhibiting arabinsyl transferase (encoded by *embB*). Mutations in *embB* seem to be the primary form of resistance seen clinically.

1.1.2.3 Inhibitors of nucleic acid synthesis

Inhibitors of nucleic acid synthesis include rifampicin and the fluoroquinolones. Rifampicin binds to the DNA-dependant RNA polymerase β-subunit encoded by *rpoB*, mutations in which cause resistance to rifampicin. The fluoroquinolones, including ciprofloxacin and ofloxacin, inhibit DNA synthesis by binding to the A and B subunits of the DNA gyrase. Inhibition of nucleic acid sythesis effectively paralyses the cell; however removal of these drugs may allow growth to continue. Rifampicin is particularly important since it has activity against both growing and non-growing bacteria.

1.1.2.4 Design of new antimicrobials against *M. tuberculosis*

The design of new drugs, to new targets in *M. tuberculosis*, avoids problems with cross-resistance that may occur when existing drugs are sequentially modified for improvements. Furthermore, the addition of novel drugs to existing multidrug therapies may prevent the development of resistance to the new drugs altogether. Ideally these new antimicrobials would target essential aspects of mycobacterial growth, be bactericidal, and would have the ability to be modified to give good pharmacokinetic properties. Certainly it may be possible to target those key aspects of mycobacterial growth and virulence that we know are sufficiently different to eukaryotic systems as to minimise problems of toxicity to the host. The regulatory systems of *M. tuberculosis* are likely to fit this category: for example, the histidine-aspartate phosphotransfer that occurs in two-component systems is rare in eukaryotes, and may represent a good target for inhibition by a designed small molecule (Stephenson & Hoch 2004). PAS domains may

have sufficiently specific structure to allow the selective inhibition of ligand-binding or sensor activation. Any genes that have been designated as essential for growth, for example by high-density mutagenesis (Sassetti *et al.* 2003), can be screened for utility as a drug target. This can be done by screening libraries of compounds for inhibitory effects, or a lifting of repression of inhibitory effects, against mutant strains of the bacteria; isolation of natural chemical partners of mycobacterial proteins and synthesis of toxic analogues; or by the use of crystallographic data to specifically design inhibitory molecules.

1.1.3 Immunological responses to TB infection

The study of the immune responses to *M. tuberculosis* is a very active and complex field. In the majority of infections the mycobacteria initially contact the innate immune response consisting of the alveolar macrophages lining the lungs. If these macrophages are unactivated, it may be possible for the bacteria to enter without stimulating the oxidative burst, using TACO protein amongst others (see Pieters & Gatfield (2002) for a review). It is known that mycobacteria can be seen to reside in phagosomes after macrophage infection, and these are prevented by the bacteria from fusing with lysosomes (Armstrong & Hart 1971). There is also evidence that the bacteria can escape the phagosomes and reside within the cytoplasm. Certainly, the bacteria can survive the innate immune response in immunocompetent hosts.

In terms of the acquired immune response, *M. tuberculosis* elicits both Th1 and Th2 responses. The latter tend to be reactions to the secreted antigens of *M. tuberculosis*, and although this response may be useful in diagnosis of infected individuals it is not

thought to be protective of the host. As a predominantly intracellular pathogen, much of the inflammatory pathology of the *M. tuberculosis* granuloma is produced by the Th1 arm of the acquired immune response, to which the secretion of the cytokine interferon-γ is involved. Stimulation of the Th1 response may also occur due to the interaction of *M. tuberculosis* secreted proteins with pattern-recognitions receptors, such as the toll-like receptor proteins, as part of the innate immune response. A great deal of *M. tuberculosis* pathology is caused by the host immune responses to the bacterium, and there is a great deal of evidence for *M. tuberculosis* components such as lipoarabinomannan being able to modulate the immune response, reviewed in Flynn & Chan (2003).

As the bacteria multiply, a large variety of cells build up around the focus of infection, including activated macrophages, CD4+ and CD8+ T-lymphocytes and multi-nucleated giant cells. As the inflammatory response continues, a granuloma forms around the focus of infection and the tissue within the granuloma becomes necrotic, forming a caseous, 'cheesy' lesion. If the bacteria are contained, then the granuloma will eventually calcify unless the disease later reactivates. If the bacteria and the lesion continue to grow however, then death may eventually occur when the lesion erodes the lining of a major artery.

1.1.4 The *M. tuberculosis* genome

In a landmark paper, Cole *et al.* (1998)described the sequence of the *M. tuberculosis* genome. The *M. tuberculosis* genome has a very high GC content, and contains a large number of genes encoding proteins involved in lipid metabolism as well as a series of genes encoding glycine-rich proteins that have been designated as the PE

and PPE families that may be involved in antigenic variation. Approximately half of the open reading frames (ORFs) identified coded for proteins that bear little homology to those identified in other organisms.

The publishing of the *M. tuberculosis* sequence allowed genetic manipulation of the bacterium on a much larger scale than before, and attracted a large amount of attention to a field that had been considered stagnant due to a lack of tools to allow genetic manipulation of the mycobacteria. The new genetic information made it easier to create targeted gene knockouts, and polymerase chain reaction (PCR) amplification of sections of each ORF identified in the sequencing project allowed the construction of an *M. tuberculosis* microarray. The ability to create targeted gene mutations using reverse genetics meant that it was possible to create strains without a phenotype. This was particularly important since *M. tuberculosis* is difficult to grow, in terms of requiring category III containment and having a very slow generation time, which meant that traditional methods of inducing mutation in the bacteria and selecting for phenotypes was not an attractive option for generating mutants in *M. tuberculosis*.

1.1.4.1 Types of regulatory mechanism in *M. tuberculosis*

In order to respond efficiently to changes in their environment, bacteria use a number of different protein-based systems to sense stimuli and co-ordinately regulate gene expression. Generally considered the most important of these are the two-component signal transduction systems, although these are outnumbered numerically by the one-component signal transduction systems, a review of which can be seen in Ulrich *et al.* (2005).

One component signal transduction systems consist of a single protein with both a sensory domain and a DNA-binding domain. These are activated directly or indirectly by a signal, and either repress or activate transcription of genes. As such they are cytoplasmic proteins. Two component signal transduction systems consist of a sensor protein that contains a signal input domain and a histidine kinase domain, which phosphorylates a response regulator that contains a DNA binding domain. Activation of the sensor protein, which is usually membrane-anchored, causes it to autophosphorylate on a histidine residue, followed by phosphotransfer to an aspartate residue on the response regulator. Phosphorylation of the response regulator alters its DNA binding activity and allows it to activate or repress the translation of other genes. For a review see Hoch (2000). Since the phosphorylation of histidine residues in two-component systems is specific for bacteria, they represent a potential target for the design of novel antibacterial agents (Stephenson & Hoch 2004).

Sigma factors are another mechanism of protein-level regulatory control in bacteria. Sigma factors form part of the RNA polymerase holoenzyme and provide recognition of sequences upstream of genes, allowing transcriptional control of these genes. Whereas in *E. coli* sigma factor F has a role as a general sigma factor, in *M. tuberculosis* there are 11 sigma factors (Gomez *et al.* 1997; Cole *et al.* 1998; Wu *et al.* 2004), with sigma factor A being considered as the major sigma factor since it has so far been impossible to create null mutants for this gene (Gomez *et al.* 1998). Use of different sigma factors is likely to influence the response of the bacterium to environmental signals and influence the transcriptome of the organism; and current knowledge suggests that the sigma factors of *M. tuberculosis* have a great deal of regulatory overlap, producing a

complex system of regulation (reviewed in Manganelli *et al*, (2004)). Of particular relevance to the oxidative stress defence of *M. tuberculosis* is sigma factor H, SigH, mutants of which are more susceptible than the wild-type to the superoxide stress and heat stress, which may itself cause oxidative stress of the cells (Raman *et al.* 2001).

Unusually, the genome sequence of *M. tuberculosis* appears to encode 11 eukaryotic-like serine-threonine protein kinases (Cole *et al.* 1998). These have a variety of proposed regulatory roles, from membrane transport to cell division (reviewed in Av-Gay & Everett (2000)) and at least one serine-threonine protein kinase is required for growth of *M. tuberculosis in vivo* (Curry *et al.* 2005). Although serine-threonine protein kinases are an interesting class of regulator in *M. tuberculosis*, there are currently no roles for them described for defence against oxidative stress.

1.2 Oxygen and metabolic sensing in bacteria

Many of the mechanisms that bacteria use to sense their environments have been elucidated in *E. coli*, and it is hoped that systems discovered in *M. tuberculosis* will allow comparison by homology and similarities of regulons in an effort to accelerate our understanding of *M. tuberculosis* pathogenesis. In *E. coli* much of the machinery for sensing the presence of oxygen and other metabolites is also utilised for detoxification of ROI and RNI used in host defence; it is difficult to separate the two in *E. coli*, let alone *M. tuberculosis* where many of the regulatory systems are still to be defined. The following sections describe our current knowledge of relevant systems in *E. coli*, and comparisons have then been drawn with data produced from studies in *M. tuberculosis*.

1.2.1 The role of bacterial metabolic sensory systems

In bacteria there are often several layers of regulation governing the expression of genes involved in metabolism and redox states (Green & Paget 2004). The existence of a dynamic network of interactions between these systems means that bacteria can 'fine tune' their energy metabolism to give them the maximum possible energy yield for a given energetic outlay. This is done by deploying components of the aerobic, anaerobic and fermentative metabolism, such as cytochrome oxidases and nitrate reductases, to maximise the potential difference between electron donors and the terminal electron acceptors of the electron transport chain. Oxygen is the preferred electron donor, with an redox potential midpoint of E^{1o}= +820mV, followed by nitrate (E^{1o}+ +430mV) and finally organic molecules such as fumarate (E^{1o}= +30mV) that are used in the fermentative metabolism. The use of oxygen and nitrogen as terminal electron acceptors means that, in contrast to fermentation, redox balance does not have to be maintained in the cell. The regulation of the electron transport chain is also linked to glycolysis and the TCA cycle, where absence of oxygen will induce the formation of the glyoxylate shunt pathway, enabling the bacterium to derive energy from pyruvate as a metabolic product of fatty acid degradation. The manipulation of these regulatory mechanisms in *M. tuberculosis* would be likely to aid its survival in nutrient or oxygen starvation conditions.

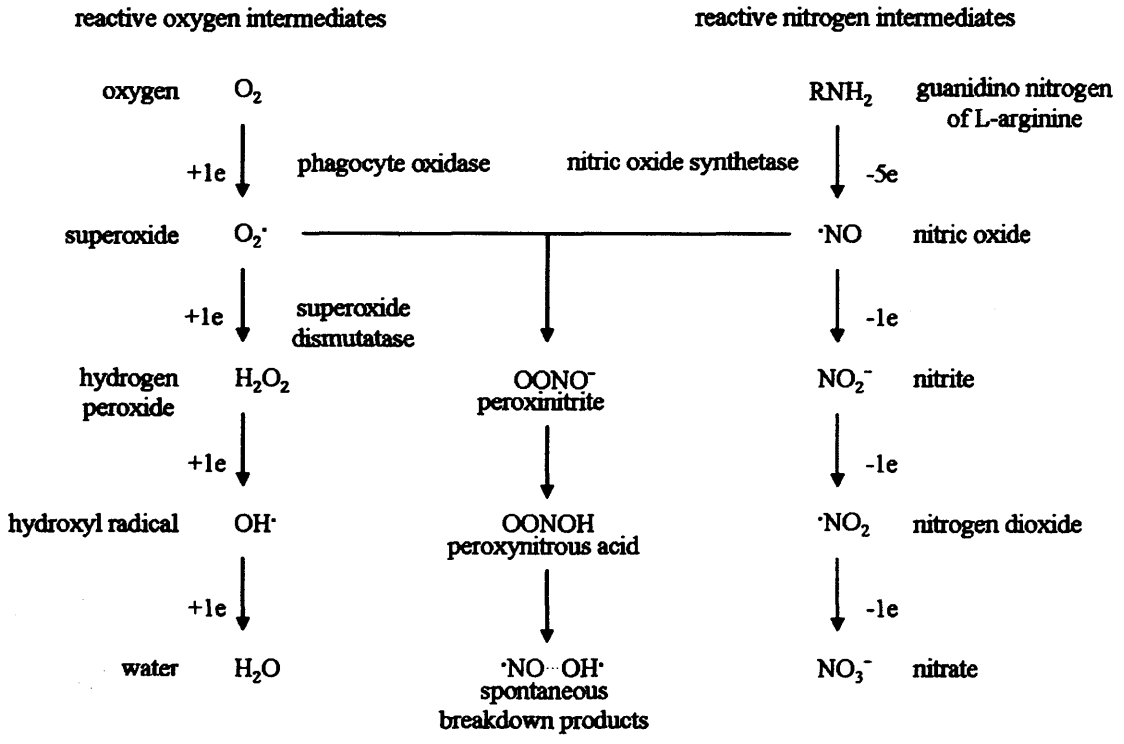
The utilisation of the high-redox potential oxygen in energy metabolism however does have a cost. Molecular oxygen degrades to a free radical state comparatively easily, since its two unpaired electrons have parallel spin and occupy separate orbitals in the gas, giving the molecule a paramagnetic nature. Unless both of these electrons pair with

electrons of the complementary spin at the same time, a one-electron reduction will take place yielding the superoxide radical, $O2^{\cdot-}$; see Demple (1991) for a review. Bacteria can detoxify superoxide to hydrogen peroxide using the enzyme superoxide dismutase to produce the relatively unreactive hydrogen peroxide; however if the hydrogen peroxide is not reduced further by catalase to yield water and molecular oxygen it forms the highly reactive hydroxyl radical in the presence of transition metals such as Fe^{2+} (Imlay & Linn 1988). Organic hydroperoxides may form organic peroxide radicals, which are analogous to hydrogen peroxide and must be reduced by alkyl hydroperoxidases. Hydroxyl radicals will attack almost all biological macromolecules, including DNA and lipids. Superoxide radicals are also produced by the enzyme phagocyte oxidase in host cells as a bactericidal mechanism during an infection, in which case the detoxification mechanisms of the bacteria are induced as a method of surviving within the host. Furthermore, in the absence of carbohydrate sources or oxygen, as is hypothesised to occur in human *M. tuberculosis* granuloma environments, the bacterium would be at a greater advantage if it were able to switch off the then unnecessary and potentially deleterious aerobic metabolism and its concomitant superoxide production (reviewed in Nystrom (1999)). Reactive nitrogen intermediates may also be formed by a variety of mechanisms and may combine with reactive oxygen intermediates to further damage the cell. See Figure 1.1 (adapted from Nathan & Shiloh (2000)) for a schematic diagram of the reactive oxygen and nitrogen intermediates involved in causing cellular damage.

A number of global regulatory pathways interact to control the bacterial response to oxidative stress. These have been principally studied in Gram-negative bacteria such

electrons of the complementary spin at the same time, a one-electron reduction will take place yielding the superoxide radical, $O2^{\cdot-}$; see Demple (1991) for a review. Bacteria can detoxify superoxide to hydrogen peroxide using the enzyme superoxide dismutase to produce the relatively unreactive hydrogen peroxide; however if the hydrogen peroxide is not reduced further by catalase to yield water and molecular oxygen it forms the highly reactive hydroxyl radical in the presence of transition metals such as Fe^{2+} (Imlay & Linn 1988). Organic hydroperoxides may form organic peroxide radicals, which are analogous to hydrogen peroxide and must be reduced by alkyl hydroperoxidases. Hydroxyl radicals will attack almost all biological macromolecules, including DNA and lipids. Superoxide radicals are also produced by the enzyme phagocyte oxidase in host cells as a bactericidal mechanism during an infection, in which case the detoxification mechanisms of the bacteria are induced as a method of surviving within the host. Furthermore, in the absence of carbohydrate sources or oxygen, as is hypothesised to occur in human *M. tuberculosis* granuloma environments, the bacterium would be at a greater advantage if it were able to switch off the then unnecessary and potentially deleterious aerobic metabolism and its concomitant superoxide production (reviewed in Nystrom (1999)). Reactive nitrogen intermediates may also be formed by a variety of mechanisms and may combine with reactive oxygen intermediates to further damage the cell. See Figure 1.1 (adapted from Nathan & Shiloh (2000)) for a schematic diagram of the reactive oxygen and nitrogen intermediates involved in causing cellular damage.

A number of global regulatory pathways interact to control the bacterial response to oxidative stress. These have been principally studied in Gram-negative bacteria such


reactive oxygen intermediates
reactive nitrogen intermediates
oxygen
O_2
RNH_2
guanidino nitrogen of L-arginine
+1e
phagocyte oxidase
nitric oxide synthetase
-5e
superoxide
$O_2^{\cdot}$
$^{\cdot}NO$
nitric oxide
+1e
superoxide dismutatase
-1e
hydrogen peroxide
H_2O_2
$OONO^-$ peroxinitrite
NO_2^-
nitrite
+1e
-1e
hydroxyl radical
$OH^{\cdot}$
OONOH peroxynitrous acid
$^{\cdot}NO_2$
nitrogen dioxide
+1e
-1e
water
H_2O
$^{\cdot}NO\cdots OH^{\cdot}$ spontaneous breakdown products
NO_3^-
nitrate

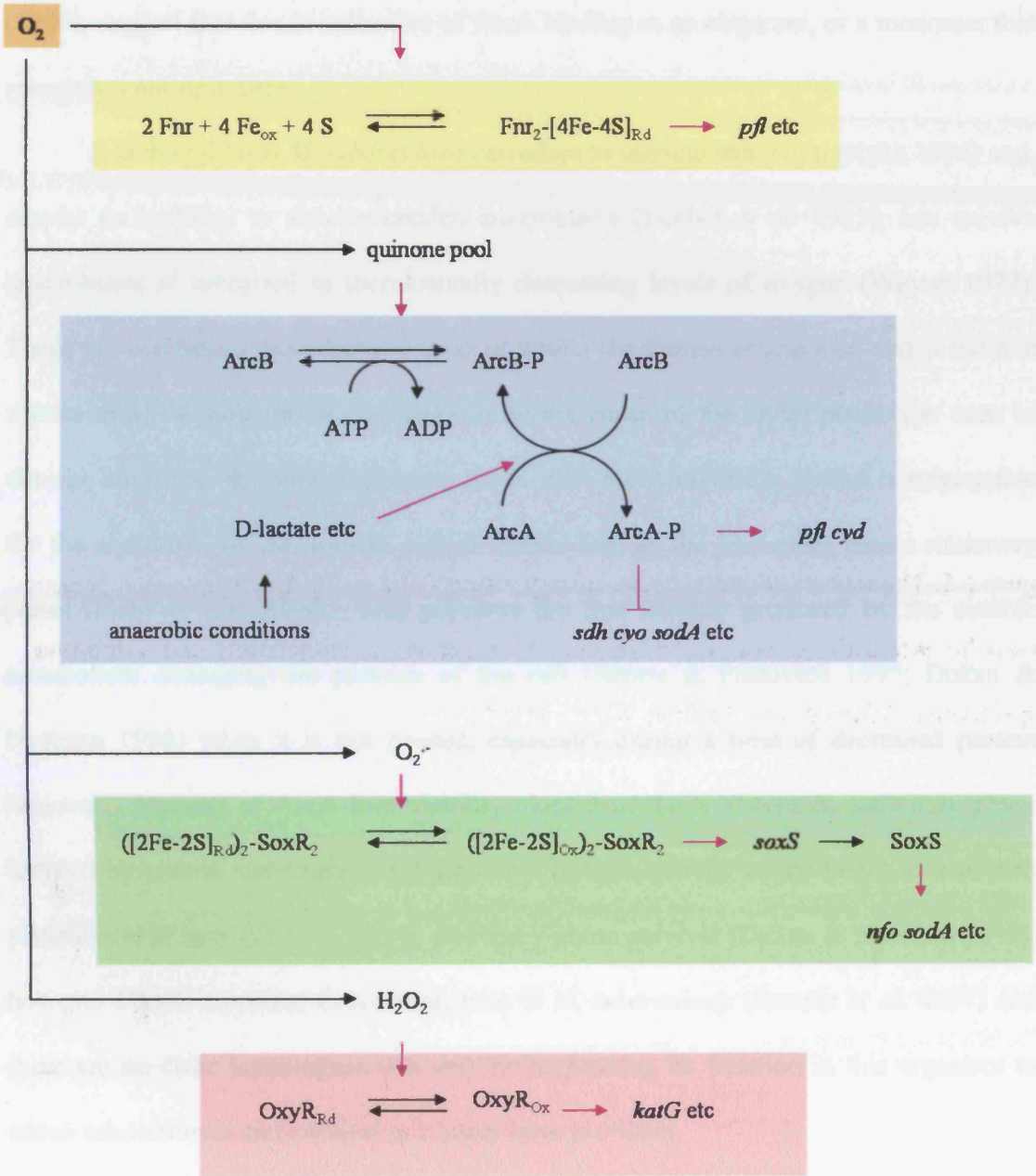
as *E. coli*, in which the SoxS, Fnr, OxyR, NarL and ArcB-ArcA systems respond to changes in redox; and Fur and Crp sense iron concentration and catabolites, respectively. See Figure 1.2 (adapted from Iuchi & Weiner (1996)) for a diagram of the mechanisms of action of the four principal oxygen-sensing systems.

1.2.1.1 ArcB-ArcA

One system that may be relevant when considering a role for the SenX3-RegX3 two component signal transduction system is the ArcB-ArcA two component signal transduction system of *E. coli*.

ArcB is a membrane-bound sensor protein in *E. coli* whose ability to autophosphorylate is inhibited by oxidation of two critical cysteine residues (Malpica *et al.* 2004) caused by the presence of a large pool of oxidised quinone electron carriers (Georgellis *et al.* 2001). Under microaerobic conditions the pool of oxidised quinones decreases, enabling autophosphorylation of ArcB, subsequent phosphorylation of ArcA and modulation of the bacterium's transcriptional activity. ArcB is constitutively synthesised in *E. coli*; however *arcA* is expressed approximately four-fold higher anaerobically (Iuchi & Lin 1992). The kinase activity of ArcB can be enhanced by the presence of other metabolic intermediates such as D-lactate (Rodriguez *et al.* 2004), suggesting at least two levels of control for this system. There are several transcriptional start sites in the ArcA promoter, one of which is dependant on Fnr for activity (Compan & Touati 1994), and promoter studies using ArcA itself showed that ArcA displays high affinity for a 141 bp region downstream of the pyruvate formate-lyase promoter

Figure 1.2: The four principle oxygen sensing systems of *E. coli*. Black arrows represent metabolic interactions, purple arrows and bars represent stimulation and inhibition, respectively. Adapted from Iuchi & Weiner (1996).



From Iuchi and Weiner 1996

(Drapal & Sawers 1995). A 65 bp region of the sodA promoter was protected by lysate of cells overproducing ArcA (Tardat & Touati 1993). Iuchi & Weiner, in their review (1996), suggest that this is indicative of ArcA binding as an oligomer, or a monomer that recognises multiple sites.

It is thought that *M. tuberculosis* can adapt to nutrient starvation (Nyka 1974) and, despite its inability to survive sudden anaerobiasis (Loebel *et al.* 1933), can survive anaerobiasis if subjected to incrementally decreasing levels of oxygen (Wayne 1977). These are conditions hypothesised to exist within the human granuloma, and metabolic shutdown of *M. tuberculosis* is likely to be the cause of the latent phenotype seen in clinical infection. In other organisms, viz *E. coli*, the ArcB-ArcA system is responsible for the shutdown of the aerobic carbon metabolism as the bacterium enters stationary phase (Iuchi & Lin 1988). This prevents the free radicals produced by the aerobic metabolism damaging the proteins of the cell (Benov & Fridovich 1995; Dukan & Nystrom 1998) when it is not needed, especially during a time of decreased protein turnover. Mutants of ArcA lose viability more than the wild-type in stationary phase induced by carbon starvation (Nystrom *et al.* 1996). Activity of the OxyR system also plays a role in survival of *E. coli* to stationary-phase survival (Dukan & Nystrom 1999), however OxyR is present as a pseudogene in *M. tuberculosis* (Deretic *et al.* 1997) and there are no clear homologues that may be performing its function in this organism to which adaptation to and survival in a latent form is critical.

1.2.1.2 Fnr

Fnr senses oxygen directly in *E. coli*, and was discovered by mutations that produced strains deficient in nitrate, nitrite and fumarate reductases. Sensing occurs by reduction of an iron-sulphur cluster in the presence of oxygen, upon which the protein's DNA binding activity alters, causing up- and down-regulation of a number of operons, although it is predominantly a transcriptional activator of operons involved in anaerobic metabolism. Key operons regulated by Fnr are pyruvate formate lyase, and the cytochrome oxidases *o* and *d*. Fnr also repressed its own synthesis under anaerobic conditions (Spiro & Guest 1987) and has strong homology to Crp (catabolite repressor protein) (Guest 1992), which extends to the similarity of their DNA-binding sequences and gave rise to the Crp-Fnr family of transcriptional regulators (Shaw *et al.* 1983).

1.2.1.3 OxyR

OxyR was discovered as a result of the study of *E. coli* and *S. typhimurium* mutants resistant to high levels of hydrogen peroxide, which showed high levels of catalase and alkyl hydroperoxide reductases that are critical in the response of the bacterium to oxidative stress (Demple & Halbrook 1983; Christman *et al.* 1985). OxyR was subsequently shown to bind to the catalase I (*katG*) promoter under oxidising, but not reducing conditions (Toledano *et al.* 1994). In *E. coli* OxyR is constitutively expressed but becomes active after oxygen exposure. It appears to bind to a motif of four ATAG motifs in adjacent major grooves on one face of the DNA helix (Toledano *et al.* 1994).

1.2.1.4 SoxR

SoxR (superoxide resistance) was originally identified as a mutant that conferred resistance to superoxide generating agents (Greenberg *et al.* 1990). Upon oxidation of its two iron-sulphur clusters it upregulates transcription of SoxS, which then regulates transcription of a number of genes including *sodA* (Hidalgo & Demple 1994; Hidalgo *et al.* 1995; Wu *et al.* 1995). Although this system contains two components, one being a sensor and another being a regulator, it does not constitute a *bona fide* two component signal transduction system since interaction between the two proteins does not occur at the phosphorylation level. It is likely that SoxR does not sense superoxide directly, since the system can also be activated by nitric oxide (Nunoshiba *et al.* 1993). SoxS can also interact with its own promoter, causing a two-fold decrease in expression (Nunoshiba *et al.* 1993).

1.2.1.5 Fur and Crp

Neither Fur (ferric uptake regulator) nor Crp respond to oxygen; however these proteins serve to sense the nutrient levels in the external environment and can modify the responses of the bacterium to other agents. Fur homologues in other organisms such as *S. typhimurium* use iron availability as an indicator of their presence inside another organism, and upregulate the production of siderophores and a number of other virulence factors in preparation for their encounter with the innate immune response. Fur exhibits a level of interaction with the more direct agents of the oxidative stress response, as discussed below.

Crp functions to sense the levels of cyclic AMP, which is a direct measure of energy levels within the cell (Pastan & Perlman 1970). Crp regulates the transcription of

over 100 operons, and analysis of the functions of the Crp-Fnr family of regulators shows that they govern diverse responses within the cell, including responses to oxidative and nitrosative stress, and redox sensing. For a review see Korner *et al.* (2003).

1.2.2 Interactions between the global regulators of *E. coli*

The key regulators of the oxidative stress responses of *E. coli* are well connected, offering a dynamic and broad response to diverse stimuli. The major elements of the oxidative stress response therefore are often controlled by a number of different regulators that may synergise or antagonise each other depending on the environmental stimuli and the nutrients available.

Cytochrome *d* oxidase, for example, is expressed maximally in *E. coli* at microaerobic levels in the presence of activated ArcA, minimal aerobic repression by the histone-like protein H-NS and minimal anaerobic repression by Fnr (Govantes *et al.* 2000). *sodA* expression is controlled by at least six different regulatory proteins, including Fur , ArcA, SoxS, SoxQ, Fnr and IHF (Compan & Touati 1993), where SoxS and SoxQ increase expression of *sodA* and the other proteins act as repressors of *sodA* transcription. This has the effect of linking the expression of *sodA* to iron availability (Fur), and the level of oxygen in the environment (ArcA) (Tardat & Touati 1991). Consistent with these results, the binding sites for ArcA and Fur overlap (Tardat & Touati 1993). At other levels the presence of intracellular iron, which activates Fur, also increases the rate at which reactive oxygen intermediates form, stimulating a number of other pathways.

Fnr is required for maximal ArcA expression, and can weakly induce transcription of ArcA under anaerobic conditions (Compan & Touati 1994). ArcA also recognises the metabolic conditions controlled by members of the Crp and Fnr regulons, creating another level of interaction.

It can be seen from these two examples that the response of *E. coli* to oxidative or nitrosative stress, and its metabolic state, are closely linked and it has been possible to elucidate many of these mechanisms due to the comparative ease by which *E. coli* can be manipulated in the laboratory.

1.3 Oxidative stress responses of *M. tuberculosis*

In contrast to the increasingly detailed understanding of the oxidative stress responses in *E. coli* and some other gram-negative bacteria, the molecular basis for resistance to oxidative stress in *M. tuberculosis* is relatively poorly understood. Furthermore, due to the considerable evolutionary differences between the mycobacteria and the gram-negatives, and the strikingly different lifestyles of the pathogenic slow-growing mycobacteria, it is conceivable that the approach that *M. tuberculosis* takes to oxidative stress defence is broadly different to that of *E. coli*. Nevertheless, our understanding of the regulatory systems of *E. coli* allows certain assumptions to be made of apparently similar systems in *M. tuberculosis*, which may subsequently decrease research time and discount unlikely hypotheses.

1.3.1 Homology of the *M. tuberculosis* oxidative stress responses to those of *E. coli*

Many of the well characterised key regulators in *E. coli*, such as SoxR, OxyR and Fnr either have no clear homologues in *M. tuberculosis*, or are present as pseudogenes (Deretic *et al.* 1995). In *M. tuberculosis* OxyR is a pseudogene (Deretic *et al.* 1995). In *M. smegmatis* and the Gram-negative bacteria OxyR controls the response to low doses of hydrogen peroxide that protects the bacteria against subsequent high-dose challenges with the same agent. In *M. tuberculosis* however, this response is limited to the induction of a single catalase, KatG (Sherman *et al.* 1995). KatG is a determinant of isoniazid sensitivity, and is non-functional in all clinical, isoniazid-resistant strains. These KatG mutant strains have acquired secondary site mutations that cause overexpression of AhpC (Sherman *et al.* 1996) which may compensate for the absence of KatG in the detoxification of organic peroxides.

AhpC was not initially seen to be produced at detectable levels in *M. tuberculosis in vitro* during aerobic growth, even though it is present as an unmutated gene (Dhandayuthapani *et al.* 1996; Zhang *et al.* 1996), and although it is possible that this is due to the loss of OxyR (Dhandayuthapani *et al.* 1997), other reports (Springer *et al.* 2001) have shown AhpC of *M. tuberculosis* to be expressed in statically grown cultures and in macrophages but not aerobically grown cultures, pointing to a second level of regulation for AhpC. AhpC was subsequently shown to have a role in peroxynitrite resistance and survival of *M. tuberculosis* in macrophages (Master *et al.* 2002), and this remains consistent with previous work on AhpC with other reactive oxygen intermediates when one considers that peroxynitrite is a product of reactive oxygen intermediate breakdown and may be one of the most damaging products of oxidative and nitrosative stress.

Other possible homologues of global regulatory proteins in *E. coli* in *M. tuberculosis* are Fur, which is divergently transcribed with *katG* (Pagan-Ramos *et al.* 1998) and two potential Crp homologues, Rv3676 (Rickman *et al.* 2005) and Rv1675c (A. Blockley, personal communication). The ArcB-ArcA two component signal transduction system of *E. coli* may have a homologue in the SenX3-RegX3 two component signal transduction system of *M. tuberculosis*; this was examined this thesis, and a publication arising from work partially contained within this thesis (Rickman *et al.* 2004).

1.3.2 Two component signal transduction systems

Two component systems in bacteria, as typified by the ArcB/A system of *E. coli*, often respond to signals from the external environment, and in pathogenic bacteria will increase the bacterium's chance of survival during infection (Cheung *et al.* 2004). As such, it is possible to see that bacteria that ordinarily encounter diverse environments tend to have a greater number of two-component systems encoded in their genome. *E. coli* and *B. subtilis*, for example, contain between 30 and 50 two component systems. *M. tuberculosis* contains 11 two-component systems, perhaps as a reflection of the fact that it exists almost entirely within a host. It is possible, however, that *M. tuberculosis* can 'compensate' its lack of two-component signal transduction systems with the serine-threonine protein kinases (Galperin & Gomelsky 2005).

The *M. tuberculosis* two-component systems, then, could be assumed to respond to the host environment that it encounters in the human body. *M. leprae* has an even more limited environmental range than *M. tuberculosis*, in that it can only infect a few

different tissues in the human body. *M. leprae* is considered to have evolved from *M. tuberculosis* by reductive evolution (Monot *et al.* 2005), and although its genome is approximately the same size as that of *M. tuberculosis*, it codes for approximately half the functional genes that are encoded in the *M. tuberculosis* genome (Cole *et al.* 2001). The remainder of the genes encoded in *M. tuberculosis* are present as pseudogenes or are deleted in *M. leprae*, giving rise to the assumption that the genes still remaining in *M. leprae* reflect the minimum geneset necessary for the mycobacterial intracellular lifestyle.

1.3.2.1 The SenX3-RegX3 two component signal transduction system of *M. tuberculosis*.

One of the 11 two component systems of *M. tuberculosis*, and one of the four two component systems conserved in *M. leprae*, is the SenX3-RegX3 two component signal transduction system. Previous *in silico* analysis of this system by R. Buxton and J. Millar (NIMR) used BLAST searches to identify RegX3 as a possible orthologue of ArcA (Drury & Buxton 1985; Iuchi & Lin 1988) of *E. coli*, and SenX3 as a possible orthologue of Mak2p (Buck *et al.* 2001) of *Schizosaccharomyces pombe* (Buck *et al.* 2001; Rickman *et al.* 2004). ArcA is part of the PAS-domain containing ArcB-ArcA two component system of *E. coli* that is involved in the shutdown of aerobic metabolism genes when the bacterium enters a microaerobic environment (Compan & Touati 1994; Alexeeva *et al.* 2003; Levanon *et al.* 2005). Mak2p is a PAS-domain containing histidine kinase that is important in the defence of *S. pombe* to peroxide damage.

An isogenic mutant of the *regX3* gene of *M. tuberculosis* (called Tame15) was made contemporaneously with this study and was shown to be attenuated in a variety of

macrophage and macrophage-like infection assays, and immunocompetent and immunodeficient mice (Parish *et al.* 2003). It should be noted that in this study that Tame 15 did not grow normally *in vitro*, which may be responsible for the apparent attenuation of the strain when compared to the growth of the wild-type in the infection models. The authors did not, however, complement Tame 15 with the wild-type allele. It seems that the deleted region of Tame 15 included part of the region between *senX3* and *regX3* that contains three mycobacterial interspersed repeat units (MIRUs) (T. Parish, personal communication). MIRUs are used for typing strains of mycobacteria and may have a regulatory function (Supply *et al.* 1997).

1.3.2.2 The DosR-DosS two component signal transduction system of *M. tuberculosis*.

It is entirely plausible that the oxidative stress responses of *M. tuberculosis* are at least partially governed by systems that have no equivalent in *E. coli* or other Gram-negative bacteria. One of the few two component signal transduction systems of *M. tuberculosis* of which we have a clear understanding of the stimulus and potential relevance to the oxidative stress responses of *M. tuberculosis* is the DosR-DosS system. The DosRS two-component system is likely to be involved in the sensing of oxygen, although probably not directly (Park *et al.* 2003).

DosR is a response regulator of *M. tuberculosis* controlled by two sensor histidine kinases, DosS and DosT (Roberts *et al.* 2004) that is required for dormancy survival of *M. tuberculosis* in an *in vitro* model of oxygen starvation-induced metabolic shutdown

(Boon & Dick 2002) and is expressed in standing shallow broth cultures and *in vivo* granuloma models (Florczyk *et al.* 2003; Karakousis *et al.* 2004; Kendall *et al.* 2004).

It is likely that DosR is involved in the adaptation of *M. tuberculosis* to stationary phase (Voskuil *et al.* 2004); however a study of the transcription of the two-component systems of *M. tuberculosis* during infection of human peripheral blood monocyte-derived macrophages shows that the SenX3-RegX3 two component signal transduction system is induced partially at 18 hours post infection and strongly at 48 hours post infection; yet disappears entirely by the time DosR is induced at 110 hours post infection (Haydel & Clark-Curtiss 2004). This suggests that a role for DosR in stationary-phase adaptation does not preclude a role for the SenX3-RegX3 two component signal transduction system in oxygen sensing, and raises the possibility that the two systems may be acting in concert. In addition to the DosR system, it has been suggested that a universal stress protein (USP) response may be implicated in the adaptation of *M. tuberculosis* to low-oxygen conditions, and may play a role in latency (see O'Toole & Williams (2003) for a review of USPs).

Work on a DosR mutant has shown them to be hypervirulent (Parish *et al.* 2003), which may be due to disregulation of the alpha-crystallin gene (Hu *et al.* 2005). A DosR homologue, DevR, exists in *M. smegmatis* where it is required for adaptation to oxygen heat stress resistance (O'Toole *et al.* 2003), indicating that this organisms may simplify the study of the DosR response in *M. tuberculosis*.

1.3.3 Two component systems of other bacteria

1.3.3.1 FixL of *Rhodobacter spheroides*

FixL homologues are found in the bacteria *Rhizobium meliloti*, *Bradyrhizobium japonicum* and *Azorhizobium caulinodans*, and in these organisms forms part of a hierarchy of regulatory genes, constituting one half of a two-component signal transduction system (along with FixJ) that responds to environmental nitrogen and oxygen conditions, controlling nitrogenase synthesis and the transcription of genes when the bacterium encounters microaerobic environments. Nitrogenase is the principle method by which nitrogen is fixed to ammonia in these bacteria, and needs to be tightly regulated.

FixL is a membrane-bound haem-containing protein that is able to phosphorylate the FixJ response regulator. It has three domains; one N-terminal membrane anchor, a haem-binding central domain and a C-terminal domain with autophosphorylation and kinase activity. The haem prosthetic group in the central domain is associated with a PAS domain, operates to sense oxygen conditions, and controls the kinase and phosphatase activity of the c-terminal domain (Gong *et al.* (1998).

Under low oxygen conditions, the kinase activity of the N-terminal domain of FixL is increased, resulting in a larger proportion of phosphorylated FixJ, probably at its N-terminal domain. The C-terminal DNA-binding domain of FixJ then allows it to initiate transcription of target genes. When oxygen levels increase, the level of phosphorylated FixJ is decreased as the phosphatase activity of the N-terminal domain of FixL increases. For reviews please see Fischer (1994) and Gilles-Gonzalez & Gonzalez (2004).

1.3.3.1 NifL of *Azotobacter vinlandii*

NifL has been studied as a regulatory protein of *Azotobacter vinlandii* and *Klebsiella pneumoniae*. It may also be part of a two-component signal transduction system with NifA, and also controls regulation of the fixation of atmospheric nitrogen to ammonia in plant root nodules under conditions of nitrogen deprivation. NifL responds to oxygen and fixed nitrogen by means of either one or two (depending on the organism) PAS domains, one of which contains flavin adenine dinucleotide as a prosthetic group. NifL is hypothesised to regulate NifA by a number of methods. The transcriptional activator NifA is constitutively active in both *A. vinlandii* and *K. pneumoniae*, and its activity is lessened when NifL is present. For a review please see Martinez-Ardudo *et al.* (2004).

1.3.4 PAS domains of sensory proteins in *M. tuberculosis*

PAS domains are sensory input domains first identified as a motif in the *Drosophila melanogaster* period clock protein, the *E. coli* aerotaxis receptor, and the *D. melanogaster* single-minded protein (Huang *et al.* 1993) that typically sense light, oxygen and redox potential (Hefti *et al.* 2004). Although PAS domains are always located on the inside of the cell, it is possible to monitor changes in external stimuli through, for example, alterations in the concentration of metabolites arising from the electron transport chain, as is the case for ArcB (Georgellis *et al.* 2001; Malpica *et al.* 2004). Interestingly, the ArcB sensor from *Haemophilus influenzae* is able to signal redox states even after its PAS domain has been removed (Georgellis *et al.* 2001).

PAS domains often do not have a conserved primary structure and are difficult to identify by sequence alone. For reviews see Taylor & Zhulin (1999) and Gilles-Gonzalez

& Gonzalez (2004). Given the homology of SenX3 to other protein containing PAS domains, J. Saldanha (Mathematical Biology, NIMR) used structural modelling to find that SenX3 contains an atypical, or PAS-like, domain towards the membrane-anchored N-terminus of the protein with homology to the PAS domain of FixL (Miyatake *et al.* 2000), a nitrogen fixation protein from *Rhizobium meliloti* that regulates genes that allow the utilisation of nitrogen as a terminal electron acceptor in the absence of oxygen (Figures 1.2 and 1.4). The PAS domain of FixL senses oxygen concentration through association with a haem prosthetic group, and bears strong homology to a recently-solved PAS domain in the *E. coli* direct oxygen sensor, Dos (Park *et al.* 2004). Dos has a role in mediating the transition of *E. coli* from aerobic to anaerobic metabolism (Delgado-Nixon *et al.* 2000). The homology of the SenX3 PAS domain to the PAS domains in other oxygen-sensing systems further supports a hypothesis for the SenX3-RegX3 two component signal transduction system in oxygen sensing. Attempts have been made by C. Joseph and A. Pastore (Division of Molecular Structure, NIMR) to purify both SenX3 and the SenX3 PAS domain in isolation in soluble form but were unsuccessful, although it has been shown to be possible to express and purify two-component sensory proteins from other organisms (Potter *et al.* 2002).

No definite sensor domains have been proved for DosS, the cogent histidine kinase for DosR, however two potential GAF domains have been identified through homology searching, although the N-terminal GAF domain may in fact be a PAS domain (N. Stoker, personal communication). Recently, two orphan (i.e. not adjacent) gene products have been shown to be able to transphosphorylate, and may be a twelfth two

Figure 1.3: The PAS domain of *Rhizobium meliloti* FixL (A), and the predicted PAS domain of SenX3 (B). Secondary structure is coloured red for helices and yellow for strands, and labelled according to the sequence alignment in Figure 1.4. Note the loss of helices aD and aE leading to a more open structure in the modelled SenX3 PAS domain. Adapted from Rickman *et al.* (2004).

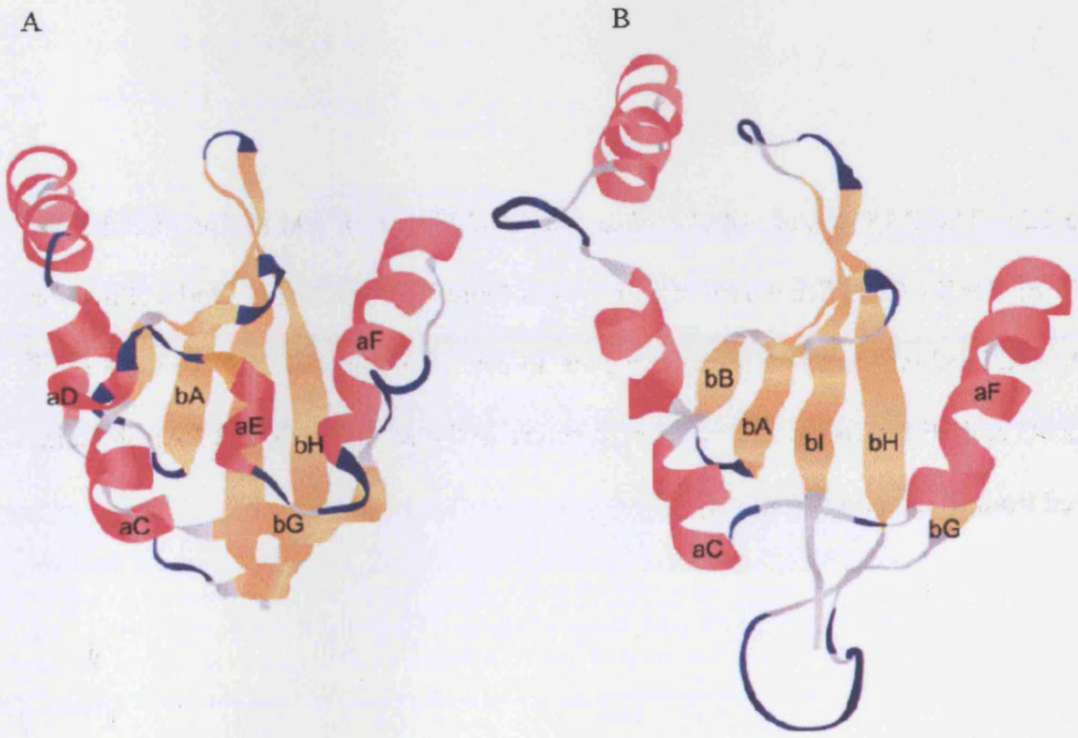
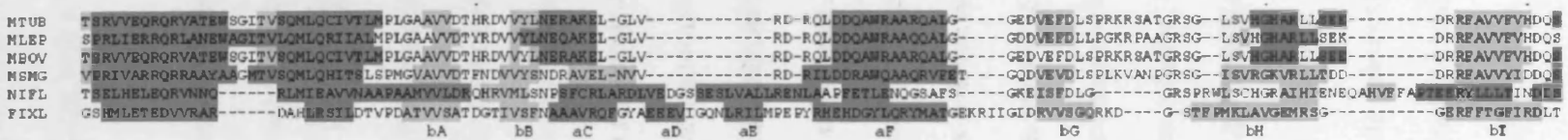

A
aD
bA
aE
bH
aF
aC
bG
B
bB
bA
bI
bH
aF
aC
bG

Figure 1.4: Alignment of the PAS domain regions of SenX3 from *M. tuberculosis*, *M. leprae*, *M. bovis*, *M. smegmatis*, NifL from *Azotobacter vinlandii* and FixL from *R. meliloti*. α-helices, dark shading; β-strands, light shading; A-I, PAS domain region. The structure of the PAS domain of FixL has been solved (Miyatake *et al.* 2000) and represents the actual secondary structure of the domain; secondary structure of the other PAS domains was predicted. Note the missing α-helical elements aD and aE in the mycobacterial PAS domains. Adapted from Rickman *et al.* (2004).


MTUB
MLEP
MBOV
MSMG
NIFL
FIXL
bA
bB
aC
aD
aE
aF
bG
bH
bI

component system of *M. tuberculosis* (Morth *et al.* 2005). The putative sensor protein, encoded by Rv3220c, contains both a PAS and a GAF domain.

1.3.5 *M. tuberculosis* can avoid the normal bactericidal mechanisms of the host

After infection of a host the bacterium is engulfed by alveolar macrophages. Generally with phagocytes, engulfed bacteria are destroyed by acidification of the phagosome in which it resides, fusion of the phagosome with lysosomes and exposure to reactive oxygen and nitrogen radicals in the oxidative burst. Later in the infection other mechanisms act on the bacteria such as oxygen limitation and nutrient starvation within the granuloma, where surviving bacteria encounter low oxygen and carbon levels.

Live tuberculous bacilli have mechanisms that allow them to survive and replicate within the alveolar macrophages that are designed to protect against infection. The first mechanism of mycobacterial immune evasion to be discovered was the ability of the bacillus to prevent phagosome-lysosome fusion (Armstrong & Hart 1971). *M. tuberculosis* can also enter macrophages without activating them, avoiding some or all of the potential damage from the oxidative burst that is stimulated by normal phagocytosis of bacteria. For reviews of the entry mechanisms of *M. tuberculosis* into macrophages and the avoidance of phagosome-lysosome fusion, including recent advances in the field, please see Vergne *et al.* (2004), Pieters & Gatfield (2002) and Pieters (2001).

1.4 The role of oxidative stress during tuberculosis infection

During the later stages of infection *M. tuberculosis* is able to enter a period of metabolic shutdown in this environment, which allows it to remain viable for later

reactivation disease and transmission (Wayne & Sohaskey 2001). There are also suggestions that exposure to low-oxygen conditions increases the virulence of *M. tuberculosis* during infection (Bacon *et al.* 2004).

1.4.3 Oxidative and nitrosative stress in the macrophage

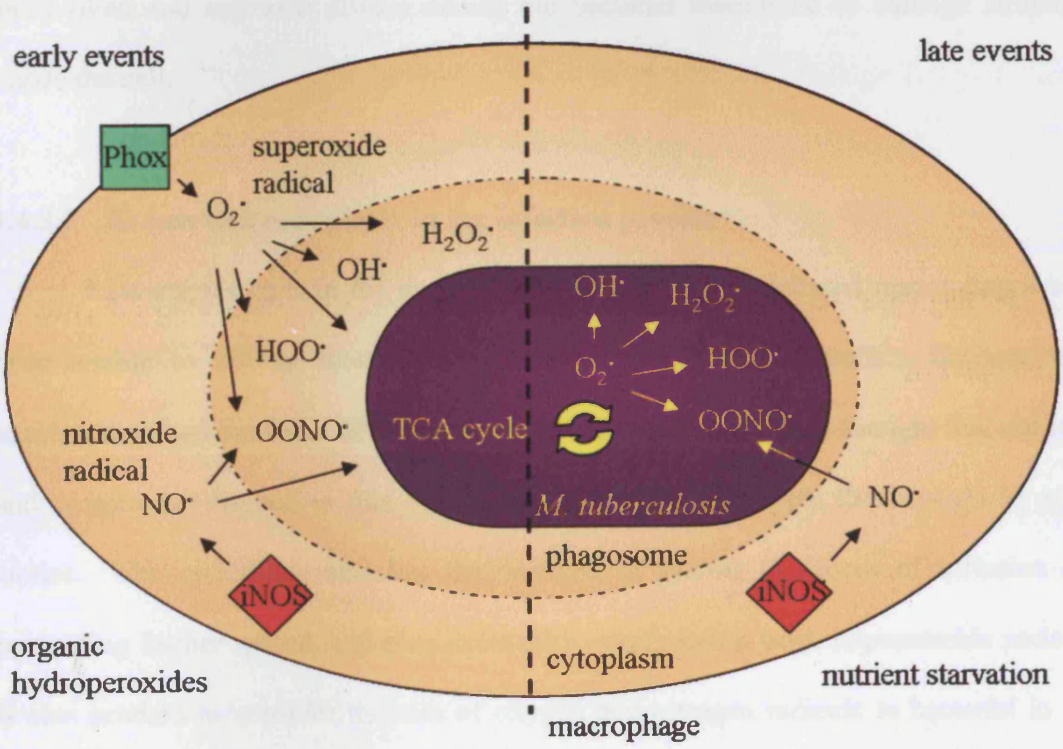
The occurrence of oxidative stress in the macrophage environment can be divided into two broad categories: i) events immediately after infection of the macrophage and ii) late-stage oxidative stress resulting from a multi-cellular response to the infected macrophage. A cartoon representation of the events may be seen as Figure 1.5.

1.4.3.1 Events that occur immediately after infection

The events that occur immediately after infection form the oxidative burst, consisting of short-lived oxygen and nitrogen radicals that primarily damage structures on the surface of the bacterium. The response of the mycobacterium to these agents is critical to its ability to create a productive infection. For an in-depth review of this subject see Nathan & Shiloh (2000).

The cell-wall associated macrophage protein phagocyte oxidase (Phox) is capable of generating the superoxide radical $O_2^{\cdot}$ that can also form reactive oxygen intermediates such as hydrogen peroxide radicals (H_2O_2) and hydroxyl radicals ($OH^{\cdot}$), which interact with organic molecules (often on the bacterial cell surface) to create the highly damaging organic hydroperoxides. The highly charged superoxide radical cannot pass the cell wall and may only damage surface structures when released in the oxidative burst, whereas

Figure 1.5: Diagram illustrating stresses that occur within the macrophage after infection by *M. tuberculosis*.


early events
late events
Phox
superoxide
radical
$O_2^{\cdot}$
$H_2O_2^{\cdot}$
OH·
HOO·
OONO·
nitroxide
radical
NO·
TCA cycle
OH·
$H_2O_2^{\cdot}$
$O_2^{\cdot}$
HOO·
OONO·
M. tuberculosis
NO·
phagosome
iNOS
iNOS
organic
hydroperoxides
cytoplasm
nutrient starvation
macrophage

some of its polar breakdown products can diffuse into the bacterium and degrade the cellular machinery.

Macrophages also rely on the cytoplasmic protein inducible nitric oxide synthase (iNOS) to create the nitroxide radical ($NO^{\cdot}$). The nitroxide radical will also damage structures at the bacterial cell surface, and may combine with the superoxide radical to form peroxynitrite, which is a highly damaging reactive oxygen intermediate that is very short lived and can also diffuse across the bacterial membrane to damage structures inside the cell.

1.4.3.2 Events that occur later in the infection process

Late-stage events in the macrophage occur when the infected macrophages have been unable to kill or control the replication of the mycobacteria. Immune cells aggregate around the focus of infection and form a granuloma. It is thought that nutrients and oxygen are limited in this environment, in an effort to kill the bacteria by siege tactics. The granuloma also has the function of sealing the focus of infection and preventing further spread, and may eventually calcify into a hard, impenetrable node. It is also prudent to consider the role of oxygen and nitrogen radicals to bacterial in this situation.

The tricarboxilic acid (TCA) cycle of bacteria is the favoured source of energy liberation during aerobic growth on carbon sources, since it provides the most efficient yield of energy; however it produces a large amount of superoxide radicals as a byproduct. The damaging effect of the superoxide radicals and their intermediates on the proteinaceous cellular machinery can be fixed by a number of specialised enzymes such

as methionine sulphoxide reductase (Nathan & Shiloh 2000), but is also mitigated by the high turnover of proteins in the rapidly dividing cells. Clearly, if the ability to use this method of energy liberation is not available, through either lack of carbon sources or lack of oxygen to use as a terminal electron acceptor, then the bacterial will grow more slowly and alternative metabolic pathways used. Although an obligate anaerobe, *M. tuberculosis* is able to survive oxygen withdrawal in a nonreplicative state (Wayne & Lin 1982). Interestingly, the fermentation pathway is conserved in *M. tuberculosis* (Cole *et al.* 1998), hinting that the bacillus may have methods of generating energy in the absence of oxygen. The alternative routes to the electron transport chain produce fewer radicals, and since the slower growth of the bacteria means that the protein turnover is lower and there is less energy available for protein fixing, it is important for the bacteria to switch off the TCA cycle, otherwise it will gradually lose viability under these conditions. For a review of this subject please see Nystrom (1998; 2002).

The mycobacteria are also able to evade the late-stage responses by entering a state of latency, of which very little is known *in vivo*. During this time the bacteria are hypothesised to be in a period of metabolic shutdown that has been referred to as "viable but non-culturable". For a review of nonreplicating persistence in *M. tuberculosis* please see Wayne & Sohaskey (2001).

This form of persistence accounts for much of the great burden of tuberculosis on the world population, since around two billion people are infected by *M. tuberculosis* by skin test and yet only a small fraction of these have active tuberculous disease (WHO 2001). Latent tuberculosis may reactivate to spread and cause disease if the individual

carrying latent tuberculosis becomes immunosupressed in some way, such as through HIV infection, malnutrition or immunosupressive cancer therapy.

The specific macrophage-like agents will now be discussed as they apply to the in vitro assessment of *M. tuberculosis*.

1.4.3.3 Low oxygen

One method of discovering whether the SenX3 and RegX3 proteins are acting like ArcA-like is to assess their growth in oxygen-limited culture. Ordinarily, given excess oxygen, the bacteria will grow until all of the available nutrients in the media are exhausted and the byproducts of the metabolism have built up to toxic levels. In rolling bottles this will typically occur at an OD_{600} of four to six. However if the bacteria are maintained in standing cultures with little air included (low headspace ratio, or HSR), then the bacteria will enter stationary phase at a much lower optical density due to a lack of oxygen. For a review of the *E. coli* adaptations to stationary phase please see (Spiro & Guest 1991)

It is possible to hypothesise that if the *senX3* and *regX3* null mutants are acting like the ArcA-B system of *E. coli*, then they will be unable to shut down their TCA cycle, despite the lack of oxygen. In *E. coli*, bacteria lacking the ArcA-B system are unable to shut down their TCA cycle and gradually lose viability when forced to shift to stationary phase by anaerobic growth or carbon starvation (Nystrom *et al.* 1996), and a similar phenotype is seen in *E. coli* lacking superoxide dismutase A (Spiro & Guest 1991; Nystrom *et al.* 1996; Dukan & Nystrom 1999). If superoxide radicals are building up in the *senX3* and *regX3* null mutants of *M. tuberculosis* more than the wild-type as they

enter stationary phase due to oxygen starvation, then one would expect to see slower growth as the bacteria become damaged.

1.4.3.4 Nutrient starvation

Another method of inducing a phenotype in an ArcA mutant of *E. coli* is to withdraw its carbon source and in that way limit aerobic respiration (Nystrom *et al.* 1996). Mutants of *arcA* in *E. coli* dramatically lose viability more than the wild-type after prolonged incubation without nutrients, as the bacteria are no longer replicating quickly and the internal structure of the cells becomes progressively damaged by oxidative stress generated by the still-active TCA cycle. A review of this concept can be seen in Nystrom (1999). The same technique can be applied to the *senX3* and *regX3* null mutant strains of *M. tuberculosis*, in which one would expect the mutants to lose viability more than the wild-type after washing and prolonged incubation in phosphate-buffered saline.

1.4.3.5 Superoxide stress

The most commonly used superoxide generator is paraquat, also known as methyl viologen, gramoxone and 1,1′-dimethyl-4,4′-bipyridinium dichloride. This is dissolved in water and added to liquid cultures of *M. tuberculosis*, where it attacks the outside of the cell. Due to its charge it cannot cross the membrane, but it may be reduced further to the whole variety of reactive oxygen intermediates and therefore represents a very broad attack on the oxidative stress defences of the bacterium. Previous work at NIMR has shown that *M. tuberculosis* exposed to a final concentration of 50mM for 24 hours will

cause a 60-80% drop in the viability of the culture (L. Rand, personal communication). Like most stressing agents, lower concentrations can be used to study the effect on transcription without the large drop in culture viability.

The effect of superoxide on the *senX3* and *regX3* null mutants of *M. tuberculosis* can be studied to see if there is any difference in the ability of either of these strains to resist the stress more than the wild-type. Any difference can be taken as an indication that the gene products have a role in either sensing or coordinating the response to damage by superoxide or its breakdown products and downstream effects.

1.4.3.6 Nitrosative stress

Although there is no strong evidence for the involvement of the SenX3/RegX3 system of *M. tuberculosis* in the direct sensing and response to nitroxide radicals, it is possible to hypothesise that the mutants would be more susceptible to damage by nitroxide radicals if the SenX3/RegX3 system were acting like ArcA of *E. coli*, overproducing superoxide which then interacted with the nitroxide radicals to form peroxynitrite which damaged the cells. The effects of nitrosative stress are thought to be particularly important in *M. tuberculosis,* data for which is reviewed in (Nathan & Shiloh 2000). One way of assessing the ability of the *senX3* and *regX3* null mutants to resist nitrosative stress is to add silver nitrite to bacteria in acidified media to a final concentration of 3mM. Although the bacteria can be grown in normal liquid media, they must be washed and resuspended in media adjusted to pH5.5 to allow the silver nitrite to become labile and release nitrogen.

Assuming the *senX3* and *regX3* gene products do not specifically interact with the reactive oxygen intermediate protection mechanisms (such as superoxide dismutase) then there would be no reason to expect to see a phenotype difference in normal aerobic growth of the *senX3* and *regX3* null mutants. During normal aerobic growth the wild-type, and strains that may hypothetically be deficient in TCA cycle shutdown, would generate the normal amount of superoxide radicals. When nitroxide radicals were added, some of the superoxide radicals from the TCA cycle would be detoxified, and some would combine with nitroxide radicals to damage the cells.

However, after microaerobic growth in which the wild-type had shut down its TCA cycle, mutants deficient in this ability would be expected to generate more internal oxidative stress than the wild-type if the system were acting like ArcA of *E. coli* (Nystrom *et al.* 1996). Excess superoxide radicals would combine with the nitroxide radicals to produce the highly damaging peroxynitrite that would damage the *senX3* and *regX3* null mutants more than the wild-type. Therefore it could be hypothesised that the *senX3* and *regX3* null mutants, if they are generating larger amounts of internal oxidative stress than the wild-type under microaerobic conditions, would be more susceptible to damage by nitrosative stress than the wild-type after microaerobic growth.

1.4.3.7 Organic hydroperoxide stress

The products of the oxidative burst will ultimately combine with organic molecules of the bacteria to form the highly-damaging organic hydroperoxides. The addition of *tert*-butyl hydroperoxide to cultures of *M. tuberculosis* seeks to replicate this end-stage effect of free radical damage. Although the general pathway of cellular

damage is maintained by this method, in other bacteria there are specific responses and regulators that react to these particular methods of damage (Fuangthong *et al.* 2001; Sukchawalit *et al.* 2001; Fuangthong & Helmann 2002; Mongkolsuk *et al.* 2002; Panmanee *et al.* 2002). If SenX3 or RegX3 could be acting in this way then it would be expected for the *senX3* and *regX3* null mutants of *M. tuberculosis* to be more susceptible to damage by *tert*-butyl hydroperoxide than the wild-type.

There are a number of other reagents that form organic hydroperoxides, such as cumene hydroperoxide and plumbagin. *tert*-butyl hydroperoxide is soluble in water at most concentrations, and relatively easy to work with. Cumene hydroperoxide, for example, is supplied dissolved in an organic solvent (cumene) that may itself damage the bacteria. Cumene hydroperoxide is also insoluble in water at the concentrations needed for bactericidal levels against *M. tuberculosis*.

1.4.4 Measurement of stress studies

To look at the growth of a culture after stress it is necessary to take viable counts or optical density readings of stressed cultures, and also of unstressed controls. Paired cultures are used to eliminate problems caused by inherent generation time differences between strains and experiments are performed in triplicate to gain an idea of the reliability of the results.

Taking viable counts involves making ten-fold dilutions of an aliquot of the culture, spreading of samples on agar plates and counting the colonies that grow. Measuring the optical densities of the cultures can also give an indication of the bactericidal level of an agent against a particular strain of *M. tuberculosis* but it is not

possible to tell whether the bacteria are dead or in bacteriostasis unless the cell have lysed, thus lowering the optical density of the culture. Furthermore, some documented responses of *M. tuberculosis* to stress conditions include a thickening of the cell wall (Cunningham & Spreadbury 1998), in which case the optical density of the culture would rise giving the appearance of growth, when in fact bacterial numbers may have stayed stable.

It is important in these types of study to use an appropriate concentration of the damaging agent, enough to cause a reliable drop in the viability of wild-type cultures. Often this can be ascertained by reading the previous literature for the appropriate organism, or by protocols acquired from other workers.

It is also important to ensure that all strains are at the same phase of growth since it is possible that two strains, although at the same optical density, will be at different growth phases and will react differently to the stress. For example, if one culture had recently been diluted from stationary phase then there would be likely to be a large quantity of bacteria in that culture that are still in stationary phase, or that are shifting out of stationary phase and are not growing evenly. This population of bacteria will respond differently to another culture that has grown from a small inoculum to the same optical density but is genuinely in exponential-phase growth. All of the experiments with aerobic cultures of *M. tuberculosis* in this thesis have been performed with early exponential-phase cells.

Although it is often possible to simply add the damaging reagent to the stressed cultures and leave the unstressed controls unmolested, the unstressed controls should properly have the solvent used to dilute the damaging agent added at an equal volume to

that of the stressed cultures. Whereas adding a few hundred microlitres of water to a 100ml culture of *M. tuberculosis* would be expected to make little difference to the viability of the control cultures, some damaging agents are soluble only in organic solvents, such as cumene or ethanol. These solvents would be expected to make a contribution to the growth or viability of the cultures, and must strictly be added to the control cultures.

It is desirable to try and keep the volumes of added reagent similar in an experiment. For example, with a titration of damaging agent against multiple cultures of the same strain, the concentrations of stocks should be adjusted so that equal volumes are added to cultures to give the appropriate range of final concentrations.

1.4.5 Growth of mycobacteria in mouse models of infection

The mouse model of infection is a stalwart of studies of *M. tuberculosis*, enabling screening of strains for virulence. The use of mice allows sufficiently large populations of animals to be housed in a laboratory, both through cost and space considerations. This means that it is possible for a single experiment to easily contain enough animals to allow it statistical validity. Mice are easier to work with than guinea pigs or primates, and are sufficiently inbred to allow consistency of results. Mice are also unable to cough, reducing the potential hazard from the release of airborne *M. tuberculosis* when working with these animals.

The murine immune system and physiology has been the subject of extensive research, which makes the model attractive to those studying pathogenesis and the immunology of *M. tuberculosis* infection. Knockout mice are also available to allow

assessment of mycobacterial virulence in, for example, a mouse deficient in the nitric oxide synthesis pathway.

The early cataloguing of the growth of *M. tuberculosis* in mice was performed by (Pierce *et al.* 1953), who showed that virulent strains of *M. tuberculosis* grow preferentially in the lungs independently of the route of infection. A review of the mouse model of *M. tuberculosis* infection can be seen by Orme & Collins (1994). The infection process as considered by the author follows.

Infection in the system used at NIMR is typically though intravenous injection into the tail vein, after which the mycobacteria disseminate through the murine vascular system to the organs. After 24 hours the bacteria disappear from the bloodstream, having been removed principally by the reticuloendothelial system. Large numbers of bacteria enter the spleen, since it has a high exposure to blood-borne particles and large quantities of phagocytes in which the mycobacteria can replicate. Bacteria also find their way to the lungs, where they replicate rapidly due the high oxygen levels present in lung tissue. Eventually the mice develop loose, poorly structured granulomas in the lung. Some mice are more susceptible to infection by *M. tuberculosis* than others; Balb/c and C57/Bl6 are commonly used, where Balb/C are generally resistant to infection displaying a Th1-type intracellular response to the infection.

Like all models of infection there are a number of factors to be considered when interpreting results.

i) Mice do not develop tight granulomas as seen in human infection. In fact, there is evidence to suggest that the predominant form of mycobacterial control in mice is exposure of the mycobacteria to nitric oxide (Chan & Flynn

1999), whereas in humans it is considered that oxygen and nutrient limitation control mycobacterial spread as a function of the dense granulomas that form around the focus of infection.

ii) The wild-type strains of *M. tuberculosis* used in the laboratory and for mouse infection were originally isolated from human infections (Steenken & Gardner 1946). They were chosen as laboratory reference strains due to their maintenance of colony morphology during repeated passage, something that was unusual amongst the clinical strains of *M. tuberculosis* examined for this purpose. Many laboratories routinely passage their reference strains through mice so that they retain virulence in this model, although it is likely that the genomic structure of the bacterium in changing somewhat.

iii) Human infection by *M. tuberculosis* is predominantly by aerosol infection, whereas mice are usually experimentally infected by the intravenous route. Several groups are making efforts to create reliable aerosol infection models for mice, but the size, expense and safety considerations of the necessary equipment means that this model is not routinely used. It is also worth bearing in mind that *M. tuberculosis* given to mice by the aerosol route are more virulent than when given intravenously (North 1995).

iv) Mice are typically infected intravenously with much larger doses of mycobacteria than a normal human aerosol infection; at least hundreds and frequently orders of magnitude more than that. Humans, however, may have an infectious dose of just a few mycobacteria inhaled from the environment.

This may have implications for the pathogenesis and host response to the bacilli.

Although the murine model of infection is not a perfect mirror of human tuberculosis, it is necessary, and adequate, for initial screening of mutants of *M. tuberculosis* for hypervirulence or attenuation.

1.4.6 Macrophage models of *M. tuberculosis* infection

Macrophage infections act a fast, convenient assay for screening strains of *M. tuberculosis* for attenuation, virulence or other *in vivo* phenotypes when there is not the time, space or desire to set up a long-running mouse infection. Macrophage infections can use any one of a variety of cells.

Ideally macrophages would be isolated from human lung tissue by alveolar lavage; these cells would be responding to tuberculosis infection in the most natural way although they would still be isolated from other compartments of the immune system, such as components of the acquired immune response. Unfortunately human alveolar macrophages obtained by lavage are difficult to get hold of and one would expect cells from different individuals to have different responses. Alternately macrophages can be isolated from the periphery by taking blood samples, but this may not yield a satisfactory number of cells and is likely to suffer from the same person-to-person variation in response.

The use of cell lines is a convenient method of obtaining large numbers of macrophage-like cells at low cost. They are available from human and mouse backgrounds, and cell lines are available that are deficient in various biosynthesis and

immunological pathways. These knockout cell lines enable very specific questions to be asked of the interaction between the eukaryotic cell and the mycobacterium. The downside of using cell lines is that they do not represent a good model of infection. The cells have usually been through multiple passages, and may not have had the developmental cues necessary to allow them to form a proper response to infection. Furthermore, some cell lines are multinucleate or have abnormal numbers of chromosomes as a result of transformation, and will not act like their mortal parent cell types.

A compromise between convenience and likeness to a true infection can be reached by using primary cells from mice. Large numbers of cells can be recovered from the thigh bones of mice and cultured to enrich them for macrophages. These cells will have a normal genetic background and respond largely in the same way as they would *in vivo*.

However, since they are derived from the thigh of the mouse they may not be fully developed into the sort of cell that would encounter a macrophage *in vivo*. In addition, they are still removed from the acquired immune response and may lack some of the infection cues that would be necessary for completely normal *in vivo*-like reactions. Cell layers also do not develop granulomas, so it is unlikely that any infecting mycobacteria would encounter oxygen stress. The infection is short-lived as bone marrow derived macrophages seldom survive more than two weeks *in vitro*. The latter two points mean that the bacteria would not experience events that occurred later on in the infection.

Due consideration is seldom given to the activation state of these macrophages – whether they can be activated in the normal way following their liberation from the thigh bones and possible immaturity, and whether the extensive physical manipulation of the macrophage population could cause them to activate or move them into an anergic state.

However for the purposes of initial screening for attenuation of mutant phenotypes this system is perfectly adequate as long as one remembers that certain phenotypic defects may be missed. If a null mutant of *M. tuberculosis* is known to be attenuated in macrophage models of infection, then this form of experiment is a convenient assay for complementation and allows further, detailed examination of the interaction between the mycobacteria and the host cell.

1.5 Allelic exchange

Genetics relies on the study of the function of genes through the isolation of mutants. Traditional methods of disrupting genes in bacteria relied on mutagenisis of the chromosome by chemicals or transposons and analysing for the desired phenotypes amongst the pool of heterogeneous bacteria. The discovery ways of using homologous recombination for alleleic exchange meant that genes could selectively be 'knocked out' genes to create stable mutant strains. Current techniques made possible by whole-genome sequencing allow the use of 'reverse genetics', where the sequence of the gene to be knocked out is known and the mutants do not need to be 'selected' in random screens. This allows isolation of strains carrying mutations in genes that do not present a selectable phenotype. The sequencing of the genome of *M. tuberculosis*, in conjuction with the development of powerful selective markers for use in the mycobacteria, now

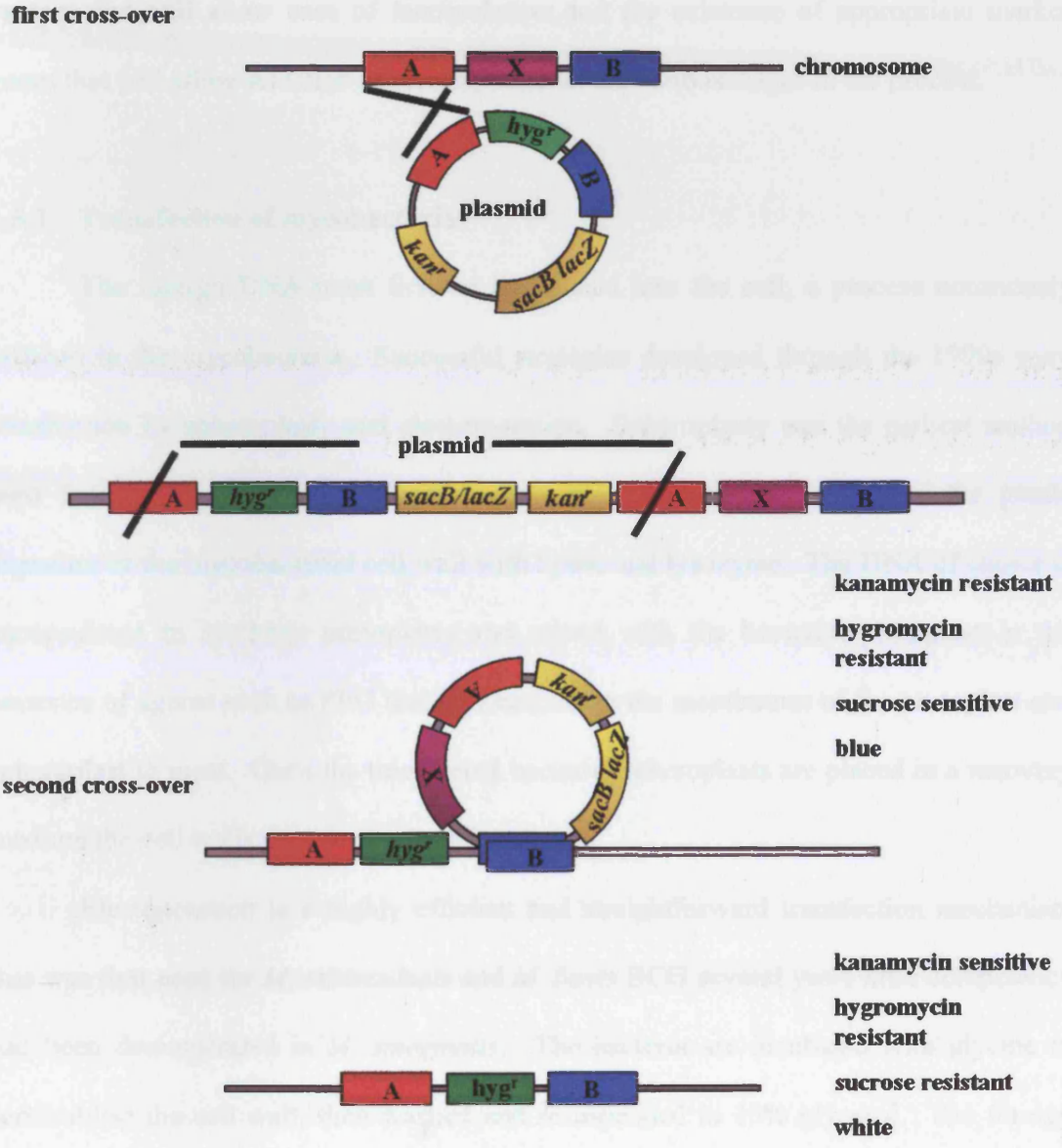
allow knockouts to be made to study the contribution of particular genes. This was not possible before with *M. tuberculosis* since the organism grows too slowly for the traditional genetic methods to be used, and selective markers were not identified until fairly late. Even in *E. coli*, using traditional genetic approaches, it was only possible to identify ~30% of the genes using traditional mutant screens.

1.5.1 Homologous recombination

Homologous recombination can occur when a cell is diploid for two identical copies of a stretch of DNA. It is an extension of the DNA damage repair mechanisms of bacteria, where damaged alleles are replaced with functional copies. Some species such as *Streptococcus pneumoniae* use the mechanism to develop a kind of 'community gene pool' where certain virulence factors are flanked by highly homologous regions and the bacteria are naturally competent.

In molecular cloning homologous recombination is used to exchange the normal chromosomal allele of a cell for one chosen by the investigator. The replacement allele may have been disrupted, mutagenised, marked in some way or may be a functional allele to complement a gene that has already been knocked out. In its simplest form, homologous recombination is a single cross-over event mediated by an assortment of enzymes. The cell will now be diploid for the region introduced to the cell, with the two alleles adjacent to each other on the chromosome and both reading in the same direction (Figure 1.6). To be of greatest use to the investigator, however, a second cross-over event will allow allelic exchange to take place.

Figure 1.6: Diagram representing genotypic changes of the chromosome of *M. tuberculosis* during selection of an allelic exchange mutant.


first cross-over
A
X
B
chromosome
hygr
A
B
plasmid
kanr
sacB lacZ
plasmid
A
hygr
B
sacB/lacZ
kanr
A
X
B
kanamycin resistant
hygromycin resistant
sucrose sensitive
blue
second cross-over
A
X
kanr
sacB lacZ
A
hygr
B
kanamycin sensitive
hygromycin resistant
sucrose resistant
white
A
hygr
B

1.5.2 Development of allelic exchange in the mycobacteria

There are three major barriers to the development of routine allelic exchange in an organism: the difficulty of introducing foreign DNA into the host cell, the existence of vectors that will allow ease of manipulation and the existence of appropriate marker genes that will allow selection of recombinants at the various stages in the process.

1.5.3 Transfection of mycobacteria

The foreign DNA must first be introduced into the cell, a process notoriously difficult in the mycobacteria. Successful strategies developed through the 1990s were transfection by spheroplasty and electroporation. Spheroplasty was the earliest method used for studies of DNA introduction into *M. smegmatis*, and involves the gentle digestion of the mycobacterial cell wall with lipase and lysozyme. The DNA of choice is encapsulated in synthetic protoplasts and mixed with the bacterial protoplast in the presence of agents such as PEG that will encourage the membranes of the protoplast and spheroplast to meet. Once the transfected bacterial spheroplasts are placed in a recovery medium the cell walls spontaneously reform.

Electoporation is a highly efficient and straightforward transfection mechanism that was first used for *M. tuberculosis* and *M. bovis* BCG several years after competency had been demonstrated in *M. smegmatis*. The bacteria are incubated with glycine to permeablise the cell wall, then washed and resuspended in 10% glycerol. The foreign DNA is then mixed with the cells and an electrical current passed through the DNA-cell

mix. After a recovery period the transformants may be selected by use of the appropriate markers.

1.5.4 Development of systems for genetic manipulation of alleles in the mycobacteria

Effective tools for genetic manipulation of alleles for gene replacement in the mycobacteria must be able to replicate in *E. coli*. Much of the early work on DNA introduction into mycobacteria focussed on the expression of antigens in *M. bovis* BCG, and required a mycobacterial origin of replication in addition to that of the *E. coli* origin of replication (Jacobs *et al.* 1987; Snapper *et al.* 1988). This work drove the development of markers for genetic manipulation in the mycobacteria. Allelic exchange on the other hand does not require a mycobacterial origin of replication, although such systems have recently been developed (Pashley *et al.* 2003), but the development of the allelic exchange vectors did rely very much upon the discoveries made with the shuttle vectors.

The first vectors for alleleic exchange in the mycobacteria were very simple plasmids that had few cloning sites and markers borrowed from the shuttle expression vectors (Marklund *et al.* 1995; Reyrat *et al.* 1995). In theory a vector for allelic exchange can be very simple – it requires an *E. coli* origin of replication, a cloning site that allows a marked allele to be inserted and a counterselectable marker elsewhere on the plasmid. Selective pressure will allow only plasmids that have integrated into the chromosome of the mycobacteria to grow, and remaining episomal DNA will be lost from the progeny.

The early successful reports of allelic exchange in the mycobacteria used custom-made plasmids for each occasion, and often relied on the auxotrophic effects of the

mutation being studied, such as in Husson *et al.* (1990). The development of a cloning system by Parish & Stoker (2000) was one example of a rational approach that allowed the process of making knockouts of *M. tuberculosis* to become more routine. The flexibility it offered also simplified the process of exchange with unmarked alleles, where in-frame deletions could be made on the mycobacterial chromosome that did not abolish transcription of downstream genes in an operon.

The key feature of the pNIL system developed by Parish & Stoker (2000)was that the manipulation of the alleles was separated from the cloning of marker genes, which could be added later in the process from a series of pre-prepared plasmid cassettes. Multiple cloning sites allowed a wide choice of restriction enzymes to be used, eliminating problems with restriction sites in the cloned *M. tuberculosis* sequences and allowing directional cloning of fragments as a matter of course. The negative marker genes were designed to be added last to the construct and were cloned into a *Pac*I site, for which no recognition sequences exist in *M. tuberculosis*.

The two-step selection strategy allowed the isolation of mutants with no obvious phenotype and increased the likelihood of generating a second cross-over. With a two-step strategy only one colony with a single cross-over was necessary for clonal amplification and selection for the second cross-over event. By this stage the various markers used in early experiments had been verified and optimised by the addition of, for example, strong mycobacterial promoters.

1.5.5 Marker genes used in generating gene knockouts in *M. tuberculosis*

In making a mutation of *M. tuberculosis*, both positive and negative selective markers are used. In a marked mutation, an antibiotic resistance gene is cloned in between amplified fragments up- and downstream of the gene to be knocked out to create a mutant allele that contains both the antibiotic resistance gene and portion of the vector sequence. Hygromycin and gentamycin resistanc cassettes are often used for this purpose and can easily be obtained commercially or from other plasmids. Kanamycin resistance markers should not be used for creating marked mutations of *M. tuberculosis* since the background level of mutation to kanamycin resistance is quite high in mycobacteria and would be likely to obscure rare recombination events.

The counterselective markers used for assessing the elimination of the vector after the second recombination event are inserted into the plasmid away from the cloning sites. An example is the *sacB*/*lacZ* fragment from pGOAL17 where both genes are cloned into the final vector simultaneously. *sacB* is a gene from *Bacillus subtilis* coding for levansucrase, which catalyses sucrose hydrolysis and levan synthesis and is also capable of adding fructosyl residues to a variety of other molecules (Dedonder 1966). Expression of levansucrase causes a lethal phenotype in the presence of sucrose and this allele has been fused to the Hsp60 promoter, which is a strong constitutive promoter in *M. tuberculosis*.

lacZ codes for the β-galactosidase gene of *E. coli*, which cleaves the colourless substrate X-gal (5-bromo-4-chloro-3-indolyl-β–galactopyranoside) into galactose and an insoluble blue product that gives expressing bacterial colonies an intense blue colour. A *lacZ* homologue is not present in the chromosome of *M. tuberculosis*, which would preclude its use as a marker in this organism. The *lacZ* gene in the pGOAL system has

the mycobacterial antigen 85 promoter to drive its expression, since the native *E. coli* promoter is non-functional in *M. tuberculosis*.

1.5.6 The knockout process in practice

Making allelic exchange mutants is an extremely long process, and it is important to be diligent throughout the process. The use of a viable strain of *M. tuberculosis* and certainty of the plasmid structure prevent any additional loss of time, as does careful planning during the selection process, the freezing of colonies at all stages and the careful and thorough analysis of potential alleleic exchange mutants at the end of the process. It is important to add supplements that might be necessary in the media to allow growth of a potential alleleic exchange mutant. For example, a lysine auxotroph would require supplementary lysine in minimal media otherwise it will be impossible to isolate alleleic exchange mutants. By the same reasoning deletion mutants of genes essential for the growth of *M. tuberculosis* will be impossible to isolate, although high-density mutagenesis experiments on *M. tuberculosis* allow this to be predicted in advance.

The process described here is specific for generating allelic exchange mutants by the p2NIL two-step strategy, although it is roughly applicable for other knockout systems of *M. tuberculosis* and one-step allelic exchange strategies. It may be helpful to refer to Figure 1.6.

1.5.7 Electroporation of the vector into *M. tuberculosis*.

The plasmid construct is prepared as appropriate (see above), and then sequenced at the cloned regions to confirm the orientation and identity of the inserts. A high-quality

DNA prep, such as a maxiprep (Qiagen), is then made of the plasmid since transformation and selection for knockouts has shown greatly reduced efficiency with vector DNA purified by small-scale preparation.

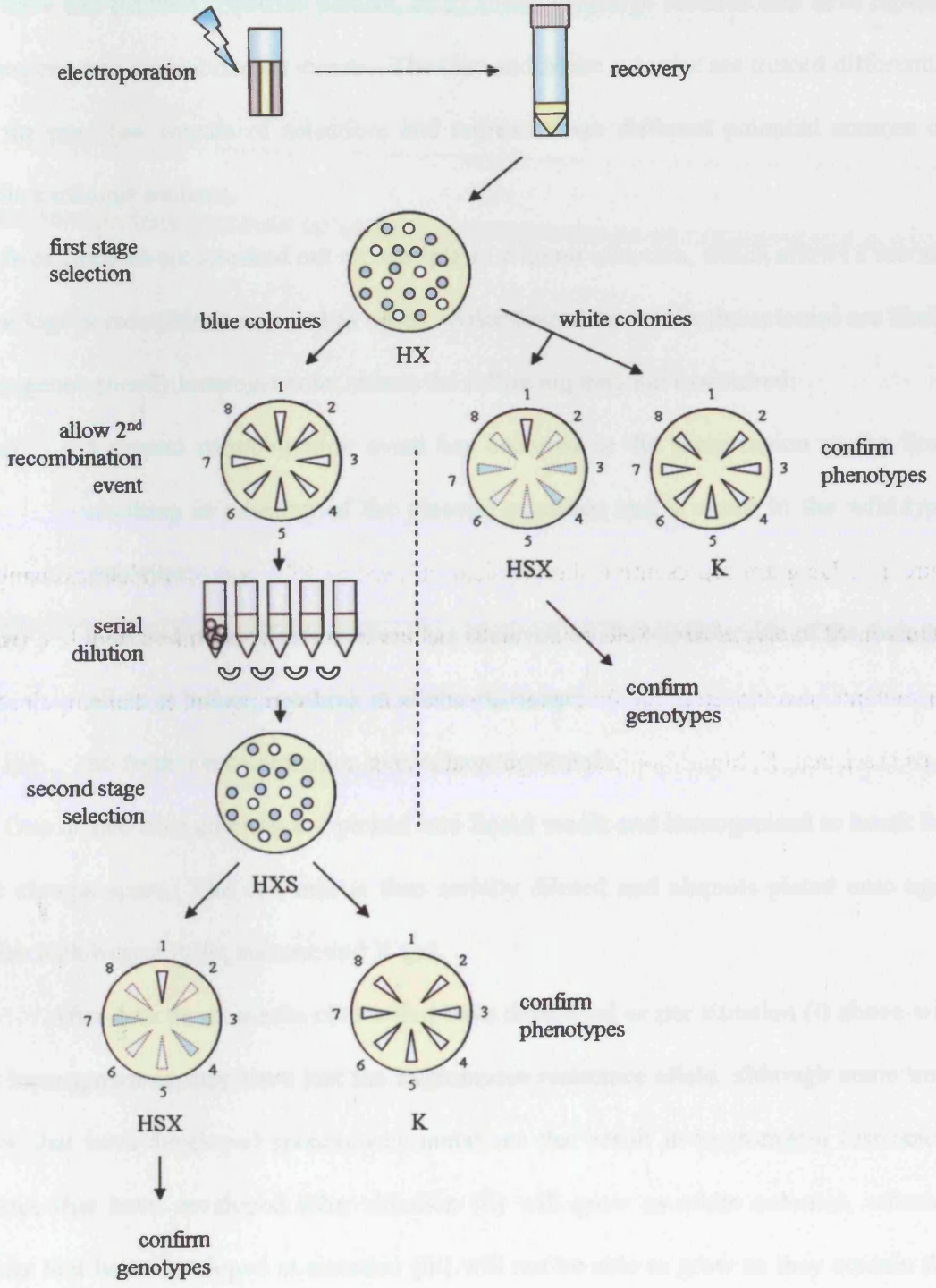
The wild-type *M. tuberculosis* is grown from a single colony obtained from stocks, and made electrocompetent (see 2.1.1.4.8). The wild-type *M. tuberculosis* used for constructing new mutants was the fully sequenced H37Rv strain, obtained directly from Dr. S. Cole at the Institut Pasteur (Paris). See figure 1.7 for reference.

Once the vector DNA and competent *M. tuberculosis* are prepared, the *M. tuberculosis* cell suspension is mixed with plasmid DNA and electroporated. The transformed cells are then allowed to recover overnight, before plating on agar plates containing hygromycin and X-gal. This selection should allow the growth of all cells that have undergone homologous recombination, either by one or two recombination events.

1.5.8 Selection for recombinants

After four to six weeks blue and white colonies are present. The blue colonies carry the integrated plasmid as a result of one homologous recombination event, meaning that they are sucrose sensitive, kanamycin and hygromycin resistant and metabolise X-gal. Although kanamycin could also be used as a selective agent at this stage, its use would prohibit the growth of colonies that have undergone two recombination events and resulted in allelic exchange. The white colonies are either a) spontaneous mutants in the

Figure 1.7: Diagram representing the physical process of selection of alleleic exchange mutants in *M. tuberculosis*. Selection usually occurs in two stages, unless double cross-over mutants have arisen at the first stage of selection. H, hygromycin, K, kanamycin, S, sucrose resistant, X, blue colonies on X-gal agar plates.


electroporation
recovery
first stage selection
blue colonies
white colonies
HX
allow 2nd recombination event
confirm phenotypes
HSX
K
serial dilution
confirm genotypes
second stage selection
HXS
confirm phenotypes
HSX
K
confirm genotypes

wild-type *M. tuberculosis* that have resulted in a hygromycin-resistant phenotype and do not have any plasmid sequence present, or b) allelic exchange mutants that have rapidly undergone two recombination events. The blue and white colonies are treated differently for the next few rounds of selection, and represent two different potential sources of allelic exchange mutants.

Blue colonies are streaked out on agar plates without selection, which allows a second homologous recombination event to occur. After four to six weeks the colonies are likely to be genotypically heterogeneous, where the following may have occurred:

i) a second recombination event has occurred in the same region as the first, resulting in excision of the plasmid sequence and a return to the wild-type genotype;

ii) a second recombination event has occurred on the opposite side of the mutated allele as before, resulting in allelic exchange;

iii) no further recombination events have occurred.

One or two blue colonies are picked into liquid media and homogenised to break the cell clumps apart. The cell mix is then serially diluted and aliquots plated onto agar plates with hygromycin, sucrose and X-gal.

After four to six weeks clones that have developed as per situation (i) above will not have grown as they have lost the hygromycin resistance allele, although some may grow that have developed spontaneous mutations that result in hygromycin resistance. Clones that have developed after situation (ii) will grow as white colonies, whereas clones that have developed at situation (iii) will not be able to grow as they contain the counterselection gene for sucrose sensitivity. It is also possible that clones from scenario

(i) and (iii) have developed spontaneous mutations for hygromycin resistance in the case of (i) and (iii) and/or sucrose resistance, kanamycin sensitivity and *lacZ* activity in the case of situation (iii).

Colonies that appear to display the correct phenotype arising from situation (ii) are checked by streaking on separate agar plates containing either hygromycin, X-gal and sucrose or kanamycin alone. True allelic exchange mutants should grow as white colonies on the hygromycin, X-gal and sucrose agar plates and should not grow on the kanamycin agar plates. Potential allelic exchange mutants are then examined by genotypic methods.

White colonies are assayed directly for allelic exchange by again streaking the colonies onto agar plates containing either hygromycin, X-gal and sucrose or kanamycin alone. Clones that grow well as white colonies on the first selection plate but not on the latter can be assumed to be potential allelic exchange mutants, and are examined further by genotypic methods.

It can be seen here that the two-step selection strategy employed with the colonies that are blue from the initial screen of transformants represents a more logical and 'safer' approach to creating allelic exchange mutants. It is however still desirable to examine potential spontaneous allelic exchange mutants that appear at the initial screen due to the considerable amount of time saved.

1.5.9 Genotypic methods of analysis of potential allelic exchange mutants

1.5.9.1 Polymerase chain reaction

Potential allelic exchange mutants are first examined by polymerase chain reaction (PCR). PCR needs no introduction since its use is widespread and there are exhaustive references on the subject. Crude DNA preparations can be made directly from the agar plates used to confirm the colony phenotypes using a rapid and inexpensive process (Instagene). With well-designed primers to amplify regions of the region in which homologous recombination was expected to occur it is possible to use the different sizes of the products as a cheap, rapid and effective screen for appropriate genotypes.

1.5.8.2 DNA microarray analysis

Clones that appear to have the correct genotype are grown in liquid media, and high-quality DNA preparations are made (see 2.4.1.2). Whole-genome analysis can then be carried out using DNA microarrays (see 2.5), where the genomic DNA from the potential allelic exchange mutant is labelled using a cyanine dye (Amersham) and hybridised on a microarray against wild-type genomic DNA labelled with a complementary cyanine dye. The resultant hybridisation levels are then visualised by a dual-wavelength laser scanner equipped with appropriate filters and linked to a computer.

If the clone has arisen by allelic exchange then there will be no labelled DNA for the exchanged gene, and there will be a striking difference in hybridisation intensities for that gene when compared to the wild-type, provided that the DNA probe on the microarray hybridises within the deleted region. This will reveal itself as a deviation away from one when the relative hybridisation intensities of the wild-type genes are compared to those of the potential allelic exchange mutant and normalised to the majority of other genes present.

Clones that have arisen from a single legitimate recombination event will show no difference in the hybridisation levels of the target gene when compared to the wild-type, but features on the microarray that were present in the *M. tuberculosis* sequences amplified and cloned into the plasmid would be expected to hybridise to the appropriate spots on the microarray at twice the level of the wild-type, since the clone is diploid for these regions. Clones resulting from single illegitimate recombination events can also be distinguished by looking at any diploid regions that appear in the microarray.

Clones that have arisen by spontaneous mutation to the allelic exchange mutant phenotype would not be expected to show any significant differences in hybridisation intensity for any of the genes when compared to the wild-type, unless the mutation abolished hybridisation of the mutant gene cDNA to the appropriate wild-type allele probe present on the microarray slide. In either event, these clones would be clearly distinguishable from the desired mutant phenotype.

1.5.8.3 Sequencing

The *de facto* confirmation of an allelic exchange mutant is provided by amplifying the exchanged region and sequencing along its length until the conversion of wild-type sequence to the vector sequence is found and *vice versa*. It is possible to use sequencing alone to confirm the knockout by beginning the sequencing in regions outside of those cloned into the knockout vector, but in practice one also relies on the PCR and microarray data to minimise the amount of sequencing needed.

Suffice to say, the agreement of PCR, microarray and sequencing data means that it is possible to be absolutely certain of a correct allelic exchange mutant genotype, should one arise.

1.6 Project aims

The aim of this research project is to examine the oxidative stress responses of *M. tuberculosis* through the phenotype and transcriptomic study of mutants in the SenX3-RegX3 two component signal transduction system, the interaction of purified recombinant RegX3 protein with potential genetic regulatory regions, and the isolation of mutants for two other genes thought to be involved in the oxidative stress responses of *M. tuberculosis*, Rv0465c and Rv1049.

2 Materials and methods

2.1 Strains and growth conditions

2.1.1 *M. tuberculosis* strains

2.1.1.1 Wild-type *M. tuberculosis*

The wild-type *M. tuberculosis* designated H37Rv was isolated from a human infection in 1908 and became popular since it did not lose infectivity for mice or colony morphology during repeated passage (Steenken & Gardner 1946). Due to subsequent divergence of the use of this strain, two variants commonly now used are H37Rv (Paris) and Erdman. The H37Rv strain used in this study came from two sources. One strain of H37Rv (Paris) has been maintained at the NIMR for several years and was used to generate both the *senX3* and *regX3* null mutants. Experiments using these two strains compare them to this parent strain. During 2003 there was concern over the mouse infectivity of the wild-type H37Rv (Paris) kept at the NIMR, which was probably due to a change in the media used to prepare the strains for infection. A second H37Rv (Paris) strain that had been used for the *M. tuberculosis* sequencing project was obtained directly from S. Cole (Institut Pasteur) to act as a parent strain for the generation of knockouts in the Rv0465c and Rv1049 genes. Experiments using the Rv0465c and Rv1049 null mutants compare them to this wild-type parent strain. No phenotypic differences have been seen between the two wild-type H37Rv strains used in this study, although differences have been seen by others (A. O'Garra, personal communication). See table $$ for a list of mycobacterial strains used in this study.

2.1.1.2 Null mutants and complementing strains

2.1.1.2.1 Δ*senX3*

Table 2.1 Mycobacterial strains used in this study

strain	genotype	reference
H37Rv	wild-type	Cole *et al* (1998)*
senX3 null mutant	H37Rv Δ*senX3*	Rickman *et al* (2004)
regX3 null mutant	H37Rv Δ*regX3 hyg*^{r}	Rickman *et al* (2004)
Rv0465c null mutant	H37Rv ΔRv0465c	this study
Rv1049 null mutant	H37Rv ΔRv1049	this study
Tame15	H37Rv Δ*regX3*	Parish *et al* (2003)**
complementing *senX3* null mutant	H37Rv Δ*senX3*/pDH12	Rickman *et al* (2004) and this study
complementing *regX3* null mutant	H37Rv Δ*regX3 hyg*^{r}/pDH12	Rickman *et al* (2004) and this study
Rv0465c complemented mutant	H37Rv ΔRv0465c/pDH30	this study
Rv1049 complemented mutant	H37Rv ΔRv0465c/pDH31	this study
regX3 overexpression with acetamidase promoter	H37Rv Δ*regX3*/pDH27	this study
regX3 overexpression with acetamidase promoter control strain	H37Rv Δ*regX3*/pACE	this study
regX3 overexpression with HSP60 promoter	H37Rv Δ*regX3*/pDH18	this study
regX3 overexpression with HSP60 promoter control	H37Rv Δ*regX3*/pMV261	this study

*Kindly supplied by S. Cole (Institut Pasteur) **Kindly supplied by T. Parish (Royal London Hospital)

A hygromycin-marked *senX3* null mutant had been constructed previously (L. Rickman, personal communication). In this strain the region from 579333 to 580542 bp on the *M. tuberculosis* genome containing the *senX3* coding sequence is replaced by a hygromycin resistance cassette that creates a polar mutation in the downstream, co-transcribed *regX3* gene.

2.1.1.2.2 Δ*regX3*

An unmarked *regX3* null mutant was previously constructed in our laboratory (Rickman *et al.* 2004). In this strain the region from 580658 to 581724 bp on the *M. tuberculosis* genome containing the coding sequence for *regX3* and 254 bp of the 3' end of the Rv0492c gene was replaced by around 300 bp of vector sequence creating a gene deletion.

2.1.1.2.3 ΔRv0465c

A hygromycin-marked Rv0465c null mutant is constructed in this study. The region between 554896 and 556312 bp on the *M. tuberculosis* chromosome containing the coding sequence of the Rv0465c gene was exchanged with a hygromycin resistance cassette.

2.1.1.2.4 ΔRv1049

A hygromycin-marked Rv1049 null mutant is constructed in this study. The region between 1172388 and 1172846 bp on the *M. tuberculosis* chromosome containing

the coding sequence of the Rv0465c gene was exchanged with a hygromycin resistance cassette.

2.1.1.2.5 Tame15

Tame 15 is a *regX3* null mutant strain of *M. tuberculosis* created in another laboratory (Parish *et al.* 2003). This strain carried an unmarked deletion that encompasses the 3' end of *senX3*, most of the coding sequence for *regX3* and the MIRU (Supply *et al.* 1997) region separating the two genes (Figure 2.1).

2.1.1.3 Overexpressing strains

2.1.1.3.1 ΔregX3/P_{hsp60}-*regX3*$^+$

To overexpress RegX3 in a *regX3* null strain, the coding sequence of *regX3* was amplified using the primers 152 and 153 (Appendix 3) and the Pfu Ultra enzyme (see 2.3.1) and blunt-end-cloned using the Zero Blunt kit (Invitrogen) into pCR4-TOPO (see Table 2.2 Plasmids used in this study The insert was sequenced before being subcloned into the *Bam*HI-*Cla*I site of pMV261 to create pDH18. pMV261 is a plasmid that allows the strong, constitutive P_{hsp60}-driven expression of foreign genes in mycobacteria (Stover *et al.* 1991, and Table 2.2). pDH18 was then electroporated into the *regX3* null strain of *M. tuberculosis* described in 2.1.1.4.8. The plasmid was maintained by growth of the strain under kanamycin selection.

2.1.1.3.2 ΔregX3/P_{hsp60}-*regX3*$^-$

Figure 2.1: Genetic arrangement of the Tame15 mutant of *M. tuberculosis*. Δ represents the region of deletion. Adapted from Parish *et al.* (2003).

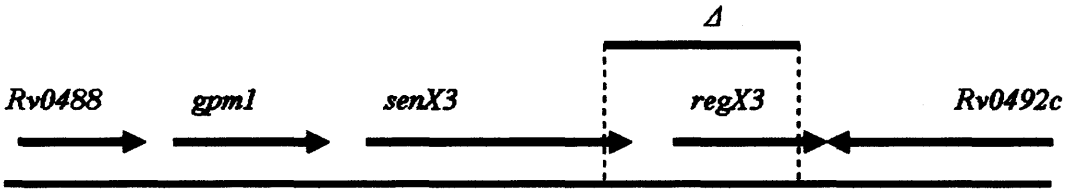

Δ
Rv0488
gpm1
senX3
regX3
Rv0492c

Table 2.2 Plasmids used in this study

plasmid	genotype	reference
pDH1	pCR4-TOPO *senX3 regX3*	L. Rickman
pDH2	pCR4-TOPO + (1619bp 3' of Rv1049)	This study
pDH3	pCR4-TOPO + (1674bp 5' of Rv0465c)	This study
pDH4	pCR4-TOPO + (1601bp 3' of Rv0465c)	This study
pDH5	pCR4-TOPO + (1500bp 5' of Rv1049)	This study
pDH6	1656bp *Bam*HI fragment from pDH3 cloned into the *Bam*HI site of p2NIL-H	This study
pDH7	1512bp *Bam*HI fragment of pDH5 cloned into the *Bam*HI site of p2NIL-H	This study
pDH9	1609bp *Not*I fragment from pDH4 cloned into the NotI site of pDH6	This study
pDH10	1623bp *Not*I fragment of pDH2 cloned into the *Not*I site of pDH7	This study
pDH12	*attP kan*^{r} *senX3 regX3*	This study
pDH14	pDH10 + *lacZ sacB Pac*I fragment of pGOAL 17	This study
pDH18	pMV261 + *regX3*	This study
pDH19	pDH9 *lacZ sacB* (*Pac*I fragment of pGOAL 17)	This study
pDH27	pACE *regX3*	This study
pDH28	pCR4-TOPO Rv0464c Rv0465c	This study
pDH29	pCR4-TOPO Rv1049 Rv1050	This study
pDH30	pKP186 + (Rv0464c Rv0465c from pDH28)	This study
pDH31	pKP186 + (Rv1049 and Rv1050 from pDH29)	This study
pKP186	*oriE attP kan*^{r}	K.G. Papavinasasunderam
pBS-Int	*int amp*^{r}	B. Springer
p2NIL-H	*oriE kan*^{r} *hyg*^{r}	Parish & Stoker (2000)
pGOAL17	P_{Ag85}-*lacZ* P_{hsp60}-*sacB* flanked by *Pac*I sites, *amp*^{r}	Parish & Stoker (2000)
pACE	*oriE oriM* acetamidase promoter, *hyg*^{r}	Triccas *et al* (1998)
pMV261	*oriE, oriM, groEL2* (Hsp60) promoter, *kan*^{r}	Stover *et al* (1991)
pCR4-TOPO	*oriE lacZ ccdB kan*^{r} *amp*^{r}	Invitrogen (www.invitrogen.com)

pET 15b	*oriE amp^r lacI*	Novagen (www.emdbiosciences.com)

This strain is identical to ΔregX3/P_{hsp60}-*regX3*$^+$ but does not contain the coding sequence for *regX3*.

2.1.1.3.3 ΔregX3/P_{ace}-*regX3*$^+$

To overexpress RegX3 in a *regX3* null strain using a different promoter, the coding sequence of *regX3* was amplified using the primers 151 and 168 (Appendix 3) and the Pfu Ultra enzyme (see 2.3.1) and blunt-end cloned using the Zero Blunt kit (Invitrogen) into pCR4-TOPO (Table 2.2), after which the insert was sequence and then subcloned into the *Bam*HI-*Cla*I site of pACE to create pDH27. pACE is a plasmid that allows acetamide-inducible expression of foreign genes in *M. smegmatis* (Triccas *et al.* 1998) but produces constitutive expression of genes at around half the level of that provided by P_{hsp60} in *M. tuberculosis* (E. Davis, personal communication). This plasmid was electroporated into the *regX3* null mutant of *M. tuberculosis* described in 2.1.1.4.8 and the plasmid was maintained by growth of this strain under hygromycin selection.

2.1.1.3.4 ΔregX3/P_{ace}-*regX3*$^-$

This strain is identical to ΔregX3/P_{ace}-*regX3*$^+$ but does not contain the coding sequence for *regX3*.

2.1.1.4 Media and growth conditions for *M. tuberculosis*

2.1.1.4.1 Media

All liquid cultures of *M. tuberculosis* strains referred to in this work were grown in Dubos broth (Becton Dickinson) supplemented with 4% v/v albumin (Becton

Dickinson), 0.2% v/v glycerol, 0.05% Tween and selective agents where mentioned. All cultures of *M. tuberculosis* were grown at 37°C whether in liquid culture or on agar plates. Growth of *M. tuberculosis* strains on solid media used Middlebrook 7H11 agar (Difco) supplemented with 10% oleic acid, dextrose, catalase and albumin mix (Difco) when the molten agar had cooled below 65°C.

2.1.1.4.2 Antibiotics and selective agents

In agar plates and in liquid cultures the following final concentrations for selective agents were used: sucrose, 2%; kanamycin, 25μg/ml; hygromycin, 50μg/ml; X-gal, 100μg/ml. Antibiotics were added to liquid agar once it had cooled below 65°C.

2.1.1.4.3 Rolling bottles

M. tuberculosis strains were prepared for all experiments in this study by inoculation of 100ml aliquots of liquid Dubos media as described above in 2.1.1.4.1 at a 1/500 dilution from exponential-phase cultures in sterilins (Bibby-Sterilin) to a 1 litre polycarbonate bottle (Techmate) and incubation at 37°C at ~2rpm in a Bellco roll-in incubator to allow maximum aeration, unless stated otherwise. All experiments, unless otherwise stated, were performed in biological triplicate.

Cultures of *M. tuberculosis* were then incubated until early exponential-phase, at which point experiments were started. Typically after 1/500 dilution and growth under the conditions described above the cultures would reach an OD_{600} of 0.3-0.4 after five days. The OD_{600} was checked regularly and cultures were diluted with pre-warmed

liquid media to keep the strains growing together. Cultures were not diluted within at least one generation of the start of the experiment.

2.1.1.4.4 Storage of *M. tuberculosis* strains as frozen stocks

Exponential-phase liquid cultures, or cultures on agar plates, of *M. tuberculosis* were mixed with freezing mix either by 50:50 dilution of a liquid culture or by the homogenisation of a fresh bacterial colony with sterile 2.5-3.5mm glass beads (BDH) in 1ml freezing mix using a vortexer (Labtech) before storage at -80°C. The components of freezing mix can be found in Appendix 4.

2.1.1.4.5 Growth of *M. tuberculosis* strains on agar plates

Strains were streaked or spotted on agar plates and left to dry overnight. The plates were then sealed with gas-permeable Nescofilm and incubated upside-down in a Bellco roll-in incubator at 37°C for 2-6 weeks.

2.1.1.4.6 Initial growth of *M. tuberculosis* strains prior to rolling bottle inoculation

To adapt colonies grown on plates to growth in liquid media, and to minimise clumping, strains were grown as 5 ml cultures in sterilins (Bibby-Sterilin) before inoculation into rolling bottles. A colony was picked from an agar plate and inoculated into a sterilin containing 1 ml 2.5-3.5mm sterile glass beads (BDH) and 5ml liquid media. The colony was homogenised by vortexing the culture and was left to incubate at 37°C in a Swallow incubator until the culture reached exponential phase.

2.1.1.4.7 Optical density reading of bacterial cultures

Spectrophotometric measurements of microbial biomass were expressed as optical densities after measurement of the scattering of 600nm wavelength light through a 1cm pathlength transparent plastic cuvette (Elkay) in a Cecil spectrophotometer. The cuvette was stoppered for biosafety reasons. Generation (doubling) times of the bacteria were established using the formula

$$G = \frac{t}{3.3 \log (\text{new } OD_{600}/\text{old } OD_{600})}$$

Any cultures that gave OD_{600} measurements of over 0.8 were serially diluted below this value and then multiplied up to generate the final measurement. This was done because most spectrophotometers do not accurately record light scattering at higher culture densities.

2.1.1.4.8 Preparation and transformation of competent *M. tuberculosis*

M. tuberculosis strains were grown as 100ml cultures in rolling bottles (see 2.1.1.4.1) to an OD_{600} of 1 before the addition of 10ml 2M sterile-filtered glycine (Sigma). After a further 24 hours of incubation the cells were pelletted and washed in glycerol by repeated centrifugation at 10 000rpm in a GSA rotor pre-cooled to 4°C in an high speed centrifuge (Sorvall) at 13000rpm, before resuspension in an equal volume of 10% glycerol (BDH). After three washes the cells were resuspended in 10ml 10% glycerol and stored for no more than a week at 4°C before electroporation.

400μl of competent *M. tuberculosis* was mixed with an appropriate amount of plasmid DNA, usually 1μg, in a 0.2cm pathlength stoppered electroporation cuvette (Bio-

Rad). The cuvette was then placed in a Bio-Rad Gene Pulse Controller and electroporated at 25μF, 2.5kV and 1000Ω. The cells were then transferred using a sterile needle-nosed pastette to 3.6ml prewarmed liquid media in a sterilin, without antibiotic selection, which allows for expression of the marker genes contained on the plasmid. After 24 hours the electroporated mycobacteria are diluted as appropriate and spread onto agar plates containing selective agents.

2.1.2 *E. coli* strains

Commercially available strains of *E. coli* were used throughout this work (see **Table 2.3**). Library efficiency DH5α cells (Invitrogen) were used for routine cloning and subcloning of plasmids and supercompetent *E. coli* (Invitrogen) were used for cloning more difficult sources, such as those generated during site-directed mutagenesis (see **Table 2.3 List of strains and genotypes used in this study** . TOP10 cells (Invitrogen) were used for blunt-end cloning PCR products into the vectors from the TOPO cloning kit (Invitrogen).

BL21-RIL-codonplus cells (Novagen) were used for protein expression, since they are protease negative and contain a plasmid that allow translation of codons rarely found in *E. coli* but common in *M. tuberculosis*. This is necessary since the *E. coli* strain may not otherwise produce the appropriate amounts of the tRNA for these codons and would translate the mycobacterial sequences inefficiently. Strains were manipulated as directed by the manufacturers' supplied protocols.

Table 2.3 List of strains and genotypes used in this study

strain	genotype	reference
DH5-α (DNA manipulation)	F- ϕ80*lacZ*ΔM15 Δ(*lacZYA-argF*)U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *phoA supE44 thi*-1 *gyr*A96 *rel*A1 λ^-	Invitrogen (www.invitrogen.com)
BL21-CodonPlus(DE3)-RP (protein expression)	*E. coli* B F– *ompT hsdS*(r_B– m_B–) *dcm*+ Tet^r *gal* λ (DE3) *endA* Hte [*argU proL* Cam^r]	Stratagene (www.stratagene.com)
TOP10	F- *mcrA* Δ(*mrr-hsd*RMS-*mcrBC*) ϕ80*lacZ*ΔM15 Δ*lacX*74 *recA*1 *araD*139 Δ(*ara-leu*)7697 *galU galK rpsL* (Str^R) *endA*1 *nupG*	Invitrogen (www.invitrogen.com)

2.1.2.1 Media and growth conditions for *E. coli* strains

Luria broth or agar was used to grow all cultures of *E. coli* according to the method in Sambrook *et al.* (2001). 3ml liquid cultures were grown overnight, and then used for small-scale DNA or protein extraction or inoculated into larger vessels (Sambrook *et al.* 2000). Liquid cultures were grown at 37°C with shaking at 300rpm in an incubator (New Brunswick Scientific). Agar plates were left to dry before being incubated upside-down overnight at 37°C in a Bellco roll-in incubator. OD_{600} readings for *E. coli* were taken using 1cm pathlength plastic cuvettes (Elkay) and a Unicam spectrophotometer.

2.1.2.2 Antibiotics and selective agents for *E. coli*

Ampicillin, carbenicillin and X-gal were used at 100μg/ml final concentration, and kanamycin and hygromycin at 50μg/ml and 250μg/ml final concentrations, respectively. IPTG was used at 1mM final concentration in liquid cultures to induce expression of recombinant polyhistidine-tagged RegX3; alternatively 100μl was spread onto the surface of a 90m diameter agar plate and allowed to dry before inoculation with *E. coli* that may be expressing β-galactosidase.

2.2 Stress experiments

2.2.1 *In vitro* experiments

2.2.1.1 Stress experiments after aerobic growth

Strains of *M. tuberculosis* were grown as 120ml cultures in rolling bottles to an OD_{600} of 0.3-0.4, after which the culture was split into six 20ml aliquots in fresh rolling bottles. The stressing agent dissolved in a small volume of solvent was then added to three of these and equal volumes of the solvent alone were added to the other three cultures as controls. All cultures were then incubated (with rolling at 37°C) for the specified amount of time, after which samples were removed from each culture and assessed for OD_{600} and viability. Viability was assessed by plating out aliquots of each culture that had been vortexed with 2.5-3.5mm glass beads (BDH) and serially diluted in a 1:1 mix of DMEM (Gifco) and foetal calf serum (Sigma). The purpose of the foetal calf serum is to buffer the mycobacteria against any stressing agents that may still be active in the culture media. Typically 25μl of the 10^{-2}, 10^{-3}, 10^{-4} and 10^{-5} dilutions were plated in duplicate onto separate halves of agar plates, and dried before incubation. After 10-14 days the plates were counted by eye using a strong light source and a dark background in order to visualise the colonies. Values for each culture were taken as the mean of the duplicate readings at the most appropriate dilution and multiplied by the dilution factor.

2.2.1.2 Microaerobic growth of *M. tuberculosis* strains

Early exponential-phase rolling cultures of *M. tuberculosis* were inoculated at 1/10 dilution into 262.5cm^2 volume non-vented tissue culture flasks containing 90ml liquid media, to make a final volume of 100ml, and the lids tightened firmly by hand. The cultures were then left to stand vertically in an incubator (Swallow) at 37°C until the cultures were at the appropriate OD_{600} as required before being processed as described in

the relevant experiments. Excess cultures were set up and sacrificed daily to measure the approximate increase in OD_{600}.

2.2.1.3 Nutrient starvation of *M. tuberculosis* strains

Early exponential-phase cultures were grown as 100ml cultures in rolling bottles after which they were washed three times with equal volumes of sterile PBS by pelleting and resuspending using a high speed centrifuge (Sorvall) and a GSA rotor at 10 000rpm for 20 min. After washing, the pelletted cells were resuspended in a further equal volume of PBS and 5 ml aliquots made into sterilin tubes (Bibby-Sterilin) that were incubated at 37°C without shaking or other disturbance for 43 days, after which the viable counts and OD_{600} measurements were taken and compared to those taken on day one.

2.2.1.4 Growth of *M. tuberculosis* strains at acid pH

Dubos media was prepared as normal (2.1.1.4.1), however the media was adjusted to pH5.5 with hydrochloric acid and a digital pH meter before being autoclaved and supplemented with the albumin. A sample of the sterile supplemented media was then re-checked to ensure the pH had not changed. Strains were inoculated from early exponential-phase cultures into 100ml aliquots of the acidified liquid media and cultured aerobically (see 2.1.1.4.3). OD_{600} was measured throughout the growth of the strains, and after the final time point was taken the pH of the media was measured using indicator strips (Merck).

2.2.1.5 Agents used for viability studies

Paraquat dichloride (Sigma) is a superoxide radical generator that was used at 50mM concentration to cause a bactericidal effect in *M. tuberculosis* (L. Rand, personal communication). The supplied powder was made up to the correct concentration with distilled water, before being sterilised by passing through a .22μm filter (Millipore). Silver nitrite (Sigma) was supplied as a powder and treated in the same was as the paraquat, and added to a final concentration of 3mM to cultures resuspended in media acidified to pH 5.5 in order to produce bactericidal levels of nitroxide radicals (L. Rand, personal communication). *tert*-butyl hydroperoxide (Sigma) was supplied as a liquid, and diluted to the required concentrations using distilled water before being sterilised by passing through a .22μm filter (Millipore). The prepared tert-butyl hydroperoxide was added to cultures over a range of concentrations and produced bactericidal levels of organic hydroperoxide stress at concentrations over 100μM (see 3.2.2.3).

2.2.1.6 Statistics for stress experiments

Homoscedastic, two-tailed Student's t-tests were used to compare the log losses in viability of a strain after stress compared to the wild-type. Calculations were made in MICROSOFT EXCEL, and graphs plotted in SIGMAPLOT. Error bars for the graphs represent standard error of three experiments as calculated by SIGMAPLOT. See Appendix 5 for raw data and statistical treatment of all stress experiments mentioned in this thesis.

2.2.2 *In vivo* experiments

2.2.2.1 Isolation and infection of macrophages

8-month old female Balb/C mice were killed and the thigh bones removed and flushed with CO_2-independent media (Gibco). After centrifugation of the flushed cells in sterilins at 1200rpm for 5 min in a swing-out rotor in a Heraeus megafuge the cell pellet was resuspended in complete IMDM media (Gibco) supplemented with 0.0004M β-mercaptoethanol (Sigma), 4mM L-glutamine (Sigma), 5% FCS and 10% MCSF (supplied in-house), and aliquoted into Petri dishes (Nunclon). After 2 days incubation at 37°C with 5% CO_2 in an incubator (LEEC) the cells were washed and the media replaced. After a further three days the cells are scraped from the surface of the Petri dish, counted using a haemocytometer and 100 000 cells were aliquoted into each well of a 12-well tissue culture dish (Nunclon).

After a further 24 hours of incubation to allow the macrophages to adhere to the wells, *M. tuberculosis* cells diluted in DMEM (Gifco) from early-exponential-phase cultures were added to the wells containing the macrophages at a multiplicity of infection of 1 mycobacterium per every 2 macrophages. After a 5 hour infection period the infected macrophages were washed to remove the extracellular bacteria, and saponin (Sigma) was added to a 0.1% final concentration to the wells of the first time point, time 0. After an hour of saponin treatment, lysis of the macrophages was confirmed by microscopy and serial dilutions of the supernatant were made in saline. 10μl aliquots of the serially diluted supernatant were spotted onto agar plates and the bacteria were left to grow at 37°C for 10-20 days before colony numbers were counted.

2.2.2.2 Murine infection experiments

M. tuberculosis strains were grown in rolling bottles to early-exponential phase, were diluted in sterile saline to give a suspension of approximately 10^6 cfu/ml and 200μl of the cell suspension was inoculated intravenously into the tail vein of 6-8 week old female BALB/c mice. The infection was monitored by removal of the lungs and spleen of groups of 3-5 mice at intervals, followed by homogenisation of the removed organs using a Mini-Bead Beater (Biospec Products). Serial 10-fold dilutions of the homogenates were made in chilled saline and spotted onto agar plates. After 10-20 days the plates were counted and the cfu/g of each tissue calculated. A paired Student's T-test was used to assess the p values of the means of different samples.

2.3 Cloning

2.3.1 Polymerase chain reaction

A variety of PCR protocols were used throughout the project. Pfu Ultra (Stratagene) is a modified high-fidelity proofreading polymerase from *Pyrococcus furiosus*, and was used to amplify regions of *M. tuberculosis* DNA used for allelic exchange and protein expression purposes where it was important not to generate errors in the amplified DNA. The disadvantages of Pfu Ultra are its relatively high cost and low processivity when compared to other enzymes. Kod Hi-Fi polymerase (Novagen) is a modified high-fidelity proofreading polymerase from *Thermococcus kodakaraensis* that is complexed with two monoclonal antibodies that confer hot-start properties on the enzyme. It produced consistent results, has very high processivity and is relatively cheap; however the accuracy of the enzyme is not as great as that of Pfu Ultra. It was

predominantly used for amplifying short lengths of DNA for cloning, protein expression purposes or sequencing when it was not possible to obtain a product with Pfu Ultra.

When errors in the sequence were less important, i.e. for assessing the presence of inserts in a plasmid by the size of the PCR products, HotStart Taq polymerase (Qiagen) or Ready-2-Go PCR beads (Amersham) were used. HotStart polymerase is a hot-start polymerase that shows increased specificity for the target due to its activation only at the first denaturation step of the PCR cycle, for which it requires 15 min incubation at 95°C. Ready-2-Go PCR beads contain the majority of the reagents required for PCR in a lyophilised bead, to which water, primers and DNA must be added. Ready-2-Go PCR beads showed poor specificity for the substrate but were fast and simple to use. Components were included in the reactions as described in the manufacturers' instructions.

All amplifications took place in an MWG Primus thermocycler or a Perkin Elmer thermocycler, using thin-walled 0.2ml tubes (ABgene) and heated lids. A touchdown cycling protocol was used for the majority of amplifications, with the following parameters: 95°C denaturation step for 2 min, followed by 10 rounds of synthesis with progressively decreasing annealing temperatures (95°C, 45 sec, 72°C-1°C per cycle, 45 sec, 72°C, 2 min), 25 rounds of synthesis at a lower, constant annealing temperature (95°C, 45 sec, 55°C, 45 sec, 72°C, 2 min) and a final extension step at 72°C, after which the samples were cooled to 30°C at the end of the reaction. For further details on PCR, see Sambrook *et al.* (2000)

2.3.2 Reverse transcriptase polymerase chain reaction

RT-PCR was carried out using the one-step RT-PCR system (Invitrogen) according to the manufacturer's directions.

2.3.3 Restriction enzymes

Restriction endonucleases were purchased from Roche and New England Biolabs, and were used in accordance with the manufacturers' directions. Reactions typically took place in a 10μl volume with the appropriate 1x buffer and were incubated for 1 hour at 37°C before the addition of glycerol running buffer (30% glycerol, 0.5% bromophenol blue) and electrophoresis of the samples through an agarose gel (see below at 2.3.4). If greater quantities of cut DNA were required the reaction was scaled up to take this into account.

2.3.4 Agarose gel electrophoresis

Separation and visualisation of DNA was performed by electrophoresis through a TBE agarose gel (Sigma, molecular biology grade) containing 0.00016mg/ml ethidium bromide and visualised and photographed on a UV transilluminator fitted with a digital camera (UVP). Gels were prepared at 1% agarose concentration for analysis for most purposes or 2% for analysis of DNA fragments smaller than 500 bp.

2.3.5 Ligation of DNA fragments

'Sticky-ended' ligations of DNA with compatible overhangs at the 5' and 3' ends were performed by overnight incubation at room temperature of the 'cut' insert and vector at a 7:1 ratio with T4 ligase (Promega) in the supplied buffer. Blunt-ended

ligations were performed using the Zero Blunt TOPO kit from Invitrogen according to the manufacturer's instructions.

2.3.6 Transformation of *E. coli* strains

Transformations were carried out using commercial strains discussed above (Table 2.2) and according to the instructions supplied by the manufacturer. Further information on transformation of competent cells can be seen in Sambrook *et al.* (2000).

2.3.7 Design and supply of oligonucleotides

Primers for PCR (Appendix 3) were designed using the program PRIMER DESIGNER FOR WINDOWS to fit the default parameters specified within that program, were checked for the presence of hairpins and runs of single bases, and varied between around 17 and 22 bp in length. All primers designed for amplification of *M. tuberculosis* target sequences were checked using the ARTEMIS program (Rutherford *et al.* 2000) for uniqueness. Oligonucleotides were ordered from Oswel-Eurogentec at 0.04M scale synthesis and were desalted, or purified by HPLC if they were to be used for radiolabelling in electromobility shift assays.

2.4 Nucleic acid extraction

2.4.1 DNA

2.4.1.1 *E. coli*

DNA was extracted from 3ml cultures *E. coli* using either Qiagen miniprep or maxiprep kits as recommended by the manufacturer. Qiagen plasmid isolation kits rely

on the adsorption of DNA onto silica membranes after alkaline lysis of the *E. coli* cells; for further information see the Qiagen handbooks (supplied) or Sambrook *et al.* (2000).

2.4.1.2 *M. tuberculosis*

For small-scale preparations of *M. tuberculosis* DNA from cultures grown on plates, the Instagene matrix (Biorad) was used according to the manufacturer's directions. The plate-grown mycobacteria were incubated with the Instagene matrix at high temperatures causing lysis of the cells. The matrix serves to remove some of the cell debris that would otherwise interfere with downstream applications.

For large-scale high quality DNA preparations (protocol supplied by R. Curtis), rolling cultures were centrifuged at 10 000 rpm for 20 min at 4°C in a GSA rotor in a Sorvall centrifuge. The pellet was washed with 20ml SET solution (see Appendix 4) and resuspended in 2ml set solution containing 2mg/ml lysozyme (Sigma) and 2mg/ml lipase (Sigma) and was incubated for 30 min at 37°C. 4 volumes of GSE solution were added before a further 90 min incubation, after which the nuclaeic acids were isolated using chloroform-ethanol precipitation (Sambrook *et al.* (2000)). The pellet was then dissolved in 400μl TE buffer with 2μl RNase (Invitrogen), 20μl of proteinase K (Sigma) at 10mg/ml and 20μl of 10% SDS, and incubated overnight at 37°C until dissolved.

2.4.2 RNA

RNA was extracted from cultures of *M. tuberculosis* by use of the RNA Pro kit (Q-Biogene) according to the manufacturer's directions. The extracted RNA was purified by incubation with 2μl 20 000u/μl RNAse-free DNAse I (Invitrogen) and two

rounds of phenol/chloroform extraction (Sambrook *et al.* (2000)). The quality and integrity of the RNA was then analysed by running 1μl on a gel and quantified by spectrophotometry at 260nm and 280nm in a UV-visible 1cm pathlength disposable cuvette (Eppendorf) in a Unicam spectrophotometer, or 1μl was analysed on a Nanodrop spectrophotometer. Representative samples were also analysed by RT-PCR (see 2.3.2) and gel chromatography in an Agilent Bioanalyser (service performed by R. Butler, NIMR).

2.5 Microarrays

2.5.1 Microarray construction

cDNA microarrays (TB version 1) covering 98% of the *M. tuberculosis* predicted genes were kindly provided by Dr. J. Hinds at the BμGS facility, St. George's Hospital Medical School, London. Arrays used in all experiments barring the RegX3 overexpression experiment (section 4.2.3) were printed on poly-L-lysine coated slides prepared in our laboratory according to a protocol from BμGS; arrays for the RegX3 overexpression experiment were printed on GAPS amino-sialine coated slides purchased from Corning. The poly-L-lysine or GAPS coating gives the slide a positively charged surface to which the negatively-charged DNA can bind. All protocols in section 2.5 were supplied by or adapted from those supplied by BμGS, with assistance from J. Hinds gratefully received.

2.5.2 Poly-L-Lysine coating

To prepare slides for microarray printing, standard glass microscope slides were placed vertically in metal racks in an appropriately sized chamber. The slides were washed with an alkaline wash solution containing 70g NaOH dissolved in 280ml dH_2O and 420ml 95% ethanol for two hours with stirring. The slides were then rinsed five times for five minutes per wash with clean dH_2O, after which they were coated by submersion in a solution containing 70ml 0.1% w/v solution poly-L-lysine (Sigma), 70ml 10x PBS and 560ml dH_2O. Slides were then given further washes in water, and allowed to dry before storage in dust-free boxes at room temperature.

2.5.3 Post-processing

The printed poly-L-lysine slides were hydrated by being held face down over boiling water for approximately 2 seconds to create a light vapour coating, after which the slides were snap-dried on a 100°C hot plate for approximately 3 seconds. Care was taken not to over-hydrate the slides, which causes the spots to run. The slides were then exposed to 100mJ cm^{-2} ultraviolet light using a UV crosslinker (Spectronics Corporation), after which the slides were chemically blocked using a succinic anhydride/sodium borate solution (according to the BμGS protocol) to minimise non-specific hybridisation of probe to the poly-L-lysine coating of the slides. The slides were then washed for 2 minutes in boiling water with agitation to denature the spotted DNA and make it accessible for hybridisation (Eisen & Brown 1999) followed by a 1 minute wash in 95% ethanol, after which the slides were dried and stored in the dark at room temperature prior to hybridisation.

2.5.4 Labelling

2.5.4.1 RNA-RNA

RNA-RNA hybridisation experiments are a sensitive means of assessing small differences in the transcriptomes of two strains of *M. tuberculosis*. They consist of the hybridisation of labelled cDNA generated from RNA extracted from the wild-type and labelled cDNA generated from RNA extracted from a mutant strain on the same slide. This results in the analysis of two samples simultaneously. Disadvantages of this method are that batches of the wild-type RNA used to generate labelled cDNA are likely to be inconsistant between, or possibly within, experiments making comparisons between different biological replicates or experiments difficult. Furthermore, some of the software for microarray analysis such as GENESPRING was originally developed with one-sample-per-array data in mind, such as that generated by Affymetrix array experiments. This can cause difficulties with later analysis when one wants to examine the control samples independantly from the test samples. In addition, 'dye-swapped' technical replicates have to be performed to prevent artefacts in the data arising from the inequalities in hybridisation and fluorescence of the Cy3- and Cy5-labelled cDNA. The method for RNA-RNA microarray hybridisation is adapted from a protocol provided by the BμGS group (St. George's Hospital Medical School, London).

cDNA samples of the RNA extracted from the wild-type and a mutant strain are generated by reverse transcription of the randomly-primed sample RNA. This is achieved by heating 2-5 μg RNA with 3μg random hexameric primers (Invitrogen) in an 11μl volume to 95°C for 5 min followed by snap cooling on ice, before addition of the following reagents to the stated final concentrations in a 25μl volume: 1x first strand

buffer (Invitrogen), 10mM DTT, 0.46mM dATP, dTTP and dGTP (Invitrogen), 0.184mM dCTP (Invitrogen), 0.116mM Cy3- or Cy5- labelled dCTP (Amersham) and 500u SuperScript III (Invitrogen). The reaction mix was then incubated in the dark for 5 min at 25°C followed by incubation in the dark at 50°C for 60 min, after which the labelled cDNA samples were mixed and purified through a Minelute column (Qiagen) breaking the wash step into two washes of 500μl and 250μl but otherwise according to the manufacturer's instructions. The purified labelled cDNA was eluted from the Minelute column with 15.9μl water.

2.5.4.2 DNA-RNA

DNA-RNA hybridisation consists of hybridisation of labelled cDNA from a standard sample of genomic *M. tuberculosis* DNA and labelled cDNA from RNA extracted from either the wild-type or a mutant strain on the same slide. Separate slides therefore contain information on either the wild-type or the mutant, and the RNA/DNA ratios of the separate slides must be compared to generate results. The advantage of this approach is the use of a universal standard (genomic DNA kindly supplied by Colorado State University) that allows comparisons to be drawn between experiments. The one-sample-per-slide method also facilitates analysis in GENESPRING, where all slides can be directly normalised to the values for each gene present on the arrays hybridised with wild-type sample. Furthermore, because the RNA-derived cDNA under test is only ever labelled with Cy5, there is no need to perform 'dye swap' technical replicates. The method used in this study for DNA-RNA hybridisation is adapted from a protocol provided by the BμGS group (St. George's Hospital Medical School, London).

3.5 μg RNA extracted from the wild-type or a mutant strain is labelled with Cy5 as described in section 2.5.4.1. 1μg DNA is randomly primed by the addition of 3μg random hexameric primers in a 41.5μl total volume and heating to 95°C for 5 min before snap-cooling on ice. Labelled cDNA is then created by incubation in the dark for 1 hour at 37°C of the primed DNA with 1x DNA polymerase buffer (Promega), 0.23 mM dATP, dTTP and dGTP (Invitrogen), 0.092 mM dCTP (Invitrogen), 0.045mM Cy3 labelled dCTP (Amersham) and 3-9u DNA polymerase Klenow fragment (Promega) in a 50μl total volume. After labelling both DNA-derived cDNA and RNA-derived cDNA are mixed and purified through a Minelute column (Qiagen).

2.5.4.3 DNA-DNA

Two samples of DNA are separately labelled with Cy3 or Cy5 (Amersham) as described in section 2.5.4.2 for one sample. The labelled cDNA from the two samples is then mixed and purified through a Minelute column (Qiagen) as described above in section 2.5.4.1.

2.5.5 Hybridisation

The slide is prepared for hybridisation by incubation in 50ml of prehybridisation buffer containing 3.5x SSC (Sigma), 0.1% SDS (w/v) (BioRad) and 10mg/ml fraction V BSA (Sigma) for 20 min at 65°C, followed by consecutive washes in 400 ml water and 400 ml propan-2-ol (Fisher Scientific) before centrifugation at room temperature in 50ml capacity Falcon tubes (Becton Dickinson) at 1500rpm in a Sigma centrifuge.

The labelled cDNA samples were prepared for hybridisation by the addition of 4.6μl 0.22μm filtered 20x SSC (Sigma) and 3.5μl 0.22μm filtered SDS. The samples were then heated to 95°C for 2 min before brief centrifugation in an MSE microcentrifuge and application onto the prepared array. 22mm x 22mm LifterSlips (Erie Scientific) were placed over the array to assist the even distribution of the sample after application.

After application of the sample, the array is sealed in a slide chamber with two 15μl aliquots of water and incubated in a water bath at 65°C for 18-24 hours. After hybridisation the array is washed in 400ml of Wash A solution containing 1x SSC (Sigma) and 0.05% SDS (BioRad) at 65°C, followed by two washes in 0.06x SSC, the first of which is at 65°C and the second at room temperature. The array is then dried by centrifugation in 50ml capacity Falcon tubes (Becton Dickinson) in a Sorvall centrifuge with a SS-34 rotor at 1500 rpm for 5 min and removed to a dark, dust-free box until scanning.

2.5.6 Microarray scanning and data generation

Microarrays were scanned in an Axon 4000A dual-wavelength scanner at 635nm (Cy5 channel) and 532nm (Cy3 channel) providing the red and green aspects, respectively, to the visualisation of the array using the GENEPIX PRO 5.0 software. The GENEPIX software was used to detect the locations of the spots on each array using the gridmap supplied by J. Hinds (BμGS facility, St. George's Hospital Medical School, London), the location and size of which were then assessed by eye and adjusted where necessary. The intensity of each spot was given by the median brightness of the pixels

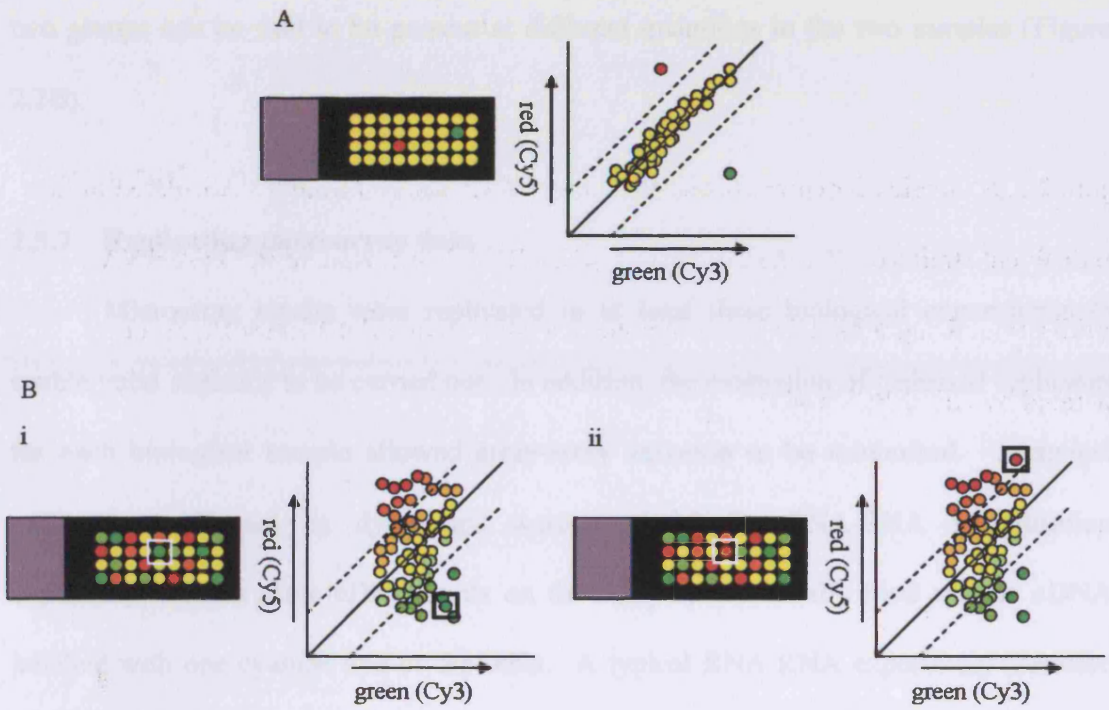
within the spot boundary less the median brightness of the pixels in the area adjacent to the spot for each given wavelength, abbreviated to V_{635} and V_{532} respectively.

In a DNA-DNA hybridisation, all genes from both tested strains would be expected to be represented in the appropriate cDNA populations at the same frequency, with the exception of genes that are present as more than one copy on the chromosome or have been deleted. If the wild-type-derived cDNA is labelled with Cy5 and the mutant-derived cDNA labelled with Cy3, then spots on the microarray specific for the null gene of the mutant will hybridise only the Cy5-labelled cDNA from the wild-type and will appear red on the scanned microarray image. The V_{635}/V_{532} ratio of the gene will be greater than 1. Conversely, if a potential mutant strain is diploid for any genes, then the corresponding cDNA will be present at twice the frequency of that in the wild-type-derived cDNA and the spot will appear green with a V_{635}/V_{532} ratio of less than one.

In an RNA-RNA hybridisation it is likely that transcripts for most genes are present at similar levels in both samples, with only a few genes differentially regulated between the two. A cDNA spot hybridised to by samples that contain labelled Cy3 and Cy5 transcript in equal quantities, as a result of equal numbers of transcripts in the parent RNA samples, will therefore display a yellow colour when both wavelengths are visualised together and produce a V_{635}/V_{532} ratio of one provided both lasers are at comparable strengths. If a gene is over- or under-expressed in either parent RNA sample then the spot will correspondingly appear red or green as the V_{635}/V_{532} ratio departs from one (Figure 2.2A).

In a DNA-RNA hybridisation, Cy3-labelled cDNA from the genomic DNA source should represent all genes equally, whereas the Cy-5 labelled cDNA from an RNA

Figure 2.2: Strategies for comparison of labelled cDNA derived from RNA isolated from two different strains of *M. tuberculosis*. A, RNA-RNA hybridisation; B, DNA-RNA hybridisation. Diagrams refer to the hypothetical slide image (left) and scatter plot generated by GENEPIX (right).


A
red (Cy5)
green (Cy3)
B
i
red (Cy5)
green (Cy3)
ii
red (Cy5)
green (Cy3)

source should contain different levels of transcript for each gene depending on its expression level. This results in uniform intensities for each gene in the 532nm channel, and a variety of intensities for each gene in the 635nm channel. The results are expressed as an RNA/DNA ratio, with arrays hybridised with cDNA derived from RNA extracted from the wild-type compared with arrays hybridised with cDNA derived from RNA extracted from the test strain. Genes that show a change in RNA/DNA ratio between the two groups can be said to be present at different quantities in the two samples (Figure 2.2B).

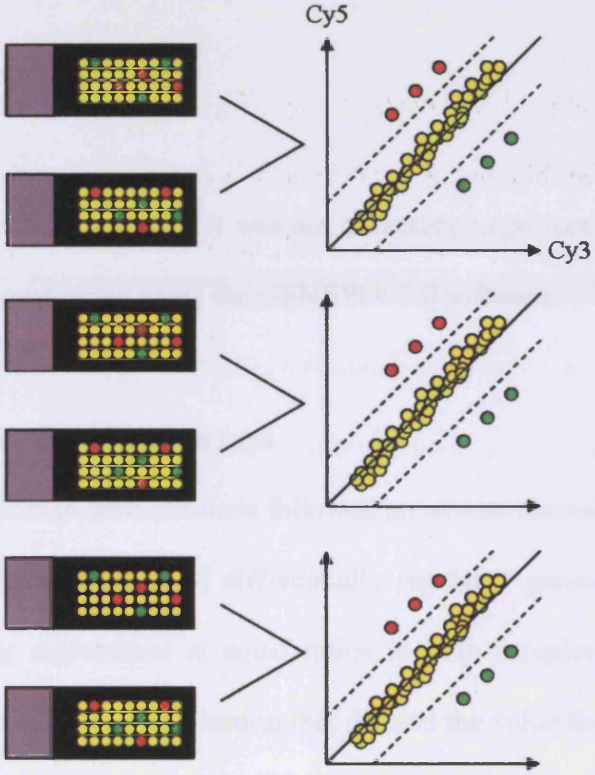
2.5.7 Replicating microarray data

Microarray results were replicated in at least three biological experiments to enable valid statistics to be carried out. In addition, the evaluation of technical replicates for each biological sample allowed array-array variation to be minimised. Technical replicates performed as dye-swaps were essential for RNA-RNA hybridisation experiments, since some cDNA spots on the arrays preferentially bind sample cDNA labelled with one cyanine dye or the other. A typical RNA-RNA experiment therefore consisted of 6 slides, with the wild-type and a mutant compared directly on each slide. For each biological experiment 2 slides were used, one of which the cDNA extracted from the wild-type RNA was labelled with Cy3, and the other of which the cDNA extracted from the wild-type RNA was labelled with Cy5 (Figure 2.3).

A typical DNA-RNA experiment does not need dye-swap technical replicates, and so the technical replication of samples is simply done by performing the same hybridisation twice. In a simple wild-type versus mutant DNA-RNA experiment, 12

Figure 2.3: Analysis strategy for RNA-RNA hybridisations. Sets of slides representing three biological experiments are analysed and compared to isolate genes that are consistently and significantly above or below a Cy5/Cy3 ratio of one.

slide	replicate	cDNA	
		Cy3	Cy5
1	1	wild-type	*senX3*
2	1	*senX3*	wild-type
3	2	wild-type	*senX3*
4	2	*senX3*	wild-type
5	3	wild-type	*senX3*
6	3	*senX3*	wild-type



slides were used: two slides for each RNA sample, three RNA samples per strain (Figure 2.4).

Data files for transcriptional microarrays involving the hybridisation of cDNA samples derived from RNA were imported from GENEPIX to GENESPRING specifying which columns contained the "gene identifier", "flags" (whether a spot was present or absent) and "F635-B635" and "F532-B532" . The latter two columns refer to the V_{635} and V_{532} respectively.

2.5.8 Analysis of microarray results

2.5.8.1 DNA-DNA microarrays

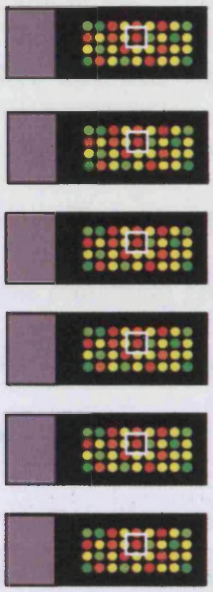
When examining DNA from potential knockouts it was not necessary to perform statistical analysis, and visualisation was performed using the GENEPIX 5.0 software.

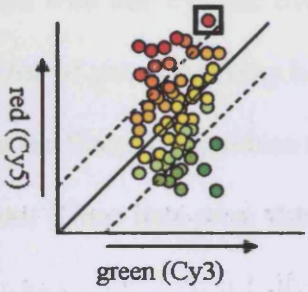
2.5.8.2 GENESPRING analysis of RNA-RNA microarrays

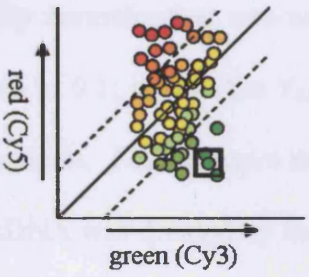
The range of V_{635} and V_{532} signal intensities in both channels followed an almost normal distribution with the exception of a small proportion of differentially regulated genes, resulting in the majority of genes being represented at equal ratios in both samples. Therefore it was appropriate to include a per-chip normalisation that divided the value for each spot by the 50^{th} percentile of the measurements on the array. Control sample values below 0.1 were transformed to 0.1; if both the V_{635} and V_{532} were below 0.1 then the data was excluded from the analysis. For each spot the value for the mutant-derived value was divided by the wild-type-derived value, and dye-swapped technical replicate data

Figure 2.4: Analysis strategy for DNA-RNA hybridisations. Hybridisation of a control labelled cDNA source (generated from genomic DNA) against labelled cDNA derived from RNA generates a 'fingerprint' of relative RNA transcript levels in the mutant strain. Groups of slides representing three biological experiments for both the wild-type and mutant are compared, and analysed to isolate genes that have changed their DNA/RNA ratio significantly between the wild-type and mutant groups.

array	replicate	cDNA source: Cy3	cDNA source: Cy5
1	1	DNA	wild-type RNA
2	1	DNA	wild-type RNA
3	2	DNA	wild-type RNA
4	2	DNA	wild-type RNA
5	3	DNA	wild-type RNA
6	3	DNA	wild-type RNA
7	1	DNA	*senX3* mutant RNA
8	1	DNA	*senX3* mutant RNA
9	2	DNA	*senX3* mutant RNA
10	2	DNA	*senX3* mutant RNA
11	3	DNA	*senX3* mutant RNA
12	3	DNA	*senX3* mutant RNA







was reversed to take account of the use of different wavelength channels for the measurements.

After normalisation, the cross-gene error model is fitted to the experiment parameters, based on replicates, and the average of the base/proportional error estimates is noted. Values with a control intensity less than this were excluded from the data.

Genes that appeared to preferentially bind cDNA labelled with one cyanine dye over the other were also excluded from the data. To generate a list of genes that may be differentially regulated in the mutant strain, the remaining genes are filtered to produce a list of genes that are up- or down-regulated by two-fold or more. Gene lists were then exported to MICROSOFT EXCEL for subsequent editing.

2.5.8.3 GENESPRING analysis of DNA-RNA microarrays

Since the range of intensities for the DNA control data (V_{532}) would not necessarily be expected to follow a normal distribution, and that the majority of genes were not represented in both samples at equal ratios, a per-chip normalisation was not performed. Control sample values below 0.1 were transformed to 0.1; if both the V_{635} and V_{532} were below 0.1 then the data was excluded from the analysis. For each spot the V_{635} representing the mutant strain or wild-type RNA-derived cDNA was divided by that of the wild-type DNA-derived V_{532}. All arrays were normalised to the V_{635}/V_{532} ratios present on the arrays hybridised with the cDNA derived from the RNA extracted from the control strain.

Genes that were present at different ratios between the arrays hybridised with cDNA derived from RNA extracted from the wild-type or the mutant were found by one-

way ANOVA comparison, at a level of $p<0.05$ between the two groups using the tool contained within the GENESPRING 7.1 program. These genes, their p-value and fold change were copied into MICROSOFT EXCEL for subsequent editing.

2.6 Protein production

2.6.1 Considerations for the expression of recombinant polyhistidine-tagged RegX3

2.6.1.1 Tools used for the expression and purification of recombinant RegX3

There are a number of different strategies that may be taken to produce purified soluble protein. The most popular is the addition of a peptide tag to one end of the protein being expressed that allows it to be isolated from a heterogeneous protein source, and which may also aid the solubility characteristics of the protein. For example, a glutathione-S-transferase (GST) tag tends to increase the soluble characteristics of the peptide it is fused to, but its large size may affect the behaviour of the protein to which it is fused.

In the expression of recombinant RegX3 in this study it was decided to add a polyhistidine (6-His) tag to the N-terminus of the RegX3 sequence. The small size of the polyhistidine tag makes it unlikely to alter the function of RegX3 compared to other tag types, and it is usually unnecessary to cleave the tag from the protein before use. The presence of the polyhistidine tag allows purification of the protein on an immobilised metal anion chromatography (IMAC) column. IMACs contain a matrix holding a divalent cation which can interact with the imidazole group of histidine residues and form

a stable structure. The interaction can be reversed at acid pH, or by washing the column with high concentrations (in excess of 150mM) of imidazole.

The polyhistidine tag is also unlikely to alter the solubility of RegX3. It is reasonable to assume that RegX3 is a cytoplasmic protein since it interacts with DNA and does not appear to have any transmembrane sections when analysed by the Pfam (Bateman *et al.* 2004) and JPRED (Cuff *et al.*) servers. These data suggest that RegX3 is likely to be soluble in *M. tuberculosis*. Himpens *et al.* (2000) used a denaturing purification protocol that does not distinguish between soluble and insoluble protein, suggesting that they were unable to isolate soluble protein in their system.

To purify recombinant RegX3 the coding sequence of the gene was cloned into the multiple cloning site 3' of a polyhistidine-encoding sequence in the replicating, non-integrative plasmid vector pET15b (Novagen). The completed construct was transformed into BL21-RIL codonplus *E. coli* (Stratagene) and grown in liquid media with carbenicillin. Transcription of the recombinant *regX3* allele is repressed during the growth of the *E. coli* and derepressed by the addition of 1mM IPTG when the culture reached late-exponential phase. Strong expression of the recombinant RegX3 sequence in this vector derives from the transcription activity of the *lac* promoter.

Initial experiments to study the characteristics of the protein of interest involved the addition of IPTG to 10ml cultures of transformed *E. coli*. This allowed some of the variables influencing solubility of the protein to be assessed without the expense of making large-scale protein preparations. Once the most useful parameters for expression were found, the process was scaled up to 1 litre volume cultures.

After 18-24 hours of expression at 18°C the soluble fraction of the cell was harvested by centrifugation at 13 000 rpm in a microcentrifuge (Heraeus) and resuspension in a nondenaturing buffer such as PBS. The cells were lysed by sonication and the recombinant polyhistidine-tagged RegX3 isolated.

2.6.1.2 Visualisation and identification of recombinant proteins

The separated proteins on the gel were visualised by staining with Coomassie Blue R-250 (Bio-Rad Laboratories), after which the protein of interest was distinguished by its predicted molecular weight and comparison with proteins from vector-only control cells.

Polyhistidine-tagged proteins were visualised by Western blotting (Sambrook *et al.* 2000) with an anti-polyhistidine antibody (Novagen), or a commercial InVision stain (Invitrogen). The InVision stain was a fast and easy method of staining gels for his-tagged proteins, since the protocol time of one hour and requires minimal attention. The InVision stain is also compatible with Coomassie staining. Unfortunately it was found in this study to bind promiscuously to other proteins in *E. coli* lysates.

Western blotting involved the transfer of the proteins in the gel to a positively charged PVDF membrane (Invitrogen) which was then probed with a murine anti-polyhistidine primary antibody (Novagen). Addition of horseradish peroxidase-conjugated goat anti-mouse antibody (Jackson ImmunoResearch) specific for the primary antibody allowed visualisation of the complex by production light from lumophoric substrates (Pierce) that were captured on Amersham Hyperfilm as an autoelectrochemilumograph.

The use of Western blots allows the protein of interest to be identified by epitope and apparent molecular mass. Use of an antibody specific for the peptide of the cloned region of the protein can confirm its identity; however these can be expensive to produce. When studying recombinant tagged proteins it is often feasible to probe with a commercially available monoclonal antibody specific for the tag. These commercial antibody preparations are consistent, specific and inexpensive. However it is important to be aware that it is the tag, and not the protein of interest that is being detected. This may cause difficulties if the cloned sequence contained mutations, or if the wrong sequence was cloned; this can be checked by sequencing.

Mass spectrometry can be used to unambiguously identify the recombinant protein. Electrospray ionisation mass spectrometry provides an extremely accurate estimate of the mass of the protein, which may be compared to the predicted mass and is accurate enough to allow identification of mutations in the sequence. Matrix-assisted laser desorption ionisation time-of-flight (MALDI-TOF) mass spectrometry involves digestion of the protein with trypsin and estimation of the sizes of the resulting fragments. This 'fingerprint' can then be submitted to a database containing predicted sequence digests and the protein identified qualitatively.

2.6.2 Methods for expression, identification and purification of recombinant polyhistidine-tagged RegX3

2.6.2.1 Acrylamide gels used for electrophoresis

Gels used for the electrophoresis and visualisation of protein used in this work were commercially available, pre-cast Bis-Tris gels from Invitrogen of 1mM thickness

with wells that held ~15µl sample. These were used in conjuction with the denaturing SDS buffers contained in the kit and following the manufacturer's instructions. Denaturing gels separate proteins based on molecular mass, and were used to analyse products of protein expression and purification experiments. Nondenaturing TBE gels also purchased from Invitrogen at 6, 8, 10 or 4-20% (gradient) acrylamide concentration (with the same dimensions) were used for the study of DNA-protein interactions, with either 1x or 0.5x TBE as an electrophoresis buffer. The use of 0.5x TBE (Sambrook *et al.* (2001), Appendix 4) as an electrophoresis buffer strengthens potential DNA-protein interactions, and requires the TBE acrylamide gel to be pre-run with this buffer for 30 min at 15mA and 200 V to replace the 1x TBE buffer that the gel is supplied in.

2.6.2.2 Visualisation of proteins in SDS-polyacrylamide gels

Gels were stained overnight in Coomassie R250 stain (Bio-Rad Laboratories) and destained for 1-2 hours in a 5% methanol (Sigma), 5% acetic acid (Sigma) solution (Sambrook *et al.* (2001). When the bands became clear the gels were dried using an alcohol-buffer based gel-drying kit (Invitrogen) and photographed. Alternatively, the InVision stain (Invitrogen) was used to stain gels containing recombinant polyhistidine-tagged proteins according to the manufacturer's protocol and were visualised and photographed on a UV transilluminator (UVP) as for agarose gels (2.3.4).

2.6.2.3 Visualisation of polyhistidine-tagged proteins by Western blot

Western blot analysis of gels was performed by semi-dry transfer of the proteins from an acrylamide gel onto an Invitrolon PVDF membrane (Invitrogen) by sandwiching

the PVDF membrane, prewetted in methanol, water and semi-dry transfer buffer (Appendix 4), against the acrylamide gel. Three sheets of Whatman 3M paper cut to the size of the gel were wetted in semi-dry transfer buffer and placed on each side of the PVDF membrane/acrylamide gel sandwich. The sandwich was then placed in a semi-dry transfer apparatus unit (Biometra) with the PVDF membrane facing the cathode and transferred at 100V and 15mA per gel for one hour. The PVDF membrane was then blocked by overnight incubation in TTBS buffer (Appendix 4) with 10% w/v milk powder (Marvel), before being washed three times in TTBS for five minutes each wash. The primary antibody was diluted at 1/1000 concentration into 5ml antibody buffer (TTBS buffer with 3% w/v milk) and incubated with shaking at room temperature for 1 hour, before three further washes in TTBS. The secondary antibody was diluted at 1/2500 into 5ml antibody buffer and used to blot the washed membrane for an hour with shaking at room temperate, after which the membrane was given four further washes in TTBS, one wash in TBS, and one wash in PBS. The protein-antibody complexes were detected by incubation with the substrates contained within the SuperSignal West Pico kit (Pierce) according to the manufacter's instructions and the autochemiluminescent signal visualised by exposure to Hyperfilm (Amersham).

2.6.2.4 Autoradiography

Gels were placed on Whatman 3MM filter paper, covered with Saranwrap and dried under a vacuum with heating. When dry the filter paper/gel/Saranwrap sandwich was transferred to an autoradiography cassette and exposed against X-Omat-AR film

(Kodak) at -80°C for an appropriate amount of time. Autoradiographs were developed using an automatic X-ray processing machine (Fuji).

2.6.2.5 Small scale assay for the solubility of recombinant polyhistidine-tagged RegX3

Frozen stocks of expression-plasmid containing *E. coli* were inoculated into 10ml LB broth (see Appendix 4 and Sambrook *et al.* 2000) under the appropriate antibiotic selection pressure (ampicillin, see 2.1.2.2) and grown to an OD_{600} of 0.6 – 0.8. IPTG (Sigma) was then added to a final concentration of 1mM and the cultures incubated with shaking at ~280 rpm for a further three hours. The cells were pelletted by centrifugation at 10 000g in a high speed centrifuge (Sorvall) using an SS34 rotor. After decanting the supernatant, the cells were resuspended in 1ml PBS and sonicated on ice using a sonicator fitted with a narrow (2mm) tip (Dawe). Lysis and genomic DNA shearing was confirmed by observing the fluidity of each sample through a 20μl pipette tip. After centrifugation of the lysed bacteria for 5 min at 13 000rpm in a microcentrifuge (Heraeus) the supernatant, representing the soluble fraction of the cells, was removed. 10μl of each soluble fraction was added to 4μl of SDS loading buffer. The insoluble pellet is then resuspended in 500μl SDS loading buffer (Novagen), and both the soluble and insoluble fractions were heated to 95°C for 5 min, followed by 30 sec vortexing and 5 min centrifugation at 13 000 rpm in a microcentrifuge (Heraeus). 5-15μl of each fraction was then loaded into the wells of a 12% Bis-Tris gel (Novagen) before electrophoresis at 15mA/200 V for 1 hour, or until the dye front reached the bottom of the gel.

2.6.2.6 IMAC purification of recombinant polyhistidine-tagged RegX3

Immobilised metal anion chromatography used in this work utilised the His-Bind kit and proprietary buffers (Novagen). 2ml supplied resin was loaded into a 7.5ml plastic column fitted with a glass fibre filter disc to retain the resin, and allowed to dry by gravity to its settled bed volume of 1ml (capacity 5mg/ml). The resin (matrix) was then washed twice with 10 bed volumes (b.v.) of water before charging with 10 b.v. of the supplied 50mM $NiSO_4$ solution and equilibration with 10 b.v. of binding buffer (20mM Tris-Cl, 0.5M NaCl, pH7.9, 5mM imidazole). The protein sample prepared in the supplied binding buffer was then applied to the column and allowed to drain by gravity flow. Aliquots of all flowthroughs were taken for later analysis and stored at 4°C. After washing with 10 b.v. supplied binding buffer and a 10 b.v. wash with the low-concentration imidazole buffer (20mM Tris-Cl, 0.5M NaCl, pH7.9, 60mM imidazole) to elute weakly-bound proteins, the remaining protein bound to the column was eluted in 10 b.v. of elution buffer (20mM Tris-Cl, 0.5M NaCl, pH7.9, 1M imidazole). The eluate was collected as a single fraction and stored at 4°C overnight before analysis.

2.6.2.7 Dialysis of recombinant polyhistidine-tagged RegX3

Dialysis of protein samples allows the equilibration of the concentration of small molecules between one liquid sample and another. This is useful for protein chemistry since it allows the protein buffer to be changed after protein purification, where the elution buffer may contain high concentrations of urea or imidazole that may interfere with downstream applications. The following protocol was used for the dialysis of the

recombinant polyhistidine-tagged RegX3 protein, and its RegX3 negative control sample, after elution from an IMAC column.

Three 5 litre aliquots of PBS with 10% glycerol (see Appendix 4) were prepared and chilled to 4°C in 5 litre beakers and covered with aluminium foil. Two 10ml volume Slide-a-Lyser cassettes (Pierce) with 3.5 kDa pore size cut-off were pre-wetted with PBS-glycerol buffer before the samples were injected into them using large bore needles (Kendall). The cassettes were then suspended in 5 litres of the PBS-glycerol aliquots using the supplied floatation devices, and left for 18-24 hours with stirring before two further changes of buffer.

2.6.2.8 **Storage of recombinant polyhistidine-tagged RegX3**

To store the dialysed protein, 25µl aliquots were made into 0.5ml tubes (Eppendorf) that were then snap-frozen in a dry-ice ethanol bath. The protein was subsequently stored at –80°C.

2.6.2.9 **Phosphorylation of recombinant polyhistidine-tagged RegX3**

Phosphorylation was carried out by a method adapted from Zahrt *et al.* (2003). 30 pmol recombinant polyhistidine-tagged RegX3 was incubated with 1M $MgCl_2$ (Sigma), 10mM DTT (Sigma) and 200mM acetyl phosphate (Sigma) for 30 min at 37°C. Acetyl phosphate is a high-energy phosphate donor that under these conditions is likely to phosphorylate all possible sites on the protein. This is also an equilibrium reaction, so that both phosphorylated and unphosphorylated species of protein will be present after the incubation, albeit with the phosphorylated form representing the greatest portion of

the protein. An alternative, but more difficult method is to use a cognate or promiscuous sensor kinase to phosphorylate the recombinant polyhistidine-tagged RegX3, which would be more likely to recreate conditions *in vivo*. However due to time constraints this approach was not attempted in this study.

2.6.2.10 Ultrafiltration

Protein samples were concentrated at 4°C using an Ambion chamber (Millipore) with a 3.5kDa pore size cutoff membrane according to the manufacturer's instructions. Approximately 8ml of dialysed eluate from the purification column (2.6.2.7) was placed in the chamber and concentrated with pressure through the membrane until the volume was approximately 800µl. The concentration of the protein was then measured using a BSA reagent kit (Pierce) according to the manufacturer's instructions.

2.6.2.11 Preparation of recombinant polyhistidine-tagged RegX3 for MALDI-TOF mass spectrometry

15µl of the recombinant polyhistidine-tagged RegX3 protein sample was electrophoresed on a Tris-borate gel (Novagen) and an approximate 1mm^2 volume of the appropriate band removed using a clean scalpel blade, keeping the amount of excess gel to a minimum. In a clean 0.5ml PCR tube (Starlabs) the gel slice was given a 2-5 min wash with water, two washes with 25mM ammonium bicarbonate (Sigma) and three washes with a 50:50 mix of acetonitrile (Sigma) and 50mM ammonium bicarbonate. If it was still stained, the gel slice was given two more washes in the acetonitrile/ammonium bicarbonate mix. After two further washes with 50mM ammonium bicarbonate the

protein was reduced with 20mM DTT (Sigma) dissolved in 50mM ammonium bicarbonate and incubated at 37°C for one hour. After two further washes with 50mM ammonium bicarbonate the protein was alkylated by incubation for 30 min in the dark with 25mM iodoacetamide (Sigma) dissolved in 50mM ammonium bicarbonate. After two further washes with 50mM ammonium bicarbonate the gel is dehydrated by washing with 200μl acetonitrile and air-dried.

Trypsin digestion of the protein sample was carried out on the dehydrated gel slices, which were incubated on ice for at least 15 min before addition of ice-cold 2μg/ml solution sequencing grade modified trypsin (Promega) in 5mM AB. Trypsin digestion was then allowed to occur overnight at 32°C and 0.5μl of the supernatant was then used for MALDI-TOF mass spectrometry. The mass spectrometry results were kindly provided by S. Howell (NIMR).

2.6.3 Electrophoretic mobility shift assays

2.6.3.1 Radiolabelling of probes

Labelling of the probes for use in the electrophoretic mobility shift assays was accomplished by the addition of a radiolabelled phosphate group from 3μl adenosine 5' γ-^{32}P triphosphate triethylammonium salt at 400 Ci/mmol (Amersham) by incubation with T4 polynucleotide kinase at 37°C for 30 min in the supplied buffer in a final volume of 30μl. Components unincorporated into the probes were removed by centrifugation of the reaction through Sephadex G25 columns (Amersham).

2.6.3.2 Binding buffer

The binding buffer to allow the interaction of recombinant polyhistidine-tagged RegX3 was taken from Himpens *et al.* (2000), and adjusted from 5x concentrate to 10x concentrate, to allow more space in the reaction mix for other components. The composition can be seen below for the final 1x solution. Some preliminary experiments were carried out using the 5x binding buffer described in Himpens *et al.* (2000), however the alteration of the stock buffer to 10x concentrate allowed more flexibility in the inclusion of other reagents in the binding reaction. No differences in results were seen during preliminary experiments using the 5x or the 10x buffer. The final concentrations of components in the reaction are Tris/Cl, 2mM; $MgCl_2$, 0.4mM; KCl, 10mM; DTT, 0.2mM; glycerol, 10% v/v; Nonidet P40, 0.01% v/v.

2.6.3.3 Protein-DNA binding reactions

Volumes were adjusted in the protein-DNA binding reaction to allow for other components such as antibody or unlabelled specific DNA. The molar ratio of protein to probe in the final reactions are specified in each experiment, but were typically 50:1, which was similar to that used in Himpens *et al.* (2000) and contained 33.4 pmol of recombinant polyhistidine-tagged RegX3 and 0.66 pmol radiolabelled probe, and the reaction additionally contained 0.2μg poly dIdC bulk carrier DNA (Sigma) in binding buffer (described above in section 2.6.3.5) in a 10μl final volume. Further discussion on protein-DNA binding reactions can be seen in Asusubel *et al.*

2.6.3.4 Electrophoresis of protein-DNA binding reactions for gelshift assays

The radiolabelled probe was added to the reaction as the final component, before the reaction was gently mixed and incubated at 37°C for 30min. BlueJuice loading buffer (Invitrogen) was added to 3x concentration, the sample gently mixed by tapping and the reactions loaded onto a nondenaturing 0.5x TE polyacrylamide gel (Invitrogen) and electrophoresis carried out at ~100V/gel for 40-60 min. Where necessary the amount of protein added was adjusted to allow for the inclusion of other components such as unlabelled competitor probe (annealed oligonucleotides 160 and 161, see Appendix 3). 1μl was added of molar preparations of the unlabelled probe to give amounts relative to the labelled DNA (see individual experiments).

2.6.3.5 Supershift assays

Supershift assays were carried out according to standard published protocols (Asusubel *et al.*, 2005). Briefly, an electrophoretic mobility shift assay was carried out as above, with the inclusion of 1µl antibody specific for the protein of interest (anti-polyhistidine antibody, Novagen) or the same amount of nonspecific antibody (anti-GST antibody, Novagen) into the binding reaction either before the protein-probe binding reaction or after. When antibody was added after the protein-probe binding reaction, all reactions in the experiment were then incubated for a further 30 min at 37°C.

2.6.3.6 Shift-Blot assays

Shift-blot assays were carried out according to the method in Demczuk *et al.* (1993). For further discussion on the shift-blot protocol see 5.1.3.3 and Figure 5.1.

2.7 Allelic exchange

2.7.1 Construction of a vector for the deletion of the Rv0465c gene of *M. tuberculosis*

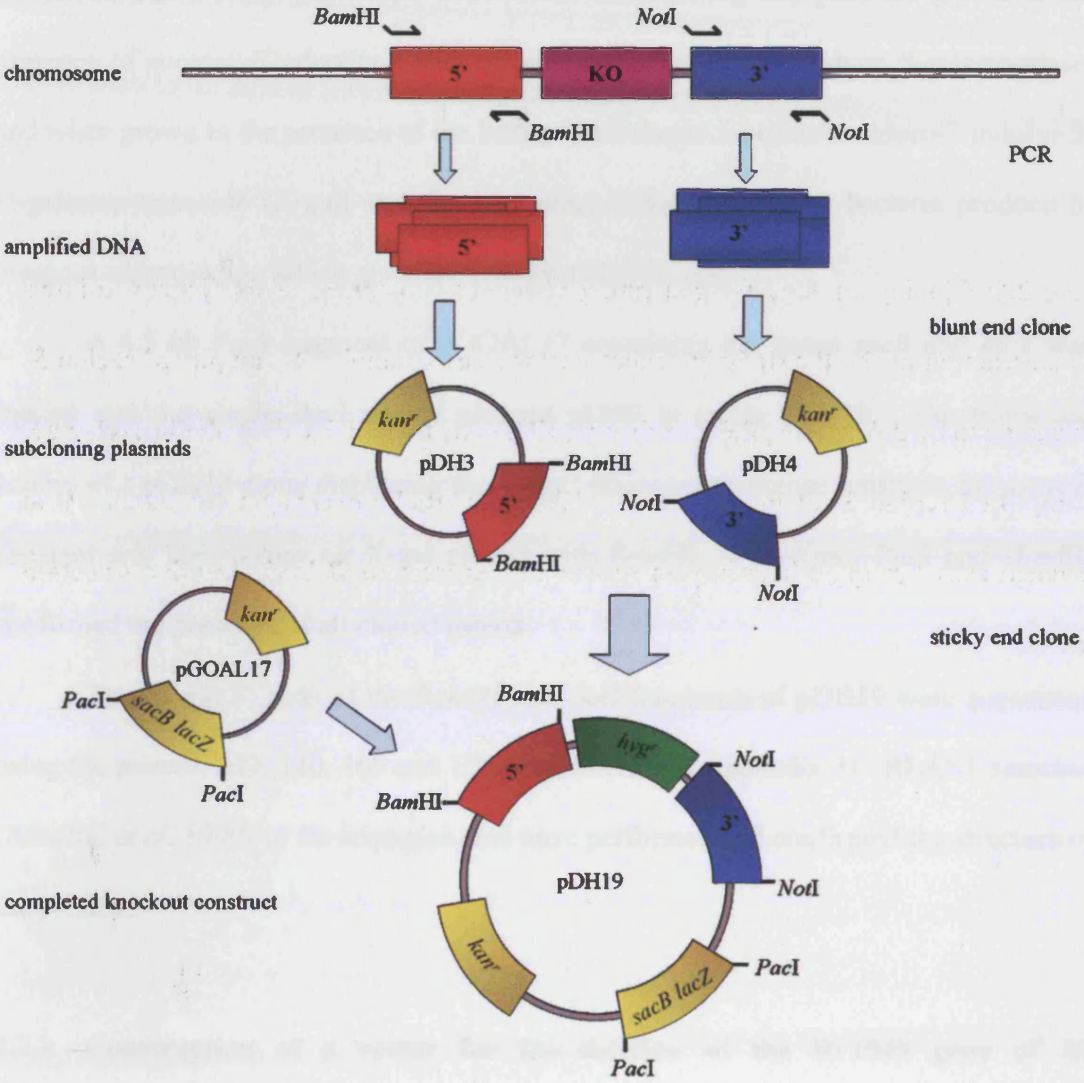
PCR was used to amplify regions of DNA flanking the gene Rv0465c (Figure 2.5). Primers 133 and 134 (Appendix 3) were used to amplify a region 1650 bp 5' of Rv0465c to produce a 1674 bp fragment which was then cloned into pCR-4 using a blunt-ended ligation kit (Invitrogen), creating pDH3. A region 1601 bp 3' of Rv0465c was amplified with primers 135 and 135 (Appendix 3) to produce a 1645 bp fragment which was cloned into pCR-4, again with blunt-ended ligation, to produce pDH 4.

The plasmid p2NIL-H contains a hygromycin resistance cassette in the *Kpn*I site, flanked by a *Bam*HI site 5' of it and *Not*I site 3' of it. The amplified region 5' of Rv0465c had to be cloned into p2NIL-H first since the amplified region 3' of Rv0465c contains a *Bam*HI site (Figure 6.1).

p2NIL-H and pDH3 were digested with *Bam*HI, and the 1656 bp fragment from pDH3 ligated into the linearised p2NIL-H to create pDH6. Seven clones of pDH6 were digested using the restriction enzyme *Sal*I to determine the orientation of the insert, and a clone with the insert in the correct orientation was chosen for further manipulation. A sample of the correct clone of pDH6 was also digested with *Kpn*I, *Pac*I and *Not*I to confirm the integrity of these sites.

The 1609 bp *Not*I fragment from pDH4 ligated into the *Not*I site of pDH6 to create pDH9. pDH9 was digested with *Pst*I and *Not*I to determine the orientation of the *Not*I fragment. The *Bam*HI, *Kpn*I and *Pac*I sites were also confirmed. To generate the

Figure 2.5: Cloning strategy for generating knockout constructs. This diagram follows the construction of the Rv0465c knockout vector as an example. 5', amplified region 5' of the gene to be deleted by allelic exchange; KO, gene to be deleted by allelic exchange; 3', amplified region 3' of the gene to be deleted by allelic exchange; *kan*^{r}, kanamycin resistance cassette; *hyg*^{r}, hygromycin resistance cassette; *sacB/lacZ*, negative selection markers.


BamHI
NotI
chromosome
5'
KO
3'
BamHI
NotI
PCR
amplified DNA
5'
3'
blunt end clone
subcloning plasmids
kanr
pDH3
BamHI
5'
BamHI
kanr
pDH4
NotI
3'
NotI
sticky end clone
kanr
pGOAL17
PacI
sacB lacZ
PacI
BamHI
hygr
NotI
5'
BamHI
3'
pDH19
NotI
kanr
PacI
sacB lacZ
PacI
completed knockout construct

final knockout construct, a *sacB/lacZ* marker cassette was cloned into the *Pac*I site of pDH9. *Pac*I has the recognition sequence TTAATTAA and this site is not present anywhere in the GC-rich genome of *M. tuberculosis*. The *sacB* gene from *Bacillus subtilis* creates a lethal phenotype when bacteria containing this gene are grown in the presence of sucrose (Dedonder 1966). Bacteria that are $lacZ^+$ produce β-galactosidase, and when grown in the presence of the lactose homologue 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) and the gratuitous inducer IPTG the bacteria produce 5-bromo-4-chloroindigo which gives the colony a blue colour.

A 4.5 kb *Pac*I fragment of pGOAL17 containing the genes *sacB* and *lacZ* was ligated into the single *Pac*I site of plasmid pDH9 to create pDH19. The restriction pattern of a pDH19 clone displaying the correct phenotype (sucrose sensitive, kanamycin resistant and blue colour on X-gal plates) with *Bam*HI, *Not*I, *Kpn*I, *Pac*I and *Hind*III confirmed the presence of all cloned inserts.

The 5' and 3' ends of the *Bam*HI and *Not*I fragments of pDH19 were sequenced using the primers 139, 140, 169 and 170 (Figure 6.1 and Appendix 3). BLAST searches (Altschul *et al.* 1990) of the sequence data were performed and confirmed the structure of the plasmid.

2.7.2 Construction of a vector for the deletion of the Rv1049 gene of *M. tuberculosis*

The knockout vector for Rv1049 was produced by the same method as in 2.7.1 and Figure 2.4. Regions of DNA flanking the gene Rv1049 were amplified by PCR. Primers 137 and 138 (Appendix 3) were used to amplify a region 1500 bp 5' of Rv1049

to produce a 1524 bp fragment which was blunt-end cloned into pCR-4 (Invitrogen) to create pDH5. This 3' region is truncated to avoid the IS1081 insertion element upstream of it (Figure 2.6).

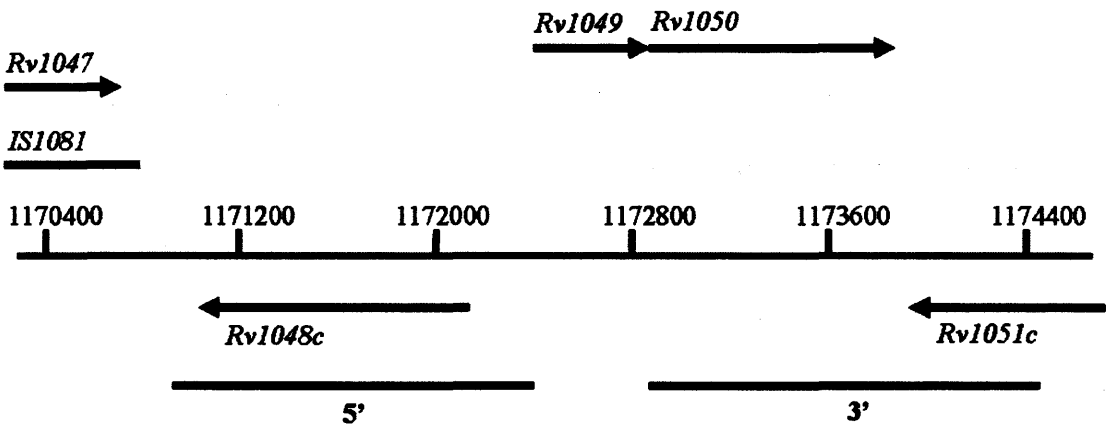
A region 1619 bp 3' of Rv1049 was amplified with primers 119 and 122 (Appendix 3) to produce a 1635 bp fragment that was cloned into pCR-4, again using a blunt-end cloning kit (Invitrogen), to produce pDH2. As with the Rv0465c knockout construct, the amplified region 5' of Rv1049 had to be cloned into the hygromycin-marked manipulation vector p2NIL-H first since the amplified region 3' of Rv1049 contains a *Bam*HI site close to the 3' end (Figure 6.2). Cloning the fragments in the opposite order would have resulted in a large deletion in the final vector.

The 1512 bp *Bam*HI fragment from pDH5 was ligated into the *Bam*HI site of p2NIL-H to create pDH7. The orientation of the insert was determined by a *Hind*III digestion.

The 1623 bp *Not*I fragment from pDH2 was ligated into the *Not*I site of pDH7 to create pDH10. Three clones of pDH10 were analysed by PCR with the primer pairs 122 and d153, and 122 and 137 (Appendix 3), which should result in products of ~3kb and ~4.5kb respectively if the orientation of the inserts was correct, and no products if the insert orientation was incorrect. Three clones appeared to have both inserts orientated in the correct direction.

A 4.5kb *Pac*I fragment from pGOAL17 was ligated into the *Pac*I site of pDH10 to create pDH14. The restriction pattern of a pDH14 clone displaying the correct phenotype (sucrose sensitive, kanamycin resistant and blue colour when grown on X-gal plates) with the enzymes *Bam*HI, *Not*I, *Kpn*I, and *Pac*I confirmed the presence of all

Figure 2.6: The Rv1049 gene region of *M. tuberculosis*. The amplified regions 5' and 3' are marked, as is the location of the insertion sequence IS1081. Rv1047 is a probable transposase, Rv1048 is a hypothetical protein, Rv1050 is a probable oxidoreductase and Rv1051c is a conserved hypothetical protein. All annotations were taken from the Tuberculist web server (http://genolist.pasteur.fr/TubercuList/)


Rv1049
Rv1050
Rv1047
IS1081
1170400
1171200
1172000
1172800
1173600
1174400
Rv1048c
Rv1051c
5'
3'

inserts. The 5' and 3' ends of the *Bam*HI and *Not*I fragments of pDH14 were sequenced using the primers 147 and 148 (Appendix 3) from the 3' end of the 5' fragment and the 5' end of the 3' fragment, respectively, which read into the hygromycin resistance gene (Figure 6.1), BLAST searches (Altschul *et al.* 1990) of which were used to confirm the identity and orientation of the inserts and the presence of the hygromycin resistance gene.

2.7.3 Electroporation and selection of a Rv0465c and a Rv1049 null mutant

Wild-type *M. tuberculosis* H37Rv (a gift from Dr. S. Cole, Institut Pasteur) was grown and made electrocompetent. 1μg of pDH19 in a total volume of 4.8μl, or the same amount of pDH14 in a total volume of 4.052μl, was used to transform 400μl of competent *M. tuberculosis* for the Rv0465c knockout. pDH19 and pDH14 were not pre-treated with UV exposure as recommended in some published protocols (Hinds *et al.* 1999) since it was considered that this may increase the frequency of base-pair mutations and illegitimate recombination.

The competent *M. tuberculosis* electroporated with either pDH19 or pDH14 were then diluted into 3.6ml pre-warmed liquid media and incubated for 24 hours at 37°C in a Swallow incubator to allow expression of the plasmid-encoded genes before aliquots were plated onto agar containing hygromycin and X-gal (Figure 1.1). After 46 days there were 13 blue and 8 white colonies on the 200μl and 400μl platings of the cells transformed with pDH19, and 13 blue and 15 white colonies on the 200μl and 400μl platings of the cells transformed with pDH14. Control plates of cultures transformed with water had no colonies, suggesting that the number of spontaneously hygromycin-

resistant mutants were fewer than hygromycin-resistant mutants arising from some interaction with pDH19 or pDH14.

The blue colonies were streaked onto further agar plates containing hygromycin to allow for a second recombination event to occur. The white colonies were streaked onto agar plates containing kanamycin and on separate plates containing hygromycin, sucrose and X-gal, since the white colonies may represent (a) mutants that have arisen spontaneously from two recombination events and contain the desired mutant genotype, or (b) spontaneous kanamycin resistant mutants of either wild-type or single-recombination event parentage.

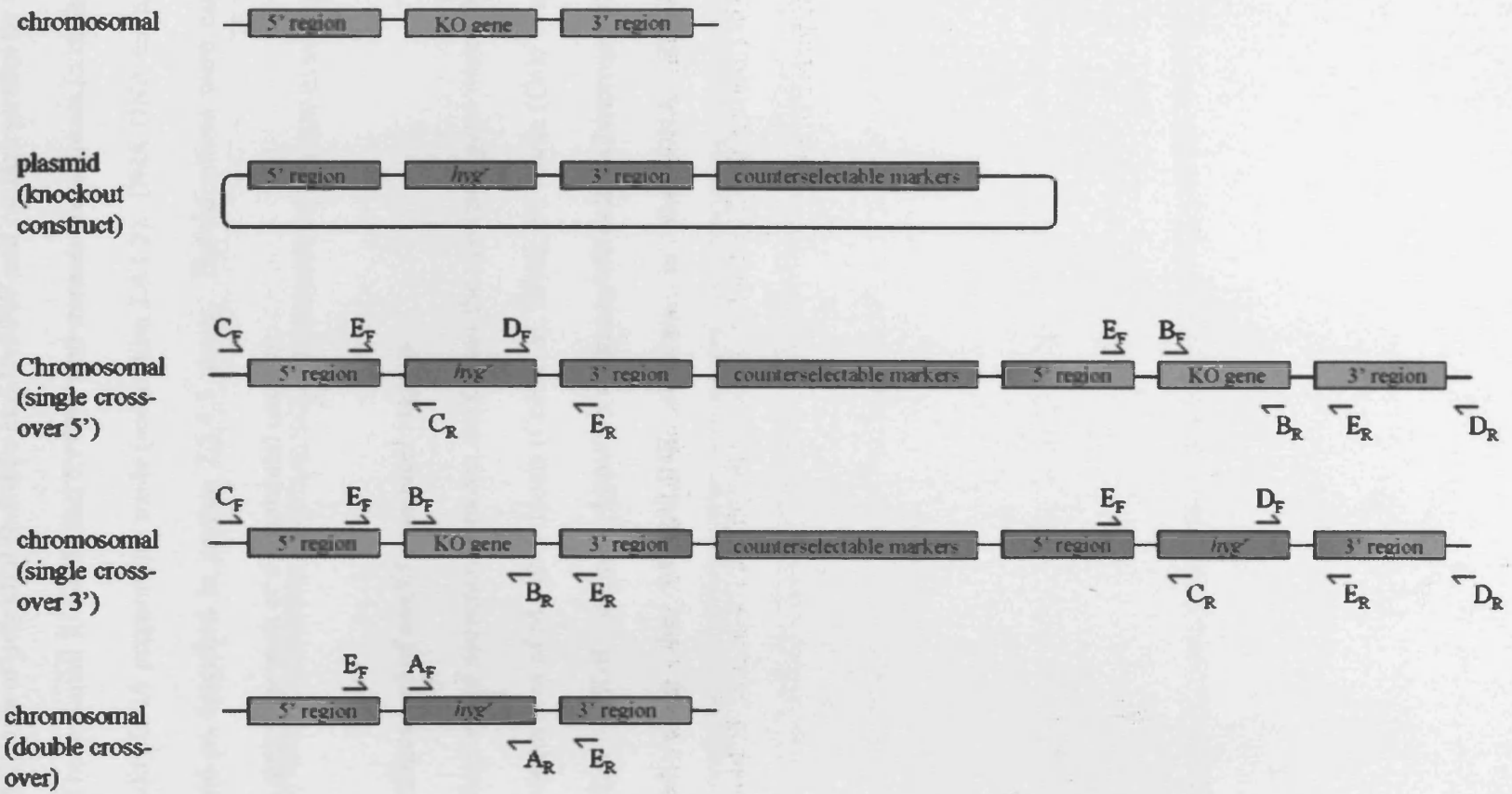
Colonies were counted after 62 days, and those displaying the correct phenotype were frozen. The clones giving rise to blue colonies after the first-round selection were also frozen as stocks, and the plates stored at 4°C. The potential mutant clones were frozen and analysed further.

2.7.4 Analysis of potential Rv0465c and Rv1049 null mutant strains by PCR

Primers were designed to distinguish between the different possible genotypes of the products of either single or double recombination events after electroporation of *M. tuberculosis* with pDH19 or pDH14 (Figure 2.7). The size of the primer product pairs was predicted for the different scenarios (Table 6.1).

The six potential mutant clones were grown and DNA extracted using the Instagene prep kit (Bio-Rad). PCR was used to assess the genotypes of the potential mutant clones, and one clone of *M. tuberculosis* electroporated with pDH19 and one clone of *M. tuberculosis* electroporated with pDH14 selected for further analysis.

Figure 2.7: Locations of primer pairs used to distinguish between genotypes of strains potentially arising during the knockout process. The plasmid sequence is blue; the chromosomal sequence is red.


chromosomal
5' region
KO gene
3' region
plasmid (knockout construct)
5' region
lyg'
3' region
counterselectable markers
Chromosomal (single cross-over 5')
C_F
E_F
D_F
E_F
B_F
5' region
lyg'
3' region
counterselectable markers
5' region
KO gene
3' region
C_R
E_R
B_R
E_R
D_R
chromosomal (single cross-over 3')
C_F
E_F
B_F
E_F
D_F
5' region
KO gene
3' region
counterselectable markers
5' region
lyg'
3' region
B_R
E_R
C_R
E_R
D_R
chromosomal (double cross-over)
E_F
A_F
5' region
lyg'
3' region
A_R
E_R

2.7.5 Analysis of potential Rv0465c and Rv1049 null mutant strains by microarray

The potential Rv0465c and Rv1049 null mutants were grown in roller bottles and high-quality DNA preparations made (see section 2.4.1.2). DNA-DNA microarrays were performed as described in section 2.5.4.3 above. Hybridisations were carried out on separate arrays for each of the potential mutants.

2.7.6 Alignment of *senX3* promoter regions

Regions of sequence between senX3 and the upstream gene were taken from the genome sequences of M. tuberculosis (Cole et al. 1998), M. leprae (Cole et al. 2001) and M. avium (TIGR online, http://www.tigr.org/tdb/mdb/mdbinprogress.html), and compared using the MEGALINE software in the DNA STAR package.

3 Assessment of the role of SenX3 and RegX3 using *in vitro* and *in vivo* assays of infection and cell viability

3.1 INTRODUCTION

When *M. tuberculosis* enters the body it is exposed to reactive oxygen and nitrogen intermediates within the macrophage. The reactive oxygen and nitrogen intermediates can be generated either by the macrophage using phagocyte oxidase and inducible nitric oxide synthase, or by the bacterium as a product of aerobic respiration. These ROI and RNI cause damage to the bacterium, which must detoxify or repair the damage caused if they are to survive.

Several lines of reasoning suggest that the SenX3/RegX3 two component system of *M. tuberculosis* is involved in the response of the bacterium to oxidative stress. Firstly, both proteins have high homology to proteins from other organisms that are involved in oxidative stress resistance. It is possible that the SenX3/RegX3 system is either acting in direct response to the stress stimuli, or that it is acting like the ArcA/B system of *E. coli* where it is involved in shutting down the ROI-generating aerobic respiration system when the bacterium encounters anaerobic conditions (Nystrom 1998).

Secondly, SenX3 contains a PAS domain that may bind a haem prosthetic group (Rickman *et al.* 2004). PAS domains are signal input domains that act as internal sensors of, e.g., redox, oxygen and light, depending on the ligand they contain. Haem-containing PAS domains are often involved in oxygen sensing (reviewed in (Taylor & Zhulin 1999)). Thirdly, microarray data tentatively supports a role for the SenX3/RegX3 system in sensing oxygen (see 4.3)

Further to the above arguments, it is also reasonable to hypothesise that the SenX3/RegX3 two component system plays a role in the infection process. Two component systems are often key regulators of virulence in other organisms, such as

Salmonella typhimurium and *Haemophilus influenzae,* and the conservation of SenX3 and RegX3, one of the eleven two component systems of *M. tuberculosis* (Cole *et al.* 1998) and one of only four in *M. leprae* (Cole *et al.* 2001), which suggests an important role in the lifestyle of the mycobacteria. Recent work by Haydel & Clark-Curtiss (2004) showed that the *regX3* transcript was upregulated at 48 hours after infection of human peripheral blood monocyte-derived macrophages.

The following experiments were designed to determine if there were any differences in the ability of SenX3 and RegX3 null mutants of *M. tuberculosis* to resist damage by ROI and other macrophage-like stresses compared to the wild-type. By adding specific agents to *in vitro* cultures it is possible to examine particular routes of damage – for example, nitrosative stress or superoxide stress. The strains were also grown in mouse and cell culture models of *M. tuberculosis* infection, which allows assessment of the contribution that SenX3 and RegX3 play in the virulence of this organism.

Previous work on these strains has shown no difference in the susceptibility of the SenX3 and RegX3 null mutants to hydrogen peroxide or cumene hydroperoxide stress (Rickman 2002).

Raw data for the stress experiments contained within this chapter, including the statistical treatment of them, can be seen in Appendix 5.

3.2 RESULTS

3.2.1 *In vitro* assessment of the *senX3* and *regX3* null mutants

3.2.1.1 Aerobic growth curve

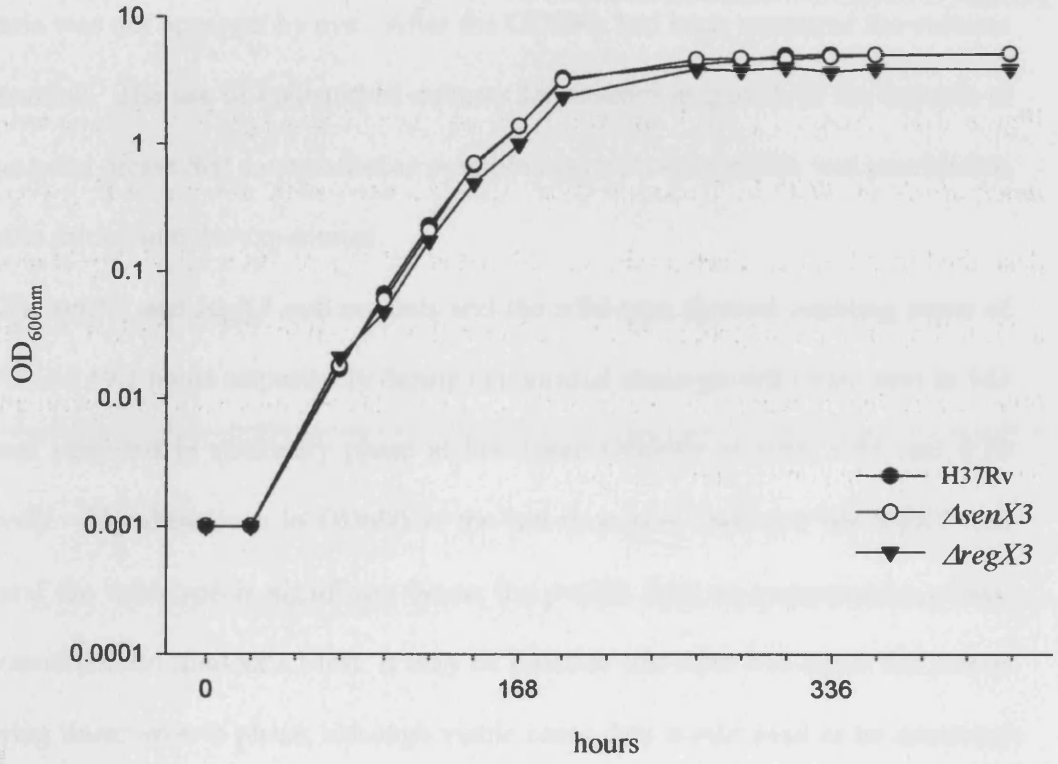
In order to assess the growth phenotype of *senX3* and *regX3* null mutants of *M. tuberculosis in vitro*, the growth of the mutants and wild-type were followed in a standard aerobic *in vitro* model, growing aerobically in Dubos medium with albumin and glycerol at 37°C (see 2.1.1.4). Strains were grown to early exponential phase in Dubos media (2.1.1.4.1) and passaged at 1/500 dilution into a further set of identical rolling bottles. The procedure was as discussed in 2.1.1.4.3. The OD_{600} was measured daily (Figure 3.1). The data in Figure 3.1 is representative of a number of experiments carried out with identical results. Each curve therefore represents a single culture per strain, with each culture measured at different times.

The wild-type and the *senX3* null mutant grew at the same rate with doubling times of 13.63 hours during exponential phase (between 96 and 144 hours post-dilution). The *regX3* null mutant grew at a rate approximately 83% that of the wild-type and the *senX3* null mutant, with a doubling time of 16.45 hours. The *regX3* null mutant also entered stationary phase with an OD_{600} of approximately 80% that of the wild-type and the *senX3* null mutant.

3.2.1.2 Microaerobic growth curve

In order to assess the growth phenotype of the *senX3* and *regX3* under oxygen-limited conditions, the mutants and wild-type were grown aerobically in Dubos medium with albumin and glycerol at 37°C before inoculation into Dubos media (see 2.1.1.4.1) in oxygen-limited, standing bijou tubes and grown at 37°C. The *senX3* and *regX3* mutants and the wild-type were grown as before (2.1.1.4.3) to early exponential phase before 1/250 dilution into 5 ml pre-warmed liquid media in 7 ml plastic bijou tubes with a

Figure 3.1: Aerobic growth of wild-type, and *senX3* and *regX3* null mutants of *M. tuberculosis* at 37°C in Dubos medium supplemented with albumin and glycerol as described in 2.1.1.4.1. Each strain was diluted at 1/500 into 100ml fresh liquid media in a 1L rolling bottle (Nalgene) and allowed to roll at ~2rpm at 37°C. Optical densities at 600nm wavelength were taken daily of the cultures. Data shows a single experiment that has been repeated on several occasions.


10
1
0.1
0.01
0.001
0.0001
OD_{600nm}
0
168
336
hours
H37Rv
ΔsenX3
ΔregX3

air:media ratio of 0.4. The caps of the bijou tubes were tightly sealed by hand to minimise gaseous exchange.

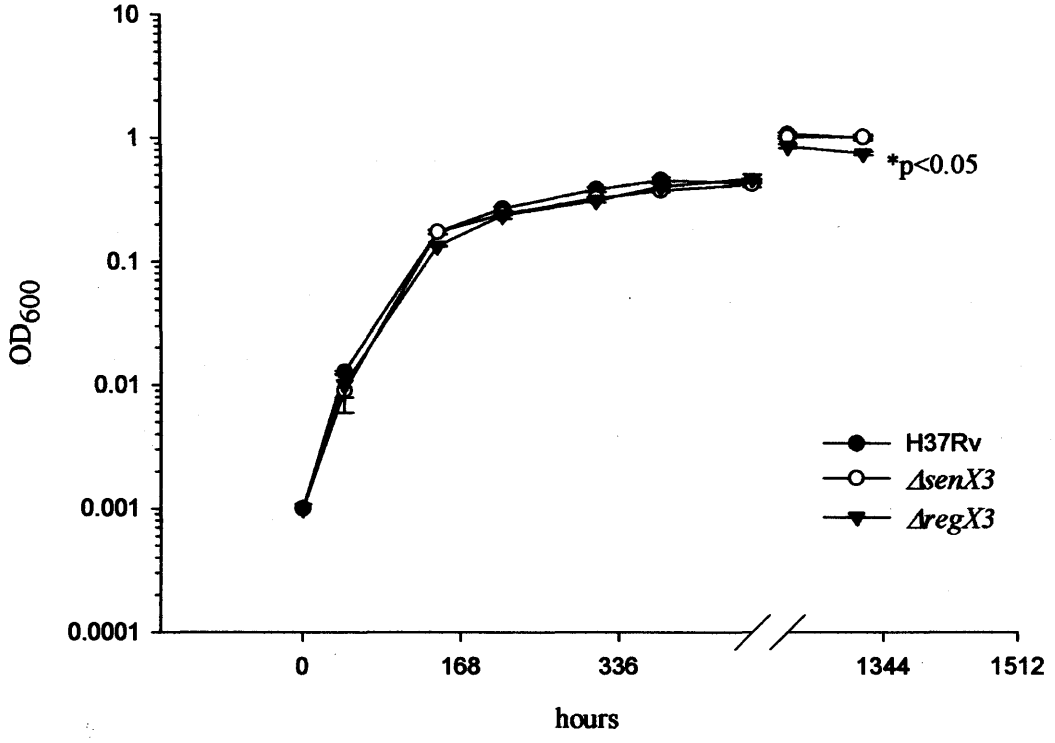
Sufficient bijous were used to allow three to be used per strain per time point for the OD600 readings (Figure 3.2). To take an OD600 reading three bijou cultures per strain were removed from the incubator and vortexed to resuspend the sediment. 'Clumping' of the bacteria was not apparent by eye. After the OD600s had been measured the cultures were discarded. The use of undisturbed cultures for measuring growth of the bacteria at each time point meant that a consistent as possible oxygen concentration was provided to the bacteria throughout the experiment.

The senX3 and regX3 null mutants and the wild-type showed doubling times of 18.5, 19.9 and 19.1 hours respectively during exponential phase growth (from zero to 143 hours) and persisted in stationary phase at low mean OD600s of 0.68, 0.65 and 0.70 respectively. The difference in OD600 at the last time point between the regX3 null mutant and the wild-type is significant below the p=0.05 limit as compared by a two-tailed homoscedastic Student's t-test. It may be possible that after 336 hours the strains are entering linear growth phase, although viable count data would need to be measured to examine this properly.

3.2.1.3 Nutrient starvation

To assess whether the *senX3* and *regX3* null mutants were less viable than the wild-type under nutrient starvation condition, exponential cultures of the mutants and wild-type grown in Dubos medium with albumin and glycerol at 37°C (2.1.1.4.1) were washed several times and resuspended in equal volumes of phosphate-buffered saline

Figure 3.2: Microaerobic growth of wild-type and *senX3* and *regX3* null mutants of *M. tuberculosis* in Dubos medium supplemented with albumin and glycerol as described in 2.1.1.4.1. All strains were grown to mid exponential phase as per 2.1.1.4.3 in Dubos media with albumin and glycerol at 37°C before dilution at 1/500 into Dubos media with albumin and glycerol in multiple standing Bijou cultures and further cultured at 37°C. The data points represent the mean of three used cultures per strain per time point. The *regX3* null mutant was at an OD_{600} significantly lower than the wild type at a level of $p<0.05$ as compared by a two-tailed homoscedastic Student's t-test.


10
1
0.1
0.01
0.001
0.0001
OD600
0
168
336
1344
1512
hours
*p<0.05
H37Rv
ΔsenX3
ΔregX3

before four aliquots of 10 ml per strain were made into 25ml universal tubes. An initial culture was used for each strain on day one to provide a baseline viable count. The nutrient-depleted cultures were then left to stand without disturbance for 43 days at 37°C, after which viable counts were made again (Figure 3.3). Statistics were not performed on these data since the base line measurement at day one was performed once for each strain and was therefore statistically unviable. No difference in viability was apparent between the survival of the wild-type or the *senX3* or *regX3* null mutants at days one and 43.

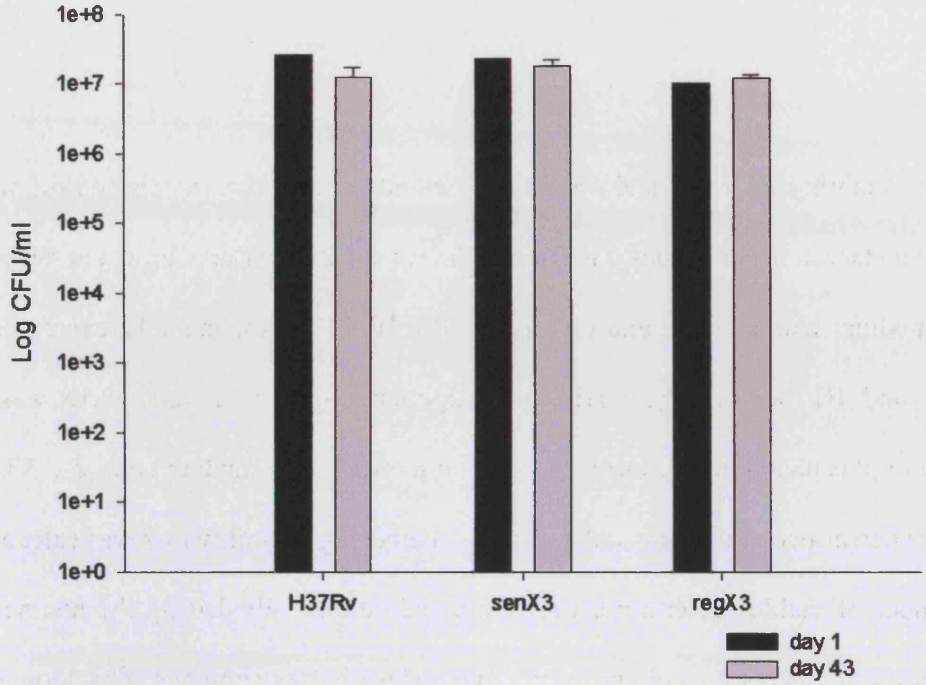
3.2.1.4 pH growth curve

In order to assess the ability of the *senX3* and *regX3* null mutants to grow in a low-pH environment, early-exponential phase cultures grown aerobically in Dubos medium with albumin and glycerol at 37°C were inoculated at 1/500 dilution into triplicate rolling bottles containing media that had been acidified to pH 5.5 with HCl. The growth of the strains was followed daily by measuring the optical density (Figure 3.4). Each data point represents three cultures per time point. Sequential readings were taken from each culture.

There was no difference in growth rates between the *senX3* and *regX3* null mutants and the wild-type. After the last time point the pH of the media was assessed with pH strips to be between pH 5 and pH 6. Growth rates were not calculated due to the non-linear growth of the strains.

Figure 3.3: Viability of *senX3* and *regX3* null mutants of *M. tuberculosis* and the wild-after nutrient starvation for 43 days, following the washing of cultures grown aerobically in Dubos medium with albumin and glycerol at 37°C in PBS (Appendix 4), expressed as A) cfu/ml and B) % viability. Triplicate exponential-phase cultures were washed thoroughly in phosphate-buffered saline before aliquotting into standing cultures. Viable counts were performed at day one and at day 43. Percentage viability loss was calculated as the number of viable bacteria per ml of stressed cultures divided by the number of bacteria per ml of unstressed cultures, times 100. This gives the number of bacteria after stress as a percentage of the unstressed cultures. Statistics were not performed on these data since the base line measurement at day one was performed once for each strain and was therefore statistically unviable.

A


1e+8
1e+7
1e+6
1e+5
1e+4
1e+3
1e+2
1e+1
1e+0
Log CFU/ml
H37Rv
senX3
regX3
day 1
day 43

B

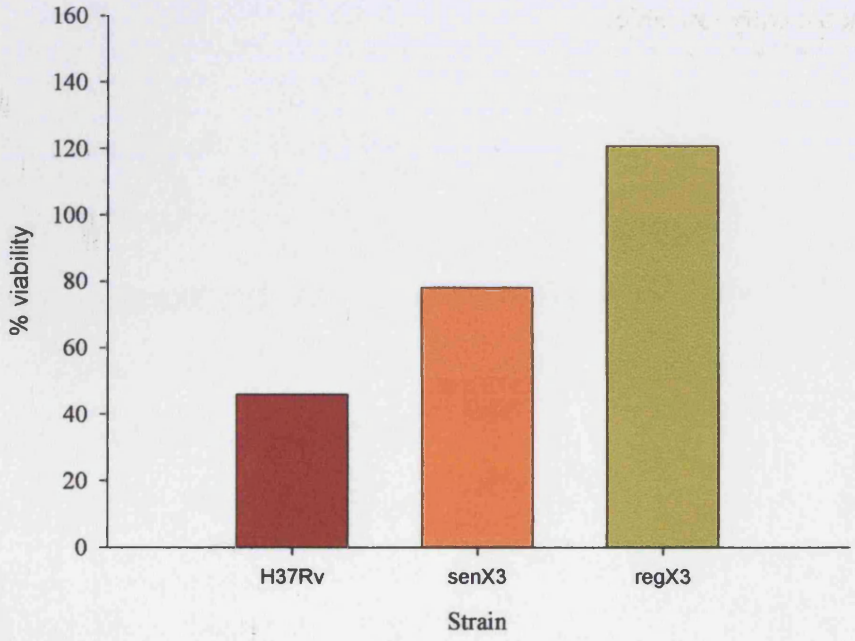
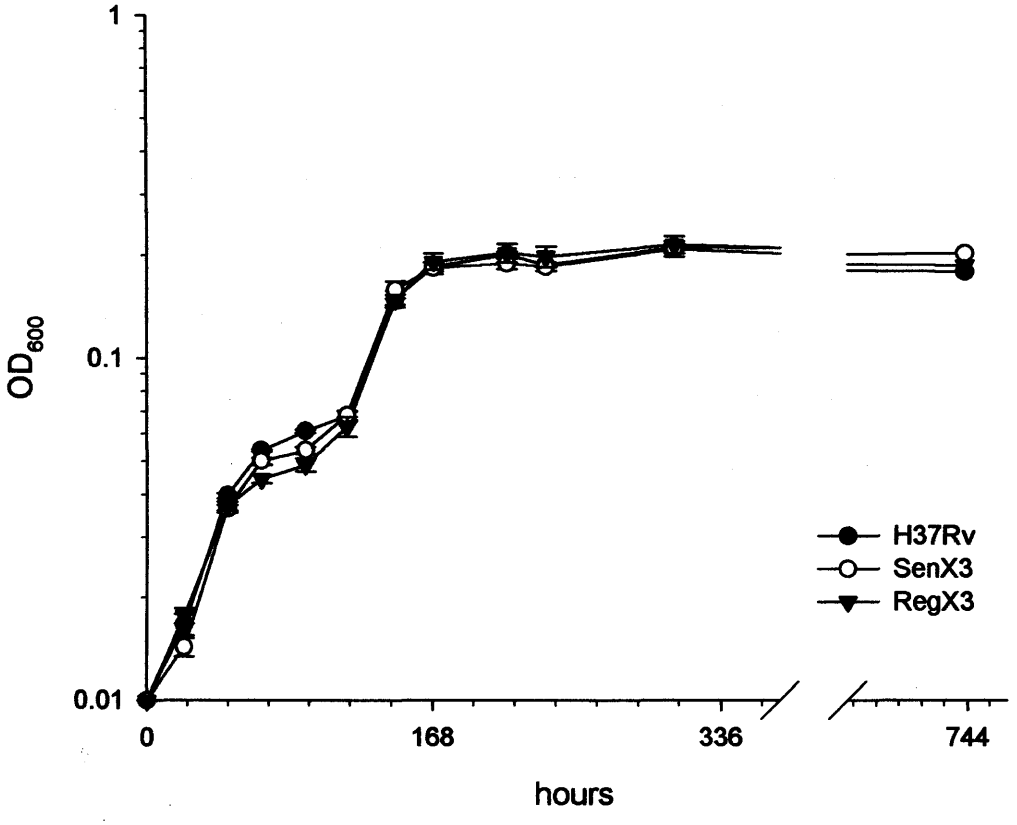

160
140
120
100
80
60
40
20
0
% viability
H37Rv
senX3
regX3
Strain

Figure 3.4: Growth of *senX3* and *regX3* mutants in media acidified to pH5.5 with HCl. Exponential-phase cultures grown in Dubos media with albumin and glycerol at 37°C in aerobic culture as per 2.1.1.4.1 were diluted at 1/500 into triplicate rolling cultures containing 100ml of Dubos media with albumin and glycerol acidified to pH5.5 with HCl and incubated at 37°C. Optical densities were followed over a period of one month. At the final time point the pH was between pH5 and pH6 using indicator strips (Merck). There was no statistical difference between the three strains as measured by a two-tailed homoscedastic Student's t-test.


1
0.1
0.01
OD600
0
168
336
744
hours
H37Rv
SenX3
RegX3

3.2.1.5 Toluidine blue assay

Mutants of ArcA in *E. coli* are hypersensitive to the dye toluidine blue (Roeder & Somerville 1979; Buxton & Drury 1983; Buxton *et al.* 1983; Buxton & Drury 1984). To assess whether the *senX3* and *regX3* mutants were more sensitive to the dye toluidine blue than the wild-type, bacteria from cultures grown aerobically in Dubos media with albumin and glycerol at 37°C (2.1.1.4.1) were streaked onto agar plates containing increasing concentations of the dye. All strains were unable to grow on agar plates containing 100μg/ml or more. The three strains were also inoculated into triplicate standing cultures containing titrations of toluidine blue in liquid media. These conditions are similar to those in 3.1.1.2 in which oxygen is thought to be the limiting growth factor. Again, no difference in sensitivity to the dye between the wild-type and the *senX3* and *regX3* null mutants was apparent (data not shown).

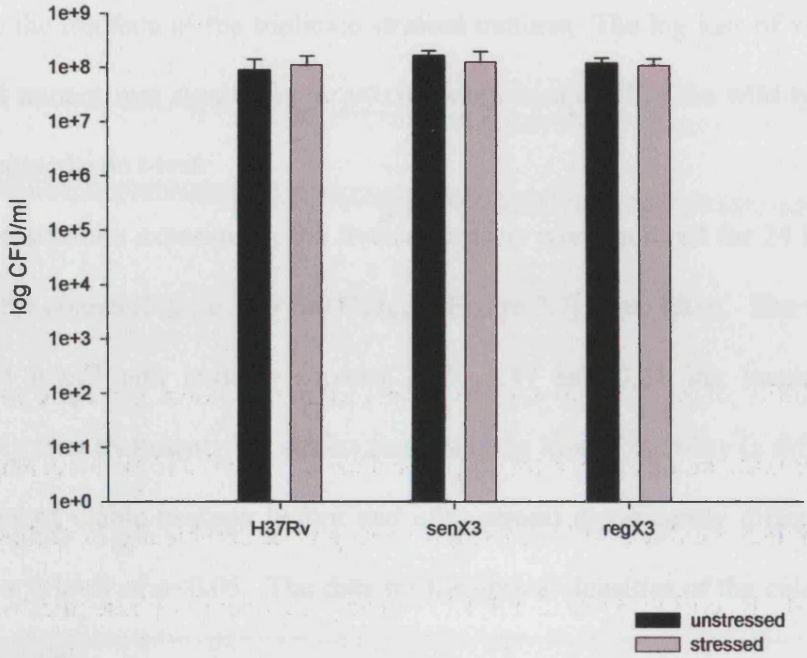
3.2.2 Stress studies

3.2.2.1 Superoxide stress

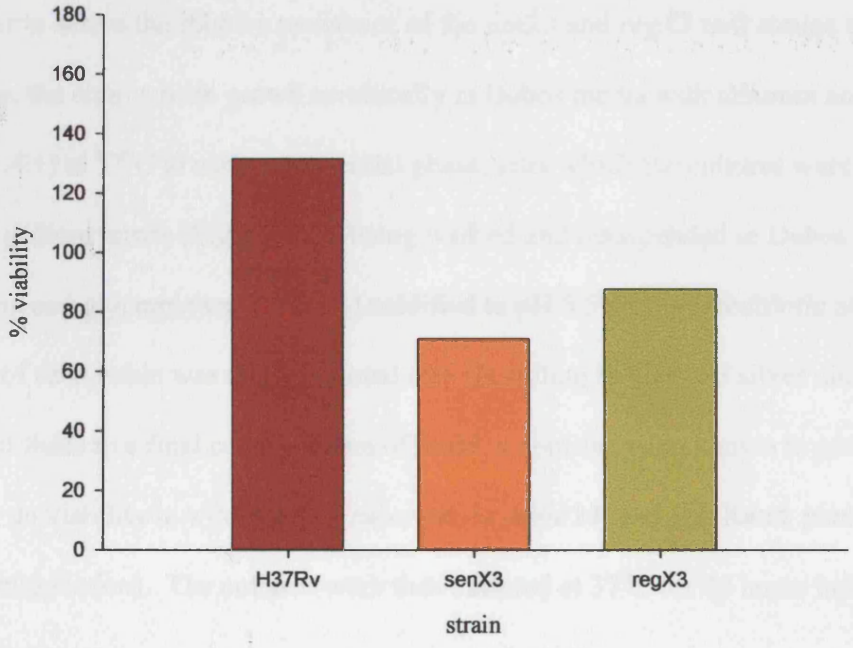
In order to assess the relative resistance of the *senX3* and *regX3* null strains and wild-type to superoxide stress, the strains were exposed to paraquat, a superoxide radical generator. The strains were grown aerobically to early exponential phase in Dubos media with albumin and glycerol at 37°C (see 2.1.1.4.1) then split into six aliquots of 20ml per strain in rolling bottles. Three 20ml aliquots of each strain were exposed to paraquat to a final concentration of 50mM and the three controls were exposed to volumes of solvent (in this case water) equal to that of the paraquat. After exposure the cultures were cultured for two further hours at 37°C before viable counts were taken (Figure 3.5). The

Figure 3.5: Viability of *senX3* and *regX3* mutants after two hours exposure to 50mM paraquat, a superoxide radical generator, expressed as A) cfu/ml and B) percentage drop in viability. Triplicate cultures of each strain were grown aerobically in Dubos media with albumin and glycerol at 37°C (2.1.1.4.1) to early exponential phase and were then split into two rolling bottles, half of which were treated with paraquat to a final concentration of 50mM. After two hours of further incubation at 37°C viable counts were taken. The log loss of viability of the *regX3* null mutant was significant at $p<0.03$ when compared to the wild-type by a two-tailed homoscedastic t-test.

A


1e+9
1e+8
1e+7
1e+6
1e+5
1e+4
1e+3
1e+2
1e+1
1e+0
log CFU/ml
H37Rv
senX3
regX3
unstressed
stressed

B


180
160
140
120
100
80
60
40
20
0
% viability
H37Rv
senX3
regX3
strain

wild-type and *senX3* and *regX3* null mutants showed -0.09, 0.13 and 0.05 log losses in viability respectively after treatment, comparing the log data of the triplicate unstressed cultures to the log data of the triplicate stressed cultures The log loss of viability of the *regX3* null mutant was significant at $p<0.03$ when compared to the wild-type by a two-tailed homoscedastic t-test.

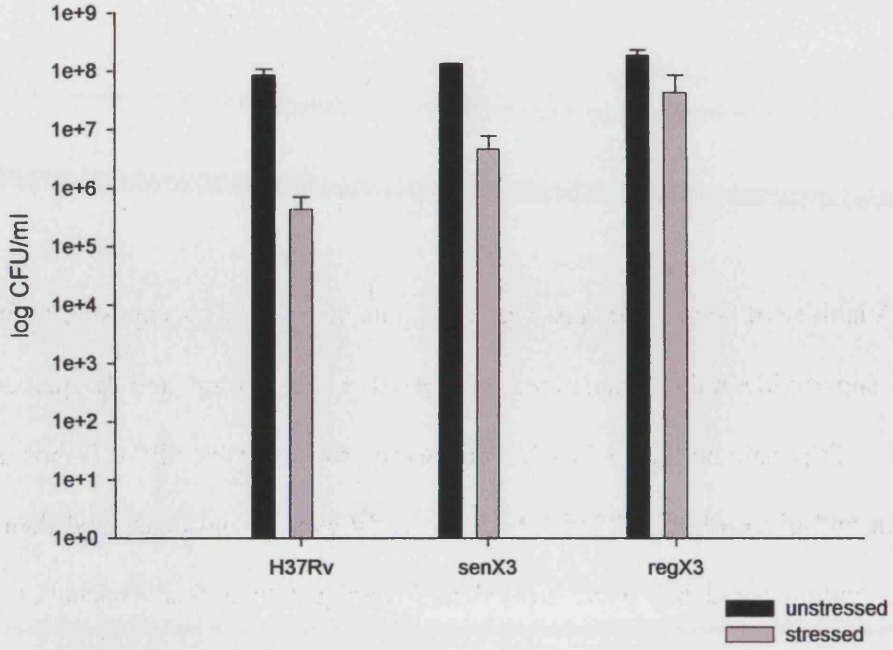
In a separate experiment, the treated bacteria were cultured for 24 hours at 37°C before viable counts (Figure 3.6) and OD_{600}s (Figure 3.7) were taken. The wild-type and *senX3* and *regX3* null mutants showed 2.29, 1.47 and 0.64 log losses in viability respectively after treatment. No strains showed a log loss of viability (a difference in the log number of viable bacteria before and after stress) significantly different from any other below a level of $p<0.05$. The data for the optical densities of the cultures supports this observation.

3.2.2.2 Nitrosative stress

In order to assess the relative resistance of the *senX3* and *regX3* null strains to nitrosative stress, the strains were grown aerobically in Dubos media with albumin and glycerol (2.1.1.4.1) at 37°C to early exponential phase, after which the cultures were exposed to 3mM sodium nitrite (Sigma) after being washed and resuspended in Dubos media with albumin and glycerol (see 2.1.1.4.1) acidified to pH 5.5 with hydrochloric acid (Sigma). 20 ml of each strain was then aliquoted into six rolling bottles and silver nitrite added to three of these to a final concentration of 3mM, a concentration known to produce a drop in viability in wild-type *M. tuberculosis* after 24 hours (L Rand, personal communication). The cultures were then cultured at 37°C for 24 hours before viable

Figure 3.6: Viability of *senX3* and *regX3* null mutants after 24 hours exposure to 50mM paraquat, a superoxide radical generator, expressed as A) cfu/ml and B) percentage viability loss. Triplicate cultures of each strain were grown aerobically in Dubos media with albumin and glycerol at 37°C (2.1.1.4.1) to early exponential phase and then split into separate cultures and half were treated with paraquat to a final concentration of 50mM. After 24 hours further incubation at 37°C viable counts were taken. No strains showed a log loss of viability significantly different from any other below a level of $p<0.05$.

A


1e+9
1e+8
1e+7
1e+6
1e+5
1e+4
1e+3
1e+2
1e+1
1e+0
log CFU/ml
H37Rv
senX3
regX3
unstressed
stressed

B

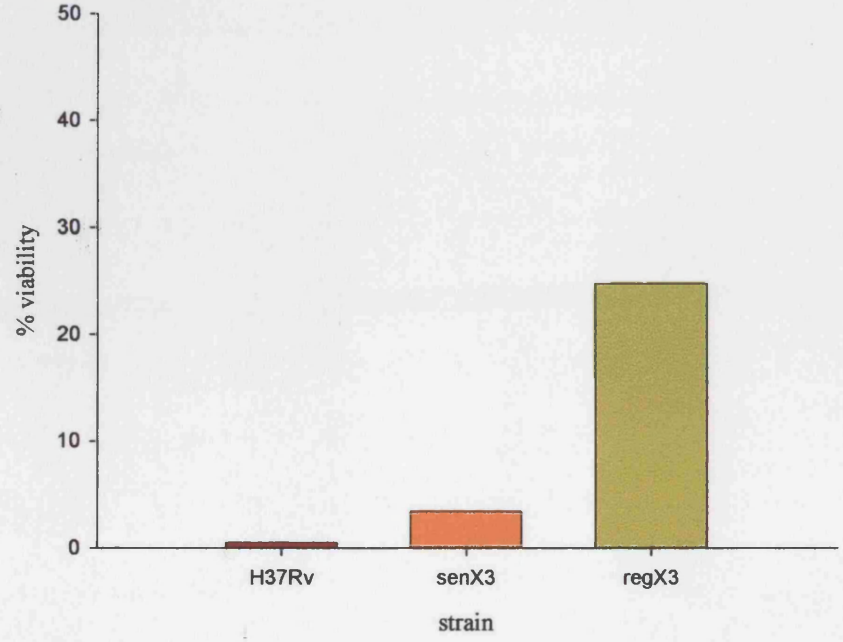
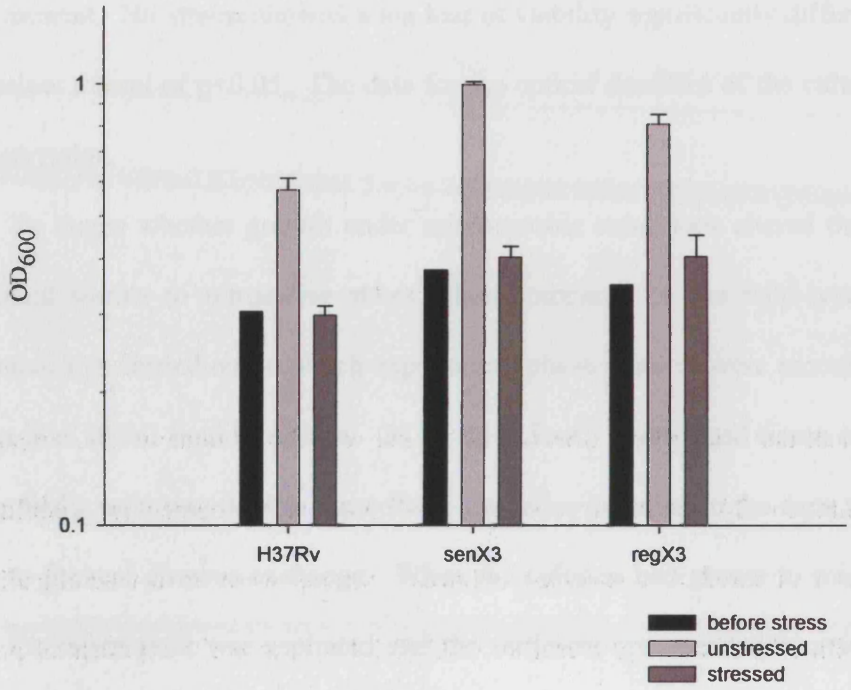
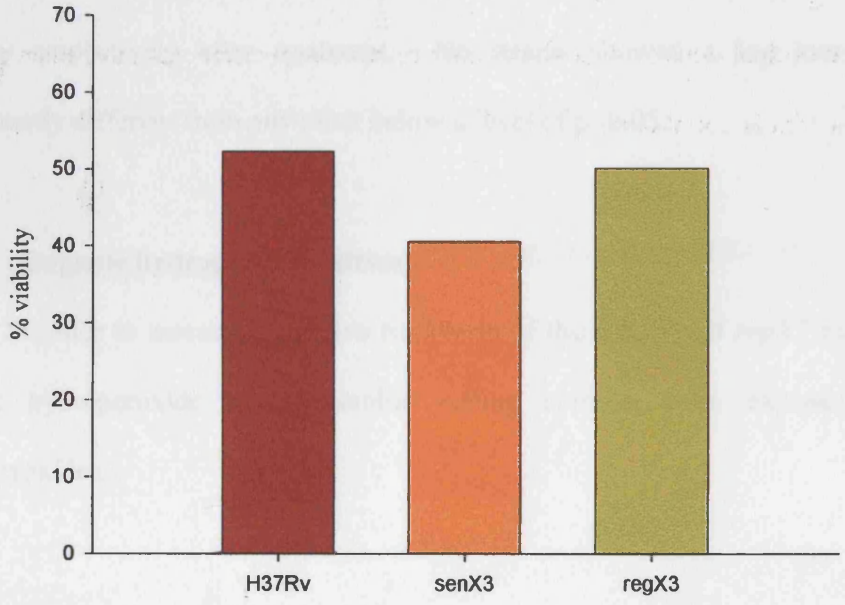

50
40
30
20
10
0
% viability
H37Rv
senX3
regX3
strain

Figure 3.7: Optical densities of cultures of *senX3* and *regX3* null mutants after exposure to 50mM paraquat for 24 hours, expressed as A) OD_{600} and B) percentage viability of stressed cultures compared to unstressed cultures. Strains were grown aerobically in Dubos media with albumin and glycerol at 37°C (2.1.1.4.1) to early exponential phase, the OD_{600} taken and triplicate cultures of each strain were split and half were treated with paraquat to a final concentration of 50mM. After 24 hours of further incubation at 37°C the OD_{600} of the stressed and unstressed cultures were taken again. Statistics were not performed on the optical density data since more reliable measurements have been taken in the form of the viable count data.

A


1
0.1
OD_{600}
H37Rv
senX3
regX3
before stress
unstressed
stressed

B


70
60
50
40
30
20
10
0
% viability
H37Rv
senX3
regX3

counts (Figure 3.8) and optical densities (Figure 3.9) were measured. The wild-type and *senX3* and *regX3* null mutants showed 0.1, 0.9 and 0.6 log losses in viability respectively after treatment. No strains showed a log loss of viability significantly different from any other below a level of $p<0.05$. The data for the optical densities of the cultures supports this observation.

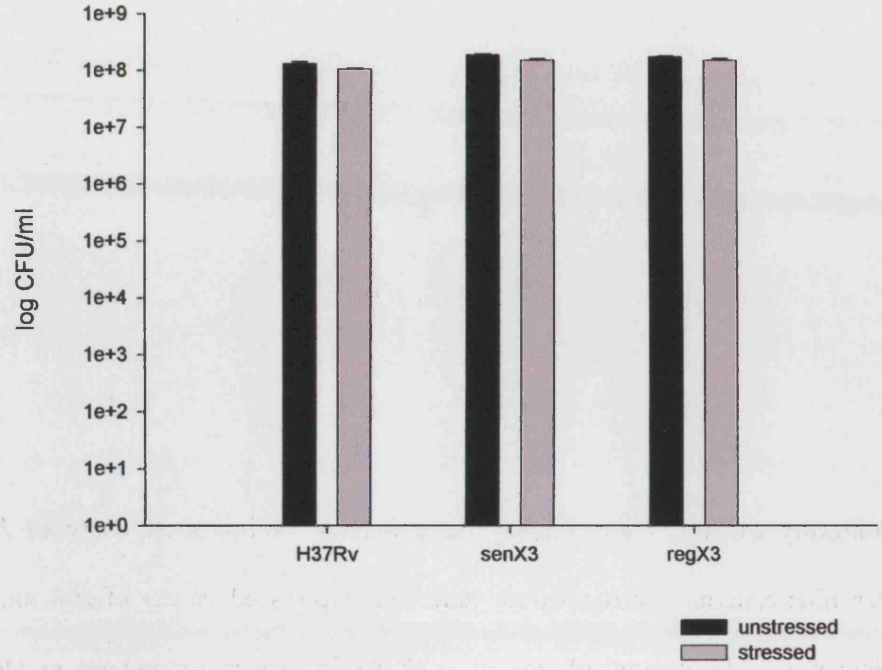
To assess whether growth under microaerobic conditions altered the viability of the mutant strains to nitrosative stress when compared to the wild-type, a separate experiment was carried out in which exponential-phase cultures were inoculated at 1/500 dilution into 100ml liquid media in 262.5 ml ($75cm^2$) non-vented tissue culture flasks. The air:media ratio was 0.38 in these flasks and after inoculation the caps were screwed tightly to prevent gaseous exchange. When the cultures had grown to mid-exponential phase, the supernatant was aspirated and the sediment resuspended in acidified media. For each strain, two cultures had silver nitrite added to a final concentration of 3mM, and two cultures left as controls. Viable counts were taken after 24 hours (Figure 3.10). The wild-type and *senX3* and *regX3* null mutants showed 0.49, 0.48 and 0.43 log losses in viability respectively after treatment. No strains showed a log loss of viability significantly different from any other below a level of $p<0.05$.

3.2.2.3 Organic hydroperoxide stress

In order to assess the relative resistance of the *senX3* and *regX3* null mutants to organic hydroperoxide stress, aerobic rolling cultures were exposed to organic hydroperoxides.

Figure 3.8: Viability of *senX3* and *regX3* mutants after 24 hours exposure to 3mM acidified silver nitrite, a nitrosative stress generator, expressed as A) cfu/ml and B) percentage viabilities. Triplicate cultures of each strain grown aerobically in Dubos media with albumin and glycerol at 37°C (2.1.1.4.1) to early exponential phase were washed and resuspended in media acidified to pH5.5. The cultures were then split and half were treated with silver nitrite to a final concentration of 50mM. After 24 hours further incubation at 37°C viable counts were taken. No strains showed a statistically different log loss in viability compared to the wild-type using a two-tailed homoscedastic Student's t-test.

A


1e+9
1e+8
1e+7
1e+6
1e+5
1e+4
1e+3
1e+2
1e+1
1e+0
log CFU/ml
H37Rv
senX3
regX3
unstressed
stressed

B

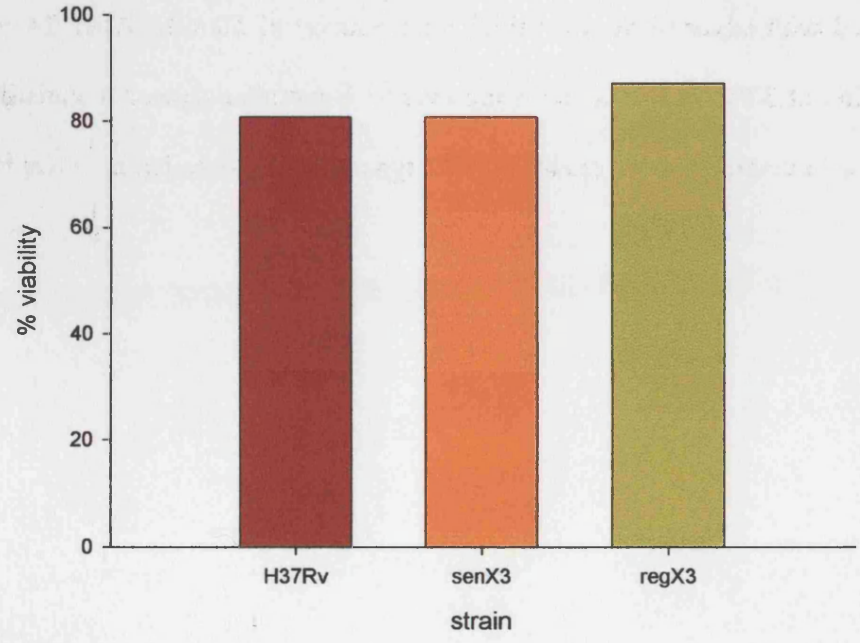
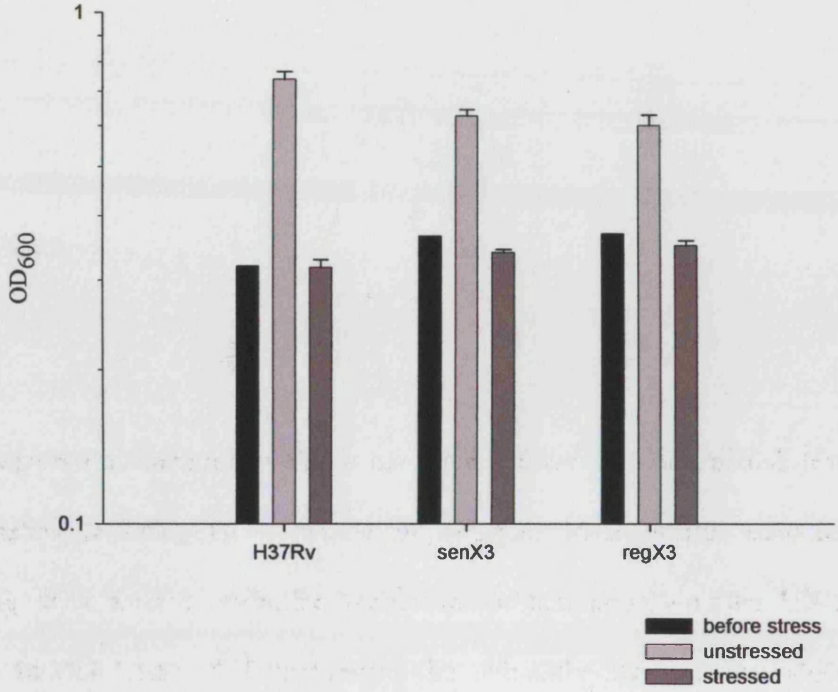

100
80
60
40
20
0
% viability
H37Rv
senX3
regX3
strain

Figure 3.9: Optical densities of cultures of *senX3* and *regX3* null mutants after exposure to 3mM acidified silver nitrite for 24 hours, expressed as A) OD_{600} and B) percentage viability of stressed cultures compared to unstressed cultures. Strains were grown aerobically in Dubos media with albumin and glycerol at 37°C (2.1.1.4.1) to early exponential phase, then washed and resuspended in media acidified to pH5.5. The OD_{600} was taken before triplicate cultures of each strain were split and half were treated with silver nitrite to a final concentration of 3mM. After 24 hours of further incubation at 37°C the OD_{600} of the stressed and unstressed cultures were taken again. Statistics were not performed on the optical density data since more reliable measurements have been taken in the form of the viable count data.

A


1
0.1
OD600
H37Rv
senX3
regX3
before stress
unstressed
stressed

B

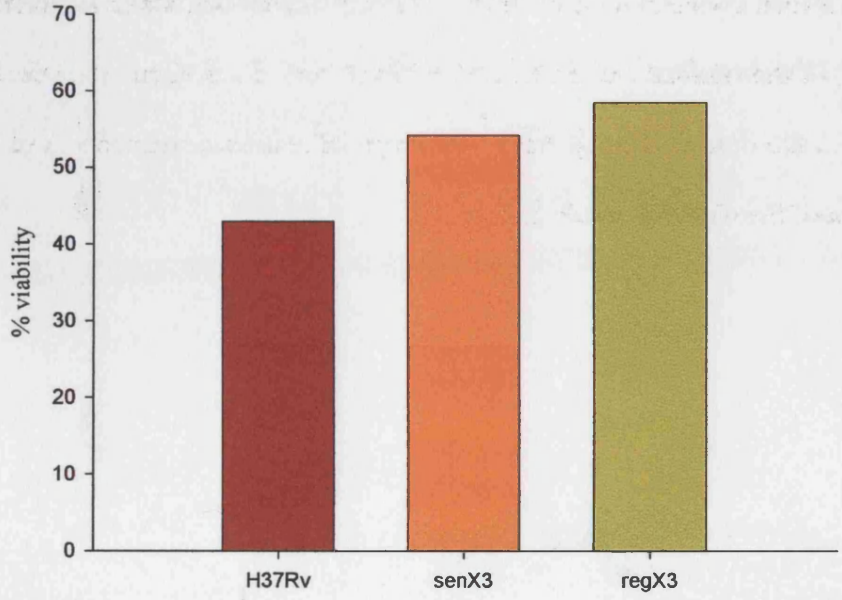
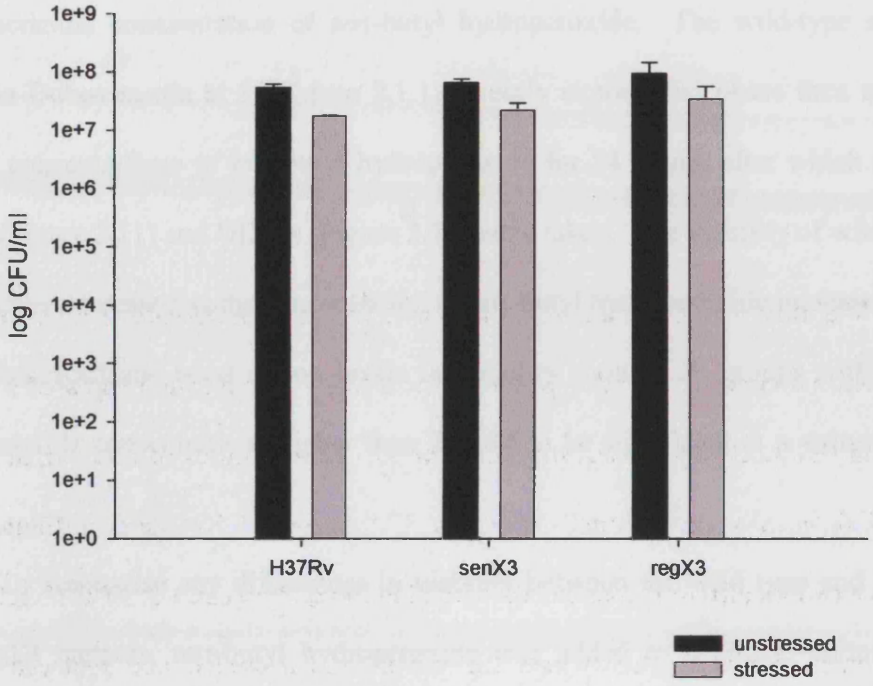
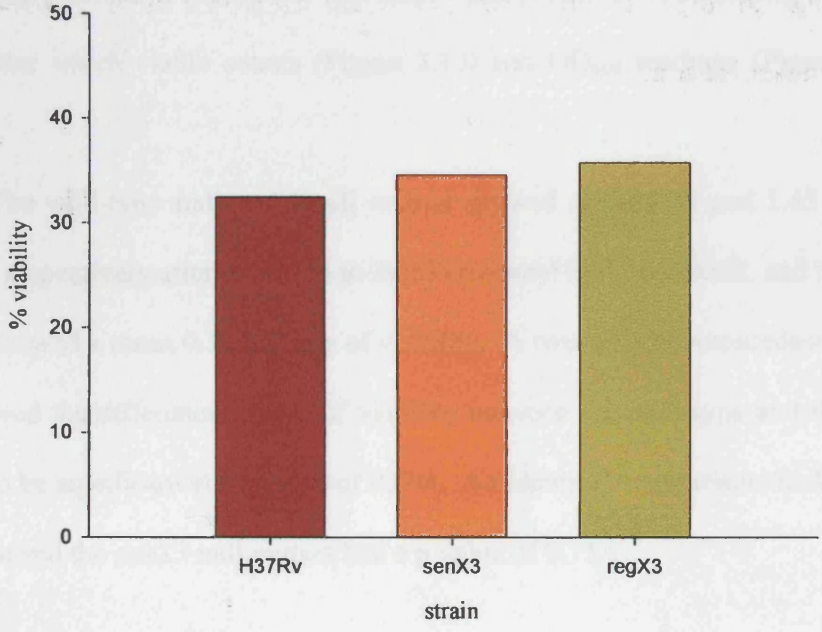

70
60
50
40
30
20
10
0
% viability
H37Rv
senX3
regX3

Figure 3.10: Viability of *senX3* and *regX3* null mutants of *M. tuberculosis* after growth under microaerobic conditions and exposure to acidified silver nitrite, expressed as (A) log cfu/ml and (B) % change in viability after stress. Four cultures of each strain were grown for 17 days in oxygen-limited culture at 37°C from a 1/500 starting dilution of bacteria grown aerobically in Dubos media with albumin and glycerol at 37°C (2.1.1.4.1) to early exponential phase, after which the media was aspirated and replaced with acidified media (pH5.5). Two cultures for each strain were then exposed to silver nitrite. After 24 hours of further incubation viable counts were taken. No strains showed a significant change in log loss of viability when compared with a two-tailed homoscedastic Student's t-test.

A


1e+9
1e+8
1e+7
1e+6
1e+5
1e+4
1e+3
1e+2
1e+1
1e+0
log CFU/ml
H37Rv
senX3
regX3
unstressed
stressed

B


50
40
30
20
10
0
% viability
H37Rv
senX3
regX3
strain

Since the author was unaware of any similar studies on *M. tuberculosis*, it was decided to titrate *tert*-butyl hydroperoxide against wild-type *M. tuberculosis* to establish the bactericidal concentration of *tert*-butyl hydroperoxide. The wild-type strain was grown in Dubos media at 37°C (see 2.1.1) to early exponential phase then exposed to varying concentrations of *tert*-butyl hydroperoxide for 24 hours, after which the viable counts (Figure 3.11) and OD_{600}s (Figure 3.12) were taken. The viability of wild-type *M. tuberculosis* decreased as the concentration of tert-butyl hydroperoxide increased. A one-tailed homoscedastic t-test of log losses in viability showed all groups with *tert*-butyl hydroperoxide concentrations higher than 250μM to be significant at p values less than 0.05.

To emphasise any differences in viability between the wild type and the *senX3* and *regX3* mutants, *tert*-butyl hydroperoxide was added to triplicate cultures of the strains to a final concentration of 2mM, which was the maximum concentration titrated against the wild-type. The strains were grown as normal (see 2.1.1.4) to early exponential phase and exposed to *tert*-butyl hydroperoxide while rolling at 37°C for 24 hours, after which viable counts (Figure 3.13) and OD_{600} readings (Figure 3.14) were taken.

The wild-type and *senX3* null mutant showed mean 1.47 and 1.45 log losses of viability respectively after exposure to 2mM *tert*-butyl hydroperoxide, and the *regX3* null mutant showed a mean 0.53 log loss of viability. A two-tailed homoscedastic student's t-test showed the difference in loss of viability between the wild-type and the *regX3* null mutant to be significant at a p value of 0.008. An identical comparison made between the wild-type and the *senX3* null mutant had a p value of 0.75.

Figure 3.11: Titration of *tert*-butyl hydroperoxide against wild-type *M. tuberculosis*. The strain was grown aerobically in Dubos media with albumin and glycerol at 37°C (2.1.1.4.1) to early exponential phase before being split into separate triplicate cultures. Each group of three cultures was exposed to either 0 (water control), 100, 250, 500, 1000 or 2000 μM concentration of *tert*-butyl hydroperoxide and incubated for a further 24 hours at 37°C, after which viable counts were taken. At all concentrations of tert-butyl hydroperoxide over 100μM the log loss in viability was significant at a level of $p<0.05$ when compared to the untreated control.

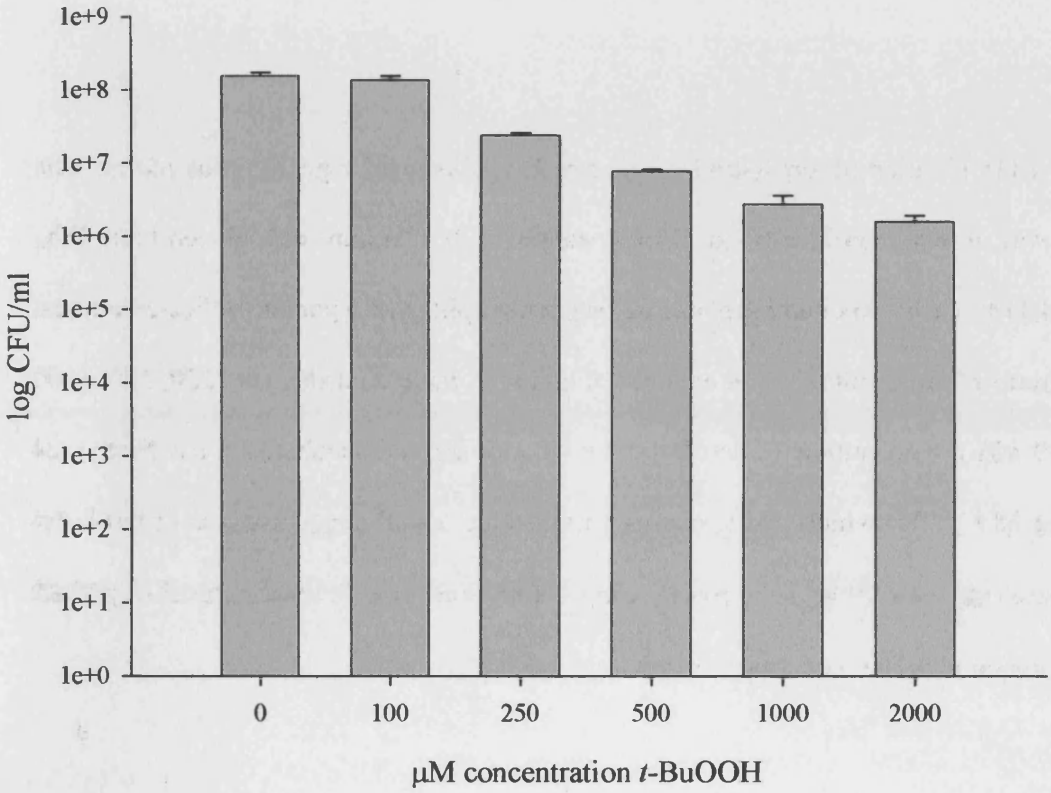

1e+9
1e+8
1e+7
1e+6
1e+5
1e+4
1e+3
1e+2
1e+1
1e+0
log CFU/ml
0
100
250
500
1000
2000
μM concentration t-BuOOH

Figure 3.12: Titration of *tert*-butyl hydroperoxide against wild-type *M. tuberculosis*. The strain was grown aerobically in Dubos media with albumin and glycerol at 37°C (2.1.1.4.1) to early exponential phase before being split into separate triplicate cultures. Each group of three cultures was exposed to either 0 (water control), 100, 250, 500, 1000 or 2000 μM concentration of *tert*-butyl hydroperoxide and incubated for a further 24 hours at 37°C, after which OD_{600} readings were taken. At all concentrations of tert-butyl hydroperoxide over 100μM the log loss in viability was significant at a level of $p<0.05$ when compared to the untreated control. Input: the OD_{600} of the strain before the *tert*-butyl hydroperoxide was added.

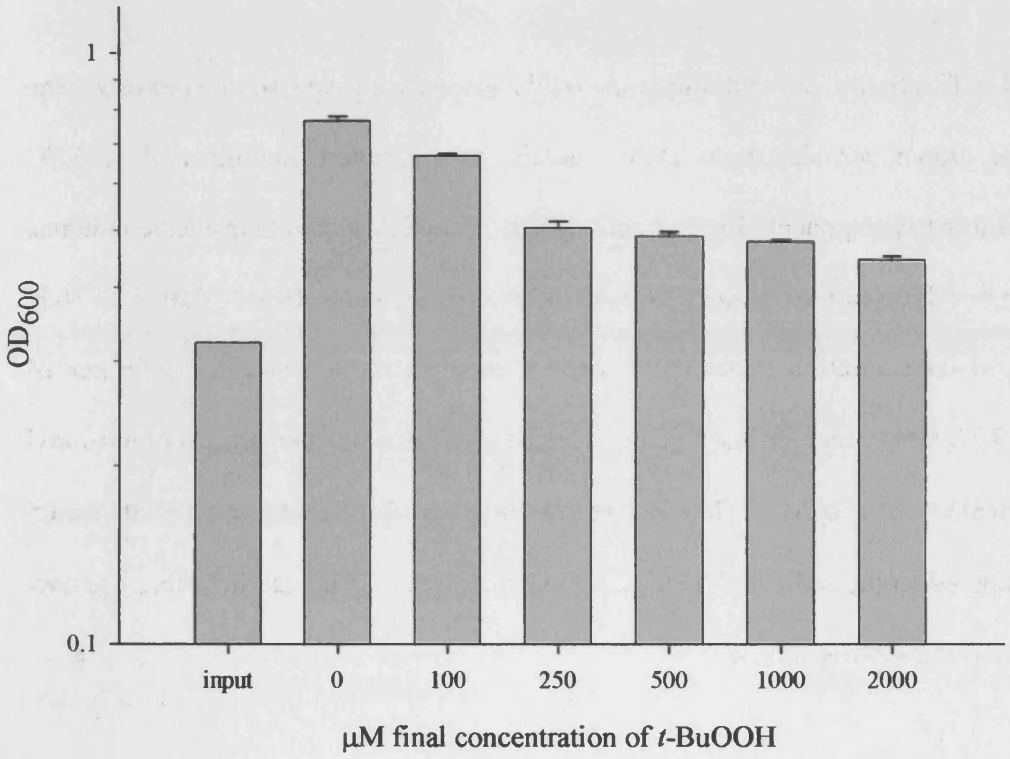
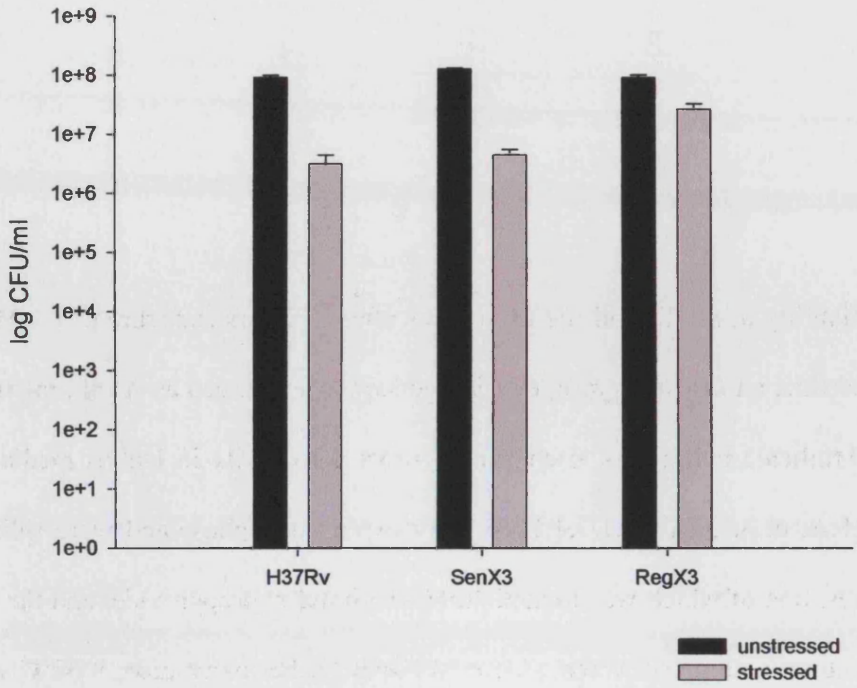

1
0.1
OD_{600}
input
0
100
250
500
1000
2000
μM final concentration of *t*-BuOOH

Figure 3.13: Viability of *senX3* and *regX3* mutants after 24 hours exposure to 2 mM *tert*-butyl hydroperoxide, an organic hydroperoxide generator, expressed as A) cfu/ml and B) % viability. Triplicate cultures of each strain grown aerobically in Dubos media with albumin and glycerol at 37°C (2.1.1.4.1) to early exponential phase and were split into separate cultures, one of which was treated 2mM *tert*-butyl hydroperoxide, and the other exposed to an equal volume of water. After 24 hours further incubation at 37°C viable counts were taken. The dashed lines represent the limits of the assay at using the dilutions described in 2.2.1.1. The log loss in viability of the *regX3* null mutant was significant at a level of $p<0.008$.

A


1e+9
1e+8
1e+7
1e+6
1e+5
1e+4
1e+3
1e+2
1e+1
1e+0
log CFU/ml
H37Rv
SenX3
RegX3
unstressed
stressed

B

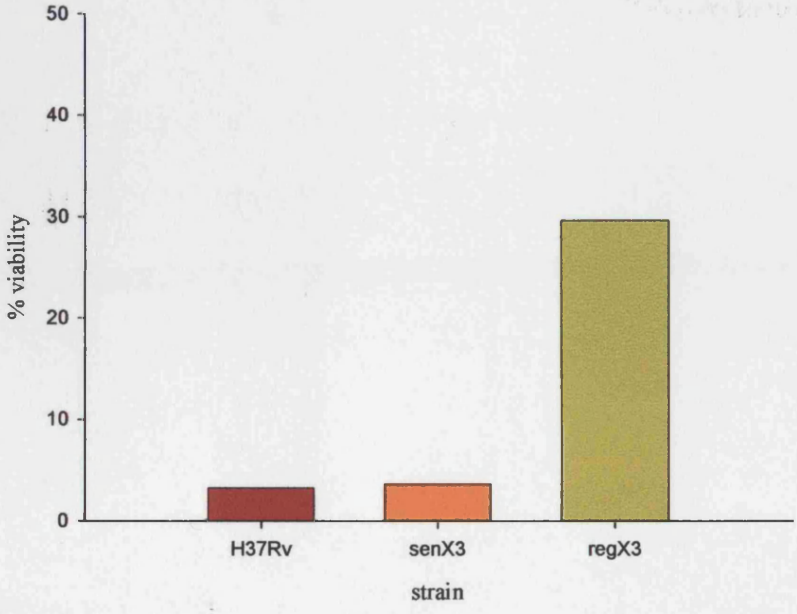
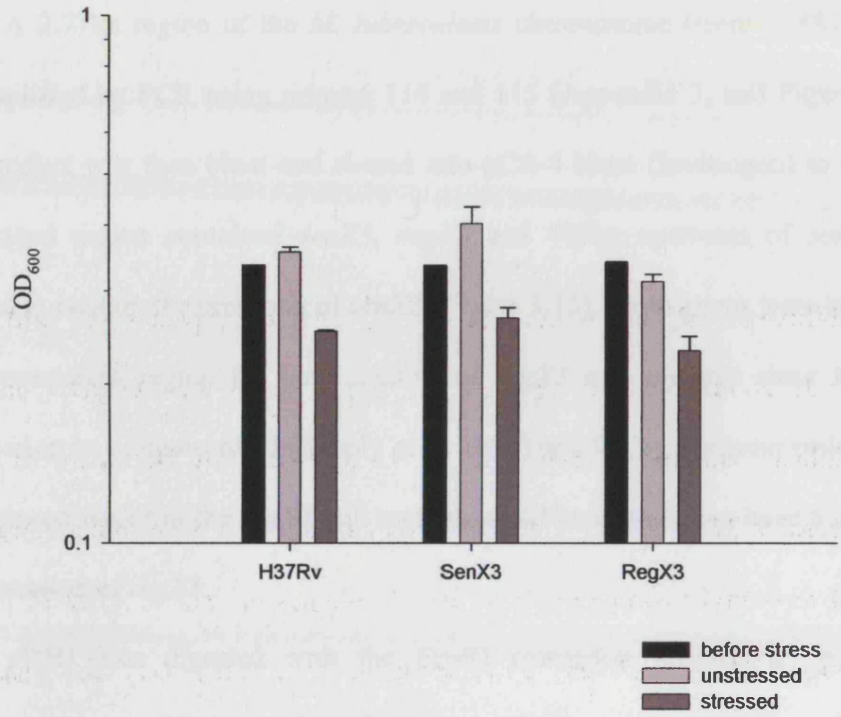
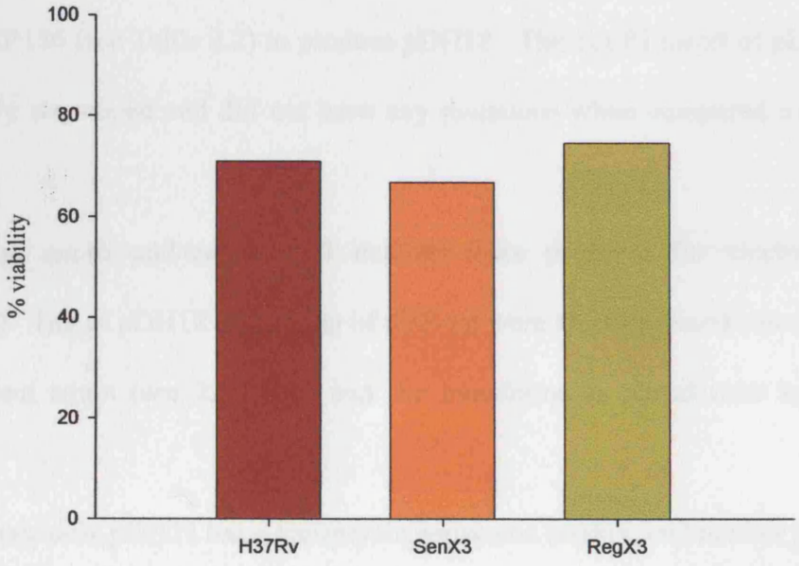

50
40
30
20
10
0
% viability
H37Rv
senX3
regX3
strain

Figure 3.14: Viability of *senX3* and *regX3* mutants after 24 hours exposure to 2 mM *tert*-butyl hydroperoxide, an organic hydroperoxide generator, expressed as A) OD_{600} and B) % viability. Triplicate cultures of each strain grown aerobically in Dubos media with albumin and glycerol at 37°C (2.1.1.4.1) to early exponential phase were split and treated with either 2mM *tert*-butyl hydroperoxide, or equal volumes of water. After 24 hours further incubation at 37°C optical density measurements were taken. Statistics were not performed on these data since more reliable figures were provided by the viability data shown in Table 3.13.

A


1
0.1
OD_{600}
H37Rv
SenX3
RegX3
before stress
unstressed
stressed

B


100
80
60
40
20
0
% viability
H37Rv
SenX3
RegX3

3.2.3 *In vivo* characterisation of the *senX3* and *regX3* null mutants

3.2.3.1 Construction of complemented *senX3* and *regX3* null mutants

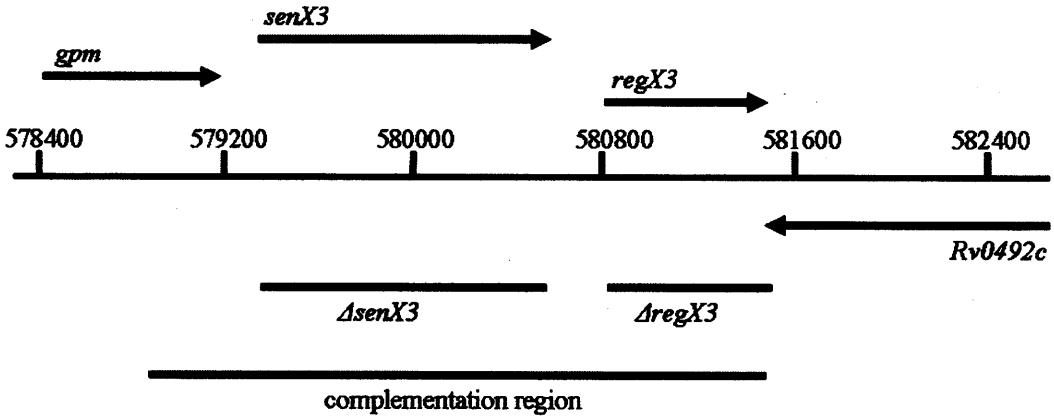
A 2.71kb region of the *M. tuberculosis* chromosome (from 578814 to 581523) was amplified by PCR using primers 114 and 115 (Appendix 3, and Figure 3.15). The PCR product was then blunt-end cloned into pCR-4 blunt (Invitrogen) to create pDH1. The cloned region contained *senX3*, *regX3* and 469bp upstream of *senX3* that was assumed to contain the promoter of *senX3* (Figure 3.15). Both genes were included in the complementation region for both *senX3* and *regX3* null mutants since the genes are presumed to be co-transcribed (Supply *et al.* 1997) and the hygromycin resistance marker that replaced *senX3* in the *senX3* null mutant would be expected to have a polar effect of the expression of *regX3*.

pDH1 was digested with the *Eco*RI restriction enzyme to yield an insert containing the amplified *senX3*, *regX3* and a 223 bp region upstream of *senX3*. The region upstream of *senX3* was shorter here due to the presence of an *Eco*RI site in the cloned sequence. The insert was then cloned into the *Eco*RI site of the $attP^{+}/int^{-}$ suicide vector pKP186 (see Table 2.2) to produce pDH12. The *Eco*RI insert of pDH12 was then completely sequenced and did not have any mutations when compared to the wild-type sequence.

The *senX3* and *regX3* null mutants were prepared for electroporation (see 2.1.1.4.8). 1μg of pDH12 and 300ng of pBS-int were electroporated simultaneously into each mutant strain (see 2.1.1.4.8), and the transformants plated onto kanamycin agar plates.

Since only pDH12 has a kanamycin resistance marker, and neither pDH12 or

Figure 3.15: The *senX3* and *regX3* region of the *M. tuberculosis* chromosome. Δ*senX3* and Δ*regX3* represent the regions deleted in the *senX3* and *regX3* null mutants, respectively. The complement region shows the relative position of the functional allele cloned into the complementing vector pDH12. gpm is a probable phosphoglycerate mutase, and Rv0492c encodes a probable oxidoreductase (Tuberculist, 2005).


senX3
gpm
regX3
578400
579200
580000
580800
581600
582400
Rv0492c
ΔsenX3
ΔregX3
complementation region

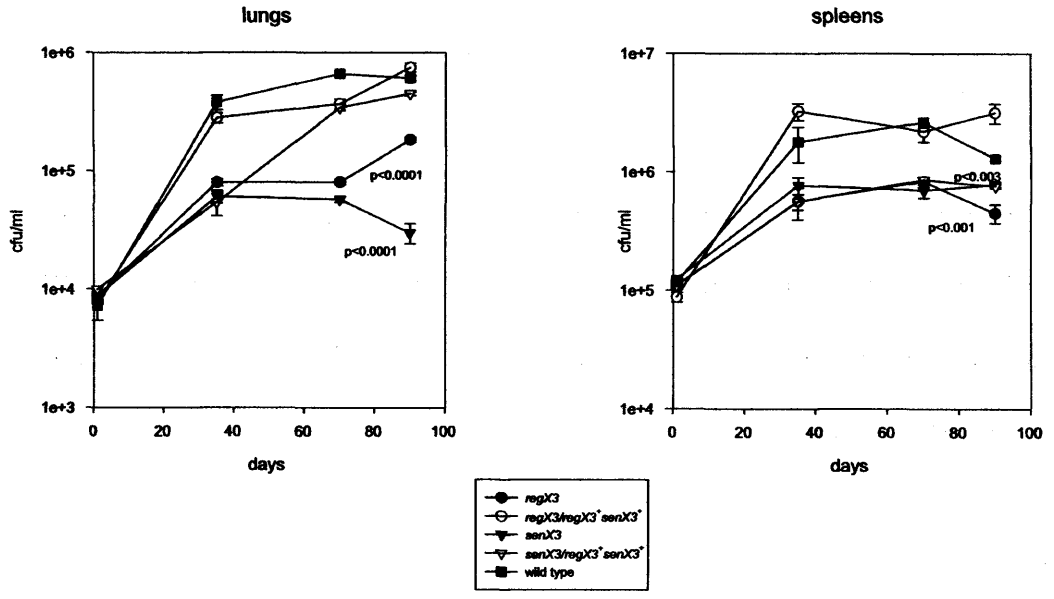
pBS-int have mycobacterial origins of replication, only bacteria that have acquired a stable copy of pDH12 integrated into the chromosome should be able to form colonies (Stover *et al.* 1991). Kanamycin-resistant colonies were then grown without selection in standing cultures and frozen stocks prepared.

3.2.3.2 Growth and complementation of *senX3* and *regX3* null mutants in a murine model of infection

To determine whether the *senX3* or *regX3* genes play a role in the virulence of *M. tuberculosis*, the *senX3* and *regX3* null mutants were assayed in a murine intravenous injection model of infection. The two mutant strains, the wild-type and the two mutant strains containing pDH12 were grown aerobically (see 2.1.1). When the cultures reached mid-exponential phase they were diluted to an OD_{600} of 0.024 in sterile phosphate-buffered saline pH 7.4 and injected into the tail veins of 6-8 week old female Balb/C mice. Groups of three to five mice were killed over a period of 140 days and the number of viable *M. tuberculosis* bacilli determined in the lungs and spleen (Figure 3.16).

Both the *senX3* and *regX3* null mutants grew to about $5x10^4$ CFU/gram tissue in the lungs at 36 days, after which they did not appear to continue growing. The wild-type, in comparison, grew to $4x10^5$ CFU/gram tissue in the lungs at day 36, gradually increasing to about $1x10^6$ at day 140. The difference between the *senX3* and *regX3* null mutants and the wild type was significant at a level of $p<0.0002$ and $p<0.00001$, respectively. The *regX3* null mutant containing the complementing allele on the chromosome via pDH12 grew to the same levels as the wild-type in the lungs, however

Figure 3.16: Growth of wild-type *M. tuberculosis*, *senX3* and *regX3* null mutants of *M. tuberculosis* and the *senX3* and *regX3* null mutants carrying the plasmid pDH12 integrated into the chromosome. The infection was monitored by removing the lungs and spleens of 3-5 infected mice per strain at intervals, and viable counts performed after serial dilution in 0.15% saline.


lungs
spleens
cfu/ml
days
1e+3
1e+4
1e+5
1e+6
1e+7
0
20
40
60
80
100
p<0.0001
p<0.0001
p<0.003
p<0.001
regX3
regX3/regX3+senX3+
senX3
senX3/regX3+senX3+
wild type

the growth of the *senX3* null mutant did not reach similar levels as the wild-type until after 60 days.

The growth of the five strains in the spleens was largely similar to the growth in the lungs, although the rate of growth between the different time points for the different strains was less predictable. At 140 days the difference in CFU/gram tissue in the spleens between the *senX3* and *regX3* null mutants and the wild-type and complementing strains was about one log, and was significant at levels of $p<0.003$ and $p<0.002$ for the *senX3* and *regX3* null mutants respectively as compared by paired t-test. The CFU/gram tissue were approximately one log higher in the spleens, which is most likely due to the greater exposure the spleen has to bacilli during the initial stages of an intravenous infection. A notable difference in the spleens is that the growth of the *senX3* null mutant carrying pDH12 is that restoration of virulence by the functional allele does not seem to occur until 140 days post infection.

The complementation of the *senX3* and *regX3* null mutants shows conclusively that it is the mutations in these particular genes that give rise to the attenuated phenotype in the mouse model, rather than some genetic polar effect or other fortuitous mutation elsewhere on the genome. Other studies of isogenic mutants in this signal transduction system have not performed this complementation test (Parish *et al.* 2003).

3.2.3.3 Complementation of the *senX3* and *regX3* null mutants in bone-marrow derived macrophages

To further examine the role of the *senX3* and *regX3* genes in a more defined *in vitro* model of infection, the *senX3* and *regX3* null mutants, the wild-type and the

complementing strains were used to infect primary macrophages that had not been activated by the addition of exogenous interferon-γ.

The five strains were grown as before (see 2.1.1) to mid-exponential phase before dilution to an OD_{600} of 0.04 in Iscove's Modified Dulbecco's media (IMDM). The bacilli were then used to infect bone marrow-derived macrophages from 8 month old female Balb/C mice at a multiplicity of infection of two macrophages per one bacillus. Infections were carried out in triplicate at each time point (Figure 3.17).

All strains were present at about $8x10^4$ bacteria per millilitre of lysed macrophage supernatant at day zero. The viable count of the wild-type and the complementing strains increased over 12 days to approx $7x10^5$, whereas the *senX3* null mutant showed a slight loss of viability over the same period of time to a viable count of around $4x10^4$. The *regX3* null mutant showed slight increase in viable count during the first week to around $0.5x10^5$, after which the growth remained static.

3.2.3.4 PCR analysis of the complementing *senX3* and *regX3* strains used in the murine infection (3.2.3.2)

DNA was extracted from a plate culture of the *senX3* null mutant carrying the complementing allele recovered from the mice and two co-transformants (see 3.2.3.2, and method at 2.4.1.2). *regX3* and part of *senX3* was amplified from the complementing allele by using a primer from the *senX3* sequence (primer 104) and one specific for the plasmid (primer PMV306-4)(Figure 3.18 and Appendix 3). All products were of the expected size, and when sequenced, products from all three *senX3* null mutant strains

Figure 3.17: Growth of *senX3* and *regX3* null mutants of *M. tuberculosis* and their respective complementing strains in a macrophage infection assay. Strains were grown to early exponential phase before dilution to an OD_{600} of 0.04 in IMDM. Cells obtained from the bone marrow of BALB/c mice were infected at a multiplicity of infection of 2:1, and the infection followed by lysis and viable counts of triplicate cultures at intervals. Comp: complemented by the insertion of pDH12. The y axis units represent $\log_{10}$ colony forming units per ml lysate. H37Rv is the wild-type strain.

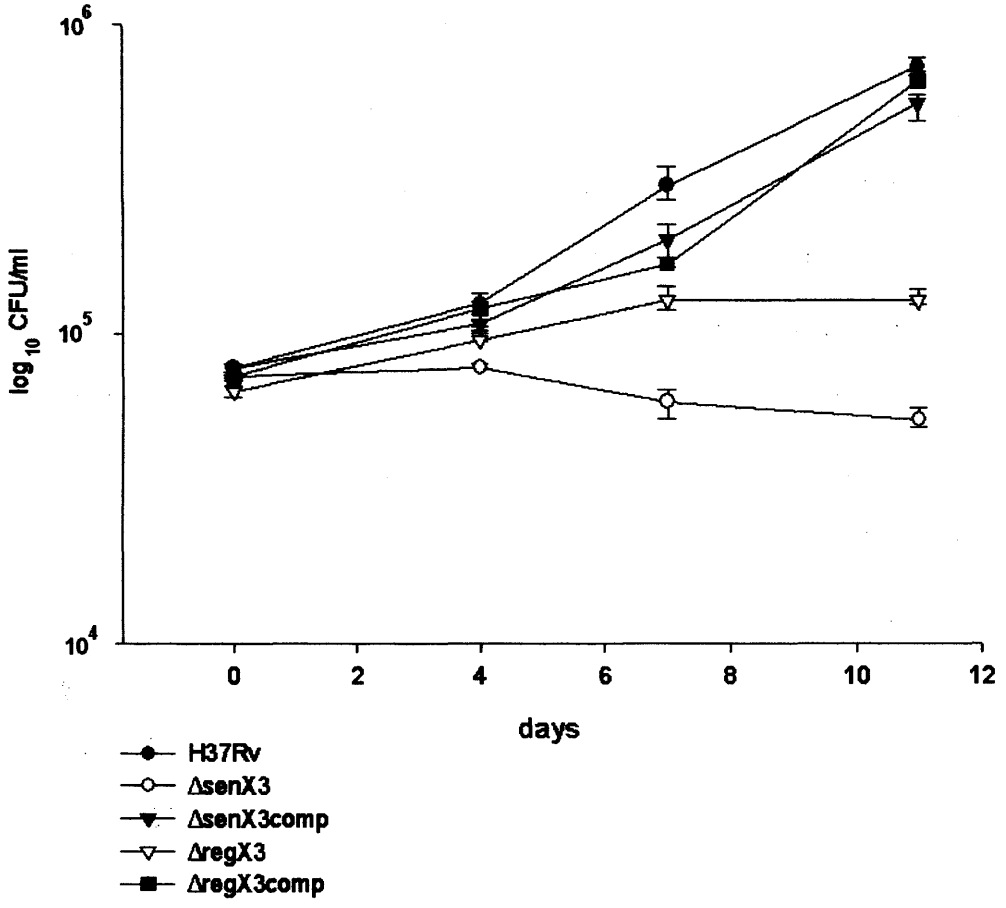
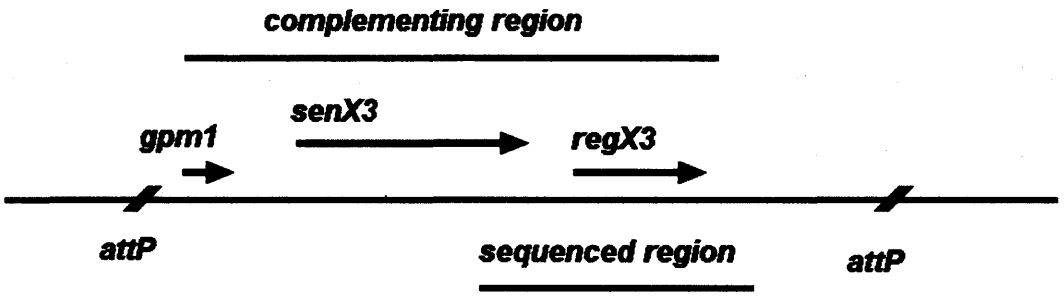

10^6
10^5
10^4
log 10 CFU/ml
0
2
4
6
8
10
12
days
H37Rv
ΔsenX3
ΔsenX3comp
ΔregX3
ΔregX3comp

Figure 3.18: Sequenced region of the chromosome of the *senX3* null mutant carrying a functional allele on the plasmid pDH12.


complementing region
senX3
gpm1
regX3
attP
sequenced region
attP

carrying the complementing plasmid were identical to the published sequence of *M. tuberculosis*.

3.2.3.5 Transcriptomic microarray analysis of the complementing *senX3* and *regX3* null mutant strains

To examine the transcription levels of *senX3* and *regX3* in the complementing *senX3* and *regX3* null mutant strain, one culture each of the wild-type, the complementing strains and the mutants were grown aerobically (2.1.1.4.3) and RNA was extracted (2.4.2). The RNA of each mutant or complementing strain was then compared to the RNA of the wild-type (2.5.4.1). The microarray results were analysed using the GENESPRING software by the method described in 2.5.8.2. It appears that the complementing allele in the *senX3* null mutant is unable to produce a *regX3* transcript (Figure 3.19), whereas the same allele in the *regX3* null mutant restores wild-type levels of the transcript (Figure 3.20).

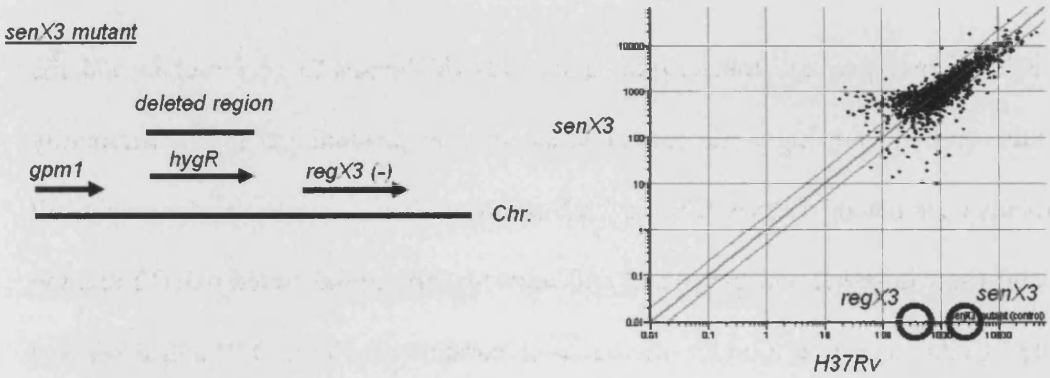
3.3 DISCUSSION

3.3.1 *In vitro* assessment of the *senX3* and *regX3* null mutants

3.3.1.1 Aerobic growth curve of the *senX3* and *regX3* null mutants

Both the *senX3* and *regX3* null mutants grow at approximately the same rate as the wild-type under aerobic growth conditions (see 3.2.1.1), a phenotype that would be expected of strains of *M. tuberculosis* mutated in non-essential genes. An average generation time of ~ 14 hours during exponential phase in rolling bottles, and this was approximately the case here in this study (Figure 3.1).

Figure 3.19: Genotypes and transcriptional profiles of the *senX3* null mutant and the *senX3* null mutant carrying a functional allele on the plasmid pDH12. Microarray comparisons were made by RNA-RNA hybridisation (see 2.5.4.1) between the *senX3* null mutant and the wild-type, and the *senX3* null mutant with the integrated pDH12 and the wild-type. Cells were grown under microaerobic conditions (see 2.2.1.2) and RNA was extracted when the bacteria had reached stationary phase. Each scatter plot represents the results from two dye-swapped slides, derived from a single biological experiment. IP, integrated plasmid; hygR, hygromycin resistance cassette. The scatter plot represents the relative normalised ratios of transcript abundance between two strains. The units on the axes are the median normalised intensity of the spot for each gene on the microarray, after subtraction of the median normalised intensity of the background pixels in the area adjacent to the spot.


senX3 mutant
deleted region
gpm1
hygR
regX3 (-)
Chr.
senX3
regX3
senX3
H37Rv

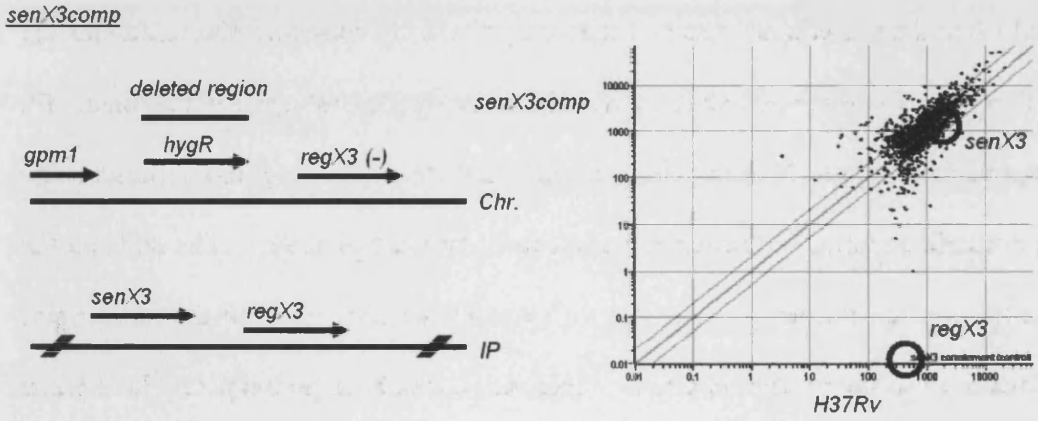
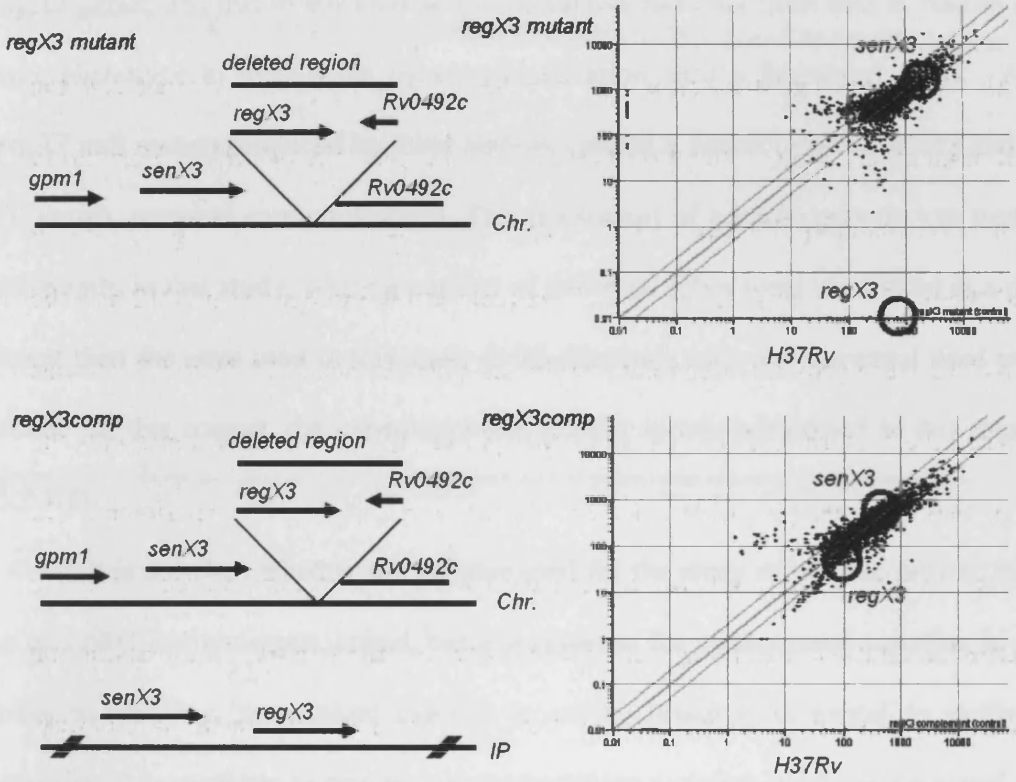

senX3comp
deleted region
gpm1
hygR
regX3 (-)
Chr.
senX3
regX3
IP
senX3comp
senX3
regX3
H37Rv

Figure 3.20: Genotypes and transcriptional profiles of the *regX3* null mutant and the *regX3* null mutant carrying a functional allele on the plasmid pDH12. Microarray comparisons were made by RNA-RNA hybridisation (see 2.5.4.1) between the *regX3* null mutant and the wild-type, and the *regX3* null mutant with the integrated pDH12 and the wild-type. Cells were grown under microaerobic conditions (see 2.2.1.2) and RNA was extracted when the bacteria had reached stationary phase. Each scatter plot represents the results from two dye-swapped slides, derived from a single biological experiment. IP, integrated plasmid; hygR, hygromycin resistance cassette. The scatter plot represents the relative normalised ratios of transcript abundance between two strains. The units on the axes are the median normalised intensity of the spot for each gene on the microarray, after subtraction of the median normalised intensity of the background pixels in the area adjacent to the spot.


regX3 mutant
deleted region
regX3
Rv0492c
gpm1
senX3
Rv0492c
Chr.
regX3 mutant
senX3
regX3
H37Rv
regX3comp
deleted region
regX3
Rv0492c
gpm1
senX3
Rv0492c
Chr.
senX3
regX3
IP
regX3comp
senX3
regX3
H37Rv

This contrasts sharply with the findings in Parish *et al.* (2003) where two different, abrogated forms of growth were seen in an isogenic *regX3* null mutant. There are doubts about the validity of the strain used in that study (see 1.3.2.1), namely that the *regX3* null mutant also lacks some of the 228 bp intergenic region between the *senX3* and *regX3* genes, and that to our knowledge the authors have not been able to restore a wild-type phenotype to their strain by complementation with a functional allele. Another *regX3* null mutant produced by these authors created a fusion between *senX3* and *regX3* (T Parish, personal communication). The assessment of aerobic growth was performed differently in that study, where a number of universal tubes were inoculated at a dilution lower than the ones used in this study (1/10 dilution), with one universal used per time point. In this respect, the experiment was similar to one performed in this thesis (see 3.2.1.2).

It is not clear whether the cultures used for the study of aerobic growth in Parish *et al.* (2003) were oxygen limited, but it is common for a substantial variation in growth rates in standing, non-shaken cultures to occur, hence it is useful to perform the experiment in triplicate to prevent culture-to-culture variation masking the actual growth rate of the bacilli. The assessment by Parish *et al.* (2003) of the aerobic growth rate of their *regX3* null mutant was performed twice, with different growth patterns each time. The variation in growth between the two experiments is likely to be an artefact of their growth system, as will be the pattern of growth within an experiment. In addition, the strain used in that study has been kindly provided to the author of this thesis, who found that it would not grow *in vitro* from a low dilution, unlike the *regX3* and *senX3* null mutants used in the majority of this study. It can be suggested, therefore, that normal

growth *in vitro*, as shown in 3.2.1.1, represents the true phenotype of a *senX3* or *regX3* null mutant of *M. tuberculosis*.

3.3.1.2 Microaerobic growth

It has been hypothesised that the SenX3/RegX3 system might be acting like ArcB/A of *E. coli*, in which situation null mutants for these genes would be expected to lose viability during oxygen-limited stationary phase (see 1.2.1.1 and Dukan & Nystrom (1998)). This was not the case here in 3.2.1.2, suggesting that in this assay *senX3* and *regX3* does not appear to play a role vital to the growth or survival of this organism. In the current study it was impossible to measure exact oxygen concentrations in the media due to biosafety restrictions in our laboratory. However, collaborative work is addressing this issue (see 7.5).

It is possible to draw a number of suppositions from the data in 3.2.1.2:

i) the bacilli are entering stationary phase due to metabolism geared to microaerobic growth, since nutrient limitation does not occur in this media until a much higher OD_{600} (see 3.2.1.1)

ii) the model for microaerobic growth used here is untested for efficacy in the ArcA/B mutants of *E. coli* and may not represent the proper system for development of this phenotype

iii) it is not possible to comment on the viability of the bacilli in stationary phase when assaying growth by OD_{600} measurements; only differences in the time between the limitation of oxygen beginning and the cessation of bacterial replication will be apparent.

iv) the SenX3/RegX3 two component system may not be acting like ArcA/B of *E. coli*

In 4.2.2.1 the same mutants are examined in a more controlled microaerobic environment and the *senX3* and *regX3* null mutants show no loss of viability in what is assumed to be oxygen-induced stationary phase, negating point (iii) above. Therefore we must assume that either the model is not a good one, or that the SenX3/RegX3 system is not acting like the ArcA/B system of *E. coli*. It is possible that the time-frame of the experiment in 3.2.1.2 was too short, and that a phenotype may have become apparent at later time points.

It may also be the case here that the strains are entering a stage of arithmetic growth after 336 hours, although viability data would be required to show this.

3.3.1.3 Nutrient starvation

Another method of testing the phenotype of potential ArcA/B- like mutants is by causing the bacilli to enter stationary phase by starving them of carbon sources. Like the experiments above, if the *senX3* and *regX3* null mutants are acting like ArcA/B, then prolonged incubation of the strains without a carbon source should result in a gradual loss of viability when compared to the wild-type (see 1.2.1.1 and Dukan & Nystrom (1998, 1999)). This was not apparent with the *senX3* and *regX3* null mutant strains. Since the experiment is a simple one, and the cells were washed thoroughly, one can only conclude that either the time-frame of the experiment was insufficient to see a phenotypic difference between the mutants and the wild-type, or that SenX3 and RegX3 are not acting in a similar manner to ArcA/ArcB of *E. coli*.

3.3.1.4 Growth of mutants at pH 5.5

Another environment experienced by *M. tuberculosis in vivo* is the lowering of pH in the phagocytic vacuole. Rather than exposing exponential-phase cultures to acidic conditions, it was felt that observing the growth of the mutant and wild-type strains would be a more sensitive assay for the strains' ability to resist acid stress. Previous work on an OmpA mutant used a similar technique (Raynaud *et al.* 2002).

There was no difference in growth between the *senX3* and *regX3* null mutant strains and the wild-type, indicating that *senX3* and *regX3* do not play a role in the adaptation of *M. tuberculosis* to acid stress. Curiously, all cultures displayed a biphasic growth curve that could represent a period of adjustment of the bacilli. Other bacteria can regulate the pH of acid growth media to a level more suitable for survival, however pH indicator strips used at the last time-point showed that the acidity of the media had not changed appreciably. It is likely that the lowered pH is making some of the metabolic pathways of the bacterium less viable, requiring the adaptation period to alter the bacterium's metabolism to one more favourable. Future work may wish to repeat this experiment in glycerol-free media.

The results of this study differed from those of Raynaud *et al.* (2002) in the occurrence of the biphasic growth and the low OD_{600} at which the bacteria enter stationary phase. In Raynaud *et al.* (2002) the bacteria underwent a long lag phase, then grew as normal to a high OD_{600}. This may be due to the use of an H37Rv substrain, "1424" in that study, or different methods in preparing the strains. However for the

purposes of this study it is sufficient only to note that there is no difference in the growth of the *senX3* and *regX3* mutants and the wild-type *M. tuberculosis*.

3.3.1.5 Toluidine blue assay

ArcA mutants of *E. coli* are exquisitely sensitive to the dye toluidine blue, although the mechanism for this sensitivity is not known. Although there is little reason to believe that toluidine blue would cause an equally dramatic phenotype in a strain of *M. tuberculosis* lacking ArcA-like function, it would represent a very fast and simple method of screening for complementation of the *senX3* and *regX3* mutants if this approach proved successful.

Growth of the *senX3* and *regX3* null mutants and the wild-type all stopped at a concentration of 100μg/ml toluidine blue when it was present in both solid and liquid media. This indicates that SenX3 and RegX3 are either not acting like the ArcA/B system, or that the ArcA/B orthologous system in *M. tuberculosis* is not sensitive to disruption by toluidine blue.

3.3.2 Stress studies

3.3.2.2 Superoxide stress

At two hours, but not 24 hours exposure to paraquat, the *regX3* null mutant showed a significant drop in viability when compared to the wild-type. Viable counts were taken and are a reliable indicator of bacterial killing. At 24 hours there were large losses of viability in all three strains, confirming that the bacilli are damaged in this system.

Counterintuitively, the *senX3* and *regX3* null mutants appear to be more, rather than less, resistant to killing by superoxide at 24 hours, with the *regX3* null mutant more so than the *senX3* null mutant. It is possible that *senX3* and *regX3* act as a negative regulator of genes involved in the detoxification of superoxide and ROI, and that these responsive genes are now constitutively expressed resulting in a resistant phenotype. There is no reason to believe that this is the case, especially when one considers work later in this chapter (3.2.3.2) demonstrating reduced virulence of the *senX3* and *regX3* null mutant strains during infection. Furthermore, there should be no reason for the *regX3* and *senX3* null mutants to differ in their responses to superoxide stress, unless the few base pairs missing from the 3' end of *Rv0492c* in the *regX3* null mutant are involved in producing this phenotype. Complementation of the mutants *in vivo* (see later) would suggest that this is not the case. Since the mutants do not show increased sensitivity *in vitro* to superoxide stress, it is difficult to envisage that they have a role in the survival of the bacterium to such stress generated *in vivo*.

3.3.2.3 Nitrosative stress

There was no significant difference in the ability of the *senX3* or *regX3* null mutants to resist killing by nitrosative stress in either aerobic conditions or microaerobic conditions. This suggests that SenX3 and RegX3 are not involved in the response to nitrosative stress, and also detracts from the hypothesis that the SenX3/RegX3 system could be acting like ArcA/B from *E. coli*, whereby the mutants would generate excess oxidative stress under microaerobic conditions and therefore be more susceptible to

killing by nitrosative stress since the ROI could react with the RNI to form the highly damaging peroxynitrite.

3.3.2.4 Organic hydroperoxide stress

Organic hydroperoxides represent some of the later stage products in the breakdown of ROI in the cell. The *regX3* null mutant alone appeared to be more resistant to killing by bactericidal levels of organic hydroperoxide than either the wild-type or the *senX3* null mutant. This discrepancy between the *senX3* and *regX3* null mutants can perhaps be tentatively explained by the slightly slower growth rate of the *regX3* null mutant, perhaps allowing it more time to repair damage. In conclusion, it is unlikely that SenX3 or RegX3 play a role in resistance to organic hydroperoxide stress.

3.3.3 *In vivo* characterisation of the *senX3* and *regX3* null mutants

3.3.3.1 Growth of the *senX3* and *regX3* null mutants in a mouse model of infection

Previous work in our laboratory had shown that the *regX3* and *senX3* null mutants were attenuated in mice (Rickman 2002) when compared to the wild-type. The dynamics of the attenuation followed a pattern described as 'growth *in vivo*' attenuation in Hingley-Wilson *et al.* (2003). To confirm the attenuation of the *senX3* null mutant, and to prove that the attenuation of both strains was due solely to the absence of *senX3* or *regX3*, functional alleles were inserted into the *attP* site on the chromosome of each of the mutants in an attempt to restore wild-type virulence. The same region was used to complement both strains.

The results in mice showed that the growth of the *senX3* and *regX3* null mutant strains was attenuated by ~10-fold compared with the wild type, but the cells still persist. The mutant strains carrying the complementing allele were restored, at least partially, to wild-type growth. Of note is the growth of the complementing strain of the *senX3* null mutant, which does not reach wild-type levels in the lungs until 65 days and in the spleens until 140 days. The late complementation of the *senX3* null mutant is examined further below.

Overall the data show that SenX3 and RegX3 are essential for growth of *M. tuberculosis* in mice.

Genotypically, the *senX3* null strain carrying the complementing allele has two copies of *regX3*, one chromosomal (*regX3*^{c}) and one on the integrated pDH12 (*regX3*^{p}). We already suspected that *regX3*^{c} might have been inactivated due to a polar effect of the hygromycin resistance marker present upstream of it. This led us to think that the *senX3* null mutant carrying the complementing allele was deficient in expression of *regX3*^{p}, which may have acquired a deletion or mutation that prevented its expression. To test this, *regX3 and part of senX3* was amplified from a plate culture of the *senX3* null mutant carrying the complementing allele recovered from the mice and two co-transformants (see 3.2.3.4). All products were of the expected size, and when sequenced, products from all three *senX3* null mutant strains carrying the complementing plasmid were identical to the published sequence of *M. tuberculosis*. Therefore, deletions or mutation was not evident in the complementing *regX3*^{p} gene in the *senX3* complementation strain.

Microarray analysis of the complementing strains used in the murine infection (3.2.3.5) enabled the transcriptotypes of the strains to be examined (3.2.3.5). This

confirmed that the complementing allele in the *senX3* null mutant is unable to produce levels of *regX3* transcript comparable to that of the wild-type (Figure 3.19), whereas the same allele in the *regX3* null mutant restores wild-type levels of the transcript (Figure 3.20).

There seem to be two possible explanations for this partial complementation of the *senX3* null mutant. Firstly, there may have been some error in the experiment that caused either the *senX3* null mutant to grow aberrantly or some of the mice to behave in an immunologically different way, although this is unlikely. Secondly, RegX3 may bind to the promoter of *senX3* and regulate its expression (Himpens *et al.* 2000) and the presence of atypical copy numbers of *regX3* and the *senX3* promoter may by causing an imbalance in the regulation of transcription of the two genes, with a subsequent downstream effect on virulence.

It would be difficult to analyse this discrepancy further because the transcriptional analysis of the complement situations by microarray was only performed once, and some of the values for the *senX3* and *regX3* gene spot intensities on the microarrays were low and possibly unreliable. Furthermore, Figure 3.16 shows that the growth of the *senX3* null mutant is inconsistant *in vivo*, and the true cause of this may not be detectable *in vitro* due to differences in the model of growth.

Our concern was that the lack of a reliable complementation model for the *senX3* null mutant would make more detailed analysis of the role of SenX3 in the virulence of *M. tuberculosis* difficult. It was decided to examine the growth of the strains in a macrophage model of infection, due to the shorter experiment time and fewer mice used. The benefits and drawbacks of a macrophage model of infection are discussed in 1.4.6.

3.3.3.2 Growth of the *senX3* and *regX3* null mutants in bone-marrow derived macrophages

The growth of the wild-type, the *senX3* and *regX3* null mutants and their respective complementing strains confirmed the attenuation of the mutants and the restoration to wild-type growth of the complementation strains in this model. These data suggest that RegX3, and probably SenX3, are essential for growth of *M. tuberculosis* in primary murine macrophages. This allows a number of conclusions to be drawn regarding the nature of the stimulus for the SenX3/RegX3 and the mechanism by which the *senX3* and *regX3* null mutants are attenuated in macrophages:

i) the stimulus/attenuation mechanisms are not part of the acquired immune response

ii) the SenX3/RegX3 system is not responding to granuloma-like conditions in this model, since granulomas do not form in cell layers.

Although the mechanism for attenuation in macrophages discounts granuloma-like conditions such as long-term nutrient starvation and oxygen deprivation, this data does not preclude it. It does, however, suggest that the *senX3* and *regX3* null mutants are sensitive to control by the host in the early stages of infection when they encounter the innate immune response of the phagocytes. These would typically be the stresses discussed in this chapter, although it is possible that macrophages have antibacterial mechanisms that we are unaware of and it is one of these that accounts for the attenuation of the mutants. Another possibility is that the SenX3/RegX3 system is responding to a combination of stresses, where the application of one stress alone is not enough to

produce a phenotype, and that this combination has not been reproduced in the *in vitro* conditions described in these experiments.

3.3.4 Summary

The data discussed in this chapter highlight the importance of the SenX3/RegX3 system in the growth of *M. tuberculosis*, and have begun to elucidate the stresses or mechanisms to which strains of *M. tuberculosis* lacking this system are sensitive. Such information would allow hypotheses to be formed of the role of *senX3* and *regX3* within the bacillus. Unfortunately we have so far been unable to single out one particular agent that is convincingly responsible for the attenuation of the *senX3* and reg*X*3 null mutants, for reasons described above. It may be possible to further examine these promising results by testing a wider range of superoxide-producing agents in the same system. However, the apparent subtleties of the sensitivities *in vitro* may be consistent with a hypothesis that multiple signals are required for the full activity of the SenX3-RegX3 two component signal transduction system.

There is no doubt, from the mouse infection experiments described, that the SenX3/RegX3 two component signal transduction system is absolutely required for the normal, wild-type growth of *M. tuberculosis* in this infection model. The complementation experiments prove that mutation of *regX3*, and possibly also *senX3* gives rise to the observed phenotype. By analogy with orthologous genes in other bacteria, viz. *E. coli*, it was hypothesised that the SenX3/RegX3 system was involved in the response of the bacterium to oxidative and/or nitrosative stress. However, no convincing *in vitro* stress conditions have been found to which the *senX3* or *regX3* null

mutants were more sensitive than the wild-type, and thus the reason for the reduced growth of the mutants *in vivo* remains unknown.

4 Transcriptional microarray studies of mutants in the SenX3/RegX3 two-component signal transduction system.

4.1 INTRODUCTION

4.1.1 Use of microarrays to study bacterial genetic regulation

Microarray is a recently established technology that allows comparative analysis of the genomes or transcriptomes of all genes in an organism simultaneously. This immensely powerful tool relies on unique probes to each gene being synthesised, either by PCR or as synthetic oligonucleotides, for which the genome sequence of the organism needs to be known. The *M. tuberculosis* genome sequence was completed in 1998 (Cole *et al.* 1998) and a cDNA microarray platform was developed at St. George's Hospital Medical School, London (Hinds *et al.* 2001). This resource has since been available to researchers and is often a key step in the elucidation of gene function by comparing different strains of *M. tuberculosis* (Stewart *et al.* 2002).

4.1.2 Comparison of mutants and wild-type under defined growth conditions

Two approaches have proved to be particularly useful in studying mutants of *M. tuberculosis*. Firstly, a null mutant of the gene under study can be compared to the wild-type and any differences in the apparent quantities of transcripts between the two strains marked as a fold change in regulation. This approach is only effective if the gene would ordinarily interact with other components of the cell under the conditions in which the two strains are compared. For example, if a regulatory protein is active only during exposure to UV radiation, then one would only expect to see a difference between the transcriptomes of a null mutant in that gene and the wild-type if they had been exposed to UV radiation. It is possible to hypothesise that superoxide may fill this role if indeed the SenX3-RegX3 two component signal transduction system is sensing superoxide stress.

Alternatively the lack of a constitutively active regulator would show a transcriptome difference to the wild-type under all conditions tested. Therefore to study regulatory genes that are only active in specific circumstances, one must know the signal to activate the system. This approach was used to identify the operon of DosR, since it was already known that the DosRS system responded to hypoxia (Park *et al.* 2003).

4.1.3 Overexpression of regulatory genes

A second method to identify signalling system regulons is to overexpress the regulator at levels that overcome the usual unresponsiveness of the system in the absence of the specific signal, and compare this overexpressing strain to its parent, either the wild-type or a null mutant. The results gained by this method depend much on the mode of action of the regulator, whether it is a repressor or an activator. For instance, if a regulator prevents transcription of its regulon, then increased amounts of the repressor are not likely to decrease the expression of those genes any further. In two-component systems a further complication is the phosphorylation state of the regulator – is phosphorylation of the regulator the rate-limiting step in signal transduction, rather than rate of expression? Another issue may be the potential toxic effects of overexpressing the regulator. However, these experiments are not necessarily performed on well-characterised regulators, so in practice one can only really attempt to express the regulator at different levels, using different expression systems, to minimise any potential problems with toxicity.

4.1.4 Caveats with gene list comparisons

Both of these approaches were attempted with the SenX3-RegX3 two component signal transduction system in this study. Although the two methods may produce different results, it is important to consider that the stimulus in the different experiments differs greatly and this may affect the regulatory response of the system. Furthermore, it is difficult to emulate growth conditions in other laboratories, so differences to published data may often be due to the use of different growth media and conditions.

4.2 RESULTS

4.2.1 Transcriptional microarray analysis after superoxide stress

The wild-type and *regX3* and *senX3* null mutants were grown in four rolling bottles per strain to early exponential phase when paraquat was added to half of the cultures. Incubation was continued for a further two hours before RNA was extracted from both the stressed and non-stressed cultures (see 2.2.1.5). Paraquat was used at a 25mM concentration to reduce the bactericidal effect seen with paraquat used at 50mM concentration (see 3.2.2.1).

Microarray hybridisation directly compared RNA from the wild-type *M. tuberculosis* and RNA extracted from a null mutant, enabling small differences in the relative abundance of transcript between the two strains to be measured (see 2.5.4.1). Analysis of the microarray results failed to show any reproducible differences between the wild-type and *senX3* and *regX3* null mutants after superoxide stress

4.2.2 Transcriptional microarray analysis after growth in oxygen-limited conditions

4.2.2.1 The microaerobic model

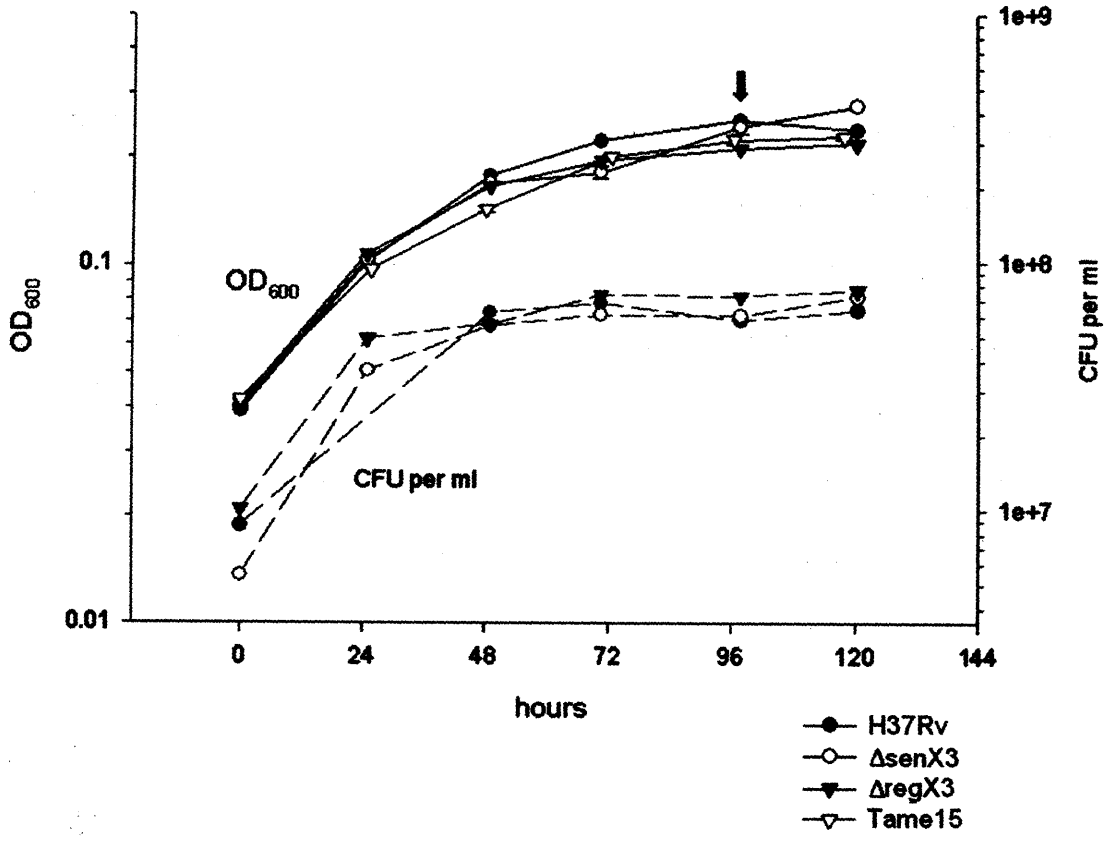
To investigate whether there were any difference apparent between the transcriptomes of the *senX3* and *regX3* null mutants and the wild-type when grown under oxygen-limited conditions, the wild-type, *senX3* and *regX3* null mutants were grown in standing culture as described in 2.2.1.2, although at a lower (1/500) dilution than the one stated in that section. This preliminary data showed that there was a difference between the transcriptomes of the mutants and the wild-type, suggesting that the SenX3-RegX3 two component signal transduction system was active under these conditions (data not shown). These data were not, however, reproducible.

To examine the above responses in a more controlled fashion, it was decided to use a higher dilution (1/10) of the wild-type, *senX3* and *regX3* null mutants and the Tame15 mutant inoculated from early exponential-phase rolling cultures to make a large stock culture that was subsequently aliquoted into flasks and grown in an identical manner to that described in 2.2.1.2.

Sufficient cultures were inoculated to allow one culture per strain to be used daily so that the OD_{600} and viable counts of the strains could be followed and RNA extracted as soon as the OD_{600} similar to that seen in the extractions for 4.2.2.1 was reached, representing early stationary phase (figure 4.1). The strains showed uniform growth and the OD_{600} at which RNA was extracted were more consistent than the OD_{600} measurements taken from the strains in 4.2.2.1.

The relative abundance of transcript levels between different cultures were compared indirectly on microarray slides using DNA-RNA hybridisation (see 2.5.4.2). Pairwise analysis of the wild-type and each of the mutants yielded lists of genes whose

Figure 4.1: Optical densities and viable counts of mutants in the SenX3-RegX3 system when grown from a 1/10 dilution of early exponential-phase strains in the standing culture microaerobic model as described in 2.2.1.2. Solid line, optical densities, left axis; dashed line, colony forming units per millilitre of culture, right axis. The growth of Tame15 was done on a separate occasion, and viable counts were not measured. The arrow represents the time at which RNA extractions were carried out on these cultures or cultures that had been treated in an identical manner.


1e+9
1e+8
1e+7
0.1
0.01
OD600
CFU per ml
OD600
CFU per ml
0
24
48
72
96
120
144
hours
H37Rv
ΔsenX3
ΔregX3
Tame15

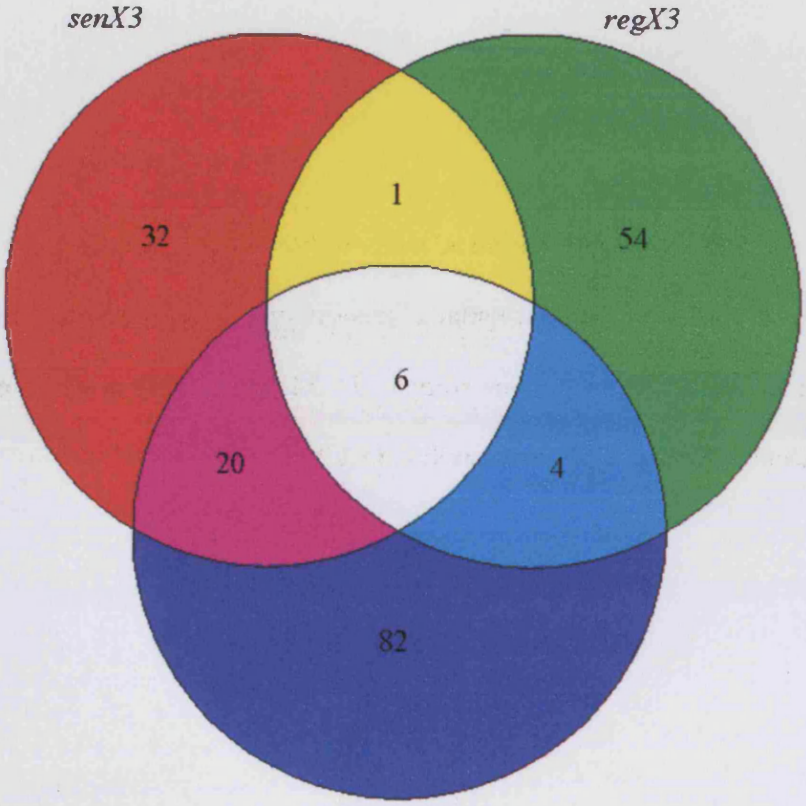
transcript levels differed significantly ($p < 0.05$) between the groups (see 2.5.8.3). The full gene lists can be seen in Appendix 1. The transcript levels of 59 genes were significantly different in the *senX3* null mutant compared to the wild-type, and the transcript levels of 65 and 112 genes were significantly different in the *regX3* null mutant and the Tame 15 mutant respectively. Of the genes identified as being differentially expressed in the three mutant strains, one gene was common to the *senX3* and *regX3* null mutants only, 20 were common between the *senX3* null mutant and the Tame15 mutant only, four were common between the *regX3* null mutant and the Tame15 mutant only, and six genes were common to all three strains (Figure 4.2 and Table 4.1).

4.2.3 Overexpression of *regX3* in a *regX3* null background

Strains of *M. tuberculosis* were constructed in a *regX3* null background that contained an extrachromosomal copy of *regX3* driven by either an Hsp60 (Stover *et al.* 1991) or acetamidase (Triccas *et al.* 1998) promoter. Identical strains containing the empty vector without the *regX3* coding sequence were also constructed as negative controls. Strains were grown aerobically in rolling bottles under antibiotic selection to early exponential phase before RNA extraction, after which DNA-RNA hybridisation microarrays were carried out (see 2.5.4.2).

Cursory analysis of the arrays using the GENEPIX 5.0 program showed that the strain ΔregX3/P_{ace}-*regX3*$^{+}$ did not produce any measurable levels of *regX3* transcript, consistent with other reports of problems with this promoter in *M. tuberculosis* (S. Kendall, personal communication). However ΔregX3/P_{hsp60}-*regX3*$^{+}$ produced high levels of *regX3* transcript, with *regX3* as the most highly expressed gene when the data was

Figure 4.2: Venn diagram showing the incidence of homology between gene lists generated by assessment of differentially regulated genes in mutants in the SenX3-RegX3 two component signal transduction system compared with the wild-type when grown under microaerobic conditions at 37°C, as described in 2.2.1.2.



Tame15

Table 4.1: Genes that appear to be differentially regulated in more than one SenX3-RegX3 two component signal transduction system mutant after growth in microaerobic conditions. The colours correspond to those seen on the Venn diagram in Figure 4.2.

senX3 and *regX3* null mutants

Rv1632c	Rv1632c

senX3 and *regX3* null mutants and Tame15

Rv2706c	Rv2706c
Rv1736c	narX
Rv2658c	Rv2658c
Rv1952	Rv1952
Rv3301c	phoY1
Rv0637	Rv0637

senX3 null mutant and Tame15

Rv0883c	Rv0883c
Rv0343	Rv0343
Rv1988	Rv1988
Rv3853	menG
Rv1398c	Rv1398c
Rv0491	regX3
Rv3854c	Rv3854c
Rv0932c	pstS
Rv2197c	Rv2197c
Rv3089	fadD13
Rv2820c	Rv2820c
Rv0297	PE PGRS
Rv2461c	clpP
Rv0483	Rv0483
Rv1955	Rv1955
Rv3680	Rv3680
Rv0174	Rv0174
Rv1172c	PE
Rv2116	lppK
Rv2007c	fdxA

regX3 null mutant and Tame15

Rv0497	Rv0497
Rv0715	rplX
Rv2935	ppsE
Rv1601	hisB

represented on the GENEPIX scatter plot (data not shown). ΔregX3/P_{hsp60}-*regX3*$^-$ did not appear to produce any *regX3* transcript when analysed using GENEPIX.

GENESPRING analysis showed 199 genes whose expression levels were significantly different in the ΔregX3/P_{hsp60}-*regX3*$^+$ strain compared to the ΔregX3/P_{Hsp60}-*regX3*$^-$ strain. Of these, 21 sets of adjacent genes were apparent, suggesting that they may be co-transcribed. These can be seen marked in bold in Appendix 2. 13 genes up or downregulated in the ΔregX3/P_{hsp60}-*regX3*$^+$ strain were also present in a published list of genes up- and down-regulated in Tame15 (Parish *et al.* 2003). These can be seen marked with an asterisk in appendix 2.

4.2.3.1 MEME motif analysis of potential promoter regions of genes identified to be differentially expressed between a ΔregX3/P_{hsp60}-*regX3*$^+$ strain and a ΔregX3/P_{hsp60}-*regX3*$^-$ strain of *M. tuberculosis*

The DNA regions upstream of all 199 genes identified to be differentially expressed between a ΔregX3/P_{hsp60}-*regX3*$^+$ strain and a ΔregX3/P_{hsp60}-*regX3*$^-$ strain of *M. tuberculosis* (see Appendix 2) were analysed by visual inspection using the ARTEMIS program (Rutherford *et al.* 2000). Genes lacking a stop codon of a gene transcribed in the same direction within 100 bp upstream of the potential start site were considered to contain potential RegX3 binding sites within their upstream region. The sequence of the upstream region for these genes was recovered from ARTEMIS, between either the feature immediately upstream of the differentially regulated gene if that distance was less than 200 bp, or a region 200 bp upstream of the start site of the differentially regulated gene. 113 of the 199 genes matched the criteria, and the appropriate sequence regions as

described above were then submitted to the MEME server (Bailey & Elkan 1994) for motif analysis.

No similar sequences of nucleotides, making 'motifs', were found that occurred in more than ten of the 113 sequences.

4.3 DISCUSSION

4.3.1 Superoxide stress

A lack of differentially regulated genes between the wild-type and the *senX3* and *regX3* null mutants after superoxide stress suggests that the SenX3-RegX3 two component signal transduction system is not responding to exogenous superoxide stress. This conclusion is backed up by the viability data in 3.2.2, and although the superoxide microarrays were only performed in duplicate it was felt on the strength of this not to continue with the third replicate.

4.3.2 Growth in a standing culture model of microaerobic adaptation

A number of genes appeared to be up- or down-regulated in the *senX3* null mutant after growth under microaerobic conditions, suggesting that the SenX3-RegX3 system is responding in some way to the environment in the microaerobic culture. The preliminary result was important in that it suggested a possible stimulus for SenX3, although the variation in growth rates in this high-dilution microaerobic culture may be to blame for an inability to repeat these experiments enough to gain statistically significant results.

4.3.3 Further microaerobic analysis

In the absence of statistically reliable results from the preliminary model, refinement of the microaerobic model and acquisition of the Tame15 mutant (Parish *et al.* 2003) allowed us to further examine and compare the transcriptional responses of the different mutants in a more reliable system. Following the growth curve of the strains enabled the extraction of RNA at the point where the cells have just entered stationary phase, presumably through the limitation of oxygen since nutrients are generally not limiting in an *M. tuberculosis* culture until a much higher OD_{600} is reached. Of note is the apparent increase in cell density after 48 hours while the numbers of viable bacteria remain constant. This could be explained by the thickening of the bacterial cell wall, which Wayne & Hayes (1996) referred to as non-replicating persistence. Cell wall thickening under oxygen-limited conditions is associated with the upregulation and co-localisation of HspX with the cell wall (Cunningham & Spreadbury 1998), and may be controlled by DosR (Park *et al.* 2003).

Comparison of differentially regulated genes between each of the mutants and the wild-type showed that while there was a small degree of overlap in the gene lists between the *senX3* and *regX3* null mutants, the greatest similarity between the strains under these condition was between the *senX3* null mutant and Tame15. This suggests that the Tame15 mutant may be acting more like the *senX3* null mutant than the *regX3* null mutant in this model.

Interestingly, *fdxA*, encoding a ferridoxin, can be seen to be differentially regulated in both the *senX3* null mutant and Tame15. *narX* is one of the few genes that appears to be differentially regulated in all three mutants, and is also reported to be induced by hypoxia (Sherman *et al.* 2001) as part of the DosR response. Also of interest

is the upregulation in the *regX3* null mutant of the mycobactin synthesis genes (*mbtB*, *C*, *D* and *G*) that code for siderophore biosynthesis enzymes, and *mce3*, which is involved in macrophage cell entry.

Overall the data produced by this method shows a number of potentially interesting genes that may be regulated by the SenX3-RegX3 two component signal transduction system or some other component that interacts with it. The appearance of apparently co-transcribed genes up- or down-regulated by similar fold changes suggests that the microarray method and analysis are working well, and these data provide a solid foundation for the transcriptional study of the mutants in other, more defined conditions.

4.3.4 Overexpression of RegX3 in a *regX3* null mutant

The high levels of *regX3* transcript in the ΔregX3/P_{hsp60}-*regX3*$^{+}$ strain compared to the ΔregX3/P_{hsp60}-*regX3*$^{-}$ strain appeared to overcome the need for the correct stimulus of the SenX3-RegX3 two component signal transduction system in order to initiate transcription of genes either controlled by RegX3 or another factor that is regulated by RegX3. This is in contrast to microarray studies comparing the *regX3* null mutant to the wild-type in aerobic rolling culture (Rickman 2002) and after superoxide stress (this study) where no significant difference between the strains was found, presumably because the SenX3-RegX3 two component signal transduction system was not active as a result of lack of the proper signal for SenX3.

The presence of 20 sets of apparently co-transcribed genes again suggests that the gene list generated by overexpression of RegX3 is not entirely random. Five other genes

also appear after overexpression of RegX3 that are differentially regulated in the *regX3* null mutant after microaerobic growth (see Appendices 1 and 2).

4.3.4.1 Comparison of genes upregulated after overexpression of *regX3* to other gene lists

Many of the genes that are differentially regulated in the *regX3* overexpressing strain appear in published microarray data. 13 of the genes differentially regulated in the *regX3* overexpressing strain can be seen in the published list of potentially RegX3-regulated genes in the Tame15 mutant grown in rolling cultures in Middlebrook broth, representing 13.3% of the published list. This is not altogether surprising since the comparison in that paper is between a *regX3* null mutant and the wild-type with normal levels of *regX3*. Of greater interest is appearance of many of the differentially-regulated genes in the RegX3 overexpressing strain in Betts *et al.* (2002), where transcriptional microarrays of the wild-type *M. tuberculosis* were carried out after 96 hours of nutrient starvation. 14 genes that are differentially regulated in the RegX3 overexpressing strain appear in their list of 114 genes, representing 12.3% of the genes discussed in that paper. Although not a great proportion, it is the key and representative genes mentioned in the discussion of Betts *et al.* (2002) that also appear in Appendix 2 here. Such genes include regulatory genes (*pknD*, *ideR*), genes involved in energy metabolism (ATP synthase genes *atpB, F, H*; *nuoC*; *pdhC*), polyketide synthesis (*ppsD*, *pks18*), translation apparatus (*rplV*), antibiotics (Rv2036, *lat*), and lipid biosynthesis (*desA3*). *ald* is also in this published list, and was seen in other microarray results in this study and used for RegX3-DNA binding experiments in 5.2.2.5.

Taken together, this supports a hypothesis that overexpression of RegX3 can mimic some of the conditions that the SenX3-RegX3 two component signal transduction system may be responding to some *in vitro* test condition, and may represent a more representative view of the RegX3 regulon than that obtained by the *ad hoc* testing achieved by stressing rolling cultures of the bacterium.

4.3.4.2 Genes differentially regulated between a ΔregX3/P_{hsp60}-*regX3*$^{+}$ strain and a ΔregX3/P_{hsp60}-*regX3*$^{-}$ strain of *M. tuberculosis*

4.3.4.2.1 Regulatory genes

The most upregulated gene in the strain overexpressing RegX3 is *regX3* itself, whose normalised fold upregulation is ~340 000 with a p value of 0.00000002. SenX3 is upregulated two-fold, suggesting that RegX3 interacts with the *senX3* promoter and increases transcription. This result supports the interaction between RegX3 and the promoter for *senX3* suggested in Himpens *et al.* (2000) and elsewhere in this study. The overexpression of *regX3* has not caused an equal upregulation in se*n*X3 transcript; however since the interaction between the SenX3 and RegX3 proteins occurs at the phosphorylation level (Himpens *et al.* 2000) it is likely that transcriptional control of the SenX3-RegX3 two component signal transduction system is not the major factor controlling this system's activity.

Other upregulated regulatory genes in the *regX3* overexpressing strain that have been shown to be important in *M. tuberculosis* are *pknD,* encoding protein kinase D, a eukaryotic-like serine-threonine protein kinase (Peirs *et al.* 1997), and *ideR*, encoding an iron-dependant repressor protein.

PknD is thought to regulate phosphate transport since it is likely to be co-transcribed with Rv0932c that encodes a protein involved in a phosphate-uptake system (Av-Gay & Everett 2000). PknD phosphorylates the FHA-A domain of Rv1747 (Grundner *et al.* 2005), which contains an ABC transporter domain. A null mutant of Rv1747 is attenuated in three models of infection, and it can therefore be assumed that the protein encoded by Rv1747 plays a key role in the *M. tuberculosis* infection process (Curry *et al.* 2005).

ideR encodes an iron-dependant repressor protein of the DtxR (diphtheria toxin repressor) family that is required for siderophore repression in iron-abundant environments (Rodriguez *et al.* 2002) and may be important for survival in macrophages (Gold *et al.* 2001). Gold *et al.* (2001) based this latter observation on the upregulation of, amongst others, the *mbtB* gene in macrophage-like THP-1 cell infection (De Voss *et al.* 2000) and in iron-limited culture. *mbtB* was seen to be downregulated in the *regX3* null mutant after microaerobic growth (see Appendix 1) and it is conceivable that the SenX3-RegX3 two component signal transduction system is acting as a level of control above that of *ideR*.

ptpA is downregulated 10-fold in the *regX3* overexpressing strain. *ptpA* encodes a low-molecular weight secreted tyrosine phosphatase that is upregulated in *M. bovis* BCG when it enters stationary phase or infects primary human monocytes (Cowley *et al.* 2002), and is potentially acting as a virulence factor affecting host proteins. Potential regulation of this gene by the SenX3-RegX3 two component signal transduction system further suggests a role in controlling the molecular pathogenesis of *M. tuberculosis*.

4.3.4.2.2 Genes involved in energy metabolism

A number of genes involved in energy metabolism are differentially regulated in the *regX3* overexpressing strain. *atpB*, *F* and *H* encode parts of the ATP synthase which generates ATP from ADP in the presence of a proton gradient and are upregulated in the *regX3* overexpressing strain. *nuoC* encodes NADH dehydrogenase I, which forms part of the NADH dehydrogenase operon that is the major aerobic electron transport chain in *M. tuberculosis*. This gene was upregulated 10-fold in the *regX3* overexpressing strain. In contrast to Betts *et al.* (2002), both the ATP synthase genes and part of the NADH dehydrogenase operon were seen to be upregulated in the *regX3* overexpressing strain, suggesting that RegX3, or another component that interacts with RegX3, acts as a transcriptional activator of these genes.

pdhC, encoding dihydrolipoamide acetyltransferase was also upregulated. PdhC catalyses the production of acetyl-CoA and carbon dioxide from pyruvate and is critical in regulating carbon flux into the TCA cycle.

cydB, seen to be upregulated, encodes a cytochrome D ubiquinol oxidase that also plays a role in the cellular respiration. Taken together, it could be that RegX3 can act as a transcriptional activator, allowing the maximal use of the aerobic respiratory mechanisms while sufficient oxygen is present.

4.3.4.2.3 Genes involved in polyketide synthesis

ppsD, also known as *pksE*, encodes part of the phthiocerol dimycocerosate (PDIM) locus (Azad *et al.* 1997) whose synthesis and transport are involved in *M. tuberculosis* pathogenicity and cell wall permeability (Camacho *et al.* 2001). *pks18*

encodes a calchone synthase-like condensing enzyme that is also involved in polyketide synthesis (Saxena *et al.* 2003). The *ppsD* gene was downregulated around two-fold in the *regX3* overexpressing strain, whereas *pks18* was upregulated 35-fold. The potential regulation of these genes by RegX3 reinforces a hypothesis for a role for RegX3 during the infection process, since the products of both of these genes are involved in synthesis of cell wall lipids that can interact with the host. The differences in up- and down-regulation of these genes may represent the complex nature of the response to this interaction.

4.3.4.2.4 Genes differentially regulated in a strain of *M. tuberculosis* overexpressing *regX3* and after growth of a *regX3* null mutant in microaerobic conditions

Two genes in particular are shared between the gene lists for the *regX3* overexpressing strain and the list generated by comparing the *regX3* null mutant against the wild-type under growth in low oxygen conditions: *sahH* and *nadA*. *sahH* encodes a thioester hydrolase which may be involved in methionine and selenoamino acid metabolisms (Tuberculist,(2005) and was identified by proteomic methods to reside in the cytoplasm (Rosenkrands *et al.* 2000). *nadA* is a probable quinolinate synthetase that catalyzes the formation of quinolinate from aspartate. The appearance of these genes in both lists is compelling evidence for an interaction with RegX3 and a role for the SenX3-RegX3 two component signal transduction system in regulation of the intermediary metabolism.

4.3.4.2.5 Genes involved in the oxidative stress response

Upregulated about 1.5-fold compared to the average were the co-transcribed *ahpC* and *D* genes, encoding alkyl hydroperoxidases C and D respectively. AhpC and D are key components in the oxidative stress defence of *M. tuberculosis*, particularly in the absence of KatG which is inactivated in the majority of isoniazid-resistant clinical strains. AhpD is required for the reduction of AhpC and also has its own alkylhydroperoxidase activity (Bryk *et al.* 2002; Koshkin *et al.* 2003); AhpC is particularly important in the detoxification of the highly damaging peroxynitrite radical (see 1.4.3 and Master *et al.* (2002)) and also has thioredoxin-dependant peroxidase activity (Wieles *et al.* 1995). The level of upregulation of these genes is not great in the *regX3* overexpressing strain compared to induction of *ahpC* expression in *M. tuberculosis* (Sherman *et al.* 1996) and other organisms (Vattanaviboon *et al.* 2003) upon exposure to organic hydroperoxides; however the small level of upregulation may be biologically significant as in the growth model used in this study there was no specific inducer of AhpC/D expression. Perhaps the SenX3-RegX3 two component signal transduction system has a role in priming or otherwise contributing to the expression of AhpC and D *in vivo*.

4.3.4.2.6 MEME motif analysis of potential promoter regions of genes identified to be differentially expressed between a ΔregX3/P_{hsp60}-*regX3*$^+$ strain and a ΔregX3/P_{hsp60}-*regX3*$^-$ strain of *M. tuberculosis*

The lack of apparent motifs in the upstream sequences of differentially regulated genes between the ΔregX3/P_{hsp60}-*regX3*$^+$ and ΔregX3/P_{hsp60}-*regX3*$^-$ strains does not compromise the validity of that gene list, since regulatory motifs are notoriously difficult to find (Rand *et al.* 2003). Where this approach has been successful, for example in

Gamulin *et al.* (2004), prior experimental knowledge of the regulation of the analysed genes (Gopaul *et al.* 2003) allowed the search for motifs to be restricted to much shorter lengths of DNA than those analysed here.

4.3.5 Summary

Overall, given the apparent upregulation of a number of genes thought to be involved in the oxidative stress response and adaptation to oxygen and nutrient limiting stationary phase in a *regX3* overexpressing strain of *M. tuberculosis* and after growth of a *regX3* null mutant in microaerobic conditions, it is possible to envisage a role for the SenX3-RegX3 two component signal transduction system in the adaptation of the bacterium to conditions within the lung, perhaps during the initial replication of the organism. This would be likely to be as a result of sensing oxygen levels, but at an intermediate level compared to DosR. Furthermore, given the diverse nature of the possible regulons of RegX3, it is also possible to hypothesise that the input signal for SenX3 requires a combination of environmental factors, either direct or indirect, for the system to become active. This may be a possible explanation for the lack of an obvious *in vitro* phenotype of the null mutants in the SenX3-RegX3 two component signal transduction system to a single agent seen in Chapter 3.

5 Expression, purification and analysis of the DNA-binding interactions of recombinant RegX3

5.1 INTRODUCTION

Both structure prediction and previous experimental evidence suggest that RegX3 is a DNA-binding protein. Pfam analysis (Bateman *et al.* 2004) of the RegX3 sequence from Tuberculist (2005) indicates that RegX3 contains a receiver domain at the N-terminus and a DNA-binding domain at the C-terminus similar to one described for OmpR of *E. coli* (Martinez-Hackert & Stock 1997).

5.1.1 Previous work on the DNA-binding properties of RegX3

RegX3 has been shown to act as the response regulator for the sensor protein SenX3 by Himpens *et al.* (2000), where a renatured form of recombinant, polyhistidine-tagged RegX3 expressed in *E. coli* was shown to interact with the region upstream of RegX3. There are two major caveats to this work. Firstly, recombinant RegX3 was purified under denaturing conditions and renatured through two successive rounds of dialysis. The polyhistidine-tagged RegX3 may not have refolded correctly by this method, meaning that the structure of the DNA binding site may not represent the natively folded RegX3 as it occurs in *M. tuberculosis*, altering its recognition of the binding sequence. Secondly, the identity of the purified RegX3 was confirmed by size and reaction with anti-polyhistidine antibody. The apparent molecular weight of the recombinant RegX3 was 4.2 kDa higher than expected in that paper, and it is possible, although extremely unlikely, that this polyhistidine tag was not attached to the correct peptide. The paper makes no mention of sequencing the expressing construct; it is possible that frameshift or substitution mutations could have occurred during the amplification of the *regX3* sequence.

This paper (Himpens *et al.* 2000) attributed the RegX3 binding sequence to a region of 42 bp upstream of the *senX3* gene by out-competing the DNA-protein interaction of a labelled, amplified region between the stop codon of *gpm1* and the start codon of *senX3* with smaller, unlabelled DNA fragments derived from this region (Himpens *et al.* 2000). They did not systematically map the region, and relied on *ad hoc* evidence derived by eye from the sequence of this region from *M. tuberculosis* to identify this 42 bp fragment.

5.1.2 Potential difficulties with the expression of mycobacterial proteins in *E. coli*

When attempting to express and purify mycobacterial proteins, there are a number of factors that have to be considered. Ideally, *M. tuberculosis* proteins would be overexpressed in *M. tuberculosis*; meaning that the purified proteins would be most likely to be in their native conformation. Unfortunately biosafety considerations make this difficult, as does the long generation times of *M. tuberculosis* and *M. bovis*. There are expression systems available in mycobacteria, but none are as versatile as those developed for the expression of proteins in *E. coli*.

Problems with protein solubility may arise due to the speed at which the induced *E. coli* strains translate the recombinant RegX3 peptide. The protein may not have time to fold properly at this level of expression, so forms aggregate 'inclusion bodies' that are insoluble. These may be recovered by denaturing the inclusion bodies although it is less likely that the protein, when renatured by dialysis, will be folded correctly in an active state. It is also possible to decrease the rate of translation by reducing the temperature of

the expressing culture, or by inducing expression of the peptide with lower levels of IPTG.

In conclusion, *E. coli* is a safe, fast and efficient method for the production of proteins and a number of commercial kits are available for this purpose. The disadvantage is that heterologous proteins from sources other than *E. coli* may be misfolded. Furthermore, *M. tuberculosis* has a high GC-content genome and translation of genes in this organism may require the use of some codon-specific tRNA more often than would be needed for the translation of *E. coli* sequences in *E. coli*. To prevent a lack of the tRNA in the appropriate quantities limiting the translation rates of *M. tuberculosis* sequences during protein expression, *E. coli* strains that carry a plasmid encoding the supplementary tRNA genes are used as standard in our laboratory.

5.1.3 The use of electrophoretic mobility shift assays to analyse protein-DNA interactions

The electrophoretic mobility shift assay (EMSA, also referred to as a bandshift or gelshift assay) provides a method of analysing directly the interactions between protein and DNA. Gelshift assays rely on the ability of protein to bind radiolabelled DNA and retard its progress through a nondenaturing gel during electrophoresis (see 2.6.3). The results are visualised by autoradiography, where an interaction will be seen as a band present in the sample where protein and probe have been allowed to mix before electrophoresis, which is not present in a sample to which no protein has been added. Alternatively, protein extract from a source not containing the protein of interest may be used as a negative control.

Further experiments can be used to confirm the nature of the interaction: i.e. specificity of the probe sequence for binding, or the identity of the protein that is retarding the electrophoretic mobility of the probe.

5.1.3.1 Cold-competition assay

To confirm the importance of the specific sequence of the probe in the DNA-protein interaction, it is preferable to show that the 'shift' – i.e. the visualisation of the retardation of the labelled probe by the protein – can be out competed by identical, specific unlabelled probe, but not a comparable, unlabelled probe that should not bind the protein.

5.1.3.2 Supershift assay

To confirm the identity of the protein forming the complex with the DNA, antibodies to the protein can be used to increase the molecular mass of the DNA-protein complex and further retard its passage through the gel, causing an apparent 'supershift' (see 2.3.6.7 and Asusubel *et al.* (2005) for method). In practice the effects of the addition of antibody depends on the way it interacts with the protein of interest. If the epitope on the protein is away from the DNA-binding site, then it is likely that a supershift will occur whether the antibody is added before or after the DNA-protein complex is allowed to form. If the binding of the epitope to the antibody obscures the DNA-binding site, then addition of the antibody before the DNA-protein complex is allowed to form will abolish the original shift. Alternatively if the DNA-binding site is obscured after the DNA-protein complex

is allowed to form, then the supershift should still occur provided the epitope has not been destroyed.

5.1.3.3 Shift-blot

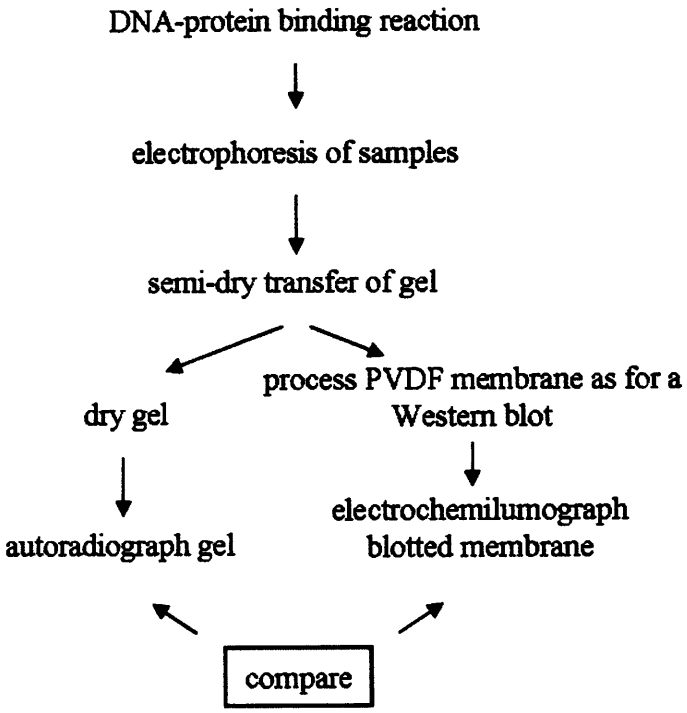
Another method of confirming the interaction of the protein of interest with the probe is to visually co-localise the protein with the shifted band. This may be done by performing a shift-blot assay where the complexes contained within the gel after a standard gel-shift assay are electroblotted onto a membrane, which is then processed as for a Western blot (see 2.6.3.6 and Demczuk *et al.* (2003) for method). The gel still contains enough radiolabelled probe to allow it to be dried and autoradiographed. The autoradiograph and the autoelectrochemilumograph from the Western blot are then aligned to co-localise the protein with the shift (Figure 5.1).

These control experiments are particularly necessary when using impure protein preparations; less so when the protein is purified and its identity confirmed. They do, however, contribute to our understanding of the interaction. It is possible to use the shift-blot and supershift assays to perform investigations of the interactions of multiple proteins that may be involved in DNA binding; however for this to be effective a good working method must be established and the identity of other proteins in the DNA-protein complex usually need to be identified by other methods.

5.1.4 Protein sources for gelshift assays

There are two possible sources of protein for gelshift assays. The preferable method involves the use of enriched or pure protein preparations, which means that it is

Figure 5.1: Diagram illustrating the steps involved in a shift-blot experiment.


DNA-protein binding reaction
electrophoresis of samples
semi-dry transfer of gel
dry gel
process PVDF membrane as for a Western blot
autoradiograph gel
electrochemilumograph blotted membrane
compare

more likely that DNA-protein interactions seen will be the correct ones attributable to that particular protein. It must also be considered as to whether the protein of interest is not folded in the correct manner and that the interaction does not represent that which would happen *in vivo*.

An alternative protein source for gelshift assays is to use whole cell protein extracts from the native bacterium, and rely on comparisons between the wild-type protein extract and that taken from a mutant shown to be lacking only the protein of interest. Although this does have the advantage that the proteins are more likely to be in their native, correctly folded state, this approach requires a knockout strain for the protein of interest, and it is possible that a number of interaction will take place between the DNA probe and non-sequence-specific DNA-binding proteins. This approach was not used in this study due to the large quantity of *M. tuberculosis* that needs to be cultured to obtain sufficient protein for an experiment (~10 litres for comparable experiments with *M. bovis*, R. Curtis, personal communication) which represents a significant drain on the resources of the laboratory and raises biosafety concerns.

Prior understanding of the potential interaction between the protein of interest and its mode of action may also be relevant when considering its interaction with DNA. For example, a regulatory protein that requires phosphorylation to activate transcription of genes may not bind to the probe unless it is phosphorylated. Fortunately, in the absence of the purified cognate phosphate donor protein that would phosphorylate the protein *in vivo*, phosphorylation of proteins can be carried out in a relatively simple, crude manner by incubation of the protein with acetyl phosphate, a high-energy phosphate donor. This

will cause phosphorylation at most likely phosphoacceptor sites on the protein (E. Davis, personal communication).

5.1.5 DNA sources for gelshift assays

A sequence of DNA that the protein of interest is thought to bind must first be identified. This is then either amplified by PCR (for fragments >50bp), or ordered as complementary oligonucleotides. Longer products are more likely to contain secondary structure that will obscure the bandshift, but are more likely to contain protein binding sites if the exact position is not known. Complementary oligonucleotides must be annealed in equimolar concentrations to produce short lengths of double stranded DNA.

5.1.6 Labelling of DNA for gelshift assays

The most versatile method of visualising the shift of a probe is by radiolabelling the probe with ^{32}P before the DNA and protein are mixed, followed by autoradiography of the dried gel once the electrophoresis has finished. A non-radioactive method of labelling the DNA is to label using dioxygenin, which is visualised by electrochemiluminescence, although that method has not been used in this study.

Blunt-ended probes generated by the annealing of oligonucleotides or PCR amplification with a Pfu-based polymerase may be radiolabelled by the addition of a ^{32}P phosphate group from ^{32}P-γ-ATP by a polynucleotide kinase. All probes labelled by this method should show uniform radioactivity regardless of probe size.

Probes generated with overhangs (by cleavage by restriction enzymes, addition of polylinkers or by the annealing of oligonucleotides mismatched at the 5' or 3' and) may

also be labelled in this fashion. Alternatively, the ends may be 'polished' by the Klenow fragment of DNA polymerase I using ^{32}P-γ-ATP as a substrate. The intensity of the radioactive signal in this case will depend on the number of adenosine bases on the overhanging regions.

To generate highly radioactive probes it is possible to amplify the target sequence of interest using the ^{32}P-γ-ATP as a replacement for the usual dATP in the deoxynucleotide mix of the PCR reaction. The radioactivity of a probe labelled in this fashion will therefore depend on the number of adenosine bases in the total sequence and is therefore likely to increase with the length of the probe. However, it is preferable to work with the minimum amount of radioactivity necessary for an experiment, and the end-labelling method described above is usually sufficient for adequate exposure times in autoradiographs.

5.2 RESULTS

5.2.1 Purification of recombinant polyhistidine-tagged RegX3

5.2.1.1 Construction of an *E. coli* strain expressing recombinant polyhistidine-tagged RegX3

The open reading frame of *regX3* was amplified by PCR and cloned into a commercial vector (pET15b, see Table 2.2) in-frame and downstream of the polyhistidine tag sequence by D. Hunt. A protease-negative strain of *E. coli* was used to study the expression characteristics of small-scale cultures that had been induced to transcribe the recombinant RegX3 gene (see 2.6.2). Whole cell extracts of all analysed cultures were able to express a polyhistidine-tagged protein of approximately the expected molecular

weight (25.67kDa) for the cloned allele (Figure 5.2) that was assumed to be recombinant polyhistidine-tagged RegX3.

5.2.1.2 Assessment of the solubility of recombinant polyhistidine-tagged RegX3 expressed at different temperatures

The solubility of the expressed recombinant RegX3 was assessed after induction of expression of an expressing strain of *E. coli* at 18°C, 25°C, 30°C and 37°C and isolation of the soluble and insoluble portions of the cell pellets at each temperature (see 2.6.2.5), which were obtained after sonication by the centrifugation of the lysed cells at 13000rpm in a Heraeus microcentrifuge and separation of the fluid phase (soluble protein) from the pellet (insoluble protein). The proteins in the pellet were recovered by resuspension in an SDS-containing buffer (Novagen) and heating at 95°C for 5 min, with vortexing for 5 sec every minute. Analysis of the soluble and insoluble samples by a commercial, histidine-selective stain showed that the expressing clones produce proportionally more soluble recombinant protein at 18°C than at 37°C (Figure 5.3), although the band was quite faint. Although it appears that a greater proportion of recombinant protein is in the insoluble fraction of the cell lysate, it was felt that sufficient soluble protein had been produced to allow the experiment to proceed. Recombinant polyhistidine RegX3 was identified by comparison with the empty vector control since the stain proved to be poorly selective for polyhistidine sequences; however it was felt that the results were clear enough to remove the need for a Western blot of the same gel.

Figure 5.2: Autoelectrochemilumograph of recombinant polyhistidine-tagged RegX3 expressing clones of *E. coli* after Western blot analysis with an anti-polyhistidine primary antibody. M1, coloured molecular weight markers, M2, polyhistidine-tagged molecular weight markers, C, empty vector control strain, 1-7, separate expressing clones. The molecular weight marker sizes expressed on the left hand side correspond with lane M1, and estimate the molecular weight of the polyhistidine tagged RegX3 (band marked with an arrow on the image) to be around 24 or 25 kDa. The expected size of the recombinant protein is 25.67 kDa. Approx 60 ng of whole cell extract in a volume of 10µl was added to the sample wells of a precast 10% Bis-Tris gel (Novagen), and was electrophoresed and blotted as described in 2.6.2.2 and 2.6.2.3. Markers were added as detailed in the manufacturers' instructions.

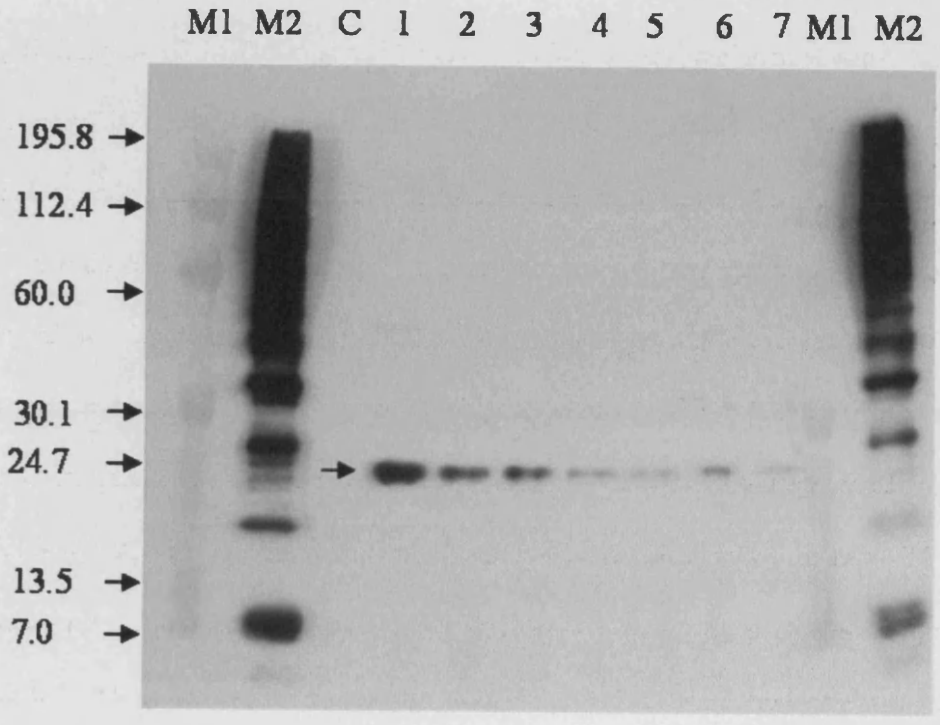
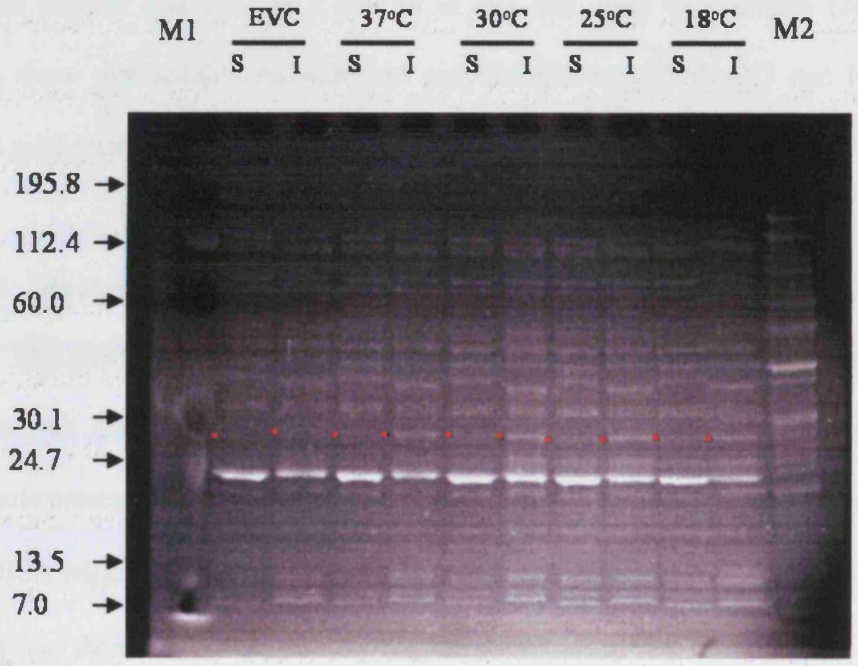

M1 M2 C 1 2 3 4 5 6 7 M1 M2
195.8
112.4
60.0
30.1
24.7
13.5
7.0

Figure 5.3: Photograph of a gel stained with a commercial polyhistidine-specific stain and visualised under UV light as part of a solublility assay, showing increased solubility of the recombinant polyhistidine-tagged RegX3 when expressed at lower temperatures. M1 and M2, coloured and polyhistidine-tagged molecular weight markers, respectively. EVC, empty vector control (expressed at 37°C), S, soluble cell fractions, I, insoluble cell fractions. Red spots mark the band thought to be recombinant polyhistidine-tagged RegX3. The molecular weight marker sizes expressed on the left hand side correspond with lane M1 and are measured in kDa. Approx 60 ng of whole cell extract in a volume of 10µl was added to the sample wells of a precast 10% Bis-Tris gel (Novagen), and was electrophoresed as described in 2.6.2.2. Markers were added as detailed in the manufacturers' instructions. The separated proteins were stained and visualised using the InVision stain (Invitrogen) according to the manufacturer's instructions (see 2.6.2.2).


M1
EVC
37°C
30°C
25°C
18°C
M2
S I S I S I S I S I
195.8
112.4
60.0
30.1
24.7
13.5
7.0

5.2.1.3 Purification of recombinant polyhistidine-tagged RegX3

A large culture of *E. coli* expressing the recombinant RegX3 allele was grown and induced at 18°C (see 2.6.2.5). After lysis of the cells under nondenaturing conditions the soluble fraction was passed through an IMAC column and the flow-through was analysed by Coomassie stain and Western blot (Figure 5.4) after each stage in the process (see 2.6.2.6). An identical strain without the cloned *regX3* sequence was treated in an identical manner and used as a control in this and other experiments (see later). The results show that soluble recombinant polyhistidine-tagged RegX3 can be isolated in almost pure form from the *E. coli* cell lysate.

5.2.1.4 Manipulation of soluble recombinant polyhistidine-tagged RegX3

The purified recombinant polyhistidine-tagged RegX3 recovered from the IMAC column was dialysed into an appropriate buffer (PBS, Appendix 4) to remove the imidazole present in the elution buffer. An estimation of the concentration of the protein using BSA reagent (Pierce) gave a concentration of ~6μg/ml (0.2 pmol/μl). The protein sample was then concentrated to 0.17mg/ml (5.5pmol/μl) by ultrafiltration (see 2.6.2.8), the RegX3 negative control IMAC eluate volume adjusted by the same amount, and the two samples and their flow-through analysed on a Coomassie stained gel (Figure 5.5), which shows that the protein had indeed been concentrated successfully, producing a stronger band, had not visually degraded and had not been lost through the pores in the ultrafiltration membrane,

Figure 5.4: SDS-PAGE analysis of soluble extracts of *E. coli* BL21 DE3-RP/pET15b-*regX3* cultured at 18°C, stained with Coomassie R250 (A) and an anti-polyhistidine Western blot (B) of IMAC column fractions. Approx 60 ng of whole cell extract in a volume of 10µl was added to the sample wells of a precast 10% Bis-Tris gel (Novagen), and was electrophoresed and blotted as described in 2.6.2.2 and 2.6.2.3. Markers were added as detailed in the manufacturers' instructions. M1 and M2, coloured 'kaleidoscope' markers (Bio-Rad Laboratories) and polyhistidine-tagged molecular weight markers (Invitrogen), respectively. The numbers to the left of the images represent the molecular weights of the M1 markers; the numbers on the right of the images represent the molecular weights of the M2 markers. 1, empty vector control (*regX3*$^-$) flow-through (5mM imidazole), 2, *regX3*$^+$ flow-through (5mM imidazole), 3, binding buffer flow-through of *regX3*$^+$ protein sample (5mM imidazole), 4, imidazole buffer flow-through of *regX3*$^+$ protein sample (60mM imidazole), 5, imidazole elution of matrix-bound protein of *regX3*$^+$ protein sample (1M imidazole), 6, imidazole elution of matrix-bound protein of *regX3*$^-$ protein sample (1M imidazole). Molecular weights indicated on the left refer to the coloured markers, M1. 15µl aliquots of each sample were loaded into each sample well of two pre-cast 10% Bis-Tris polyacrylamide gel and were electrophoresed and blotted identically according to 2.6.2.2 and 2.6.2.3.

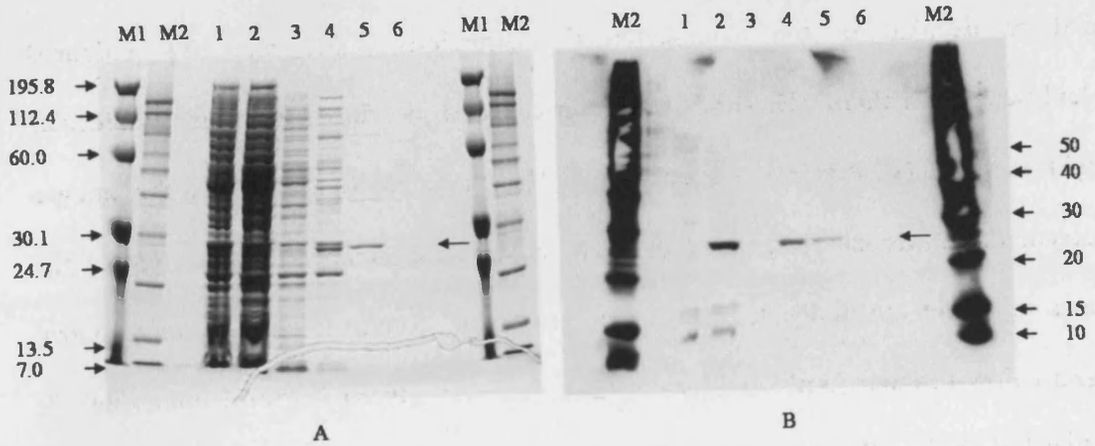
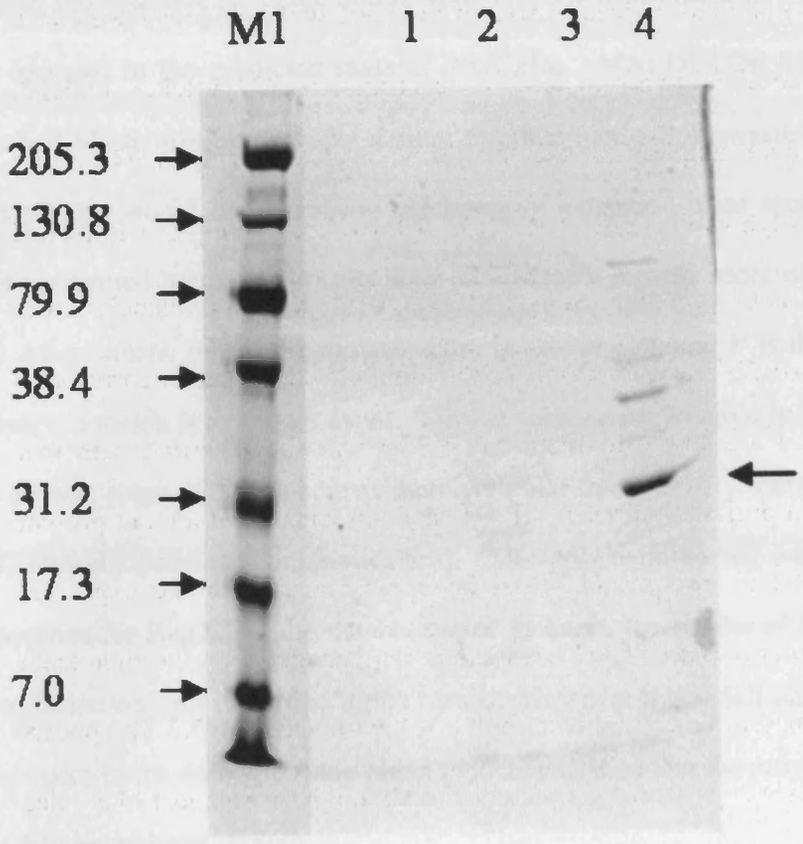

M1 M2 1 2 3 4 5 6 M1 M2
195.8
112.4
60.0
30.1
24.7
13.5
7.0
A
M2 1 2 3 4 5 6 M2
50
40
30
20
15
10
B

Figure 5.5: Coomassie stained SDS-polyacrylamide gel of concentrated protein samples after ultrafiltration. M1, coloured molecular weight markers, with molecular weight in kDa, 1, flowthrough of *regX3*$^-$ protein sample, 2, concentrated *regX3*$^-$ protein sample, 3, flowthough of *regX3*$^+$ protein sample, 4, concentrated *regX3*$^+$ protein sample. The arrow marks the recombinant polyhistidine-tagged RegX3. 10µl of each sample was loaded into the wells of a 12% Tris-Cl gel (supplied by E. McGovan), with approximately 1.8µg of purified protein in the *regX3*$^+$ sample. Some of the higher bands in lane 4 correspond with those in lane 2, and are likely to represent proteins in the *E. coli* lysate that themselves bind the nickel purification resin. Those bands in lane 4 that are not present in lane 2 are likely to be *E. coli* proteins that have associated with the expressed heterologous *M. tuberculosis* protein. The band marked with an arrow on the right hand side of the image is assumed to be recombinant polyhistidine-tagged RegX3. It appears to have a higher molecular weight than seen previously, although this is most likely due to the use of Tris-Cl gel (rather than the Bis-Tris gels used elsewhere in this chapter), and a higher concentration of that protein in the sample, resulting in slower migration of the protein during electrophoresis.


M1
1
2
3
4
205.3
130.8
79.9
38.4
31.2
17.3
7.0

5.2.1.5 Identification of recombinant RegX3

The protein assumed to be recombinant polyhistidine-tagged RegX3 was excised from a gel and prepared for MALDI-TOF mass spectrometry (see 2.6.2.11), the results of which showed 68% coverage of a protein with a nominal mass of 24.891kDa (Table 5.1), as opposed to the predicted mass of 25.67kDa. MALDI-TOF MS is predominantly a method of identifying proteins; a more accurate method of measuring the actual mass of the protein would be to perform electrospray ionisation mass spectrometry, which was not performed here. The results showed a Mascot mowse score of 153 for RegX3 from *M. tuberculosis*, where the mowse score is -10*log(P), and P is the probability that the observed match is a random event. Mowse scores over 76 are significant below p=<0.05. A mowse score of 153 is approximately similar to an $e=10^{-15}$ value for a BLAST search (S. Howell, personal communication). The polyhistidine tag was not included in the sequence for RegX3 in the database used to query the results of the mass spectrometry and hence was not identified in the results. However it was felt unnecessary to re-submit the query to the database since direct evidence tells us that the polyhistidine tag is present in the recombinant protein.

5.2.2 Gelshift assays

The recombinant polyhistidine-tagged RegX3 was used to study DNA-protein interactions with regions of DNA thought to contain RegX3 binding sites.

5.2.2.1 Titration of recombinant RegX3 against a previously identified binding site

Table 5.1: MALDI mass spectrometry results for the purified RegX3 protein showing matched peptides in black and unmatched H, K and R residues in red.

```
  1 MTSVLIVEDE ESLADPLAFL LRKEGFEATV VTDGPAALAE FDRAGADIVL LDLMLPGMSG TDVCKQLRAR
 71 SSVPVIMVTA RDSEIDKVVG LELGADDYVT KPYSARELIA RIRAVLRRGG DDDSEMSDGV LESGPVRMDV
141 ERHVVSVNGD TITLPLKEFD LLEYLMRNSG RVLTRGQLID RVWGADYVGD TKTLDVHVKR LRSKIEADPA
211 NPVHLVTVRG LGYKLEG
```

Observed	Mr(expt)	Mr(calc)	± ppm	Start	End	Miss	Peptide
701.419	700.411	700.387	34.5	176	181	0	GQLIDR
727.485	726.477	726.486	-13.4	112	117	1	IRAVLR
811.416	810.408	810.460	-64.3	193	199	0	TLDVHVK
836.383	835.375	835.444	-81.9	220	227	1	GLGYKLEG
967.513	966.505	966.561	-57.8	193	200	1	TLDVHVKR
1159.611	1158.603	1158.643	-34.3	71	81	0	SSVPVIMVTAR
1175.594	1174.587	1174.638	-43.6	71	81	0	SSVPVIMVTAR 1 Oxidation (M)
1210.525	1209.517	1209.567	-40.9	182	192	0	VWGADYVGDTK
1630.826	1629.818	1629.884	-40.2	205	219	0	IEADPANPVHLVTVR
1845.956	1844.948	1845.011	-34.0	203	219	1	SKIEADPANPVHLVTVR
1892.855	1891.847	1891.943	-50.8	176	192	1	GQLIDRVWGADYVGDTK
1924.856	1923.848	1923.933	-43.8	176	192	1	GQLIDRVWGADYVGDTK 1 Oxidation (W)
2052.982	2051.974	2052.053	-38.4	88	106	0	VVGLELGADDYVTKPYSAR
2093.836	2092.828	2092.897	-32.7	118	137	1	RGGDDDSEMSDGVLESGPVR 1 Oxidation (M)
2094.886	2093.878	2093.990	-53.5	24	43	0	EGFEATVVTDGPAALAEFDR
2740.221	2739.214	2739.360	-53.5	82	106	1	DSEIDKVVGLELGADDYVTKPYSAR
2902.306	2901.299	2901.531	-80.1	143	167	1	HVVSVNGDTITLPLKEFDLLEYLMR

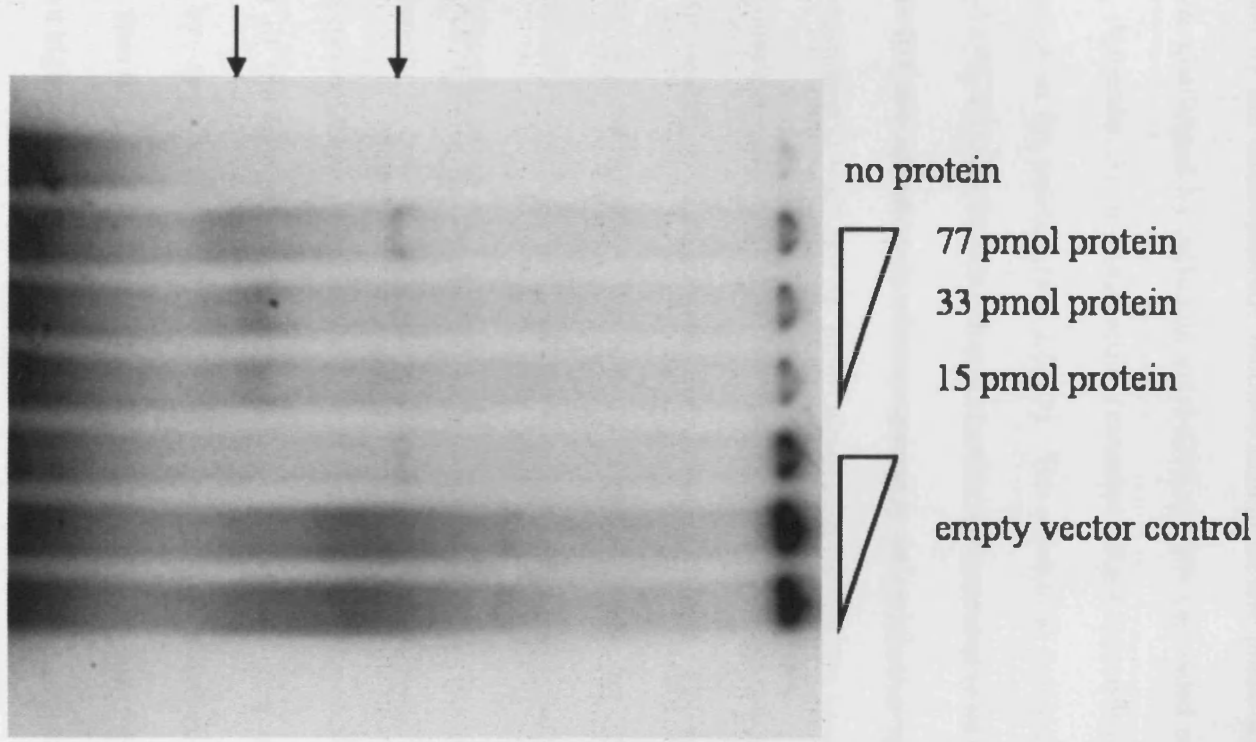
Himpens *et al.* (2000) identified a 42bp sequence that lies upstream of *senX3* as a potential binding site for RegX3 (given as oligonucleotide 106 in Appendix 3; sequence 5'-atgtgaacggtaaccgaacagctgtggcgtagtgtgtgactt-3'). We initially sought to replicate the results of that paper to confirm the interaction between soluble recombinant RegX3 and this potential binding site.

A titration of different concentrations of the recombinant RegX3 bracketed around that described in Himpens *et al.* (2000), and equivalent volumes of the empty vector control preparation, were allowed to bind to labelled oligonucleotides of the same sequence as described in Himpens *et al.* (2000) (primer 160 annealed to its complementary primer 161, see Appendix 3). The reaction conditions were as described in 2.6.3.3. Autoradiography of the gel after electrophoresis showed two apparent shifts in the electrophoretic mobility of the probe when incubated with recombinant RegX3, both of which grew less intense with reducing concentrations of RegX3 (Figure 5.6). The upper shift seemed to be visible, although less intense, in the empty vector control strain. This could represent contamination of the empty vector control strain. The lower shift, however, did not seem to be present in the empty vector control strain and may represent a second binding site on the protein. The shift was not present in a reaction with no protein added.

5.2.2.2 Specificity of the interaction of recombinant RegX3 with a previously identified binding site

To identify the specific nature of the DNA sequence in the interaction between recombinant RegX3 and the probe, a cold-competition assay was carried out. The

Figure 5.6: Bandshift analysis of RegX3. Autoradiograph of a titration of recombinant polyhistidine-tagged RegX3 and the empty vector control protein source against the 42 bp fragment upstream of the *senX3* start codon identified in Himpens *et al.* (2000). Reactions were performed as described in 2.6.3.3, after which 10µl of each reaction was loaded onto a precast 10% TE gel (Invitrogen). At 77pmol of protein, the molar ratio of protein to DNA is 7.7:1; at 15pmol protein the molar ratio of protein to DNA is 1.5:1. The two apparent bandshifts are marked with arrows on the left hand side of the image.


no protein
77 pmol protein
33 pmol protein
15 pmol protein
empty vector control

bandshift caused by the electrophoretic mobility change of the probe when complexed with protein was challenged by including unlabelled specific (annealed oligonucleotides 160 and 161, Appendix 3) or nonspecific (annealed oligonucleotides 162 and 163, Appendix 3) probe in the reaction (see 2.6.3.3). The unlabelled nonspecific probe was double stranded 43bp oligonucleotide (oligonucleotide 162 annealed to its complimentary oligonucleotide 163, see Appendix 3) whose sequence is derived from a region internal to *senX3* and is mismatched from the *M. tuberculosis* sequence in the central nine bases. This sequence does not appear on the *M. tuberculosis* genome and would not be expected to be involved in binding to an *M. tuberculosis* response regulator. The sequences of the forward reading oligonucleotides of the specific (oligonucleotide 160) and nonspecific (oligonucleotide 162) regions are atgtgaacggtaaccgaacagctgtggcgtagtgtgtgactt (42bp) and gatccgggctatcagtgcgtgggaaagcccggttgctgagcga (43bp) respectively.

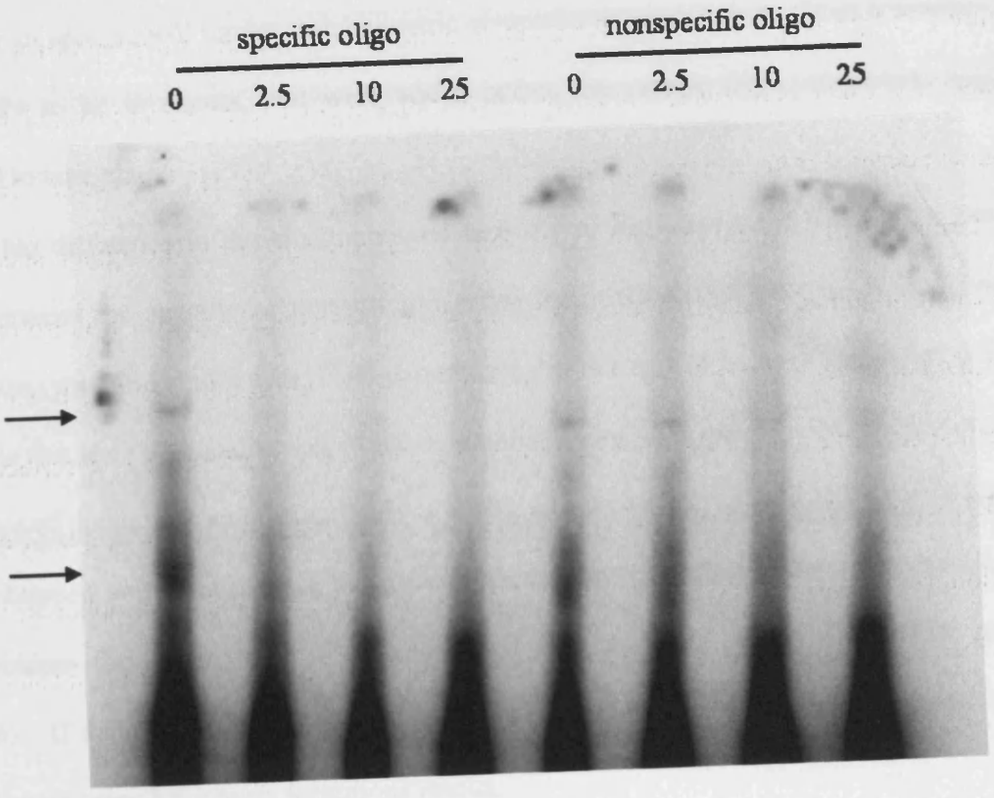
The intensity of both of the shifts in electrophoretic mobility of the probe was reduced as the concentration of unlabelled specific probe in the reaction increased over 2.5 times that of the labelled probe (Figure 5.7). The lower shift appeared to be outcompeted by specific unlabelled double stranded oligonucleotide at higher concentrations than the upper shift; this suggests that one of the putative binding sites on the protein has a higher affinity for the specific target sequence than the other.

5.2.2.3 Confirmation of the identity of the protein involved in DNA-binding

5.2.2.3.1 Supershift assay

To identify the protein in the DNA-protein complex causing the electrophoretic mobility shift, a supershift assay was carried out. The binding reactions were set up as before but

Figure 5.7: Specificity of RegX3 for a putative binding site. Autoradiograph of a cold-competition assay between recombinant polyhistidine-tagged RegX3 and labelled 42 bp fragment derived from oligonucleotides 160 and 161 (identified in Himpens *et al.* (2000), the formation of a complex between which is being outcompeted by either specific (annealed oligonucleotides 160 and 161, Appendix 3) oligonucleotide or a nonspecific (annealed oligonucleotides 162 and 163, Appendix 3) oligonucleotide at 0, 2.5, 10 and 25 times the concentration of the labelled probe. Reactions were performed as described in 2.6.3.3, with the inclusion of unlabelled DNA as described above. 10µl aliquots of each reaction were loaded onto precast 10% TE gels (Invitrogen) and electrophoresed as described previously. The two bandshifts are marked with arrows at the left hand side of the image.


specific oligo
nonspecific oligo
0
2.5
10
25
0
2.5
10
25

with the inclusion of anti-polyhistidine tag monoclonal antibody (Novagen), or nonspecific anti-glutathione-S-transferase tag (GST)(Novagen) monoclonal antibody at different concentrations (see 2.6.3.5). It was not felt necessary to estimate the exact concentrations of the antibodies, since the different affinities each antibody is likely to have for an epitope will render stoichiometric comparisons unnecessary. Both antibodies are likely to be in excess, and were added before the protein-DNA interaction was allowed to take place.

No difference in the electrophoretic mobility of the protein-DNA complex was seen between the specific or non-specific antibodies, with a shift present in all lanes containing protein (Figure 5.8). The experiment did not appear to work correctly; it is possible that the polyhistidine tag of the recombinant RegX3 is buried somewhere in the structure of the protein and the antibody is unable to gain access. Anti-RegX3 antibodies were obtained on two occasions, but did not seem to detect RegX3 either in the purified recombinant form seen here, or in whole cell extracts of *M. tuberculosis* (data not shown). If a functional anti-RegX3 antibody were obtained, this experiment could be repeated and may have more fortuitious results.

In an attempt to get the protocol to work, the experiment was repeated with a no-antibody control reaction included, as well as reactions where specific (anti-polyhistidine) and nonspecific (anti-GST) antibodies were added after the DNA- protein interactions were allowed to occur. A shift was observed in the no-antibody control reaction, along with comparable shifts in reactions with specific (anti-polyhistidine) and nonspecific (anti-GST) antibody added before or after the DNA- protein binding reaction (Figure 5.9). There appeared to be an additional, higher molecular weight shift in the

Figure 5.8: Attempt to identify recombinant polyhistidine-tagged RegX3 in the DNA-protein binding complex using a supershift assay (2.6.3.5). Autoradiograph of a supershift assay, where specific (anti-His) or nonspecific (NS) anti-glutathione (Novagen) antibody is used to increase the apparent molecular mass of the DNA-protein complex, further retarding the electrophoretic mobility of the DNA-protein complex through the gel. Reactions were performed as described in 2.6.3.3, with the inclusion of antibody as described above. 10µl aliquots of each reaction were loaded onto precast 10% TE gels (Invitrogen) and electrophoresed as described previously. The two bandshifts are marked with arrows at the left hand side of the image.

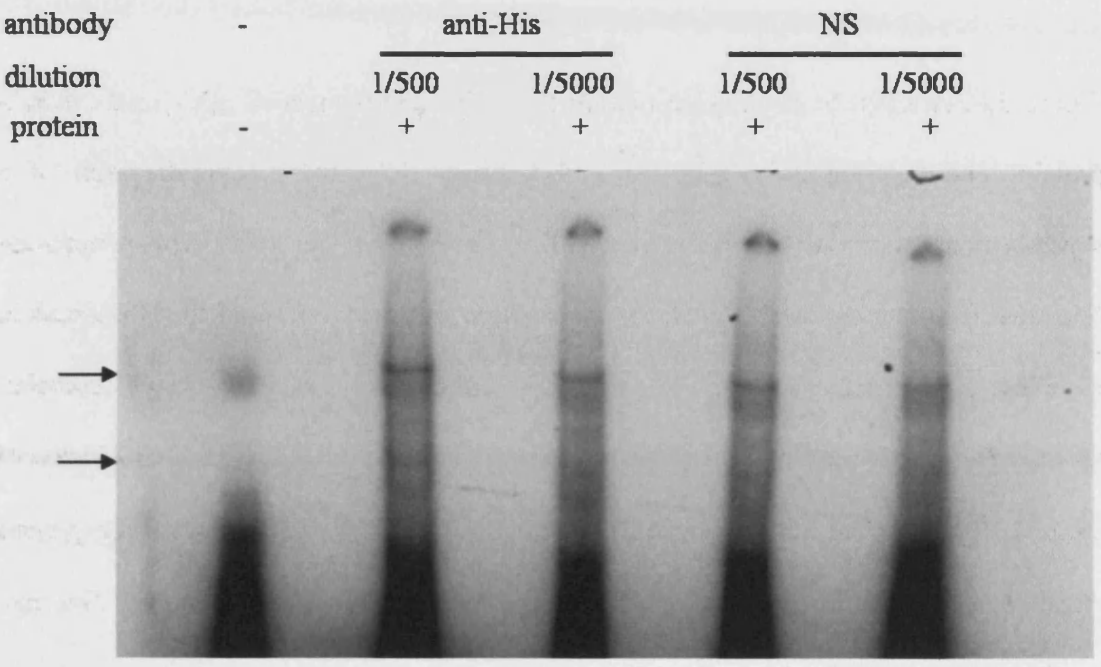
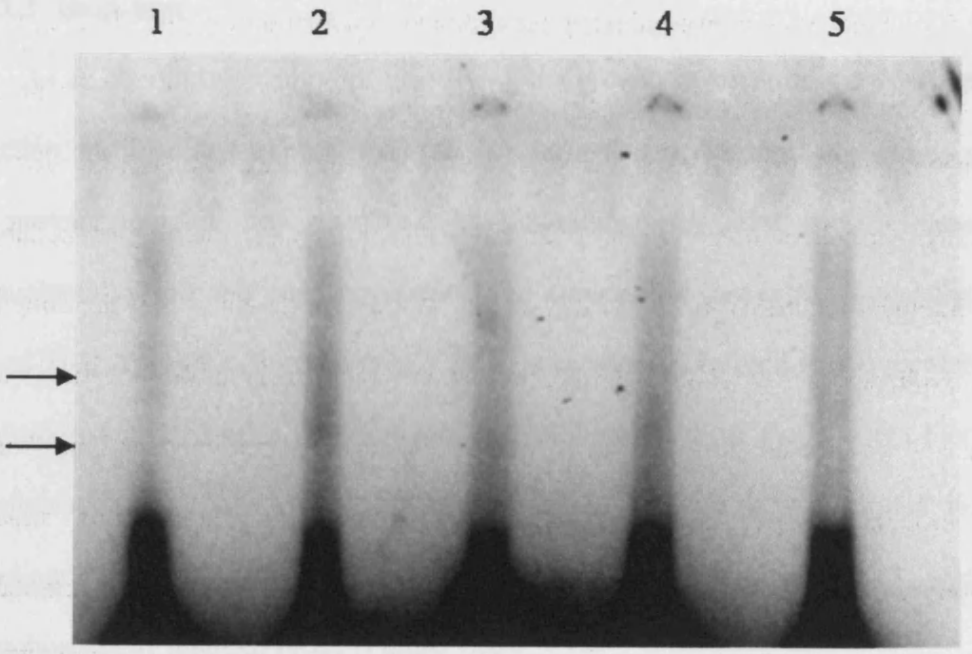

antibody
-
anti-His
NS
dilution
1/500
1/5000
1/500
1/5000
protein
-
+
+
+
+

Figure 5.9: Further attempt to identify recombinant polyhistidine-tagged RegX3 as part of the protein-DNA complex. Autoradiograph of a supershift assay, where specific (anti-His) or nonspecific (anti-GST) antibody is used to increase the apparent molecular mass of the DNA-protein complex, further retarding the electrophoretic mobility of the labelled 42 bp fragment identified in Himpens *et al.* (2000) through the gel. Lane 1, probe only, lane 2, probe and protein, lane 3, probe, protein and nonspecific antibody, lane 4, probe, protein and specific antibody added after the DNA-protein binding reaction. Lane 5, probe, protein and specific antibody added before the DNA-protein binding reaction. Reactions were performed as described in 2.6.3.3, with the inclusion of antibody as described above. 10µl aliquots of each reaction were loaded onto precast 10% TE gels (Invitrogen) and electrophoresed as described previously. The two bandshifts are marked with arrows at the left hand side of the image. See Asusubel *et al* (2005) for further discussion on potential electrophoretic mobility changes of DNA-protein complexes in supershift assays.


1
2
3
4
5

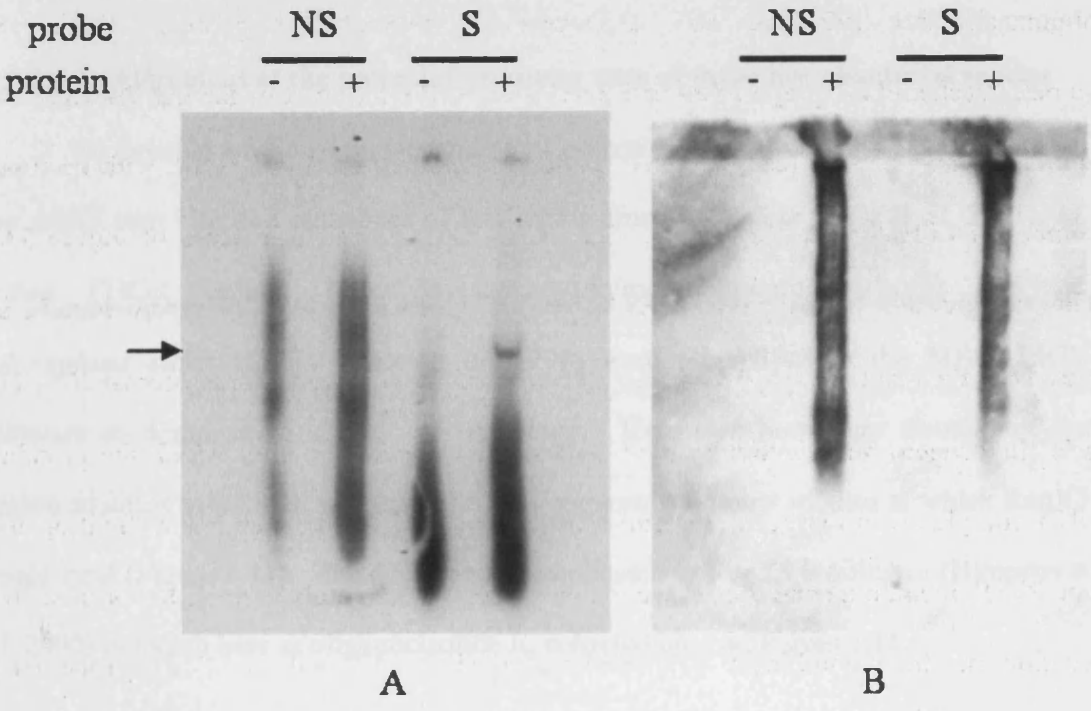
reaction containing the nonspecific antibody, and a possible increase in the mobility of the shift when specific antibody was added before the DNA-protein binding reaction.

5.2.2.3.2 Shift-blot

As an alternative method of identifying the protein involved in the DNA-protein interaction, a shift-blot was performed (see 2.6.3.6). The electrophoretic mobility of the DNA-protein complex was visualised by autoradiography after specific (annealed oligonucleotides 160 and 161, Appendix 3) or nonspecific (annealed oligonucleotides 162 and 163, Appendix 3) probe (see 5.2.2.2) was allowed to bind to the recombinant polyhistidine-tagged RegX3. An autoelectrochemilumograph of the Western blot with anti-polyhistidine antibody of the proteins contained in the same gel was aligned with the autoradiogram in an attempt to co-localise the recombinant RegX3 and the electrophoretically retarded probe (Figure 5.10).

The recombinant polyhistidine-tagged RegX3 appeared on the autoelectrochemilumograph as two long smears in the reactions containing protein. There were no bands in the interaction between the protein and the specific probe that co-localised with the shift apparent on the autoradiogram, although there did appear to be a band in the reaction between the non-specific probe and the protein in the autoelectrochemilumograph. This latter band did not co-localise with any band on the autoradiogram, suggesting that it may have been a background artefact of the autoelectrochemilumograph.

Figure 5.10: Shift-blot of the recombinant polyhistidine-tagged RegX3 interaction with the 42 bp fragment identified in Himpens *et al.* (2000). A is an autoradiograph of the interaction of recombinant polyhistidine-tagged RegX3 with either specific (S, annealed oligonucleotides 160 and 161, Appendix 3) or nonspecific (NS, annealed oligonucleotides 162 and 163, Appendix 3) radiolabelled probe. B is an autoelectrochemilumograph of an anti-polyhistidine Western blot of the gel shown as an autoradiograph in A. Reactions were performed as described in 2.6.3.3, with reaction components identical to those in 5.2.2.1 and the second and fifth lanes of Figure 5.6. This corresponds with a protein:DNA ratio of approximately 7.7:1 where protein is present in the reaction. 10µl aliquots of each reaction were loaded onto precast 10% TE gels (Invitrogen) and electrophoresed as described previously, after which the protein-DNA complexes in the gel were semi-transferred onto a PVDF membrane (Invitrogen). The PVDF membrane and the gel were then processed separately in order to visualise the location of the recombinant polyhistidine-tagged RegX3 and the DNA probe (see 2.6.3.6). The arrow at the left of the image marks a clear band in the fourth lane of image A.


probe
protein
NS
S
NS
S
- + - +
- + - +
A
B

5.2.2.4 Mapping of the interactions between recombinant RegX3 and the region upstream of *senX3*.

Since previous published data (Himpens *et al.* 2000) did not systematically map the potential promoter region of the *senX3-regX3* operon for interactions with recombinant polyhistidine-tagged RegX3, it was decided to map the interactions between the recombinant polyhistidine-tagged RegX3 and overlapping oligonucleotides covering the region between *gpm1* and *senX3*.

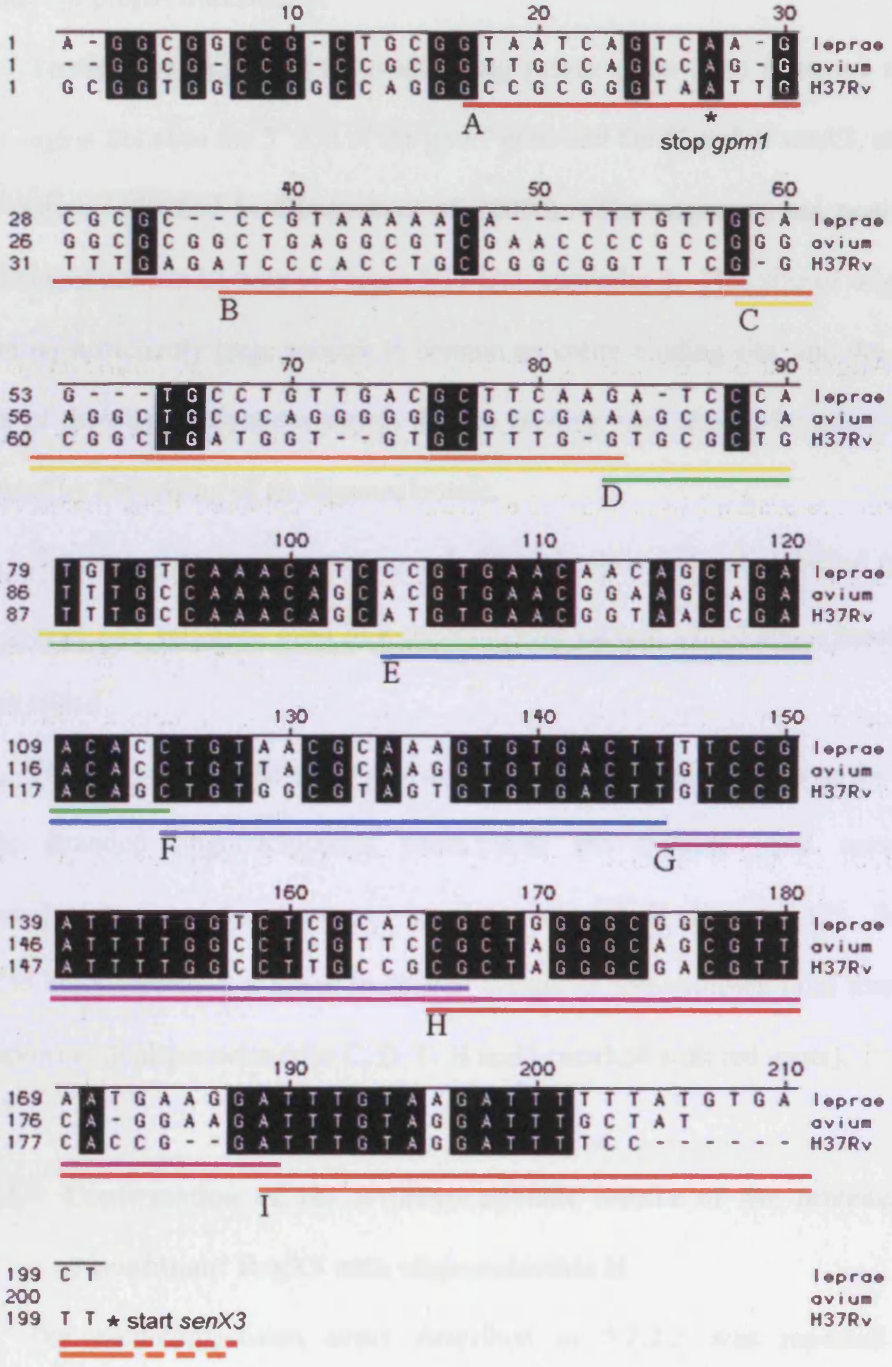
5.2.2.4.1 Alignment of the potential promoter sites of three mycobacterial species

To develop a rational understanding of potential RegX3 binding sites upstream of the *senX3* start site, the sequences of this region from *M. leprae* (Cole *et al.* 2001), *M. avium* (TIGR online; http://www.tigr.org/tdb/mdb/mdbinprogress.html) and *M. tuberculosis* strain H37Rv (Cole *et al.* 1998) were aligned using the MEGALIGN software to demonstrate regions of homology. Extensive homology throughout the region adjacent to the start of the *senX3* gene suggests a number of sites at which RegX3 might bind (Figure 5.11). The 42 bp region implicated in RegX3 binding in (Himpens *et al.* 2000) is shown here as oligonucleotide E, coloured blue on Figure 5.11.

5.2.2.4.2 Design of oligonucleotides to map the potential RegX3 binding sites upstream of *senX3*

Given the extensive nature of the sequence homology between *M. leprae*, *M. avium* and *M. tuberculosis* strain H37Rv, especially between base pairs 579173 and 579348, it was decided to physically map the interactions between recombinant RegX3

Figure 5.11: Alignment of the region upstream of *senX3* in *M. leprae*, *M. avium* and *M. tuberculosis* generated from the program MEGALIGN, annotated by the author. Matching bases are marked by a black background. The coloured lines represent the oligonucleotides designed for mapping the interactions of recombinant polyhistidine-tagged RegX3 with this region, and the corresponding labels for each of the oligonucleotides is shown at the 5' end of each sequence. The 42 bp fragment identified in Himpens *et al.* (2000) to interact with RegX3 is shown in blue as oligonucleotide E. Oligonucleotides H and I extend 4 and 25 bp into the coding region of *senX3*, respectively. All oligonucleotides with the exception of oligonucleotide E are 43bp long.


10
20
30
1 A - G G C G G C C G - C T G C G G T A A T C A G T C A A - G lepra
1 - - G G C G - C C G - C C G C G G T T G C C A G C C A G - G avium
1 G C G G T G G C C G G C C A G G G C C G C G G G T A A T T G H37Rv
A
* stop gpm1
40
50
60
28 C G C G C - - C C G T A A A A A C A - - - T T T G T T G C A lepra
26 G G C G C G G C T G A G G C G T C G A A C C C C G C C G G G avium
31 T T T G A G A T C C C A C C T G C C G G C G G G T T T C G - G H37Rv
B
C
70
80
90
53 G - - - T G C C T G T T G A C G C T T C A A G A - T C C C A lepra
56 A G G C T G C C C A G G G G G G C G C C G A A A G C C C C A avium
60 C G G C T G A T G G T - - G T G C T T T G - G T G C G C T G H37Rv
D
100
110
120
79 T G T G T C A A A C A T C C C G T G A A C A A C A G C T G A lepra
86 T T T G C C A A A C A G C A C G T G A A C G G A A G C A G A avium
87 T T T G C C A A A C A G C A T G T G A A C G G T A A C C G A H37Rv
E
130
140
150
109 A C A C C T G T A A C G C A A A G T G T G A C T T T T C C G lepra
116 A C A C C T G T T A C G C A A G G T G T G A C T T G T C C G avium
117 A C A G C T G T G G C G T A G T G T G T G A C T T G T C C G H37Rv
F
G
160
170
180
139 A T T T T G G T C T C G C A C C G C T G G G G C G G C G T T lepra
146 A T T T T G G C C T C G T T C C G C T A G G G C A G C G T T avium
147 A T T T T G G C C T T G C C G C G C T A G G G C G A C G T T H37Rv
H
190
200
210
169 A A T G A A G G A T T T G T A A G A T T T T T T A T G T G A lepra
176 C A - G G A A G A T T T G T A G G A T T T G C T A T avium
177 C A C C G - - G A T T T G T A G G A T T T T C C - - - - - - H37Rv
I
199 C T
200
199 T T * start senX3
lepra
avium
H37Rv

and one of its potential binding sites. It is worth noting that the 42bp region identified in Himpens *et al.* (2000) does not properly cover this region and may be missing sequence essential for proper interactions.

To this end, 42 or 43 bp overlapping primer pairs were designed to cover the whole region between the 3' end of the *gpm1* gene and the 5' end of *senX3*, starting at the 42bp region identified in Himpens *et al.* (2000). The sequence and positions of the oligonucleotides can be seen in Figure 5.11 and Appendix 3. This size of oligonucleotide should be sufficiently large enough to contain an entire binding site, and the overlapping nature of the oligonucleotides means that no binding sites should be missed if they are disrupted by the ending of an oligonucleotide.

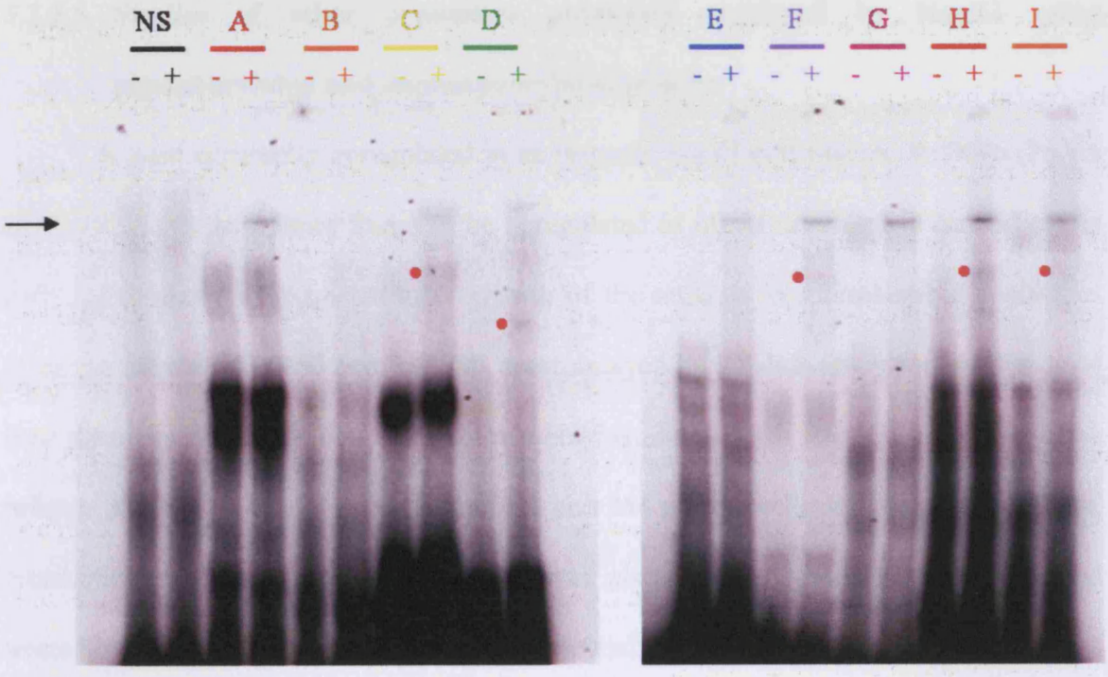
The nine primer pairs were annealed in 1:1 ratios and radiolabelled (see 2.6.2.4). Each primer pair was then used in a reaction with protein added and a reaction without protein added.

The autoradiographs show that an electromobility shift occurs in the majority of double stranded oligonucleotides pairs, with the faintest band occurring with oligonucleotide C and the strongest with oligonucleotide H (Figure 5.12). Second shifts seem to be occurring at a lower molecular weight of the complex (and therefore faster migration) with oligonucleotides C, D, F, H and I (marked with red spots).

5.2.2.4.3 Confirmation of the sequence-specific nature of the interaction of the recombinant RegX3 with oligonucleotide H

The cold-competition assay described in 5.2.2.2 was repeated here with oligonucleotide H. Increasing concentrations of unlabelled oligonucleotide H were used

Figure 5.12: Electrophoretic mobility shifts of double-stranded oligonucleotides covering the region between *gpm1* and *senX3*, and recombinant polyhistidine-tagged RegX3. NS, nonspecific oligonucleotide (annealed oligonucleotides 162 and 163, Appendix 3). 10µl of each binding reaction (described in 2.6.3.3) with or without protein was added to each well of an 8% precast TBE gel (Invitrogen) prerun with 0.5x TBE buffer, followed by electrophoresis at 100v for 40 min. After electrophoresis the gel was dried and autoradiographed. The letters A-I and their colours correspond with those in Figure 5.11. One potential shift is marked by an arrow on the left hand side of the image, which seems to occur with oligo regions A-I. Potential second shifts occurring in some of the lanes are marked on the image with red spots.


NS
A
B
C
D
E
F
G
H
I
- + - + - + - + - +
- + - + - + - + - +

to out-compete the interaction between labelled oligonucleotide H and the protein as can be seen from a reduction in intensity of the shift. Only high concentrations of nonspecific oligonucleotide (annealed oligonucleotides 162 and 163, Appendix 3), 50x more than oligonucleotide H, were able to slightly reduce the intensity of the upper shift; and had a similar effect on the lower shift (Figure 5.13).

5.2.2.5 Studies of other promoters putatively regulated by RegX3 using phosphorylated and unphosphorylated protein

A gene apparently upregulated in an isogenic *regX3* null mutant, Rv0096 (Parish *et al.* 2003) and two genes found to be upregulated in initial microarrays carried out in early, preliminary experiments after growth of the mutants in microaerobic conditions (data not shown), Rv1460 and Rv2780, were assayed by gelshift assay for the ability of their upstream regions to bind recombinant polyhistidine-tagged RegX3 (amplified using primers 237 and 238, 241 and 242, and 243 and 244 respectively, see Appendix 3). The recombinant polyhistidine-tagged RegX3 was also phosphorylated (see 2.6.2.9) and treated in the same manner to compare the potential binding of the protein-DNA complex in these two different states.

The results show that the electrophoretic mobility of radiolabelled DNA representing 200bp upstream of each of the genes could be retarded at a low level by binding to RegX3, since bands, although faint, were present in samples containing recombinant polyhistidine-tagged RegX3 that were not present in samples without protein (Figure 5.14). Phosphorylation of the recombinant polyhistidine-tagged RegX3 appeared to induce a reduction in the level of binding to DNA, although this may have

Figure 5.13: Autoradiograph of a cold-competition assay between recombinant polyhistidine-tagged RegX3 and labelled double-stranded oligonucleotide H, similar to that described in 5.2.2.2. The formation of a complex between the labelled probe and the protein is being outcompeted by either unlabelled oligonucleotide H or a nonspecific oligonucleotide (annealed oligonucleotides 162 and 163, Appendix 3) at 5, 10 and 50 times the concentration of the labelled probe. 10µl of each binding reaction (described in 2.6.3.3) with specific or nonspecific double stranded oligonucleotide was added to each well of an 8% precast TBE gel (Invitrogen) prerun with 0.5x TBE buffer, followed by electrophoresis at 100v for 40 min. After electrophoresis the gel was dried and autoradiographed. The potential shifts are marked by an arrow.

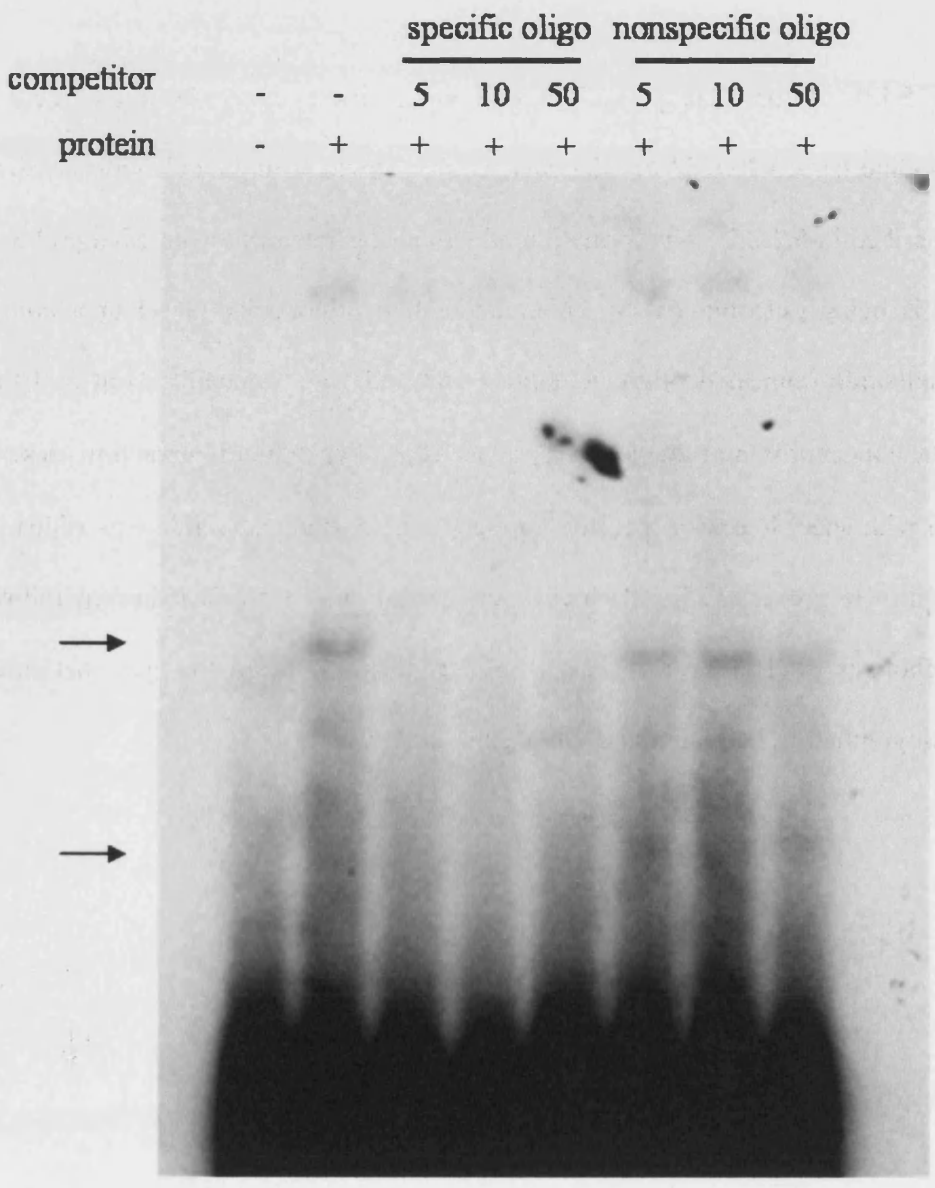
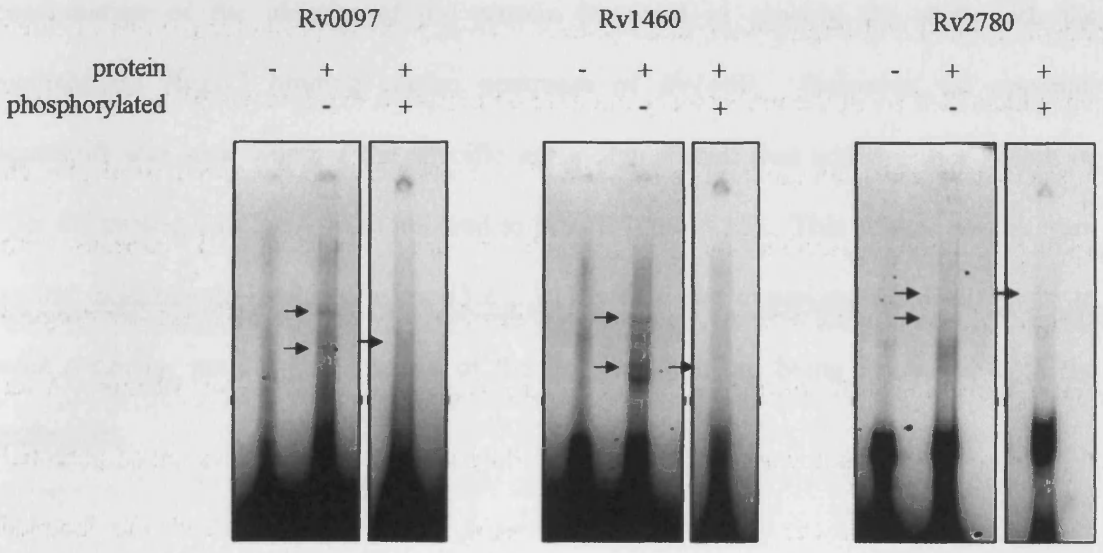

specific oligo
nonspecific oligo
competitor - - 5 10 50 5 10 50
protein - + + + + + + +

Figure 5.14: Autoradiograph of an electrophoretic mobility shift assay, using both phosphorylated and unphosphorylated protein, of the putative promoter regions of three genes identified by in the preliminary microaerobic growth experiments (data not shown) by microarray to be upregulated in *regX3* null mutants of *M. tuberculosis*. 10µl of each binding reaction (described in 2.6.3.3) with or without protein, and reactions with phosphorylated protein, was added to each well of an 8% precast TBE gel (Invitrogen) prerun with 0.5x TBE buffer, followed by electrophoresis at 100v for 40 min. After electrophoresis the gel was dried and autoradiographed. The arrows mark potential interactions between the recombinant polyhistidine-tagged RegX3 and the labelled probes that do not occur in protein negative controls.


Rv0097
Rv1460
Rv2780
protein
phosphorylated
- + +
- +

been due to a reduction in the quantity of protein added to those DNA-protein binding reactions as a result of the phosphorylation reaction. However, in all three promoter regions examined the phosphorylation of the recombinant polyhistidine-tagged RegX3 seemed to reduce the number of shifts apparent from two to one.

5.2.2.5.1 Supershift assay

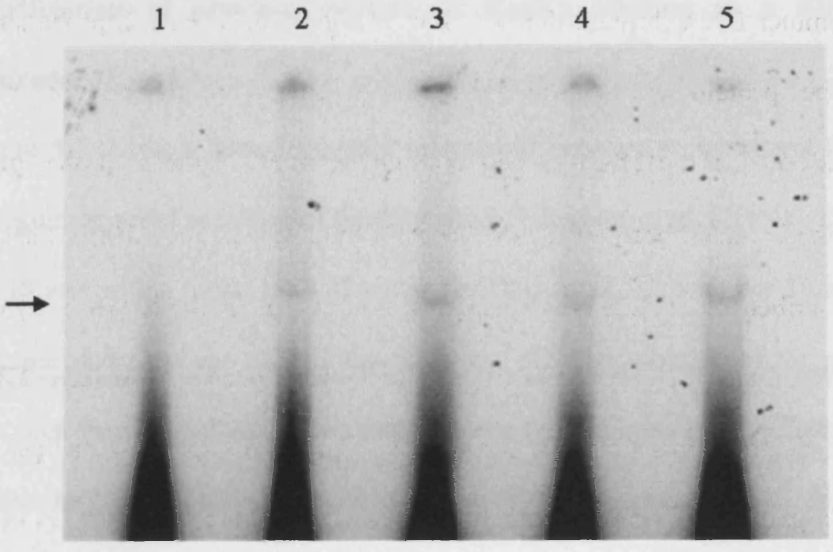
Further retardation of the DNA-protein shift by a specific antibody allows confirmation of the identity of the protein involved in causing the shift with the recombinant RegX3 binding region upstream of *Rv1460*. However, no apparent supershift was seen whether the specific anti-polyhistidine was added either before or after the protein and DNA were allowed to bind (Figure 5.15). This echoes results from similar experiments in this chapter (5.2.2.3.1), where the experiment did not seem to work properly, possibly as a result of the polyhistidine tag being inaccessible to the antibodies.

5.3 DISCUSSION

5.3.1 Production and purification of recombinant RegX3

It has been shown that it is possible to express recombinant polyhistidine-tagged RegX3 in a soluble form from *E. coli*, in contrast with a previous report (Himpens *et al.* 2000). Although it was only possible to purify 0.137mg of protein from a 1 litre culture, it was more than sufficient to conduct experiments similar to those described previously (Himpens *et al.* 2000), and observe an electrophoretic mobility shift of a specific DNA probe. Since the protein was purified under non-denaturing conditions, it is more likely

Figure 5.15: Autoradiograph of a supershift assay, where specific antipolyhistidine or nonspecific anti-GST antibody is used to increase the apparent molecular mass of the DNA-protein complex, further retarding the electrophoretic mobility of the labelled Rv1460 promoter DNA through the gel. Lane 1, probe only, lane 2, probe and protein, lane 3, probe, protein and nonspecific antibody, lane 4, probe, protein and specific antibody added after the DNA-protein binding reaction. Lane 5, probe, protein and specific antibody added before the DNA-protein binding reaction. 10μl of each binding reaction (described in 2.6.3.3) with specific or nonspecific antibody (see text) was added to each well of an 8% precast TBE gel (Invitrogen) prerun with 0.5x TBE buffer, followed by electrophoresis at 100v for 40 min. After electrophoresis the gel was dried and autoradiographed. The experiment was conducted in the same way as in 5.2.2.3.1. The potential shift is marked by an arrow, and seems to show no alteration of electrophoretic mobility when specific antibody is added.


1
2
3
4
5

to be folded correctly than protein that has been purified under denaturing conditions and renatured by dialysis. It may be possible to increase the yield of the recombinant RegX3 for use in future experiments by varying the imidazole concentration of the buffer used to wash the column in 2.6.2.6 and 0, or by extracting the protein from a larger volume of starter culture, however this was felt to be unnecessary for the current study.

5.3.2 Clarification of previous reports of RegX3 binding to a 42bp region upstream of *senX3*

Figure 5.6 shows a dose-dependant interaction between recombinant RegX3 and the 42 bp region reported as a RegX3 binding site in Himpens *et al.* (2000). The greatest intensity shift was with a molar ratio of protein to DNA of 7.7:1, whereas Himpens *et al.* (2000) indicate that they use half of this amount. The appearance of the shift in the experiment seen here was sharper than those shown by Himpens *et al.* (2000), however the only data they show on this subject is with a 189bp fragment, which in my hands retarded its own passage through the gel in the absence of protein (data not shown). This was probably due to possible loops and secondary structure in the sequence which interfered with interpretation of the results. This spurious banding pattern was not evident in Himpens *et al.* (2000).

The evidence for the binding of recombinant RegX3 to the 42 bp fragment upstream of *senX3* in Himpens *et al.* (2000) is based on ability of the unlabelled 42 bp fragment to displace the binding of the labelled 189 bp fragment, although they do not show this data. Since they also do not show retardation of the labelled 42bp region it is impossible to directly compare images with those published.

Himpens *et al.* (2000) also show that the interaction between recombinant RegX3 and the upstream region of *senX3* can be outcompeted by specific, but not nonspecific, unlabelled DNA probes. The unlabelled nonspecific DNA probe used in that study was poly-dIdC DNA, a highly branched bulk carrier DNA which does not represent a very good negative control since its properties are so unlike those of the binding, specific DNA probe. It is generally used (as it is here) to reduce the levels of non-specific interaction between protein and probe in the reactions. The evidence shown here in Figure 5.7 represents more direct evidence for a sequence-specific interaction of the 42bp region upstream of *senX3* with recombinant RegX3.

Unfortunately it has not been possible to show interaction of an anti-polyhistidine antibody with the shift complex. The 'smear' of protein shown in the results of the shift-blot in Figure 5.10 suggest that the interaction between the protein and DNA is a brief one, where dissociation and reassociation occurs quickly as the complex moves through the gel during electrophoresis. In this case protein would be left in the matrix of the gel as the complex migrates, and its subsequent blotting and binding with the anti-His antibody would leave the 'smear' marks apparent in Figure 5.10. However, a positive control was not included in this experiment and may illuminate any technical difficulties if included in further experiments. The results of Figure 5.10 confirm that the anti-polyhistidine antibody will still bind to the recombinant RegX3 after electrophoresis, suggesting that the lack of interaction between the antibodies and recombinant RegX3 in Figures 5.8, 5.9 and 5.15 is due to the transitory nature of the interaction between the recombinant RegX3 and the DNA.

Overall the results in section 5.2.2.1 - 5.2.2.3 prove and clarify the sequence-specific interaction between RegX3 and a 42bp region upstream of *senX3* suggested by Himpens *et al.* (2000).

5.3.3 Mapping the interactions of recombinant RegX3 with the region upstream of *senX3*

The nature of the evidence for binding of RegX3 to the 42bp region by Himpens *et al.* (2000) relies on the ability of the 42bp region investigated above to out-compete the interaction of recombinant RegX3 with a larger labelled fragment. This does not preclude the possibility that the 42 bp fragment contains only part of the binding site for RegX3, and is blocking the interaction between RegX3 and the full binding site contained on the 189 bp larger fragment (see Figure 5.11). Indeed, an alignment of the putative *senX3* promoter regions from three closely related mycobacterial species shows that potential regulatory motifs as regions of homology that occur throughout the region. The site of the 42 bp fragment studied in Himpens *et al.* (2000) and in this study does not contain any obvious isolated regulatory motif and we can assume that the proof of interaction by blocking binding of the larger fragment does not necessarily mean that the sole binding motif lies within this 42 bp region.

To examine the interactions of recombinant RegX3 with the potential *senX3* promoter region in a more systematic fashion, overlapping double stranded oligonucleotides that cover the region from the stop codon of *gpm1* to the start codon of *senX3* were designed and assayed for their ability to bind recombinant RegX3 (Figure 5.12). Surprisingly, the addition of RegX3 seemed to cause a shift in the majority of

oligonucleotides, with the strongest interaction occurring with oligonucleotide H, which is positioned immediately 5' of the *senX3* start codon.

A perfunctory search for homologous sequences in these regions did not show any motifs that the oligonucleotide regions had in common (data not shown). It might be reasonable to assume that the recombinant RegX3 is interacting with these regions in a non-specific manner. To test this hypothesis, a cold-competition assay of the interaction of recombinant RegX3 with oligonucleotide H was performed, and showed that the interaction could be out-competed by unlabelled specific oligonucleotide H, but not by unlabelled nonspecific oligonucleotide (annealed oligonucleotides 162 and 163, Appendix 3)(Figure 5.13). There was a slight reduction in the intensity of the shift when the binding of oligonucleotide H and recombinant RegX3 was challenged with a 50-fold excess of nonspecific unlabelled probe. Taken together, these results suggest that RegX3 can bind to different sequences specific for the upstream region of *senX3* with an affinity at most 50-fold greater than for nonspecific sequences.

5.3.4 Analysis of recombinant RegX3 binding of putative regulatory regions for genes upregulated by microarray

One gene, Rv0098, from previous work on an isogenic *regX3* null mutant (Parish *et al.* 2003) and two genes from this work, Rv1460 and Rv2780, were identified to be upregulated by microarray analysis in *regX3* null mutants in various conditions. Rv0096 belongs to the PPE family of genes, whereas *Rv1460* is annotated as a probable regulatory protein. *Rv2780* codes for alanine dehydrogenase, an enzyme that catalyzes the reversible deamination of L-alanine to pyruvate.

All three genes are likely to contain a promoter site at their 5' end since none of them appear to be co-transcribed with genes immediately upstream of them, and have clear intervals before their transcriptional start site (Figures 5.16, 5.17 and 5.18).

It was considered likely that the regulatory elements for these genes would occur within the 200 bp immediately upstream from each start codon. The gelshift results with both phosphorylated and unphosphorylated recombinant polyhistidine-tagged RegX3 show a possible weak interaction of the unphosphorylated protein with the promoters of all three genes. However, these interactions appeared not to occur when the protein was phosphorylated. If there is an interaction of RegX3 with the promoter regions for these genes then it could be hypothesised that *in vivo*, the phosphorylation of RegX3 could cause it to unbind from the DNA and allow transcription. This hypothesis would be consistent with the apparent upregulation of these genes in mutant strains of *M. tuberculosis* in the preliminary microaerobic microarray work (data not shown, see also Parish *et al.* (2003)) where the absence of RegX3 is associated with increased transcription of these genes. Activation of genes by the kinase-dependant lifting of gene repression through unbinding of a regulatory protein is a common mechanism of genetic control and occurs in the ArcB-ArcA system of *E. coli* (Georgellis *et al.* 2001).

The implications of these results are that RegX3 may regulate the expression of diverse proteins, and that this interaction may be phosphorylation-dependant. Potential binding of recombinant RegX3 to the region upstream of *Rv2780* (alanine dehydrogenase) suggests that RegX3 may have a role in controlling part of the intermediary metabolism of *M. tuberculosis*. The potential regulation by RegX3 of a likely regulatory protein (Rv2780) suggests that the SenX3-RegX3 two component signal

Figure 5.16: Diagram showing the region surrounding the start site of the coding sequence of Rv0096 on the chromosome of *M. tuberculosis*.

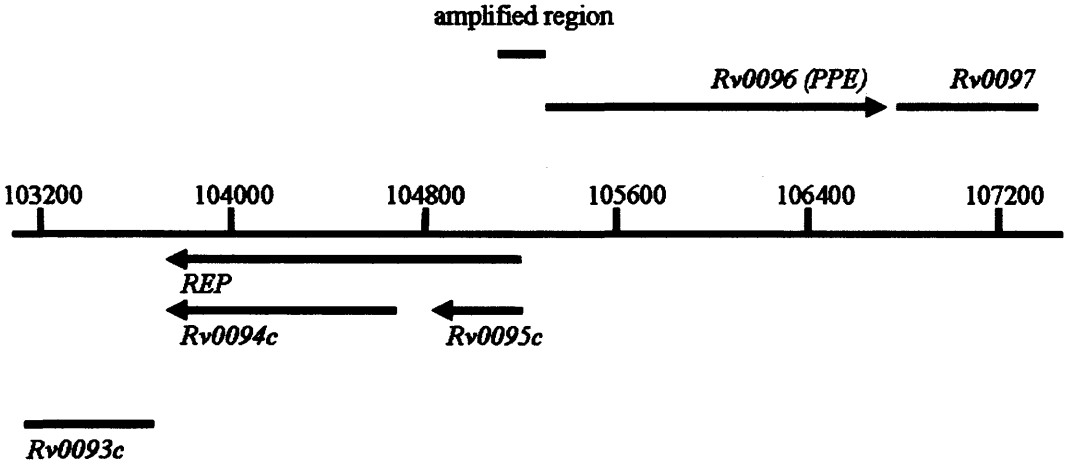

amplified region
Rv0096 (PPE)
Rv0097
103200
104000
104800
105600
106400
107200
REP
Rv0094c
Rv0095c
Rv0093c

Figure 5.17: Diagram showing the region surrounding the start site of Rv2780 on the chromosome of *M. tuberculosis*.

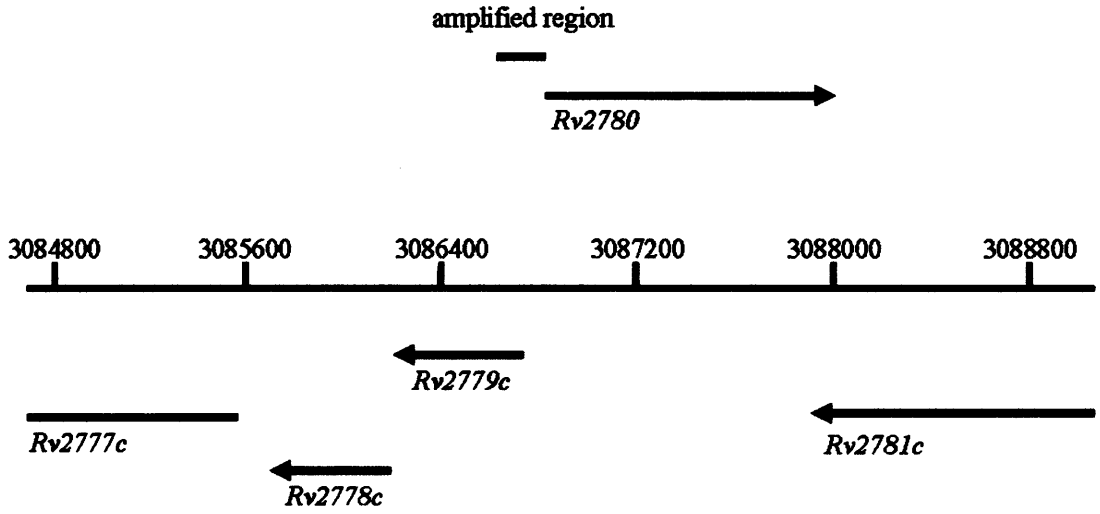
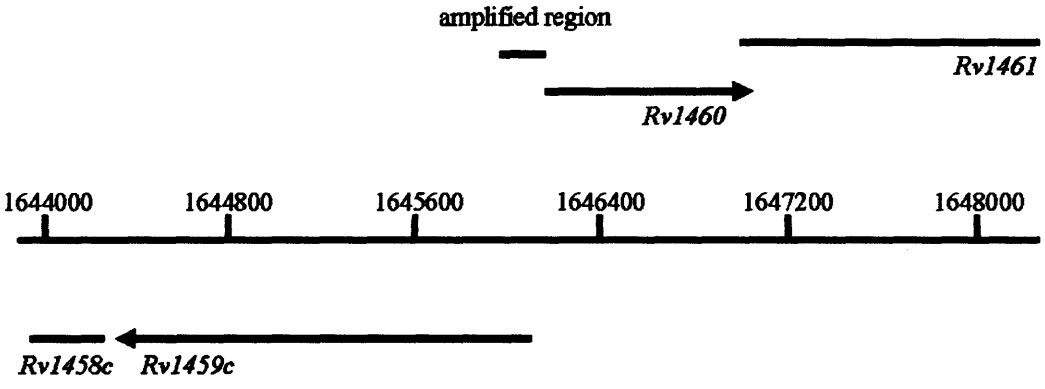

amplified region
Rv2780
3084800
3085600
3086400
3087200
3088000
3088800
Rv2779c
Rv2777c
Rv2781c
Rv2778c

Figure 5.18: Diagram showing the region surrounding the start site of Rv1460 on the chromosome of *M. tuberculosis*.


amplified region
Rv1461
Rv1460
1644000
1644800
1645600
1646400
1647200
1648000
Rv1458c
Rv1459c

transduction system may be part of a signal transduction network of regulators. Similarly, if the expression of *Rv0096* (a PPE gene) is being regulated by RegX3, then this indicates a role for the SenX3-RegX3 two component signal transduction system in the avoidance of the host immune response by antigenic variation (Voskuil *et al.* 2004).

5.3.5 Evidence for two DNA binding sites on RegX3?

In this study, but not in Himpens *et al.* (2000), two potential shifts were observed. These were seen in the initial experiments with the 42bp sequence identified in that paper, with some of the oligonucleotides used for mapping the region upstream of *senX3*, and crucially with the longer upstream regions of promoters of other genes used in the phosphorylation assays. This is indicative of RegX3 containing two binding sites for DNA, with the upper band representing RegX3 complexed with two molecules of double stranded probe, and the lower band representing RegX3 complexed with a single molecule of double stranded probe. The two sites may be specific for different sequences, which might explain the promiscuous binding of the protein to a number of sites upstream of *senX3*, and may have different affinities for DNA, which would explain why the upper shift appears to fade first in the cold-competition assays.

5.3.6 Summary

Overall, the results presented suggest that RegX3 may bind to its own promoter and regulate its own expression. Furthermore, systematic analysis of the promoter region of *senX3* has shown that RegX3 may bind at multiple sites in a sequence-specific manner. It has been shown that the nature of RegX3 binding is phosphorylation-dependant, and

studies of promoter regions of three other genes suggest that RegX3 may be a global regulator, controlling a variety of processes within *M. tuberculosis*.

6 Analysis of the contribution of Rv0465c and Rv1049 to the pathogenesis of *M. tuberculosis* through the isolation of null mutations

6.1 INTRODUCTION

6.1.1 Identification of Rv0465c and Rv1049 as genes potentially involved in the oxidative stress response of *M. tuberculosis*

Both Rv0465c and Rv1049 appeared to be highly upregulated in microarray studies after stress of the wild-type *M. tuberculosis* by an organic hydroperoxide generator, cumene hydroperoxide (Rickman 2002). Since it is feasible to consider that these genes may be involved in the response of *M. tuberculosis* to oxidative stress, the potential properties of these genes were investigated by homology and literature searching.

6.1.2 Identification of Rv0465c as a potential hydroperoxide stress-responsive gene

The predicted open reading frame of the Rv0465c peptide indicates that the protein is 474 amino acids long and annotated as a probable transcriptional regulatory protein in Tuberculist (2005). It has been suggested as a non-essential gene as defined by *Himar*-1 transposon mutagenesis (Sassetti *et al.* 2003), where a library of transposon-based mutants were screened for their ability to grow *in vitro*.

The Pfam domain prediction (Figure 6.1) shows that Rv0465c appears to contain a helix-turn-helix domain (HTH 3) between residues 10 and 64; this domain is part of a large family of DNA binding proteins. Pfam also predicts a domain of unknown function (DUF955) between residues 187 and 310, suggesting similarity to metalloproteases. Metalloprotease domains often use a divalent cation in the active site to activate a water molecule and hydrolyse peptide bonds. Rv0465c lies directly upstream of Rv0464c, with which it is likely to be co-transcribed

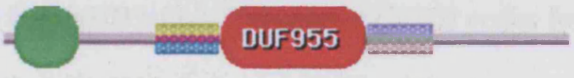
Figure 6.1 Rv0465c and Rv1049. The protein sequences (A) for each gene were taken from Tuberculist (http://genolist.pasteur.fr/TubercuList/). The diagrams for each protein (B) were obtained by submitting the sequences of each protein to the Pfam server (http://www.sanger.ac.uk/Software/Pfam/)

Rv0465c

A

```
1 -   VSKTYVGSRV RQLRNERGFS QAALAQMLEI SPSYLNQIEH DVRPLTVAVL LRITEVFGVD
61 -  ATFFASQDDT RLVAELREVT LDRDLDIAID PHEVAEMVSA HPGLACAVVN LHRRYRITTA
121 - QLAAATEERF SDGSGRGSIT MPHEEVRDYF YQRQNYLHAL DTAAEDLTAQ MRMHHGDLAR
181 - ELTRRLTEVH GVRINKRIDL GDTVLHRYDP ATNTLEISSH LSPGQQVFKM AAELAYLEFG
241 - DLIDAMVTDG KFTSAESRTL ARLGLANYFA AATVLPYRQF HDVAENFRYD VERLSAFYSV
301 - SYETIAHRLS TLQRPSMRGV PFTFVRVDRA GNMSKRQSAT GFHFSSSGGT CPLWNVYETF
361 - ANPGKILVQI AQMPDGRNYL WVARTVELRA ARYGQPGKTF AIGLGCELRH AHRLVYSEGL
421 - DLSGDPNTAA TPIGAGCRVC ERDNCPQRAF PALGRALDLD EHRSTVSPYL VKQL (474aa)
```

B



Rv1049

A

```
1 -   MGKGAAFDEC ACYTTRRAAR QLGQAYDRAL RPSGLTNTQF STLAVISLSE GSAGIDLTMS
61 -  ELAARIGVER TTLTRNLEVM RRDGLVRVMA GADARCKRIE LTAKGRAALQ KAVPLWRGVQ
121 - AEVTASVGDW PRVRRDIANL GQAAEACR (148aa)
```

B

6.1.3 Identification of Rv1049 as a potential hydroperoxide stress-responsive gene

The predicted sequence of the Rv1049 protein of *M. tuberculosis* was identified to have strong homology to the organic hydroperoxide resistance regulator (OhrR) proteins of *Xanthomonas campestris* and *Bacillus subtilis* by BLAST comparison (E. Davis, personal communication). Rv1049 is also defined as a non-essential gene after *Himar*-1 transposon-based mutagenesis of wild-type *M. tuberculosis* by Sassetti *et al.* (2003). Rv1049 is located directly upstream of Rv1050, and these genes may be co-transcribed.

6.1.4 *in silico* analysis of the Rv1049 predicted protein sequence

The predicted ORF for Rv1049 from Tuberculist (2005) codes for a 148 amino-acid protein. PSI-BLAST analysis (Altschul *et al.* 1997) of the sequence followed by a conserved domain search (Marchler-Bauer & Bryant 2004) shows that Rv1049 contains a helix-turn-helix MarR domain from amino acids 17 to 134. Rv1049 therefore belongs to the MarR family of transcriptional regulators, which among other things are implicated in multiple antibiotic resistance (Grkovic *et al.* 2002). The *marR* operon of *E. coli* can also be induced by the redox cycling agents plumbagin and menadione (Seoane & Levy 1995; Alekshun & Levy 1999), and its transcriptional mechanisms have been comprehensively examined (Martin & Rosner 2004).

6.1.5 Previous studies on OhrR in *Xanthomonas campestris* and *Bacillus subtilis*

OhrR from *X. campestris* is an organic peroxide-inducible negative regulator of genes responsible for orchestrating the resistance of that organism to organic hydroperoxides that arise from oxidative stress. It is induced by *tert*-butyl hydroperoxide

and cumene hydroperoxide, both of which generate organic hydroperoxide radicals, and only weakly by peroxide itself. Other byproducts of oxidants and metabolites of organic peroxide metabolism did not induce the expression of an OhrR-responsive promoter in *X. campestris* (Panmanee *et al.* 2002), and OhrR may be responding to *tert*-butyl hydroperoxide stress only, in *B. subtilis* (Helmann *et al.* 2003). More recently, Klomsiri *et al.* (2005) also showed phenotypic evidence for a role of OhrR in the resistance of *X. campestris* to lipid hydroperoxides.

It seems from the published literature that the oxidation of a single cysteine residue in the OhrR protein is essential for the function of OhrR in both *B. subtilis* and *X. campestris* (Fuangthong & Helmann 2002; Panmanee *et al.* 2002). There is an equivalent cysteine residue present in Rv1049.

In *X. campestris*, *tert*-butyl hydroperoxide has been used for induction of promoters and operons (Sukchawalit *et al.* 2001; Mongkolsuk *et al.* 2002; Panmanee *et al.* 2002) *in vivo*, whereas cumene hydroperoxide and *tert*-butyl hydroperoxide were used for examination of the activity of purified OhrR and *lacZ* promoter-fusion strains of *X. campestris* containing mutated *ohrR* complementing alleles (Fuangthong & Helmann 2002).

In *B. subtilis*, *tert*-butyl hydroperoxide and cumene hydroperoxide was used to demonstrate either induction of Ohr-region genes, or was used in a disk diffusion assay to demonstrate inhibition of growth in various *B. subtilis* mutant strains (Fuangthong *et al.* 2001). It is worth consideration that in these studies, cumene hydroperoxide-stressed cultures are compared with untreated controls. As has previously been mentioned (1.4.3), cumene hydroperoxide is supplied dissolved in cumene, which itself is likely to be toxic

to bacteria. Furthermore cumene hydroperoxide is not readily soluble in water, which may cause uneven distribution of the agent throughout the bacterial culture during any experiments.

6.1.6 A possible role for Rv0465c and Rv1049 in the oxidative stress response of *M. tuberculosis*?

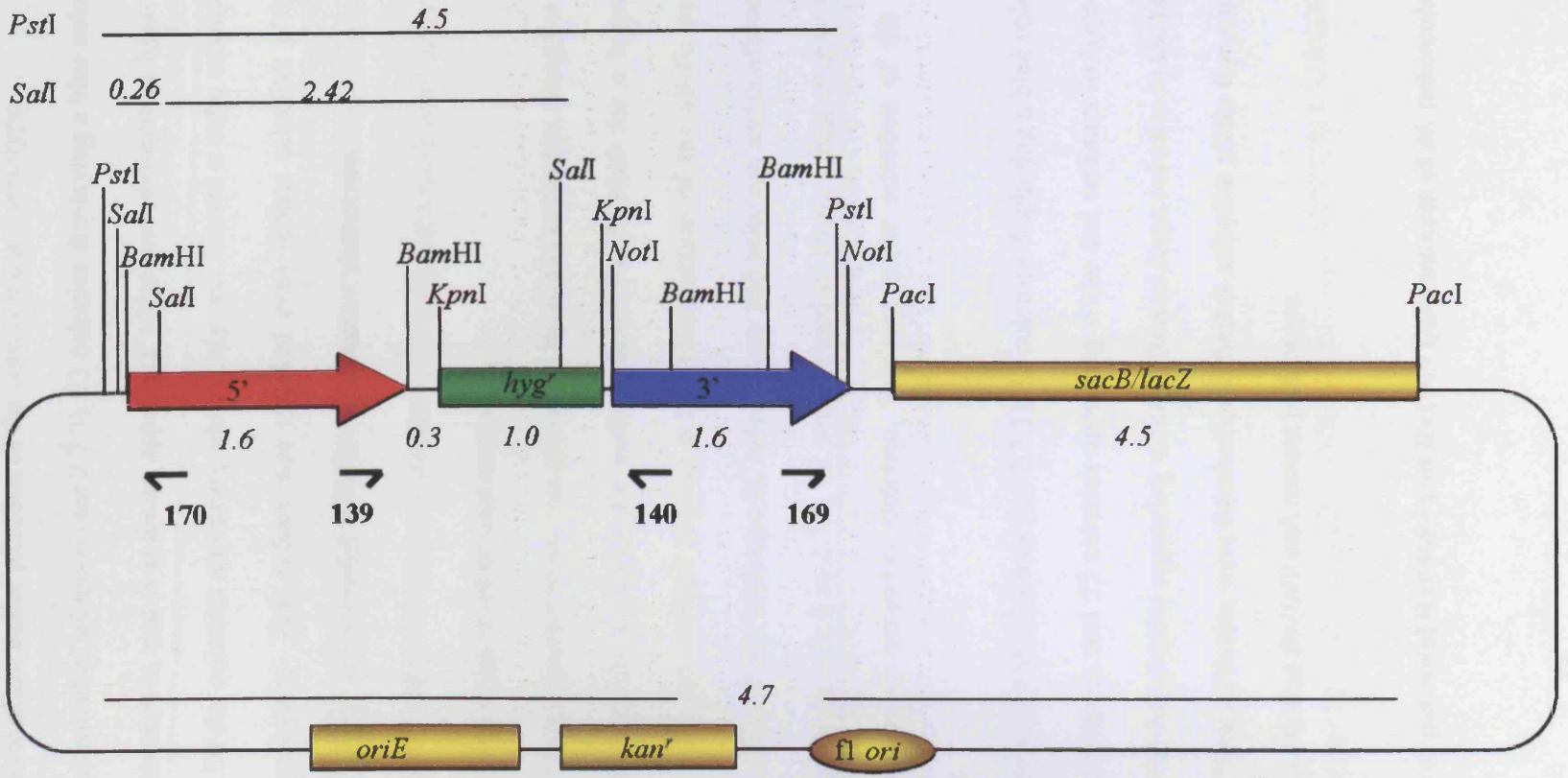
It was hypothesised that if Rv0465c and Rv1049 were playing roles in the detoxification of organic hydroperoxides, then the Rv0465c and Rv1049 null mutant strains of *M. tuberculosis* would be more sensitive to killing by these agents than the wild-type *M. tuberculosis*. In addition, if the proteins encoded by Rv0465c and Rv1049 were playing roles in the infection process, then one might expect to see mutants strains deficient in these genes to behave differently from the wild-type *M. tuberculosis* during infection. The work described in this chapter attempts to further the understanding of the roles of these genes in the oxidative stress response of *M. tuberculosis* by characterising the phenotypes of null mutants of these genes.

6.2 RESULTS

6.2.1 Construction of an Rv0465c null mutant.

An Rv0465c null mutant strain was generated by allelic exchange, using the p2NIL suicide delivery vector to substitute the coding region for Rv0465c for a hygromycin resistance allele (Figure 6.2). Construction of the allelic exchange vector, electroporation of the wild-type *M. tuberculosis* and selection and confirmation of the mutant strains is discussed in 2.7. It was decided hygromycin-

Figure 6.2: The completed Rv0465c knockout construct, showing the size of the *Pst*I and *Sal*I restriction fragments of the gene used to confirm the orientation of the 5' and 3' fragments, the locations of the various restriction sites mentioned in the text, the sizes of the plasmid features, and the locations of the primers used to confirm the orientation of the 5' and 3' inserts by sequencing. Red colour, amplified region 5' of the gene to be knocked out; blue colour, amplified region 3' of the gene to be knocked out; green colour, hygromycin resistance marker; yellow colour, other selectable markers.


PstI
4.5
SalI
0.26
2.42
PstI
SalI
BamHI
SalI
BamHI
KpnI
SalI
KpnI
NotI
BamHI
BamHI
PstI
NotI
PacI
PacI
5'
hyg^r
3'
sacB/lacZ
1.6
0.3
1.0
1.6
4.5
170
139
140
169
4.7
oriE
kan^r
f1 ori

pDH19 14kb

mark the mutation to allow easier selection of the correct mutant genotype.

At the first round of selection (see 2.7.3), 13 colonies displaying a blue colour and resistance to hygromycin, and 8 colonies appearing white and resistant to hygromycin were observed. Of the colonies appearing white, only one formed a large colony in the presence of hygromycin. This colony was streaked onto further selective plates that confirmed the presence of a potential Rv0465c null mutant phenotype.

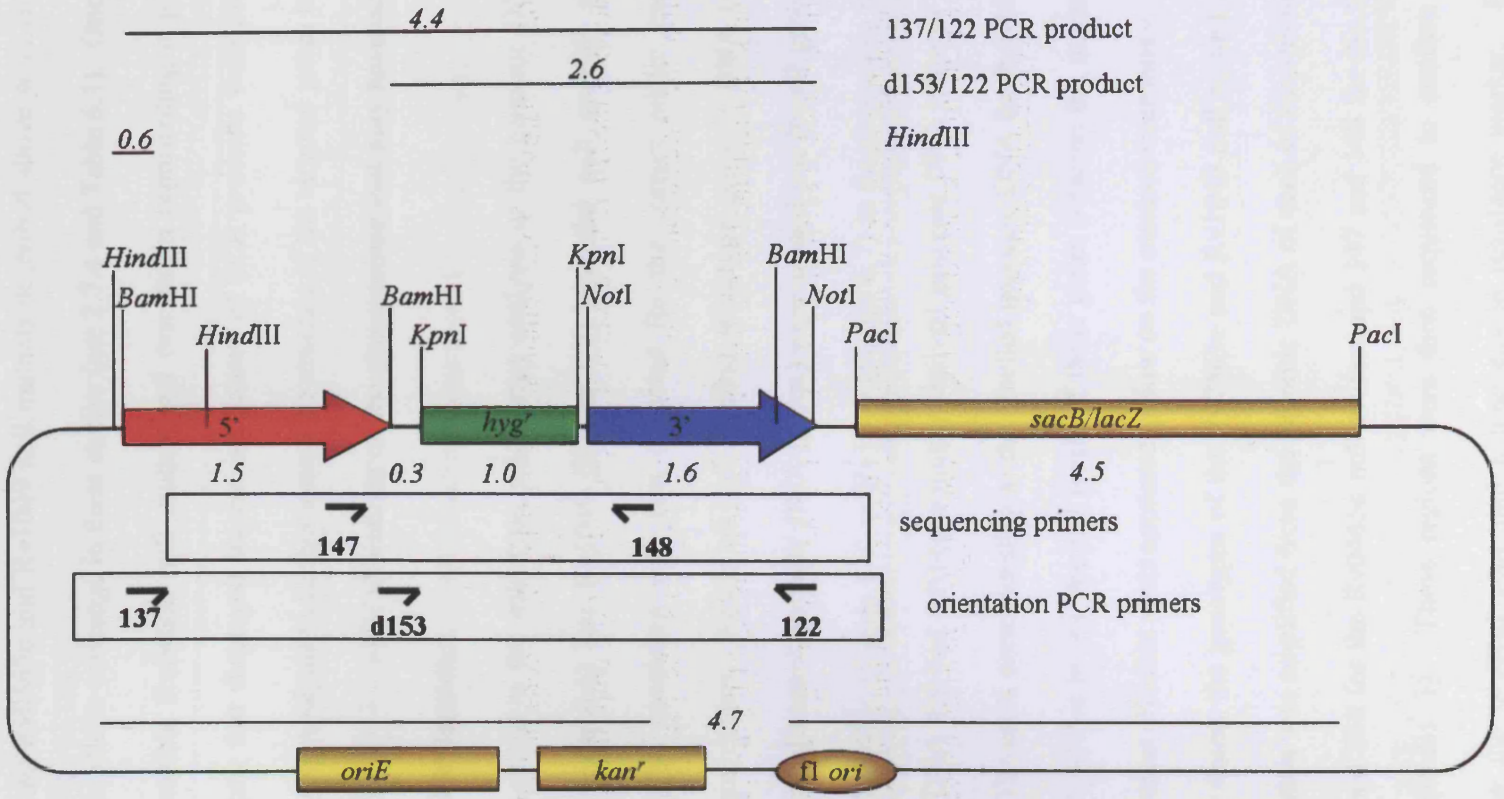
6.2.2 Construction of an Rv1049 null mutant.

An Rv1049 null mutant strain was generated by allelic exchange, using the p2NIL suicide delivery vector to substitute the coding region for Rv1049 for a hygromycin resistance allele to create pDH14 (Figure 6.3). Construction of the allelic exchange vector, electroporation of the wild-type *M. tuberculosis* and selection and confirmation of the mutant strains is discussed in 2.7.2. It was decided to create a hygromycin-marked mutation, rather than an unmarked deletion, to allow easier selection of the correct mutant genotype.

At the first round of selection (see 2.7.3), 13 colonies displaying a blue colour and resistance to hygromycin, and 15 colonies appearing white and resistant to hygromycin were observed. Of the colonies appearing white, 5 formed large colonies in the presence of hygromycin. These colonies were streaked onto further selective plates that confirmed the presence of a potential *Rv1049* null mutant phenotype.

6.2.3 Analysis of potential Rv0465c and Rv1049 null mutants of *M. tuberculosis*.

Figure 6.3: Diagram of the completed Rv1049 knockout construct, showing the size and location of the *Hind*III restriction fragment and the PCR amplification products used to confirm the orientation of the 5' and 3' fragments, the locations of the various restriction sites mentioned in the text, the sizes of the plasmid features, and the locations of the primers used to confirm the orientation of the 5' and 3' inserts by sequencing. Red colour, amplified region 5' of the gene to be knocked out; blue colour, amplified region 3' of the gene to be knocked out; green colour, hygromycin resistance marker; yellow colour, other selectable markers.


4.4
137/122 PCR product
2.6
d153/122 PCR product
0.6
HindIII
HindIII
BamHI
HindIII
BamHI
KpnI
KpnI
NotI
BamHI
NotI
PacI
PacI
5'
hygr
3'
sacB/lacZ
1.5
0.3
1.0
1.6
4.5
147
148
sequencing primers
137
d153
122
orientation PCR primers
4.7
oriE
kanr
f1 ori

pDH14 14kb

The potential Rv0465c and Rv1049 null mutants described above were analysed by PCR to confirm allelic exchange in these strains (see 2.7.4 and Table 6.1). One strain representing a possible Rv0465c null mutant and one strain representing a possible Rv1049 null mutant that displayed the correct pattern of PCR products were analysed further by microarray analysis to further assess the absence of the targeted genes in these strains and look for other genetic alterations on the chromosome that may have occurred by illegitimate recombination.

When compared to the wild-type, microarray analysis of the genomic DNA for both Rv0465c (Figure 6.4) and Rv1049 (Figure 6.5) potential null mutants showed hybridisation profiles consistent with that expected for the correct allelic exchange mutants of *M. tuberculosis*. Ratios of Cy5-labelled wild-type genomic DNA (red) to Cy3-labelled potential mutant genomic DNA (green) were around one for all genes, with the exception of Rv0465c (Figure 6.6) or Rv1049 (Figure 6.7) in the respective mutant-derived genomic DNA that had wild-type:mutant-derived genomic DNA ratios of 58 and 14 respectively. The spots corresponding to the labelled genomic DNA for Rv0465c and Rv1049 appeared red due to the unequal levels of these genes between the mutants and the wild-type, whereas all other spots appeared yellow on the scanned microarray.

To further assess the genotypes of the Rv0465c and Rv1049 null mutant strains, the region of deletion was amplified from the genomic DNA of each of the strains using the primers 139 and 140 for the Rv0465c null mutant and 147 and 148 for the Rv1049 null mutant (Appendix 3). These regions were then sequenced to confirm allelic exchange between the targeted gene and a hygromycin resistance marker. BLAST searches were used to confirm the deleted regions of the two strains.

Table 6.1 Expected product sizes of primer pairs designed to distinguish between the genotypes of strains potentially arising during the knockout process to construct Rv0465c and Rv1049 insertionally inactivated mutants. The letters denoting the different primer pairs refer to Figure 2.7.

primer pair	product		pDH14::*M. tuberculosis*		Rv0465c::*M. tuberculosis*	
			primer	product size (kb)	primer	product size (kb)
A	intergenic *hyg*^{r}	F	181	0.659	181	0.659
		R	182		182	
B	intergenic	F	183	0.287	185	0.627
	Rv1049/Rv0465c	R	184		186	
C	chromosome	F	177	1.949 (5' crossover)	179	2.067 (5' crossover
	upstream of hygr	R	157	>14.5 (3' crossover)	157	>16 (3'crossover)
D	chromosome	F	156	>14.8 (5' crossover)	156	>15.5 (5' crossover)
	downstream of hygr	R	178	1.842 (3' crossover)	180	1.803 (3' crossover)
E	region between the	F	147	1.243 (*hyg*^{r})	139	1.361 (*hyg*^{r})
	cloned flanking sequence	R	148	571 (*Rv1049*)	140	1.485 (*Rv0465c*)

Figure 6.4: Adapted GENEPIX 5.0 screenshot of the microarray analysis of the potential Rv0465c null mutant. Cy3-labelled cDNA (shown as green) derived from the potential Rv0465c null mutant was hybridised against Cy5-labelled cDNA (shown as red) derived from wild-type *M. tuberculosis*. Genes that are present in both the wild-type and the potential Rv0465c null mutant appear as yellow, and red spots indicate genes that are present only in the wild-type and not in the mutant. The control spots are generally used for normalisation in RNA-based microarray experiments and locating the corners of each grid, and are not relevant to this experiment. The apparent green spot toward the top left of the image is likely to be a technical artefact, as it is not seen in the numerical analysis in Figure 6.6. For further discussion of genomic DNA microarrays see 2.5.8.1.

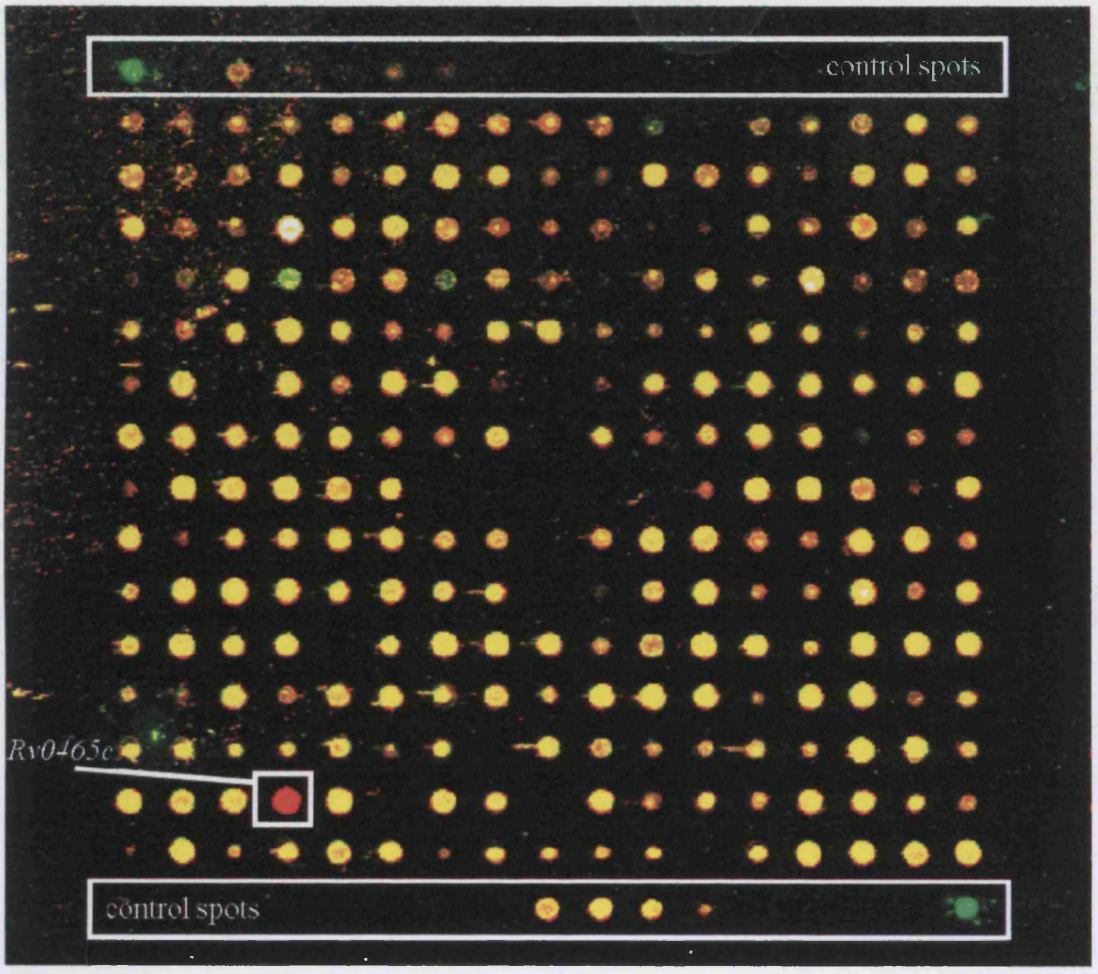

control spots
Rv0465c
control spots

Figure 6.5: Adapted GENEPIX 5.0 screenshot of the microarray analysis of the potential Rv1049 null mutant. Cy3-labelled cDNA (shown as green) derived from the potential Rv1049 null mutant was hybridised against Cy5-labelled cDNA (shown as red) derived from wild-type *M. tuberculosis*. Genes that are present in both the wild-type and the potential Rv1049 null mutant appear as yellow, and red spots indicate genes that are present only in the wild-type and not in the mutant. The control spots are generally used for normalisation in RNA-based microarray experiments and locating the corners of each grid, and are not relevant to this experiment. The apparent green spot toward the top left of the image is likely to be a technical artefact, as it is not seen in the numerical analysis in Figure 6.7. For further discussion of genomic DNA microarrays see 2.5.8.1.

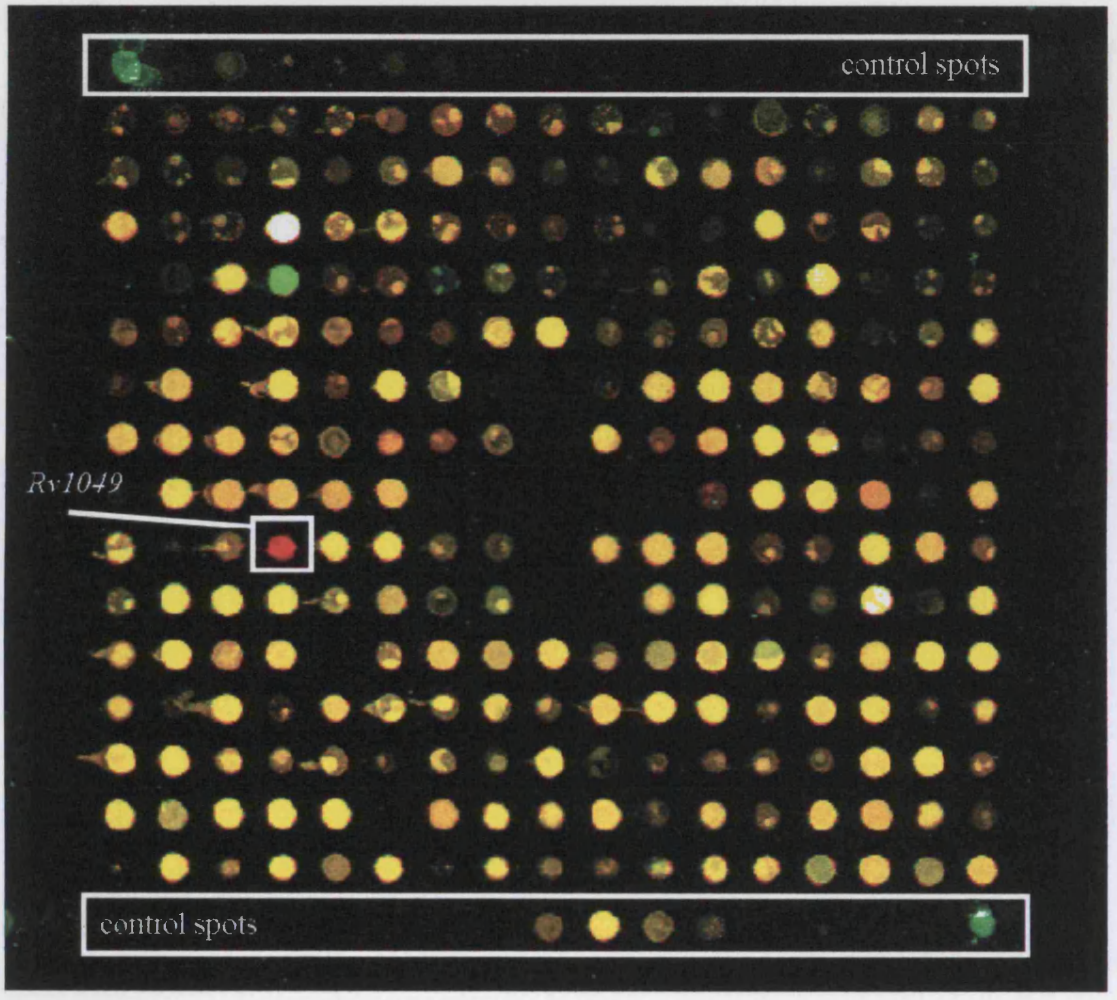

control spots
Rv1049
control spots

Figure 6.6: Adapted GENEPIX 5.0 scatter plot of a microarray comparing the DNA of wild-type *M. tuberculosis* to that of a potential Rv0465c null mutant. Part of this microarray is shown in Figure 6.2 above. The features of the microarray are plotted sequentially along the x-axis and compared to the median of the Cy5 values divided by the median of the Cy3 values for each spot. A median of ratios above 1 represents a gene that is missing in the Cy3-labelled cDNA from the potential Rv0465c null mutant.

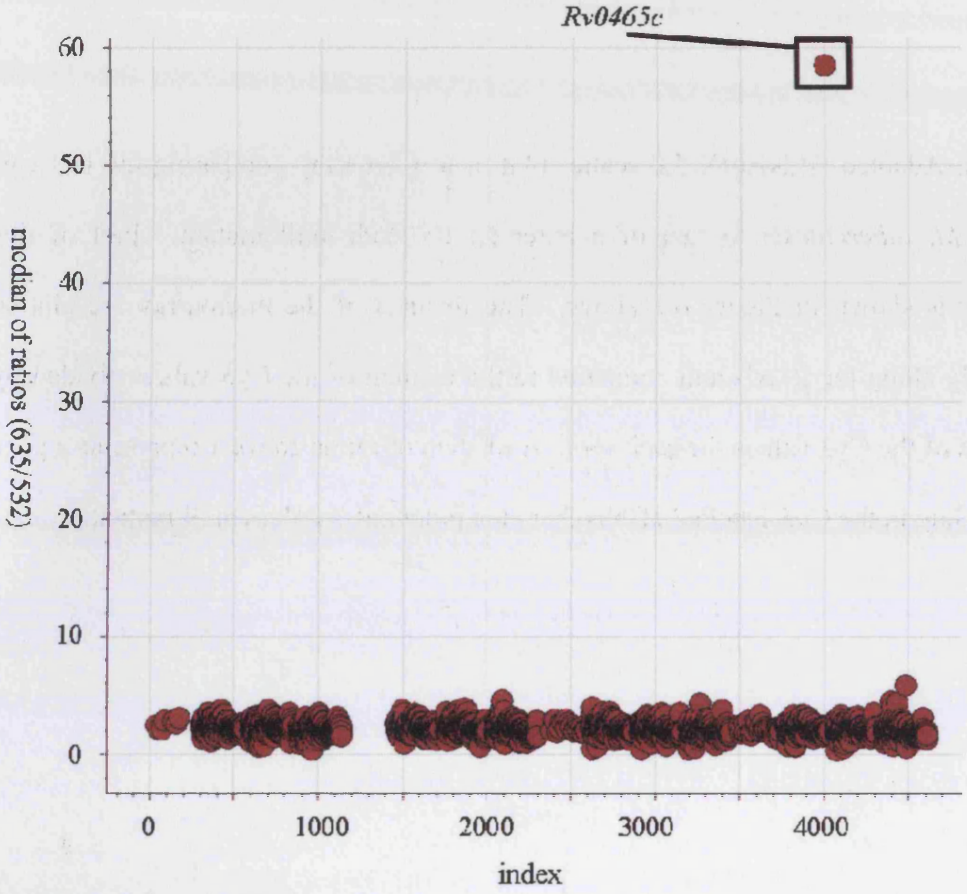
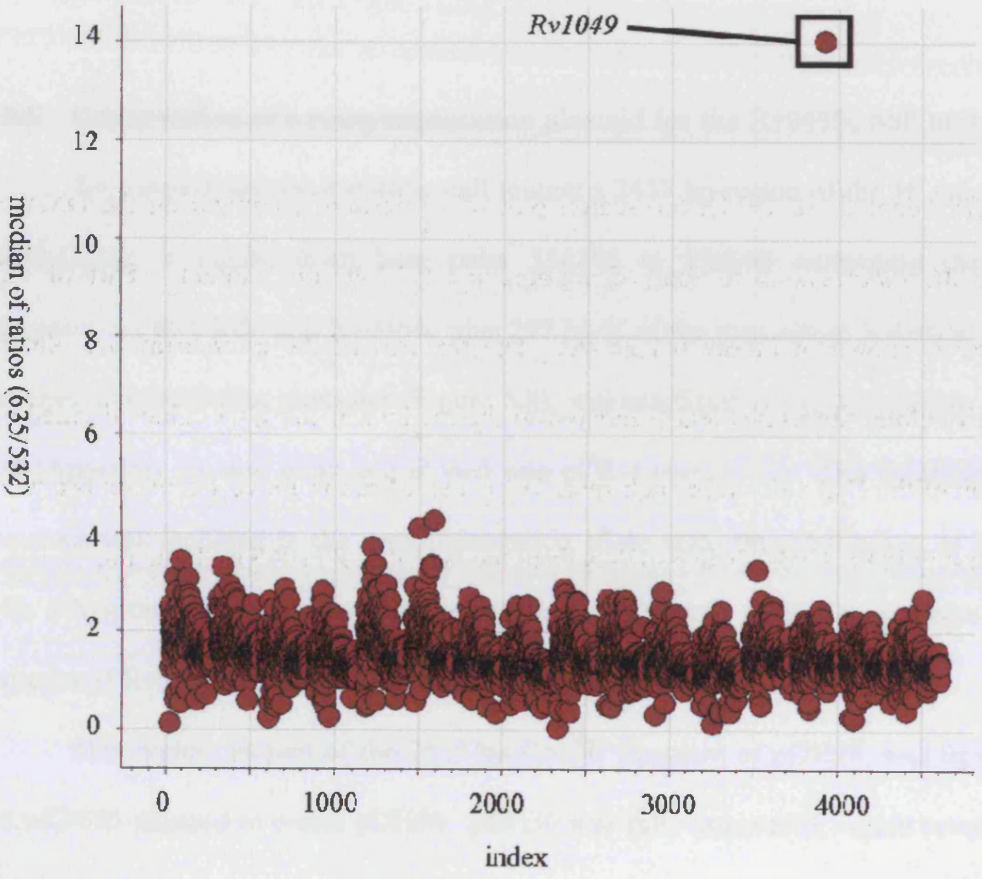

Rv0465c
60
50
40
30
20
10
0
median of ratios (635/532)
0
1000
2000
3000
4000
index

Figure 6.7: Adapted GENEPIX 5.0 scatter plot of a microarray comparing the DNA of wild-type *M. tuberculosis* to that of a potential Rv1049 null mutant. Part of this microarray is shown in Figure 6.4 above. The features of the microarray are plotted sequentially along the x-axis and compared to the median of the Cy5 values divided by the median of the Cy3 values for each spot. A median of ratios above 1 represents a gene that is missing in the Cy3-labelled cDNA from the potential Rv1049 null mutant.


Rv1049
14
12
10
8
6
4
2
0
median of ratios (635/532)
0
1000
2000
3000
4000
index

6.2.4 Construction of Rv0465c and Rv1049 null mutant strains carrying complementing alleles.

It was decided to complement the Rv0465c and Rv1049 null mutant strains with intact alleles on the integrating vector pKP186 before the mutants were assessed for a phenotype *in vivo*.

6.2.5 Construction of a complementation plasmid for the Rv0465c null mutant

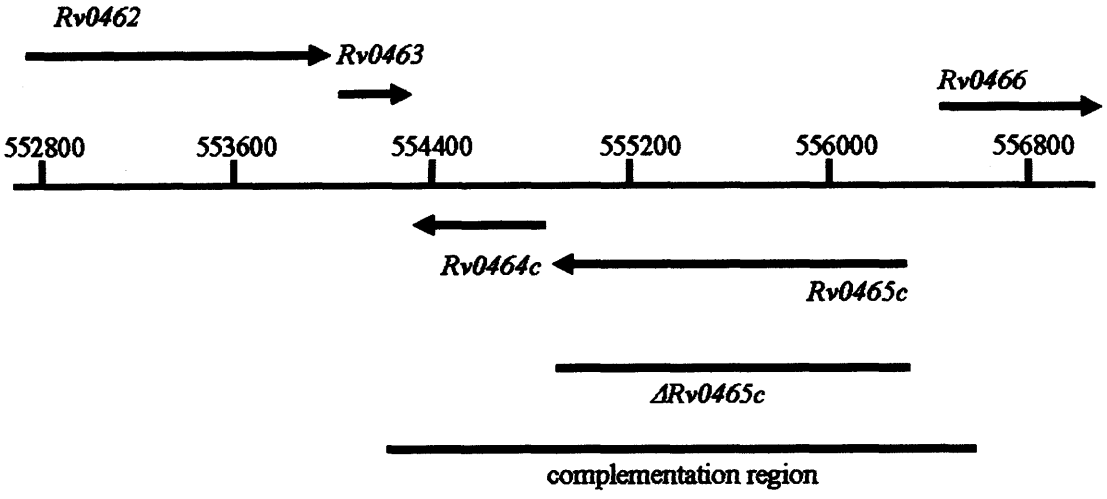
To complement the Rv0465c null mutant a 2457 bp region of the *M. tuberculosis* chromosome, a region from base pairs 554146 to 556603 containing the coding sequences for Rv0465c and Rv0464c plus 297 bp 5' of the start site of Rv0465c that was assumed to contain the promoter (Figure 6.8), was amplified using the primers 189 and 190 (Appendix 3), and blunt-end cloned into pCR-4 (see 2.7.1). The Rv0464c coding sequence was included in the complementation allele since the substitution of Rv0465c with a hygromycin resistance cassette in the Rv0465c null mutant may create a polar mutation if Rv0465c and Rv0464c are co-transcribed.

This region, as part of the 2463 bp *Bam*HI fragment of pDH28, was ligated into the pKP186 plasmid to create pDH30. pDH30 was fully sequenced, which revealed that no mutations were present.

6.2.6 Construction of a complementation plasmid for the Rv1049 null mutant

To complement the Rv1049 null mutant, a 2075 bp region of the *M. tuberculosis* chromosome from 1172019 to 1174094 that contained the coding sequences for Rv1049

Figure 6.8: Complementation of the Rv0465c null mutant of *M. tuberculosis*. The complementation region covers both Rv0465c and Rv0464c and the region between Rv0465c and Rv0466 that is likely to contain the promoter for Rv0465c.


Rv0462
Rv0463
Rv0466
552800
553600
554400
555200
556000
556800
Rv0464c
Rv0465c
ΔRv0465c
complementation region

and Rv1050 plus a 367 bp region 5' of the start site of Rv1049 that was assumed to contain the promoter (Figure 6.9), was amplified using the primers 191 and 192 (Appendix 3), and blunt-end cloned into pCR-4 TOPO (see 2.7.2) using the Zero Blunt kit (Invitrogen). Rv1050 was included in the complementation construct since the substitution of Rv1049 for a hygromycin resistance cassette in the Rv1049 null mutant may create a polar mutation if Rv1049 and Rv1050 are co-transcribed.

This region, as part of the 2093 bp *Eco*RI fragment of pDH29, was ligated into pKP186 to create pDH31. *Eco*RI was used to digest the restriction sites of the polylinker of pDH29 since the insert contained a *Bam*HI site, digestion by which would have cleaved the complementation region. Sequencing revealed that no mutations were present.

Both pDH30 and pDH31 were electroporated into their respective mutant strains as described in 2.1.1.4.8.

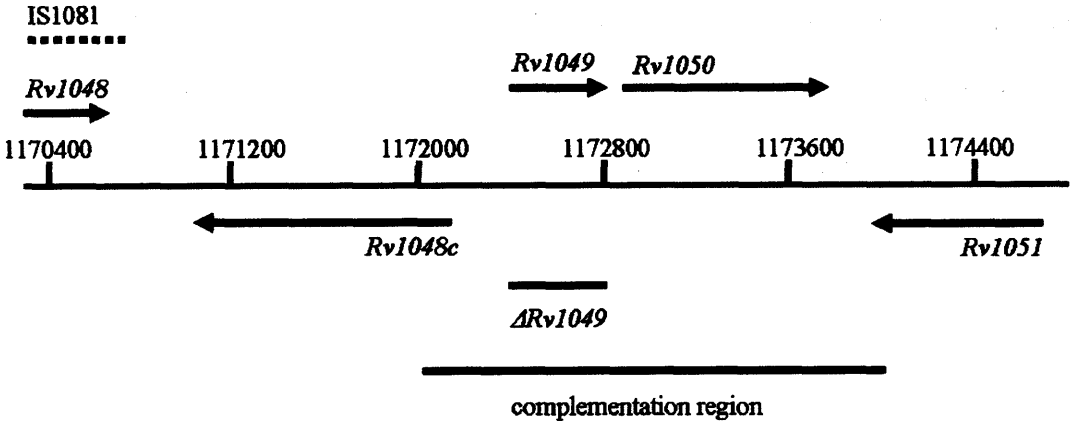
6.2.7 Phenotypic assessment of Rv0465c and Rv1049 null mutants

It was necessary to establish the normal growth phenotype of the Rv0465c and Rv1049 null mutants *in vitro* before examining the response of these strains to *tert*-butyl hydroperoxide and their growth of these strains in macrophage and murine models of infection.

6.2.8 *In vitro* growth

To assess the growth rate of *M. tuberculosis* strains lacking the genes Rv0465c and Rv1049 *in vitro*, wild-type *M. tuberculosis*, the Rv0465c null mutant and the Rv1049

Figure 6.9: Complementation of the Rv1049 null mutant of *M. tuberculosis*. The complementation region covers both Rv1049 and Rv1050 and the region between Rv1049 and Rv1048c that is likely to contain the promoter for Rv1049


IS1081
Rv1048
Rv1049
Rv1050
1170400
1171200
1172000
1172800
1173600
1174400
Rv1048c
Rv1051
ΔRv1049
complementation region

null mutant were grown to early exponential phase, diluted 1/500 into further roller bottles, three per strain, and the growth followed daily (Figure 6.10)(see 2.1.1.4.3). During exponential phase (93-164 hours) the wild-type, Rv0465c null mutant and the Rv1049 null mutant showed mean generation times of 16.66, 16.41 and 16.36 hours respectively. All three strains entered stationary phase around 237 hours with respective mean OD_{600}s of 4.64, 4.60 and 4.72 for the wild-type and Rv0465c and Rv1049 null mutants.

To assess the ability of wild-type *M. tuberculosis* to resist organic hydroperoxide stress, a titration of *tert*-butyl hydroperoxide was made against the wild-type *M. tuberculosis* using OD_{600} and viable counts measured (see 2.2.1.1 and 3.2.2.3). This data was used to determine bactericidal concentrations of *tert*-butyl hydroperoxide; 250μM was found to be a suitable concentration (3.2.2.3).

To examine the ability of Rv0465c and Rv1049 mutants to survive exposure to organic hydroperoxides, they were grown alongside the wild-type to early exponential phase and then each strain split into six aliquots (see 2.2.1.1). Three aliquots per strain were exposed to 250μM *tert*-butyl hydroperoxide, and the control strains exposed to volumes of water identical to that of the *tert*-butyl hydroperoxide. After 24 hours viable counts (Figure 6.11) and OD_{600}s (Figure 6.12) were measured. The wild-type exhibited a 0.85 mean log drop in viability. The Rv0465c and Rv1049 null mutants showed log losses in viability of 1.00 and 1.31 respectively. The loss in viability of the Rv1049 null mutant was significantly greater than that of the Rv0465c null mutant and the wild-type at a level of $p<0.005$. The OD_{600} data (Figure 6.12) shows a similar trend.

Figure 6.10: Aerobic growth of Rv0465c and Rv1049 null mutants of *M. tuberculosis*. The wild-type and the Rv0465c and Rv1049 null mutants were grown aerobically in Dubos media with albumin and glycerol at 37°C (2.1.1.4.1) to early exponential phase, before being diluted 1/500 in triplicate into fresh Dubos media with albumin and glycerol (2.1.1.4.1) and cultured aerobically at 37°C. Growth of the cultures was followed daily by measuring the OD_{600}. Each data point represents the mean of three cultures with standard error bars plotted. Numbers above the data points refer to the calculated generation time of the wild-type; numbers below the data refer to the mean OD_{600}s at each data point for the wild-type.

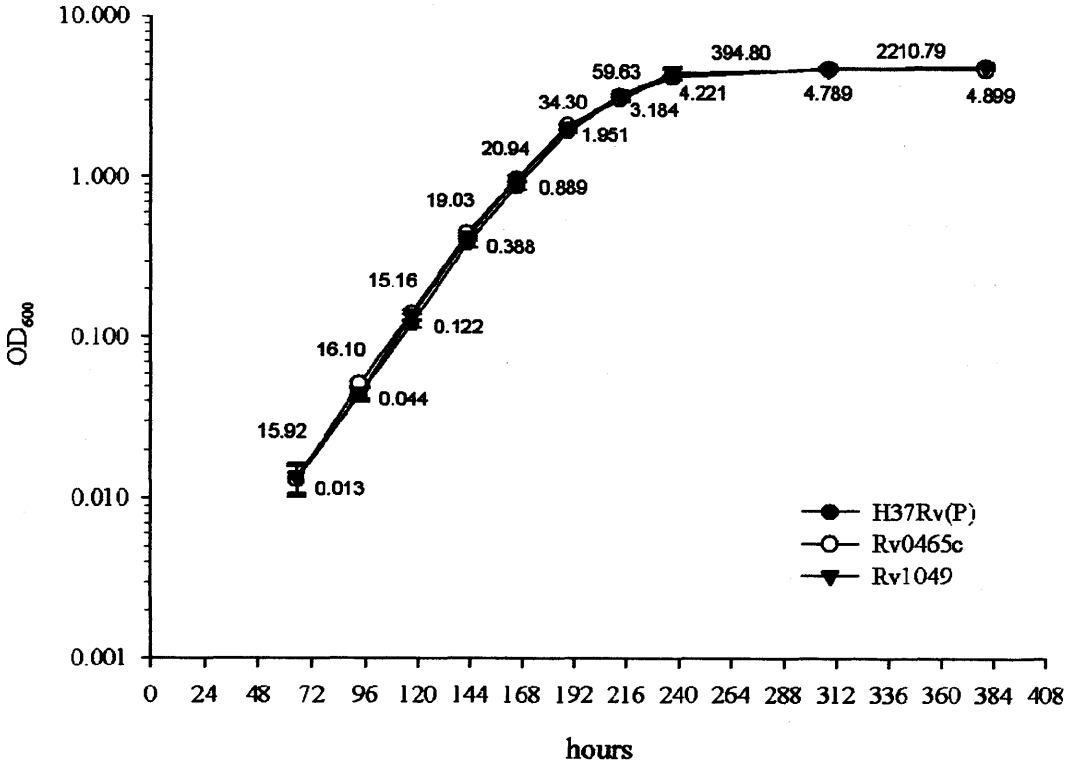
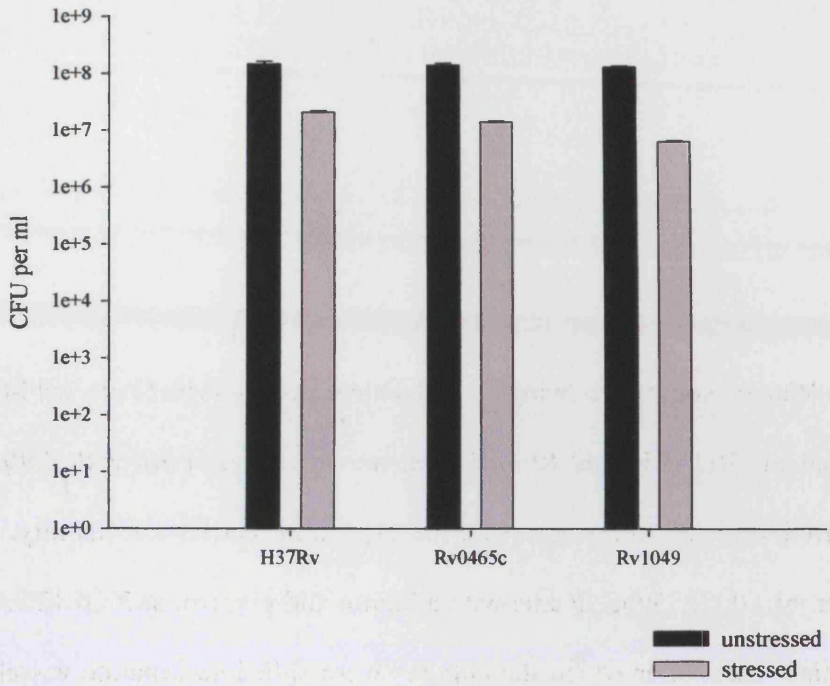

10.000
1.000
0.100
0.010
0.001
OD600
0
24
48
72
96
120
144
168
192
216
240
264
288
312
336
360
384
408
hours
15.92
0.013
16.10
0.044
15.16
0.122
19.03
0.388
20.94
0.889
34.30
1.951
59.63
3.184
394.80
4.221
2210.79
4.789
4.899
H37Rv(P)
Rv0465c
Rv1049

Figure 6.11: Viable counts showing the comparative survival of wild-type *M. tuberculosis* and the Rv0465c and Rv1049 null mutants after stress with 250μM *tert*-butyl hydroperoxide for 24 hours, expressed as A) cfu/ml and B) % viability. Strains were grown aerobically in Dubos media with albumin and glycerol at 37°C (2.1.1.4.1) to early exponential phase, after which the cultures were split into separate vessels. *tert*-butyl hydroperoxide was added to a final concentration of 250μM to half of the cultures, and an equal volume of water was added to the control cultures. Viable counts of the cultures were taken after 24 hours further incubation at 37°C. The experiment was performed in triplicate as described in 2.2.1.1, and standard errors are shown. The log loss of viability of the Rv0465c null mutant was significant at a level of $p<0.005$.

A


1e+9
1e+8
1e+7
1e+6
1e+5
1e+4
1e+3
1e+2
1e+1
1e+0
CFU per ml
H37Rv
Rv0465c
Rv1049
unstressed
stressed

B

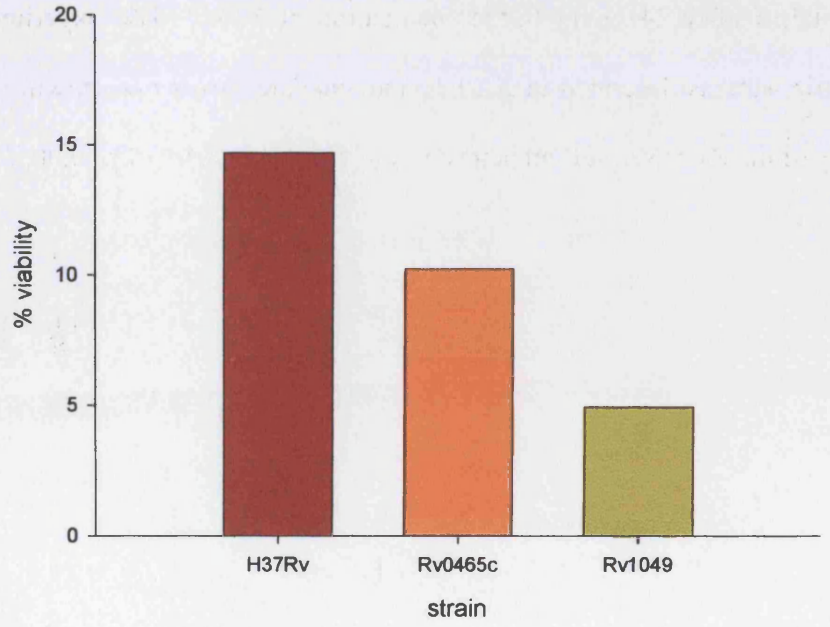
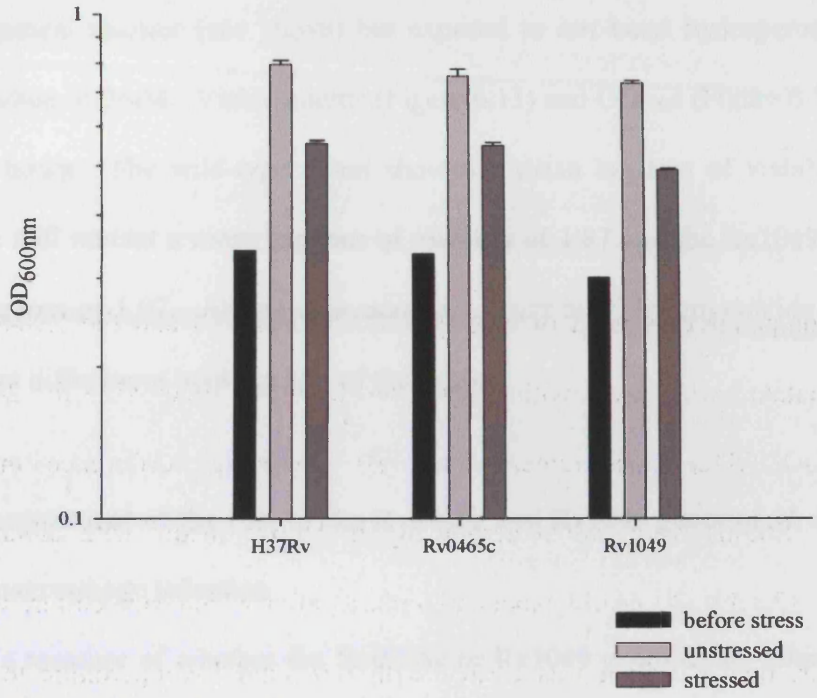
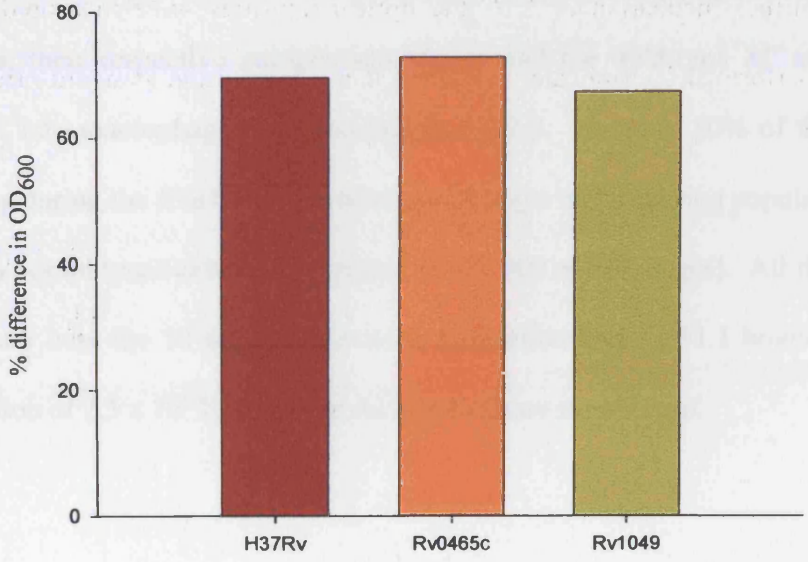

20
15
10
5
0
% viability
H37Rv
Rv0465c
Rv1049
strain

Figure 6.12: Comparative survival of wild-type *M. tuberculosis* and the Rv0465c and Rv1049 null mutants after stress with 250μM *tert*-butyl hydroperoxide for 24 hours, expressed as A) OD_{600} measurements, and B) percentage difference in OD_{600} measurements. Strains were grown aerobically in Dubos media with albumin and glycerol at 37°C (2.1.1.4.1) to early exponential phase, after which each culture was split into two vessels, one of which was treated with *tert*-butyl hydroperoxide to a final concentration of 250μM. The unstressed control cultures were treated with a water control of the same volume as the *tert*-butyl hydroperoxide used in the stressed cultures. After 24 hours further incubation at 37°C the optical densities were measured. The experiment was performed in triplicate as described in 2.2.1.1, and standard errors are shown.

A


1
0.1
OD600nm
H37Rv
Rv0465c
Rv1049
before stress
unstressed
stressed

B


80
60
40
20
0
% difference in OD600
H37Rv
Rv0465c
Rv1049

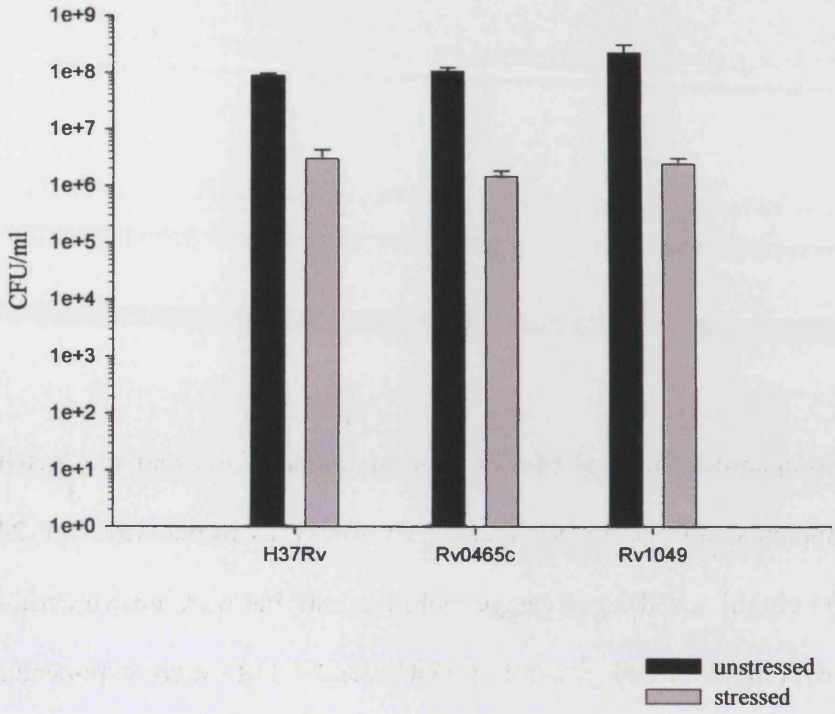
To examine the responses of the Rv0465c and Rv1049 null mutants to higher concentrations of organic hydroperoxides, the two mutants and the wild-type were treated in an identical manner (see above) but exposed to *tert*-butyl hydroperoxide at a final concentration of 2mM. Viable counts (Figure 6.13) and OD_{600}s (Figure 6.14) were taken after 24 hours. The wild-type strain showed a mean log loss of viability of 1.5, the Rv0465c null mutant a mean log loss of viability of 1.87 and the Rv1049 null mutant a mean log loss of 1.97. At this concentration of *tert*-butyl hydroperoxide there were no significant differences between any of the strains.

6.2.9 Assessment of the role of the Rv0465c and Rv1049 genes of *M. tuberculosis* in macrophage infection

To gain a measure of whether the Rv0465c or Rv1049 genes of *M. tuberculosis* play a role in the growth of this organism *in vivo*, the Rv0465c and Rv0149 null mutants, and the Rv0465c and Rv1049 null mutants carrying the integrated complementation plasmids, were assayed for growth in a macrophage model of infection (Figure 6.15). The two null mutants, their respective complement strains and the wild-type *M. tuberculosis* were infected into macrophages as described in 2.2.2.1. Roughly 10% of the bacteria were taken up during the five hour infection, presenting a mean starting population of 3.2×10^2 bacteria per ml lysed culture supernatant (=~80 000 macrophages). All three strains grew constantly over the 10 days with a mean generation time of 31.1 hours to a final mean population of 2.3×10^6 bacteria per ml lysed culture supernatant.

Figure 6.13: Comparative survival of wild-type *M. tuberculosis* and the Rv0465c and Rv1049 null mutants after stress with 2mM *tert*-butyl hydroperoxide for 24 hours, expressed as A) cfu/ml, and B) percentage viability. Strains were grown aerobically in Dubos media with albumin and glycerol at 37°C (2.1.1.4.1) to early exponential phase, after which each culture was split into two further vessels, one of which was treated with *tert*-butyl hydroperoxide to a final concentration of 2mM. The unstressed control cultures were treated with a water control of the same volume as the *tert*-butyl hydroperoxide used in the stressed cultures. After 24 hours further incubation at 37°C the viable counts of each strain were measured. The experiment was performed in triplicate as described in 2.2.1.1, and standard errors are shown.

A


1e+9
1e+8
1e+7
1e+6
1e+5
1e+4
1e+3
1e+2
1e+1
1e+0
CFU/ml
H37Rv
Rv0465c
Rv1049
unstressed
stressed

B

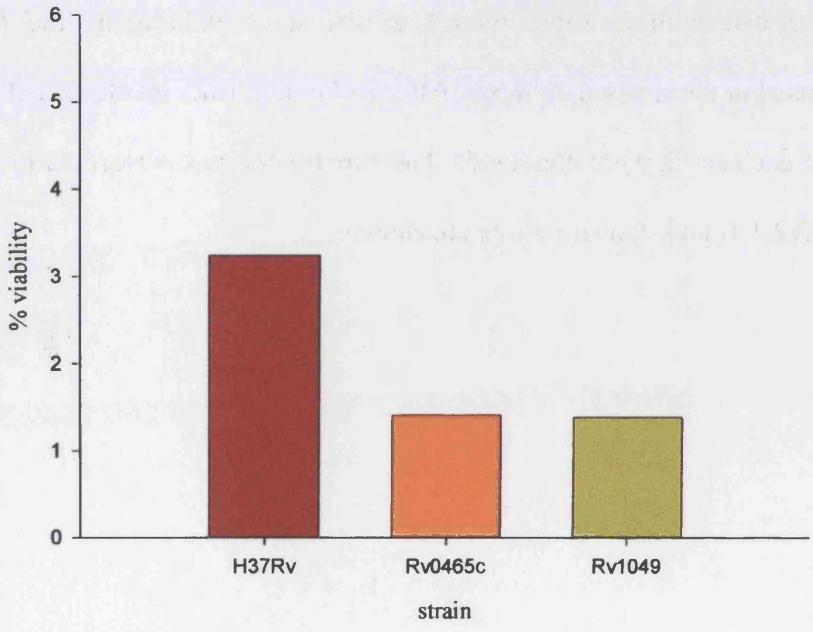
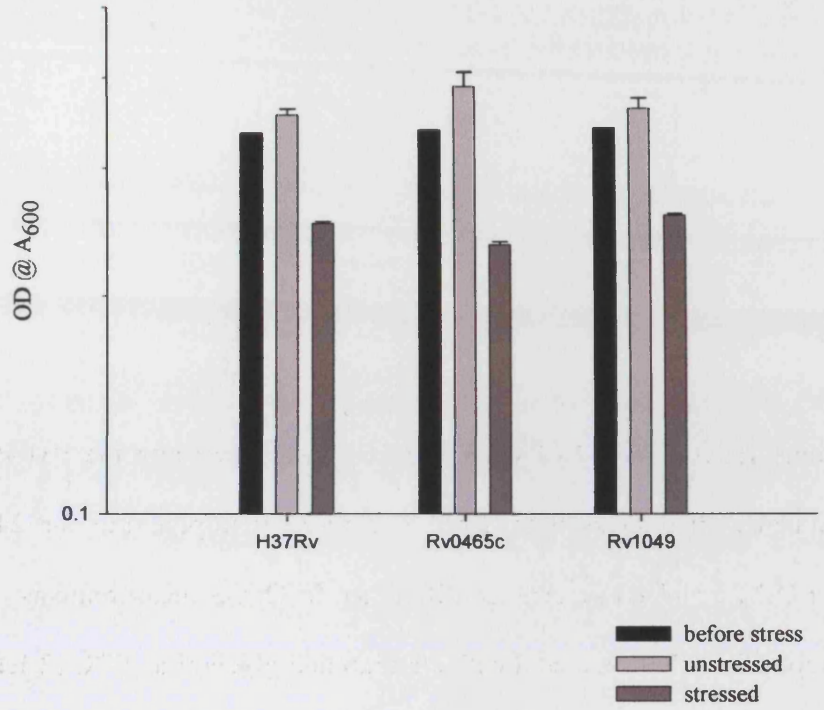

6
5
4
3
2
1
0
% viability
H37Rv
Rv0465c
Rv1049
strain

Figure 6.14: Comparative survival of wild-type *M. tuberculosis* and the Rv0465c and Rv1049 null mutants after stress with 2mM *tert*-butyl hydroperoxide for 24 hours, expressed as A) OD_{600}, and B) percentage difference in OD_{600} measurements. Strains were grown aerobically in Dubos media with albumin and glycerol at 37°C (2.1.1.4.1) to early exponential phase, after which each culture was split into two further vessels, one of which was treated with *tert*-butyl hydroperoxide to a final concentration of 2mM. The unstressed control cultures were treated with a water control of the same volume as the *tert*-butyl hydroperoxide used in the stressed cultures. After 24 hours further incubation at 37°C the viable counts of each strain were measured. The experiment was performed in triplicate as described in 2.2.1.1, and standard errors are shown.

A


OD @ A_{600}
0.1
H37Rv
Rv0465c
Rv1049
before stress
unstressed
stressed

B

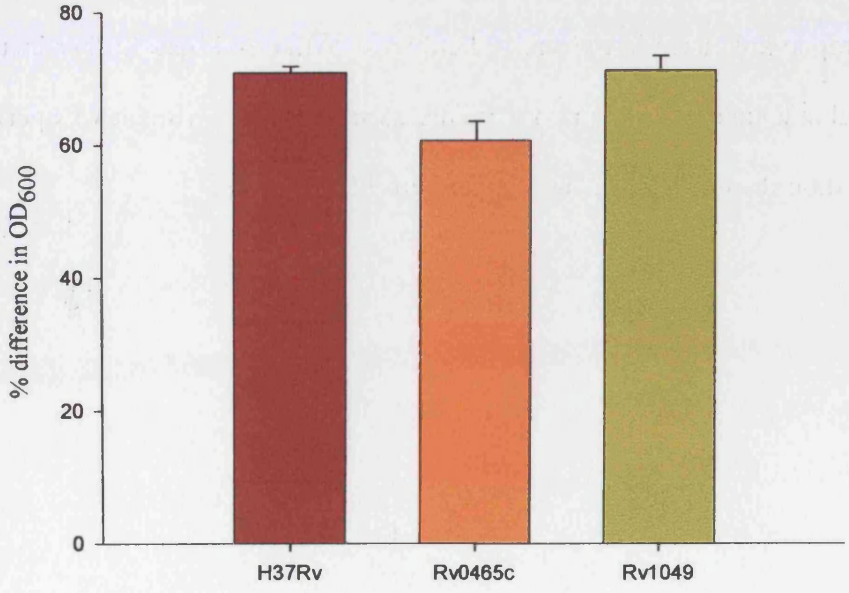
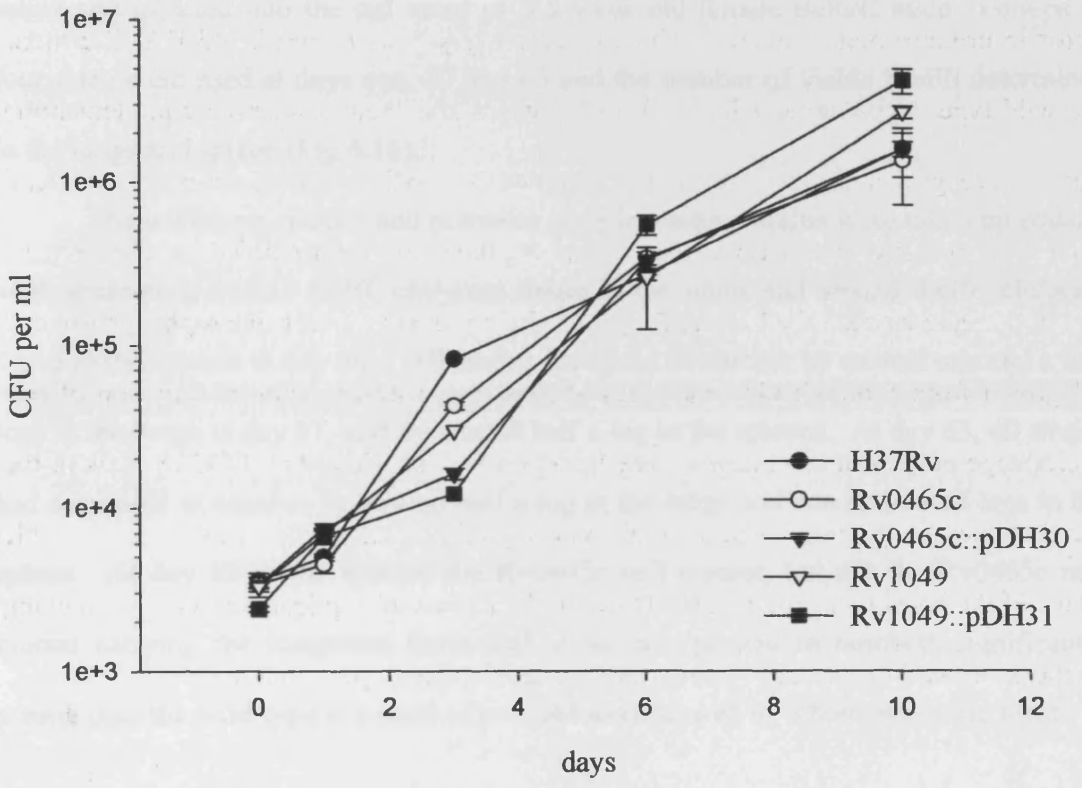

% difference in OD_{600}
80
60
40
20
0
H37Rv
Rv0465c
Rv1049

Figure 6.15: Growth of the wild-type *M. tuberculosis*, the Rv0465c and Rv1049 null mutants and their respective strains carrying functional alleles on the integrating pKP186 vector in primary bone marrow derived macrophages from 8 month old BALB/c mice. The wild-type, Rv0465c and Rv0149 null mutants and their respective complementing strains were grown aerobically at 37°C in Dubos media with albumin and glycerol (2.1.1.4.1) to early exponential phase, then washed and resuspended in macrophage infection media as described in 2.2.2.1 before being used to infect macrophages isolated and grown from 8 month old female Balb/C mice at a multiplicity of infection of one macrophage per two macrophages. The infection was monitored at 37°C by lysis of the infected macrophage cultures at intervals with saponin (2.2.2.1) and measuring the viable counts of the lysed cell supernatant. Data are the results of duplicate infections, showing the mean of two infected cell cultures and the standard error per time point.


1e+7
1e+6
1e+5
1e+4
1e+3
CFU per ml
0
2
4
6
8
10
12
days
H37Rv
Rv0465c
Rv0465c::pDH30
Rv1049
Rv1049::pDH31

6.2.10 Assessment of the role of the Rv0465c and Rv1049 genes of *M. tuberculosis* in a murine model of infection

To determine whether the Rv0465c or Rv1049 genes play a role in the growth of *M. tuberculosis in vivo*, the Rv0465c and Rv1049 null mutants, and the Rv0465c and Rv1049 null mutants carrying the integrated complementation plasmids, were grown in rolling bottles as in other experiments (see 2.1.1.4.3) until they reached exponential phase. The bacteria were then diluted to an OD_{600} of 0.020 in sterile phosphate-buffered saline and injected into the tail veins of 6-8 week old female Balb/C mice. Groups of four mice were used at days one, 37 and 63 and the number of viable bacilli determined in the lungs and spleen (Fig 6.16).

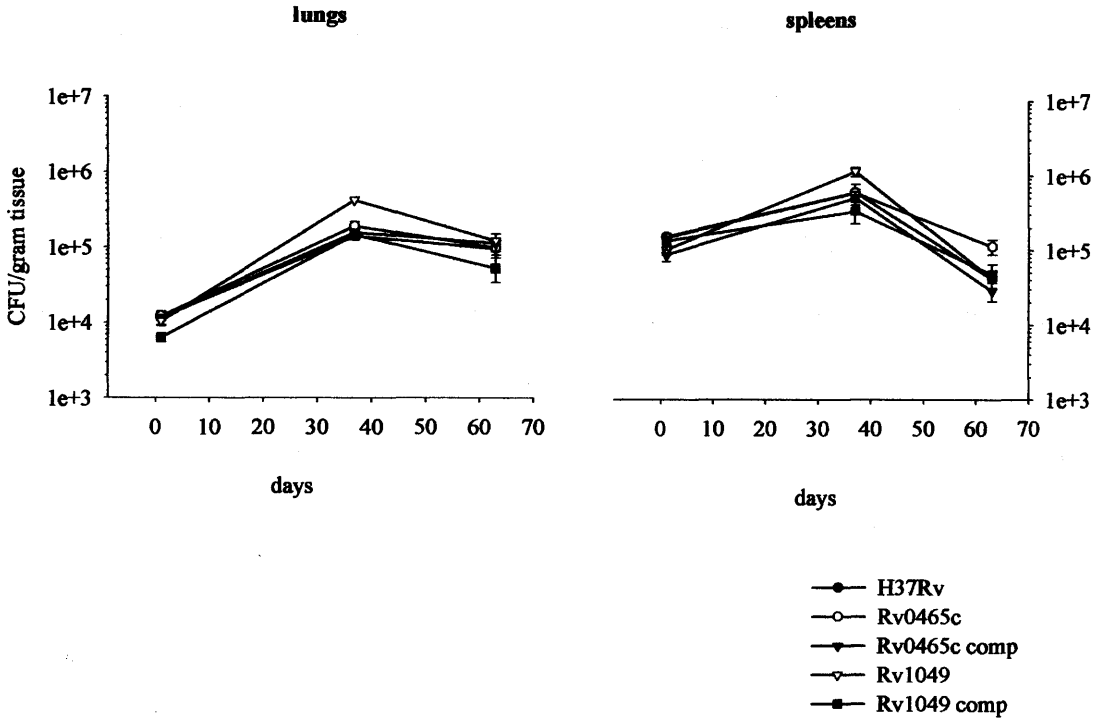
The wild-type, mutant and potential complementing strains were taken up equally well, presenting around $1x10^4$ cfu/gram tissue in the lungs and around $1x10^5$ cfu/gram tissue in the spleens at day one. All strains increased in number by around one and a half logs in the lungs at day 37, and by around half a log in the spleens. At day 63, all strains had decreased in number by around half a log in the lungs and one and a half logs in the spleen. At day 63 in the spleens the Rv0465c null mutant, but not the Rv0465c null mutant carrying the integrated functional allele, was present in numbers significantly greater than the wild-type at a level of $p=0.044$ as compared by a homoscedastic t-test.

6.3 DISCUSSION

6.3.1 Construction of Rv0465c and Rv1049 null mutants of *M. tuberculosis*.

The data presented here shows conclusive proof of the construction of two strains of *M. tuberculosis* lacking the genes *Rv0465c* and *Rv1049*. Phenotypic testing,

Figure 6.16: Growth of wild-type *M. tuberculosis*, Rv0465c and Rv1049 null mutants of *M. tuberculosis* and the Rv0465c and Rv1049 null mutants carrying integrated copies of their respective functional alleles in a murine intravenous injection model of infection. The wild-type, Rv0465c and Rv1049 null mutants were grown aerobically to early exponential phase in Dubos media with albumin and glycerol (see 2.1.1.4.1) at 37°C before washing in sterile PBS. Approx. 2 x 10^6 bacteria in a volume of 0.2ml were then injected intravenously into the tail vein of 6-8 week old Balb/C mice and the infection followed by removing the lungs and spleens of four mice per strain at intervals, and measuring the viable counts of tissue homogenate after serial dilution in saline. See 2.2.2.2 for further details on the method.


lungs
spleens
CFU/gram tissue
1e+7
1e+6
1e+5
1e+4
1e+3
0
10
20
30
40
50
60
70
days
H37Rv
Rv0465c
Rv0465c comp
Rv1049
Rv1049 comp

complementation and further analysis will allow the role that these two genes play in infection or in resistance to oxidative stress to be assessed.

Both *Rv0465c* and *Rv1049* appear to be the upstream members of co-transcribed two-gene operons (Figure 6.17 and 6.18). It is reasonable to assume that the substitution of *Rv0465c* and *Rv1049* for hygromycin resistance markers in the two strains are creating polar mutations of the downstream genes *Rv0464c* and *Rv1050*, respectively. *Rv0464c* is annotated as a conserved hypothetical protein, and *Rv1050* is annotated as a probable oxidoreductase on the Tuberculist web server (2005).

6.3.2 *in vitro* growth of Rv0465c and Rv1049 null mutants of *M. tuberculosis*

The growth of the Rv0465c and Rv1049 null mutants of *M. tuberculosis* was indistinguishable from that of the wild-type, consistent with previous data suggesting that Rv0465c and Rv1049 are non-essential genes (Sassetti *et al.* 2003). The growth phenotype here suggests that Rv0465c and Rv1049 are not involved in normal cellular growth and division of *M. tuberculosis* in aerobic rolling cultures.

6.3.3 Resistance of Rv0465c and Rv1049 null mutants to *tert*-butyl hydroperoxide stress

Following the titration of *tert*-butyl hydroperoxide against the wild-type *M. tuberculosis* (3.2.2.3) , it was decided to stress the Rv0465c and Rv1049 null mutant strains with *tert*-butyl hydroperoxide at a 250μM final concentration to cause a significant, measurable drop in viability of the wild-type to which the viabilities of the null mutants could be compared. A small but significant drop in the viability of the

Figure 6.17: Tuberculist (2005) representation of the genome around Rv0465c.

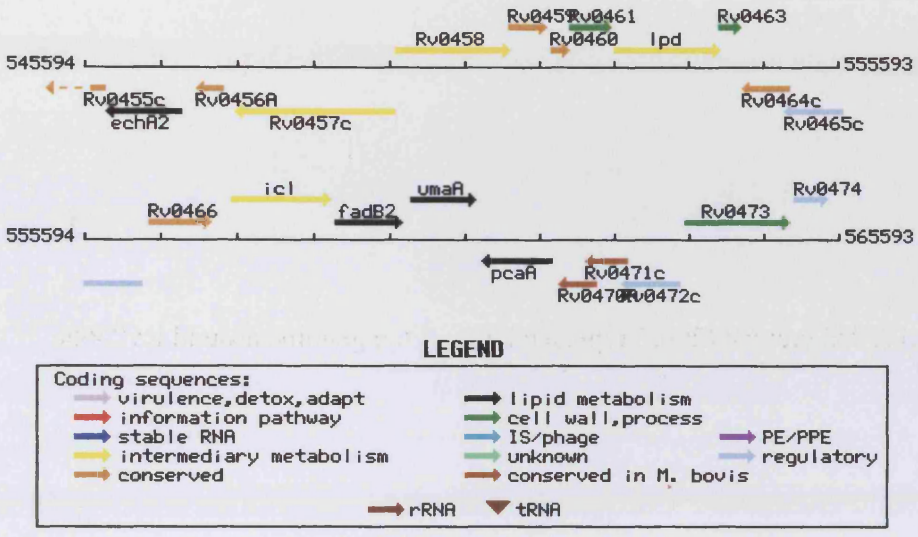
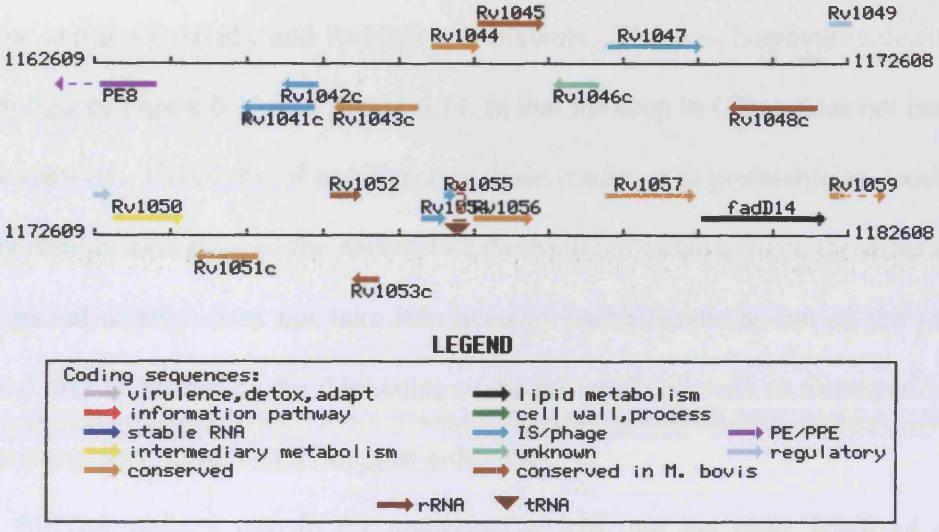

Rv0459
Rv0461
Rv0463
Rv0458
Rv0460
lpd
545594
555593
Rv0455c
Rv0456A
Rv0464c
echA2
Rv0457c
Rv0465c
icl
umaA
Rv0474
Rv0466
fadB2
Rv0473
555594
565593
pcaA
Rv0471c
Rv0470
Rv0472c
LEGEND
Coding sequences:
virulence,detox,adapt
information pathway
stable RNA
intermediary metabolism
conserved
lipid metabolism
cell wall,process
IS/phage
unknown
conserved in M. bovis
PE/PPE
regulatory
rRNA
tRNA

Figure 6.18: Tuberculist (2005) representation of the genome around Rv1049.


Rv1045
Rv1049
Rv1044
Rv1047
1162609
1172608
PE8
Rv1042c
Rv1046c
Rv1041c Rv1043c
Rv1048c
Rv1052
Rv1055
Rv1057
Rv1059
Rv1050
Rv1056
fadD14
1172609
1182608
Rv1051c
Rv1053c
LEGEND
Coding sequences:
virulence,detox,adapt
information pathway
stable RNA
intermediary metabolism
conserved
lipid metabolism
cell wall,process
IS/phage
unknown
conserved in M. bovis
PE/PPE
regulatory
rRNA
tRNA

Rv1049 null mutant when compared to the wild-type suggested a role for this gene in the resistance of *M. tuberculosis* to *tert*-butyl hydroperoxide.

To extend upon this finding, it was decided to stress the cultures with a much higher concentration of *tert*-butyl hydroperoxide in order to maximise any differences between the Rv1049 null mutant and the wild-type. A 2mM final concentration of *tert*-butyl hydroperoxide was used, representing the highest concentration tested against the wild-type (3.2.2.3). At this concentration there was no statistical difference between the wild-type and the Rv0465c and Rv1049 null mutants. There is, however, a discrepancy between figures Figure 6.13 and Figure 6.14, in that the drop in OD_{600} does not match the drop in viability. Given the of conflict with these results, it is preferable to consider the viability data as indicative of the survival of the bacterial culture, since the measurement of the optical density does not take into account bacteriostatic action of the stressing agent and may be misled by the thickening of the bacterial cell wall, as discussed in 1.4.4.

Taken together, these results suggest either that:

a) Rv1049 plays a role in the resistance of low, but not high, levels of organic hydroperoxide stress in *M. tuberculosis* and that some other mechanism is responsible for high-level resistance of *M. tuberculosis* to *tert*-butyl hydroperoxide, or

b) Rv1049 plays a central role in the response of *M. tuberculosis* to *tert*-butyl hydroperoxide, but a 2mM concentration of *tert*-butyl hydroperoxide is too high a concentration for the wild-type to resist killing, or

c) The behaviour of the Rv1049 null mutant after exposure to 250μM *tert*-butyl hydroperoxide was an experimental anomaly, and the gene plays no role in the resistance of *M. tuberculosis* to *tert*-butyl hydroperoxide.

Since the response of the Rv1049 null mutant to 250μM *tert*-butyl hydroperoxide was performed in triplicate with consistent results, it may be possible to discount scenario (c). Scenario (b) is likely to be untrue, since the drop in viability of the Rv1049 null mutant was only 0.31 when compared to the wild-type and, although statistically significant, is not of the scale one would expect to see if the sole regulator of *tert*-butyl hydroperoxide responses in *M. tuberculosis* was rendered non-functional. Scenario (a) is a reasonable hypothesis, since MarR-type regulators tend to be part of a general response in other bacteria and one would expect a more specialised system to be acting in conjunction with this in an organism in which resistance to organic hydroperoxides was important for survival.

6.3.4 Construction of potential complementation strains of Rv0465c and Rv1049 null mutants of *M. tuberculosis*.

The alleles that were cloned to complement the Rv0465c and Rv1049 null mutant strains included what appeared to be the full operon (of two genes) for each of the mutants, plus regions that are assumed to contain the native promoter of the operons.

In the event of either the Rv0465c or Rv1049 null mutant displaying a phenotype, the ability to complement the mutant and restore the wild-type virulence would allow confirmation of the role of that gene in the physiology of *M. tuberculosis*, which may be particularly important if the allelic exchange of Rv0465c and Rv1049 has lead to polar

mutations in downstream, cotranscribed genes. Complementation of a mutant phenotype would also allow structure-function analysis of the gene product. However, under the conditions tested the mutants did not behave differently from the wild-type, although this does not preclude the use of the complemented strains in the future should a mutant phenotype present itself.

Due to the reasons discussed in 6.1, it is possible to hypothesise that either or both of Rv0465c or Rv1049 may play a role in the virulence of *M. tuberculosis*. This hypothesis is supported by the response of the Rv0465c and Rv1049 null mutants to *tert*-butyl hydroperoxide stress, however it is possible that the two genes are responding to more diverse stimuli than were replicated *in vitro* and that *in vivo* analysis of the two mutants would show stronger phenotypes.

6.3.5 Assessment of the role of the Rv0465c and Rv1049 genes of *M. tuberculosis* in a macrophage infection model.

There appeared to be no difference between the growth of the Rv0465c and Rv1049 null mutants and the wild-type in macrophages. Correspondingly, there was also no difference seen between the potential complementing strains, and the Rv0465c and Rv1049 null mutants. Since no morphology changes were seen in cells infected with any of the strains under test, it can only be concluded that the Rv0465c and Rv1049 genes of *M. tuberculosis* do not play a role in the infection of murine macrophages in this model. It is, however, not possible at this stage to conclude that either of these genes do not play a role in infection.

6.3.6 Assessment of the phenotypes of Rv0465c and Rv1049 null mutants of *M. tuberculosis* in a murine model of infection.

To examine the *in vivo* responses of Rv0465c and Rv1049 null mutants in a more complex model of infection it was decided to assay them in the same mouse model of infection as used in 3.2.3.3. Neither of the mutants grew in a manner particularly different to that of the wild-type, although the number of viable bacteria in the spleens of the mice at day 63 was significantly higher than the wild type for the Rv0465c null mutant. This may represent a subtle *in vivo* phenotype since the mutant strain carrying the integrated functional copy of the allele did not exhibit this behaviour, however there were no differences in the growth of the Rv0465c null mutant to the wild-type in the lungs.

7 Discussion

7.1 Phenotypic examination of *senX3* and *regX3* null mutants of *M. tuberculosis*

By similarity of *senX3* and *regX3* to orthologous genes in other organisms, i.e. *S. pombe* and *E. coli*, it was possible to hypothesise a role for the SenX3-RegX3 two component signal transduction system in the response of *M. tuberculosis* to oxidative stress. Furthermore, structural modelling of SenX3 by J. Saldanha (Mathematical Biology, NIMR) showed a PAS-like domain to be the likely signal input domain of SenX3. PAS domains are often sensors of redox levels, and the structural similarity of the SenX3 PAS-like domain to PAS domains in other organisms suggests that it may bind a haem prosthetic group, further supporting a role for the SenX3-RegX3 two component signal transduction system in oxygen sensing.

In contrast with another report (Parish *et al.* 2003) it has been shown here that null mutants in the SenX3-RegX3 two component signal transduction system display normal growth in aerobic culture. This is consistent with data arising from mutagenesis work on *M. tuberculosis* that suggests neither *senX3* nor *regX3* are required for mycobacterial growth *in vitro* (Sassetti *et al.* 2003). Furthermore, two-component signal transduction systems are often only active under specific environmental conditions, and would not be active during growth *in vitro*. In this situation it is possible to imagine that their absence would not affect the growth rate in the absence of the environmental signals that require their activity.

The importance of the SenX3-RegX3 two component signal transduction system in the *M. tuberculosis* infection process was shown here by demonstrating significantly reduced growth of the *senX3* and *regX3* null mutants in a mouse model of infection. Complementation of the null mutants with functional copies of *senX3* and *regX3* proved

conclusively that the attenuated phenotype of the mutants was solely due to the absence of these genes.

To further study the role of SenX3 and RegX3 during infection, the growth of the *senX3* and *regX3* null mutants was followed in macrophages derived from murine bone marrow. The mutants were attenuated in this *in vitro* model of infection, showing that *M. tuberculosis* requires the SenX3-RegX3 two component signal transduction system to be active during the early stages of infection. This allows us to conclude that the SenX3-RegX3 two component signal transduction system is responding to at least some of the environmental changes that occur during macrophage infection, such as the oxidative burst or nutrient starvation, although it is not possible to rule out involvement of this system to later stages during infection, such as those that occur in the granuloma.

In an effort to try and understand if a specific single agent was responsible for the attenuation of the *senX3* and *regX3* null mutants, these strains were exposed to a number of different agents *in vitro*. The mutants appeared to be more resistant than the wild-type to killing by high concentrations of superoxide ions, and organic hydroperoxides that would normally form *in vivo* as a breakdown product of the superoxide ions. No difference was seen in viability between the mutants and the wild-type after nitrosative stress, nutrient starvation or growth in acidic conditions. From these data it may be possible to conclude that the SenX3-RegX3 two component signal transduction system may play a role in repression of genes that are involved in the defence of the bacterium to superoxide stress and its breakdown products, or metabolic stress as a result of oxygen limitation. In the absence of the SenX3-RegX3 two component signal transduction

system these genes may be constitutively produced, which would provide a selective advantage to the mutant strains under superoxide stress conditions.

In conditions of *in vitro* oxygen or nutrient starvation *M. tuberculosis* would be likely to encounter internal oxidative stress as the free radicals produced by aerobic metabolism progressively damage the cellular machinery unless these processes were shut down. In *E. coli* this function is performed by ArcB, to which the SenX3-RegX3 two component signal transduction system has homology. Although no difference in viability between the mutant strains and the wild-type was seen here after growth in low-oxygen conditions or in the absence of a carbon source, it may be that the time-frame for this phenotype to become apparent in *M. tuberculosis* is greater than that assessed in those experiments.

7.2 Microarray examination of the responses of *senX3* and *regX3* null mutants to environmental stimuli

Previous work has shown no great differences in the transcriptome between the mutants and the wild-type after growth in rolling bottles (Rickman 2002). This is in agreement with the hypothesis that the SenX3-RegX3 two component signal transduction system is not active during normal growth *in vitro*. In the presence of a stimulus for the SenX3-RegX3 two component signal transduction system it would be possible to analyse differences in the ratios of gene transcripts between the mutants and the wild-type and thus deduce a role for the system based on the identity and function of the differentially regulated genes.

No differences in the transcriptome between the wild-type and the mutants were seen after superoxide stress or growth in sealed cultures containing very little air. This suggests that these conditions were not sufficient to cause activation of the SenX3-RegX3 two component signal transduction system. Initial experiments with a standing 'microaerobic' model of *M. tuberculosis* growth showed a number of operons to be differentially regulated between the wild-type and the *senX3* null mutant; however the differentially expressed gene list varied between experiments. It is likely that the bacteria are being subjected to reduced oxygen levels since cultures enter stationary phase before reaching a high enough OD_{600} to be indicative of nutrient limitation in aerobic cultures.

Refinement of this model allowed examination of the low-oxygen transcriptomes of the *senX3* and *regX3* null mutants and another *regX3* null mutant, Tame15, to be compared to the wild-type. A number of genes were common to the lists of differentially regulated genes of all three mutants, suggesting that there are subtle differences in the regulatory behaviour of the three strains. A ferridoxin gene, *fdxA,* is seen to be apparently regulated by the SenX3-RegX3 two component signal transduction system in the *senX3* null mutant and Tame15, as is *narX*, a nitroreductase, which is also associated with the DosR regulon. A number of mycobactin genes are upregulated in the *regX3* null mutant. These point to a possible role for the SenX3-RegX3 two component signal transduction system in the pathogenesis of *M. tuberculosis*, since genes of these types are often associated with virulence in other organisms.

With respect to the variation in gene lists between the three strains grown under microaerobic conditions, it is possible that the SenX3-RegX3 two component signal transduction system is very sensitive to slight variation in the culture environment, and

that slight variations in the handling of the strains accounts for the differences in the transcriptomes between the three strains.

Another method of analysing genes controlled by a regulatory system is to overexpress the regulator in a null mutant strain of *M. tuberculosis* and compare the transcriptional profile of this strain to the mutant carrying the overexpression vector without the coding sequence of the regulator. The high levels of expression of the regulator may be able to make the system constitutively active, even in the absence of the specific signal for the system.

Expression of *regX3* driven by a strong constitutive Hsp60 promoter showed high levels of *regX3* transcript being produced in the overexpressing strain. Comparison of microarray data for this strain against a *regX3* null mutant that did not appear to express any *regX3* transcript showed 199 genes that were differentially regulated between the two strains. A number of these genes also appear in the published data for the Tame15 mutant of *regX3*, as could be expected. Another group of genes also appear in the published data for the nutrient starvation stimulon of the wild-type *M. tuberculosis*, indicating that the SenX3-RegX3 two component signal transduction system is playing a role in the adaptation of nutrient starvation. Other genes in the list code for other regulatory genes, and genes known to be involved in the response of *M. tuberculosis* to components of the oxidative burst of the macrophage.

Given i) the environments in which it is possible to see transcriptional differences between mutants in the SenX3-RegX3 two component signal transduction system and the wild-type, ii) the genes that are differentially regulated between these strains and the wild-type, and iii) the similarity of the lists of differentially-regulated genes to published

data, it is possible to conclude that the SenX3-RegX3 two component signal transduction system may be acting as a global regulator, sensing redox levels or metabolic states and co-ordinating expression of pathogenicity factors that interact with the host and gene products that adapt the metabolism of the bacillus to an environment that it may encounter within a granuloma. Although the *senX3* and *regX3* null mutants were no more sensitive than the wild-type after incubation in nutrient starvation conditions (see section 3.2.1.3), it may be that the mutant bacteria were lacking critical components of the response to the host, but in the absence of any exogenous damaging compounds *in vitro* the mutants did not show any difference in viability. It is, however, possible to hypothesise that these mutants would be more attenuated in infection models after growth in nutrient starvation conditions.

7.3 Protein-DNA interactions of RegX3

A previous report (Himpens *et al.* 2000) studied the interaction between renatured, recombinant RegX3 and a 42 bp region upstream of *senX3*. Our study shows that recombinant polyhistidine-tagged RegX3 can be expressed and purified in soluble form that is more likely to be folded correctly and display the DNA-binding activity of the protein *in vivo*.

We have been able to confirm, by a more direct method, the interaction between RegX3 and the 42bp region upstream of *senX3* reported in Himpens *et al.* (2000). This interaction implicates RegX3 in regulating its own expression. Further to the confirming the results in Himpens *et al.* (2000), the region upstream of *senX3* has been mapped using overlapping oligonucleotides; contrary to expectation RegX3 appears to bind a number of

sites in this region, suggesting that this region may form secondary structure that interacts with RegX3 at a number of sites along its length. An alignment of the upstream region of *senX3* with that of two related mycobacterial species showed a large degree of homology, which does not preclude the promoter-binding region being scattered throughout this region.

One potential promoter region from a gene suggested to be regulated by RegX3 in Parish *et al.* (2003), and two from genes in this study that appeared to be differentially regulated in a *senX3* null mutant after growth in low oxygen conditions, were assayed for their ability to bind RegX3. RegX3 appeared to bind all three promoters, the electrophoretic mobility shift of which altered when RegX3 was phosphorylated. These results suggested that the phosphorylation state of RegX3 alters its ability to bind to promoters. The interactions specifically with the upstream regions of the gene encoding a probably transcriptional regulatory protein Rv1460 again suggests that RegX3 may be controlling other transcriptional regulators, and the interaction of the upstream region of Rv2780 (alanine dehydrogenase, encoded by *ald*) with RegX3 confirms that the SenX3-RegX3 two component signal transduction system has a role in regulating the metabolic state of *M. tuberculosis*.

7.4 Initial study of other genes that may be affecting the responses of *M. tuberculosis* to oxidative stress

Rv0465c and Rv1049 are potential regulatory genes that were implicated in the oxidative stress response by their upregulation in the wild-type following organic hydroperoxide stress. Rv1049, a MarR-type regulator, also bears homology to genes in

other organisms, namely *B. subtilis* and *X. campestris*, which are involved in the resistance of those organisms to oxidative stress.

Null mutants of these genes have been created in this study, and it has been possible to demonstrate that these genes are not essential for the growth of *M. tuberculosis* in vitro as shown by their growth rate in rolling culture being similar to that of the wild-type. The Rv1049 null mutant is significantly more susceptible to killing by intermediate, but not high, levels of *tert*-butyl hydroperoxide, an organic hydroperoxide. These results suggest that Rv1049 is necessary for the proper response of *M. tuberculosis* to components of the oxidative burst produced during infection. The lack of a reduced viability of the Rv1049 null mutant compared to the wild-type after exposure to high levels of *tert*-butyl hydroperoxide suggests that Rv1049 may be part of a generalised stress response of the bacterium, as has been reported for other MarR-type regulatory genes, or may simply be a case of the bacteria being overwhelmed by high doses of damaging agent to which they are unable to respond.

Neither the Rv0465c or the Rv1049 null mutants were significantly attenuated in a macrophage model of infection, suggesting either that these genes do not play a role in the growth of *M. tuberculosis* in macrophages or that the macrophage model of infection used did not provide the stimulus to which these potential regulators would react *in vivo*. Growth of the Rv0465c and Rv1049 null mutants, and potential complementing strains, in an intravenous mouse model of infection showed that the Rv0465c null mutant was present at significantly higher numbers than the wild-type at day 63 in the spleen. This may represent a subtle *in vivo* growth phenotype for the Rv0465c null mutant.

7.5 Future work

The interaction of RegX3 with a number of the promoters seen to be differentially regulated in the strain overexpressing the *regX3* transcript (section 4.2.3) could be confirmed with electrophoretic mobility shift assay and DNA footprinting using the purified recombinant RegX3 protein. The microarray results could also be confirmed by quantitative PCR.

To further investigate the role of the SenX3-RegX3 two component signal transduction system in sensing low levels of oxygen, two collaborations have been set up in order to better control the oxygen levels in cultures of the *senX3* and *regX3* null mutants. Dr. J. Bacon (HPA, Porton Down) has grown the *senX3* null mutant in a chemostat and subjected it to decreasing levels of oxygen. Future microarray analysis of samples of the culture at defined oxygen tensions and comparison to the wild-type data previously generated at CAMR should provide a definitive analysis of the low-oxygen regulon of the SenX3-RegX3 two component signal transduction system and suggest the oxygen concentration at which the system becomes active. Dr. T. Parish (St. Barts, London) is currently growing both the *senX3* and *regX3* null mutants in the well-documented 'Wayne' model of anaerobic shutdown (Wayne & Hayes 1996).

The role of SenX3 and RegX3 during infection could be analysed by assessing the growth of the *senX3* and *regX3* null mutants in mice deficient in aspects of the innate immune defence, such as iNOS and phox mice, which are unable to produce inducible nitric oxide synthase and phagocyte oxidase, respectively. The *senX3* and *regX3* null strains could also be grown in models of infection, such as the guinea pig model, that produce a granulatomous response to infection. Furthermore, the viability model used

throughout chapter 3 (see 2.2.1.1) could be used with a variety of different superoxide generating agents not tested in this work to more thoroughly examine the promising *in vitro* results seen here. If the PAS domain of SenX3 contains a haem prosthetic group, then it may be able to conduct experiments using cyanide to inhibit the activity of the haem group and further study the system in *in vitro* models where a phenotype can be observed for the *senX3* and *regX3* null mutants.

To further investigate the role of Rv0465c and Rv1049 in the response of *M. tuberculosis* to reactive oxygen intermediates, the growth of these strains is currently being assessed in a mouse model of infection. Further work could include microarray or proteomic analysis of the Rv0465c and Rv1049 null mutant strains and their complementing strains, or assessment in other models of infection.

7.6 Conclusion

Overall, the results of this study show that the SenX3-RegX3 two component signal transduction system is essential for the full virulence of *M. tuberculosis* in mice and macrophages. Null mutants in this system are more resistant than the wild-type to killing by superoxide and organic hydroperoxide stress *in vitro*, suggesting that SenX3 and RegX3 are involved in responding to oxidative stress during infection. Microarray and protein-DNA interaction studies build on homology data suggesting that SenX3 may be sensing oxygen, indicating that the SenX3-RegX3 two component signal transduction system is active under low oxygen conditions and regulates a number of other regulatory genes, as well as sets of genes known to be involved in oxidative stress defence and survival in conditions thought to occur within granulatomous lesions. A collaboration

has been set up to enable a more controlled examination of the low-oxygen regulon of the SenX3-RegX3 two component signal transduction system.

Creation and analysis of other genes that may be involved in the oxidative stress response of *M. tuberculosis* has shown that null mutants of both Rv0465c and Rv1049 grow normally *in vitro* aerobically, and that the Rv1049 gene product is likely to be involved in the resistance of *M. tuberculosis* to killing by organic hydroperoxides. Neither the Rv0465c nor the Rv1049 null mutants were attenuated in a macrophage model of infection, suggesting that these genes are not crucial of the survival of *M. tuberculosis* in the early stages of infection when they encounter the oxidative burst of the macrophage. It may be that the protein encoded by Rv1049 is acting like some other regulators to which it has homology and forms part of a generalised stress response of *M. tuberculosis*; alternatively Rv1049 or Rv0465c may be acting like previously described proteins that have a function in the oxidative stress response of *E. coli*. Further data on these two mutants will provide a clearer picture of their potential role in the oxidative stress response of *M. tuberculosis*.

This study has increased our knowledge of the oxidative stress responses of *M. tuberculosis*, which form part of the interaction between *M. tuberculosis* and the host. Understanding of this relationship may allow further development of antimycobacterial drugs and identification of potential vaccine targets, which is particularly relevant for the SenX3-RegX3 two component signal transduction system since two component signal transduction systems represent a promising target for potential drug design (Stephenson & Hoch 2004). Furthermore, mutations in the SenX3-RegX3 two component signal

transduction system such as those described here could be used a component of a future live attenuated vaccine strain.

Appendix 1: Genes differentially regulated in null mutants of the SenX3-RegX3 two component signal transduction system after growth in low oxygen conditions

senX3 null mutant vs. H37Rv after microaerobic growth

negative fold change values are upregulated in the mutant

Rv number	fold change	p value	common name
Rv0042c	0.416	0.00201	Rv0042c
Rv0111	0.19	0.0436	Rv0111
Rv0174	0.605	0.00797	Rv0174
Rv0232	30.21	0.0483	Rv0232
Rv0297	0.33	0.013	PE PGRS
Rv0343	0.484	0.0395	Rv0343
Rv0383c	0.313	0.042	Rv0383c
Rv0408	376.4	0.00342	pta
Rv0483	0.613	0.011	Rv0483
Rv0490	107.8	0.0228	senX3
Rv0491	157.5	0.0265	regX3
Rv0637	0.287	0.000775	Rv0637
Rv0883c	0.542	0.0426	Rv0883c
Rv0896	0.575	0.0111	gltA2
Rv0932c	0.727	0.0233	pstS
Rv0968	0.34	0.00513	Rv0968
Rv1104	0.207	0.0303	Rv1104
Rv1141c	0.192	0.0453	echA11
Rv1172c	0.299	0.00748	PE
Rv1347c	0.48	0.05	Rv1347c
Rv1398c	0.391	0.0308	Rv1398c
Rv1460	0.54	0.0429	Rv1460
Rv1471	0.396	0.0284	trxB
Rv1518	22.12	0.0254	Rv1518
Rv1632c	0.679	0.00293	Rv1632c
Rv1642	0.151	0.0211	rpmI
Rv1736c	62.84	0.0216	narX
Rv1894c	0.461	0.027	Rv1894c
Rv1952	39.88	0.0138	Rv1952
Rv1955	0.48	0.0101	Rv1955
Rv1980c	0.576	0.0281	mpt64
Rv1988	0.0303	0.0332	Rv1988
Rv2007c	0.481	0.000621	fdxA
Rv2094c	0.454	0.0497	Rv2094c
Rv2116	0.444	0.00544	lppK
Rv2197c	0.3	0.0219	Rv2197c
Rv2461c	0.524	0.0121	clpP
Rv2626c	0.595	0.0425	Rv2626c
Rv2658c	0.291	0.0157	Rv2658c
Rv2706c	0.548	0.0246	Rv2706c
Rv2820c	0.194	0.0194	Rv2820c
Rv2954c	0.363	0.0435	Rv2954c
Rv3049c	0.498	0.0129	Rv3049c
Rv3051c	0.525	0.0317	nrdE
Rv3089	0.301	0.0194	fadD13
Rv3301c	8.757	0.00551	phoY1
Rv3409c	0.545	0.00469	choD
Rv3585	0.351	0.0442	radA
Rv3679	0.56	0.022	Rv3679
Rv3680	0.478	0.00818	Rv3680
Rv3682	0.498	0.0106	ponA'
Rv3741c	0.0884	0.0318	Rv3741c
Rv3750c	0.357	0.0238	Rv3750c
Rv3853	0.343	0.031	menG
Rv3854c	0.431	0.0255	Rv3854c
Rv3857c	0.428	0.0242	Rv3857c
Rv3886c	0.343	0.0285	Rv3886c
Rv3893c	0.521	0.0218	PE
Rv3924c	0.233	0.0181	rpmH

regX3 null mutant vs. H37Rv after microaerobic growth

negative fold change values are upregulated in the mutant

Rv number	fold change	p value	common name
Rv0064	7.367	0.0429	Rv0064
Rv0065	0.418	0.0119	Rv0065
Rv0332	0.201	0.0133	Rv0332
Rv0410c	4.821	0.0242	pknG
Rv0489	0.195	0.046	gpm
Rv0497	0.21	0.027	Rv0497
Rv0511	4.337	0.00188	cysG
Rv0637	0.578	0.0437	Rv0637
Rv0654	0.362	0.00191	Rv0654
Rv0678	0.516	0.0178	Rv0678
Rv0715	0.156	0.0259	rplX
Rv0748	0.144	0.0485	Rv0748
Rv0790c	0.232	0.0423	Rv0790c
Rv0823c	4.97	0.0138	Rv0823c
Rv1165	2.746	0.00801	Rv1165
Rv1327c	4.611	0.0218	Rv1327c
Rv1382	3.215	0.0223	Rv1382
Rv1493	3.768	0.00244	mutB
Rv1566c	0.429	0.0287	Rv1566c
Rv1584c	0.273	0.0489	Rv1584c
Rv1594	10.14	0.0214	nadA
Rv1595	6.948	0.0076	nadB
Rv1599	3.326	0.0372	hisD
Rv1600	3.134	0.0386	hisC
Rv1601	2.26	0.00105	hisB
Rv1602	2.261	0.012	hisH
Rv1632c	0.297	0.00799	Rv1632c
Rv1654	57.84	0.0485	argB
Rv1655	3.569	0.00119	argD
Rv1658	3.747	0.0184	argG
Rv1736c	75.55	0.0325	narX
Rv1952	30.28	0.0151	Rv1952
Rv1966	19.6	0.0354	mce3
Rv1970	3.965	0.0353	lprM
Rv1996	16.96	0.0248	Rv1996
Rv2035	0.438	0.0433	Rv2035
Rv2318	0.0332	0.0215	uspC
Rv2378c	5.386	0.033	mbtG
Rv2381c	5.202	0.000271	mbtD
Rv2382c	5.9	0.00141	mbtC
Rv2383c	3.363	0.0251	mbtB
Rv2484c	2.974	0.0269	Rv2484c
Rv2496c	5.111	0.0401	pdhB
Rv2498c	2.241	0.0382	citE
Rv2499c	6.953	0.0382	Rv2499c
Rv2500c	2.865	0.00146	fadE19
Rv2504c	3.297	0.00626	scoA
Rv2553c	3.149	0.0329	Rv2553c
Rv2628	42.24	0.0107	Rv2628
Rv2630	6.399	0.0019	Rv2630
Rv2631	7.444	0.00212	Rv2631
Rv2658c	0.525	0.0383	Rv2658c
Rv2706c	0.327	0.0028	Rv2706c
Rv2895c	0.0768	0.039	viuB
Rv2935	0.152	0.00591	ppsE
Rv3042c	0.471	0.0459	serB2
Rv3205c	0.575	0.0459	Rv3205c
Rv3206c	0.256	0.0178	moeZ
Rv3248c	4.98	0.0137	sahH
Rv3301c	18.73	0.026	phoY1
Rv3329	0.623	0.0491	Rv3329
Rv3551	8.455	0.0241	Rv3551
Rv3736	0.0834	0.0252	Rv3736
Rv3785	0.567	0.0374	Rv3785
Rv3800c	6.244	0.000756	pks13

Tame15 mutant vs. H37Rv after microaerobic growth

negative fold change values are upregulated in the mutant

Rv number	fold change	p value	common name
Rv0174	0.617	0.0405	Rv0174
Rv0275c	0.455	0.000596	fadD27
Rv0283	0.374	0.00553	Rv0283
Rv0289	0.512	0.0123	Rv0289
Rv0297	0.457	0.0438	PE PGRS
Rv0343	0.559	0.0166	Rv0343
Rv0352	0.589	0.0173	dnaJ
Rv0393	0.622	0.0229	Rv0393
Rv0407	2.064	0.00889	Rv0407
Rv0483	0.517	0.0408	Rv0483
Rv0491	160.7	0.00207	regX3
Rv0497	0.493	0.00128	Rv0497
Rv0502	0.633	0.049	Rv0502
Rv0569	0.459	0.0127	Rv0569
Rv0572c	0.258	0.00761	Rv0572c
Rv0637	0.356	0.0457	Rv0637
Rv0715	0.338	0.0169	rplX
Rv0757	0.44	0.0195	phoP
Rv0848	2.316	0.0445	cysM3
Rv0876c	0.527	0.00366	Rv0876c
Rv0883c	0.492	0.0294	Rv0883c
Rv0886	0.468	0.0163	fprB
Rv0932c	0.591	0.0478	pstS
Rv0996	0.668	0.0495	Rv0996
Rv1015c	0.476	0.00287	rplY
Rv1017c	0.547	0.0199	prsA
Rv1026	2.159	0.0428	Rv1026
Rv1078	0.507	0.0201	pra
Rv1130	0.152	0.00991	Rv1130
Rv1148c	0.493	0.0201	Rv1148c
Rv1172c	0.312	0.00567	PE
Rv1178	0.426	7.61E-05	Rv1178
Rv1237	0.353	0.0183	sugB
Rv1361c	0.436	0.00366	PPE
Rv1374c	0.445	0.0265	Rv1374c
Rv1387	0.508	0.00705	PPE
Rv1398c	0.28	0.0465	Rv1398c
Rv1428c	0.462	0.0224	Rv1428c
Rv1462	0.68	0.0394	Rv1462
Rv1464	0.615	0.0243	Rv1464
Rv1590	14.05	0.0499	Rv1590
Rv1601	0.652	0.0389	hisB
Rv1639c	0.594	0.0134	Rv1639c
Rv1733c	0.438	0.00467	Rv1733c
Rv1736c	28.35	0.00164	narX
Rv1791	0.461	0.0396	PE
Rv1885c	0.416	0.0106	Rv1885c
Rv1925	0.668	0.0393	fadD31
Rv1952	47.77	0.0139	Rv1952
Rv1954c	0.387	0.00724	Rv1954c
Rv1955	0.231	0.00428	Rv1955
Rv1957	0.297	0.0481	Rv1957
Rv1988	0.0448	0.043	Rv1988
Rv1992c	0.722	0.038	ctpG
Rv2004c	0.501	0.0233	Rv2004c
Rv2005c	0.469	0.0059	Rv2005c
Rv2007c	0.333	0.00192	fdxA
Rv2031c	0.141	0.03	hspX
Rv2056c	0.423	0.0364	rpsN2
Rv2057c	0.204	0.0214	rpmG
Rv2116	0.488	0.0491	lppK
Rv2159c	0.507	0.0441	Rv2159c
Rv2193	0.671	0.0458	ctaE
Rv2197c	0.389	0.0498	Rv2197c
Rv2216	0.665	0.0438	Rv2216
Rv2243	0.418	0.000864	fabD
Rv2280	0.594	0.027	Rv2280
Rv2373c	0.628	0.0486	dnaJ2
Rv2461c	0.599	0.0383	clpP
Rv2462c	0.487	0.0317	tig
Rv2465c	0.532	0.0388	rpi
Rv2497c	0.631	0.0271	pdhA
Rv2569c	0.573	0.0174	Rv2569c
Rv2609c	0.58	0.0474	Rv2609c

Rv2623	0.592	0.0296	Rv2623
Rv2624c	0.645	0.022	Rv2624c
Rv2658c	0.336	0.00913	Rv2658c
Rv2706c	0.411	0.0021	Rv2706c
Rv2782c	2.542	0.0132	pepR
Rv2814c	0.258	0.0444	Rv2814c
Rv2820c	0.359	0.00621	Rv2820c
Rv2841c	0.62	0.00959	nusA
Rv2935	0.197	0.00943	ppsE
Rv2936	0.444	0.00233	drrA
Rv2940c	5.79	0.0212	mas
Rv2963	2.26	0.00797	Rv2963
Rv2990c	0.614	0.0387	Rv2990c
Rv3043c	0.59	0.0494	ctaD
Rv3069	6.484	0.0255	Rv3069
Rv3089	0.469	0.00755	fadD13
Rv3121	0.374	0.0279	Rv3121
Rv3133c	0.436	0.00513	Rv3133c
Rv3134c	0.553	0.0173	Rv3134c
Rv3150	0.52	0.0173	nuoF
Rv3239c	7.442	0.0453	Rv3239c
Rv3270	1.973	0.0216	ctpC
Rv3271c	2.01	0.00242	Rv3271c
Rv3301c	12.02	0.00242	phoY1
Rv3371	0.582	0.0351	Rv3371
Rv3398c	200.2	0.0351	idsA
Rv3422c	3.821	0.0343	Rv3422c
Rv3479	0.51	0.0024	Rv3479
Rv3615c	0.123	0.0068	Rv3615c
Rv3616c	0.551	0.033	Rv3616c
Rv3648c	0.222	0.0269	cspA
Rv3680	0.553	0.011	Rv3680
Rv3764c	0.559	0.0476	Rv3764c
Rv3801c	0.483	0.0169	fadD32
Rv3814c	0.386	0.0158	Rv3814c
Rv3853	0.282	0.0223	menG
Rv3854c	0.401	0.000926	Rv3854c
Rv3874	0.321	0.0323	Rv3874

Appendix 2: Genes differentially regulated between a ΔregX3/P_{hsp60}-*regX3*$^{+}$ strain and a ΔregX3/P_{hsp60}-*regX3*$^{-}$ strain of *M. tuberculosis*.

regX3 overexpression gene list

negative fold change values are upregulated in the overexpressing strain

Rv number	fold change	p value	common name		Rv number	fold change	p value	common name	
Rv0080	0.665	0.0149	Rv0080		Rv1372	35.88	0.000821	pks18	
Rv0093c	**0.521**	**0.0301**	**Rv0093c**		Rv1387	67.97	0.0375	PPE	
Rv0094c	**0.694**	**0.0139**	**Rv0094c**		Rv1409	0.693	0.0366	ribG	
Rv0110	8.225	0.0206	Rv0110		Rv1491c	0.368	0.000486	Rv1491c	
Rv0115	0.613	0.0493	Rv0115		Rv1556	12.7	0.0369	Rv1556	
Rv0127	0.398	0.0283	Rv0127		Rv1587c	0.717	0.0148	Rv1587c	
Rv0160c	**0.544**	**0.0403**	**PE**		Rv1588c	0.702	0.00664	Rv1588c	
Rv0161	**0.341**	**0.00553**	**Rv0161**		Rv1594	0.617	0.0403	nadA	
Rv0166	4.722	0.00208	fadD5		Rv1610	0.017	0.0232	Rv1610	
Rv0172	**1.695**	**0.0351**	**Rv0172**		**Rv1622c**	**0.0857**	**0.0209**	**cydB**	*
Rv0173	**1.927**	**0.0162**	**lprK**		**Rv1623c**	**0.1**	**0.00407**	**appC**	
Rv0211	0.615	0.0117	pckA		Rv1632c	0.76	0.0233	Rv1632c	
Rv0222	9.548	0.00911	echA1		**Rv1731**	**16.54**	**0.0303**	**gabD1**	
Rv0250c	**2.717**	**0.00957**	**Rv0250c**		**Rv1732c**	**1.408**	**0.0343**	**Rv1732c**	
Rv0251c	**2.314**	**0.00712**	**hsp**		Rv1775	2.713	0.0416	Rv1775	
Rv0256c	0.363	0.0457	PPE		Rv1784	0.58	0.0188	Rv1784	
Rv0296c	1.687	0.00448	atsG		Rv1789	0.378	1.71E-06	PPE	
Rv0308	1.72	0.0466	Rv0308		Rv1795	0.452	0.00411	Rv1795	
Rv0344c	1.472	0.0283	lpqJ		Rv1797	0.36	0.0196	Rv1797	
Rv0361	0.725	0.0143	Rv0361		Rv1813c	2.878	0.0312	Rv1813c	
Rv0369c	0.386	0.00419	Rv0369c		Rv1830	0.82	0.0483	Rv1830	
Rv0377	0.454	0.0157	Rv0377		Rv1849	7.766	0.02	ureB	
Rv0411c	**0.199**	**0.00431**	**glnH**		Rv1857	7.092	0.028	modA	
Rv0412c	**0.368**	**1.15E-05**	**Rv0412c**		Rv1890c	0.229	0.0194	Rv1890c	
Rv0438c	6.158	0.00186	moaA3		**Rv1898**	**1.836**	0.0186	Rv1898	
Rv0479c	0.726	0.015	Rv0479c		Rv1956	1.913	0.0387	Rv1956	
Rv0485	0.588	0.00968	Rv0485		Rv2006	6.938	0.0431	otsB	
Rv0490	**0.492**	**0.00245**	**senX3**		Rv2015c	0.789	0.0449	Rv2015c	
Rv0491	**2.94E-06**	**1.73E-08**	**regX3**	*	Rv2036	10.91	0.029	Rv2036	
Rv0513	0.101	0.0328	Rv0513		Rv2056c	0.694	0.00629	rpsN2	
Rv0516c	0.392	0.0064	Rv0516c	*	Rv2082	0.734	0.000917	Rv2082	
Rv0557	0.0486	0.00557	Rv0557		Rv2148c	0.233	0.0443	Rv2148c	
Rv0559c	1.31	0.0123	Rv0559c		Rv2177c	0.559	0.0345	Rv2177c	
Rv0561c	22.29	0.0484	Rv0561c		Rv2193	0.568	0.0384	ctaE	
Rv0578c	0.343	0.00198	PE PGRS		Rv2202c	0.57	0.0412	cbhK	
Rv0621	13.68	0.0127	Rv0621		Rv2234	10.78	0.0455	ptpA	
Rv0650	0.6	0.0426	Rv0650		Rv2243	1.779	0.038	fabD	
Rv0683	1.907	0.0197	rpsG		Rv2276	0.633	0.00389	Rv2276	
Rv0706	1.258	0.0124	rplV	*	Rv2419c	21.06	0.019	Rv2419c	
Rv0749	0.535	0.0309	Rv0749		**Rv2428**	**0.632**	**0.00153**	**ahpC**	*
Rv0784	0.025	0.000132	Rv0784		**Rv2429**	**0.615**	**0.00256**	**ahpD**	*
Rv0823c	0.636	0.0205	Rv0823c		**Rv2430c**	**0.704**	**0.00306**	**PPE**	*
Rv0891c	0.548	0.0407	Rv0891c		**Rv2492**	**2.01**	**0.043**	**Rv2492**	
Rv0905	**0.462**	**0.0256**	**echA6**		**Rv2493**	**7.488**	**0.0203**	**Rv2493**	
Rv0906	**0.362**	**0.0236**	**Rv0906**		Rv2495c	0.282	0.0348	pdhC	
Rv0928	**0.00356**	**0.00326**	**phoS2**		Rv2510c	7.095	0.000933	Rv2510c	
Rv0929	**0.0205**	**0.00102**	**pstC2**		Rv2543	0.135	0.0401	lppA	
Rv0931c	0.731	0.0467	pknD		Rv2577	0.0151	0.00613	Rv2577	
Rv1008	3.592	0.0361	Rv1008		Rv2582	1.499	0.00174	ppiB	
Rv1011	4.544	0.026	Rv1011		Rv2591	1.394	0.00532	PE PGRS	
Rv1038c	0.67	0.015	Rv1038c		Rv2594c	1.991	0.0186	ruvC	*
Rv1041c	1.543	0.00781	Rv1041c		Rv2645	5.019	0.0309	Rv2645	
Rv1067c	0.0301	0.0121	PE PGRS		Rv2661c	2.46	0.0431	Rv2661c	
Rv1102c	1.568	0.0277	Rv1102c		**Rv2692**	**66.55**	**0.0147**	**trkB**	
Rv1130	0.351	0.0172	Rv1130		**Rv2693c**	**19.12**	**0.0118**	**Rv2693c**	
Rv1147	1.43	0.00037	Rv1147		Rv2698	21.04	0.0236	Rv2698	
Rv1150	1.491	0.00458	Rv1150		**Rv2711**	**1.291**	**0.0264**	**ideR**	
Rv1158c	2.064	0.000347	Rv1158c		**Rv2712c**	**20.13**	**0.0369**	**Rv2712c**	
Rv1197	0.662	0.000337	Rv1197		Rv2719c	1.942	0.0159	Rv2719c	
Rv1224	1.378	0.035	Rv1224		Rv2728c	6.429	0.0164	Rv2728c	
Rv1228	0.522	0.0395	lpqX		Rv2739c	2.63	0.0277	Rv2739c	
Rv1230c	1.581	0.0165	Rv1230c		Rv2757c	21.79	0.0073	Rv2757c	*
Rv1277	0.66	0.0102	Rv1277		Rv2829c	2.176	0.0382	Rv2829c	
Rv1284	11.56	0.0286	Rv1284	*	Rv2831	22.87	0.032	echA16	
Rv1295	0.68	0.0357	thrC		Rv2832c	7.322	0.0037	ugpC	
Rv1297	0.782	0.011	rho		Rv2838c	1.68	0.0171	rbfA	
Rv1300	0.636	0.0116	hemK		Rv2840c	1.449	0.00617	Rv2840c	
Rv1304	0.652	0.0284	atpB		Rv2864c	19.3	0.0152	Rv2864c	
Rv1306	**0.728**	**0.0201**	**atpF**		Rv2894c	10.39	0.00175	xerC	
Rv1307	**0.719**	**0.0153**	**atpH**		Rv2930	0.612	0.00268	fadD26	
Rv1316c	1.319	0.0256	ogt		**Rv2934**	**0.572**	**0.0338**	**ppsD**	
Rv1320c	6.182	0.00582	Rv1320c		**Rv2935**	**0.725**	**0.00827**	**ppsE**	
Rv1339	0.502	0.00389	Rv1339		**Rv2944**	**1.797**	**0.0455**	**Rv2944**	
Rv1360	0.664	0.0262	Rv1360		**Rv2945c**	**0.738**	**0.0327**	**lppX**	
Rv1361c	3.418	0.0133	PPE		Rv2984	0.496	0.0353	ppk	

Rv number	fold change	p value	common name	
Rv2990c	0.665	0.0429	Rv2990c	
Rv3045	0.652	0.00203	adhC	
Rv3062	19.55	0.0411	ligB	
Rv3095	1.388	0.00661	Rv3095	
Rv3147	1.613	0.0102	nuoC	
Rv3165c	3.562	0.0489	Rv3165c	
Rv3175	11.77	0.0351	Rv3175	
Rv3180c	15.41	0.043	Rv3180c	
Rv3182	6.328	0.0272	Rv3182	
Rv3191c	**1.383**	**0.0231**	**Rv3191c**	
Rv3192	**2.236**	**0.0161**	**Rv3192**	
Rv3195	4.153	0.0474	Rv3195	
Rv3224	1.75	9.91E-05	Rv3224	
Rv3229c	20.77	0.0123	desA3	*
Rv3248c	0.642	0.000566	sahH	
Rv3285	1.511	0.0162	accA3	
Rv3290c	0.232	6.07E-07	lat	
Rv3305c	13.93	0.0299	amiA	
Rv3355c	21.29	0.0338	Rv3355c	
Rv3370c	0.0449	0.0059	dnaE2	
Rv3381c	17.33	0.00292	Rv3381c	
Rv3395c	2.822	0.043	Rv3395c	
Rv3408	1.345	0.0482	Rv3408	
Rv3418c	1.976	0.000761	groES	
Rv3466	**0.708**	**0.00554**	**Rv3466**	
Rv3467	**0.714**	**0.00501**	**Rv3467**	
Rv3478	**2.556**	**0.0399**	**PPE**	
Rv3479	**26.22**	**0.00596**	**Rv3479**	
Rv3486	1.362	0.00571	Rv3486	
Rv3524	0.512	0.00253	Rv3524	
Rv3587c	1.386	0.00664	Rv3587c	*
Rv3619c	0.649	0.0149	Rv3619c	
Rv3620c	0.66	0.00136	Rv3620c	
Rv3634c	1.733	0.0174	rmlB2	
Rv3640c	37.45	0.00673	Rv3640c	
Rv3642c	2.025	0.0159	Rv3642c	
Rv3707c	0.604	0.0247	Rv3707c	
Rv3720	0.731	0.0238	Rv3720	
Rv3745c	0.0521	0.0457	Rv3745c	
Rv3750c	1.752	0.0384	Rv3750c	*
Rv3839	**2.662**	**0.00237**	**Rv3839**	
Rv3840	**5.009**	**0.0297**	**Rv3840**	
Rv3842c	0.779	0.0436	glpQ1	
Rv3867	0.623	0.0462	Rv3867	
Rv3870	0.842	0.0369	Rv3870	
Rv3873	0.699	0.0146	PPE	
Rv3888c	0.575	0.0331	Rv3888c	
Rv3914	1.611	0.0347	trxC	
Rv3919c	0.677	0.0204	gid	

Appendix 3: Oligonucleotide primers

Restriction sites are underlined

primer	sequence
104	cgacgagcgacccgagc
114	tctagagcagtcagttcagccagga
115	atcgataccagacagtcgccaaggtt
119	gcggccgcatatgtgtagttacacccaactg
122	gcggccgcgtgaccatcgaccggatcct
133	aattcgggatccacgccgatgtgtccaaggagat
134	aattcgggatcctgaaacaactatgaccggtcagaacg
135	aattaaggaaaaaagcggccgcgagcgtaccgggtgttgtgaaga
136	aattaaggaaaaaagcggccgcgctgcagtgcgttgttgatcc
137	aattcgggatccaagtcgagccgtcgcagtagtga
138	aattcgggatcccattactccagactcgtgtagctacacgtattt
139	acttggccgttctgacc
140	aatggcgtcttcacaacacc
147	ggctttgacgggtcttgaca
148	gatcaacaccacctggtcac
151	cgggatccatgaccagtgtgttgattgtggaggac
152	ccatcgatctagccctcgagtttgtagc
153	cgggatccaatgaccagtgtgttgattgtggaggac
156	ctgctcgccttcaccttcct
157	gcaggctcgcgtaggaatca
160	atgtgaacggtaaccgaacagctgtggcgtagtgtgtgactt

161 aagtcacacactacgccacagctgttcggttaccgttcacat

162 gatccgggctatcagtgcgtgggaaagcccggttgctgagcga

163 tcgctcagcaaccgggctttcccacgcactgatagcccggatc

168 ccatcgatgctagccctcgagtttgtagc

169 ttcgtggcggtcaggatcgt

170 tgaaggccatctacctgtcg

177 ctcgtatcgtcgtcacagtc

178 gtgcgagtcgctatgtc

179 acggcgttgacgtgaccatc

180 cagttctccaccgaagtatc

181 aagctgttcggcgagcactg

182 aagtcctcggtccgcttcca

183 gttcgacgaatgcgcttgct

184 tcagctcgatgcgcttgcac

185 taggtgtggccgcagtattc

186 gccagcaggtgttcaagatg

189 aattggatccgtcgcgtggtgttattgctg

190 aattggatcctgaccgatgtcctggatgtg

191 ttccagcgcactcac

192 gagatgatcggcgttcc

237 ctgacaaagagcgccccgctgat

238 ccatatcgtggccgccagga

241 agctcggtagtggtcagcgg

242	tgacaaaattccgttgatcg
243	gccatgcagcaagctcagga
244	gattacccctctcgatactg
PMV306-4	ccagcaacgcggccttttacg
d153	ggtcagcgaaccaatca

Appendix 4: Media and solutions

Media

Dubos broth

K_2HPO_4	1g
$Na_2HPO_4.12H_20$	6.25g
Na_3 citrate	1.25g
$MgSO_4.7H_2O$	0.6g
Asparagine	2g
10% Tween 80	5ml
Casamino acids (Difco)	2g

pH to 7.2 (NaOH), make up to 960ml with dH_2O, autoclave 121°C for 15 min. Add 40ml albumin supplement (Difco) once cool.

Freezing medium

K_2HPO_4	12.6g
Na_3 citrate	0.9g
$MgSO_4.7H_2O$	0.18g
$(NH_4)_2SO_4$	1.8g
KH_2PO_4	3.6g
glycerol	96g
dH_2O	to 1 litre

Autoclave 121°C 15 min.

Luria-Bertani agar (L-agar)

Tryptone (Difco)	10g
Yeast extract	5g
NaCl	10g
Agar (Difco)	15g
dH_20	to 1 litre

Autoclave 121°C 15 min.

Luria-Bertani broth (L-broth)

Tryptone (Difco)	10g
Yeast extract	5g
NaCl	10g

pH to 7.5 (NaOH), make up to 1 litre with dH2O, autoclave 121°C for 15 min.

Middlebrook 7H11 agar

Glycerol	5ml
Tween 80	0.05% w/v
7H11 medium powder (Difco)	21g

Autoclave 121°C for 15 min. Add 100ml OADC (olaeic acid, albumin, dextrose and catalase) supplement (Difco) once cool.

Solutions

TBE buffer (10x)

Tris base	121g
Boric acid	61.83g
EDTA	18.612g

pH to 8.0 (HCl), make up to 1 litre with dH_2O

TE buffer

Tris-HCl, pH8.0	10mM
EDTA	1mM

GSE solution

Guanidinium chloride	6M
N-lauroylsarcosine	1% v/v
EDTA	20mM

SET solution

Sucrose	0.3M
Tris-Chloride pH8	50mM

PBS

Sodium phosphate	10mM
Sodium chloride	0.8%

Appendix 5: Raw data for stress experiments

strains	H37Rv, Δ*senX3* , Δ*regX3*
agent	paraquat
concentration	50mM
time	2 hours
culture	aerobic

	raw data		log values		
strain	**unstressed**	**stressed**	**log 10 unstressed**	**log 10 stressed**	**log change**
H37Rv	3.45E+07	5.05E+07	7.54E+00	7.70E+00	-1.65E-01
H37Rv	9.70E+07	1.50E+08	7.99E+00	8.18E+00	-1.89E-01
H37Rv	1.38E+08	1.29E+08	8.14E+00		
Δ*senX3*	1.20E+08	6.20E+07	8.08E+00	7.79E+00	2.87E-01
Δ*senX3*	1.78E+08	1.04E+08	8.25E+00	8.02E+00	2.33E-01
Δ*senX3*	1.93E+08	1.99E+08	8.29E+00	8.30E+00	-1.33E-02
Δ*regX3*	8.65E+07	7.90E+07	7.94E+00	7.90E+00	3.94E-02
Δ*regX3*	1.22E+08	8.60E+07	8.08E+00	7.93E+00	1.50E-01
Δ*regX3*	1.45E+08	1.46E+08	8.16E+00	8.16E+00	-4.49E-03

Student's T-test

comparison of log changes

2 tailed homoscedastic, equal variance

	H37Rv	Δ*senX3*	Δ*regX3*
H37Rv		0.062767315	0.02833629
Δ*senX3*			0.357314703
Δ*regX3*			

strains	H37Rv, Δ*senX3* , Δ*regX3*
agent	paraquat
concentration	50mM
time	24 hours
culture	aerobic

	raw data		log values		
stain	**unstressed**	**stressed**	**log 10 unstressed**	**log 10 stressed**	**log change**
H37Rv	1.01E+08	6.30E+05	8.00E+00	5.80E+00	2.20E+00
H37Rv	5.73E+07	2.40E+05	7.76E+00	5.38E+00	2.38E+00
H37Rv	9.82E+07		7.99E+00		
Δ*senX3*	1.29E+08	5.44E+06	8.11E+00	6.74E+00	1.37E+00
Δ*senX3*	1.35E+08	7.29E+06	8.13E+00	6.86E+00	1.27E+00
Δ*senX3*	1.37E+08	8.90E+05	8.14E+00	5.95E+00	2.19E+00
Δ*regX3*	2.32E+08	2.55E+06	8.37E+00	6.41E+00	1.96E+00
Δ*regX3*	1.42E+08	3.77E+07	8.15E+00	7.58E+00	5.77E-01
Δ*regX3*	1.87E+08	8.74E+07	8.27E+00	7.94E+00	3.30E-01

Student's T-test
comparison of log changes
2 tailed homocedastic, equal variance

	H37Rv	Δ*senX3*	Δ*regX3*
H37Rv		0.1710196	0.1351226
Δ*senX3*			0.325488
Δ*regX3*			

strains	H37Rv, Δ*senX3* , Δ*regX3*
agent	silver nitrite
concentration	3mM
time	24 hours
culture	aerobic, pH5.5

	raw data		log values		
strain	**unstressed**	**stressed**	**log 10 unstressed**	**log 10 stressed**	**log change**
H37Rv	1.18E+08	1.06E+08	8.07E+00	8.03E+00	4.64E-02
H37Rv	1.41E+08	1.01E+08	8.15E+00	8.01E+00	1.43E-01
H37Rv	1.34E+08	1.08E+08	8.13E+00		
Δ*senX3*	1.77E+08	1.52E+08	8.25E+00	8.18E+00	6.40E-02
Δ*senX3*	1.91E+08	1.58E+08	8.28E+00	8.20E+00	8.09E-02
Δ*senX3*	1.91E+08	1.39E+08	8.28E+00	8.14E+00	1.36E-01
Δ*regX3*	1.58E+08	1.31E+08	8.20E+00	8.12E+00	8.31E-02
Δ*regX3*	1.72E+08	1.47E+08	8.24E+00	8.17E+00	6.82E-02
Δ*regX3*	1.71E+08	1.60E+08	8.23E+00	8.20E+00	3.05E-02

Student's T-test
comparison of log changes
2 tailed homoscedastic, equal variance

	H37Rv	Δ*senX3*	Δ*regX3*
H37Rv		0.9880464	0.4707302
Δ*senX3*			0.2851766
Δ*regX3*			

strains	H37Rv, Δ*senX3* , Δ*regX3*
agent	silver nitrite
concentration	3mM
time	24 hours
culture	microaerobic, pH5.5

	raw data		log values		
strain	**unstressed**	**stressed**	**log 10 unstressed**	**log 10 stressed**	**log change**
H37Rv	5.90E+07	1.70E+07	7.77E+00	7.23E+00	5.40E-01
H37Rv	4.85E+07	1.75E+07	7.69E+00	7.24E+00	4.43E-01
Δ*senX3*	7.20E+07	1.60E+07	7.86E+00	7.20E+00	6.55E-01
Δ*senX3*	5.75E+07	2.70E+07	7.76E+00	7.43E+00	3.28E-01
Δ*regX3*	5.65E+07	1.85E+07	7.75E+00	7.27E+00	4.85E-01
Δ*regX3*	1.28E+08	4.95E+07	8.11E+00	7.69E+00	4.13E-01

Student's T-test
comparison of log changes
2 tailed homoscedastic, equal variance

	H37Rv	Δ*senX3*	Δ*regX3*
H37Rv		0.9995206	0.5540622
Δ*senX3*			0.8221876
Δ*regX3*			

strains	H37Rv, Δ*senX3* , Δ*regX3*
agent	*tert*-butyl hydroperoxide
concentration	2mM
time	24 hours
culture	aerobic

	raw data		log values		
strain	**unstressed**	**stressed**	**log 10 unstressed**	**log 10 stressed**	**log change**
H37Rv	7.80E+07	1.86E+06	7.89E+00	6.27E+00	1.62E+00
H37Rv	1.04E+08	5.72E+06	8.02E+00	6.76E+00	1.26E+00
H37Rv	9.00E+07	1.64E+06	7.95E+00	6.21E+00	1.74E+00
Δ*senX3*	1.34E+08	4.51E+06	8.13E+00	6.65E+00	1.47E+00
Δ*senX3*	1.12E+08	6.15E+06	8.05E+00	6.79E+00	1.26E+00
Δ*senX3*	1.30E+08	2.64E+06	8.11E+00	6.42E+00	1.69E+00
Δ*regX3*	8.40E+07	1.33E+07	7.92E+00	7.12E+00	8.02E-01
Δ*regX3*	7.64E+07	3.51E+07	7.88E+00	7.55E+00	3.38E-01
Δ*regX3*	1.06E+08	2.88E+07	8.03E+00	7.46E+00	5.66E-01

Student's T-test
comparison of log changes
2 tailed homoscedastic, equal variance

	H37Rv	Δ*senX3*	Δ*regX3*
H37Rv		0.7493036	0.0078501
Δ*senX3*			0.00774
Δ*regX3*			

strains	H37Rv, ΔRv0465c, ΔRv1049
agent	*tert*-butyl hydroperoxide
concentration	250mM
time	24 hours
culture	aerobic

	raw data		log values		
strain	**unstressed**	**stressed**	**log 10 unstressed**	**log 10 stressed**	**log change**
H37Rv	1.09E+08	2.02E+07	8.04E+00	7.31E+00	7.32E-01
H37Rv	1.66E+08	1.90E+07	8.22E+00	7.28E+00	9.42E-01
H37Rv	1.71E+08	2.42E+07	8.23E+00	7.38E+00	8.50E-01
ΔRv0465c	1.19E+08	1.40E+07	8.08E+00	7.15E+00	9.29E-01
ΔRv0465c	1.51E+08	1.38E+07	8.18E+00	7.14E+00	1.04E+00
ΔRv0465c	1.62E+08	1.58E+07	8.21E+00	7.20E+00	1.01E+00
ΔRv1049	1.21E+08	6.80E+06	8.08E+00	6.83E+00	1.25E+00
ΔRv1049	1.46E+08	6.00E+06	8.16E+00	6.78E+00	1.39E+00
ΔRv1049	1.30E+08	6.60E+06	8.12E+00	6.82E+00	1.30E+00

Student's T-test

comparison of log changes

2 tailed homoscedastic, equal variance

	H37Rv	ΔRv0465c	ΔRv1049
H37Rv		0.09360084	0.00299508
ΔRv0465c			0.00359867
ΔRv1049			

strains	H37Rv, ΔRv0465c, ΔRv1049
agent	*tert*-butyl hydroperoxide
concentration	2mM
time	24 hours
culture	aerobic

	raw data		log values		
stain	**unstressed**	**stressed**	**log 10 unstressed**	**log 10 stressed**	**log change**
H37Rv	7.80E+07	1.86E+06	7.89E+00	6.27E+00	1.62E+00
H37Rv	1.04E+08	5.72E+06	8.02E+00	6.76E+00	1.26E+00
H37Rv	9.00E+07	1.64E+06	7.95E+00	6.21E+00	1.74E+00
ΔRv0465c	1.38E+08	1.14E+06	8.14E+00	6.06E+00	2.08E+00
ΔRv0465c	7.60E+07	1.04E+06	7.88E+00	6.02E+00	1.86E+00
ΔRv0465c	1.12E+08	2.26E+06	8.05E+00	6.35E+00	1.70E+00
ΔRv1049	1.48E+08	3.64E+06	8.17E+00	6.56E+00	1.61E+00
ΔRv1049	1.34E+08	1.52E+06	8.13E+00	6.18E+00	1.95E+00
ΔRv1049	4.02E+08	2.20E+06	8.60E+00	6.34E+00	2.26E+00

Student's T-test

comparison of log changes

2 tailed homoscedastic, equal variance

	H37Rv	ΔRv0465c	ΔRv1049
H37Rv		0.136579324	0.168810985
ΔRv0465c			0.80412749
ΔRv1049			

REFERENCES

Alekshun M.N. and Levy S.B. (1999). Alteration of the repressor activity of MarR, the negative regulator of the *Escherichia coli marRAB* Locus, by multiple chemicals *in vitro*. *J. Bacteriol.* **181**: 4669-4672.

Alexeeva S., Hellingwerf K.J. and Teixeira De Mattos M.J. (2003). Requirement of ArcA for redox regulation in *Escherichia coli* under microaerobic but not anaerobic or aerobic conditions. *J Bacteriol* **185**: 204-9.

Altschul S.F., Gish W., Miller W., Myers E.W. and Lipman D.J. (1990). Basic local alignment search tool. *J Mol Biol* **215**: 403-10.

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

Armstrong J.A. and Hart P.D.A. (1971). Response of cultured macrophages to *Mycobacterium tuberculosis*, with observations on fusion of lysosomes with phagosomes. *J Exp Med* **134**: 713.

Asusubel F., Brent R., Kingston R., Moore D., Seidman J., Smith J. and Struhl K. (2005). Current protocols in molecular biology. New York, John Wiley & Sons, Inc.

Av-Gay Y. and Everett M. (2000). The eukaryotic-like Ser/Thr protein kinases of *Mycobacterium tuberculosis*. *Trends Microbiol* **8**: 238-44.

Azad A.K., Sirakova T.D., Fernandes N.D. and Kolattukudy P.E. (1997). Gene knockout reveals a novel gene cluster for the synthesis of a class of cell wall lipids unique to pathogenic mycobacteria. *J Biol Chem* **272**: 16741-5.

Bacon J., James B.W., Wernisch L., Williams A., Morley K.A., Hatch G.J., Mangan J.A., Hinds J., Stoker N.G., Butcher P.D. and Marsh P.D. (2004). The influence of reduced oxygen availability on pathogenicity and gene expression in *Mycobacterium tuberculosis*. *Tuberculosis* **84**: 205-217.

Bailey T.L. and Elkan C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Second International Conference on Intelligent Systems for Molecular Biology, Menlo Park, California, AAAI Press.

Bateman A., Coin L., Durbin R., Finn R.D., Hollich V., Griffiths-Jones S., Khanna A., Marshall M., Moxon S., Sonnhammer E.L.L., Studholme D.J., Yeats C. and Eddy S.R. (2004). The Pfam protein families database. *Nucl. Acids Res.* **32**: D138-141.

Behr M.A. (2002). BCG--different strains, different vaccines? *Lancet Infect Dis* **2**: 86-92.

Benov L. and Fridovich I. (1995). A superoxide-dismutase mimic protects SodA SodB *Escherichia coli* against aerobic heating and stationary-phase death. *Archives of Biochemistry and Biophysics* **322**: 291-294.

Betts J.C., Lukey P.T., Robb L.C., Mcadam R.A. and Duncan K. (2002). Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. *Mol Microbiol* **43**: 717-31.

Bifani P.J., Mathema B., Kurepina N.E. and Kreiswirth B.N. (2002). Global dissemination of the *Mycobacterium tuberculosis* W-Beijing family strains. *Trends Microbiol* **10**: 45-52.

Boon C. and Dick T. (2002). *Mycobacterium bovis* BCG response regulator essential for hypoxic dormancy. *J Bacteriol* **184**: 6760-7.

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

Bryk R., Lima C.D., Erdjument-Bromage H., Tempst P. and Nathan C. (2002). Metabolic enzymes of mycobacteria linked to antioxidant defense by a thioredoxin-like protein. *Science* **295**: 1073-7.

Buck V., Quinn J., Soto Pino T., Martin H., Saldanha J., Makino K., Morgan B.A. and Millar J.B. (2001). Peroxide sensors for the fission yeast stress-activated mitogen-activated protein kinase pathway. *Mol Biol Cell* **12**: 407-19.

Buxton R.S. and Drury L.S. (1983). Cloning and insertional inactivation of the *dye* (*sfrA*) gene, mutation of which affects sex factor F expression and dye sensitivity of *Escherichia coli* K-12. *J Bacteriol* **154**: 1309-14.

Buxton R.S. and Drury L.S. (1984). Identification of the *dye* gene product, mutational loss of which alters envelope protein composition and also affects sex factor F expression in *Escherichia coli* K-12. *Mol Gen Genet* **194**: 241-7.

Buxton R.S., Drury L.S. and Curtis C.A. (1983). Dye sensitivity correlated with envelope protein changes in *dye* (*sfrA*) mutants of *Escherichia coli* K12 defective in the expression of the sex factor F. *J Gen Microbiol* **129**: 3363-70.

Camacho L.R., Constant P., Raynaud C., Laneelle M.A., Triccas J.A., Gicquel B., Daffe M. and Guilhot C. (2001). Analysis of the phthiocerol dimycocerosate locus of *Mycobacterium tuberculosis*. Evidence that this lipid is involved in the cell wall permeability barrier. *J Biol Chem* **276**: 19845-54.

Chan J. and Flynn J.L. (1999). Nitric oxide in *Mycobacterium tuberculosis* infection. Nitric oxide and infection. F. C. Fang. New York, Kluwer Academic and Plenum Publishing: 281-310.

Cheung A.L., Bayer A.S., Zhang G.Y., Gresham H. and Xiong Y.Q. (2004). Regulation of virulence determinants *in vitro* and *in vivo* in *Staphylococcus aureus*. *FEMS Immunology and Medical Microbiology* **40**: 1-9.

Christman M.F., Morgan R.W., Jacobson F.S. and Ames B.N. (1985). Positive control of a regulon for defenses against oxidative stress and some heat-shock proteins in *Salmonella typhimurium*. *Cell* **41**: 753-62.

Clarke A. and Rudd P. (1992). Neonatal BCG immunisation. *Arch Dis Child* **67**: 473-4.

Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eiglmeier K., Gas S., Barry C.E., 3rd, Tekaia F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., Mclean J., Moule S., Murphy L., Oliver K., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton J., Squares R., Squares S., Sulston J.E., Taylor K., Whitehead S. and Barrell B.G. (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**: 537-44.

Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R., Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D., Mungall K., Basham D., Brown D., Chillingworth T., Connor R., Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N., Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S., Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford

K.M., Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R. and Barrell B.G. (2001). Massive gene decay in the leprosy bacillus. *Nature* **409**: 1007-11.

Compan I. and Touati D. (1993). Interaction of six global transcription regulators in expression of manganese superoxide dismutase in *Escherichia coli* K-12. *J Bacteriol* **175**: 1687-96.

Compan I. and Touati D. (1994). Anaerobic activation of *arcA* transcription in *Escherichia coli*: roles of Fnr and ArcA. *Mol Microbiol* **11**: 955-64.

Cowley S., Babakalaiff R. and Av-Gay Y. (2002). Expression and localisation of the *Mycobacterium tuberculosis* protein tyrosine phosphatase PtpA. *Research in Microbiology* **153**: 233-241.

Cuff J.A., Clamp M.E., Siddiqui A.S., Finlay M. and Barton G.J. (1998). JPred: a consensus secondary structure prediction server. *Bioinformatics* **14**: 892-3.

Cunningham A.F. and Spreadbury C.L. (1998). Mycobacterial stationary phase induced by low oxygen tension: cell wall thickening and localization of the 16-kilodalton alpha-crystallin homolog. *J Bacteriol* **180**: 801-8.

Curry J.M., Whalan R., Hunt D.M., Gohil K., Strom M., Rickman L., Colston M.J., Smerdon S.J. and Buxton R.S. (2005). An ABC transporter containing a forkhead-associated domain interacts with a serine-threonine protein kinase and is required for growth of *Mycobacterium tuberculosis* in mice. *Infect. Immun.* **73**: 4471-4477.

De Voss J.J., Rutter K., Schroeder B.G., Su H., Zhu Y. and Barry C.E., 3rd (2000). The salicylate-derived mycobactin siderophores of *Mycobacterium tuberculosis* are essential for growth in macrophages. *PNAS* **97**: 1252-7.

Dedonder R. (1966). Levansucrase from *Bacillus subtilis*. *Methods Enzymol* **8**: 500-505.

Delgado-Nixon V.M., Gonzalez G. and Gillez-Gonzalez M. (2000). Dos, a heme-binding PAS protein from *Eschericia coli*, is a direct oxygen sensor. *Biochemistry* **39**: 2685 - 2691.

Demczuk, S., Harbers, M. & Vennstrom, B. (1993). Identification and analysis of all components of a gel retardation assay by combination with immunoblotting. *PNAS* **90**, 2574-2578.

Demple B. (1991). Regulation of bacterial oxidative stress genes. *Annu Rev Genet* **25**: 315-37.

Demple B. and Halbrook J. (1983). Inducible repair of oxidative DNA damage in *Escherichia coli*. *Nature* **304**: 466-8.

Deretic V., Philipp W., Dhandayuthapani S., Mudd M.H., Curcic R., Garbe T., Heym B., Via L.E. and Cole S.T. (1995). *Mycobacterium tuberculosis* is a natural mutant with an inactivated oxidative-stress regulatory gene: implications for sensitivity to isoniazid. *Mol Microbiol* **17**: 889-900.

Deretic V., Song J. and Pagan-Ramos E. (1997). Loss of OxyR in *Mycobacterium tuberculosis*. *Trends Microbiol* **5**: 367-72.

Dhandayuthapani S., Mudd M. and Deretic V. (1997). Interactions of OxyR with the promoter region of the *oxyR* and *ahpC* genes from *Mycobacterium leprae* and *Mycobacterium tuberculosis*. *J Bacteriol* **179**: 2401-9.

Dhandayuthapani S., Zhang Y., Mudd M.H. and Deretic V. (1996). Oxidative stress response and its role in sensitivity to isoniazid in mycobacteria: characterization

and inducibility of *ahpC* by peroxides in *Mycobacterium smegmatis* and lack of expression in *M. aurum* and *M. tuberculosis*. *J Bacteriol* **178**: 3641-9.

Drapal N. and Sawers G. (1995). Purification of ArcA and analysis of its specific interaction with the *pfl* promoter-regulatory region. *Mol Microbiol* **16**: 597-607.

Drury L.S. and Buxton R.S. (1985). DNA sequence analysis of the *dye* gene of *Escherichia coli* reveals amino acid homology between the Dye and OmpR proteins. *J Biol Chem* **260**: 4236-42.

Dukan S. and Nystrom T. (1998). Bacterial senescence: stasis results in increased and differential oxidation of cytoplasmic proteins leading to developmental induction of the heat shock regulon. *Genes Dev* **12**: 3431-41.

Dukan S. and Nystrom T. (1999). Oxidative stress defense and deterioration of growth-arrested *Escherichia coli* cells. *J Biol Chem* **274**: 26027-32.

Eisen M.B. and Brown P.O. (1999). DNA arrays for analysis of gene expression. *Methods Enzymol* **303**: 179-205.

Fine P.E.M. (1988). BCG vaccination against tuberculosis and leprosy. *British Medical Bulletin* **44**: 691.

Fine P.E.M. (1996). Randomised controlled trial of single BCG, repeated BCG, or combined BCG and killed *Mycobacterium leprae* vaccine for prevention of leprosy and tuberculosis in Malawi. *The Lancet* **348**: 17-24.

Fisher H.M. (1994). Genetic regulation of nitrogen fixation in rhizobia. *Microbiol. Rev.* **58**: 352-386

Florczyk M.A., Mccue L.A., Purkayastha A., Currenti E., Wolin M.J. and Mcdonough K.A. (2003). A family of *acr*-coregulated *Mycobacterium tuberculosis* genes shares a common DNA motif and requires rv3133c (*dosR* or *devR*) for expression. *Infect. Immun.* **71**: 5332-5343.

Flynn J.L. and Chan J. (2003). Immune evasion by *Mycobacterium tuberculosis*: living with the enemy. *Curr Opin Immunol* **15**: 450-5.

Fuangthong M., Atichartpongkul S., Mongkolsuk S. and Helmann J.D. (2001). OhrR is a repressor of *ohrA*, a key organic hydroperoxide resistance determinant in *Bacillus subtilis*. *J Bacteriol* **183**: 4134-41.

Fuangthong M. and Helmann J.D. (2002). The OhrR repressor senses organic hydroperoxides by reversible formation of a cysteine-sulfenic acid derivative. *PNAS* **99**: 6690-5.

Galperin M. and Gomelsky M. (2005). Bacterial signal transduction modules: from genomics to biology. *ASM News* **71**: 327-233.

Gamulin V., Cetkovic H. and Ahel I. (2004). Identification of a promoter motif regulating the major DNA damage response mechanism of *Mycobacterium tuberculosis*. *Fems Microbiology Letters* **238**: 57-63.

Georgellis D., Kwon O. and Lin E.C. (2001). Quinones as the redox signal for the arc two-component system of bacteria. *Science* **292**: 2314-6.

Georgellis D., Kwon O., Lin E.C., Wong S.M. and Akerley B.J. (2001). Redox signal transduction by the ArcB sensor kinase of *Haemophilus influenzae* lacking the PAS domain. *J Bacteriol* **183**: 7206-12.

Gilles-Gonzalez M.A. and Gonzalez G. (2004). Signal transduction by heme-containing PAS-domain proteins. *Journal of App. Physiol.* **96**: 774-783.

Gold B., Rodriguez G.M., Marras S.A., Pentecost M. and Smith I. (2001). The Mycobacterium tuberculosis IdeR is a dual functional regulator that controls transcription of genes involved in iron acquisition, iron storage and survival in macrophages. *Mol Microbiol* **42**: 851-65.

Gomez J.E., Chen J.M. and Bishai W.R. (1997). Sigma factors of *Mycobacterium tuberculosis*. *Tuber Lung Dis* **78**: 175-83.

Gomez M., Doukhan L., Nair G. and Smith I. (1998). *sigA* is an essential gene in *Mycobacterium smegmatis*. *Mol Microbiol* **29**: 617-28.

Gong W., Hao B., Mansy S.S. Gonzalez G., Gilles-Gonzalez M.A. and Chan M.K. (1998). Structure of a biological oxygen sensor: a new mechanism for heme-driven signal transduction. *PNAS* **95** 15177-15182.

Gopaul K.K., Brooks P.C., Prost J.F. and Davis E.O. (2003). Characterization of the two *Mycobacterium tuberculosis recA* promoters. *J Bacteriol* **185**: 6005-15.

Govantes F., Albrecht J.A. and Gunsalus R.P. (2000). Oxygen regulation of the *Escherichia coli* cytochrome d oxidase (*cydAB*) operon: roles of multiple promoters and the Fnr-1 and Fnr-2 binding sites. *Mol. Microbiol.* **37**: 1456-1469.

Green J. and Paget M.S. (2004). Bacterial redox sensors. *Nat Rev Microbiol* **2**: 954-66.

Greenberg J.T., Monach P., Chou J.H., Josephy P.D. and Demple B. (1990). Positive control of a global antioxidant defense regulon activated by superoxide-generating agents in *Escherichia coli*. *PNAS* **87**: 6181-5.

Grkovic S., Brown M.H. and Skurray R.A. (2002). Regulation of bacterial drug export systems. *Microbiol. Mol. Biol. Rev.* **66**: 671-701.

Grundner C., Gay L.M. and Alber T. (2005). *Mycobacterium tuberculosis* serine/threonine kinases PknB, PknD, PknE, and PknF phosphorylate multiple FHA domains. *Protein Sci* **14**: 1918-1921.

Guest J.R. (1992). Oxygen-regulated gene expression in *Escherichia coli*. The 1992 Marjory Stephenson Prize Lecture. *J Gen Microbiol* **138**: 2253-63.

Haydel S.E. and Clark-Curtiss J.E. (2004). Global expression analysis of two-component system regulator genes during *Mycobacterium tuberculosis* growth in human macrophages. *FEMS Microbiology Letters* **236**: 341-347.

Hefti M.H., Francoijs K.J., De Vries S.C., Dixon R. and Vervoort J. (2004). The PAS fold - A redefinition of the PAS domain based upon structural prediction. *European Journal of Biochemistry* **271**: 1198-1208.

Helmann J.D., Wu M.F., Gaballa A., Kobel P.A., Morshedi M.M., Fawcett P. and Paddon C. (2003). The global transcriptional response of *Bacillus subtilis* to peroxide stress is coordinated by three transcription factors. *J Bacteriol* **185**: 243-53.

Hidalgo E., Bollinger J.M., Jr., Bradley T.M., Walsh C.T. and Demple B. (1995). Binuclear [2Fe-2S] clusters in the *Escherichia coli* SoxR protein and role of the metal centers in transcription. *J Biol Chem* **270**: 20908-14.

Hidalgo E. and Demple B. (1994). An iron-sulfur center essential for transcriptional activation by the redox-sensing SoxR protein. *Embo J* **13**: 138-46.

Himpens S., Locht C. and Supply P. (2000). Molecular characterization of the mycobacterial SenX3-RegX3 two-component system: evidence for autoregulation. *Microbiology* **146 Pt 12**: 3091-8.

Hinds J., Mahenthiralingam E., Kempsell K., Duncan K., Stokes R., Parish T. and Stoker N. (1999). Enhanced gene replacement in mycobacteria. *Microbiology* **145**: 519-527.

Hinds J., Vass J.K., Mangan J.A., Laing K., Waddell S.J., Al-Ghusein H., Buxton R.S., Stoker N.G., Colston M.J. and Butcher P.D. (2001). Design, construction and validation of a whole genome DNA microarray for *Mycobacterium tuberculosis*. Keystone Symposium: Molecular and Cellular Aspects of Tuberculosis Research in the Post Genomic Era, Taos, New Mexico, Keystone Symposia.

Hingley-Wilson S.M., Sambandamurthy V.K. and Jacobs W.R., Jr. (2003). Survival perspectives from the world's most successful pathogen, *Mycobacterium tuberculosis*. *Nat Immunol* **4**: 949-55.

Hoch J.A. (2000). Two-component and phosphorelay signal transduction. *Curr Opin Microbiol* **3**: 165-70.

Hopewell P.C. (1992). Impact of human immunodeficiency virus infection on the epidemiology, clinical features, management, and control of tuberculosis. *Clinical Infectious Diseases* **15**: 540 - 547.

Hu Y., Movahedzadeh F., Stoker N.G. and Coates A.R. (2005). The alpha-crystallin-like *acr* gene is responsible for hypervirulence in a *Mycobacterium tuberculosis dosR* mutant. Acid Fast Club summer meeting, Birmingham, UK.

Huang Z.J., Edery I. and Rosbash M. (1993). PAS is a dimerization domain common to *Drosophila* period and several transcription factors. *Nature* **364**: 259-62.

Husson R.N., James B.E. and Young R.A. (1990). Gene replacement and expression of foreign DNA in mycobacteria. *J Bacteriol.* **172**: 519–524.

Imlay J.A. and Linn S. (1988). DNA damage and oxygen radical toxicity. *Science* **240**: 1302-9.

Iuchi S. and Lin E.C. (1988). *arcA* (*dye*), a global regulatory gene in *Escherichia coli* mediating repression of enzymes in aerobic pathways. *PNAS* **85**: 1888-92.

Iuchi S. and Lin E.C. (1992). Mutational analysis of signal transduction by ArcB, a membrane sensor protein responsible for anaerobic repression of operons involved in the central aerobic pathways in *Escherichia coli. J Bacteriol* **174**: 3972-80.

Iuchi S. and Weiner L. (1996). Cellular and molecular physiology of *Escherichia coli* in the adaptation to aerobic environments. *J Biochem (Tokyo)* **120**: 1055-63.

Jacobs W.R., Jr., Tuckman M. and Bloom B.R. (1987). Introduction of foreign DNA into mycobacteria using a shuttle phasmid. *Nature* **327**: 532-5.

Karakousis P.C., Yoshimatsu T., Lamichhane G., Woolwine S.C., Nuermberger E.L., Grosset J. and Bishai W.R. (2004). Dormancy phenotype displayed by extracellular *Mycobacterium tuberculosis* within artificial granulomas in mice. *J. Exp. Med.* **200**: 647-657.

Kendall S.L., Movahedzadeh F., Rison S.C.G., Wernisch L., Parish T., Duncan K., Betts J.C. and Stoker N.G. (2004). The *Mycobacterium tuberculosis* DosRS two-component system is induced by multiple stresses. *Tuberculosis* **84**: 247-255.

Klomsiri C., Panmanee W., Dharmsthiti S., Vattanaviboon P. and Mongkolsuk S. (2005). Novel Roles of *ohrR-ohr* in *Xanthomonas* sensing, metabolism, andphysiological adaptive response to lipid hydroperoxide. *J. Bacteriol.* **187**: 3277-3179.

Kochi A. (1991). The global tuberculosis situation and the new strategy of the World Health Organsiation. *Tubercule* **72**: 1-6.

Korner H., Sofia H.J. and Zumft W.G. (2003). Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: exploiting the metabolic spectrum by controlling alternative gene programs. *FEMS Microbiol Rev* **27**: 559-92.

Koshkin A., Nunn C.M., Djordjevic S. and Ortiz De Montellano P.R. (2003). The mechanism of *Mycobacterium tuberculosis* alkylhydroperoxidase AhpD as defined by mutagenesis, crystallography, and kinetics. *J Biol Chem* **278**: 29502-8.

Levanon S.S., San K.Y. and Bennett G.N. (2005). Effect of oxygen on the *Escherichia coli* ArcA and FNR regulation systems and metabolic responses. *Biotechnology and Bioengineering* **89**: 556-564.

Loebel R.O., Shorr E. and Richardson H.B. (1933). The influence of adverse conditions on upon the repiratory metabolism and growth of the human tubercle bacilli. *J. Bacteriol.* **26**: 139-166

Malpica R., Franco B., Rodriguez C., Kwon O. and Georgellis D. (2004). Identification of a quinone-sensitive redox switch in the ArcB sensor kinase. *PNAS* **101**: 13318-13323.

Manganelli R., Provvedi R., Rodrigue S., Beaucher J., Gaudreau L. and Smith I. (2004). sigma factors and global gene regulation in *Mycobacterium tuberculosis*. *J. Bacteriol.* **186**: 2516-2516.

Marchler-Bauer A. and Bryant S.H. (2004). CD-Search: protein domain annotations on the fly. *Nucleic Acids Res* **32**: W327-331.

Marklund B., Speert D. and Stokes R. (1995). Gene replacement through homologous recombination in *Mycobacterium intracellulare*. *J. Bacteriol.* **177**: 6100-6105.

Martin R.G. and Rosner J.L. (2004). Transcriptional and translational regulation of the *marRAB* multiple antibiotic resistance operon in *Escherichia coli*. *Mol. Microbiol.* **53**: 183-191.

Martinez-Ardudo I., Little R., Shearer N., Johnson P. and Dixon R. (2004). The NifL-NifA system: a multidomain transcriptional regulatory complex that integrates environmental signals. *J. Bacteriol.*, **186**: 601-610

Martinez-Hackert E. and Stock A.M. (1997). The DNA-binding domain of OmpR: crystal structures of a winged helix transcription factor. *Structure* **5**: 109-24.

Master S.S., Springer B., Sander P., Boettger E.C., Deretic V. and Timmins G.S. (2002). Oxidative stress response genes in *Mycobacterium tuberculosis*: role of *ahpC* in resistance to peroxynitrite and stage-specific survival in macrophages. *Microbiology* **148**: 3139-44.

Mckinney J.D., Honer Zu Bentrup K., Munoz-Elias E.J., Miczak A., Chen B., Chan W.T., Swenson D., Sacchettini J.C., Jacobs W.R., Jr. and Russell D.G. (2000). Persistence of *Mycobacterium tuberculosis* in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase. *Nature* **406**: 735-8.

Miyatake H., Mukai M., Park S.Y., Adachi S., Tamura K., Nakamura H., Nakamura K., Tsuchiya T., Iizuka T. and Shiro Y. (2000). Sensory mechanism of oxygen sensor

FixL from *Rhizobium meliloti*: crystallographic, mutagenesis and resonance Raman spectroscopic studies. *J Mol Biol* **301**: 415-31.

Mongkolsuk S., Panmanee W., Atichartpongkul S., Vattanaviboon P., Whangsuk W., Fuangthong M., Eiamphungporn W., Sukchawalit R. and Utamapongchai S. (2002). The repressor for an organic peroxide-inducible operon is uniquely regulated at multiple levels. *Mol Microbiol* **44**: 793-802.

Monot M., Honore N., Garnier T., Araoz R., Coppee J.-Y., Lacroix C., Sow S., Spencer J.S., Truman R.W., Williams D.L., Gelber R., Virmond M., Flageul B., Cho S.-N., Ji B., Paniz-Mondolfi A., Convit J., Young S., Fine P.E., Rasolofo V., Brennan P.J. and Cole S.T. (2005). On the Origin of Leprosy. *Science* **308**: 1040-1042.

Morth J.P., Gosmann S., Nowak E. and Tucker P.A. (2005). A novel two-component system found in *Mycobacterium tuberculosis*. *FEBS Lett* **579**: 4145-8.

Munoz-Elias E.J. and Mckinney J.D. (2005). *Mycobacterium tuberculosis* isocitrate lyases 1 and 2 are jointly required for *in vivo* growth and virulence. *Nature Medicine* **11**: 638-644.

Nathan C. and Shiloh M.U. (2000). Reactive oxygen and nitrogen intermediates in the relationship between mammalian hosts and microbial pathogens. *PNAS* **97**: 8841-8.

North R.J. (1995). *Mycobacterium tuberculosis* is strikingly more virulent for mice when given via the respiratory than via the intravenous route. *J Infect Dis* **172**: 1550-3.

Nunoshiba T., Derojas-Walker T., Wishnok J.S., Tannenbaum S.R. and Demple B. (1993). Activation by nitric oxide of an oxidative-stress response that defends *Escherichia coli* against activated macrophages. *PNAS* **90**: 9993-7.

Nunoshiba T., Hidalgo E., Li Z. and Demple B. (1993). Negative autoregulation by the *Escherichia coli* SoxS protein: a dampening mechanism for the *soxRS* redox stress response. *J Bacteriol* **175**: 7492-4.

Nyka W. (1974). Studies on the effect of starvation on mycobacteria. *Infect Immun* **9**: 843-50.

Nystrom T. (1998). To be or not to be: the ultimate decision of the growth-arrested bacterial cell. *FEMS Microbiology Reviews* **21**: 283 - 290.

Nystrom T. (1999). Starvation, cessation of growth and bacterial aging. *Curr Opin Microbiol* **2**: 214-219.

Nystrom T. (2002). Aging in bacteria. *Curr Opin Microbiol* **5**: 596-601.

Nystrom T., Larsson C. and Gustafsson L. (1996). Bacterial defense against aging: role of the *Escherichia coli* ArcA regulator in gene expression, readjusted energy flux and survival during stasis. *Embo J* **15**: 3219-28.

Oettinger T., Jorgensen M., Ladefoged A., Haslov K. and Andersen P (1999). Development of the *Mycobacterium bovis* BCG vaccine: review of the historical and biochemical evidence for a genealogical tree. *Tuber. Lung Dis.* 79: 243-50.

Orme I.M. and Collins F.M. (1994). Mouse model of tuberculosis. Tuberculosis: pathology, protection and control. B. B. Bloom. Washington DC, American Society of Microbiology: 115-134.

O'toole R., Smeulders M.J., Blokpoel M.C., Kay E.J., Lougheed K. and Williams H.D. (2003). A two-component regulator of universal stress protein expression and adaptation to oxygen starvation in *Mycobacterium smegmatis*. *J Bacteriol* **185**: 1543-54.

O'toole R. and Williams H.D. (2003). Universal stress proteins and *Mycobacterium tuberculosis*. *Res Microbiol* **154**: 387-92.

Pagan-Ramos E., Song J., Mcfalone M., Mudd M.H. and Deretic V. (1998). Oxidative stress response and characterization of the *oxyR-ahpC* and *furA-katG* Loci in *Mycobacterium marinum*. *J. Bacteriol.* **180**: 4856-4864.

Panmanee W., Vattanaviboon P., Eiamphungporn W., Whangsuk W., Sallabhan R. and Mongkolsuk S. (2002). OhrR, a transcription repressor that senses and responds to changes in organic peroxide levels in *Xanthomonas campestris* pv. phaseoli. *Mol Microbiol* **45**: 1647-54.

Parish T., Smith D.A., Kendall S., Casali N., Bancroft G.J. and Stoker N.G. (2003). Deletion of two-component regulatory systems increases the virulence of *Mycobacterium tuberculosis*. *Infect Immun* **71**: 1134-40.

Parish T., Smith D.A., Roberts G., Betts J. and Stoker N.G. (2003). The *senX3-regX3* two-component regulatory system of *Mycobacterium tuberculosis* is required for virulence. *Microbiology* **149**: 1423-35.

Parish T. and Stoker N.G. (2000). Use of a flexible cassette method to generate a double unmarked *Mycobacterium tuberculosis tlyA plcABC* mutant by gene replacement. *Microbiology* **146**: 1969-75.

Park H.D., Guinn K.M., Harrell M.I., Liao R., Voskuil M.I., Tompa M., Schoolnik G.K. and Sherman D.R. (2003). Rv3133c/DosR is a transcription factor that mediates the hypoxic response of *Mycobacterium tuberculosis*. *Mol Microbiol* **48**: 833-43.

Park H.J., Suquet C., Satterlee J.D. and Kang C.H. (2004). Insights into signal transduction involving PAS domain oxygen-sensing heme proteins from the X-ray crystal structure of *Escherichia coli* Dos heme domain (EcDosH). *Biochemistry* **43**: 2738-2746.

Pashley C.A., Parish T., Mcadam R.A., Duncan K. and Stoker N.G. (2003). Gene replacement in mycobacteria by using incompatible plasmids. *Appl. Environ. Microbiol.* **69**: 517-523.

Pastan I. and Perlman R. (1970). Cyclic adenosine monophosphate in bacteria. *Science* **169**: 339-44.

Peirs P., De Wit L., Braibant M., Huygen K. and Content J. (1997). A serine/threonine protein kinase from *Mycobacterium tuberculosis*. *Eur J Biochem* **244**: 604-12.

Pierce C.H., Dubos R.J. and Schaeffer W.B. (1953). Multiplication and survival of tubercle bacilli in the organs of mice. *J Exp Med* **97**: 189-205.

Pieters J. (2001). Entry and survival of pathogenic mycobacteria in macrophages. *Microbes Infect* **3**: 249-55.

Pieters J. and Gatfield J. (2002). Hijacking the host: survival of pathogenic mycobacteria inside macrophages. *Trends Microbiol* **10**: 142-6.

Potter C.A., Ward A., Laguri C., Williamson M.P., Henderson P.J. and Phillips-Jones M.K. (2002). Expression, purification and characterisation of full-length histidine protein kinase RegB from *Rhodobacter sphaeroides*. *J Mol Biol* **320**: 201-13.

Raman S., Song T., Puyang X., Bardarov S., Jacobs W.R., Jr. and Husson R.N. (2001). The alternative sigma factor SigH regulates major components of oxidative and heat stress responses in *Mycobacterium tuberculosis*. *J Bacteriol* **183**: 6119-25.

Rand L., Hinds J., Springer B., Sander P., Buxton R.S. and Davis E.O. (2003). The majority of inducible DNA repair genes in *Mycobacterium tuberculosis* are induced independently of RecA. *Mol. Microbiol.* **50**: 1031-1042.

Raynaud C., Papavinasasundaram K.G., Speight R.A., Springer B., Sander P., Bottger E.C., Colston M.J. and Draper P. (2002). The functions of OmpATb, a pore-forming protein of *Mycobacterium tuberculosis*. *Mol Microbiol* **46**: 191-201.

Reyrat J., Berthet F. and Gicquel B. (1995). The urease locus of *Mycobacterium tuberculosis* and iIts utilization for the demonstration of allelic exchange in *Mycobacterium bovis* Bacillus Calmette- Guerin. *PNAS* **92**: 8768-8772.

Rickman L. (2002). Global regulators of gene expression in *Mycobacterium tuberculosis*. Functional analysis using DNA microarrays. PhD thesis, University College London.

Rickman L., Saldanha J.W., Hunt D.M., Hoar D.N., Colston M.J., Millar J.B. and Buxton R.S. (2004). A two-component signal transduction system with a PAS domain-containing sensor is required for virulence of *Mycobacterium tuberculosis* in mice. *Biochem Biophys Res Commun* **314**: 259-67.

Rickman L., Scott C., Hunt D.M., Hutchinson T., Menendez M.C., Whalan R., Hinds J., Colston M.J., Green J. and Buxton R.S. (2005). A member of the cAMP receptor protein family of transcription regulators in *Mycobacterium tuberculosis* is

required for virulence in mice and controls transcription of the *rpfA* gene coding for a resuscitation promoting factor. *Mol. Microbiol.* **56**: 1274-1286.

Roberts D.M., Liao R.P., Wisedchaisri G., Hol W.G.J. and Sherman D.R. (2004). Two sensor kinases contribute to the hypoxic response of *Mycobacterium tuberculosis*. *J. Biol. Chem.* **279**: 23082-23087.

Roche P.W., Triccas J.A. and Winter N. (1995). BCG vaccination against tuberculosis: past disappointments and future hopes. *Trends in Microbiology* **3**: 397-401.

Rodriguez C., Kwon O. and Georgellis D. (2004). Effect of D-lactate on the physiological activity of the ArcB sensor kinase in *Escherichia coli*. *J. Bacteriol.* **186**: 2085-2090.

Rodriguez G.M., Voskuil M.I., Gold B., Schoolnik G.K. and Smith I. (2002). *ideR*, An essential gene in *Mycobacterium tuberculosis*: role of IdeR in iron-dependent gene expression, iron metabolism, and oxidative stress response. *Infect Immun* **70**: 3371-81.

Roeder W. and Somerville R.L. (1979). Cloning the *trpR* gene. *Molecular and General Genetics* **176**: 361-368.

Rosenkrands I., Weldingh K., Jacobsen S., Hansen C.V., Florio W., Gianetri I. and Andersen P. (2000). Mapping and identification of *Mycobacterium tuberculosis* proteins by two-dimensional gel electrophoresis, microsequencing and immunodetection. *Electrophoresis* **21**: 935-948.

Rutherford K., Parkhill J., Crook J., Horsnell T., Rice P., Rajandream M.A. and Barrell B. (2000). Artemis: sequence visualization and annotation. *Bioinformatics* **16**: 944-5.

Sambrook J. and Russell D. (2001). Molecular cloning: a laboratory manual. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press.

Sassetti C.M., Boyd D.H. and Rubin E.J. (2003). Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* **48**: 77-84.

Saxena P., Yadav G., Mohanty D. and Gokhale R.S. (2003). A new family of type III polyketide synthases in *Mycobacterium tuberculosis*. *J. Biol. Chem.* **278**: 44780-44790.

Seoane A. and Levy S. (1995). Characterization of MarR, the repressor of the multiple antibiotic resistance (*mar*) operon in *Escherichia coli*. *J. Bacteriol.* **177**: 3414-3419.

Shaw D.J., Rice D.W. and Guest J.R. (1983). Homology between CAP and Fnr, a regulator of anaerobic respiration in *Escherichia coli*. *J Mol Biol* **166**: 241-7.

Sherman D.R., Mdluli K., Hickey M.J., Arain T.M., Morris S.L., Barry C.E., 3rd and Stover C.K. (1996). Compensatory *ahpC* gene expression in isoniazid-resistant *Mycobacterium tuberculosis*. *Science* **272**: 1641-3.

Sherman D.R., Sabo P.J., Hickey M.J., Arain T.M., Mahairas G.G., Yuan Y., Barry C.E., 3rd and Stover C.K. (1995). Disparate responses to oxidative stress in saprophytic and pathogenic mycobacteria. *PNAS* **92**: 6625-9.

Sherman D.R., Voskuil M., Schnappinger D., Liao R., Harrell M.I. and Schoolnik G.K. (2001). Regulation of the *Mycobacterium tuberculosis* hypoxic response gene encoding alpha -crystallin. *PNAS* **98**: 7534-9.

Snapper S.B., Lugosi L., Jekkel A., Melton R.E., Kieser T., Bloom B.R. and Jacobs W.R., Jr. (1988). Lysogeny and transformation in mycobacteria: stable expression of foreign genes. *PNAS* **85**: 6987-91.

Spiro S. and Guest J.R. (1987). Regulation and over-expression of the *fnr* gene of *Escherichia coli*. *J Gen Microbiol* **133**: 3279-88.

Spiro S. and Guest J.R. (1991). Adaptive responses to oxygen limitation in *Escherichia coli*. *Trends Biochem Sci* **16**: 310-4.

Springer B., Master S., Sander P., Zahrt T., Mcfalone M., Song J., Papavinasasundaram K.G., Colston M.J., Boettger E. and Deretic V. (2001). Silencing of oxidative stress response in *Mycobacterium tuberculosis*: expression patterns of *ahpC* in virulent and avirulent strains and effect of *ahpC* inactivation. *Infect Immun* **69**: 5967-73.

Steenken W. and Gardner L.U. (1946). History of H37 strain of tubercle bacillus. *American Reviews of Tuberculosis* **54**: 62-66.

Stephenson K. and Hoch J.A. (2004). Developing inhibitors to selectively target two-component and phosphorelay signal transduction systems of pathogenic microorganisms. *Current Medicinal Chemistry* **11**: 765-773.

Stewart G.R., Wernisch L., Stabler R., Mangan J.A., Hinds J., Laing K.G., Young D.B. and Butcher P.D. (2002). Dissection of the heat-shock response in *Mycobacterium tuberculosis* using mutants and microarrays. *Microbiology* **148**: 3129-38.

Stover C.K., De La Cruz V.F., Fuerst T.R., Burlein J.E., Benson L.A., Bennett L.T., Bansal G.P., Young J.F., Lee M.H., Hatfull G.F. and Et Al. (1991). New use of BCG for recombinant vaccines. *Nature* **351**: 456-60.

Sukchawalit R., Loprasert S., Atichartpongkul S. and Mongkolsuk S. (2001). Complex regulation of the organic hydroperoxide resistance gene (*ohr*) from *Xanthomonas* involves OhrR, a novel organic peroxide-inducible negative regulator, and posttranscriptional modifications. *J Bacteriol* **183**: 4405-12.

Supply P., Magdalena J., Himpens S. and Locht C. (1997). Identification of novel intergenic repetitive units in a mycobacterial two-component system operon. *Mol Microbiol* **26**: 991-1003.

Tardat B. and Touati D. (1991). Two global regulators repress the anaerobic expression of MnSOD in *Escherichia coli*::Fur (ferric uptake regulation) and Arc (aerobic respiration control). *Mol Microbiol* **5**: 455-65.

Tardat B. and Touati D. (1993). Iron and oxygen regulation of *Escherichia coli* MnSOD expression: competition between the global regulators Fur and ArcA for binding to DNA. *Mol Microbiol* **9**: 53-63.

Taylor B.L. and Zhulin I.B. (1999). PAS domains: internal sensors of oxygen, redox potential, and light. *Microbiol Mol Biol Rev* **63**: 479-506.

Toledano M.B., Kullik I., Trinh F., Baird P.T., Schneider T.D. and Storz G. (1994). Redox-dependent shift of OxyR-DNA contacts along an extended DNA-binding site: a mechanism for differential promoter selection. *Cell* **78**: 897-909.

Triccas J.A., Parish T., Britton W.J. and Gicquel B. (1998). An inducible expression system permitting the efficient purification of a recombinant antigen from *Mycobacterium smegmatis*. *FEMS Microbiol Lett* **167**: 151-6.

Tuberculist (2005). TubercuList, Institut Pasteur. http://genolist.pasteur.fr/Tuberculist/.

Ulrich L.E., Koonin E.V. and Zhulin I.B. (2005). One-component systems dominate signal transduction in prokaryotes. *Trends Microbiol* **13**: 52-56.

Vattanaviboon P., Panmanee W. and Mongkolsuk S. (2003). Induction of peroxide and superoxide protective enzymes and physiological cross-protection against peroxide killing by a superoxide generator in *Vibrio harveyi*. *FEMS Microbiol Lett* **221**: 89-95.

Vergne I., Chua J., Singh S.B. and Deretic V. (2004). Cell biology of the *Mycobacterium tuberculosis* phagosome. *Ann. Rev. Cell Dev. Biol.* **20**: 367-394.

Voskuil M.I., Schnappinger D., Rutherford R., Liu Y. and Schoolnik G.K. (2004). Regulation of the *Mycobacterium tuberculosis* PE/PPE genes. *Tuberculosis* **84**: 256-62.

Voskuil M.I., Visconti K.C. and Schoolnik G.K. (2004). *Mycobacterium tuberculosis* gene expression during adaptation to stationary phase and low-oxygen dormancy. *Tuberculosis* **84**: 218-227.

Wayne L.G. (1977). Synchronized replication of *Mycobacterium tuberculosis*. *Infect Immun* **17**: 528-30.

Wayne L.G. and Hayes L.G. (1996). An in vitro model for sequential study of shiftdown of *Mycobacterium tuberculosis* through two stages of nonreplicating persistence. *Infect Immun* **64**: 2062-9.

Wayne L.G. and Lin K.Y. (1982). Glyoxylate metabolism and adaptation of *Mycobacterium tuberculosis* to survival under anaerobic conditions. *Infect Immun* **37**: 1042-9.

Wayne L.G. and Sohaskey C.D. (2001). Nonreplicating persistence of *Mycobacterium tuberculosis*. *Annu Rev Microbiol* **55**: 139-63.

Who (2001). Tuberculosis, WHO Regional Office for South-East Asia.

Wieles B., Nagai S., Wiker H., Harboe M. and Ottenhoff T. (1995). Identification and functional characterization of thioredoxin of *Mycobacterium tuberculosis*. *Infect. Immun.* **63**: 4946-4948.

Wu J., Dunham W.R. and Weiss B. (1995). Overproduction and physical characterization of SoxR, a [2Fe-2S] protein that governs an oxidative response regulon in *Escherichia coli*. *J Biol Chem* **270**: 10323-7.

Wu S.P., Howard S.T., Lakey D.L., Kipnis A., Samten B., Safi H., Gruppo V., Wizel B., Shams H., Basaraba R.J., Orme I.M. and Barnes P.F. (2004). The principal sigma factor SigA mediates enhanced growth of *Mycobacterium tuberculosis in vivo*. *Molecular Microbiology* **51**: 1551-1562.

Yuan Y., Crane D. and Barry C., 3rd (1996). Stationary phase-associated protein expression in *Mycobacterium tuberculosis*: function of the mycobacterial alpha-crystallin homolog. *J. Bacteriol.* **178**: 4484-4492.

Zahrt T.C., Wozniak C., Jones D. and Trevett A. (2003). Functional Analysis of the *Mycobacterium tuberculosis* MprAB two-component signal transduction system. *Infect. Immun.* **71**: 6962-6970.

Zhang Y. (2005). The magic bullets and tuberculosis drug targets. *Annual Review of Pharmacology and Toxicology* **45**: 529-564.

Zhang Y., Dhandayuthapani S. and Deretic V. (1996). Molecular basis for the exquisite sensitivity of *Mycobacterium tuberculosis* to isoniazid. *PNAS* **93**: 13212-13216.