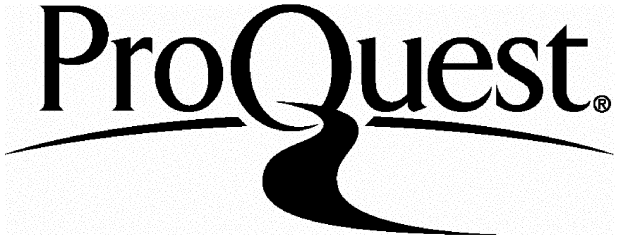
The Structure, Function and Regulation of Mycobacterial Porin-Encoding Genes

Richard Alan Speight

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

October 2000



Index of Contents

Index of Figures

NOTE ADDED AFTER EXAMINATION
A revised and additional list of figure legends can now be found as Supplementary Appendix 1.
A list of plasmids used can also be found as Supplementary Appendix 2.
Please refer to these extra appendices for additional information.

Index of Tables

Abbreviations

APS	Ammonium persulphate
BCA	Bicinchoninic acid
BCG	Bacille Calmette-Guérin
BLAST	Basic local alignment search tool
bp	Base pairs
BSA	Bovine serum albumin
cDNA	complementary single-stranded DNA
CIP	Calf intestinal phosphatase
DEPC	Diethyl pyrocarbonate
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
dNTP	2'-Deoxynucleoside 5'-triphosphate
DTT	Dithiothreitol
dUTP	Deoxyuracil triphosphate
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme linked immunosorbent assay
IPTG	Isopropyl-β-D-thiogalactoside
kb	Kilobase pairs
NIMR	National Institute for Medical Research
OD_{600}	Optical density at a wavelength of 600 nanometers

ONPG	o-Nitrophenyl β-D-galactopyranoside
PBS	Phosphate buffered saline
PCR	Polymerase Chain Reaction
PMSF	Phenylmethylsulphonylfluoride
PVDF	Polyvinylidene fluoride
r.p.m.	Revolutions per Minute
RNA	Ribonucleic acid
RT-PCR	Reverse transcription polymerase chain reaction
SDS	Sodium dodecyl sulphate
SRS	Sequence retrieval service
SSC	Saline-sodium citrate buffer
TAE	Tris(hydroxymethyl)aminomethane-acetic acid-Ethylenediamine-tetraacetic acid buffer
TBE	Tris(hydroxymethyl)aminomethane-boric acid-Ethylenediamine-tetraacetic acid buffer
TBS	Tris(hydroxymethyl)aminomethane-buffered saline
TE	Tris(hydroxymethyl)aminomethane-ethylenediaminetetraacetic acid buffer
TEMED	N, N, N', N' tetramethylene-ethylene-diamine
Tris	Tris(hydroxymethyl)aminomethane
TTBS	Tris(hydroxymethyl)aminomethane-buffered saline with 0.05 % Tween20
U	Unit(s)
X-gal	4-bromo-3-chloro-2-indolyl-β-galactoside

Abstract

Two years ago Senaratne *et al.*, (1998) published research into a gene, *ompATb*, which encodes an outer membrane protein in the genus *Mycobacterium.* It was shown that the protein, OmpATb, possesses a pore-forming ability in liposomes and can be incorporated into lipid bilayers in which it exhibits voltage-gated pore characteristics.

This study examines the distribution of this gene, and the protein product OmpATb, amongst different mycobacterial species using a PCR based method, and demonstrates immunoblotting using an antibody previously raised against the protein.

Regulation of the porin genes in *Escherichia coli* by a two-component sensor-regulator pair is discussed and parallels drawn to a putative method of regulation in the mycobacteria. Regulation of the mycobacterial porin is demonstrated by real-time RT-PCR of *Mycobacterium bovis* BCG RNA and the addition of osmotic stress to the bacterial culture is shown to have a dramatic effect on the levels of *ompATb* transcript.

Two genes, Rv0902c and Rv0903c, are proposed to have a mode of action in controlling the expression of the porin gene in *Mycobacterium tuberculosis.* These were identified by homology to their *Escherichia coli* counterparts by searching the published genome of *Mycobacterium tuberculosis* (Cole *et al.*, 1998). The regulatory component, Rv0903c, is cloned into an expression vector and the recombinant Rv0903 protein overexpressed in *Escherichia coli.* This protein product is purified using immobilised metal affinity chromatography and is used in gel retardation assays to show that two regions around the *ompATb* gene bind the protein. Phosphorylation of the protein *in vitro* is also shown to enhance the binding affinity.

The porin-encoding gene is expressed in a non-native organism, *Mycobacterium smegmatis*, both from its native promoter and from the mycobacterial *hsp60* promoter, the effect of overexpression of this gene and the protein product is investigated. The protein is also cloned into *Mycobacterium smegmatis* in the presence and the absence of the putative regulatory machinery.

An attempt to characterise the promoter region of the *ompATb* gene is made using the construction of *lacZ* reporter vectors. These are used for assays of β-galactosidase activity

in *Mycobacterium smegmatis*. It is shown that the *ompATb* promoter functions in this organism and that the length of upstream sequence included affects the promoter activity.

A suicide vector for knockout of the *ompATb* gene is constructed and transformed into *Mycobacterium tuberculosis* H37Rv in an attempt to disrupt the function of the gene and a Δ*ompATb* strain of *Mycobacterium tuberculosis* 1424 is characterised by Southern blot and PCR based methods. This strain is investigated by DNA and cDNA microarray analysis and by growth in media providing a variety of environmental stresses.

Gene disruption by homologous recombination is also attempted on the regulatory gene Rv0903c, employing a different technique for suicide vector construction that uses 3 counterselectable markers. This technique has been previously applied with some success in the mycobacteria (Parish and Stoker, 2000; Parish and Stoker, 2000). A number of colonies exhibiting the correct phenotype for gene knockouts are examined using a PCR based method.

NOTE ADDED AFTER EXAMINATION
A revised and additional list of figure legends can now be found as Supplementary Appendix 1.
A list of plasmids used can also be found as Supplementary Appendix 2.
Please refer to these extra appendices for additional information.

1. Introduction

1.1 The Mycobacteria

In 1882 Robert Koch isolated and characterised the organism that causes the disease tuberculosis, *Mycobacterium tuberculosis.* Since that time much research has been devoted to understanding this causative agent and how it may be controlled.

Significant progress has been made since the days when 'consumption' caused 20-25 % of all deaths in European cities (Daniel *et al.*, 1994). However to this day tuberculosis remains an extremely significant cause of human mortality and morbidity world-wide, with the latest World Health Organisation report estimating that 3 % of deaths globally per annum are due to tuberculosis (WHO, 2000), a total of some 1.7 million people. This makes tuberculosis the leading cause of microbial mortality. It is also estimated that around 1 in 3 of the world's population are carriers of the disease in a latent form (Preface to Bloom, (1994)) and of these around 10 % will go on to develop clinical symptoms later in life (Bloom and Murray, 1992).

Furthermore there is a continued threat from tuberculosis as the rise in immunocompromised individuals continues due to the A.I.D.S. pandemic and increasing use of chemotherapy and radiotherapy. Loss of immune function increases susceptibility to the disease and is leading to a global re-emergence of tuberculosis, in 1993 the W.H.O announced that tuberculosis was a 'global emergency' (Colston, 1997). Also of concern is the rise in so called 'multidrug resistant' forms of *Mycobacterium tuberculosis* which are able to withstand the (85 % effective) combination chemotherapy of isoniazid, rifampicin and other first line antituberculous drugs (Petrini and Hoffner, 1999). Such resistant forms arise as a result of patients not continuing therapy for the required term or by providing a selection pressure towards drug resistance due to the amount of antibiotics already in use. This is often the case with H.I.V infected patients who are surrounded by other tuberculosis sufferers and by those at risk of developing tuberculosis infection (Morse, 1994).

The term of therapy for tuberculosis is typically in the region of 8 months, so is not surprising that there is a high level of patients who do not complete the therapy. This length of treatment is necessary due to the extreme resistance of mycobacteria to

antibiotics, many of which are unable to easily permeate the cell and can be broken down by mycobacterial enzymes such as the ß-lactamases. Antibiotics which are effective, such as the lipid soluble rifampicin, are often ineffective against a small sub-population of 'persistent' mycobacteria which seem to have entered a dormant state. In this state they are not actively growing and hence are not susceptible to the actions of the drug. It is the elimination of these persistent mycobacteria which signals a successful course of therapy.

Another member of the genus *Mycobacterium* that is responsible for considerable human suffering is *Mycobacterium leprae*, the causative agent of leprosy. This is largely a disease that has been eradicated in the developed world but remains a significant threat in the 3rd world.

It is not just humans that suffer from mycobacterioses; in economic terms disease of cattle is very important. A steady increase in bovine tuberculosis in the U.K. was the focus of the Krebs report, which in 1997 advised the monitored culling of badgers, the animals which act as hosts for the causative agent *Mycobacterium bovis* (Krebs, 1999). It is thought that *Mycobacterium bovis* is the ancestor of *Mycobacterium tuberculosis* and that at some time during history passage from domesticated animals to humans occurred as a result of living and working in close proximity to cattle (Daniel *et al.*, 1994). *Mycobacterium bovis* is not solely a pathogen of cattle and badgers; it is also able to infect humans.

1.2 Taxonomy of the mycobacteria

Mycobacteriacae are members of the actinomycete clade, and form a distinct phyletic line alongside the Corynebacteriacae (Corynebacterium and Dietzia) and Nocardiacae (Nocardia, Rhodococcus, Tsukamurella and Gordona). They can be most readily distinguished from these in terms of cellular composition by the length of mycolic acid in the cell wall (Section 1.5.2), since corynebacteria have comparatively short mycolic acids (22-38 Carbon atoms) whereas Mycobacteriacae and Nocardiacae have long mycolic acids (34-90C) (Goodfellow and Magee, 1999).

Historically the identification and classification of mycobacteria was carried out by analysing individual characteristics, such as growth on media containing a variety of selective added factors (for example sodium chloride, different antibiotics) or growth at particular temperatures. It could then be determined whether the organism under scrutiny

had characteristics that were particular to the mycobacteria. A more precise means was developed in which the degree of relatedness was deduced mathematically from shared characteristics (Tsukamura, 1981). Such characteristics were frequently antigenicity, growth at different temperatures, resistance to antimicrobials and other criteria which gave a distinct 'yes' or 'no' answer.

Tsukamura concluded that more than 40 separate criteria needed to be scored and the 'judgement should be made without preconception'. Unsurprisingly there was a bias towards likening every mycobacterial species to *Mycobacterium tuberculosis* and thus the term 'non-tuberculous (or atypical) mycobacteria' is often used to describe species other than *Mycobacterium bovis* or *Mycobacterium tuberculosis.*

Since the early 1980s, due to advances in molecular biology, it has become easier to classify the mycobacteria. It is now commonplace to sequence the gene encoding the 16S ribosomal RNA subunit and to base interrelationships between species on sequence similarity of this gene. Fundamental assumptions have to be made in this approach, namely that there has been no lateral gene transfer and that extent of evolution, or dissimilarity, from a common ancestor is reflected in the difference in their genomes (Goodfellow and Magee, 1999). Such assumptions are logical and reasonable as the 16S rRNA gene is present and essential in all mycobacterial species, thus there is no demand for it to be laterally transferred as a species without the gene would be non-viable. Analysis of 16S rRNA relatedness also confirms taxonomic classification based on the more traditional methods

Broadly speaking members of the genus *Mycobacterium* can be divided into two groups, the 'slow-growers' and the 'fast-growers'. Slow growing mycobacteria have a generation time of around 24 - 48 hours, whereas the fast growing mycobacteria take around 3 hours. By comparison the generation time of *Escherichia coli* in favourable laboratory conditions is accepted to be around 20 minutes.

Goodfellow and Magee (1999) group the mycobacteria, of which 71 species have been described, into 5 distinct subsets based on fast/slow growth and whether the bacterium is a pathogen or a non-pathogen. Their final (5^{th}) group of 'non-cultivable mycobacteria' contains as its sole member *Mycobacterium leprae* as this species has never been cultured outside of a host, often the armadillo for laboratory purposes.

The slow growing division of the genus contains (although not exclusively) those mycobacteria which are responsible for disease, including the aforementioned *Mycobacterium tuberculosis* and *Mycobacterium bovis,* as well as *Mycobacterium bovis* Bacille Calmette-Guérin (BCG) which is widely used as a vaccine strain in the western world. The fast growers are generally not responsible for disease; representatives include the laboratory model organism *Mycobacterium smegmatis*, *Mycobacterium phlei* and *Mycobacterium vaccae.* One exception is the fast growing species *Mycobacterium fortuitum*, which is able to act as an opportunistic pathogen of humans. Others of these fast growing mycobacteria may be opportunistic pathogens in an immunocompromised host.

Quite why there is such a difference in the growth rate is still a matter for debate; there is a suggestion that the slow growing mycobacteria devote a great deal of metabolic energy into the synthesis of their complex cell envelope at the expense of systems which are more concerned with cell division, or that key respiratory pathways are in some way compromised. It may also be that the mycobacterial cell envelope (Section 1.5) forms such an impermeable barrier to the influx of essential nutrients that it is rate-limiting. All of these explanations seem flawed, as firstly the morphology of the cell wall is similar amongst the fast and slow growers and secondly the *Mycobacterium tuberculosis* genome is apparently complete, even so far as having genes and regions that are apparently redundant

Another distinction occurs in the number of genes encoding ribosomal RNA (rRNA); the slow growers have only one operon encoding rRNA, the *rrnA* operon, whereas fast growers have two, *rrnA* and *rrnB*. It was suggested that this marks a point where the slow growing mycobacteria diverged from their fast growing ancestors by losing a coding region of the *rrnB* operon (Ji *et al.*, 1994).

There is still considerable debate on whether certain mycobacteria are distinct species or whether they should be described as strains of a broader 'complex'; for example it is argued that *Mycobacterium tuberculosis, Mycobacterium microti, Mycobacterium bovis and Mycobacterium africanum* should all come under the umbrella of '*Mycobacterium tuberculosis* complex' based on their 16S rRNA and mycolic acid similarity.

Thus the taxonomy of the mycobacteria is still unclear, it is obvious that *Mycobacterium* is a distinct genus with separate species as its members, the relationships between these members are dependent upon the means used to assess them. Mycobacterial taxonomy is dependent on which criteria are weighted more heavily (of more importance) in the scoring of similarity. There is however no doubt that the mycobacteria are actinomycetes, and that they can be distinguished by key features such as acid-alcohol fastness (Section 1.3), the presence of mycolic acids containing 60-90 carbons (Section 1.5.2), and a Guanidine + Cytosine (G+C) ratio of 61-71 % (Goodfellow and Magee, 1999).

1.3 The Staining Properties of Mycobacteria

The mycobacteria are usually defined as 'acid fast', that is to say that they resist decolourisation by acid-alcohol after staining with carbol fuchsin. They appear as red or pink rod-shaped bacilli after such a procedure, whereas other bacteria are readily decolourised and appear either very faintly red or the colour of the counterstain, usually methylene blue.

Acid fast staining in itself is not a completely reliable marker for the identification of mycobacteria, as often it is dependent upon the stage of the bacterial growth cycle that they are in. That is to say, in many mycobacterial species acid fast staining is a phenomenon that tends to appear only in the later stages of growth, however at some stage in the growth cycle of all mycobacteria it is possible to stain them in an acid-fast manner. *Mycobacterium tuberculosis* always stains in an acid-fast manner, although the degree of acid fastness may vary in very young cultures.

A commonly applied bacterial diagnostic stain is that of the Gram stain (Neidhart *et al.*, 1990). Whilst mycobacteria are evolutionarily and biologically considered to be Gram positive, they do not stain well using crystal violet (the basis of Gram's stain). This is most likely due to the impermeable nature of the mycobacterial envelope, which is well protected by a hydrophobic lipid layer (it is the peptidoglycan that is responsible for binding the stain). The cell envelope is discussed later in (Section 1.5).

1.4 Host response to tuberculosis

Some aspects of the immune reaction to tuberculosis infection are relevant to this thesis, in particular the series of events occuring after infection by *Mycobacterium tuberculosis*.

These demonstrate the properties of a bacterial cell that is able to withstand considerable attack by the immune system of the host.

Upon infection, usually by the aerosol route, the tubercle bacillus will enter the lung and there be quickly ingested by host macrophages, the major defence of the innate immune system. The tubercle bacillus is able to survive inside the host macrophage and is there able to actively divide and grow. In the majority of cases the immune response is effective, and the clinical symptoms of tuberculosis do not develop. However, that is not to say that the bacteria are eradicated from the lung, it is here that *Mycobacterium tuberculosis* is unusual since it can persist in an essentially dormant state, often for the lifetime of the infected individual. The host response to mycobacterial infection is thought to consist of both an innate component and an acquired component (cell mediated immunity). This later immunity, brought about as a result of stimulation of cytokine production and T-cell response by the antigen presenting cells, is responsible for the survival of infected individuals although not always for the eradication of the bacteria.

Mycobacteria are remarkable in their ability to survive inside the hostile environment of the host macrophage. These cells have a range of antimicrobial responses including acidified vacuoles, hydrolases, bactericidal peptides and the production of reactive oxygen and nitrogen intermediates (ROIs and RNIs). Macrophages also function as professional antigen presenting cells (APCs) and stimulate the antimicrobial activities of the T (Thymus derived) and B (Bone marrow derived) lymphocytes (Russell *et al.*, 1997).

The exact nature of the resistance by the mycobacteria to all these conditions is not known, but is thought to be a combination of the impermeability of the mycobacterial cell and a means of preventing the phagosome-lysosome fusion in macrophages. Prevention of phagolysosome formation ensures that although the mycobacteria are phagocytosed they do not come into contact with the acidic vacuoles and are thus never attacked by antibacterial products in the lysosomes. (Goren *et al.*, 1976; Russell *et al.*, 1997). Most bacteria and other parasites ingested by macrophages are digested in phagolysosomes in the traditional manner. The MHC (Major Histocompatibility Complex) molecules then present antigens on the surface of the macrophage.

It is currently believed that the TH1 response is the dominant form of protective immune resistance to mycobacteria, resulting from the macrophage (or other APC) mediated

presentation of antigens in a MHC class II complex. Thus cytokines such as Interleukin 2 (IL-2) and Interferon gamma (IFNγ) are readily detectable after mycobacterial infection. Other cytokines such as Tumour Necrosis Factor (TNF), IL-8 and IL-6 are also thought to be important in the host response (Fine, 1994).

1.5 The Mycobacterial Cell Envelope

It has already been mentioned that the mycobacteria are unusual in their extreme impermeability to many antibiotics and chemotherapeutic agents, and that the bacteria do not stain Gram positive despite being biologically of that group. The reason for this is the nature of the envelope surrounding the bacterial cell.

Eubacteria can be divided into Gram negative and Gram positive subtypes, each sharing a number of similarities in the structure of the cell membrane and wall. Mycobacteria, though biologically in the Gram positive group, have a distinctive cell envelope that shows some resemblance to that of the Gram negative bacteria.

The most unusual feature of the mycobacterial envelope is the amount of lipid present, mainly as mycolic acids. It is estimated that 60 % of the weight of the cell wall is lipid based (Liu *et al.*, 1999), a feature peculiar to the actinomycetales. These lipids confer upon the cell an extremely hydrophobic nature, and may be the basis for the clumping observed when mycobacteria are grown in laboratory culture in the absence of high levels of detergents. They are also responsible for the commonly described 'waxy coat' (Draper, 1998).

The mycobacterial cell envelope is of particular importance as it is the interface between the host and the pathogen in an infection, with the other potentially antigenic components of the cell being well protected by this barrier (Daffe and Draper, 1998). It is worthy of note that 1 M sodium hydroxide is commonly used in the isolation of *Mycobacterium tuberculosis* from clinical samples as this is lethal to other bacteria. *Mycobacterium tuberculosis* is able to survive treatment due to the resistance afforded by the hydrophobic barrier.

The definitions of the mycobacterial cell envelope used here need to be clearly stated, as over the years controversy has existed over the exact terminology used. Here I will

discuss the mycobacterial envelope in terms of 3 gross structures: the plasma membrane, the cell wall (or cell wall skeleton, CWS) and the capsular layer.

1.5.1 The Plasma Membrane

The plasma membrane is the region of the envelope that shares most homology with other bacteria; it is a typical lipid bilayer and is physiologically active, having important systems such as electron transport integral to it. It is suggested that the function of the 'tougher' outer layers of the cell wall and capsule is to protect this delicate inner membrane (Draper, 1998).

The membrane is perhaps the simplest of the envelope components, consisting mainly of phopholipids arranged in a polar bilayer. Phosphatidylinositol-mannoside (PIM) lipoarabinomannan (LAM) and lipomannan (LM) are also assumed components of the plasma membrane (Daffe and Draper, 1998). The latter two phosphorylated lipopolysaccharides have an uncertain position in the envelope and indeed some models show LAM spanning the entire envelope (Brennan and Draper, 1994). Evidence for this comes from the fact that wall sections prepared by freeze substitution or careful chemical fixation show a plasma membrane which is asymmetrical (the outer leaflet being thicker, representing associated carbohydrates or PIMs) It is also possible that these are merely transient substances, on their way to the capsular or cell wall layers.

Apart from the proteins responsible for electron-transfer activities such as cytochromes, succinate dehydrogenase and NADH oxidase there have been few associated proteins identified in the plasma membrane (Daffe and Draper, 1998).

Between the plasma membrane and the cell wall lies a hypothetical periplasmic space, experimental evidence for this does not exist, its presence is based largely on an expectation from studying other Gram negative bacteria (Daffe and Draper, 1998; Daffe and Etienne, 1999; Pelicic *et al.*, 1996)

1.5.2 The Cell Wall

This is perhaps the best-understood layer of the mycobacterial envelope in terms of its composition, although the precise arrangement of the substances involved has not been determined with any certainty.

The cell wall consists of peptidoglycan, which is covalently linked to arabinogalactan; this in turn is linked to an outer layer of mycolic acids to form the Cell Wall Skeleton (CWS). The wall also contains other wall-associated lipids and peptides that are not covalently linked to the CWS.

Minnikin (1982) proposed a structure of the cell wall in which there was an intercalation of the wall associated lipids with the mycolates (mycolic acid residues). This assumed that the mycolic acids were arranged as a monolayer with the alkyl chains running perpendicular to the bacterial surface and parallel to each other. In this model the polar 'head' groups of the associated lipids face outwards from the cell and presents a hydrophilic surface. The interactions between the mycolic acids and the associated lipids are weak hydrophobic bonds and are thought to be responsible for the plane of weakness shown in freeze fracture electron micrographs (Minnikin, 1982).

Further support for this model came from McNeil and Brennan who in 1991 updated the model to include previously unknown substances (McNeil and Brennan, 1991).

Rastogi (1991) proposed an alternate model in which the associated lipids form a distinct and separate monolayer. Either of these theories could be correct, as current techniques are unable to distinguish between them. The mycobacterial cell wall is notoriously difficult to extract in a pure form and electron microscopy techniques are prone to artefacts. Draper (1998) argues that the distinction is rather trivial so long as it is agreed that the mycolate is present in a monolayer. Good evidence in support of the mycolate monolayer comes from both x-ray diffraction studies (Nikaido *et al.*, 1993) and differential scanning calorimetry, which measures a phase change, or 'melting point' of the mycobacterial cell wall extract (Draper, 1998, and references therein).

The peptidoglycan in the mycobacterial cell wall is similar to that found in other eubacteria. It consists of oligosaccharides (formed from *N*-acetylglucosamine (Glc-NAc) and *N*-glycolylmuramic acid) cross linked by short peptides (Draper, 1998).

The arabinogalactan is a heteropolysaccharide formed from D-arabinofuranosyl residues and D-galactofuranosyl residues and is the target of the anti-tubercular drug ethambutol, which prevents polymerisation of the two sugar subunits.

Mycobacterial mycolic acids, the mycolates, are long-chain α-branched β-hydroxy fatty acids which occur in all mycobacterial species (Daffe and Draper, 1998). These mycobacterial mycolates may contain up to 90 carbon residues, which is a distinguishing feature of the genus, and are attached to the terminal arabinose units of the arabinogalactan by ester bonds.

1.5.3 The Capsular layer

The existence of a capsular layer has been the subject of some dispute because the historical methods of preparation of cells for electron microscopy served to collapse or remove it. However advances in preparation techniques have led to the demonstration of a thick capsular layer which is external to the cell wall and the plasma membrane. Since the capsule is not covalently bound to the cell wall it is strictly a pseudo-capsule (Daffe and Etienne, 1999), this distinction is probably not important functionally.

The capsular layer is rather fragile and difficult to observe in cells that have been grown in a shaking culture, or have been subject to mechanical or detergent based extraction (commonly shaking with glass beads or the use of Tween 80 (0.05 %) in the growth media) (Daffe and Etienne, 1999) and (P. Draper, personal communication). Nevertheless bacteria prepared from *in-vivo* cultures or grown statically as surface pellicles exhibit a thick interface between the cell and the surrounding region, and growth as a static culture is arguably more similar to the situation encountered in the human lung or elsewhere *in vivo*.

The capsule is an important part of the mycobacterial cell envelope, as it is this that forms the interface between the host and the pathogen. It is therefore not surprising to find that the capsule contains many substances implicated in mycobacterial pathogenicity, such as those involved with invasion and adhesion. In addition to this compounds such as catalase/peroxidases and superoxide dismutases have been demonstrated in the capsule, these are responsible for the neutralisation of reactive oxygen intermediates produced by the macrophage. It is also not surprising to find that some components of the capsular layer are serologically active, notably the arabinomannan which is speculated to have an immunosuppressive effect (Daffe and Draper, 1998).

In contrast to the cell wall the capsule is not composed primarily of lipid, with this comprising only 2-5 % of the total capsular material (Daffe and Etienne, 1999). Instead it is composed primarily of polysaccharides and proteins, the ratio of which depends on the species, for example in *Mycobacterium tuberculosis* H37Rv the polysaccharide:protein ratio is roughly 2:3, whilst in *Mycobacterium smegmatis* protein is more prevalent (a 1:4 ratio). (Personal communication, M. Daffé, Institut Pasteur, France and Ortalo-Magne *et al.*, 1995).

Data on the lipid composition of the capsule is sparse, most work has been performed on *Mycobacterium tuberculosis Canetti* strain, an organism which has a significantly different capsular composition to H37Rv (a 9:1 polysaccharide:protein ratio makes it easily distinguishable from H37Rv). Lipids present in the capsule of *Mycobacterium tuberculosis* Canetti include phenolic glycolipids (not present in H37Rv, (Papa *et al.*, 1992)), lipooligosaccharides, dimycolyl trehalose ('cord-factor'), phosphatidyl ethanolamine and 2,3-diacyl trehaloses (Ortalo-Magne *et al.*, 1996). The lipid composition of the capsule is species specific and can even vary amongst different isolates of the same species (Ortalo-Magne *et al.*, 1996).

The capsular polysaccharides presented at the surface are mainly glucan, arabinomannan and mannan, some contend that LAM is also present (Chatterjee *et al.*, 1992) although the detection of this is uncertain since there is immunogenic cross-reactivity with the arabinomannan (AM) (P Draper, personal communication).

Determination of proteins present in the mycobacterial capsule is difficult; it is very hard to distinguish proteins that are genuinely resident in the capsule from those which are being secreted, by the bacterial cell. Similarly culture filtrates from mycobacterial broth cultures are seen to contain many proteins which are thought to be components of the capsular layer and have been shed due to mechanical force. A classical example of this is the antigen 85 complex; use of these 30-31kDa proteins as vaccines in guinea pigs and as naked DNA has proven successful in providing protection from tuberculosis (Daffe and Draper, 1998). A schematic diagram of the mycobacterial envelope is presented in Figure 1.

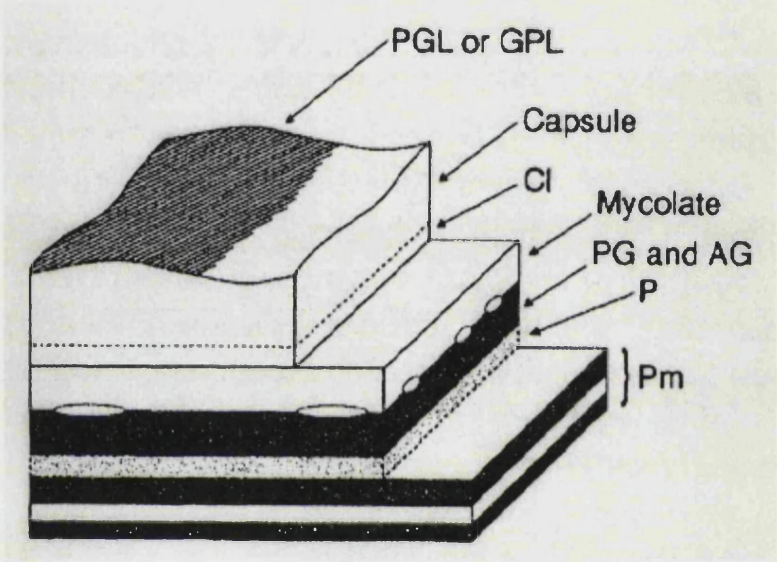


Figure 1 Schematic diagram of the mycobacterial cell envelope

PGL, phenolic glycolipid; GPL, glycopeptidolipid (both species-dependent); C1, lipid-rich lower layer of the capsule; PG, peptidoglycan; AG, arabingalactan; P, proposed periplasmic space; PM, plasma membrane.

Diagram taken from (Daffe and Draper, 1998)

1.6 Publication of the Mycobacterial genome

During the course of this project a great leap forward in the research of tuberculosis was made with the publication of the contiguous genome of *Mycobacterium tuberculosis* H37Rv (Cole *et al.*, 1998). It was shown that the genome contained nearly 4.5 million base pairs coding for 4000 separate genes, many of which were dedicated to the biogenesis of lipids and lipolytic enzymes. This is perhaps not entirely unexpected given the complex lipid envelope surrounding this bacterium.

This publication and the availability of the database containing the information therein on the World Wide Web have made it possible to search for genetic differences and similarities between other organisms with ease. The identification of genes *in silico* in this bacterium is now routine lab practice. Currently efforts are under way to sequence the genome of *Mycobacterium bovis* and that of *Mycobacterium leprae* is complete.

1.7 Bacterial Permeability

This section will discuss the problems associated with permeability of bacterial cells, considering in turn the Gram positive bacteria, the Gram negative and the mycobacteria.

1.7.1 Permeability in Gram positive bacteria

The fundamental permeability of the bacterial cell wall is governed by the lipid bilayer, the cytoplasmic membrane in Gram positive bacteria. However bacteria are unable to alter the permeability of the cell using alteration of this layer alone as its permeability is correlated with its fluidity (Nikaido, 1994); an increase in membrane fluidity brought about by increasing the amount of unsaturated hydrocarbons is accompanied by an increase in permeability. A point is reached at which the membrane would become too fluid; interfering with the proper function of membrane proteins. Such a situation is detrimental to the cell so there is a compromise position at which both fluidity and permeability are acceptable.

Whilst the Gram positive bacteria are surrounded by a thick layer of peptidoglycan outside of the cytoplasmic membrane (figure 2), this can only impede the passage of many large molecules and affords little protection against antibiotics which are able to

diffuse across it; the peptidoglycan meshwork is too coarse to exclude small molecules (Nikaido, 1994).

It seems therefore that Gram positive bacteria have little protection against molecules that are able to traverse the cytoplasmic membrane. This observation is borne out by the fact that 95 % of newly discovered antibiotics are active against Gram positive bacteria which do not have the protective outer membrane seen in Gram negative bacteria (Nikaido, 1993). Resistance to antibiotics in these organisms is often the result of mechanisms either to detoxify antimicrobials once they have entered the cell, of mutations in the genome leading to the targets of the antibiotics, or the presence of active pumps which remove the antibiotic more quickly than the any antimicrobial effect can take place.

1.7.2 Permeability in Gram negative bacteria

Gram negative bacteria are unique in their possession of a further membrane outside of the peptidoglycan layer and the cytoplasmic membrane; this is termed the *outer membrane*. The major function of this outer membrane is to act as a permeability barrier to compounds which may otherwise have access to the bacterial cell and which may be toxic (Nikaido, 1994). Thus the Gram negative bacteria are, in comparison to the majority of the Gram positives, many times more resistant to compounds or solutes that are able to traverse the outer layers of the latter group.

The exterior leaflet of the outer membrane is comprised of lipopolysaccharide (LPS) which contains saturated fatty acid chains. This is thought to lead to an arrangement that is far less fluid than the cytoplasmic membrane, which contains unsaturated fatty acids. Each LPS 'head' group also has a 'tail' of 6 or 7 covalently linked fatty acid chains serving to further decrease the fluidity (Nikaido, 1994).

The downside of this is that there obviously has to be a means of entry to the cell for molecules which are necessary for the survival of the cell, and a means of exit for waste products which may become harmful. These bacteria have adopted a number of ways to import nutrients into the cell and to exclude waste products, which will be discussed shortly.

Figure 2 shows a comparison of Gram positive, Gram negative and mycobacterial cell envelopes.

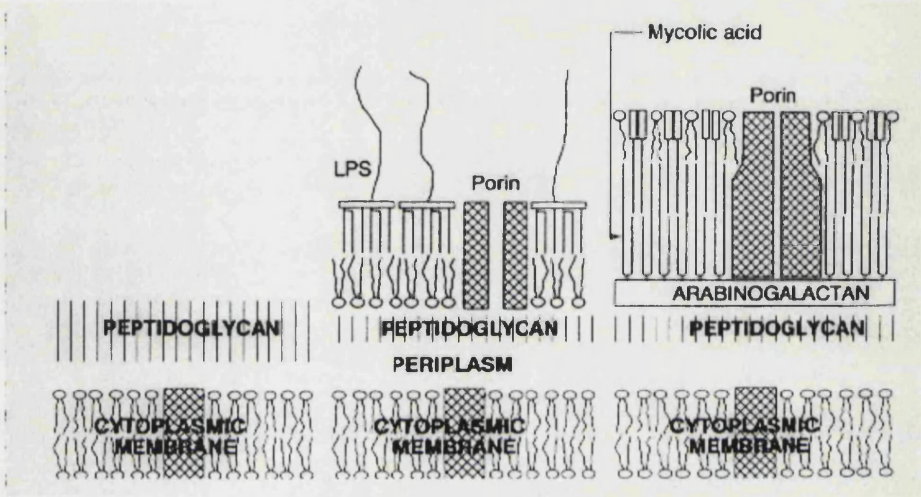


Figure 2 A comparison of bacterial envelopes

On the left is a Gram positive envelope showing a peptidoglycan layer and a typical lipid bilayer including associated proteins. In the middle a Gram negative envelope showing the additional outer barrier (LPS, lipopolysaccharide), the lesser amount of peptidoglycan and the periplasmic space. On the right a mycobacterial envelope showing extensive mycolic acids, associated lipids and the arabinogalactan layer.

Diagram taken from (Nikaido, 1994)

1.7.3 Permeability in the Mycobacteria

As mentioned previously the mycobacteria are unlike the chemically related Gram positive bacteria in that they have a layer of extreme impermeability in the form of the hydrophobic cell wall and its mycolic acids and other associated lipids. One result of this is that the mycobacteria have "a diffusion barrier 100-1000 times less permeable to hydrophilic molecules than *Escherichia coli*" (Kartmann *et al.*, 1999). Naturally this leads to a significant problem when considering drugs which may be used as chemotherapeutic agents, as no matter what the target of the drug inside the cell, it will be worthless if it cannot get there initially. It is also important to note that the capsular layer of the cell envelope is important in the exclusion of larger macromolecules, and that this layer is apparently more abundant in the slow growing members of the genus. Whether this is significant in terms of the pathogenicity of these slower growing bacteria is not fully understood, however it most certainly makes them more resistant to the macromolecular anti-microbials which would be otherwise able to gain entry to the cell.

The permeability of the mycobacterial envelope has been measured using the diffusion rates of cephalosporins (Jarlier and Nikaido, 1990). Using this method a permeability coefficient was obtained, reflecting the relative rates of permeability of several solutes. The values obtained in the species tested (*Mycobacterium chelonae*) were 1 order of magnitude lower than those for *Pseudomonas aeruginosa* and 3 orders of magnitude lower than those for *Escherichia coli*. Since the levels of β-lactamase are not unusually high in the mycobacteria this permeability problem explains the lack of efficacy of the cephalosporins. That is to say that a low level of permeability compensates for a level of β-lactamases that would otherwise be insufficient in the more permeable Gram positive bacteria. These more permeable bacteria must posses a higher level of β-lactamases to counteract the increased influx of antibiotic. Jarlier and Nikaido (1990) conclude that the hydrophobicity of the solute and the temperature have little effect on the permeability, implying that there is a hydrophilic pathway through which these compounds enter the cell. This is important as shall be demonstrated later (Section 1.10)

The permeability barrier created by the mycobacterial cell wall is without doubt the single most distinctive feature of this genus, and is the most important in terms of the treatment of mycobacterial infection. Since the permeability barrier has been discussed already in

(Section 1.5) it will not be covered again in detail here. A diagram of the mycobacterial cell wall can be seen in figure 3.

1.8 Methods of transport across the bacterial outer envelope

The outer membrane of typical Gram negative bacteria excludes, on the basis of size alone, most molecules over around 650 Da in size. Since the bacterial membrane is an impenetrable barrier for substances both harmful to the cell and also to those that are beneficial, it is necessary to have a number of means by which molecules can enter and leave the bacterium.

There are three such systems by which bacteria communicate with the outside world by the exchange of small molecules (Nikaido, 1994);

i) Specific diffusion channels

ii) active efflux pumps with high specificity

iii) non-specific channels (the porins)

These are necessary to allow the passage of all the vital solutes, which are unable to freely diffuse through the membrane.

1.8.1 Specific Diffusion Channels

An example of a specific diffusion channel is the LamB protein of *Escherichia coli* (the phage λ receptor). This is a transmembrane protein that allows the specific diffusion of maltodextrin, maltose and glucose. Indeed under carbon starvation LamB is the major route into the cell for glucose (Nikaido, 1999). The LamB protein has a β-barrel structure formed from 18 antiparallel β-strands and is trimeric - very similar in structure to the classical porins, which will be discussed in detail later. The feature distinguishing specific diffusion channels from the general porins is the presence of a high-affinity binding site for the respective ligand in the channel section of the protein, typically around the third external loop region, which is responsible for the zone of narrowest constriction (Nikaido, 1993). LamB, the 'maltoporin' has three internal loops in the β-barrel which constrict the channel to a minimum diameter of 5Å (Schirmer, 1998).

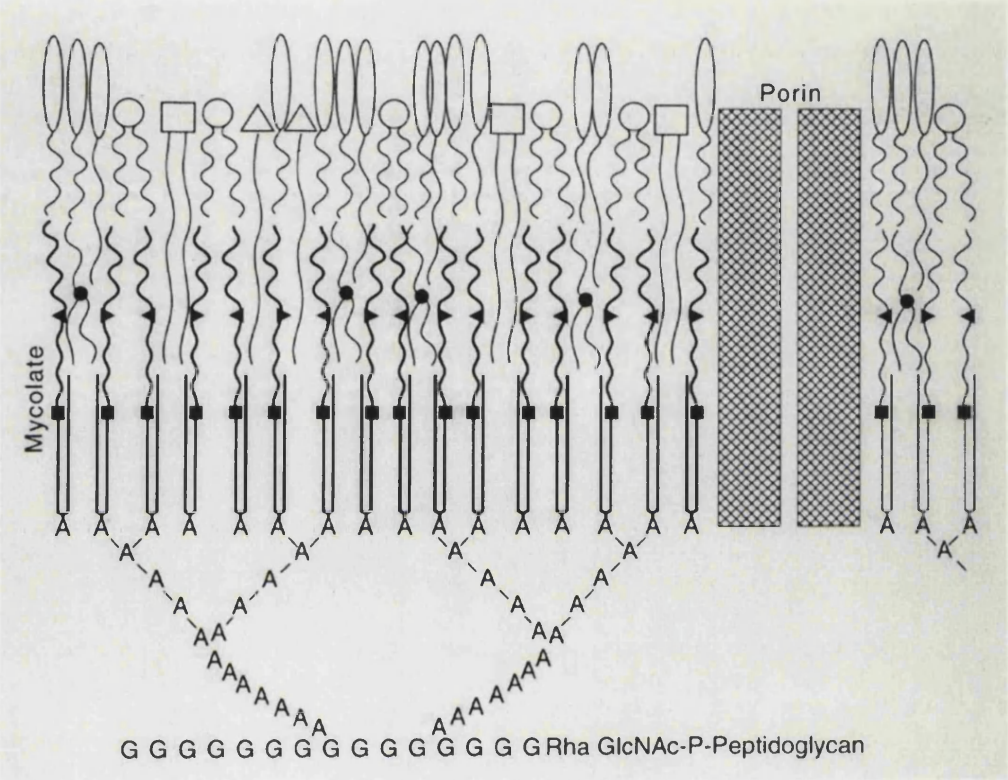


Figure 3 The mycobacterial cell wall

Solid squares represent proximal double bonds or *trans*-cyclopropane groups, solid triangles represent distal double bonds or *cis*-cyclopropane groups. A and G represent part of the arabinogalactan complex.

Triacyl glycerols are represented by solid circles, associated lipids are shown as open symbols, a mycobacterial porin is shown traversing the wall.

This diagram taken from (Liu *et al.*, 1999)

There are other known examples of specific diffusion channels in bacteria other than *Escherichia coli*, such as *Pseudomonas aeruginosa*. It was originally believed that this species lacked entirely any non-specific porins in the outer membrane and so was dependent on specific channels for solute transport, explaining the very low diffusion speed across the membrane. It has since been demonstrated that *Pseudomonas aeruginosa* does have non-specific channels, such as OprF (Rawling *et al.*, 1998). Nevertheless there are also specific channels such as OprD (D2) and the OprB (D1) protein responsible for the transport of imipenem, an antibiotic, and glucose / xylose respectively (Nikaido, 1993).

1.8.2 Active Efflux Pumps

It is important to note that where a bacterial membrane may appear to be impermeable to a substance, it is possible that it is permeable, but that a very efficient active efflux system exists, which immediately excludes the solute studied from the cell. Such a phenomenon is referred to by Nikaido (1999) in discussing the apparent impermeability of *Salmonella typhimurium* cells to nafcillin. In fact the cells are not impermeable but are able to actively pump out the antibiotic via the wide-specificity multidrug efflux pump AcrAB.

It is becoming increasingly obvious that efflux pumps in bacteria emulate the multidrug resistance (mdr) pumps found in mammalian cells. They are often sufficient for a range of resistance to antibacterials, which are able to gain entry into the cell but cannot persist for a sufficient amount of time to do any significant damage before they are exported.

Active efflux systems have been elucidated in a number of bacterial species, notably *Escherichia coli*, *Staphylococcus aureus*, *Bacillus subtilis* and in mycobacteria. *Staphylococcus aureus* and *Bacillus subtilis* have homologous efflux pumps in NorA and Bmr respectively which are demonstrated to pump out cationic dyes, puromycin and chloramphenicol (Nikaido, 1994). Interestingly chloramphenicol is an uncharged species; thus the basis of selective transport through the efflux pump cannot be based on charge alone. *Escherichia coli* has an efflux pump named EmrB, which is responsible for the transport of carbonyl cyanide *m*-chlorophenylhydrazone (CCMP), an uncoupler. This pump is also responsible for the efflux of phenylmercuric acetate, nalidixic acid and thiolactomycin (Nikaido, 1994).

The species mentioned above have a limited homology to mycobacteria; nevertheless the same systems and principles may well be present in the mycobacteria due to the nature of their cell envelope. It is proposed that the active efflux systems in Gram negative bacteria are composed of two, or more, parts. Firstly an efflux transporter (involved in transport of the compound of interest from the cytoplasm to the periplasmic space), an accessory protein (which transfers the solute through the periplasm to the outer membrane) and finally a further channel in the outer membrane which allows the free diffusion of the solute (Nikaido, 1994). Since the mycobacteria may have a periplasmic space it is not inconceivable that such systems exist in this genera.

Also worthy of note are the iron-specific siderophore transporters FhuU and FepA which utilise the active pump of TonB in order to transport iron across the outer membrane (Koebnik *et al.*, 2000).

1.8.3 Non-Specific Channels, The Porins

The third system by which bacteria communicate with the outside world is via non-specific channels which permit the free diffusion of molecules based on size alone, with some selection based on the charge of the solute and the charges present around the mouth of the channel formed by the porin.

The porins are water-filled channels which are present in the outer membrane of (predominantly) Gram negative bacteria and are responsible for providing the outer membrane with its quality of being a 'molecular sieve' (Jap and Wallian, 1996; Nikaido, 1999). They are also responsible for making the outer membrane an inherently more leaky system than the inner membrane. Porins are not only non-specific but (in contrast to the efflux pumps) are also passive in nature; there is no active system driving the passage of molecules through these pore-forming proteins.

Broadly speaking, the porins can be considered simply as holes in the outer membrane of the bacteria; they are either monomeric or trimeric in nature and have a pore diameter of around 7-10Å. Porins have a distinctive topology; instead of the more usual α-helical structure found in the cytoplasmic membrane proteins they are predominantly β-sheets (normally 16 or 18) which are pleated in an antiparallel fashion to give rise to a β-barrel structure (Jap and Wallian, 1996). It is through the pore formed by this barrel that solutes

are able to traverse the membrane. It is thought that the β-barrel structure is necessary in order to allow the porin proteins to cross the cytoplasmic membrane on their way to the outer membrane. If they were too hydrophobic they would become stuck in the inner membrane (Koebnik *et al.*, 2000). The consequences of this would be disastrous, as the contents of the cytoplasm would be able to leak out into the periplasm and through the outer membrane. The proteins adopting their tertiary and quaternary structure after they have passed through the cytoplasmic membrane overcome this. (Nikaido, 1994).

Escherichia coli has a battery of porins, including the classical trimeric porins OmpC, OmpF and PhoE ('phosphoporin') as well as more minor proteins such as the monomeric OmpA and the specific porins. This leads to an effective system for solute transport that avoids having to use the energy-expensive ATP driven pumps to actively move solutes out of the cell. These active pumps are present in *Escherichia coli*, they function in situations where are more rapid response is required or where a process is being driven across a gradient. Each of the *Escherichia coli* porins contributes a particular set of conductance characteristics to the outer membrane and by exploiting whichever of these is most appropriate to the circumstance the bacterial cell is able to tolerate a wide and varied array of environments.

The classical porins from *Escherichia coli* all exist as tightly associated trimers, with each of the (3) barrel like structures forming its own discrete channel. The bond between associated subunits is sufficiently strong to avoid even harsh denaturing treatment such as heating in 2 % sodium dodecyl sulphate (SDS) at 70°c or treatment with 5 M guanidinium hydrochloride (Koebnik *et al.*, 2000). This stability is conferred by salt bridges between neighbouring NH_2- (amino- termini) and -COOH (carboxy- termini) and by hydrogen bonds between adjacent monomers. These monomer subunits often go so far as to share external loops out of the barrel structure and thus interlock with each other (Jap and Wallian, 1996).

Jap and Walian (1996) consider the substructure of porin molecules in three parts: the external vestibule and mouth, the constriction region, and the internal vestibule and mouth. This it can be imagined that the molecules form an 'hourglass' shape. These regions will be considered in turn:

1.8.31. External vestibule

The structure of a porin monomer is such that there is a wide entrance, a wide exit and a region of much narrower girth in between. The external vestibule is the entrance region of the porin and is responsible for the ion specificity of the porin molecule. Porins are traditionally considered to be non-specific. However, this refers only to particular molecules; it is accepted that there may be a general preference for, say, anions or cations in any particular porin. The distribution of charged amino acid residues around the entrance of the porin is crucial in determining which molecules are transported preferentially. Whether this is due to the forces acting intramolecularly or upon the solute itself is unclear, for example all of the *Escherichia coli* porins are known to exclude lipophilic solutes, which is explained by Shultz as being due to one side of the central ß-barrel containing negatively charged residues and the other side containing positively charged residues. The result of this is that charged side-chains of the amino acids are maximally extended, conferring rigidity on the structure (Schulz, 1993). There is also a strong electrostatic field surrounding the residues which serves not only to hold water molecules in place but also to make it energetically unfavourable to move them (Nikaido, 1994). In contrast to the internal vestibule the loops between the β-sheets in the external vestibule are comparatively long and flexible and show little homology between porins or species. There is a suggestion (Schirmer, 1998) that this structural variability forms some kind of bacterial 'camouflage', preventing easy recognition of the porin as an antigen.

Many of the porins were first discovered due to their ability to act as bacteriophage receptors and are also targets of antibodies and proteases, however bacteriophages are highly species-specific and show little cross reactivity to the porins of other species than their host. This demonstrates how external vestibule variability serves to prevent extra-species phage recognition.

1.8.32. Constriction region

It is this region which confers the size-exclusion criteria on the porin through the action of loop L3, which is a short region of α-helical structure around 30-50 amino acids in length which folds back into the channel formed by the beta sheets (Jap and Wallian, 1996). It is important to note that the zone of constriction is not a simple circular aperture and that many other factors such as the shape of the molecule and the distribution of

charge are important in governing the rate of diffusion through the pore. When the pore diameter of porins is stated, it is accepted to be the diameter at which the pore is narrowest; for example whilst the classical *Escherichia coli* porin OmpF might be expected to have a pore diameter of around 23 Å (based on 16 transmembrane β-strands which are 4.5 Å apart, $(16 \times 4.5)/\pi = 23$) in actual fact the pore is close to 11 Å in diameter (Nikaido, 1993).

1.8.33. Internal vestibule and mouth

This region of the porin is the most well conserved in terms of topology, and is dependent on the properties of the barrel alone rather than the connecting loops which are much shorter than those on the external surface. There are fewer charged residues in this portion of the protein and it is suggested (Jap and Wallian, 1996) that the function of the internal vestibule is to provide an easy means of exit from the porin into the cell for transported solutes.

1.9 Porins from Gram Positive Bacteria

Previously porins have been discussed almost exclusively with reference to the Gram negative bacterial species, however it is now evident that some Gram positive bacteria have porins in their membrane as well.

Recent research has revealed porins in the cell envelope of *Nocardia asteroides* (Riess *et al.*, 1999), *Nocardia farcinica* (Riess *et al.*, 1998), *Corynebacterium glutamicum* (Lichtinger *et al.*, 1998) and speculatively in *Rhodococcus* (Sutcliffe, 1998). Importantly in terms of this discussion these actinomycetes are all closely related to the mycobacteria and all contain mycolic acids in their cell wall, albeit shorter ones than the mycobacteria - they can be considered as members of the supragenic taxon *mycolata.*

It can be argued that the other Gram positive bacteria, which have no mycolic acids in the cell wall and hence are naturally more permeable, have no need for porins. Indeed at the time of writing no porins have been discovered in Gram positive bacteria other than the mycolata.

1.10 Porins from the Mycobacteria

The first suggestion that there may be porins in the cell envelope of mycobacteria came from Jarlier and Nikaido (1990) whilst studying the permeability of the cell wall of *Mycobacterium chelonae*. They deduced that the change in permeability conferred by changing the temperature was not consistent with an increase in lipid fluidity alone (by comparison with a model lipid bilayer). As a result of measuring the diffusion of drugs of varying hydrophobicity it was deduced that the diffusion of small, hydrophilic molecules into the cell was occurring in an aqueous environment such as that found in porins. Conversely diffusion through a lipid bilayer occurs by dissolution into the lipid phase - hence the efficacy of hydrophobic antibiotics, this is not affected in the same way by hydrophobicity of the solute. Further evidence for mycobacterial porins was provided in 1994 by Trias and Benz (1992), again using *Mycobacterium chelonae*. This was the first direct demonstration of a porin from any Gram positive bacterium. A 59 kDa cell wall protein was isolated and demonstrated to have a pore diameter of 2.2 nanometers (22 Å) and a single channel conductance of 2.7 nanoSiemens, properties indicative of a porin-like structure. The way in which this porin was isolated was by extracting the cell walls (prepared by step gradient centrifugation) with a detergent, Zwittergent 3-12, and by measuring the rate of swelling of liposomes reconstituted with the extracted protein. It was found that the rate of swelling of the liposomes was inversely proportional to the size of the molecules that were being tested for diffusion into the liposome and was not dependent upon the temperature. This suggests permeation through a water-filled pathway and not a hydrocarbon core. The 'liposome swelling assay' is a standard technique for the detection of pore-forming proteins and has been used many times in the identification of such proteins.

A channel forming activity was identified in purified walls of *Mycobacterium smegmatis* (Trias and Benz, 1994) which showed similar properties to that previously identified by the same group in *Mycobacterium chelonae* (the degree of relatedness between *Mycobacterium smegmatis* and *Mycobacterium chelonae* is distant). This porin was shown to have a pore diameter of 3 nm (30 Å) and exhibits some selectivity for cations as a result of negative point charges around the mouth of the external vestibule.

Further work on *Mycobacterium smegmatis* has revealed a protein, hypothesised to be responsible for the channel forming activity, of 40 kDa in size (Mukhopadhyay *et al.*,

1997) and a gene, *mspA*, encoding a protein with an apparent mobility on polyacrylamide gels of 100kDa, and a mass of 19.4kDa when subjected to mass spectrometry (Niederweis *et al.*, 1999). The protein identified in (Mukhopadhyay *et al.*, 1997) formed a non-specific channel allowing the permeation of small molecules such as sugars and amino acids whereas the protein encoded by *mspA* showed a strong cation selectivity and a conductance of 2.3 nanosiemens. The mature protein encoded by mspA was shown to be a multimer of 19.4kDa subunits since boiling with 80 % dimethylsulphoxide (DMSO) for 15 minutes reduced a 100kDa homomultimer to its component monomers. The N-terminal sequence of both proteins has been determined and shows no similarity, indicating that there are possibly 2 distinct porin like proteins in *Mycobacterium smegmatis*. It is noted by (Niederweis *et al.*, 1999) that contaminating amounts of the *mspA* product are present throughout a polyacrylamide gel and may give rise to an apparent pore-forming activity of purified protein, and that channels may exist which show a very low level of pore-forming ability. Both of these suggestions are in agreement with the results of (Mukhopadhyay *et al.*, 1997).

Interestingly, an attempt to hybridise a *mspA* based probe to genomic DNA digests of various mycobacteria resulted in the finding that there was no *mspA* homologue in the slow-growing mycobacteria (Niederweis *et al.*, 1999). Nevertheless it did seem logical that there would be similar proteins in the cell wall of the slower growers, in particular *Mycobacterium tuberculosis* and Brennan and Nikaido stated in 1994 that 'Similar porins are probably distributed widely among mycobacteria' (Brennan and Nikaido, 1995).

Lichtinger and colleagues discovered two types of channel forming proteins in the cell envelope of *Mycobacterium bovis* BCG with properties similar to those in *Mycobacterium smegmatis*. These channels had a single channel conductance of 4 nanoSiemens and 780 picoSiemens in 1 M KCl (Lichtinger *et al.*, 1999) and were cation selective and anion selective respectively. These were identified by detergent extraction of the cell wall and subsequent lipid bilayer experiments.

1.10.1 Porins in the cell wall of *Mycobacterium tuberculosis*

The first porin in *Mycobacterium tuberculosis* was discovered by Senaratne *et al* in 1998 (Senaratne *et al.*, 1998). This protein, termed OmpATb due to homology with the minor porin OmpA of *Escherichia coli* and other members of the OmpA family, was identified

by searching the annotated tuberculosis genome (http://www.sanger.ac.uk/ Projects/M_tuberculosis/) for genes which looked likely candidates to give rise to pore forming proteins. The current annotation of the gene from the Sanger centre database is:

Rv0899

Starts at 1002810

Ends at 1003787

Molecular weight 33543

Gene name ompA

(MTCY31.27), len: 326. Similar to many members of the OMPA family of outer membrane proteins e.g. YIAD_ECOLI P37665 hypothetical 22.2 kd lipoprotein in b (219 aa), FASTA scores; opt: 296, z-score: 322.0 E(): 2.2e-11, 45.3 % identity in 117 aa overlap; contains PS00044 Bacterial regulatory proteins, lysR family signature.

This gene was cloned into a suitable expression vector (pET15b, Clontech) and purified by metal affinity chromatography. The purified protein was demonstrated to have pore-forming ability by liposome swelling assay, with an estimated pore size of 1.4 – 1.8 nm and lipid bilayer assay showed the conductance of the channel to be 700 ps. Thus the porin identified here could be identical in nature to the one identified in BCG the following year by Lichtinger *et al* (1999). Estimates put the amount of OmpATb protein as a proportion of total protein at around 0.02 % using enzyme linked immunosorbent assay (ELISA), or at around 2 % by activity in liposome swelling assay. This was taken to indicate that a large proportion of the recombinant protein might be present in an inactive form by comparison with the native protein. Importantly OmpATb is thought to exist as a stable monomer and not as a trimer like the classical porins of *Escherichia coli.*

Further work on *Mycobacterium tuberculosis* has revealed the presence of other porins distinct from OmpATb (Kartmann *et al.*, 1999). These are a cation-selective, 0.7

nanosiemens conductance channel composed of 15kDa subunits and a 3 nanosiemens conductance channel with an apparent molecular weight of greater than 60kDa. It is noted that the channel with a 0.7 nanosiemens conductance is not the same as OmpATb since elution of the protein from a polyacrylamide gel shows this property only in fractions of around 15kDa and not around 38kDa which is the apparent mobility of OmpATb (predicted mass is 33.5kDa). This group concluded that the 0.7 nanoSiemens channel was a minor protein of *Mycobacterium tuberculosis* and that the prevalence of both porins in this bacterium was much lower than in the fast growing species. This is in concordance with the observation that the slow growing mycobacteria are far less permeable than the fast growers.

It is evident at this point in time that there are at least 3 distinct proteins in the cell wall of *Mycobacterium tuberculosis* that exhibit pore-forming ability.

1.11 Voltage Gating of Porin Channels

Work on porins both from Gram negative bacteria and from the mycolata has shown that there exists a way in which the channels can be coaxed into an 'open' or 'closed' state. Application of a potential difference across a lipid bilayer containing the pore demonstrates that there is a threshold value at which the conformation of the channel changes. The exact significance of this is not known as some of the voltages concerned are not physiologically relevant, however it seems unlikely that the phenomenon is merely artefactual and that other factors such as ionic strength may make the measured potential at which gating occurs appear artificially high. Thus it may be that *in vivo* voltage gating is a way in which the bacteria are able to modify permeability in response to the environment.

An example of voltage gating occurs in OmpATb, where a potential difference of ±140-150 mV is sufficient to close the channel reversibly. The porin from *Mycobacterium chelonae* has also been shown to be voltage gated (Trias and Benz, 1993), switching to a closed conformation when a potential difference of 40 mV was applied across a lipid bilayer.

In other porins it has been noted that sensitivity to potential difference is heightened in the presence of acid (Schirmer, 1998). However the *in vivo* significance of this is complicated by the fact that acid is able to modulate the expression of porins at the

genetic level by altering transcription rates. This phenomenon will be discussed shortly (Section 1.12). The mechanosensitive channel MscL has been characterised from *Mycobacterium tuberculosis* (Chang *et al.*, 1998) by x-ray crystallography and shown to have a diameter of around 18 Å at the mouth leading to a region of greater constriction, down to around 2 Å. This publication is largely theoretical and is based on data from synchrotron radiation and computer modelling, however it is demonstrated that the protein has an 'open' and 'closed' state and is thus a means by which the bacterium is able to moderate the permeability of the membrane. MscL is a protein that exists in the plasma membrane, and not in the wall (as OmpATb is thought to do), it serves to demonstrate the importance of being able to close the channel if it would be hazardous to the cell for ti to be open.

1.12 Dynamic alteration of bacterial membrane permeability

Often bacteria have to be able to survive in a number of harsh environments; for example *Escherichia coli* is able to live in the gut where it may encounter physiological stresses that would be detrimental without a means of resistance. It is also able to pass through the food chain in order to enable re-colonisation of further hosts and is able to survive in fresh water. Similarly *Mycobacterium tuberculosis* is able not only to survive in the hostile environment of the host macrophage, but also to persist in soil and in the bloodstream, both very different environments.

One way in which bacteria are able to persist when under environmental stress is to modulate their permeability via the porins. It has already been explained that voltage gating may be an important way in which permeability can be controlled; a far more intensively studied and large-scale means is by altering the amount of porin present in the outer membrane at the level of transcription. That is to say that when conditions demand it, the amount of porin in the outer membrane can be increased (making the bacterium more permeable) or decreased (less permeable) by altering the mRNA levels from the gene responsible. In fact the situation is less clear-cut than this and in bacteria that have more than one porin the ratio of porins relative to each other is important. This has been most intensively studied in *Escherichia coli*; it has been shown that although the difference in pore size between the major porins OmpF and OmpC is only 0.1 nm (OmpF=1.2 nm, OmpA=1.1 nm (Pratt *et al.*, 1996)) a significant way in which *Escherichia coli* makes itself less permeable to bile salts and other harsh external

environments is to increase the ratio of OmpC to OmpF (Graeme-Cook, 1991; Liu and Ferenci, 1998; Mizuno and Mizushima, 1990; Pratt *et al.*, 1996). Thus when it is harmful to the cell for solutes in the environment to gain entry the levels of OmpF in the membrane will be low in comparison to OmpC and *vice-versa.*

1.13 Modulation of Porin levels in *Escherichia coli* - a bacterial two-component regulatory system

The two-component system, EnvZ-OmpR, is responsible for the translation of an environmental stimulus into an effect on gene expression. EnvZ is a membrane-bound sensor protein, which upon receiving an environmental stimulus is able to transfer a phosphate group to its cognate regulator OmpR (Aiba *et al.*, 1989). Their genes are transcribed together in the *ompB* operon (Hall and Silhavy, 1981; Waukau and Forst, 1992). EnvZ is also able to function as a phosphatase and to remove the phosphate group from the regulator; thus sensors such as EnvZ are termed kinase/phosphatases. OmpR is not membrane bound and is able to interact directly at the DNA level with the *ompC* and *ompF* promoter regions. The addition of phosphate to EnvZ is an energy-dependent process, yielding ADP during phosphorylation and inorganic phosphate during dephosphorylation, it has been possible to isolate mutants of EnvZ which are unable to de-phosphorylate OmpR (Mizuno and Mizushima, 1990).

Much of the work on regulation of porin levels in *Escherichia coli* has been concerned with the effect of osmotic stress, since this is an easy condition to simulate in the lab. However it is important to note that many other stresses such as pH, temperature, anaerobiosis, carbon source and turgor also affect porin levels.

Whilst it was initially thought that OmpR (phosphorylated) served only as a positive regulator of porin genes (Aiba *et al.*, 1989), and that OmpR and OmpR-P bound to promoter regions of the porin genes with varying specificity (Mizuno and Mizushima, 1990), it has since been shown that at times of high osmolarity OmpR-P serves to repress transcription of *ompF* and enhance *ompC,* whereas at low osmolarity *ompF* is activated whilst *ompC* is repressed (Pratt *et al.*, 1996). The way in which this happens is due to the presence of binding sites with different affinities for OmpR-P upstream of the porin-encoding genes.

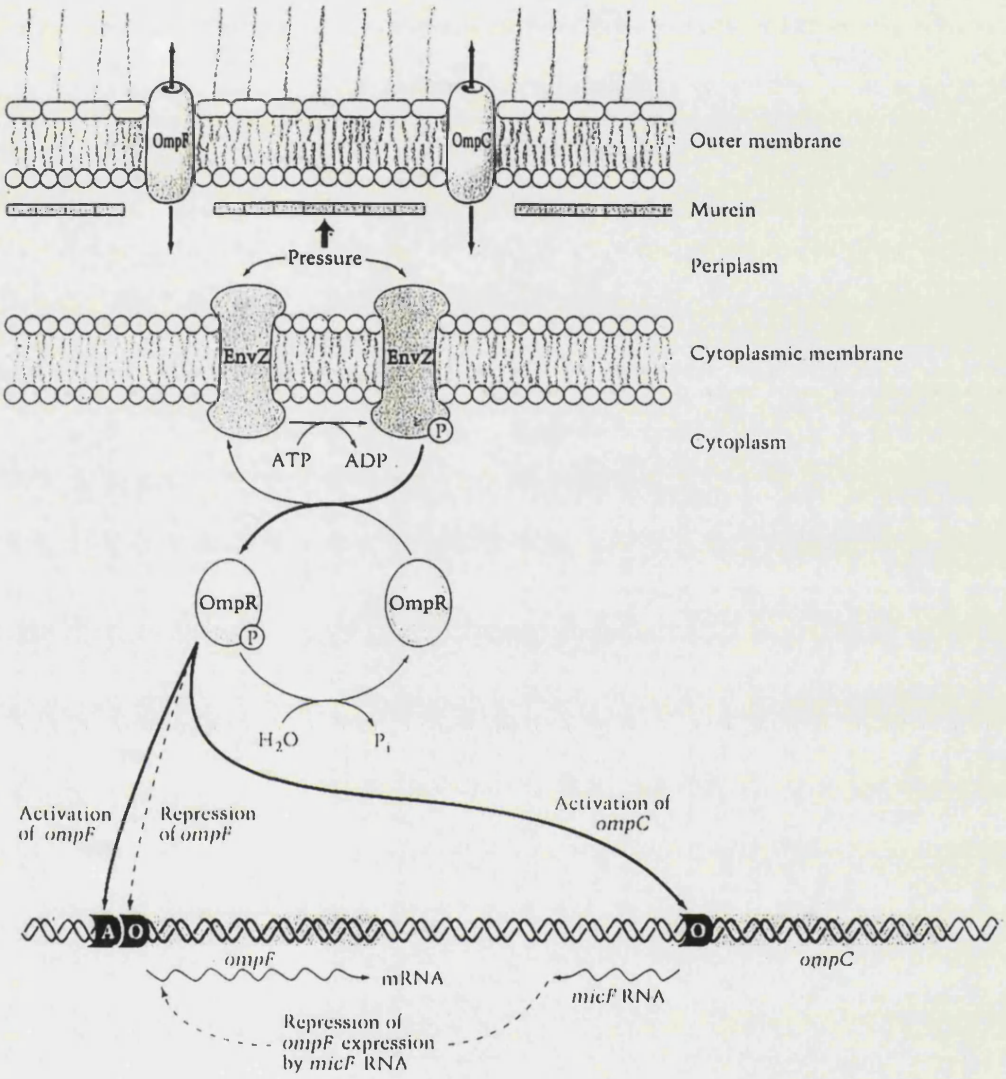


Figure 4 Regulation of porin levels in *Escherichia coli*

This diagram taken from (Neidhart *et al.*, 1990)

The level of OmpR (non-phosphorylated) in the cell is of little consequence apart from having the ability to become phosphorylated and hence become more active. OmpR is not itself able to interact as a transcription factor although it does bind to the upstream region in an unphosphorylated form at high-affinity sites (Huang and Igo, 1996; Lan and Igo, 1998).

There are currently believed to be 4 binding sites for OmpR-P upstream of *ompF* (F1 - F4) and 3 upstream of *ompC* (C1 - C4) (Head *et al.*, 1998). It is the involvement of the furthest upstream site, F4, which is -351 to -384 base pairs upstream of the *ompF* gene, which is thought to negatively regulate the expression of the gene at high osmolarity. Under such conditions occupation of this site has been proposed to induce a hairpin loop in the DNA structure, with Integration Host Factor (IHF) acting as a pivotal protein around which the loop forms (Pratt *et al.*, 1996). Since phosphorylation of OmpR increases its binding affinity not only for the DNA but also to itself (Huang *et al.*, 1997), co-operative binding of OmpR-P to other OmpR-P units when it is bound both at site F4 and at the sites closer to the gene would prevent proper binding and function of RNA polymerase.

Another model has been proposed which does not include the F4 site (Huang *et al.*, 1997), this depends on occupancy of the lower affinity sites F2 and F3 by phosphorylated OmpR. This allows RNA polymerase to bind and function properly (a direct interaction between OmpR and RNA polymerase has been proposed in (Slauch *et al.*, 1991)). However, this model appears to contradict the majority of findings, which show that at high osmolarity high levels of OmpR-P serve to repress transcription of *ompF* (Head *et al.*, 1998; Heyde *et al.*, 2000; Lan and Igo, 1998). It is also suggested that the phosphorylation of OmpR, leading to an increase in co-operative binding, is not sufficient by itself to alter the expression of porin genes significantly (Head *et al.*, 1998).

There are studies that show that DNA binding by OmpR leads to an increase in phosphorylation (Ames *et al.*, 1999) whether by the cognate sensor EnvZ or by use of artificial phospho-donors.

It is evident, therefore, that it is not simply a case of phosphorylated OmpR being a positive regulator of porin expression, but rather a complex scenario involving the affinity of binding sites, the proportion of phosphorylated OmpR, the proportion of OmpR bound

to DNA and the number of OmpR-P molecules which are bound co-operatively to each other as well as the DNA. An equilibrium situation exists that is easily perturbed by any environmental input leading to phosphotransfer from the sensor protein.

The mechanism of environmental sensing by EnvZ remains to be elucidated.

1.14 Other two-component systems in bacteria

In all respects EnvZ-OmpR in *Escherichia coli* is a typical bacterial two-component system. It has a membrane bound sensor and a cytoplasmic regulator, between which the transfer of a phosphate group is sufficient to alter the mechanism of transcription through the action of the regulator. Such systems are widespread in the bacterial kingdom and are proposed to have come from a common progenitor and to have evolved specific functions (Hoch, 2000).

In bacterial systems the transfer of phosphate groups occurs between a histidine residue on the sensor and an aspartate residue on the regulator, this is no different in EnvZ-OmpR with the residues shown to be His243 (EnvZ) (Hsing and Silhavy, 1997) and Asp55 (OmpR) by mutational analyses (Mizuno and Mizushima, 1990; Waukau and Forst, 1992). It has also been shown that it is possible to phosphorylate bacterial regulatory components artificially using low molecular weight phospho-donors such as acetyl phosphate (Lukat *et al.*, 1992; McCleary and Stock, 1994; Nakayama and Watanabe, 1998). The *in vivo* significance of this is demonstrated in a publication by Heyde *et al.*, (2000) in which at varying external pH the importance of intracellular acetyl phosphate on porin levels is shown. It is also shown in a paper by McCleary and Stock, (1994) in which acetyl phosphate levels are suggested to amplify an adaptive response rather than being responsible for it.

1.15 Two-component systems in *Mycobacterium tuberculosis*

Two-component systems in the mycobacteria have not been as comprehensively studied as in the Gram negative bacteria (e.g. *Escherichia coli*), however the publication of the *Mycobacterium tuberculosis* genome (Cole *et al.*, 1998) has made it possible to identify a number of two-component systems in this organism. Amongst those studied are TrcR/S (Haydel *et al.*, 1999), MtrA/B (Via *et al.*, 1996) and SenX3/RegX3 (Supply *et al.*, 1997). In fact there are only 11 complete sensor/regulator pairs in the tuberculosis genome

(identified by genetic homology), with a further 8 isolated genes that show homology to sensors or regulators. In *Escherichia coli* there are more than 30.

MtrA (*Mycobacterium tuberculosis* response regulator A) was identified using a probe derived from *Pseudomonas aeruginosa* PhoB, showing how there is a high degree of conservation amongst these genes. It was the first of the *Mycobacterium tuberculosis* two-component systems to be identified biochemically, prior to the availability of the genome sequence. In an *in vitro* study it was shown that it is possible to phosphorylate MtrA using CheA, a non-cognate kinase (Via *et al.*, 1996) showing that phosphotransfer occurs in the *Mycobacterium tuberculosis* systems, as it does in *Escherichia coli.*

Recently it has been shown that *mtrA* is an essential gene in *Mycobacterium tuberculosis.* This was done by employing a homologous recombination knockout strategy (Section 1.16) in which it was impossible to achieve a double crossover without first complementing the single crossover and hence making the bacterium merodiploid for the gene of interest (Zahrt and Deretic, 2000). The same study also provided evidence that the sensor component (*mtrB*) is not an essential gene at least for *in vitro* growth of *Mycobacterium tuberculosis.*

TrcR (*t*uberculosis *r*egulatory *c*omponent response *r*egulator) and TrcS (*t*uberculosis *r*egulatory *c*omponent *s*ensor) were identified using degenerate PCR primers prior to the availability of the genome sequence; this sensor/regulator pair is typical of other bacterial two-component systems. The genes are adjacent on the genome and are transcribed in the same direction, a common situation in bacterial sensor/regulator pairs. Alignment of TrcR with other bacterial response regulators reveals a likely site of the receiver aspartate (Asp52); similarly alignment of TrcS reveals a likely site for the sensor phosphotransfer site (His287). Interestingly overexpression of a full-length TrcS recombinant protein in *Escherichia coli* is lethal, presumably due to interference with the bacterial membrane rather than cross talk with the native sensor (a possibility ruled out by expressing a version of the protein truncated at the phosphotransmitter domain).

SenX3 and RegX3 have been studied in less detail than the two previous examples; whilst they are a two-component system interest in them has been mainly focussed on the intergenic region which consists of a number of repeated base-pair motifs. This has been termed the Mycobacterial Intergenic Repetitive Unit (MIRU). Such a motif is

hypothesised to be of use in the diagnosis and identification of bacteria, its function is unknown (Supply *et al.*, 1997).

To date neither the environmental stresses which trigger any of the sensor proteins, nor any of the genes under the regulation of any of the mycobacterial 2-component regulatory proteins have been identified in this genus.

1.16 Disruption of gene expression in the mycobacteria

A common technique for measuring the effects and importance of genes in bacteria is to disrupt them in a targeted fashion. In mycobacteria this has been achieved by using one of three methods:

1.16.1 Mobile genetic elements

It is possible that transposons have played a major role in the development of the *Mycobacterium tuberculosis* genome, as there are many regions that appear to be replications of each other or insertions of the same sequence at different sites (Cole *et al.*, 1998). Transposons appear in the mycobacterial genome in nature and several insertion sequences have been identified and characterised, among them IS*1096* (Cirillo *et al.*, 1991) and IS*6120* (Guilhot *et al.*, 1992) in *Mycobacterium smegmatis*; IS*1081* (Collins and Stephens, 1991) in *Mycobacterium tuberculosis* complex and IS*6100* (Martin *et al.*, 1990) in *Mycobacterium fortuitum.*

Most of the mycobacterial transposons can be classified into three families: IS3, IS110 and IS256. There are three members of the IS3 family: IS6110, IS1141 and IS1137. The similarity amongst these members exists on the basis of inverted repeat sequences at the termini that show significant homology and also the presence of two open reading frames (ORFs) (A and B, B is frame shifted by -1 relative to A). IS6110 is found in the *Mycobacterium tuberculosis* genome at up to 25 copies (Fang *et al.*, 1999) and is extensively used in restriction fragment length polymorphism (RFLP) studies to investigate the epidemiology of tuberculosis outbreaks (Guilhot *et al.*, 1999). Insertion of IS6110 into the mycobacterial genome is essentially random although it is possible to identify 'hot spots' where it is more common for recombination to occur (Fang *et al.*, 1999).

Members of the IS110 family, including IS900, IS901/902 and IS1110 are characteristic in that there are no terminal inverted repeats and that the insertion sites have been characterised and are non-random (Guilhot *et al.*, 1999), this has the effect of making these transposons less effective targets for RFLP typing as they insert in the same place in non-identical isolates.

The IS256 family contains most of the mycobacterial transposons so far identified, though this may be due to the amount of effort invested in finding suitable targets for RFLP analysis of the slow-growing species. Where characterised the insertion sites for members of this family have been shown to be random, and so far only one member (IS1081) has been demonstrated in the *Mycobacterium tuberculosis* complex (Collins and Stephens, 1991), others have been identified in *Mycobacterium avium*, *Mycobacterium smegmatis*, *Mycobacterium xenopi* and *Mycobacterium gordonae*. Recently other families have been shown to have transposon activity in the mycobacteria whilst not being native in this genus. One example of this is the mariner-based elements, one of which has been identified in the hornfly *Haematobia irritans* and shown to be able to transpose non-specifically into the mycobacterial genome (beyond a requirement for a TA dinucleotide) (Rubin *et al.*, 1999).

Transposons, insertion sequences and bacteriophages represent a means by which elements can be introduced into the mycobacterial genome at random. Naturally occurring elements with the ability to transpose into the genome are introduced, in a suitable vector, in an attempt to disrupt a gene of interest. It is then possible to screen a library of transformed bacterial cells in order to determine which have a disruption by virtue of auxotrophy or virulence determination. Other genes, which display a less distinct phenotype upon disruption, are more difficult to screen for using such a system.

Common vectors for the delivery of transposons into the mycobacterial genome are able to replicate in an *Escherichia coli* or *Mycobacterium smegmatis* host but can then be manipulated in such a way that they become 'suicide' vectors. This is either achieved by using specific origins of replication or by using temperature sensitive origins of replication. Kalpana *et al* (1991) were amongst the first to design an insertional mutagenesis strategy for mycobacteria. An *Escherichia coli* shuttle vector was used and the transformants were screened on the basis of kanamycin resistance, a property of the transposon used (Tn5). This strategy is somewhat complicated as the transposition is

carried out in *Escherichia coli* and not in the mycobacterial host. It is then dependent upon homologous recombination to re-introduce the disrupted gene into the mycobacterial genome. This strategy is therefore more useful in the fast growing members of the genus. It has however met with some success and *Mycobacterium smegmatis* mutants that are asparagine auxotrophs were isolated.

A temperature sensitive approach has the benefit of being usable directly in the mycobacteria without the need to employ an intermediate *Escherichia coli* host. This technique has been used by Guilhot *et al.*, (1994) to generate an insertional mutant library of *Mycobacterium smegmatis* based on incorporation of the Tn611 transposon into pCG63. This is a shuttle vector that is derived from pAL5000 (Labidi *et al.*, 1992) but also includes the *Escherichia coli* pUC18 replicons and is temperature sensitive. Thus transposition of the vector+transposon into the mycobacteria leads directly to transposition into the genome and ablates the need for homologous recombination. Since Tn611 carries the kanamycin resistance cassette, at a non-permissive temperature for vector replication it is simple to screen for transposition events having occurred. More recently Pelicic *et al* (1997) used the properties of the *sacB* gene (Pelicic *et al.*, 1996) under the control of a temperature sensitive promoter to achieve a similar situation in *Mycobacterium tuberculosis*. Using this approach the vector could be replicated to high levels and then efficiently selected against by growing at 39°C in the presence of sucrose. Modified IS*1096*, with a kanamycin resistance element, was transposed into the mycobacterial genome with high efficiency in a random manner.

The systems detailed above are still an inefficient way of introducing a transposon into the mycobacterial genome; a more effective method is detailed by Bardarov *et al* (1997) whereby a shuttle phasmid is used. This will replicate as a plasmid in *Escherichia coli* and a phage in *Mycobacterium tuberculosis*, representing a way to infect every cell in a population with the phage and hence with the transposon of interest. The mycobacteriophage was manipulated so as to become temperature sensitive - lysogeny could be maintained at 30°C whereas at 37°C the phage was unable to replicate. Using the same modified IS1096 construct detailed above (Tn5367) or a mini-Tn10 construct it was possible to deliver transposons efficiently into the mycobacteria. However, mini-Tn10 appeared to integrate into the *Mycobacterium smegmatis* genome in a site-specific manner and Tn5367 transposed at a very low frequency, this was not due to the 11 copies

of IS*1096* already present in the genome as strains with less copies were just as inefficient. Inability of the phage to infect other strains effectively meant that the system was of little use in *Mycobacterium fortuitum*, *Mycobacterium phlei*, *Mycobacterium chelonae* and *Mycobacterium aurum*. Nevertheless transposition was achieved in the *Mycobacterium tuberculosis* genome at random sites in an efficient manner. The main benefit of this system is that it removes the possibility of mutations in the transposon being propagated during the outgrowth stage if plasmids are electroporated into the bacterium, since this means of delivery is inefficient and only a small subset of cells are transformants. In contrast phage infection can deliver the transposon to every cell in a population and is not dependent on potential mutants being descendants of one original transformant.

1.16.2 Use of antisense RNA

This technique, pioneered by Parish and Stoker (1997), does not depend on a disruption of a target gene but rather disruption of gene function by preventing the mRNA transcript from being translated into a protein. The way in which this is done is to drive the plasmid-borne production of a mRNA molecule which is complementary to that of the gene of interest from a promoter which is inducible (the best characterised in the mycobacteria is the acetamidase promoter). The result of induction from the promoter is production of an RNA strand that is able to form an RNA/RNA hybrid inaccessible to the translation machinery.

The one obvious advantage of this system is that it can be used to study genes that are essential to the survival of the bacteria, as until there is addition of acetamide there is a normal gene function. There are disadvantages however, such as the fact that the gene cannot be totally down regulated to the point where the protein is not produced at all, and the wild-type gene is still present in an active form. At this time an antisense technique has not been applied to *Mycobacterium tuberculosis*.

1.16.3 Targeted homologous recombination

Much of the work on gene disruption has been focussed on homologous recombination. This technique allows for the targeted and specific disruption of a gene of interest by multiple crossover events during DNA replication and thus allows for the creation of a

'null' mutant - the definitive measure of gene function. The basis of homologous recombination is that if regions of identity to the mycobacterial genome are introduced on a plasmid, then due to affinity between these regions there is occasionally a crossover event whereby during replication of the DNA the polymerase will cross from one DNA strand (the chromosome) to another (the plasmid. Thus it is possible to integrate the entire plasmid into the bacterial chromosome by a single crossover event. A further crossover event between carefully chosen regions of homology on the plasmid and chromosome will result in loss of the majority of the plasmid and a chromosomal mutation, which is different to the original wild type.

Homologous recombination in mycobacteria, particularly in the slow growers and the *Mycobacterium tuberculosis* complex, is notoriously difficult and is a long process. Typically a high number of transformants are illegitimate recombinants, in which the foreign DNA has integrated in a non-specific manner in the chromosome. Such a problem is difficult to overcome using conventional selection methods, but can be minimised by pre-treatment of the vector DNA by ultraviolet light or alkali denaturation (Hinds *et al.*, 1999). There is also a problem with the transformation of mycobacteria in that introduction of plasmid DNA into the bacteria by electroporation is a rare event. In *Escherichia coli* transformation is a routine practice and results in high numbers of cells that have taken up the foreign DNA, whereas mycobacterial transformants are few. The most likely reason for this is the complex structure of the mycobacterial cell wall, which is difficult to make permeable by heat shock or electroporation (Parish and Stoker, 1995).

While it is possible to have only one region of homology introduced to the bacterial genome, and thus to depend only on a single homologous crossover event leading to insertion of the plasmid borne DNA into the open reading frame of a target gene, a more common approach is to use two regions of homology (both upstream and downstream of the target gene) (Stewart and McFadden, 1999). Between these regions of homology there may be an antibiotic resistance cassette, which effectively replaces part of the coding sequence in the bacterial genome resulting in a 'marked mutation'. Alternatively there may be no resistance cassette, leading to a so-called 'unmarked mutation'.

Whilst DNA can be delivered to the bacterial cell as a non-replicating linear strand with an appropriate disruption (usually an antibiotic resistance cassette) in the target gene (Balasubramanian *et al.*, 1996) this technique has limited efficacy. A more common

means is to use much the same system as for transposon delivery, either introducing DNA on a plasmid that cannot replicate in mycobacteria (a 'suicide' vector) or in a shuttle plasmid which has a counterselectable marker such as temperature or sucrose sensitivity. In this way the amount of suitable DNA in a cell can be increased until such time as the selection pressure is applied, reducing the effect that low transformation efficiency has on the overall success of homologous recombination strategies. A drawback of some systems for homologous recombination is the necessity to use an antibiotic resistant host strain, which is of less use in subsequent *in vivo* studies. This is most evident if the counterselectable marker on the plasmid delivery vehicle is *katG* or *rpsL*. These genes confer isoniazid and streptomycin sensitivity respectively on a mutant host strain that is recessively resistant to these antibiotics due to mutations in the wild type copies of these genes (Norman *et al.*, 1995; Sander *et al.*, 1995). When a true wild type copy of the *katG* or *rpsL* gene is present the cell will be sensitive to the appropriate antibiotic. This indicates that either a single crossover/illegitimate recombination has occurred or that the plasmid has not been lost effectively.

An extremely useful tool in homologous recombination studies is the use of the *sacB* gene. This encodes levansucrase and was initially isolated in *Bacillus subtilis* (Pelicic *et al.*, 1996). Expression of *sacB* in the presence of 10 % sucrose is lethal in the mycobacteria as the enzyme catalyses the synthesis of levans by transfer of the fructose from a sucrose molecule (leaving glucose). These levans, which are fructose polymers, are thought to either interfere with the bacterial periplasm or to transfer fructose molecules to inappropriate receptors, although the exact mechanism of toxicity is unknown (Pelicic *et al.*, 1996). If the *sacB* gene is included in a suicide delivery vector along with a target gene disrupted by an antibiotic resistance cassette then selection for sucrose resistance *and* antibiotic resistance results in growth of only cells that have lost the plasmid. These cells have not undergone a single crossover or illegitimate recombination event and whilst they could either be spontaneous antibiotic resistant mutants or reversions of the *sacB* gene, they are expected to be cells in which homologous recombination has occurred. This technique has been successfully employed by Pelicic *et al* to disrupt the *pyrF* allele in *Mycobacterium smegmatis*, creating uracil auxotrophs (Pelicic *et al.*, 1996) and the *hisD* allele in the same organism (Hinds *et al.*, 1999). Disruption of the *lysA* gene leading to lysine auxotrophy in the slow growing mycobacteria *Mycobacterium bovis* BCG, *Mycobacterium tuberculosis* and the fast

growing *Mycobacterium smegmatis* has been achieved by Pavelka and Jacobs (Pavelka and Jacobs, 1999) in a study which demonstrates the efficiency of allelic exchange in these species.

Used in conjunction with other selectable markers such as antibiotic resistance the *sacB* gene can be used in a two-step strategy whereby single crossovers are picked from a pool of transformants and subjected to a second round of growth in order for a second crossover to occur.

This procedure has the advantage of being able to produce unmarked mutations in a simple fashion and if a number of markers are included on the plasmid which is introduced then the probability of achieving a gene knockout is greatly increased. Parish and Stoker (2000) were able to use vectors incorporating the *lacZ* gene (blue/white screening on X-gal agar), the *sacB* gene (sucrose sensitivity if present) and a combination of hygromycin or kanamycin resistance to create multiple unmarked mutations in the *Mycobacterium tuberculosis* genome. This was possible by careful selection of single crossover events (blue colonies which are sucrose sensitive and antibiotic resistant) and then by growth of these colonies in the absence of antibiotic followed by plating on sucrose/X-gal agar. Only white sucrose resistant colonies need to be examined in order to determine whether they are double crossover homologous recombinants.

Homologous recombination as a technique is still in its infancy, however current methods are yielding mutants in diverse genes in the mycobacteria. Developments in counterselectable markers and means to improve efficiency of recombination and transformation will result in this technique being more widely used as a way of studying mycobacterial virulence in the future.

Aims of the project

The aim of this project was to examine the distribution of the *ompATb* gene amongst mycobacterial species and to attempt to identify those species which have genes, and hence proteins, which show homology to the porin from *Mycobacterium tuberculosis.*

It was also important to determine what happened to the bacterial cell if the porin-encoding gene (*ompATb*) was disrupted. The way in which this was done was by employing a homologous recombination based approach to gene knockout.

A separate part of the project involved trying to identify a means by which *Mycobacterium tuberculosis* regulates the level of porin in the cell envelope. It seems likely that a bacterial species that has to survive in a wide variety of environments, some of which are extremely hostile to the cell, has a means to modulate permeability in response to extracellular stress.

Firstly, regulation of the porin gene was shown to occur by using real-time RT-PCR to show that transcript levels from the *ompATb* gene varied upon the addition of environmental stress to the mycobacterial cell. It was therefore attempted to use homology to *Escherichia coli* as a basis on which to search for possible regulatory mechanisms

Having identified a likely regulatory system this was then investigated further in an attempt to prove the function of the proteins involved, the same homologous recombination based approach as for the porin gene was used to investigate how the cell responded to loss of the regulatory protein.

The structure of the porin gene in *Mycobacterium tuberculosis* is not well studied; it was therefore attempted to identify a promoter region using β-galactosidase activity in *Mycobacterium smegmatis.* This was transformed with appropriate integrating vectors comprised of a region upstream of the *Mycobacterium tuberculosis ompATb* gene fused to a *lacZ* reporter gene.

2. Materials and Methods

2.1 Bacterial strains and growth conditions used

Escherichia coli DH5α (Clontech) was grown in L-broth at 37°C in a shaking incubator at 250 r.p.m. Vessels used were conical flasks of a volume 5 times that of the broth. *Escherichia coli* BL21 (DE3) (pLysS) (Novagen) was used as a strain for inducible expression of proteins and was grown in Terrific broth in the presence of chloramphenicol at 34 μg/ml under the same conditions. Where appropriate carbenicillin was included at a final concentration of 200 μg/ml.

Mycobacterium smegmatis mc^2155 (Snapper *et al.*, 1990) was grown in shaking conical flasks (250 r.p.m.) in Dubos broth (Difco) supplemented with Dubos medium albumin (4 %) and glycerol (0.2 %). *Mycobacterium tuberculosis* H37Rv (National Collection of Type Cultures, Colindale, #7416), *Mycobacterium tuberculosis* 1424 (Dr P Sander, Institüt für Medizinische Mikrobiologie, Hannover, Germany) or *Mycobacterium tuberculosis* 1424 Δ*ompATb* were grown in 100 ml cultures in roller incubator bottles (2 r.p.m.) at 37°C, again in Dubos broth supplemented with albumin and glycerol as above.

Mycobacterium bovis Bacille Calmette-Guérin (BCG, Vaccine strain, Glaxo) was grown in 500 ml volumes of Tween-glutamate broth (See appendix) in 2 L conical flasks, shaking at 250 r.p.m. at 37°C, or in 250 ml conical flasks with 100 ml or 50 ml media.

Dubos 7H11 Agar (Difco), supplemented with albumin and glycerol as above, was used for the culture of all mycobacterial species and where appropriate antibiotic was added at the following concentration; Kanamycin 25 μg/ml, Streptomycin (100 μg/ml). Sucrose was added at 2 % and 4-bromo-3-chloro-2-indolyl-β-galactoside (X-gal, Melford) was used at 100 μg/ml when necessary.

2.2 PCR amplification of DNA

All polymerase chain reaction amplification was carried out using a Perkin Elmer Applied Biosystems GeneAmp PCR system 9700. Reactions (50 µl) were composed of 5 µl 10× PCR buffer with $MgCl_2$ (15 mM, Perkin Elmer), 4 µl dNTP mix (2.5 mM), 2 µl of forward and reverse primer (7.5 µM) and 35 µl sterile water. Either 1 µl of Taq Polymerase Gold (Perkin Elmer) was used, in which case an initial denaturation step of 7 minutes at 95°C was necessary, or 1 µl High Fidelity enzyme (Expand kit, Roche). High Fidelity enzyme was used in all procedures that gave products subsequently used in cloning applications and does not require an extended initial heating step: 4 minutes at 94°C was used. DNA template was added at a final concentration of 10-30 ng per reaction, negative controls used 1 µl sterile water.

More detailed descriptions of PCR reaction conditions are given in the appropriate section when applicable.

2.3 Agarose gel electrophoresis of DNA

Agarose gels of either 1 % (DNA products of 0.5 kb-7 kb) or 1.2 % (products smaller than 0.5 kb) were prepared by boiling molecular biology grade agarose (Bio-Rad) in tris-acetic acid-EDTA (TAE) buffer, this contains 40 mM tris-acetate and 2 mM EDTA, at pH 8.5.

Ethidium bromide (Bio-Rad) was added to the melted agarose to a final concentration of 0.8 µg/ml, gels were allowed to set for at least 1 hour at room temperature and run at a constant voltage of 65 V (mini-gels), 80 V (midi-gels) or 100 V (maxi-gels).

Visualisation of bands was performed using an ultraviolet transilluminator at 302 nm and the photograph taken using a Foto-Analyst video camera (FotoDyne) and a thermal printer (Mitsubishi).

2.4 Excision of DNA bands from agarose gels

DNA was extracted from agarose gels using a Qiagen gel extraction kit. Bands were excised using a clean scalpel and 3 gel volumes of proprietary buffer QG added. This was incubated at 50°C until all traces of solid agarose were dissolved. For fragments smaller

than 500 bp or larger than 4 kb, an additional 1 gel-volume of propan-2-ol was added. The dissolved agarose mixture was then passed through a QIAquick spin column by centrifugation at 11600 × *g* for 1 minute. Columns were washed by the addition of 750 µl buffer PE and centrifugation for 1 minute at 11600 × *g*, and were dried by centrifugation for a further 1 minute at 11600 × *g*.

DNA was eluted by the addition of 30 µl sterile water, or 30 µl DNA dilution buffer (Rapid DNA Ligation Kit, Roche) if fragments were to later be used in construction of plasmids, followed by centrifugation at 11600 × *g* for 2 minutes.

2.5 Ligation of plasmid and insert DNA

Vector and insert DNA were ligated using a Rapid DNA Ligation Kit (Roche). Reactions of 21 µl final volume included 1 µl Rapid DNA Ligase, 10 µl ligation buffer, 0.3 µg plasmid DNA and 1 µg insert DNA (as determined by visualisation in an agarose/ethidium bromide gel). Alternatively if sufficient DNA was not available a 1:3 ratio of plasmid:insert was established in the final reaction. Reactions were incubated for 1 hour at room temperature, prior to transformation into *Escherichia coli* as described in (Section 2.7).

2.6 Preparation of competent *Escherichia coli*

An overnight culture of *Escherichia coli* (0.5 ml) was inoculated into 50 ml of pre-warmed L-broth in a 250 ml conical flask. The culture was shaken at 37°C (250 r.p.m.) for 1.5-2 hours until the OD_{600} was 0.3-0.4.

The culture was then kept on ice for 10 minutes prior to centrifugation at 653 × *g* for 5 minutes at 4°C. Cells were resuspended in 20 ml of ice-cold 0.1 M $CaCl_2$ and centrifuged once more as before. Finally cells were resuspended in 2 ml cold $CaCl_2$ and stored on ice for 1.5-2 h.

If cells were not used immediately glycerol was added as a cryo-preservant to 20 % final concentration and cells were stored at -80°C.

2.7 Transformation of *Escherichia coli* by heat shock

Competent cells (100 µl) prepared as described above were added to a 15 µl ligation reaction (Rapid DNA Ligation Kit, Roche) following manufacturer's instructions) and kept on ice for 30 minutes. Cells were then subjected to heat shock for 2 minutes at 42°C before returning to ice for 2 minutes.

Pre-warmed L-broth (1 ml) was added to the transformation reaction and incubated for 1 hour at 37°C. Cells were centrifuged at 4000 × *g* for 5 minutes and 0.9 ml supernatant removed. The remaining 200 µl was plated on L-agar containing the appropriate antibiotic and/or supplement.

2.8 Extraction of plasmids from *Escherichia coli*

Plasmids were extracted from *Escherichia coli* DH5α using a miniprep spin column kit (Qiagen), proprietary buffers P1, P2, N3 and PE are supplied as part of this kit.

Overnight cultures (10 ml) which had been shaken at 250 r.p.m. in L-broth (See appendix) were spun down at 4000 × *g* for 10 minutes. Pellets were resuspended in 250 µl buffer P1 and transferred to 1.5 ml microfuge tubes. 250 µl buffer P2 was added and the suspension mixed by inverting the tube six times. 350 µl buffer N3 was then added and the tube again inverted for six times. The precipitate was removed by centrifugation at 11600 × *g* for 10 minutes and the supernatant loaded onto a Qiaprep spin mini-column (Qiagen). This was spun once at 11600 × *g* for one minute, the flow-through discarded and 750 µl buffer PE added to wash the column. After a 1 minute spin at 11600 × *g* flow-through was again discarded and the column spun to dryness for a further minute. Plasmid DNA was eluted by the addition of 30 µl sterile water onto the column and a final spin at 11600 × *g* for 1 minute. Plasmids were stored at -20°C.

2.9 Preparation of electrocompetent mycobacteria

Cultures at an OD_{600} of around 0.6 were centrifuged in a Sorvall GSA rotor for 30 minutes at 10444 × *g*, resuspended in 100 ml 10 % glycerol and spun down again a total of four times. Finally cells were resuspended in 10 ml 10 % glycerol and stored on ice until use. *Mycobacterium tuberculosis* competent cells were used on the day of

preparation, whereas *Mycobacterium smegmatis* cells were stored at 4°C for a period of up to 4 weeks.

2.10 Electroporation of mycobacteria

Mycobacterial competent cell suspension, prepared as above (400 µl), was added to a 0.5 ml volume (0.2 cm path length) electroporation cuvette (BioRad), along with 5-10 µg of plasmid DNA (1 µg/ml in water). A GenePulser unit (BioRad) was used to deliver a pulse at 25 µF, 2.5 kV and 1000 Ω. The entire culture was then transferred into 2.6 ml pre-warmed Dubos broth with albumin and glycerol supplements using a sterile Pasteur pipette.

Mycobacterium tuberculosis cultures were then incubated at 37°C in a static incubator overnight before being split into 1.5 ml aliquots and spun down at 11600 × *g* for 10 minutes. Supernatant (1.3 ml) was removed and the remaining 200 µl plated onto 7H11 agar (Middlebrook) containing the appropriate antibiotics and supplements. This procedure is suitable for non-replicating vectors (or suicide vectors), as more cells need to be plated in order to obtain transformants

Mycobacterium smegmatis cultures were incubated for 3 hours; 100 µl of the 3 ml culture was then plated directly onto 7H11 agar (Middlebrook) containing the appropriate antibiotics and supplements. This procedure is suitable for integrating or replicating vectors in which it is less important to obtain large numbers of transformants.

2.11 Automated sequencing of plasmids

Plasmids were sequenced using an ABI Prism 377 automated sequencer and results analysed using DNA sequence analysis software V2.1.1 (Perkin Elmer).

A *d*Rhodamine terminator cycle sequencing Ready Reaction Kit (PE Applied Biosystems) was used in the following reaction: 1 µl appropriate primer (3.5 µM), 1 µl dimethylsulphoxide (DMSO, Sigma), 3 µl sterile H_2O, 7 µl plasmid miniprep (from 30 µl elution as described in section 2.8, corresponding to 200-500 ng) and 8 µl Ready Reagent *d*rhodamine sequencing mix. This was mixed in a 0.2 ml microfuge tube and the reaction performed in a Progene thermal cycler (Techne) with the following conditions: 94°C for 2

minutes, 20 cycles of 94°C 30 sec/59°C 30 sec/60°C 4 min, and finally a 4°C hold. Samples were then ethanol precipitated by addition of 74 µl 70 % ethanol/0.5 mM $MgCl_2$, incubated at room temperature for 20 minutes and spun at 11600 × *g* for 20 minutes. Supernatant was carefully removed and pellets stored at -20°C.

Pellets were resuspended in loading buffer (5:1 formamide:25 mM EDTA pH 8 with Blue Dextran (Pharmacia Biotech) 50 mg/ml) and loaded onto a 5 % Long Ranger (Flowgen) sequencing gel (6 M urea, 1× TBE (89 mM tris base, 89 mM boric acid, 2 mM EDTA pH 8)) and run at constant temperature of 51°C for 9 hours.

2.12 PCR probing of Mycobacterial species for *ompATb*

Primers U1 [5' AGG GAG TCA TAT GGT GGC TTC TAA GGC GGG TTT G 3'] and L1 [5' GAA GGA TCC CCC CCA GGA ACG CCA GCA GGT A 3'] (kindly provided by Dr Ryan Senaratne) were used in a PCR reaction using High Fidelity enzyme (Expand kit, Roche). An annealing temperature of 68°C was used, with an extension time of 1 minute for 25 cycles. Mycobacterial DNA was used at a template concentration of 20 ng per reaction, and was extracted as detailed in section 2.23. *Mycobacterium bovis* DNA was kindly supplied by Dr Glyn Hewinson (Veterinary Laboratories Agency, New Haw, Surrey); DNA from other mycobacterial species was kindly supplied by Ms Pat Brooks (NIMR) as Instagene preparations (Section 2.23.2).

2.13 Western blot analysis of mycobacterial cell lysates

2.13.1 Preparation of mycobacterial cell lysates

Cell lysates of mycobacterial cultures of *Mycobacterium chelonae*, *Mycobacterium microti* and *Mycobacterium smegmatis* were prepared by centrifugation of 200 ml culture that had reached an OD_{600} of 0.6 at 11600 × *g* for 20 minutes. Cell pellets were resuspended in 5 ml phosphate buffered saline (PBS, 137 mM NaCl, 3 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4) and were transferred to smaller tubes and spun again as before. Finally pellets were suspended in 500 µl 1 % Zwittergent 3-12 (Roche), 20 mM tris-HCl pH 8, 40 mM disodium EDTA and were transferred to Ribolyser 'blue' tubes (Hybaid). Phenylmethylsulphonylfluoride (PMSF, Sigma) was added to a final concentration of 2 mM and suspensions were lysed by three 20 second pulses in a

Ribolyser (Hybaid) at speed 6. Between pulses tubes were placed on ice for 1 minute to prevent overheating.

Tubes were spun for 5 minutes at 6000 × *g* and supernatant taken for measurement of protein concentration.

2.13.2 Determination of protein concentration

Protein concentrations were estimated using a BCA assay kit (Pierce). Reagents A and B were mixed at a 1:50 ratio and 200 µl of the reagent added to an appropriate volume of protein.

Absorbance at 550 nm was measured after a 30 minute incubation at 37°C using an EL312 ELISA plate reader (Bio-tek Kontron) and protein concentrations determined using the Kineti-Calc software (Bio-tek Kontron).

2.13.3 Polyacrylamide gel electrophoresis

Mini-gels for protein electrophoresis were used at a final acrylamide concentration of 12.5 %. In order to obtain this 6.65 ml water was added to 5 ml resolving gel buffer (1.5 M tris pH 8.8) and 8.35 ml Protogel (30 % acrylamide, 0.8 % bisacrylamide, National Diagnostics). The mixture was de-aerated under vacuum for 2 minutes and 0.2 ml sodium dodecyl sulphate (SDS, 10 %) added (Bio-Rad). Ammonium persulphate (0.2 ml, 10 %, Sigma) and 7 µl TEMED (N, N, N', N' tetramethylene-ethylene-diamine; Bio-Rad) were added and gels poured to a distance of 1cm from the top of the plates. Water-saturated butanol was then added until the resolving gel had polymerised in order to make the top surface of the gel more level. A stacking gel was prepared by the addition of 1.35 ml Protogel (National Diagnostics) to 2.5 ml stacking gel buffer (0.5 M tris, pH 6.8) and 6.15 ml water. After de-aeration 0.1 ml SDS (10 %), 50 µl ammonium persulphate (12 %) and 10 µl TEMED (Bio-Rad) were added. Butanol was removed and resolving gels rinsed with sterile water. Stacking gels were cast and allowed to polymerise overnight.

Gels were run in a Mighty Small II apparatus (Hoefer Scientific Instruments) in running buffer containing 25 mM tris, 0.2 M glycine, 0.1 % SDS. Samples were loaded by combining with an equal volume of sample buffer (125 mM tris, 2.5 % SDS, 20 %

glycerol, 10 % 2-mercaptoethanol, 0.03 % bromophenol blue) and heating at 100°C for 2 minutes. They were then rapidly cooled on ice and loaded at a volume of 20 µl per well.

Gels were run at a constant current of 15 mA (per gel) until the dye front reached the buffer in the lower tank. MultiMark coloured standards (Novex) were used as a molecular weight ruler.

2.13.4 Western transfer of protein gels

Proteins were transferred to Immobilon-P polyvinylidene fluoride (PVDF) membranes (Millipore) using a semi-dry blotting apparatus (Biometra). Three sheets of 3 MM filter paper (Whatman) wetted in transfer buffer (48 mM tris, 39 mM glycine, 0.037 % SDS, 20 % methanol) were placed on the anode, followed by a PVDF membrane that had been pre-wetted in methanol and then transfer buffer. After removal of the stacking gel the rest of the gel was then placed onto the membrane, followed by a further 3 sheets of 3 MM filter paper.

The apparatus was run at a constant 50 mA for 50 minutes and membrane removed for immunoassay.

2.13.5 Antibody based detection of OmpATb in Western blotted cell lysates

Membranes were pre-wetted in methanol and transfer buffer then blocked in 3 % 'Marvel' dried milk mix in tris-buffered saline with Tween (TTBS, 20 mM tris, pH 7.5, 0.5 M NaCl, 0.05 % Tween 20 (polyoxyethylenesorbitan-monolaurate, Sigma)). Primary antibody (polyclonal rabbit anti-OmpATb, kindly supplied by Dr Ryan Senaratne) was added to 3 % Marvel in TTBS at a 1:1000 dilution and membranes shaken gently in this solution for 1 hour at room temperature. Membranes were washed in TTBS for 5 minutes 3 times. Peroxidase conjugated secondary antibody (swine anti-rabbit; Dako) was added to 1 % Marvel mix in TTBS at 1:1600 dilution and the blot shaken for 1 hour at room temperature.

Membranes were then washed again 3 times for 5 minute periods in TTBS, and twice in tris-buffered saline without Tween20. 100 µl development solution was prepared (4 mg 3, 3'-diaminobenzidine, 8 mg NiCl, 100 µl water) and added to 10 ml phosphate buffered saline (PBS). This was added to the membrane and development allowed to occur until

bands were clearly visible. Membranes were then washed thoroughly in water and allowed to air dry.

2.14 Identification of a putative regulatory system

Escherichia coli has a complex and well-studied mechanism of porin gene regulation. This has been previously discussed in section 1.13. The amino acid sequences of *Escherichia coli* EnvZ and OmpR were retrieved from the SRS database (http://www.srs.ebi.ac.uk) and were used to search the genome of *Mycobacterium tuberculosis* at http://www.sanger.ac.uk/Projects/M_tuberculosis/ blast_server.shtml using a ***b***asic ***l***ocal ***a***lignment ***s***earch ***t***ool (protein vs translated DNA, TBLASTN) search . This shows the similarities of the *Escherichia coli* proteins to their *Mycobacterium tuberculosis* counterparts (where present).

Two genes were identified, Rv0902c and Rv0903c, which are present immediately downstream of the gene coding for the porin (*ompATb*). The protein product of these genes showed homology to the proteins which regulate porin expression in *Escherichia coli* and are were present in a position which made them likely candidates for a similar role in *Mycobacterium tuberculosis.*

2.14.1 Gene Rv0903c is transcribed in mycobacteria

Primers RAS021U [GCC GGC GCC GAC GAT TAC CT] and RAS022L [CCA CCG CCG GCC TCC TCC AGT TT] were used to amplify a 350 bp product from cDNA of *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG.

A 65°C annealing temperature was used with an extension time of 30 seconds.

2.15 Expression and purification of Rv0903 protein

2.15.1 Construction of plasmid pRAS1

Primers *Nde*Upper [5' TGC TGG ACA TAT GTC AGG CAA AAT GGG CGG 3'] (restriction site and start codon underlined) and *Bam*HILower [5' TTA AGG ATC CCA GAC GAC GGT GCC GAC AAT GAG T 3'] (restriction site underlined) were used to generate a PCR product of 865 base pairs in length. After digestion with *Nde*I and *Bam*HI

restriction enzymes (Roche) overnight at 37°C in proprietary buffer B (Roche) (Roche) the fragment (851 bp) was run through a 1 % agarose gel and extracted as in section 2.4. This fragment contains the entirety of the Rv0903c gene and additionally 4 codons upstream of the start codon.

Plasmid pET15b (Novagen), an *Escherichia coli* vector which contains the bacteriophage T7 promoter and the *lac* operator, was also digested with *Bam*H1 and *Nde*I enzymes as above and was treated with 1 U alkaline phosphatase (Roche) for 1 hour after an overnight digestion. This was then run through, and extracted from, a 1 % agarose gel (see section 2.4). The result of a ligation (see section 2.5) between the vector and the 851 bp insert was plasmid pRAS1. *Escherichia coli* DH5α was transformed with this plasmid and cultures grown overnight for plasmid extraction (see section 2.8), subsequently the plasmid was transformed into *Escherichia coli* BL21 (DE3) pLysS. Plasmid pRAS1 expresses the Rv0903 protein from the T7 promoter in an inducible fashion following the addition of isopropyl-β-D-thiogalactoside (IPTG, a lactose homologue) in this host strain.

2.15.2 Expression of recombinant Rv0903 protein

Escherichia coli BL21 (DE3) pLysS transformed with pRAS1 was grown in 20 ml L-broth with 100 μg/ml ampicillin and 34 μg/ml chloramphenicol. 15 ml of overnight culture was inoculated into 750 ml L-broth at 37°C (with ampicillin and chloramphenicol) and the culture shaken at 250 r.p.m. until the OD_{600} was 0.6

Expression of the protein was induced by addition of IPTG to a final concentration of 1 mM.

At this point the incubation temperature was lowered to 28°C in an attempt to produce more Rv0903 protein in the soluble fraction. Six hours after induction the cells were harvested by centrifugation at 2611 × *g* for 15 minutes in a Sorvall GSA rotor.

2.15.3 Purification of recombinant Rv0903 protein

Pellets were resuspended in 50 ml extraction buffer (50 mM sodium phosphate, 250 mM sodium chloride, pH 7) and an EDTA-free protease inhibitor tablet (Stratagene) added to the suspension.

Cells were lysed by sonication (five 30 second bursts at 140 W) using a Soniprobe (DAWE instruments). Lysed suspensions were then centrifuged at 13000 × *g* for 15 minutes and the supernatant taken and subjected to high-speed centrifugation (48246 × *g* for 20 minutes) in a Sorvall SS34 rotor.

The supernatant was incubated with 4 ml Talon resin (Clontech) which had been pre-washed in extraction buffer and allowed to equilibrate. Incubation was carried out at room temperature for 1 hour. Resin was then loaded onto a column and washed first with extraction buffer (flow rate 0.8 ml/min) and then wash buffer (50 mM sodium phosphate, 100 mM NaCl, 10 mM imidazole (Fluka), pH 7) until the A_{280} stabilised.

Elution buffer (50 mM sodium phosphate, 100 mM sodium chloride, 300 mM imidazole) was then added and fractions (3.2 ml) collected every four minutes. Fractions were then checked using SDS-PAGE (Section 2.13.3) and those which contained the protein of interest were pooled together and run through a Superose 12 HR 10/30 size-exclusion column (Amersham Pharmacia) with a bed volume of 240 ml. This had been previously equilibrated with 50 mM tris, 200 mM NaCl pH 7.5. Flow rate was 1 ml/min and fractions (4 ml) were collected every 4 minutes. Concentration of the fractions was determined by spectophotometry and glycerol added to a final concentration of 20 % before freezing at -20°C.

2.16 Gel Retardation assays

Non-denaturing 8 % polyacrylamide gels were prepared by mixing 19.95 ml Protogel (National Diagnostics) with 3.75 ml 10X TBE (89 mM tris base, 89 mM boric acid, 2 mM EDTA pH 8), 51.3 ml water and 112 µl each of 25 % ammonium persulphate (APS; Sigma) and TEMED (Bio-Rad). Gels were allowed to polymerise overnight at 4°C and were pre-run in 0.5X TBE for 1 hour prior to loading.

DNA fragments produced by PCR (For details of primers used see Table 1) were radiolabelled with γ^{32}P. 10 pmol oligonucleotide (usually 1-2 µl of a 50 µl PCR reaction) was incubated at 37°C for 10 minutes with 1 µl T4 polynucleotide kinase (PNK) buffer (Promega), 1 µl T4 PNK, 2.5 µl γ^{32}P Redivue adenosine triphosphate (0.925 MBq, Amersham Radiochemicals) and sterile water to 10 µl. Enzyme was inactivated by heating to 90°C for 2 minutes and 90 µl sterile water added to a final volume of 100 µl.

Binding reactions were carried out by incubation of 2 µl from the above labelling reaction with 4 µl binding buffer (20 mM HEPES pH 7.3, 10 mM ammonium sulphate, 1 mM DTT, 1 mM EDTA, 0.2 % Tween20, 30 mM potassium chloride), 1 µl poly-L-lysine (1 µg/µl, Sigma), 1 µl poly-D(I-C) (1 µg/µl, Roche), 11 µl sterile water and 1 µl (30-50 pmol) recombinant protein.

Reactions were allowed to proceed at room temperature for 15-20 minutes and 5 µl loading buffer (60 % 0.25× TBE, 40 % glycerol) added. Samples were loaded onto the pre-run gel and run for 5 hours at 160V.

Gels were dried using a vacuum gel drier and were exposed to Hyperfilm MP (Amersham Pharmacia). Films were developed after a suitable period of time using a Fuji processor. For each reaction a control was performed in which there was no recombinant protein.

2.17 Phosphorylation of Rv0903 protein prior to gel retardation assay

Thirty pmol of recombinant Rv0903 protein was phosphorylated by incubation at 37°C with 0.5 µl magnesium chloride (1 M), 1 µl DTT (10 mM) and 2.5 µl acetyl phosphate (200 mM, Sigma) in a final reaction volume of 10 µl. After 30 minutes phosphorylated protein was used in binding reactions as detailed above using 1 µl sterile water in order to maintain a 20 µl final volume.

Upper primer	Sequence (5' to 3')	Lower primer	Sequence (5' to 3')	Product length (bp)
1490UP	GCACGACGAGGGCGACAG CAAGAT	1196Low	GCGGCACGGTGGCATCAGCGAGT	1938
		RAS001L	TGACCGGCGGCAACATCTTC	354
		RAS002L	GCCACGGCGCCGAAATAGTCAC	984
RAS012U	CCAGCGGATCGTCGTTGT CAGC	RAS011L	CCCCTACCCTGCGAAGCCTACTGC	347
		RAS014L	CGATGCCGGCGCTCAGACG	282
		RAS015L	CCCCGGGCGCACAGGTCAAGGTC	514
UPLeader	CCTAAAGCTTTGACCGCC GCCAACCACGACAA	BamLow	GAAGGATCCCCCCCAGGAACGCCAG CAGGTA	1487
		Bam LowShort	AAGGGGATCCGCAAAGACAAGGACA ACGCAGAAG	689
U2	TTCGAGCGGCCCCAGTCC	BamLow	See above	895
570Upper	ACCGCCAGCGTGCCGATT CCT	570Lower	CATGCTGGGTTCTCCTTAGTTGAC	570
250Upper	CCGGCAGCGAAGGTATCA AT	250Lower	CCCAGGAACGCCAGCAGGTAGCAC	250
131Upper	ATCCGATCGCCAGCAACG CCACAC	BamLow	See above	131
U2	See above	RAS004L	ACGCTGGCGGTGAAAACTG	289
570Upper	See above	RAS005L	GTGCCTGCCCCGTAACCTCAA	168
570Upper	See above	RAS003L	CCTTCGCTGCCGGTGTTGTC	376

Table 1 Primers used to generate PCR products for gel-shift experiments

2.18 Expression of OmpATb in *Mycobacterium smegmatis*

2.18.1 Overexpression of the porin gene

Primers RAS026U [5' GCG GGG ATC CAG TGG CTT CTA AGG C 3'] and RAS027L [5' CGG CAA GCT TTC CTT AGT TGA CCA C 3'] were used in a PCR reaction using High Fidelity enzyme (Roche) (60°C annealing temperature, 1 minute 45 seconds extension time, 30 cycles) to generate a 1004 bp fragment containing the *ompATb* gene from cosmid MTCY31. This was cloned into a *Bam*HI/*Hind*III digest of pMV261 (Stover *et al.*, 1991) using a Rapid DNA Ligation Kit (Roche) according to the manufacturer's instructions. The resultant plasmid pRAS10 was then electroporated (see section 2.10) into competent *Mycobacterium smegmatis*. This placed the *ompATb* gene under the control of the mycobacterial *hsp60* promoter in a replicating vector

2.18.2 High-level expression of the porin gene

An *Xba*I/*Nhe*I digest of plasmid pRAS10 produced a fragment containing the *ompATb* gene under the control of the mycobacterial *hsp60* promoter. This was ligated into an *Xba*I digest of pMV306 which had been treated at 37°C for one hour with alkaline phosphatase (1 U, Roche) and extracted from a 1 % agarose gel (Section 2.4).

pMV306 is a mycobacterial shuttle plasmid, designed by the same lab which produced pMV261 (see above), which integrates into the mycobacterial genome.

A Rapid DNA Ligation Kit (Roche) was used at a ratio of 300 ng plasmid: 900 ng insert DNA (by visualisation) according to the manufacturer's instructions to produce a plasmid, pRAS11, which was first transformed into *Escherichia coli* DH5α and, after miniprep as described in section 2.8, into *Mycobacterium smegmatis* mc^2155 (see section 2.10). This resulted in expression of the *ompATb* gene from the *hsp60* promoter at a single copy per cell.

2.18.3 Expression of the porin gene and the putative regulatory complex under a native promoter

Three μg mycobacterial cosmid MTCY31, containing the porin gene, was digested with *Sap*I (Roche) and run through a 0.8 % agarose gel. The largest fragment from this gel

(size could not be precisely determined due to lack of availability of the entire cosmid sequence) was eluted and subjected to *Sac*I digest (Roche). A 6588 bp fragment containing the porin gene *ompATb* as well as Rv0902c and Rv0903c was eluted from an agarose gel and ligated into *Sac*I digested pBluescript KS (Stratagene). The resultant plasmid, pRAS19, was transformed into *Escherichia coli* DH5α and successful transformants were selected by plating onto L-agar containing carbenicillin (200 μg/ml) and 4-bromo-3-chloro-2-indolyl-β-galactoside (X-gal, 100 μg/ml, Melford). Colonies containing pBluescript in to which the *Sac*I fragment was successfully incorporated were white on this medium whereas vector religations were blue.

Plasmid DNA was extracted from cultures of *Escherichia coli* DH5α as described in section 2.8 and 10 μl of the 30 μl miniprep digested with *Sac*I to release the 6.588 bp fragment. This was run through a 1 % agarose gel and the relevant band excised and eluted as in section 2.4, in 30 μl sterile water. The eluted fragment (25 μl) was blunt-ended by incubation at 37°C for 8 minutes with 3 μl 10× DNA polymerase buffer and 1 μl T4 DNA polymerase, followed by the addition of 3 μl dNTP mix (0.5 mM) and incubation at 37°C for 12 minutes. Reactions were then combined with 3 volumes of proprietary buffer QG and run through a gel-extraction column (Qiagen), finally eluting in 30 μl DNA ligation buffer (Roche).

Two μg Plasmid pMV306 was digested with EcoRV (Roche) by overnight incubation in buffer B (Roche) in 20 μl final volume. This was treated with 1 U alkaline phosphatase (Roche) for 1 hour prior to running through a 1 % agarose gel. The band, corresponding to linearised plasmid, was excised and eluted in 30 μl DNA dilution buffer as described in section 2.4.

Seven μl of the digested pMV306 elution and 3 μl of the blunt-ended *Sac*I fragment elution, corresponding to approximately 200 ng and 600 ng of DNA respectively, were ligated using a Rapid DNA Ligation Kit (Roche) according to the manufacturer's instructions. This resulted in a plasmid, pRAS20, which contained the porin gene, the sensor gene (Rv0902c) and the regulator gene (Rv0903c) in a mycobacterial integrating vector.

Escherichia coli DH5α was transformed and plated (as described previously) onto L-agar containing 50 μg/ml kanamycin. Successful transformants were picked and plasmid extracted from 10 ml cultures. This was then used in a transformation of *Mycobacterium smegmatis* mc^2155 as described in section 2.10.

2.18.4 Expression of the porin gene under a native promoter

An *Eco*RI digest of pRAS19 results in a 4428 bp fragment containing all of the coding sequence of the porin gene *ompATb* and continuing into the first 281 bp of the Rv0902 coding sequence. The expected result of this is a loss of function from the sensor protein and total loss of the regulator protein.

This 4428 bp fragment was digested out of 2 μg pRAS19 by overnight incubation with 2 U *Eco*RI in 20 μl total volume including 2 μl 10X buffer H. The digest was run through a 1 % agarose gel and the 4428 bp band excised as in section 2.4.

This fragment was ligated into *Eco*RI digested, CIP treated, pMV306 as described above resulting in plasmid pRAS21. This was transformed first into *Escherichia coli* DH5α and then into *Mycobacterium smegmatis* mc^2155 as above.

2.19 Construction of a suicide vector for *ompATb* knockout by homologous recombination

Plasmid pRS5 was provided by Dr Ryan Senaratne (NIMR); this had been previously used in an attempt to knock-out *ompATb* functionality which had not been successful. pRS5 is comprised of a 4.7 kb *Nsi*I fragment from *Mycobacterium tuberculosis* cosmid MTCY31 cloned into pMCS5 (Molecular Biologische Technologie). The resulting plasmid then had a deletion of the *Sac*I site by inverse PCR with modified primers, thus the *Sac*I site present internal to the *ompATb* coding region is unique in pRS5. pRS5 then had a gentamycin resistance cassette inserted at this site.

2.19.1 Construction of plasmid pRAS4

Plasmid pRAS4 was constructed by removing the gentamycin resistance cassette from pRS5 and replacing it with a kanamycin resistance cassette. Two μg pRS5 was digested overnight at 37°C in buffer A (Roche) in a 20 μl reaction with 1 μl (10 U) *Sac*I. This was

then treated with 1 U alkaline phosphatase (Roche) for 1 hour. The resultant linearised plasmid and released insert were run through a 1 % agarose gel and the band corresponding to the plasmid vector was excised (Section 2.4) and eluted in 30 µl sterile water.

A 1.28 kb kanamycin resistance cassette was amplified from plasmid pMV261 using primers P92/*Sac*I-K-UP [5 ' TTG TGA GCT CAG ACT GGG CCT TTC GTT TTA TG 3 '] and P93/*Sac*I-K-LOW [5 ' ATC TGA GCT CGT GAA GAA GGT GTT GCT GAC 3 '] (kindly provided by Dr K.G. Papavinasasundaram, NIMR). The entire 50 µl PCR reaction was digested with 2 µl (20 U) *Sac*I overnight at 37°C in buffer A (Roche).

Both the kanamycin resistance cassette and the digested pRS5 were run through a 1 % agarose gel for the purposes of quantification and a ligation reaction was performed with approximately 500 ng plasmid and 1 µg insert (Rapid DNA Ligation Kit, (Roche). The resultant plasmid, pRAS4, was then used in the construction of pRAS5 in which a *sacB* gene confers sucrose lethality on the host strain.

2.19.2 Construction of plasmid pRAS5

The *sacB* gene was obtained by digesting plasmid pKP140 (kindly provided by Dr K G Papavinasasundaram, NIMR) with *Xba*I and *Nhe*I; 10 µl pKP140 miniprep was digested with 1 µl (10 U) *Xba*I, 1 µl (10 U) *Nhe*I, 3 µl Buffer A (Roche) and 15 µl sterile water. This released a fragment containing the *sacB* gene under the control of the mycobacterial *hsp60* promoter. Plasmid pRAS4 was digested at the unique *Xba*I site by overnight incubation at 37°C in buffer A (Roche) and the product treated with alkaline phosphatase (1 U for 1 hour at 37°C) and run through a 1 % agarose gel. The *Xba*I digested plasmid was eluted in 30 µl DNA dilution buffer (Roche) and used in a ligation reaction with the *Xba*I/*Nhe*I fragment from pKP140 containing *hsp60-sacB* at a ratio of 1:3 according to manufacturer's instructions. The resultant plasmid, pRAS5, was transformed into *Escherichia coli* DH5α and extracted by midiprep from 50 ml L-broth containing carbenicillin (200 µg/ml) and kanamycin (50 µg/ml) using a Qiagen kit according to manufacturer's instructions.

Plasmid pRAS5 was treated prior to electroporation into *Mycobacterium tuberculosis* by an alkali denaturation method, this results in the plasmid losing the supercoiled structure

and becoming more linear and has been shown to enhance transformation efficiency in the mycobacteria (Hinds *et al.*, 1999). Plasmid DNA (20 μg) was denatured by addition of 0.1 volumes sodium hydroxide (2 M) and 0.1 volumes EDTA (2 mM), this was incubated at 37°C for 30 minutes and then precipitated by the addition of 0.1 volumes sodium acetate (3 M, pH 5.2) and 2.5 volumes ethanol. After 20 minutes at -70°C DNA was pelleted by centrifugation (11600 × *g*, 15 minutes) and washed once with 70 % ethanol. Finally DNA was resuspended in sterile water to a final concentration of 1 μg/μl.

Five μg plasmid pRAS5, both alkali denatured and supercoiled, was transformed into *Mycobacterium tuberculosis* H37Rv as described in section 2.10. Cells were plated out onto 7H11 agar containing 2 % sucrose and 25 μg/ml kanamycin and colonies resistant to both these selection pressures selected for DNA extraction and analysis.

Extracted DNA was subjected to PCR using primers RAS006U [GCGCGGCGACCAGCACCTG] and RAS007L [GGCAACTCCGCGGGCAACCAG] which would give a 343 bp product in the wild type and no product in the Δ*ompATb* knockout mutant as the primer binding sites are not present.

2.19.3 Construction of pOmpA-aph-rpsL tb

Concurrent to, and in collaboration with, these attempts to disrupt the function of *ompATb* a team lead by Dr Peter Sander (Institut fur Medizinische Mikrobiologie, Hannover, Germany) constructed a different suicide vector for *ompATb* knockout. This approach used an *Eco*RI fragment from the 4.7 kb *Nsi*I fragment present in pRS5 (before disruption). This *Eco*RI fragment, of 4430 bp in length, was cloned into p*trpA*-1-*rpsL*, a suicide delivery vector containing the wild type copy of the *rpsL* gene from *Mycobacterium tuberculosis*. Digestion of the plasmid with *Bsm*I/*Hpa*I and subsequent introduction of a kanamycin resistance cassette led to the removal of 420 bp of the *ompATb* coding region and its replacement with a 1268 bp kanamycin resistance cassette.

Mycobacterium tuberculosis 1424, an *rpsL* point mutated host strain that is streptomycin resistant, was transformed with the plasmid and those colonies which were streptomycin resistant and kanamycin resistant were screened by PCR and by Southern blotting to determine if homologous recombination had occurred. Southern blotting of this strain is described in section 2.28. PCR reactions used primers RAS028U [5' CCG TCC GTG GCT

GAT TGG 3'] and RAS029L [5' GTG CCT GCC CCG TAA CCT C 3'] at an annealing temperature of 60°C and an extension time of 1 minute 20 seconds. These primers gave a PCR product of 512 bp in wild type *Mycobacterium tuberculosis* 1424 and 1360 bp if the kanamycin resistance cassette has been introduced.

2.20 Construction of a suicide vector for Rv0903c knockout by homologous recombination

A vector was constructed which took advantage of a technique for homologous recombination different to both those detailed above.

Plasmid p2NIL was kindly provided by Dr Tanya Parish (London School of Hygiene and Tropical Medicine) and was used as a manipulation vector for the introduction of two regions of homology to the Rv903c gene.

2.20.1 Construction of plasmid pRAS14

A region 3' to the coding sequence of the gene was amplified by PCR using primers RAS032U [5' TAT TAA GCT TCG TCA AGT CGG GCA CCA AGA GA 3'] and RAS033L [5' TTA CCT CGA GGA GTC GGA TTC GTG CTG CGT ATG C 3'] the PCR program was as follows:

95°C	7 minutes	x1
94°C	30 seconds	x10
64°C	30seconds	
72°C	2 minutes	
94°C	30 seconds	x20
64°C	30 seconds	
72°C	2 minutes (+10 seconds per cycle)	
72°C	7 minutes	x1
4°C	hold	

The 1779 bp product of this reaction was digested with *Hind*III and *Xho*I (restriction sites underlined in primers) at 37°C, overnight, in a 20 µl reaction with buffer B (Roche) and run through a gel extraction column (Qiagen) to clean up. Finally DNA was eluted in 30 µl DNA dilution buffer (Roche).

Plasmid p2NIL was digested with 20 U *Sal*I by overnight incubation at 37°C in buffer H (Roche) in 25 µl total volume. The reaction was run through a 1 % agarose gel and the band (linearised *Sal*I digested p2NIL) excised and eluted in 30 µl sterile water. This was then digested with 10 U *Hind*III in exactly the same way but using buffer B (Roche), and was extracted from a 1 % agarose gel.

The *Hind*III/*Xho*I digested region of homology was cloned into the *Hind*III/*Sal*I digest of p2NIL using a Rapid DNA Ligation Kit (Roche) according to manufacturer's instructions. This produced plasmid pRAS14.

pRAS14 was transformed into *Escherichia coli* DH5α, plated onto L-agar with 50 µg/ml kanamycin and extracted from 10 ml cultures by miniprep as detailed in section 2.8.

2.20.2 Construction of plasmid pRAS17

A region 5' to the coding sequence was amplified using primers RAS044U [5' TAC TGG TAC CTG CCG CCC ATT TTG CCT GAT TAC 3'] and RAS046L [5' CAT TGG TAC CTC GAC GTC GGA AAC TGG AAG AAG 3'] (both of these primers have *Kpn*I sites underlined). A 2056 bp PCR product was generated using an annealing temperature of 61°C and an extension time of 2 minutes 30 seconds for 25 cycles. This product was digested with 10 U *Kpn*I (15 µl PCR product, 2 µl buffer L, 2 µl BSA (final concentration of 100 µg/ml) and 1 µl *Kpn*I) at 37°C, overnight, and was extracted through a Qiagen gel extraction column, eluting in 30 µl DNA dilution buffer (Roche).

2 µg Plasmid pRAS14 was digested with 10 U *Kpn*I by overnight incubation at 37°C in Buffer L with 100 µg/ml final concentration of BSA, and was treated with 1 U alkaline phosphatase at 37°C for 1 hour. This reaction was run through a 1 % agarose gel and extracted using a Qiagen kit (Section 2.4), finally eluting in 30 µl DNA dilution buffer.

The 5' region of homology, flanked by *Kpn*I sites, was cloned into pRAS14 using a Rapid DNA Ligation Kit and the resultant plasmid, pRAS17, transformed into *Escherichia coli* DH5α. Transformed cells were grown in 10 ml volumes of L-broth with 50 µg/ml kanamycin and plasmid DNA extracted as in section 2.8. Orientation of the *Kpn*I fragment was determined by *Bsm*I, *Eco*RI and *Pst*I digestion, plasmids with the fragment in the 'forward' orientation (the same direction as the 3' fragment already cloned into the parent vector pRAS14) were selected.

Plasmid pRAS17 thus contains both a 5' region of homology (2038 bp) and a 3' region of homology (1769 bp) to the Rv903c gene. The majority of the coding sequence of the gene has been replaced by p2NIL plasmid DNA from between the *Sal*I and *Kpn*I sites in the multicloning sequence of this vector. Only the most 5' 8 base pairs and the most 3' 28 base pairs remain (674 base pairs of the 710 bp gene have been removed and replaced with 312 base pairs of spurious plasmid DNA).

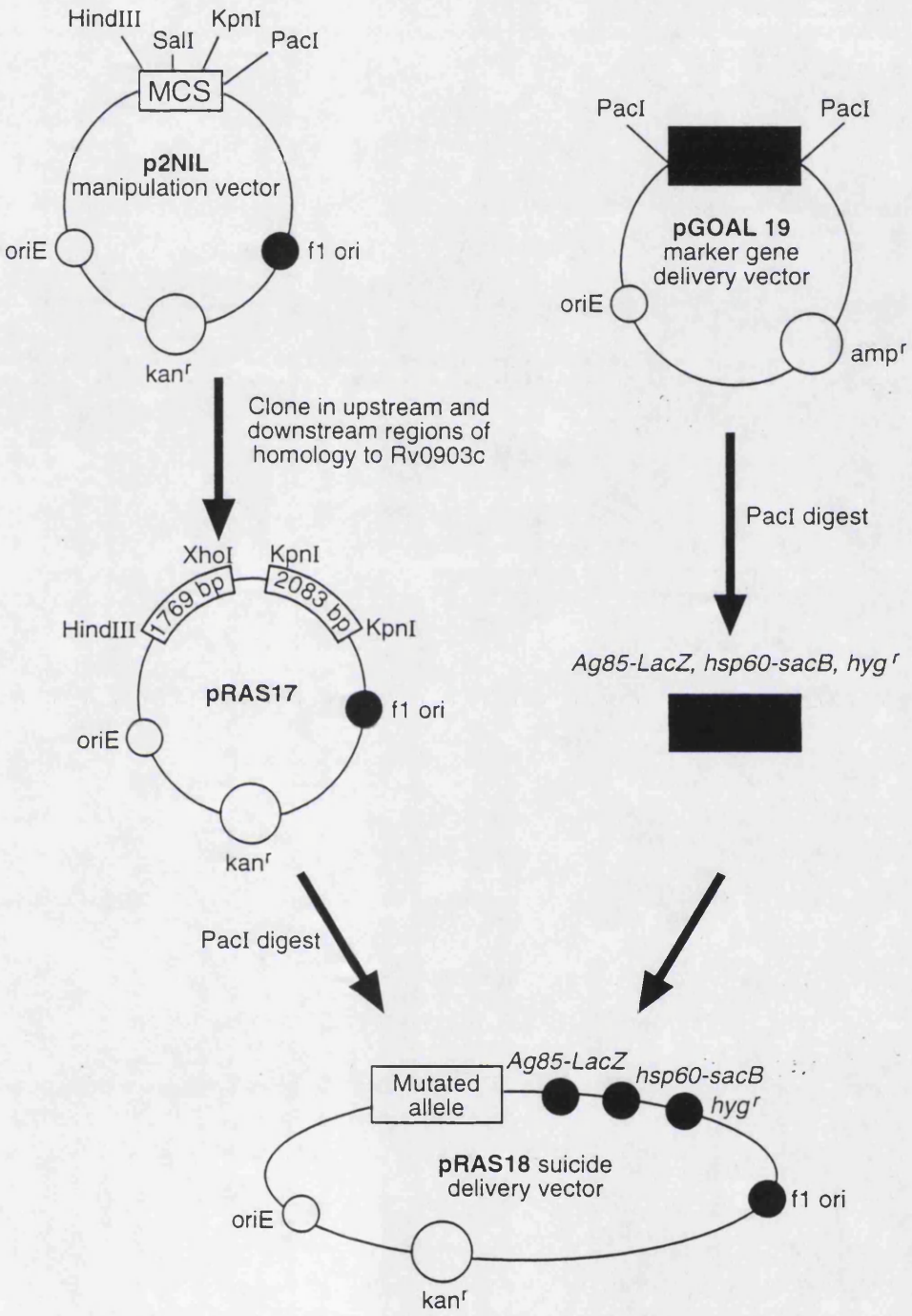


Figure 5 Construction schematic of plasmid pRAS18 – the vector used to attempt homologous recombination to knock out gene Rv0903c

2.20.3 Construction of plasmid pRAS18

Plasmid p2NIL, and hence plasmids pRAS14 and pRAS17, have a unique *Pac*I site which serves to receive a *Pac*I fragment from plasmid pGOAL19 (also kindly provided by Dr Tanya Parish). This *Pac*I fragment contains the *sacB* gene (mycobacterial *hsp60* promoter), the *lacZ* gene (mycobacterial antigen 85 promoter) and a hygromycin resistance cassette.

Three μg pGOAL19 was digested with *Pac*I (New England Biolabs) by overnight incubation with proprietary buffer 1 (New England Biolabs) and BSA in a 20 μl reaction at 37°C. The 7.9 kb fragment was eluted from a 1 % agarose gel using a Qiagen kit as before. Three μg pRAS17 was digested in exactly the same way and the 8.5 kb fragment eluted from a 1 % agarose gel.

*Pac*I digested pRAS17 and the 7.9 kb fragment from pGOAL19 were ligated together using a Rapid DNA Ligation Kit (Roche) according to the manufacturers instructions and the resultant plasmid, pRAS18, transformed into *Escherichia coli* DH5α. Cells were plated onto L-agar with 50 μg/ml kanamycin, and were also confirmed by plating onto 5 % sucrose to check functionality of the *sacB* gene.

Plasmid pRAS18 thus has both the upstream and downstream regions of homology to Rv0903c and also confers kanamycin resistance, hygromycin resistance, sucrose sensitivity and blue/white selection on host mycobacterial strains. It is therefore possible to select for both single crossover events (blue, kanamycin resistant, sucrose sensitive, hygromycin resistant colonies) and double crossover events (white, kanamycin sensitive, sucrose resistant colonies).

A schematic of the construction of plasmid pRAS18 can be seen in figure 5.

Mycobacterium tuberculosis H37Rv was transformed with UV treated (UV Stratalinker set to deliver 100 Joules, Stratagene) or supercoiled plasmid DNA at 5 μg per transformation as in section 2.10.

A two-step procedure was adopted whereby firstly single crossover transformants of *Mycobacterium tuberculosis* H37Rv were selected by plating the transformation mix on

7H11 agar with 25 μg/ml kanamycin (Sigma) and 100 μg/ml X-gal (Melford). Twelve blue, kanamycin resistant colonies were selected and plated onto 7H11 agar without antibiotic selection to allow a second crossover to occur or alternatively inoculated into 5 ml Dubos broth.

Around half of a 7H11 agar plate of colonies with a single crossover phenotype were suspended in 1 ml sterile PBS (137 mM NaCl, 3 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4) with 0.05 % Tween20 (Sigma). One hundred μl of the suspension was then combined with 900 μl sterile PBS with Tween20 (0.05 %) down to a 10^{-4} dilution. Each dilution from this series was plated at a volume of 100 μl on to 7H11 agar containing sucrose (2 %) and X-gal (100 μg/ml, Melford). Colonies which were sucrose resistant and white were selected and analysed by PCR using primers RAS063U [5' GCA TAC GCA GCA CGA ATC C 3'] and RAS064L [5' GAG TAC CCC GAC GCC GCA GAC GAG 3'] after having extracted DNA by heating suspended colonies to 100°C for 10 minutes in TE buffer (10 mM tris pH 8, 1 mM EDTA). Five μl of the heat-killed *Mycobacterium tuberculosis* was used in each PCR reaction.

Alternatively 50 μl of the broth was diluted into 450 μl sterile PBS with Tween, down to a 10^{-3} dilution, and 50 μl of each dilution was plated onto 7H11 agar with X-gal and sucrose. White colonies were selected and DNA was extracted using Instagene Matrix (Bio-Rad) as in section 2.23.2. Primers RAS063U [GCA TAC GCA GCA CGA ATC C] and RAS064L [GAG TAC CCC GAC GCC GCA GAC GAG] were used to PCR amplify a product from the potential knockouts at an annealing temperature of 59°C and an extension time of 1 minute. The product was compared with the wild type; if there was recombination the length of the product is expected to be 631 bp, whereas the wild type is 992 bp.

2.21 Construction of *lacZ* reporter vectors

Primers RAS009U [5' CTA CTC TAG ACC AGG ATG TTC ACC GTG TTG TTT G 3'] and RAS010L [5' GTC TAA GCT TAA GCC ACC TAC GCC ACT CCC TCA G 3'] were used to amplify a 2.2 kb region upstream of the *ompATb* gene, including the first 7 bases of the *ompATb* coding sequence. These primers are designed such that the PCR fragment has a 5' *Xba*I restriction site (underlined in RAS009U) and a 3' *Hind*III restriction site (underlined in RAS010L). Two μg of plasmid pEJ414, a

mycobacterial integrating vector carrying a promoterless *lacZ* gene (kindly provided by Dr Elaine Davis, NIMR), was digested with *Xba*I and *Hind*III. This was done by overnight incubation at 37°C in a 20 µl reaction containing 2 µl 10× buffer B (Roche), 1 µl (10 U) *Xba*I and 1 µl (10 U) *Hind*III (Roche). Digested plasmid was treated with 1 µl (1 U) calf intestinal phosphatase (Roche) for 1 hour and run through a 1 % agarose gel. The linearised plasmid band was extracted as described in section 2.4.

Half a µg of the 2.2 kb PCR product was also digested with *Xba*I/*Hind*III but was not alkaline phosphatase treated. This was run through a 1 % agarose gel and extracted in the same way as the plasmid.

Plasmid and insert were ligated using a Rapid DNA Ligation Kit (Roche); 50 ng plasmid DNA and 150 ng insert DNA (as determined by visualisation in an agarose gel with ethidium bromide) were mixed in a 10 µl reaction, the remainder being made up with DNA dilution buffer. DNA ligation buffer (10 µl) and 1 µl Rapid DNA ligase were added and the reaction incubated at 37°C for one hour. *Escherichia coli* DH5α was transformed as described in section 2.7 and was plated on L-agar with 50 µg/ml kanamycin.

Primers RAS023U [5' CTA CTC TAG AGG CAG GGG TGT ATG AGG AAA ATG G 3'], RAS024U [5' CTA CTC TAG ATG ACC GCC GCC AAC CAC GAC AA 3'] and RAS025U [5' CTA CTC TAG ATC GGC CAC CAG GTC ACG GAT GT 3'] were used with primer RAS010L to generate plasmids pRAS6, 7 and 8 in exactly the same way as detailed for pRAS3. These plasmids harbour regions upstream of *ompATb* of length 684 bp, 452 bp and 178 bp respectively.

2.22 β-galactosidase assays for promoter activity

Mycobacterium smegmatis cultures (20 ml) carrying pRAS3, pRAS6, pRAS7 or pRAS8 in Dubos broth with 25 µg/ml kanamycin were harvested at an OD_{600} of 0.6 by centrifugation at 4080 × *g* for 15 minutes. Cell pellets were resuspended in 1 ml Z-buffer (60 mM Na_2HPO_4, 40 mM $NaH_2PO_4 \cdot 2H_2O$, 10 mM KCl, 1 mM $MgSO_4 \cdot 7H_2O$) and were transferred to Ribolyser 'blue' tubes (Hybaid) and lysed for 30 seconds at speed 6.5 in a Ribolyser (Hybaid). Tubes were spun at 11600 × *g* for 5 minutes and supernatants transferred to clean microfuge tubes.

Bovine serum albumin standards of 0.4, 0.6, 0.8 and 1 mg/ml were prepared by dilution in Z-buffer and were used in quadruplicate in the second and third rows of a 96 well flat-bottomed microtitre plate at a volume of 15 µl.

Samples of the cell lysate were assayed for protein concentration using the BCA assay kit (Pierce), in duplicate, by adding 200 µl BCA working reagent (50A:1B) to 15 µl samples. Absorbance at 550 nm was measured after a 30 minute incubation at 37°C using an EL312 ELISA plate reader (Bio-tek Kontron) and protein concentrations determined using the Kineti-Calc software (Bio-tek Kontron).

2-Mercaptoethanol (1.5 µl) was added to the remainder of each of the cell lysates after sampling for the BCA assay, as this interferes with the assay reagents.

Samples were respun for 2 minutes at 11600 × *g* and 150 µl cell lysate supernatant taken and added to 350 µl Z-buffer containing 2.7 % 2-mercaptoethanol. Negative controls were also prepared, these were simply 500 µl Z-buffer+mercaptoethanol. Samples were equilibrated at 28°C for 5 minutes and 100 µl *o*-nitrophenyl β-D-galactopyranoside (ONPG, 4 mg/ml) added once to each sample at intervals of 10 seconds. Reactions were timed precisely until a distinct yellow colour had developed and were stopped by the addition of 250 µl Na_2CO_3 (1 M) followed by mixing.

Absorbency of the samples and the blank was measured in duplicate and a mean absorbance at 405 nm determined using the Kineti-Calc software (Bio-tek Kontron).

Units of ß-galactosidase per mg protein were calculated using the following formula (Miller, 1972):

$$\text{Units} = \frac{380 \times A_{405}}{\text{time (min)} \times \text{volume assayed (ml)} \times \text{protein concentration (mg/ml)}}$$

2.23 Extraction of Mycobacterial Nucleic Acids

2.23.1 DNA extraction from *Mycobacterium tuberculosis* grown on 7H11 agar plates

Half a plate of *Mycobacterium tuberculosis* culture grown on Middlebrook 7H11 agar for roughly 3 weeks was harvested with a sterile plastic loop into 300 µl sterile TE (10 mM tris pH 8, 1 mM EDTA). This was heat-killed by incubation at 80°C for 1 hour.

Lysosyme (Sigma) and Lipase (Sigma) were added to a final concentration of 2 mg/ml alongside 5 µl RNAse (Roche). This was incubated for 1 hour at 37°C in a water bath. The extracts were then frozen for 10 minutes in an ethanol/dry ice bath and heated to 75°C prior to the addition of ProteinaseK (Roche) to a final concentration of 500 µg/ml. SDS was then added to a final concentration of 0.05 %.

Extracts were incubated at 50°C for 1 hour and were phenol:chloroform (Sigma) extracted twice then chloroform extracted once. Chromosomal DNA was precipitated by addition of 1/50th volume 5 M NaCl and 2.5 volumes room temperature ethanol. Pellets were spun down (15 minutes at 11600 × *g*) and washed once with 70 % ethanol before re-suspension in 300 µl TE. Concentration was measured by A_{260} (ThermoUnicam UV2 spectrophotometer).

2.23.2 DNA extraction using Instagene Matrix

For isolation of mycobacterial DNA from individual colonies Instagene Matrix was used (Bio-Rad). A sterile loopful of bacteria was resuspended in 1 ml sterile water in a microfuge tube. Cells were pelleted by centrifugation at 11600 × *g* for 1 minute and the supernatant removed. Instagene Matrix (200 µl) was added to the pellet and the suspension incubated for 30 minutes at 56°C.

After incubation cells were vortexed for 10 seconds and incubated at 100°C for 8 minutes, followed by a further vortex and a spin at 11600 × *g* for 2 minutes. Twenty µl of the supernatant was used in a PCR reaction. If large amounts of bacteria were added initially, it was sometimes necessary to dilute the supernatant up to 50-fold in order to obtain a PCR product.

2.23.3 DNA Extraction from Mycobacterial broth cultures

Cultures (100 ml) were grown to an OD_{600} of between 0.5 and 1, and 0.1 volumes 2 M glycine added 1 generation time before harvest (1 h for fast growers, 12-24 h for slow growers). Cultures were harvested by centrifugation at 4080 × *g* in a Sorvall GSA rotor for 20 minutes; cell pellets were resuspended in 2 ml TE and divided into 500 ml aliquots. Cell suspensions were then processed in an identical way to bacteria harvested from plates.

2.23.4 Isolation of Mycobacterial RNA from BCG

Cultures of BCG (100 ml) were grown in conical flasks in Tween-glutamate broth (See appendix); flasks were shaken at 250 r.p.m. at 37°C until the OD_{600} of the culture was 0.6.

Cultures were spun down at 10444 × *g* for 20 minutes in a Sorvall GSA rotor and cell pellets re-suspended in 1 ml proprietary buffer RLT (Qiagen) containing 10 µl/ml 2-mercaptoethanol.

Suspended bacteria were added to a Ribolyser 'blue' tube containing matrix (Hybaid) and were lysed in a Ribolyser (Hybaid) once at speed 6 for 20 seconds. Lysates were then cooled on ice and tubes centrifuged for 10 minutes at 11600 × *g* to pellet debris. Supernatant (700 µl) was taken and an equal volume of 70 % ethanol, made with diethyl pyrocarbonate (DEPC, 1:1000 in sterile water, autoclaved, Sigma) treated water, was added.

Seven hundred µl of the suspension was added to an RNeasy column (Qiagen) which was spun for 15 seconds at 11600 × *g*, the remaining 700 µl was then loaded and the column spun again. Columns were washed with 700 µl buffer RW1 and washed twice with 500 µl of buffer RPE according to the protocol supplied with the kit.

RNA was then eluted by adding 50 µl of RNase-free water to the dry column and spinning at 11600 × *g* for 2 minutes, 1 µl of RNasin (Promega) was added to the eluent and the RNA stored at -80°C.

2.23.5 Isolation of RNA from *Mycobacterium tuberculosis*

This was performed in exactly the same way as for BCG with the following modifications: cultures were grown in 100 ml Dubos broth with glycerol and albumin (See appendix) in a rolling incubator bottle at 37°C, with an approximate rate of 2 r.p.m., until the OD_{600} of the culture was approximately equal to, but no more than, 1. Culture volumes of up to 200 ml were obtained for experiments that were performed over a time course on a single culture by mixing two identical 100 ml cultures immediately prior to taking the first 50 ml sample. In this way it was possible to take up to four 50 ml samples from a single culture without exceeding the capacity of a roller incubator bottle.

2.24 Removal of DNA contamination from mycobacterial RNA

To each 50 µl RNA sample was added: 7.5 µl magnesium sulphate (0.1 M), 5 µl sodium acetate (3 M, pH 5.2), 2 µl RNasin (80 U, Promega), 2 µl RNase-free DNase RQ1 (2 µl, Promega) and DEPC water up to 150 µl (84.5 µl).

Reactions were incubated for 1 hour and were added to an RNeasy column (Qiagen) after the addition of 350 µl buffer RLT (Qiagen) containing 10 µl/ml 2-mercaptoethanol and 200 µl ethanol. Columns were spun for 15 seconds at 11600 × *g*. Buffer RPE (500 µl, Qiagen) was then added and the column spun once more for 15 seconds at 11600 × *g*. A further 500 µl buffer RPE was added and the column spun for 2 minutes at 11600 × *g*.

RNA was eluted in 50 µl RNase-free water by spinning for 2 minutes at 11600 × *g*.

This procedure needed to be performed at least twice, with the eluent checked for DNA contamination by a PCR reaction, which would amplify a product from DNA but not from RNA.

2.25 Synthesis of cDNA by reverse transcription

Between 0.5µg and 1µg of RNA, as measured by absorbance at 260 nm, was reverse transcribed into cDNA using the following reaction: 2.5 µl random hexamer primer mix (1 µg/ml, Roche) was added to 1 µg RNA and DEPC water added to a final volume of 10 µl. Primer annealing was carried out by heating to 65°C for 10 minutes and allowing to cool to room temperature for a further 10 minutes.

To each 10 µl reaction was added 2 µl 5× first strand buffer (Life Technologies), 1 µl acetylated BSA (1 mg/ml, Promega), 2 µl dNTP mix (2.5 mM, made in DEPC water, Amersham Pharmacia), 2.5 µl dithiothreitol (DTT, 0.1 M, Life-Technologies) and 1 µl RNasin (Promega).

Either 1.5 µl (300 U) of SuperscriptII reverse transcriptase (Life Technologies) was added, or 1.5 µl DEPC water for negative controls. The final 20 µl reaction was incubated for 1 hour at 37°C then the enzyme inactivated by heating to 70°C for 15 minutes. cDNA was stored at -70°C.

2.26 NaCl shock of BCG cultures to investigate *ompATb* regulation

At an OD_{600} of 0.6, sterile 5 M NaCl was added to 50ml cultures of BCG to give a final concentration of 0.2 M, 0.4 M or 0.6 M. At a time after NaCl addition of 1 hour, 2 hours or 24 hours RNA was harvested from 50 ml cultures as described above. RNA was cleaned of DNA contamination and analysed using quantitative real-time RT-PCR as described in section 2.27.

2.27 Real-time PCR to demonstrate regulation of the porin gene

An ABI-Prism sequence detector 7700 was used to monitor real-time progression of PCR amplification from the porin gene and from a 'normalising' gene, *gnd.*

96-Well plates were loaded with 25 µl PCR reactions containing TaqMan universal master mix (Perkin Elmer), primers and probe. Primers and probe for *gnd* amplification were present in the first 48 wells of the plate, those for *ompATb* amplification were present in the last 48 wells of the plate. Optimisation of the primers was necessary to determine which amount gave the most detectable result; in the case of *ompATb* it was shown that a 300 nM forward primer:300 nM reverse primer ratio was optimal, whereas for *gnd* 50 nM forward primer:50 nM reverse primer was optimal. Probe concentration was always 100 nM and probes were labelled with FAM dye (Perkin Elmer). For details of primers and probes see Table 2.

In the case of *ompATb* primers used were:

Forward primer	5' ATG TGC CGA CCT GCA ATC A 3'
Reverse primer	5' ATT TCA TAG TCG GCT GGG ATC A 3'
Probe	5' TCG ACC CAT CGC GTT TGG CAA 3'

For *gnd* primers were:

Forward primer	5' GTC CAC AAC GGC ATC GAG TA 3'
Reverse primer	5' GCT GTC CAG ATC GCC ATT G 3'
Probe	5' TCC GAC ATG CAG CTC ATC GGT GA 3'

Table 2 Details of primers and probes used for TaqMan experiment

Standard curves were constructed by the amplification of known quantities of genomic DNA for each run of the sequence detection machine, these quantities were 1000 ng, 500 ng, 200 ng, 100 ng, 50 ng, 20 ng and 10 ng per reaction. The standard curves relate an average cycle threshold (the time at which fluorescence from the probe becomes detectable) to the quantity of DNA in the reaction; this allows interpolation of the results for the samples into a real value of cDNA present in the reaction. A ratio of *ompATb* to the normalising gene *gnd* (the level of which remains constant) could thus be obtained.

Results were analysed using ABI Prism Sequence Detection Systems V1.6.3 software (Perkin Elmer) and Excel (Microsoft).

2.28 Southern blotting of mycobacterial DNA

The validity of the Δ*ompATb* knockout strain was checked by Southern blotting. Three μg DNA was digested with *Eco*RI, *Nsi*I (Roche) or *Sty*I (New England Biolabs) in a 20 μl reaction containing either buffer B (Roche) (in the case of *Eco*RI/*Nsi*I) or NE Buffer 3 (New England Biolabs) + Bovine Serum Albumin (100 μg/ml final concentration) for *Sty*I.

Digestion reactions were incubated overnight at 37°C.

Agarose gels (0.8 %) containing 0.8 μg/ml ethidium bromide were prepared and the digests run at 90V until the first dye front approached the end of the gel. Appropriate DNA size markers were run next to the reactions and photographs were taken at a wavelength of 302 nm prior to vacuum blotting.

A VacuGene XL apparatus (Amersham) was assembled and the gel placed onto a sheet of Hybond N^+ membrane (Amersham) on a sheet of Whatman 3 MM filter paper. Both filter paper and blotting membrane were pre-wetted with 2× SSC.

A vacuum of 50 mbar was applied to the gel and 0.25 M HCl added to the gel in order to depurinate the nucleic acid. After 30-45 minutes, surplus HCl was removed and 0.4 M NaOH added.

After 2.5 hours the membrane was removed, rinsed in 2× SSC and allowed to dry completely. The membrane was then cross-linked in a UV Stratalinker (Stratagene) at 'optimal crosslink' setting.

A 132 bp probe was prepared by PCR using primers RAS030U [5' ATC CGA TCG CCA GCA ACG CCA CAC 3'] and RAS031L [5' CCC CCA GGA ACG CCA GCA GGT AGC 3'] at an annealing temperature of 61°C and an extension time of 45 seconds for 25 cycles. This was labelled using a Megaprime kit (Amersham Pharmacia) according to the manufacturer's instructions, using α^{32}P dCTP. Labelled probe was denatured at 95°C for 10 minutes and rapidly cooled on ice before addition to membrane which had been pre-hybridised for 30 minutes in Rapid-hyb buffer (Amersham) at 65°C. Care was taken to add the probe to the buffer and not directly onto the membrane. Hybridisation was carried out in a Hybaid oven at 65°C overnight.

Membranes were washed in 2× SSC, 0.1 % SDS for 5 minutes at room temperature, 1× SSC 0.1 % SDS for 15 minutes at 65°C and finally 0.1× SSC, 0.1 % SDS for 15 minutes at 65°C.

Membranes were wrapped in Saran film and exposed to Hyperfilm MP (Amersham). Films were processed in a Fuji developer.

2.29 Analysis of the growth of a porin-deficient strain of *Mycobacterium tuberculosis*

Mycobacterium tuberculosis 1424 Δ*ompATb*, having been characterised by PCR (Section 2.19) and by Southern blot (Section 2.28) was grown in Dubos media (100 ml) in roller incubator flasks (2 r.p.m.) and OD_{600} measured every 4, 6 or 8 hours until growth curves levelled off, indicating stationary phase had been reached. Alternatively raffinose (0.1 M final concentration, Sigma), sucrose (0.2M final concentration, Merck) or NaCl (0.2 M final concentration) was added prior to inoculation of the medium with bacteria in order to investigate whether osmotic shock had any effect on the growth rate of the porin knockout strain. Growth of the mutant strain was compared with *Mycobacterium tuberculosis* 1424 wild type.

2.30 Microarray hybridisation

2.30.1 Labelling of genomic DNA

Genomic DNA was labelled with either Cy3 or Cy5 dye coupled to dUTP (Amersham Pharmacia). DNA (2 μg) was added to 3 μl random hexamer primers (1 μg/ml, Roche) and sterile water to 41.5 μl. This was then heated to 95°C for 5 minutes and rapidly cooled on ice. Reactions were briefly centrifuged and to each was added: 5 μl 10X Klenow polymerase buffer (Promega), 1 μl dNTP mix (5 mM dA/G/CTP, 2 mM dTTP), 1.5 μl (1.5 nmol) Cy3 or Cy5 dUTP (Amersham Pharmacia) and 1 μl (5 U) Klenow polymerase (Promega).

Reactions were incubated at 37°C for 90 minutes then labelled DNA was precipitated by combining Cy3 and Cy5 samples and adding 0.1 volumes sodium acetate (3 M, pH 5.2, Sigma) and 0.8 volumes propan-2-ol. Precipitation was allowed to proceed in the dark at room temperature for 30 minutes and reactions were spun down at 11600 × *g* for 20 minutes. DNA pellets were washed once with 70 % ethanol and air dried for 10 minutes. Finally DNA was resuspended in 10.5 μl sterile water.

2.30.2 Labelling of RNA

RNA was labelled with either Cy3 or Cy5 dye by incorporation of labelled dUTP in a reverse transcription reaction.

RNA (5 μg) was added to 6 μl (6 μg) random hexamer primers (Roche) and DEPC water to 11 μl final volume. Random priming was carried out by heating to 95°C for 5 minutes and rapid cooling on ice.

A reverse transcription of the randomly primed DNA was performed by adding 5 μl first strand buffer (Life Technologies), 2.5 μl 0.1 M DTT, 2.3 μl dNTP mix (5 mM dA/G/CTP, 2 mM dTTP), 1.7 μl (1.7 nmol) Cy3 or Cy5 dUTP (Amersham Pharmacia) and 2.5 μl (500 U) Superscript II (Life Technologies). Reactions were incubated for 10 minutes at 25°C then 90 minutes at 42°C and cDNA precipitated by mixing the Cy3/Cy5 labelled samples, adding 0.1 volumes sodium acetate (3 M, pH 5.2, Sigma) and 0.8 volumes propan-2-ol. Precipitation was allowed to occur by incubating at room temperature for 30 minutes and cDNA was pelleted by centrifugation at 11600 × *g* for 20 minutes. cDNA pellets were washed once with 70 % ethanol and air dried for 10 minutes before resuspension in 10.5 μl sterile water.

2.30.3 Hybridisation of slides

Dr R Buxton (NIMR) in collaboration with Dr P Butcher (St George's Hospital Medical School) kindly provided microarray slides representing 95 % of the genes in the *Mycobacterium tuberculosis* genome. These were pre-hybridised by incubation at 60°C in 3.5× SSC, 0.1 % SDS, 10 mg/ml bovine serum albumin (Sigma) for 20 minutes, and were washed in water for 2 minutes and propan-2-ol for 2 minutes before being dried by centrifugation at 433 × *g* in a Sorvall HS4 rotor for 5 minutes.

Hybridisation mix was prepared by adding 3.2 μl 20× SSC and 2.3 μl 2 % SDS to 10.5 μl labelled DNA or cDNA (prepared as described in section 2.30.1, 2.30.2); this was heated to 95°C for 2 minutes and allowed to cool briefly before being added to the array. A coverslip was lowered gently onto the microarray/hybridisation solution and the slide placed in a hybridisation cassette in a 60°C water bath. The hybridisation was carried out overnight or for a period of 16-20 hours.

Slides were washed once for two minutes in a 1× SSC, 0.05 % SDS solution, in which the coverslips were carefully removed. Two further washes of two minutes each in 0.06 % SSC were performed prior to drying the slide by centrifugation at 433 × *g* for 5 minutes.

Slides were scanned in a GenePix 4000A microarray scanner (Axon Instruments) and analysed using GenePix software V3.0 (Axon Instruments).

3. Results

3.1 PCR probing of Mycobacterial species for *ompATb*

The aim of this experiment was to identify mycobacterial species that were sufficiently similar at the genetic level to give a PCR product from the porin gene. Primers U1 and L1 (see section 2.13) gave a PCR product of *circa* 1050 bp in 6 of the 12 samples of genomic DNA from the mycobacterial species tested; typical results of these experiments are presented in figure 6 and a table of results overleaf (table 3). It can be seen that a clear band corresponding to the amplified *ompATb* gene is present in *Mycobacterium tuberculosis, M. microti, M. avium, M. intracellulare, M. bovis* and *M. bovis* BCG. No product is amplified from *M. fortuitum*, *M. chelonae*, *M. smegmatis*, *M. kansasii*, *M. xenopi* and *M. marinum*; this does not necessarily mean that the gene, or a homologue of it, is not present, as the PCR based method is very specific (this is discussed later in section 4.1)

PCR amplification of the putative regulatory gene, Rv0903c, (discussed in more detail later) using primers NdeUpper and BamH1Lower (see section 2.15.1) to amplify a fragment of 865 bp in length demonstrated that the same species that have *ompATb* also have Rv0903c. In no case was it found that one of the two genes occurred independently of the other. Results for the PCR based screening for Rv0903c are shown in figure 7.

It can be seen that there is a distinction between the fast-growing and the slow-growing mycobacterial strains; the porin gene and the regulator gene are found only in slow-growing members of the genus. They are not indicative of slow growth however, as *Mycobacterium kansasii*, a slow-grower, has a PCR product from neither the porin gene nor the regulator gene. The presence of the porin gene is not therefore a universal property of slow growing mycobacteria. It is interesting to note that *Mycobacterium avium* and *Mycobacterium intracellulare*, which are closely related to each other but not to the *Mycobacterium tuberculosis* complex, also have the porin and the regulator gene.

In no case has the porin gene been identified by this PCR based method in a fast-growing member of the genus.

	ompATb	Rv0903
Mycobacterium tuberculosis	+	+
M. microti	+	+
M. avium	+	+
M. intracellulare	+	+
M. bovis	+	+
M. bovis BCG	+	+
M. fortuitum	-	-
M. chelonae	-	-
M. smegmatis	-	-
M. kansasii	-	-
M. marinum	-	-
M. xenopi	-	-

Table 3 Results of a PCR-based screen for the porin gene and the regulator gene Rv0903c

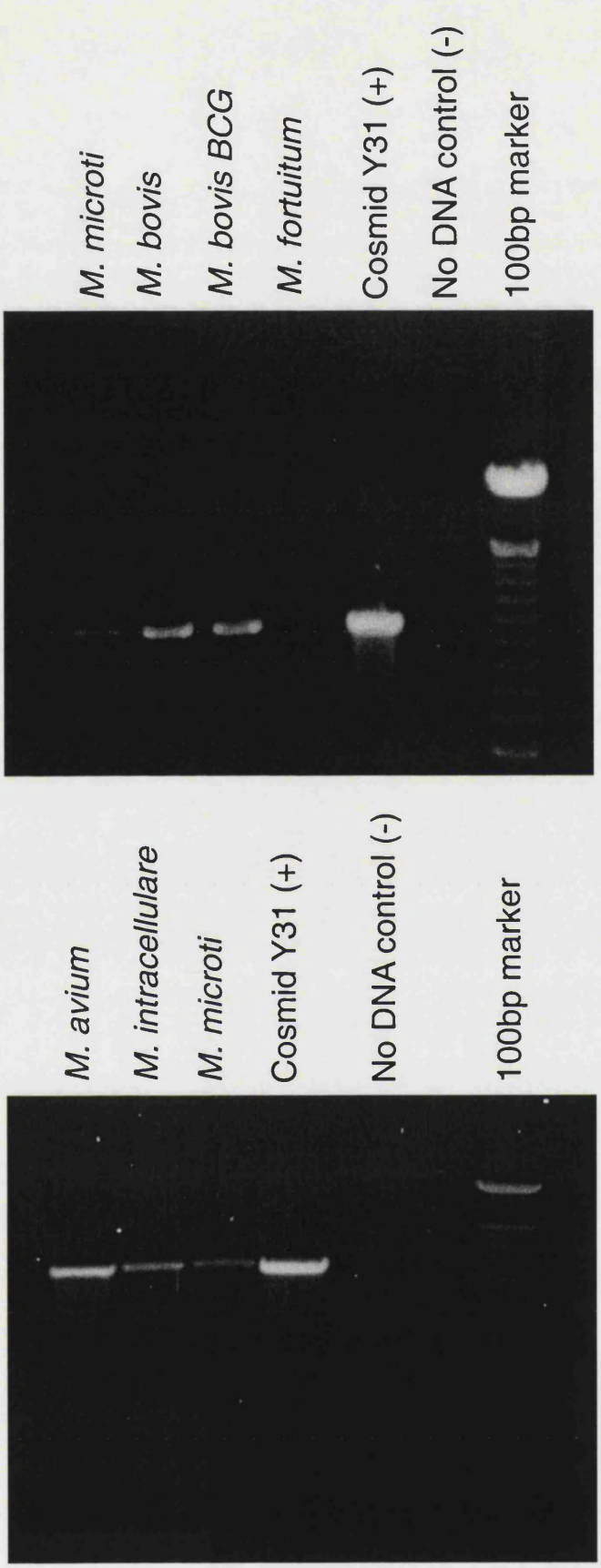


Figure 6 PCR based detection of *ompATb* gene amongst different mycobacterial species

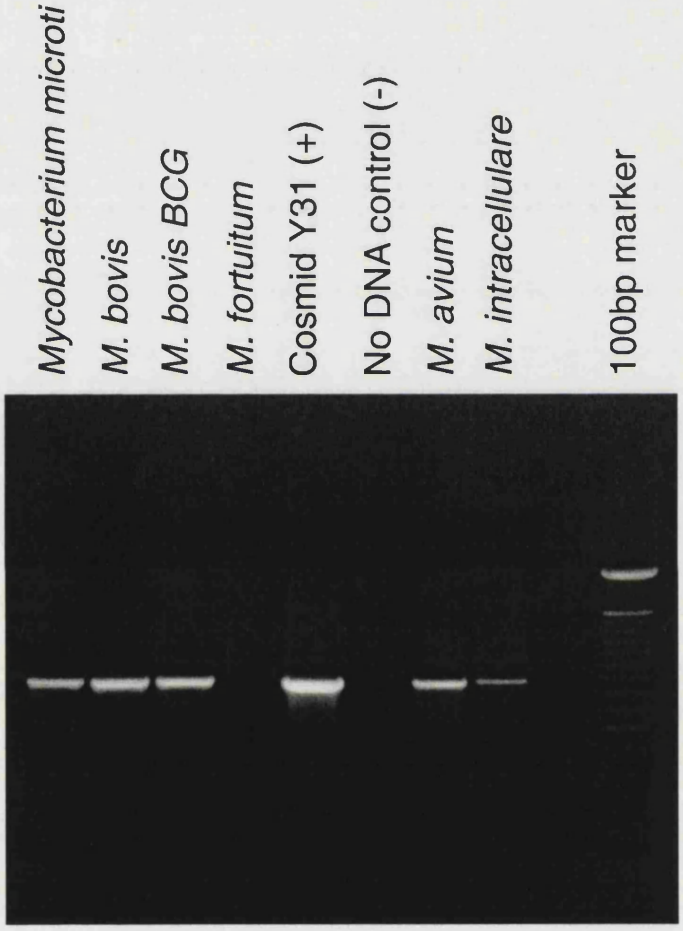


Figure 7 PCR based detection of Rv0903c gene amongst different mycobacterial species

3.2 Western blot analysis of mycobacterial cell lysates

In order to identify OmpATb protein in total cell lysates of mycobacterial cells immunoblotting with an antibody reactive to OmpATb was performed. Western blotting of cell lysates of mycobacterial cultures proved to be difficult. The amount of porin present in the cells is estimated to be very low (as little as 0.02 % - 2 %) (Senaratne *et al.*, 1998) so practical problems become apparent when loading lysate (of which the porin protein is a small component) into wells that have a maximum volume of 25 µl. Nevertheless it was possible to obtain some results, as the anti-OmpATb antibodies have a very high affinity to the porin protein. Only *Mycobacterium microti*, *Mycobacterium smegmatis* and *Mycobacterium chelonae* were assayed in this way. *M. microti* was shown to have a band corresponding to the OmpATb protein when maximal levels of cell lysate were loaded, shown in Figure 8, whilst there was no detectable band in *M. smegmatis* or *M. chelonae*.

3.3 Identification of a putative regulatory system

The sequence of *Escherichia coli* OmpR, the protein responsible for the regulation of porins in this organism, and also a good representative of a general bacterial regulator protein, was retrieved from the SRS database at the European Bioinformatics Institute (http://www.srs.ebi.ac.uk). This was used in a BLAST search against the annotated *Mycobacterium tuberculosis* genome at the Sanger centre (http://www.sanger.ac.uk/Projects/M_tuberculosis/ blast_server.shtml).

The result of the blast search is shown in figure 9; a number of proteins were identified with homology to OmpR from *Escherichia coli*; of particular interest was that encoded by gene Rv0903c. This gene is present almost immediately downstream of the porin-encoding gene *ompATb* and is transcribed in the opposite direction, from the opposite strand. The BLAST search results did not identify the Rv0903 protein as the protein with the most homology to *Escherichia coli* OmpR. There are 5 other proteins with higher levels of homology, these are encoded by genes on the *Mycobacterium tuberculosis* cosmid numbers MTY20B11, MTV044, MTCY20G9, MTCY19H5, and MTCY10G2 (as deposited at the Sanger centre (http://www.sanger.ac.uk)). It was decided to concentrate on the Rv0903c gene, as this is present in close proximity to the porin-encoding gene *ompATb*. The presence of a transcript from the Rv0903c gene was determined in

Mycobacterium bovis BCG (figure 10) and in *Mycobacterium tuberculosis* (result not shown) by reverse-transcription PCR

The corresponding *Escherichia coli* sensor protein sequence, EnvZ, was also retrieved from the SRS database and again used in a BLAST search of the *Mycobacterium tuberculosis* genome. A gene encoding a protein with homology to EnvZ, Rv0902c, was identified immediately downstream of Rv0903c (reading in the direction of the coding sequence), again transcribed in the opposite direction to the porin gene. Again this was not the gene which encodes a protein with the most homology to EnvZ; there are two other genes on cosmids MTV044 and MTCY369 which have a higher degree of homology but gene Rv0902c was concentrated on for the reason given above. Cosmid MTV044 appears to have a sensor/regulator system with higher homology to the *Escherichia coli* OmpR/EnvZ pair; this can be identified as *Mycobacterium tuberculosis* genes Rv0981 and Rv0982, However this is some distance from the coding sequence of the porin gene *ompATb* (Rv0899). Alignments of the protein sequences using MegAlign (DNAstar inc) can be seen in figures 11 and 12.

Figure 13 shows the position of the porin gene and the sensor/regulator genes on the *Mycobacterium tuberculosis* genome; it is notable that both the sensor and the regulator gene are transcribed in the opposite direction to the porin gene. The position and layout of these genes is indicative of a likely candidate system for regulation of the porin. It is a reasonable inference that due to their proximity on the chromosome the sensor and the regulator genes encode proteins form a functional pair.

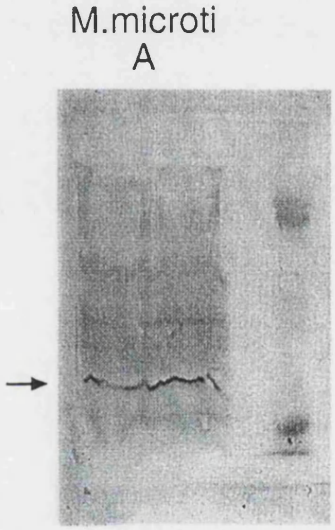


Figure 8 Western blot analysis of *Mycobacterium microti* with anti-OmpATb antibodies

MTCY31 [Full Sequence]
Length = 37,630

Minus Strand HSPs:

Score = 339 (119.3 bits), Expect = 6.1e-31, P = 6.1e-31
Identities = 83/226 (36%), Positives = 126/226 (55%), Frame = -3

```
Query:     6 KILVVDDDMRLRALLERYLTEQGFQVRSVANAEQMDRLLTRESFHLMVLDLMLPGEDGLS 65
             ++LVVDDD  + A LER L   GF+V +  +  +  R  T      +VLD+ +P  DG+S
Sbjct: 34037 RVLVVDDDSDVLASLERGLRLSGFEVATAVDGAEALRSATENRPDAIVLDINMPVLDGVS 33858

Query:    66 ICRRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKPFNPRELLARIRAVLRRQAN 125
             +   LR+  N +P+ +++A+    DR+ GLE GADDY+ KPF   EL+AR++A+LRR+ +
Sbjct: 33857 VVTALRAMDNDVPVCVLSARSSVDDRVAGLEAGADDYLVKPFVLAELVARVKALLRRRGS 33678

Query:   126 ELPGAPSQEEAVIAFGKFKLNLGTREMFREDEPMPLTSGEFAVLKALVSHPREPLSRDKL 185
                 A S  E  I  G  ++++  R        + LT  EF +L  L  H    LSR +L
Sbjct: 33677 T---ATSSSET-ITVGPLEVDIPGRRARVNGVDVDLTKREFDLLAVLAEHKTAVLSRAQL 33510

Query:   186 MNLARGREYSAMERSIDVQISRLRRMVEEDPAHPRYIQTVWGLGYV 231
             + L  G +++A    +DV I  LRR +E     PR + TV G+G+V
Sbjct: 33509 LELVWGYDFAADTNVVDVFIGYLRRKLEAGGG-PRLLHTVRGVGFV 33375
```

Figure 9 Result of a BLAST search of the *Mycobacterium tuberculosis* genome with the *Escherichia coli* OmpR sequence

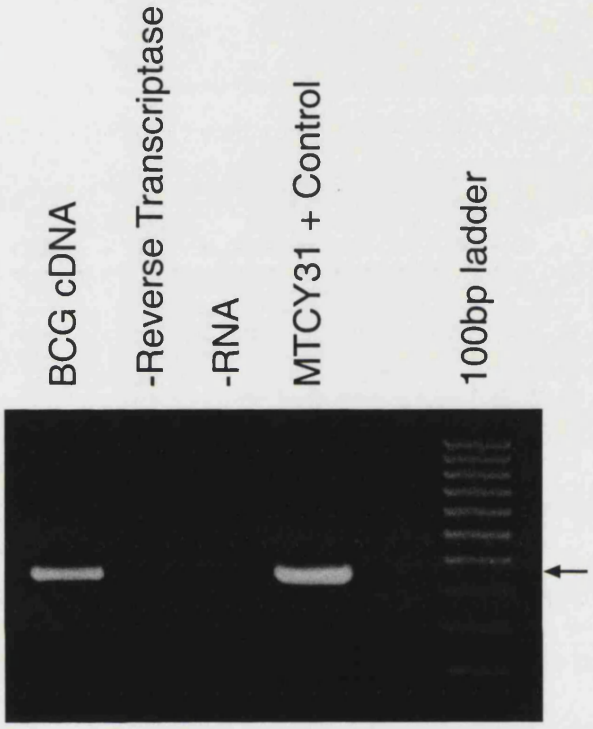


Figure 10 Rv0903 gene is transcribed in *Mycobacterium bovis* BCG

Results of RT-PCR on BCG cDNA using primers RAS021U and RAS022L (Section 2.14.1)

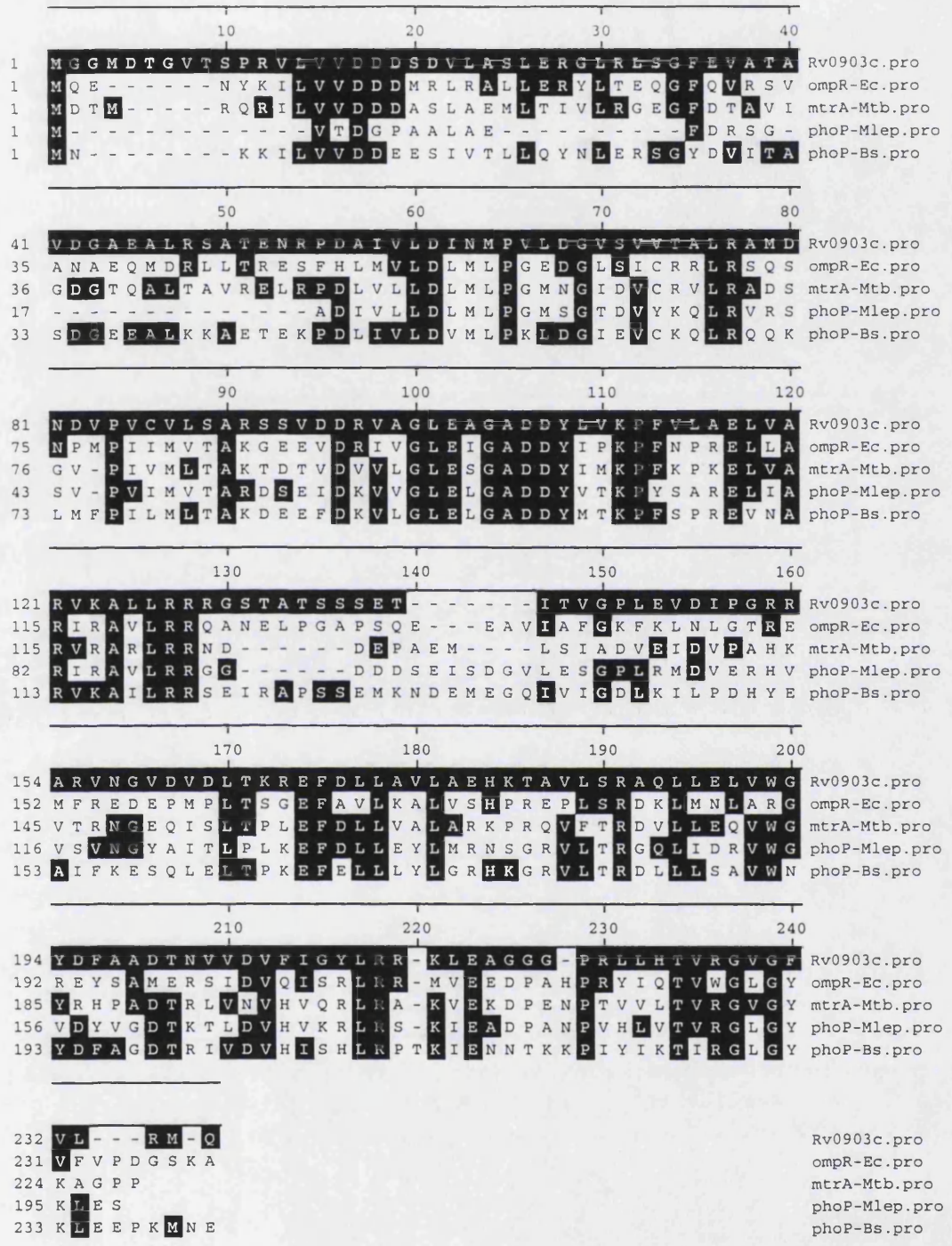


Figure 11 Alignment of other regulatory proteins with Rv0903 protein from *Mycobacterium tuberculosis*

Ec–*Escherichia coli*, Mtb–*Mycobacterium tuberculosis*, Mlep–*Mycobacterium leprae*, Bs-*Bacillus subtilis*. Shading indicates residues that match

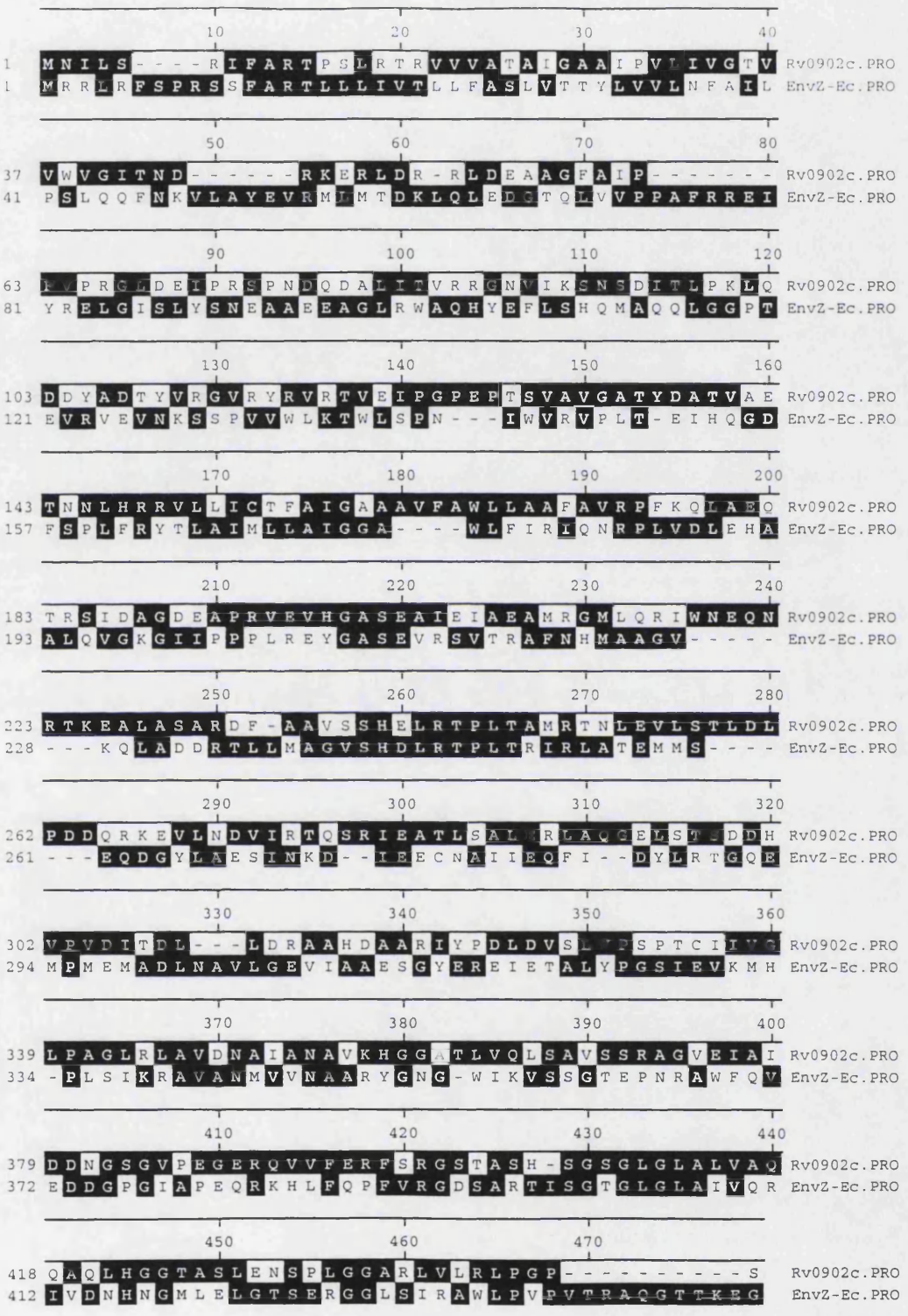


Figure 12 Alignment of *Mycobacterium tuberculosis* Rv0902 protein and *Escherichia coli* EnvZ. Shading indicates residues that match.

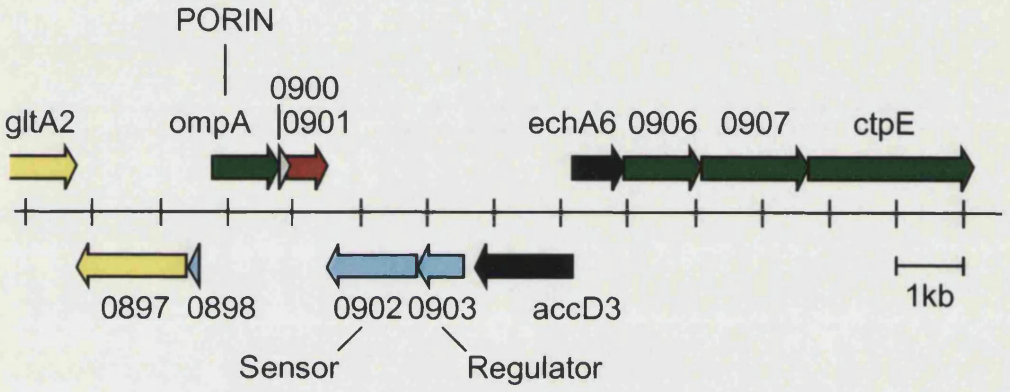


Figure 13 Position of coding sequences on the *Mycobacterium tuberculosis* genome

NOTE ADDED AFTER EXAMINATION

Supplementary Figure I at the end of this thesis shows a diagram of the *ompATb* locus and may be useful as a reference at this point

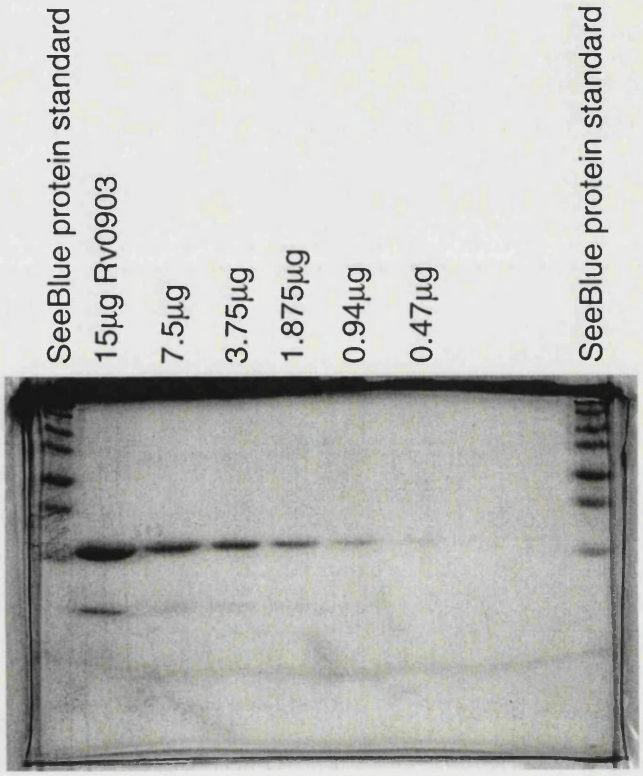


Figure 14 Purification of the recombinant Rv0903 protein

SDS-PAGE of Rv0903 protein after metal affinity chromatography and size exclusion chromatography

3.4 Expression and purification of Rv0903 protein

Figure 14 shows the result of an SDS-PAGE of the purified protein after expression, Talon column purification and Superose 12 HR 10/30 column purification. Such a procedure should remove any proteins which do not have a histidine tag (they cannot bind to metal affinity resin) and proteins of sufficiently different molecular weight to be present in a fraction different from the one which contains Rv0903. The concentration of the protein was measured by absorbance at 280 nm and shown to be 1 mg/ml.

There is a noticeable breakdown product of the protein at a size lower than that which is expected for the intact Rv0903c recombinant product (27 kDa). The amount of proteolysis was minimised by the inclusion of protease inhibitor tablets in the bacterial lysate and by freezing the purified product at -80°C after the addition of 20 % glycerol (final concentration). *Escherichia coli* BL21 (DE3) pLysS is also engineered to have lost its two major proteases and to have a minimal level of protein expression from the T7 promoter under uninduced (no IPTG) conditions; this means that basally expressed protein is not present in the cell (and therefore exposed to the remaining cellular proteases) for a prolonged period of time.

The recombinant protein is different in some respects to the native protein, there is a 6-histidine tag at the N-terminal end, a thrombin cleavage site (which would enable removal of the histidine tag) and also an additional 4 amino acids prior to the N-terminal methionine which is the start codon of the Rv0903c gene.

3.5 Gel Retardation assays

The purified recombinant Rv0903 protein was demonstrated to bind at least two regions of the *Mycobacterium tuberculosis* genome in or around the porin-encoding gene; a 450 bp region between 1781 bp and 1331 bp upstream of the coding sequence and another separate 196 bp region, internal to the porin gene between positions 594 and 790 of the coding sequence.

These sites were deduced by progressively reducing the size of the DNA fragment, a PCR product, that was radiolabelled and shown to bind to the recombinant Rv0903 protein. Typical results from the gel retardation assays can be seen in figure 15 and the fragments used can be seen in figure 16. The basis of the gel retardation (or 'band shift') is that a protein-DNA complex migrates through a polyacrylamide gel slower than DNA alone,

thus fragments which have not proceeded through the gel as fast as the corresponding DNA with no protein added are those which are deduced to form a complex.

Phosphorylation of the recombinant protein *in vitro* using acetyl phosphate as a donor of the phospho-group was shown to have a marked effect on the binding of the same regions of DNA. Figure 17 shows the result of a band-shift experiment after phosphorylation with acetyl phosphate. This demonstrates how, with an identical loading of protein in the samples, a more distinct band-shift is observed. It was possible to achieve a band-shift with an approximately two-fold lower amount of protein than using an unphosphorylated form. Quantification of such results might be possible using a system such as phosphor-imaging.

The importance of these results is discussed in more detail in section 4.5.

3.6 Expression of OmpATb in *Mycobacterium smegmatis*

The aim of this experiment was to investigate how expression of the porin protein in a non-native host (*Mycobacterium smegmatis*) would affect the growth or physiology of the cell. High level expression of *ompATb* from the mycobacterial *hsp60* promoter in a replicating vector (pRAS10) resulted in a severely compromised bacterial strain. After the initial transformation it was possible to obtain colonies from a 7H11 agar plate, but attempts to grow these further in broth failed. This is probably because high level expression of a membrane protein severely disrupts the envelope of *Mycobacterium smegmatis*. Since a replicating vector (pMV261) was used to express the OmpATb protein in this case it was decided to express the same construct in an integrating vector; this ensures that only one copy of the gene and the promoter is present per cell.

Expression of the porin gene from an integrating vector (pRAS11) was more successful as it was possible to isolate colonies and to grow them in broth culture. Expression of the porin was determined by western blotting of a *Mycobacterium smegmatis* mc^2155 pRAS11 cell lysate. Bands were very faint, though detectable, as the level of porin in relation to other cellular proteins is low. This is shown in figure 18. It was also possible to detect a transcript of the *ompATb* gene using a reverse-transcription PCR approach (results not shown)

Growth of *Mycobacterium smegmatis* harbouring pRAS11 was slower than the same bacteria containing pMV306 (the 'empty vector' parent of pRAS11). It appeared that

growth in identical media or in medium with 0.4 M NaCl was hampered by the presence of the porin. In media without any additional osmotic stress it was seen that growth of the porin-containing strain was significantly impaired, whilst if 0.4 M NaCl was added growth of both the control strain and the porin-containing strain was slower, again with the porin-containing strain growing more slowly than the pMV306 strain. Growth curves for these bacteria can be seen in figure 19. Interpretation of these results is difficult and is discussed in further detail in section 4.3. It should be noted that 400 mM NaCl is a considerable stress on the mycobacterial cell. Bergey's Manual of Systematic Bacteriology uses NaCl tolerance as a distinguishing characteristic for some mycobacteria, with a benchmark being the ability to grow in 600 mM NaCl (Bergey, 1986).

When the porin gene was included under the control of its own promoter, in vector pRAS21, growth was comparable to that of bacteria carrying the empty vector pMV306. However when the regulatory apparatus (Rv0902c/Rv0903c) was included it could be seen that growth appeared to be slower than in the control strain. This was the case both when no osmotic stress was included and when 0.2 M NaCl was included in the broth. RT-PCR was again used to verify that RNA from the transformed *Mycobacterium smegmatis* strains was expressing an *ompATb* transcript (results not shown). Further interpretation of these results is discussed in more detail in section 4.3.

Growth curves for *Mycobacterium smegmatis* expressing the porin gene (pRAS21) and the porin gene along with the regulatory apparatus (pRAS20) are shown in figures 20 and 21.

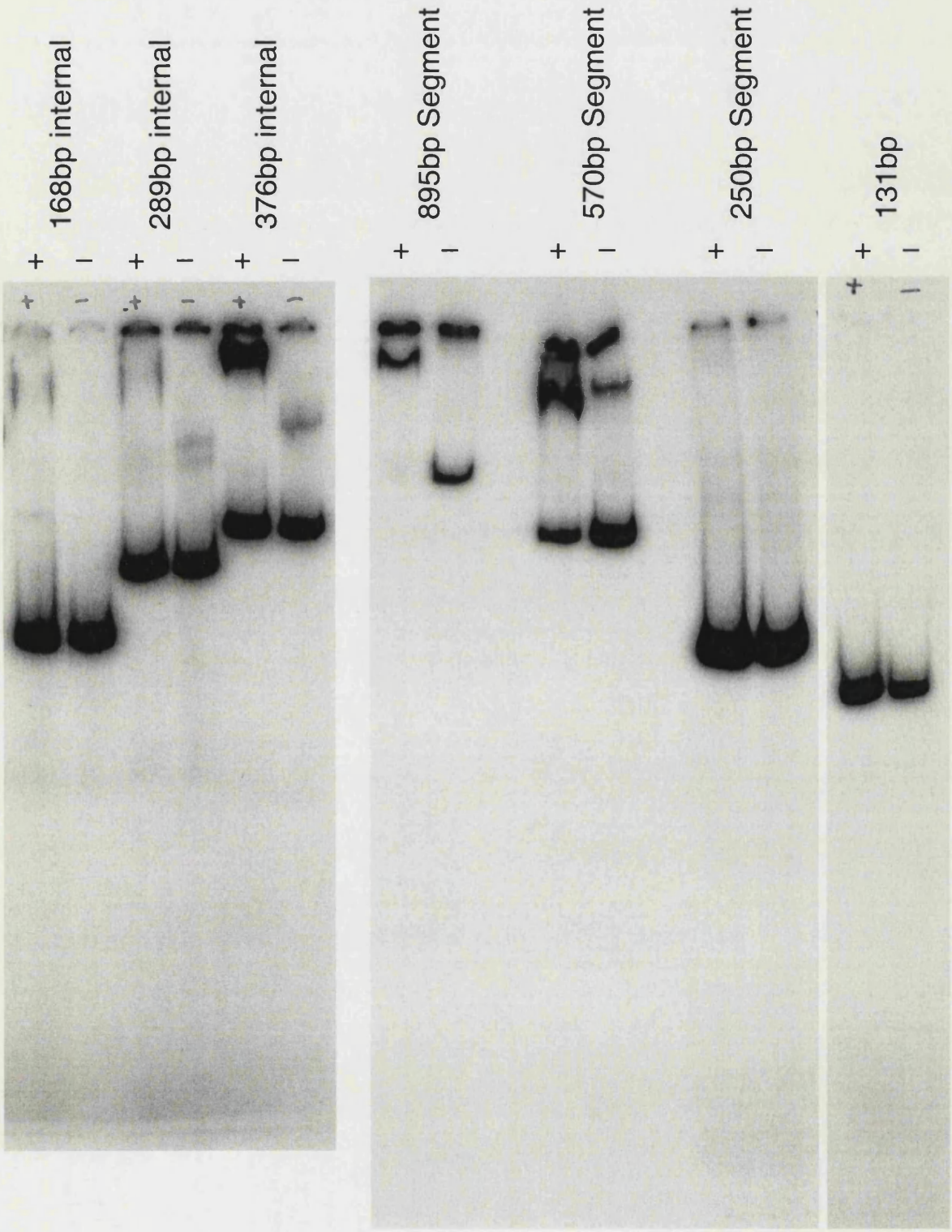


Figure 15 Sample results of a band-shift assay on regions of the mycobacterial genome using recombinant Rv0903 protein

NOTE ADDED AFTER EXAMINATION

Supplementary Figure III at the end of this thesis shows the position of the Rv0903 protein binding sites in the *M. tb* genome identified by gel retardation assay

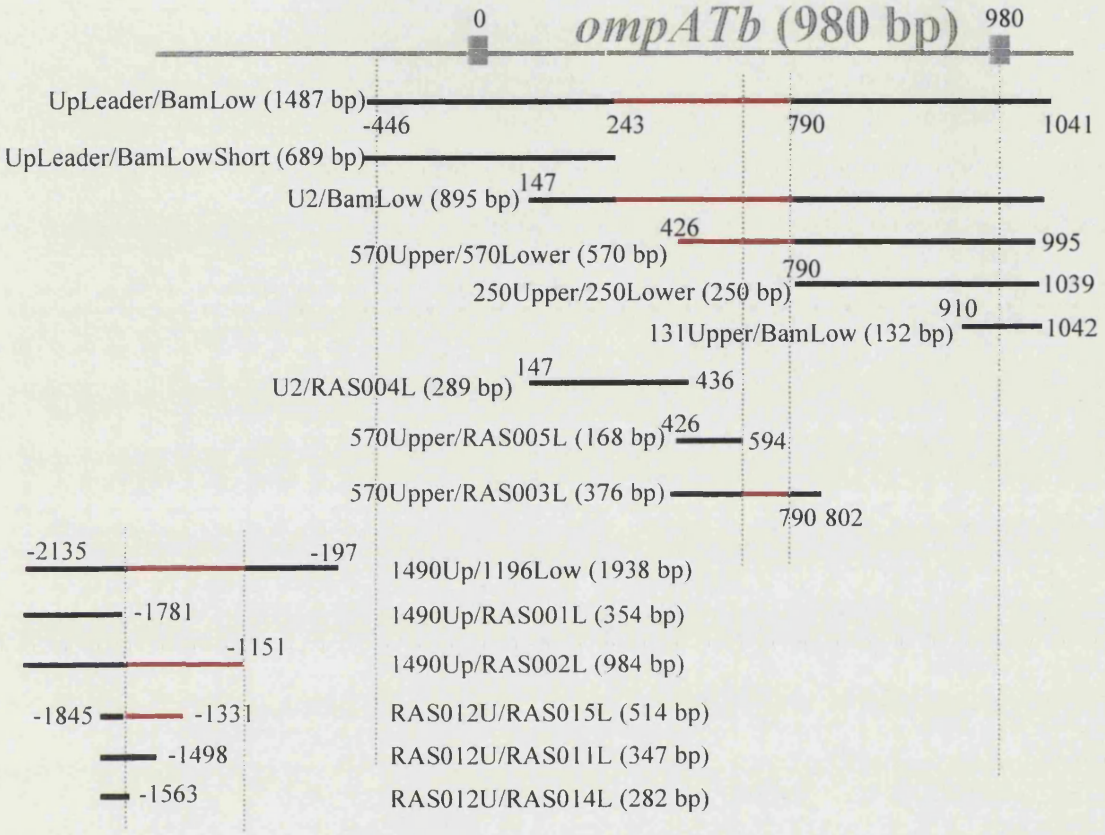


Figure 16 Schematic representation of fragments used in gel shift assay

Red areas are those which are deduced to bind protein. Diagram is not to scale, primers used and the length of products are given.

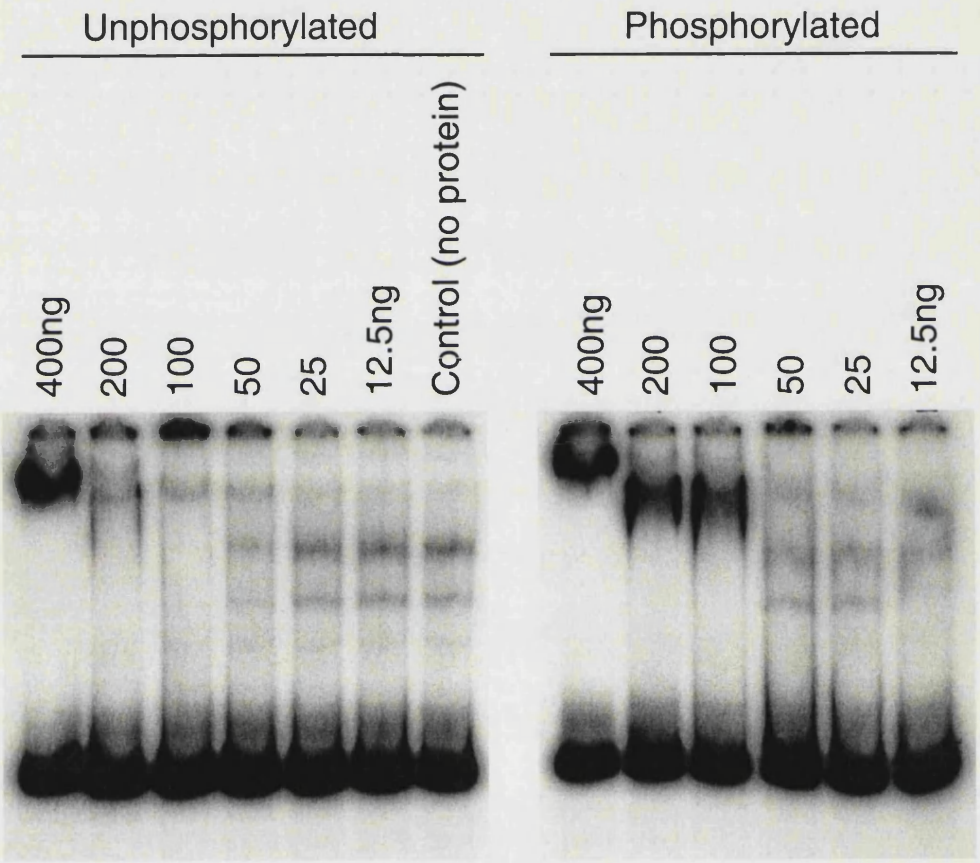


Figure 17 Phosphorylation of the recombinant protein enhances binding to DNA

Figure shows a 343 bp region generated by PCR using primers RAS006U and RAS007L (for details see section 2.19.2). After *in vitro* phosphorylation with acetyl phosphate affinity of the protein for the DNA is seen to be enhanced

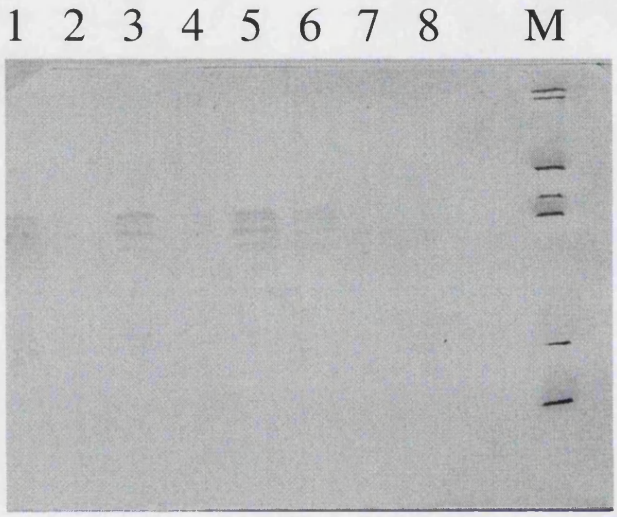


Figure 18 Expression of the OmpATb protein in *Mycobacterium smegmatis* assayed by western blot

Lane 1, 3, 5 - Cell lysate in Zwittergent 3-12 buffer (1 % Zwittergent 3-12, 40 mM EDTA, 20 mM Tris-HCl pH 8) from *Mycobacterium smegmatis* transformed with pRAS20

Lanes 2, 4 and 6 - washing of beads in urea buffer

Lanes 7 and 8 - Wild type *Mycobacterium smegmatis*, cell lysate in Zwittergent 3-12 buffer (7) and bead washing in urea buffer (8).

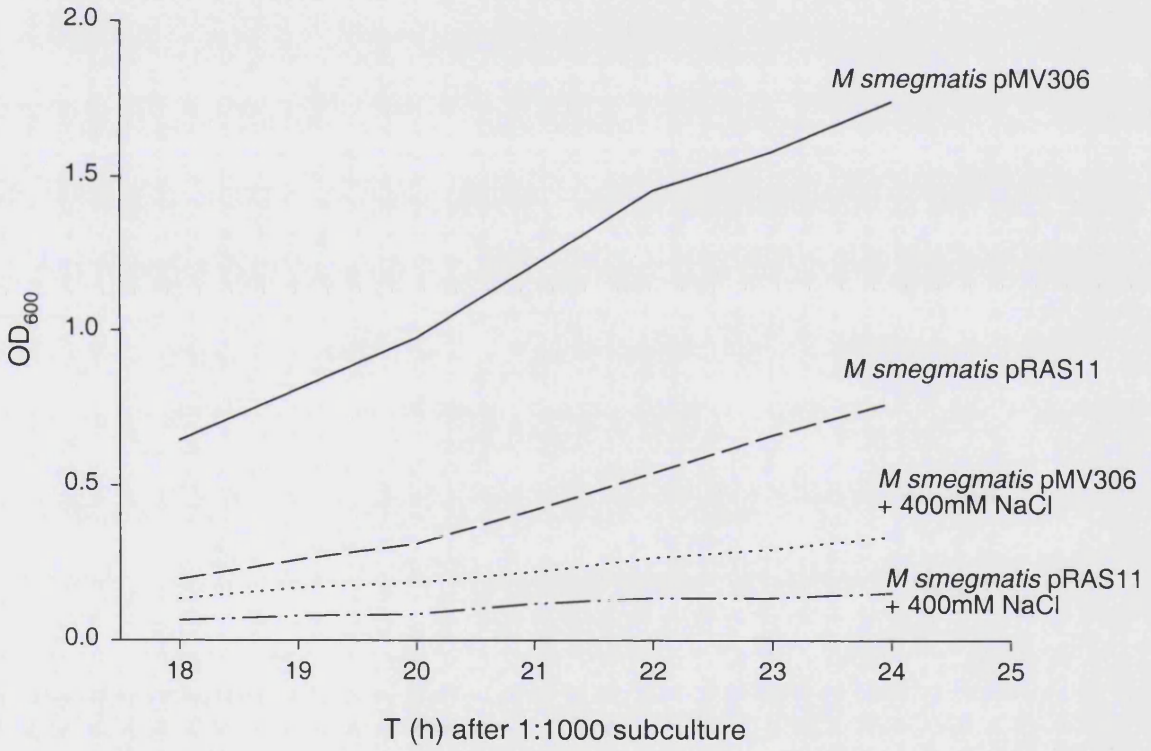


Figure 19 Growth of *Mycobacterium smegmatis* expressing the porin from a single copy of the *hsp60* promoter per cell (transformed with plasmid pRAS11)

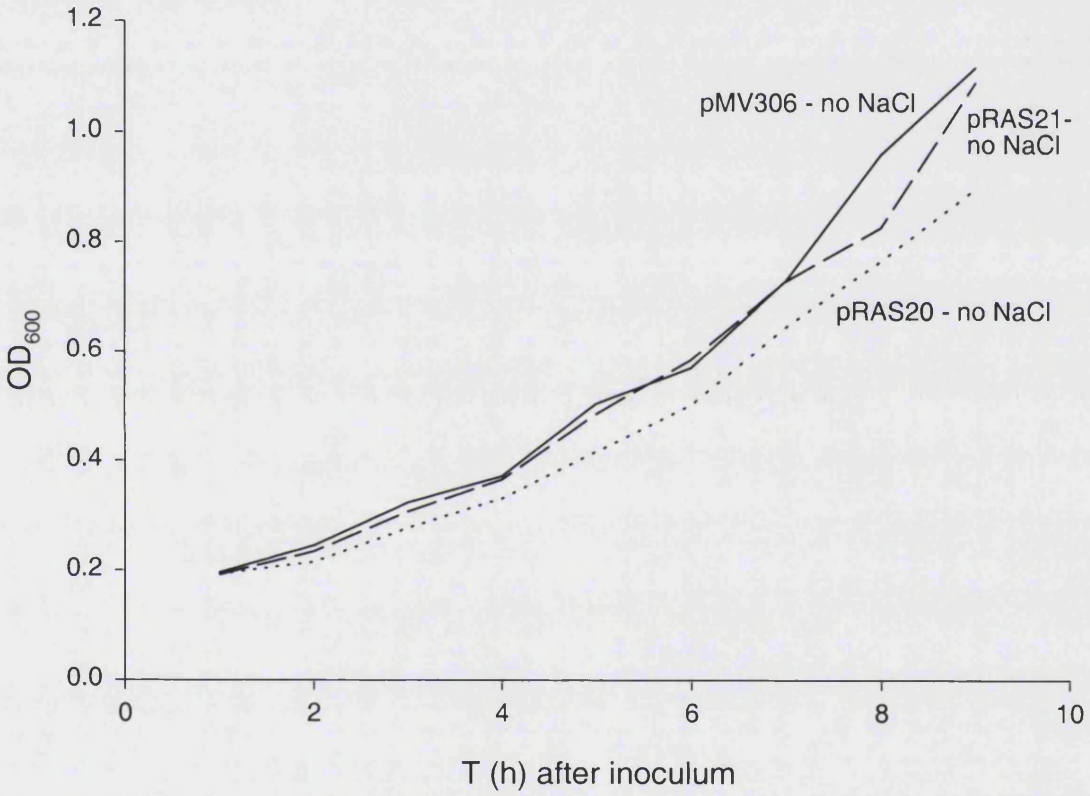


Figure 20 Growth of *Mycobacterium smegmatis* expressing the porin (pRAS21) or the porin and associated regulatory machinery (pRAS20) without additional stress

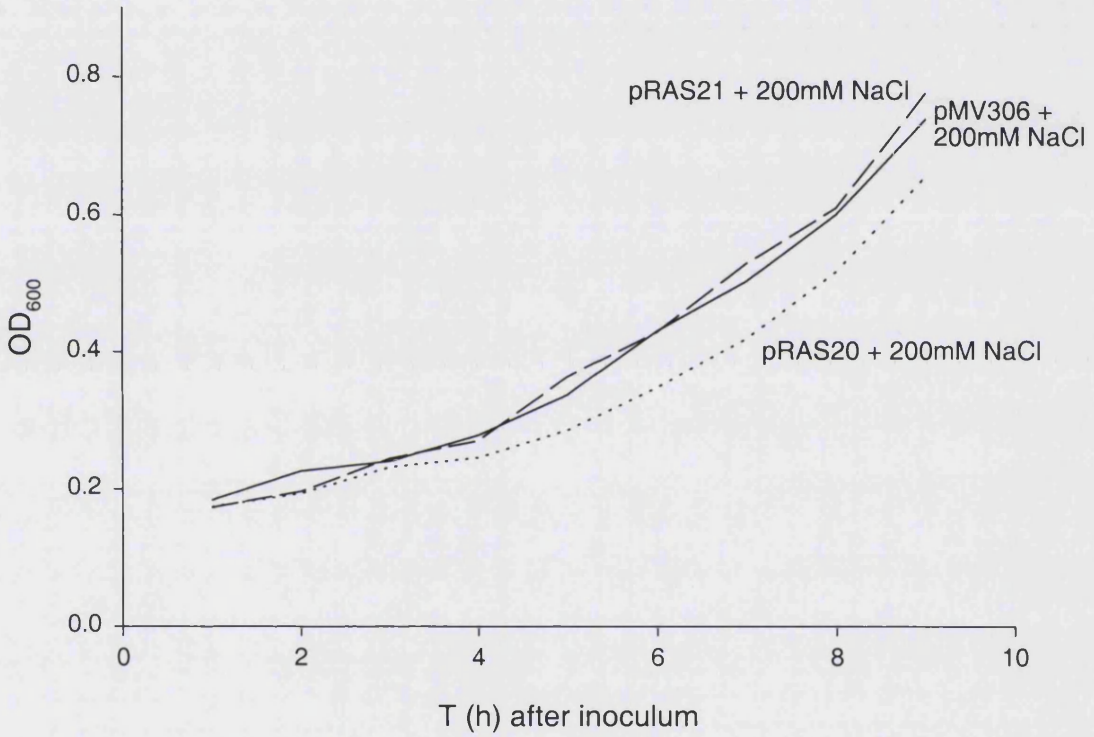


Figure 21 Expression of the porin (pRAS21) or the porin and associated regulatory machinery (pRAS20) in *Mycobacterium smegmatis* with 200 mM NaCl in the culture medium

3.7 Analysis of *Mycobacterium tuberculosis* transformed with pRAS5 or pOmpA-aph-rpsL tb for *ompATb* knockout

3.7.1 *Mycobacterium tuberculosis* transformed with pRAS5

Thirteen *Mycobacterium tuberculosis* colonies were isolated from 7H11 agar plates containing 2 % sucrose and 25 μg/ml kanamycin after plating the pRAS5-transformed bacteria. These colonies were both sucrose resistant (indicating that *sacB* was not being expressed) and kanamycin resistant (the cassette inserted in the coding sequence of the *ompATb* gene appears to be expressed), thus they displayed the correct phenotype for homologous recombinants. However upon screening by PCR it was seen that they were all wild type with respect to the porin gene. All PCR amplifications gave the 343 bp wild type product; in no case was there no product (indicating gene disruption). The reason for this is likely to be that the isolated bacteria are spontaneous mutants in the *sacB* gene, which is known to have a high rate of mutation (personal communication, Dr KG Papavinasasundaram, NIMR). It is also possible that they were spontaneous kanamycin resistant colonies although a control experiment without transformation of a kanamycin resistant plasmid did not lead to as many colonies (results not shown). It was not possible to obtain any mutants of *Mycobacterium tuberculosis* using pRAS5

3.7.2 *Mycobacterium tuberculosis* transformed with pOmpA-aph-rpsL tb

Mycobacterium tuberculosis 1424 that had been transformed with pOmpA-aph-rpsL tb grew on 7H11 agar with streptomycin (100 μg/ml) and kanamycin (25 μg/ml). This indicated that either the wild type *rpsL* gene was not being expressed from the plasmid in a randomly integrated form, or that there had been a mutation in the wild type (previously plasmid-borne) *rpsL* such that the domininant phenotype (wild type streptomycin sensitivity) was not expressed. It also indicated that the kanamycin resistance gene was present (internal to the *ompATb* coding sequence) or that spontaneous kanamycin resistance had arisen.

Eight of these isolates were screened by PCR as detailed in section 2.19.3, and Southern blot as detailed in section 2.28. The PCR products observed (see figure 22) were 1360 bp (indicating incorporation of the kanamycin resistance cassette), 512 bp (a wild type product) or both bands (if single crossover or random integration had occurred). Southern blotting using a 132 bp probe (see section 2.28) clearly showed different sized products in

the wild type and the mutant as shown in figure 23. It can be seen that a number of different genotypes were observed in the 8 colonies selected for screening and that wild type, single crossover (or random integrant) and Δ*ompATb* strains are identified. Isolate number 1, a porin deficient mutant, was inoculated into broth for further experimentation. All other isolates were frozen at -70°C after the addition of glycerol (to a final concentration of 20 %).

3.8 Analysis of *Mycobacterium tuberculosis* transformed with pRAS18 for Rv0903c knockout

A two-stage procedure was used to isolate potential Rv0903c knockouts; *Mycobacterium tuberculosis* H37Rv transformed with pRAS18 was first plated onto 7H11 media containing X-gal (100 μg/ml) and kanamycin (25 μg/ml). It was seen that pre-treatment with ultraviolet light made transformation efficiency approximately 4 fold higher. The results of this are seen in Table 4. Integration of the plasmid (a single crossover or a random integration) results in expression of the *lacZ* gene, thus on addition of X-gal the colonies will be blue in colour. Colonies which appear on 7H11 + kanamycin but are not blue in colour are either spontaneous kanamycin resistant mutants or successful single crossovers (or random integrants) which have a loss of function from the *lacZ* gene.

	UV treated plasmid				**Untreated plasmid**				**Control (no plasmid)**	
Plate number	***1***	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***
BLUE COLONIES	66	50	54	39	9	8	15	16	0	0
Average	**52**				**13**				**0**	
WHITE COLONIES	12	17	17	7	17	17	19	8	11	22
Average	**12**				**15**				**17**	

Table 4 Transformation efficiency of suicide delivery vector is enhanced by pre-treatment of plasmid DNA with ultraviolet light

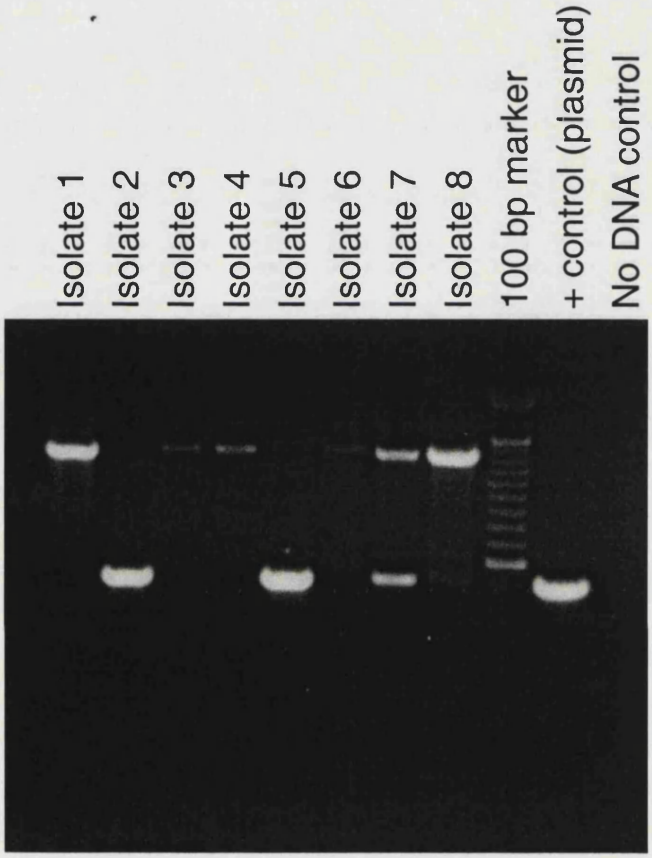


Figure 22 PCR-based analysis of pOmpA-aph-rpsL tb transformants (8 isolates)

Isolates 1, 3, 4, 6 are seen to be mutants (1360 bp)

Isolates 2 and 5 are seen to be wild type *Mycobacterium tuberculosis* 1424 (512 bp)

Isolates 7 and 8 are single crossover or random integrants (both bands)

NOTE ADDED AFTER EXAMINATION

Supplementary Figure II at the end of this thesis shows a diagram of the *M. tb* genome in the wild type and *ompATb* knockout strain, including the region of homology used in the recombination experiment and the sizes of restriction fragments used in Southern blot verification of HR.

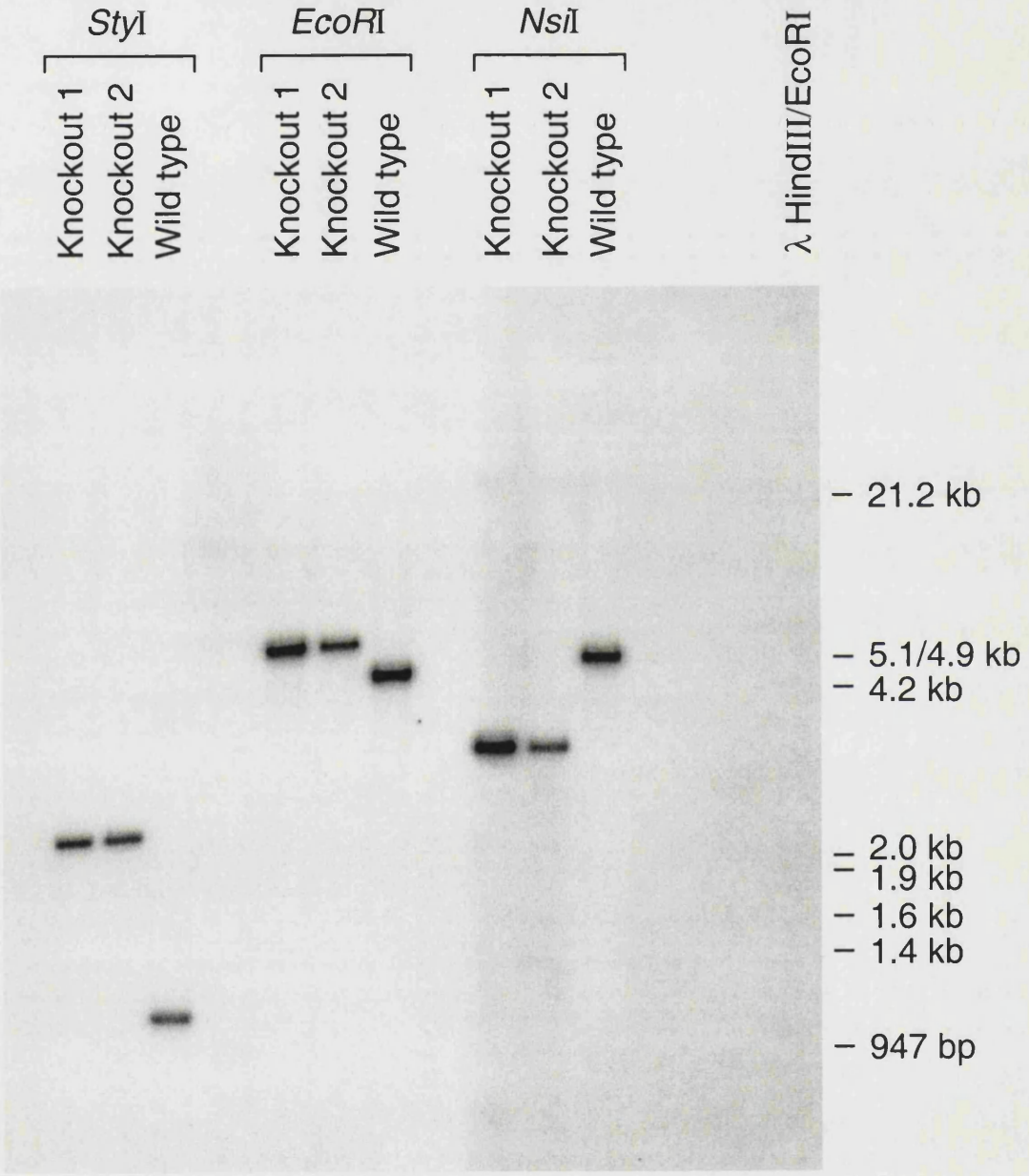


Figure 23 Southern blot analysis of *Mycobacterium tuberculosis* Δ*ompATb* knockout strains

Wild type: *Sty*I 1 kb, *Eco*RI 4.4 kb and *NsiI* 4.7 kb

Δ*ompATb* mutant: *Sty*I 2.3 kb, *Eco*RI 5.3 kb and *Nsi*I 2.6 kb

3.8.1 Attempt at second crossover of pRAS18 transformants in liquid broth

In order to attempt a second crossover (and thus a gene knockout) from the colonies isolated above 12 blue colonies were selected; 8 of these were from transformants that received UV pre-treated plasmid DNA and 4 were from transformants that received supercoiled plasmid DNA. These were inoculated into 5 ml Dubos broth (with albumin and glycerol) without any selection pressure and allowed to grow for 3 weeks at 37°C. After serial dilution of these cultures in PBS (50 µl culture into 450 µl PBS) with Tween20 (0.1 %) 100 µl was plated onto 7H11 agar containing 2 % sucrose and 100 µg/ml X-gal and incubated for 4 weeks.

Fifteen sucrose resistant white colonies were isolated; this indicates both loss of the *lacZ* gene (which preciously gave these isolates a blue colour on X-gal agar) and loss of *sacB* (which confers sucrose sensitivity). These colonies therefore had the correct phenotype for a gene knockout. All isolates were plated onto 7H11 with kanamycin (25 µg/ml) and it was shown that they were sensitive to this antibiotic; this demonstrated that the previously integrated plasmid carrying *sacB*, *kan*^{r} and *lacZ* had been lost. DNA from the bacteria was analysed by PCR. All of the 15 isolates had a PCR product identical to the wild type indicating that a knockout of Rv0903c had not been successful.

3.8.2 Attempt at second crossover on solid media

Alternatively 4 of the single crossover strains were plated onto 7H11 agar with no antibiotic or sucrose to allow a second crossover to occur on solid media, a technique which is believed to yield more successful recombinants (Personal communication, Dr T Parish, London School of Hygiene and Tropical Medicine). After growth these colonies were diluted and plated as in section 2.20.3. Forty-two colonies which grew on 7H11 agar with sucrose and X-gal and were white indicating loss of the previously integrated plasmid by single crossover were picked, plated and analysed by PCR (Section 2.20.3). All of these colonies were seen to have a PCR product identical in length to the wild type; this is shown in figure 24.

In all a total of 57 colonies with the expected phenotype for a regulatory gene knockout and appeared to have lost the integrated plasmid after a single crossover event were

screened by PCR. All of these had a product identical to the wild type; it was therefore not possible to disrupt the Rv0903c gene in this study.

Possible reasons for this and further approaches to the experiment are detailed in section 4.7.

3.9 β-galactosidase assays for promoter activity

Plasmid pEJ414 has a multicloning site upstream of a *lacZ* gene, thus any fragment of DNA placed into this multicloning site can be assayed for promoter activity since it may drive expression of the *lacZ* gene. Expression of *lacZ* (β-galactosidase) leading to its presence in the cell lysate can be measured by breakdown of the substrate ONPG, this gives a yellow colour which can be measured by spectrophotometry. It can be seen in figure 25 that the strength of the promoter varies with the length of the region from upstream of the porin gene which is incorporated; the 2.2 kb fragment upstream of *ompATb* shows much stronger promoter activity than a 178 bp fragment, however the 684 bp fragment and the 452 bp show stronger activity still. The strongest promoter activity is seen in the 452 bp fragment.

Expression from pEJ414 in the absence of any incorporated fragment is in the region of 0-5 Miller units; therefore a significant promoter activity above the background level from the *Mycobacterium tuberculosis ompATb* upstream region was observed in *Mycobacterium smegmatis* in all cases.

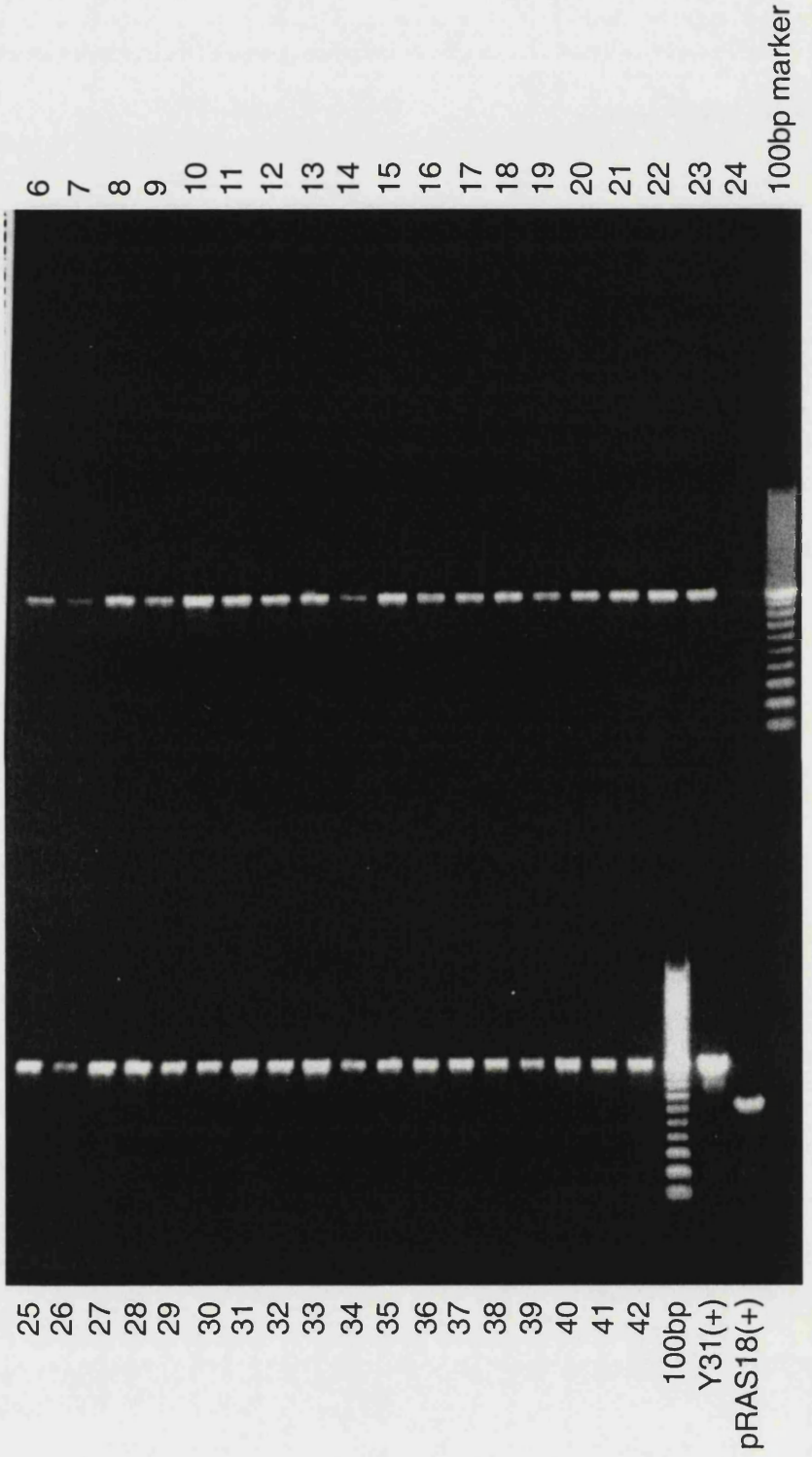


Figure 24 Results of PCR-based screen of potential Rv0903c knockouts

All isolates are observed to have the 992 bp wild type product

NOTE ADDED AFTER EXAMINATION

Supplementary Figure IV at the end of this thesis shows the position of the fragments used in the β-galactosidase assays in relation to the position of the genes around *ompATb*.

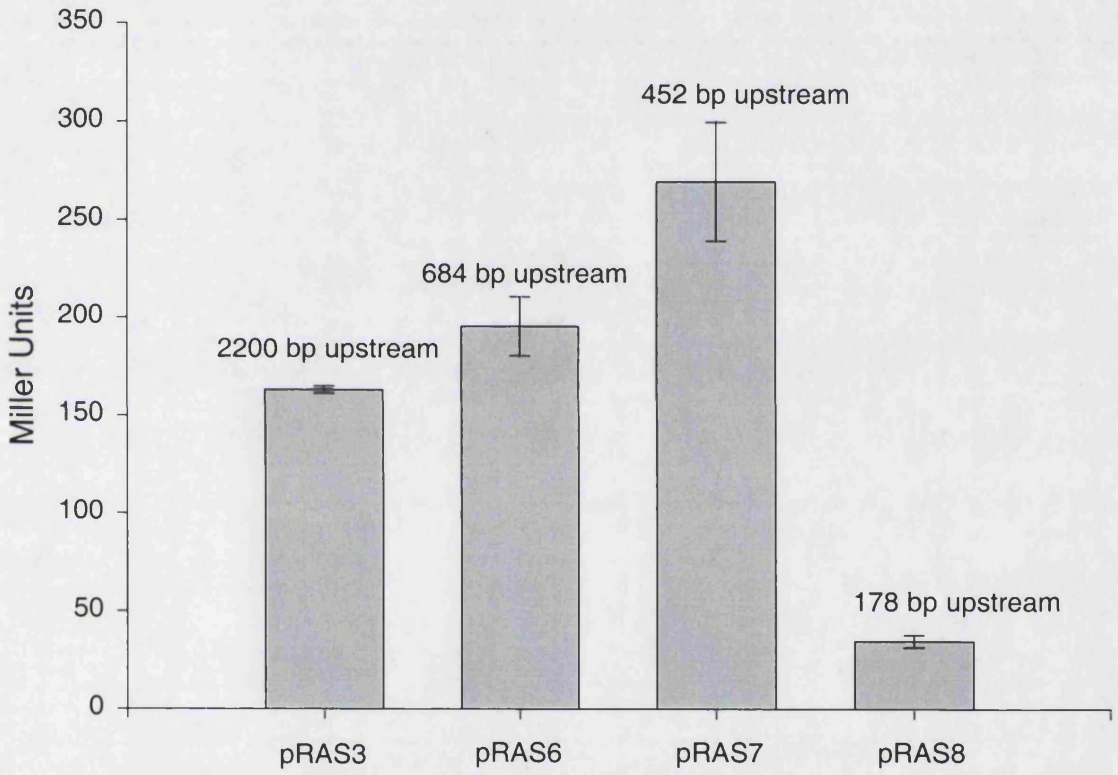


Figure 25 β-galactosidase activity of *Mycobacterium smegmatis* cell lysates after transformation with pRAS3, pRAS6, pRAS7 or pRAS8

3.10 Real-time PCR to demonstrate regulation of the porin gene

3.10.1 Real time RT-PCR of BCG

Figure 26 shows a typical amplification from one well of a 96-well plate, the cycle threshold (Ct), corresponding to the amount of template, can be seen after 13.7 cycles. Sample results of NaCl shock on BCG cultures are shown in figures 27 and 28. It was clearly shown that osmotic stress in the form of salt shock led to a decrease in the level of transcript of the porin gene reflected in a lower *ompATb/gnd* ratio. The results presented in the figures come from a single run of the TaqMan machine, though they were repeated and gave the same result in terms of overall ratio as is seen from the raw data presented in the appendix. Averaging of these results was not attempted; the real-time PCR machine is extremely sensitive and may have given different results (in terms of absolute figures) from one experiment to the next due to mechanical rather than biological variance. It is more important to examine the trends that are apparent in each data set analysed, whether between replicates of the nucleic acid extracts from a single culture, or nucleic acid extracts from several cultures.

These results demonstrate that there is an active regulation of the porin gene in response to an environmental stimulus, in the same way as that which occurs in *Escherichia coli.*

Typical results from salt-stress of the BCG cultures are presented in tables 5 and 6.

Time (h) after addition of 400 mM NaCl	Ratio *ompATb/gnd*
0	0.667
1	0.1196
2	0.0786

Table 5 Level of *ompATb* transcript varies with time after addition of 400 mM NaCl

Final molarity of NaCl in media	Ratio of *ompATb/gnd* (after 1 hour)
0	0.514
0.2	0.218
0.4	0.161

Table 6 Level of *ompATb* transcript varies with molarity 1 h after NaCl addition

Sucrose stress of BCG gave a more ambiguous result; this was only attempted once before work was concentrated on *Mycobacterium tuberculosis*, however it can be seen that no clear trend is established. A puzzling feature of this set of results is the way in which the level of the porin gene is seen to fluctuate in the absence of the addition of any stress to the culture medium, this phenomenon was not observed previously in the NaCl stress experiments and may be due simply to errors inherent in the data collection by the TaqMan machine which are prone to exaggeration by the way in which the data are analysed.

Results from the sucrose stress experiment on BCG are presented in Table 7; these results are discussed in more detail in section 4.4.

Time (no sucrose)	Ratio *ompATb/gnd* (2 experiments)		Ratio *ompATb/gnd* with addition of 200 mM sucrose (2 experiments)	
	1	**2**	**1**	**2**
0	0.87	0.95	0.76	0.65
1	1.58	0.46	n/a	0.40
2	0.9	0.38	1.02	0.32
3	0.19	0.70	3.79	0.36
4	1.71	0.33	1.96	0.33

Table 7 Result of addition of 200 mM sucrose to growth media of BCG

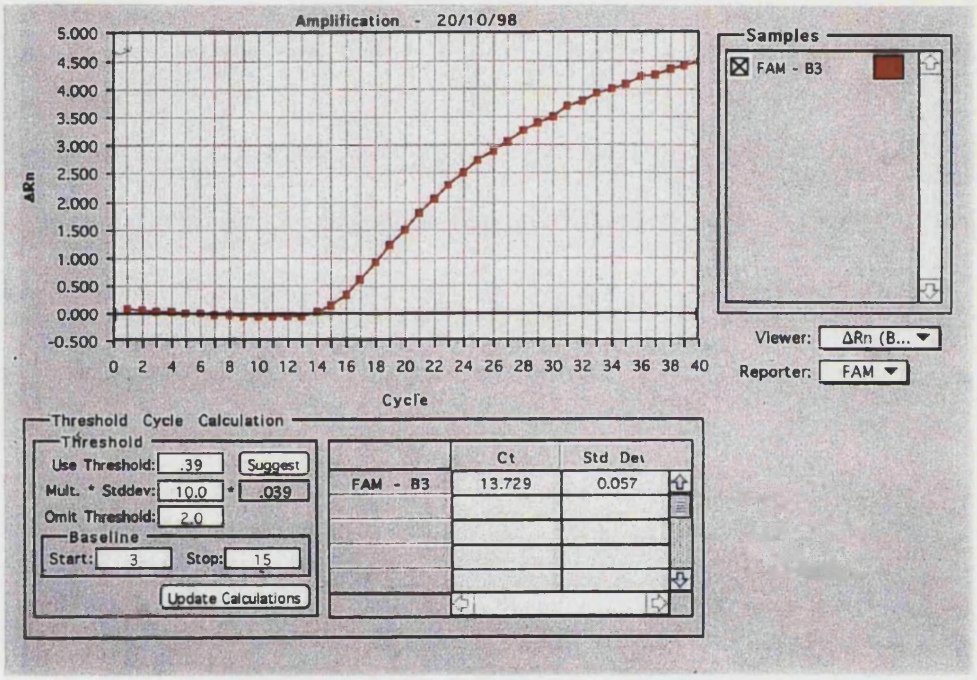


Figure 26 Sample amplification of a cDNA target in the real-time PCR machine, the cycle threshold (Ct) can be seen as the point where the curve crosses a predefined fluorescence value.

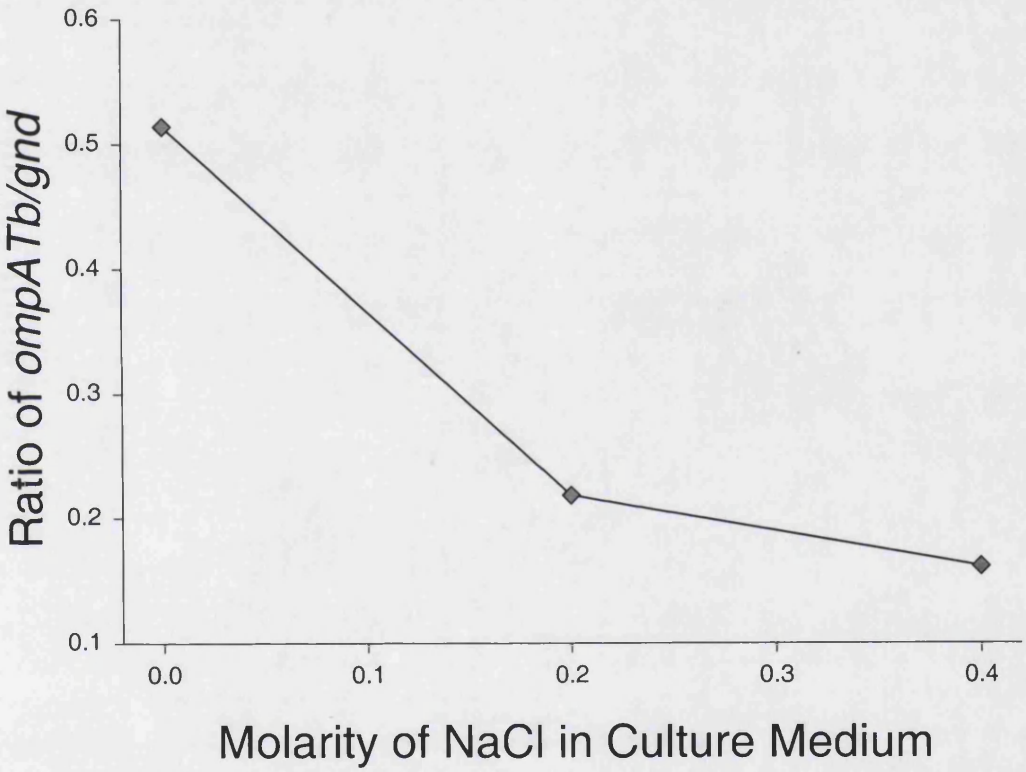


Figure 27 Molarity of culture medium affects levels of *ompATb* transcript

Values for *ompATb/gnd* ratio are shown 1 hour after NaCl addition

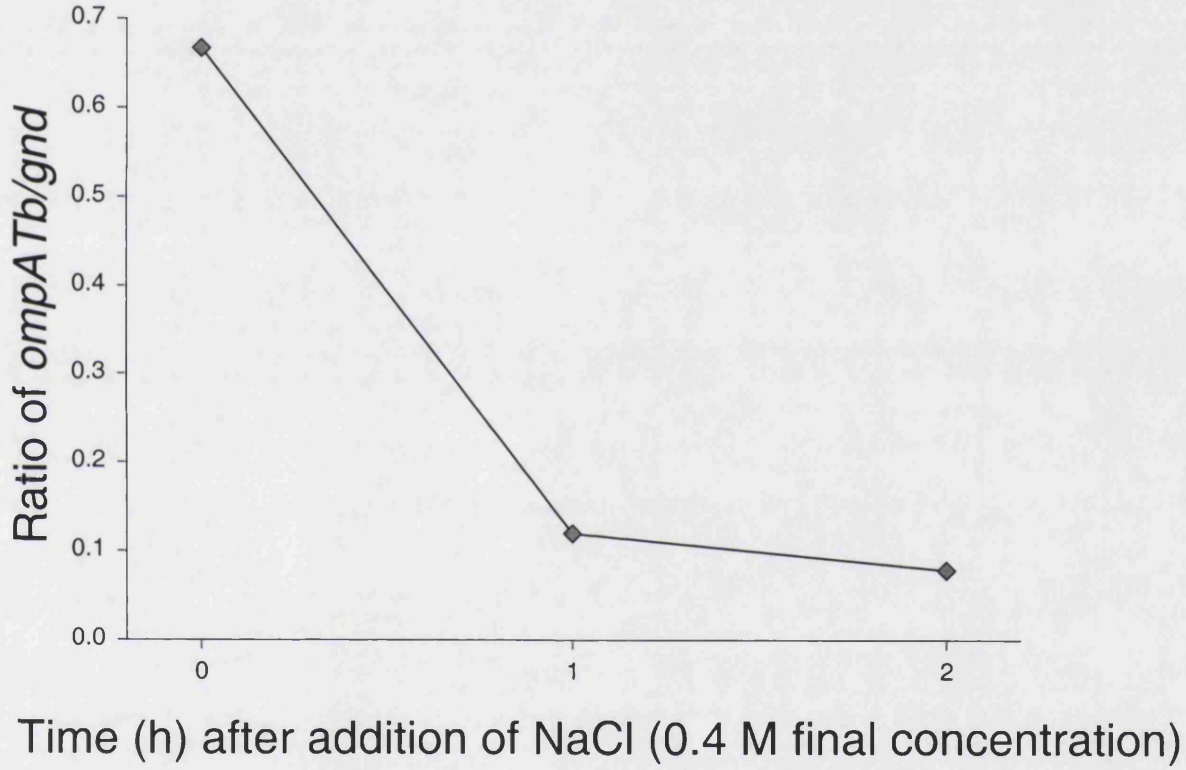


Figure 28 Time after addition of NaCl shock (400 mM) affects levels of *ompATb* transcript

3.10.2 Real time RT-PCR of *Mycobacterium tuberculosis*

The results of NaCl stress on cultures of *Mycobacterium tuberculosis* were not as clear as those for BCG. It was not evident that there was any regulation of the porin gene in response to salt shock and again the ratio of *ompATb* to *gnd* transcript was seen to change in the absence of any NaCl addition. It is important to note that if each time point is considered in turn the *ompATb/gnd* ratio is very similar, which perhaps points to a genuine fluctuation in levels rather than a fault of the machine. Interpretation of this phenomenon is discussed in section 4.4. Table 8 shows the result of NaCl addition to *Mycobacterium tuberculosis* cultures. Raw data is included in the appendix.

Time (h)	Ratio *ompATb/gnd* no additional stress	Ratio *ompATb/gnd* with 200 mM NaCl added at T=0
0	0.21	0.26
2	0.58	0.45
6	0.57	0.49
24	0.32	0.35

Table 8 Effect of addition of 200 mM NaCl on transcript levels of *ompATb* in *Mycobacterium tuberculosis*.

3.11 Microarray hybridisation

3.11.1 Comparison of genomic DNA from *Mycobacterium tuberculosis* 1424 and Δ*ompATb* mutant using a microarray

Genomic DNA from *Mycobacterium tuberculosis* 1424 was compared with that of the Δ*ompATb* knockout mutant. The main aim of this procedure was to determine if the spot corresponding to the porin gene *ompATb* (Rv0899) hybridised to DNA from the mutant strain.

Unfortunately, the way in which the mutant was constructed meant that a substantial portion of the coding sequence remained (a total of 421 bp). This was sufficient to bind to

the PCR product on the microarray slide and thus no difference was observed between the wild type and the mutant. The PCR product on the slide consists of a 661 bp fragment beginning 122 bp into the coding sequence (ending at base 783 of the coding sequence). The knockout of the gene removes only until base 571 of the coding sequence, therefore 212 bp with identity to the probe remain at the 3' terminus of the porin gene.

3.11.2 Comparison of RNA from *Mycobacterium tuberculosis* 1424 and Δ*ompATb* mutant using a microarray

RNA from both the wild type *Mycobacterium tuberculosis* 1424 and the Δ*ompATb* knockout mutant was harvested from cultures grown in Dubos medium at OD_{600} of 0.6. The aim of this experiment, as above, was to determine if there would be any binding to the spot corresponding to the porin gene by the *ompATb* transcript. It was also attempted to identify any other genes which may be regulated differently in response to the disruption of the porin, particularly those which may encode other pore-forming proteins in *Mycobacterium tuberculosis*, such as those identified by Kartmann *et al.*, (1999). These were identified only at the protein level and not at the genetic level, so it was not possible to examine directly if the genes responsible were altered in their expression.

Again it was seen that the portion of the gene that remains in the mutant was sufficient to give a transcript that binds the PCR probe, thus no difference was observed between the wild type and the mutant.

3.11.3 Comparison of *Mycobacterium tuberculosis* H37Rv with NaCl shocked bacteria using a microarray

The purpose of this experiment, since it had been shown that porin levels in BCG alter in response to NaCl shock, was to examine whether a similar alteration occured in *Mycobacterium tuberculosis*. It was seen that the spot corresponding to the porin gene hybridised equally well to both the salt-shocked bacteria and to the bacteria that were not shocked. This is in agreement with the results from the real time RT-PCR experiment, which failed to find any difference in porin gene transcript levels after salt shock.

At the same time it was however possible to examine whether salt shock had any effect on other genes. It was seen that the *GroEL2* and *GroES* genes were markedly up regulated upon addition of NaCl to growing cultures. Figure 29 shows a typical section of a microarray, demonstrating the hybridisation of both Cy3 (fluoresces red) and Cy5

(fluoresces yellow). A scatter plot of the relative fluorescence is shown in figure 30, this shows the *gro* genes as white spots, other genes which appear to be upregulated were analysed and seen to be not of significance; most relate to aftefacts on the slide and errors in processing the data.

Whilst it was disappointing to find that there was no observable difference between the wild type and porin-knockout bacteria, the microarray technique functioned well. This procedure is discussed in more detail in section 4.8.

3.12 Analysis of growth of *Mycobacterium tuberculosis* without the porin encoding gene

The growth of *Mycobacterium tuberculosis* Δ*ompATb* was examined in the presence of sucrose (0.2 M), NaCl (0.2 M) or raffinose (0.1 M). Growth was compared with the wild type *Mycobacterium tuberculosis*, strain 1424; the reason for this is that Senaratne *et al* (1998) had previously shown that sucrose is able to diffuse through the OmpATb pore whilst raffinose is not, since it is a much larger sugar. It is not possible to use sugars smaller than sucrose as these may be metabolised by the mycobacterial cell thus their stress on the cell is less easily measured. Results are presented in figures 31 to 34.

From these curves it is not possible to draw any certain conclusion regarding the growth of the porin-deficient mutant. It is obviously not seriously impaired in comparison to the wild type, and does not seem to have any noticeable advantage since although the curves do not always overlap the gradient is similar in all comparable cases.

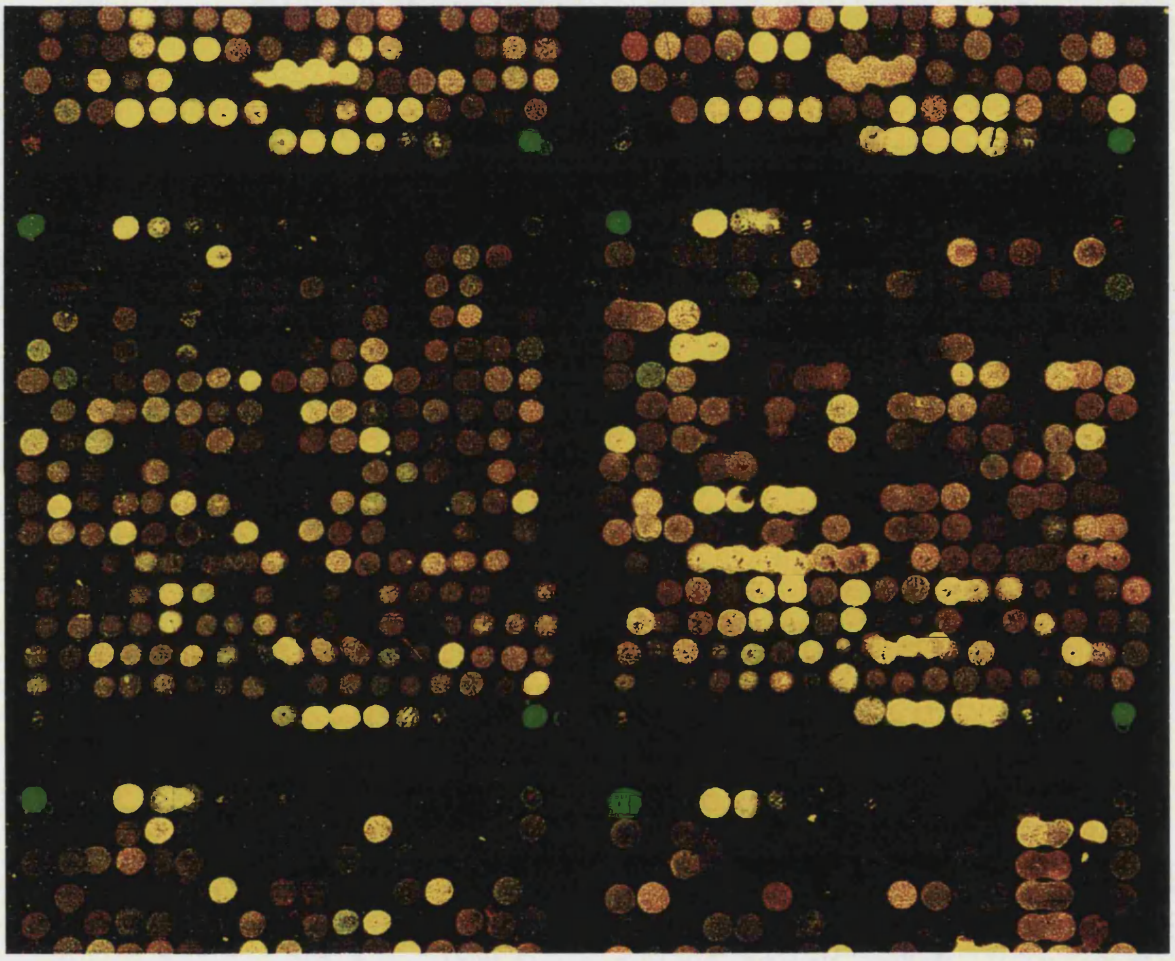


Figure 29 Sample results from microarray hybridisation using Cy3 and Cy5 labelled nucleotides

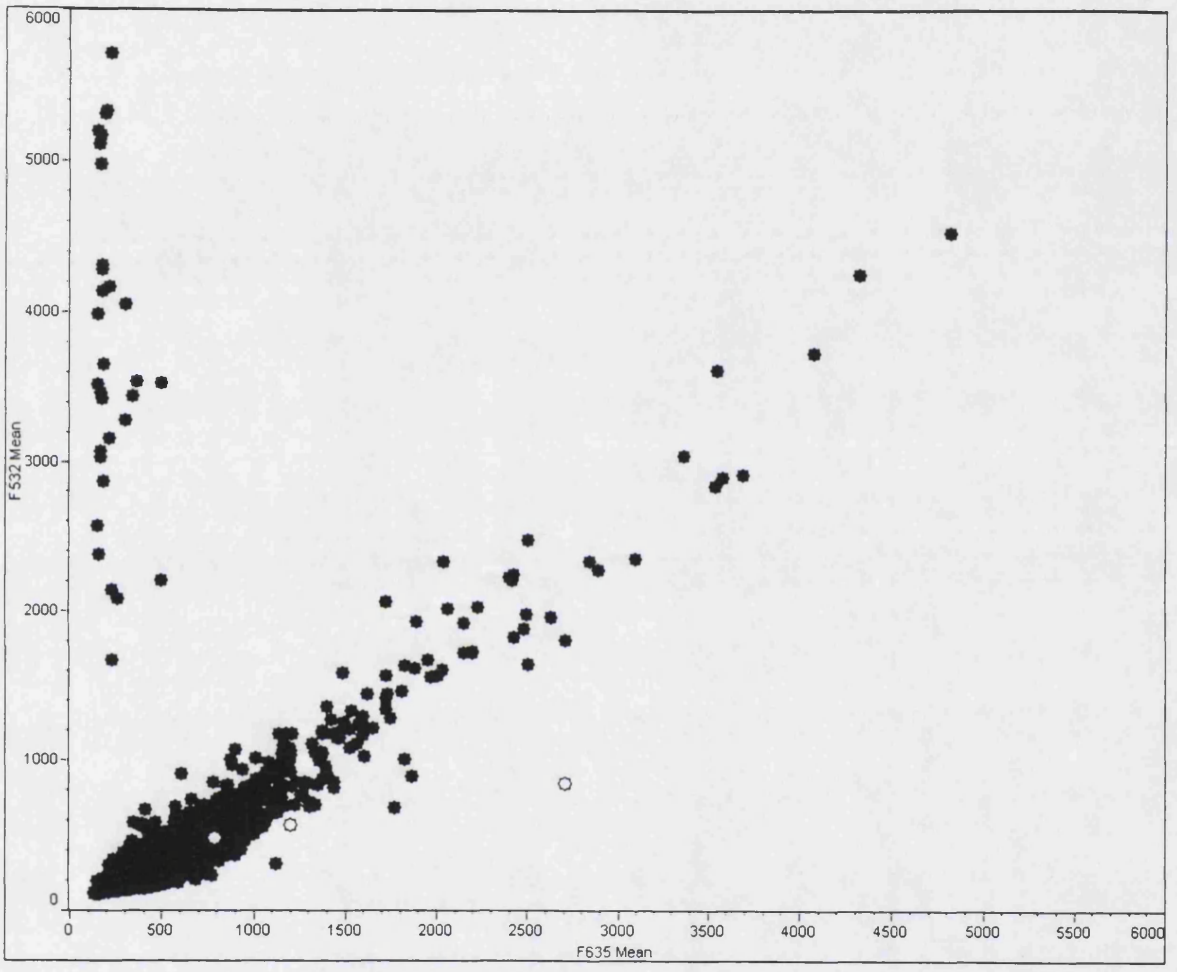


Figure 30 Scatter plot of microarray hybridisation from cDNA prepared from RNA harvested 6 hours after 0.2 M NaCl shock of *Mycobacterium tuberculosis* cultures

Plot of Cy5 fluorescence vs Cy3 fluorescence, spots shaded in light colour are the *gro* genes; from L-R *gro*EL1, *gro*ES (these two are present in an operon) and finally *gro*EL2.

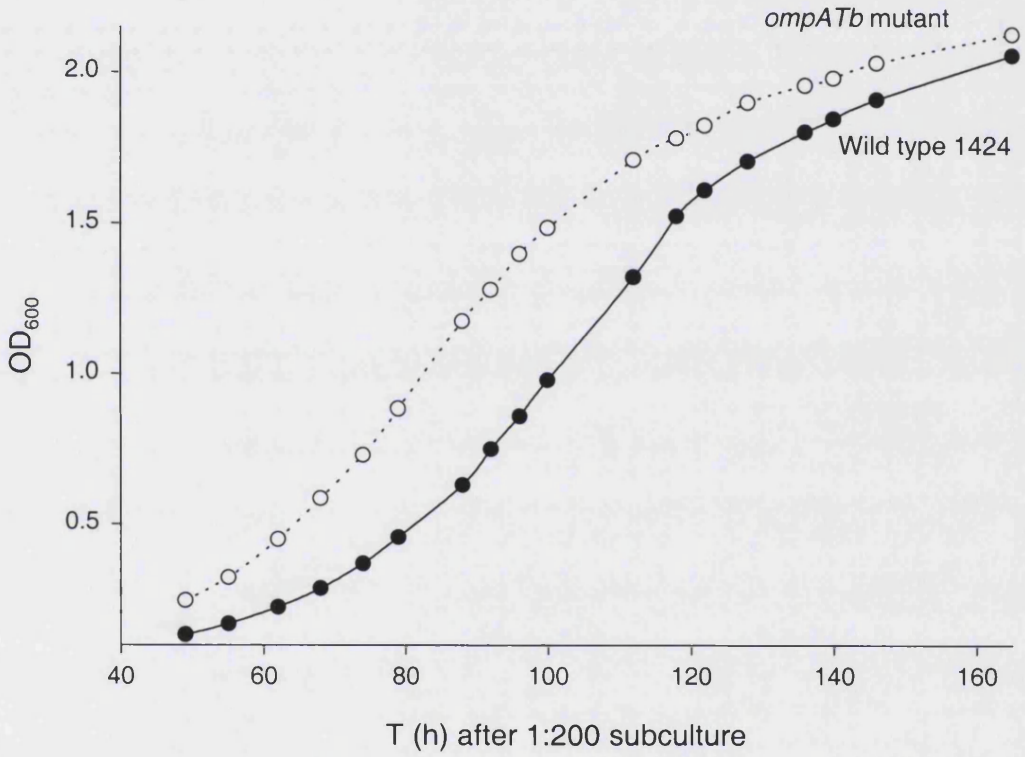


Figure 31 Growth of *Mycobacterium tuberculosis* 1424 vs Δ*ompATb* in Dubos broth

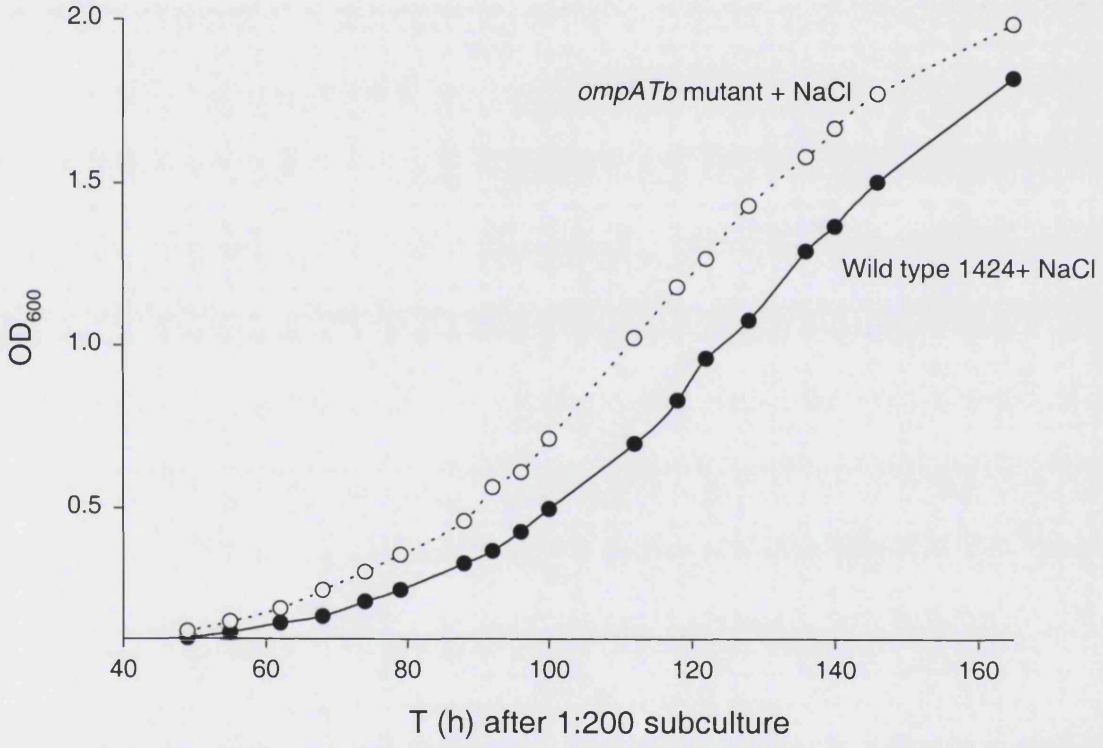


Figure 32 Growth of *Mycobacterium tuberculosis* 1424 vs Δ*ompATb* in Dubos broth with 0.2 M NaCl

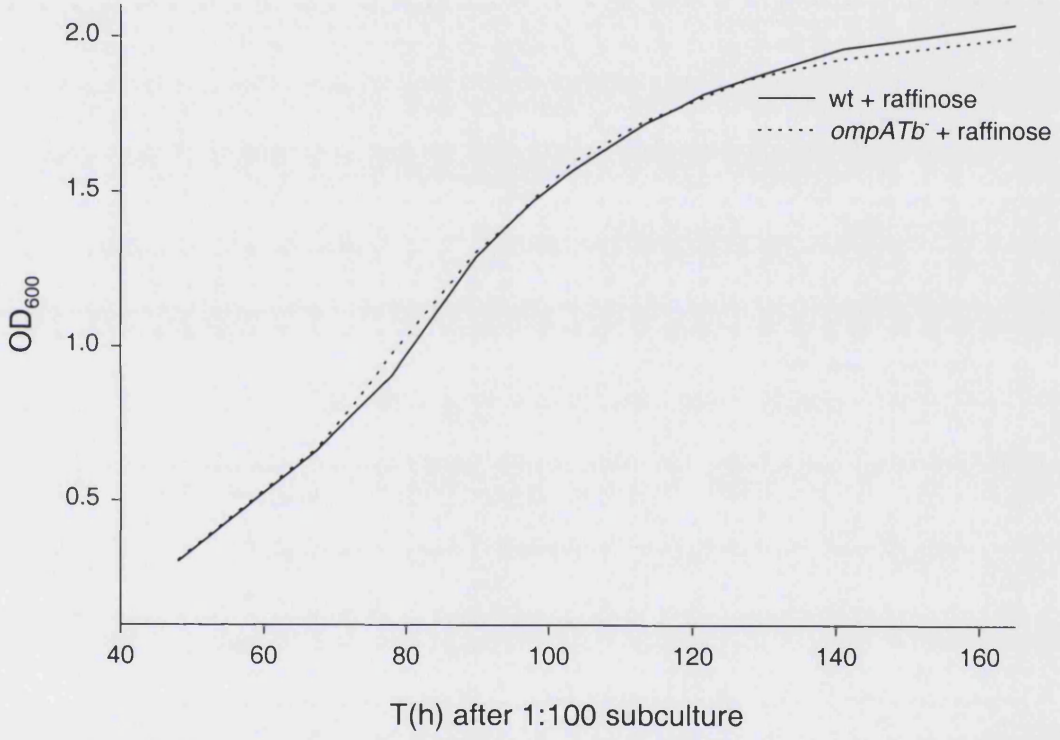


Figure 33 Growth of *Mycobacterium tuberculosis* 1424 vs Δ*ompATb* in Dubos broth with 0.1 M raffinose

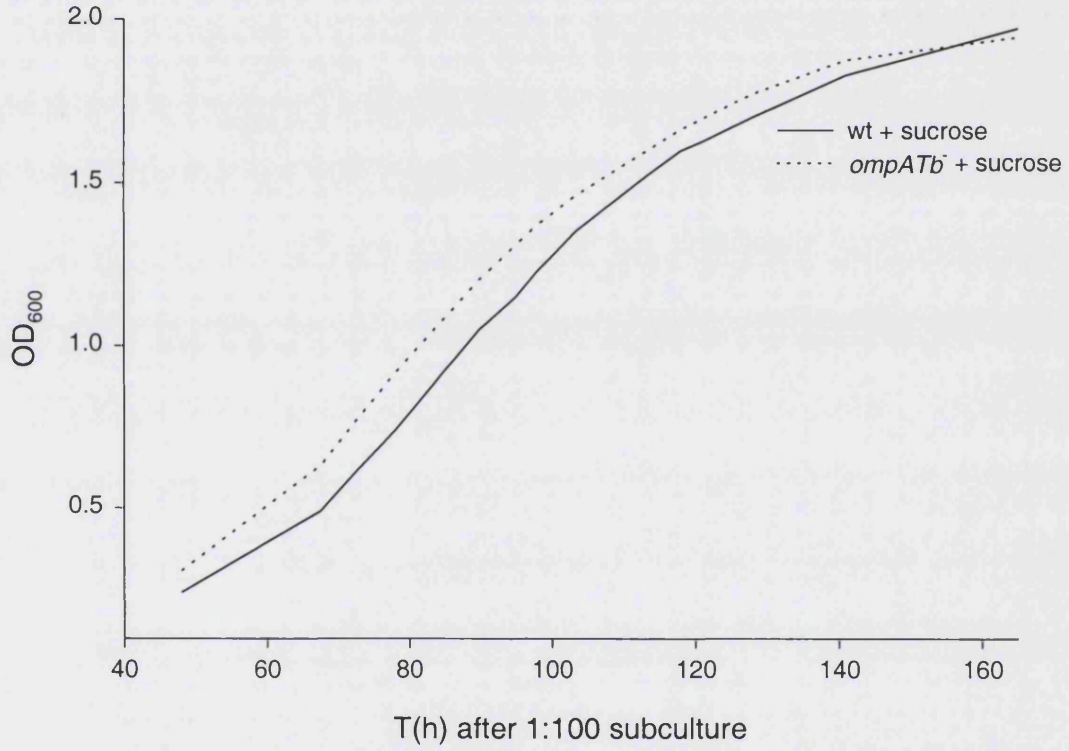


Figure 34 Growth of *Mycobacterium tuberculosis* 1424 vs Δ*ompATb* in Dubos broth with 0.2 M sucrose

4. Discussion

Overview

A number of novel findings with regard to mycobacterial porins have been made in this study; firstly it was shown that the porin *ompATb* was not solely confined to *Mycobacterium tuberculosis,* but was present in a number of other mycobacterial species. Interestingly the porin seemed to be present only in slow-growing members of the genus, although it was not diagnostic of such growth as *Mycobacterium kansasii*, a slow-growing species, did not appear to have the porin-encoding gene. This is discussed in section 4.1.

Secondly several regions upstream of the porin-encoding gene were fused to a *lacZ* reporter construct in a mycobacterial integrating vector. It was demonstrated that not only was the *Mycobacterium tuberculosis* promoter active in *Mycobacterium smegmatis*, but that the β-galactosidase activity of the cell lysate was dependent upon the length of the upstream region incorporated into the vector. The construct carrying the longest region upstream of the gene did not give the highest levels of β-galactosidase activity and activity was shown to decrease sharply when the promoter fragment became too short. β-galactosidase assays are discussed in section 4.2.

Thirdly the porin-encoding gene was expressed in an organism to which it was not native, *Mycobacterium smegmatis*, from its own promoter or from a high-level mycobacterial promoter. It was shown that at very high levels of expression of the porin encoding gene it was not possible to culture the transformed bacteria in liquid broth. Expression of the porin at lower levels led to bacteria which could be cultured and which appeared to grow at a very similar rate to the wild type *Mycobacterium smegmatis*. Inclusion of the putative regulatory apparatus in the expression construct had a minor, but observable, effect on growth rate. This is discussed in section 4.3.

Fourthly, it was shown that there was a regulation of the porin gene in response to osmotic stress; *Mycobacterium bovis* BCG was subjected to stress by the addition of sodium chloride to the growth medium and it was seen that levels of the porin gene transcript fell, according either to the time after the shock was introduced or to the

osmolarity of the medium. Attempts to produce similar results in *Mycobacterium tuberculosis* were inconclusive, the reasons for this are discussed later in section 4.4.

A putative regulatory system was also identified and shown to be present immediately downstream of the porin-encoding gene in the published genome of *Mycobacterium tuberculosis* (Cole *et al.*, 1998).

The regulatory component of this system, encoded by the gene Rv0903c, was cloned into an expression vector and the recombinant protein product was purified. This product was shown to bind to the mycobacterial DNA in a region upstream of the porin-encoding gene and also to a region internal to the porin-encoding gene. Furthermore this binding was shown to be enhanced by *in-vitro* phosphorylation of the regulatory protein using a low molecular weight phospho-donor (acetyl phosphate). This provided more evidence in favour of the protein being a regulator of the porin gene. Gel-retardation assays are discussed in section 4.5.

An attempt to knockout the porin-encoding gene using a *sacB* selection method was unsuccessful, however a vector using the dominant selectable marker $rpsL^+$ was transformed into *Mycobacterium tuberculosis* and the colonies which remained streptomycin resistant and became kanamycin resistant were shown to contain some isolates which had a disruption in the porin-encoding gene. A study of the growth of these mutant strains did not demonstrate a major effect on growth rate or susceptibility to osmotic pressures.

A suicide vector was constructed in an attempt to knockout the function of the regulatory gene Rv0903c, however this was unsuccessful despite using a system with two separate counterselectable markers (kanamycin resistance/sensitivity, sucrose resistance /sensitivity) and blue/white selection. A number of colonies which displayed the correct phenotype of kanamycin sensitivity, sucrose resistance and a white colour after a second crossover stage were obtained and subjected to PCR based screening but were shown to be identical to the wild-type. Gene knockout by homologous recombination is discussed in sections 4.6 and 4.7.

Microarray analysis of the Δ*ompATb* knockout strain, discussed in section 4.8, was unable to demonstrate that the gene was knocked out, since a sufficient portion remained to bind to the PCR product on the microarray slide. This was true for both genomic DNA

based microarray screening and for cDNA based screening which is dependent on the gene transcript being present. Screening of the cDNA from mycobacterial cultures which had been NaCl shocked did not show any marked up or down regulation of the porin gene, however other genes which were unrelated to the porin were shown to be affected.

Each of these statements shall now be discussed in turn before providing a general conclusion.

4.1 PCR-based screening of mycobacterial cultures for the *ompATb* gene and the Rv0903c gene

The results from the PCR based screen of mycobacterial DNA were somewhat surprising. PCR is a technique that relies on strict homology between the target DNA and the primers, which in this instance were originally designed to amplify the *ompATb* region from *Mycobacterium tuberculosis*. To achieve a PCR product from different species therefore means that not only is the gene present (indicating that the amplified gene is likely to be the same) but that the level of homology is sufficient to allow the primers to bind to the target DNA. As little as one base difference between the target and the primers, particularly at the 3' terminus, can cause the PCR reaction to fail.

A PCR product was obtained from all but one of the slow-growing species assayed in this way, but it is not acceptable to draw the conclusion that proteins related to the porin, *ompATb*, are present only in slow-growing bacteria for the reason detailed above. It may be that the PCR reaction is too specific to use as a screening method for the presence of homologous genes. It would therefore be more appropriate to use a Southern blot based method, in which the degree of homology to the *Mycobacterium tuberculosis* gene is less important. By a stepwise lowering of the hybridisation temperature at which a Southern blot is performed it is possible to identify genes with progressively less homology to the probe. A positive result in this experiment would identify genes that were homologous, but not necessarily identical, to the *ompATb* gene.

Unfortunately levels of genomic DNA available for the experiment were limiting and Southern blot analysis was only possible on genomic DNA from *Mycobacterium smegmatis*, *Mycobacterium chelonae* and *Mycobacterium microti*. In all cases the result from this assay was in agreement with the PCR based method. It must not be forgotten that the fast-growing mycobacteria have their own porins (Mukhopadhyay *et al.*, 1997;

Niederweis *et al.*, 1999; Trias and Benz, 1993; Trias *et al.*, 1992), so the lack of *ompATb* in these species does not mean the lack of envelope permeability altogether; such a situation would be detrimental to the cell, as the influx of nutrients would be limiting. It is also important not to make the assumption that the presence of the porin is responsible for the slow growth rate of the species in which it occurs, as many other factors are likely to govern this. Indeed presence of the porin would be hypothesised to increase the growth rate of the cell, as it would facilitate passage of molecules across the envelope (until a level at which the amount of porin became detrimental was reached).

PCR screening for the Rv0903c gene, identified as a putative regulator of the porin gene, revealed that it was present only in those species that had a highly homologous porin gene to *ompATb*. This result could mean that the regulatory gene is indeed important in regulation of the porin, or that perhaps the region of the genome which contains the porin gene and the regulatory gene is sufficiently small that the entire region has been conserved through evolution, possibly coming from a common ancestor. In this case the porin gene and the regulatory gene need not be related to each other, they have simply remained together due to their proximity on the chromosome.

A search of the genome of *Mycobacterium smegmatis*, the sequencing of which is currently underway, with the *Mycobacterium tuberculosis* Rv0903 protein sequence aganst translated DNA reveals a possible protein with 93 % identity to the tuberculosis protein (shotgun clone gms_181 as listed at http://www.tigr.org/). It is interesting that there is a protein with such homology in *Mycobacterium smegmatis*, it may be interesting to investigate whether this protein is able to interact with the porin gene from *Mycobacterium tuberculosis*. If so, this may explain why a more dramatic phenotype was not observed upon expression of the *Mycobacterium tuberculosis* porin in *Mycobacterium smegmatis*.

4.2 β-galactosidase assays for promoter activity

The results of the β-galactosidase assays are indicative that there is either some regulation of the porin gene or that DNA structure is important, it is seen that including the whole of the 2.2 kb region upstream of the porin gene, which contains the upstream binding site identified by gel retardation assays, does not give the most active promoter activity. The most active fragment is one that contains only 452 base pairs; when the length of the

upstream fragment is decreased to 178 bp the activity is reduced nearly 8-fold. Whilst the length of these fragments is large in relation to common promoters identified in *Escherichia coli*, the mycobacteria often have long and complicated regions upstream of genes which are involved with promoter activity. Examples of this occur in the *recA* gene (Movahedzadeh *et al.*, 1997) and the acetamidase gene (Parish *et al.*, 1997) although this is part of an operon.

It is possible that the 2.2 kb fragment adopts a secondary DNA structure, independent of any protein interactions, which prevents access of the transcriptional machinery. When the length of the fragment is made shorter there is more easy access for the RNA polymerase and the rest of the translation complex. Truncation of the fragment to 178 bp upstream of the start codon then disrupts any regions in the DNA which are important in binding either a component of the translation complex or an accessory protein.

It is also possible that a regulatory protein, speculated to be the product of gene Rv0903c, interacts with the DNA upstream of the porin such that transcription is affected. The results from the gel-retardation show that a site that binds the protein is present in the largest fragment used in the β-galactosidase assays. If this is the case then firstly the protein would appear to be a negative regulator of the gene, and secondly it would be necessary for the protein, or one of sufficient functional homology, to be present in *Mycobacterium smegmatis* as this is the model organism in which the assays were carried out. As already mentioned above a search of the incomplete *Mycobacterium smegmatis* genome reveals a protein with 93 % sequence homology to Rv0903. It would be interesting to investigate whether the promoter functioned in the same way if the constructs were incorporated into *Mycobacterium tuberculosis*. Whilst there is no reason to believe there would be a significant difference it may be that presence of a native transcription and regulation complex affects β-galactosidase activity.

4.3 Expression of the porin in *Mycobacterium smegmatis*

Expression of the porin in *Mycobacterium smegmatis* did not have a dramatic effect on growth; it was thought that expression of the porin in a foreign species host might have a profound effect on the viability of that species to survive or to tolerate stresses. Only in the case of high-level expression from a replicating vector was the effect so severe as to cause a growth perturbation. This may be because expression of a foreign protein at such

high levels is toxic anyway, and the fact that it is a porin is incidental, but the vector (pMV261) has been used before without causing any effects on growth and the empty vector has no detrimental effect at all. It can therefore be hypothesised that such high-level expression of the porin causes either a failure of the host cell to regulate its osmotic balance or that it interferes with the membrane to such an extent that it is no longer able to function as a barrier. Interestingly it was possible to obtain colonies from solid 7H11 agar plates immediately after transformation, but subsequent growth in liquid media was not possible. This suggests that the conditions imposed on the cell in liquid broth are more hostile that solid media, a recognised phenomenon as spheroplasts and difficult to grow mycobacteria are often more easily cultured on solid media.

Expression of the porin from the *hsp60* promoter at a single copy per cell led to cells which were viable and grew in liquid broth, however they appeared to be compromised in growth in comparison to the control cells carrying a pMV306 empty vector. This is possibly due to physical membrane disruption, as above, but may also be due to the fact that the porin is functional in *Mycobacterium smegmatis*. Expression at such a high level (far higher than would be encountered in the natural situation) is beyond the stress-response capabilities of the mycobacterial cell and the influx of toxic molecules, or the efflux of other molecules, may occur despite regulatory systems in the cell which are functioning to stop it. It has been shown that the porin functions in liposomes and in planar lipid bilayers to allow the passage of small molecules (Senaratne *et al.*, 1998). It also facilitates the passage of pyrazinamide across liposomes, and may allow the passage of the anti-tuberculous drug pyrazinamide across the envelope (Raynaud *et al.*, 1999). Possibly it is active in a similar fashion in the envelope of *Mycobacterium smegmatis* and, together with the porins that are present already in this organism the envelope reaches a 'threshold permeability' above which growth rate is affected.

Expression of the porin from its own promoter had very little effect; if the gradients of the growth curves are compared then it is seen that they are very similar. It is interesting, however, to note that *Mycobacterium smegmatis* transformed with pRAS20 and thus expressing not only the porin but also the regulatory system which is downstream of the porin encoding region, appears to be slightly impaired in growth. It can be seen that without any additional osmotic stress and also with 200 mM NaCl the curves for *Mycobacterium smegmatis* pRAS20 start in exactly the same place as those for the same

strain harbouring pRAS21 (porin alone) or pMV306 (empty vector control) yet they diverge whilst the pRAS21 and pMV306 transformed bacteria grow at almost identical rates. This phenomenon occurs in both sets of growth curves (figures 20 and 21).

It could be postulated that the presence of the regulatory genes downstream of the porin gene has some effect on the levels of porin in the cell. This is particularly true if the high-level expression of the porin from the *hsp60* promoter gave a true reduction in growth rate in the previous experiment, however if this were the case it would appear that presence of the regulatory machinery *increases* the level of the OmpATb in the cell and leads to a growth rate slower than in the wild type. This is counterintuitive, as it should be that the presence of the regulator allows the cell to respond to osmotic stress, and to reduce the levels of porin when appropriate. However it may be that there is a corresponding regulation of other pore-forming proteins and thus, as in *Escherichia coli*, it is the ratio of porins rather than the absolute amount of each which is important. Again complications are introduced into this experiment by using *Mycobacterium smegmatis* as a host organism as there could be cross talk between native regulators and the porin gene, or between the Rv0903c product and native proteins. It would be better to use a *Mycobacterium smegmatis* strain in which the native porins had been disrupted so they were no longer functional. This strain is not available at the present time, it is also possible that such a mutation may be lethal to the host. If a porin null mutant of *Mycobacterium smegmatis* is a viable strain it would serve as a more appropriate host for the *Mycobacterium tuberculosis* porin as any effects would be due to the inclusion of this foreign protein and not to any dynamic response to it. Of course this would not eliminate any action of the regulatory protein Rv0903 on genes in *Mycobacterium smegmatis*.

Since the porins of *Mycobacterium smegmatis* are not well characterised, and any regulators of these porins have not been investigated, the inclusion of a *Mycobacterium tuberculosis* porin was an attempt to generate a marked effect on the growth of the bacterium. This occured only upon very high level expression of the porin from a foreign promoter, expression of the porin from its own promoter or at a lower level from a foreign promoter (*hsp*60) did not have a profound effect on growth rate. The protein identified in *Mycobacterium smegmatis* with 93 % homology to Rv0903 (see section 4.1) is possibly important in this context.

4.4 Regulation of the porin as shown by real-time RT-PCR

Addition of osmotic stress to growing cultures of *Mycobacterium bovis* BCG showed clearly that addition of NaCl to the culture medium led to a decrease in levels of the porin gene transcript, in relation to the 'normalising' gene *gnd*. This result was repeatable and occurred depending not only on the level of the osmotic shock but also on the time after the shock was added.

Such a result is expected, as the bacteria are able to respond in many ways to environmental conditions that may otherwise be detrimental. An osmotic potential across the BCG envelope without a reduction in levels of the pore-forming protein may lead to either the influx of NaCl or the efflux of other solutes and water in order to try and stabilise the equilibrium. This demonstration of the regulation of the porin was reassuring as it suggests that there is indeed a system responsible for the dynamic alteration of porin levels in the outer envelope.

Addition of NaCl to the culture medium of *Mycobacterium tuberculosis* gave results that were not as easy to interpret; instead of seeing a clear drop in the level of porin transcript there appeared to be an increase over the 2 and 6 hour time points and then a drop down to normal levels after 24 hours. What was clear was that the levels of transcript in the absence or in the presence of salt rose in both cases with respect to time. Not only did they rise; the change was almost identical in each case (with or without NaCl). It is difficult to say what is happening here, it may be that there are other shocks being applied to the bacteria which lead to a change in porin transcript level or indeed (though less likely) a change in the level of *gnd* – a fall in *gnd* would make it appear that the level of porin is increasing even though it is really unchanged. Possible causes of this may relate to the way in which the bacteria were grown. The BCG cultures were grown in shaking flasks of Tween glutamate medium and were handled under normal laboratory conditions (category 2); however the *Mycobacterium tuberculosis* cultures were handled under containment level 3 conditions and were grown in roller-incubator bottles with Dubos medium. It is possible that the *Mycobacterium tuberculosis* cultures are shocked by their removal from the incubator, this involves both cooling the cultures (for longer than would have occurred with BCG, as the procedure for handling these dangerous pathogens is more time-consuming) and also opening the sealed bottles. It may be that the cultures become slightly oxygen-starved (anoxic) and that opening the bottles introduces fresh

oxygen, resulting in an up-regulation of the porin levels, also CO_2 levels would rise in a sealed container and this may affect porin levels. This was not a problem with BCG as it could be grown in flasks with foam bungs and thus gas-exchange would not become limiting. It could also be that taking the cultures from the rolling incubator, where they have been for a number of days or weeks, and subjecting them to the mechanical shock of shaking them and leaving them static whilst the samples are taken induces a response from the bacteria. All of these explanations are speculative yet there must be something that happens to the bacteria during the sampling process that affects either porin or *gnd* levels.

It may alternatively be that porin levels are dynamic during the growth cycle, and the change that is seen reflects a dynamic alteration of porin levels as growth moves towards the stationary phase. The growth rates of BCG and *Mycobacterium tuberculosis* are similar yet the increase in levels of the porin was seen only in *Mycobacterium tuberculosis* and not in BCG. This does not explain why the levels of porin gene transcript change in cultures that have not been subjected to any NaCl addition to the same extent as those which have received NaCl, nor does it explain why the change is apparently so dramatic over what is, in terms of mycobacterial growth, a short period of time. It should also be noted that whilst BCG and *Mycobacterium tuberculosis* are members of the same taxonomic complex, there are large differences at the genetic level between the two species. In particular BCG has lost a number of genetic regions which appear to have remained in *Mycobacterium tuberculosis* (Behr *et al.*, 1999).

There is some evidence that levels of *gnd* are not entirely static, but that they do stay very similar throughout the growth cycle and also upon the influence of DNA damaging shock to the culture (personal communication, Dr E Davis, NIMR). Unfortunately *gnd* is the best option for a normalising gene that is available as others have been assayed and shown to change to a much greater extent. If for some reason the addition of NaCl to *Mycobacterium tuberculosis* cultures led to a change in the level of *gnd*, then these results become even more difficult to interpret.

The results of sucrose stress on both BCG and on *Mycobacterium tuberculosis* were not as definitive, no clear trend upon the addition of sucrose was seen despite the fact that this must be exerting an osmotic stress on the cells. The osmotic shock resulting from NaCl addition may be different to that exerted by sucrose; NaCl dissociates into ions and is

very much smaller than sucrose, thus its passage across the envelope is not expected to be subject to the same resistance as that of sucrose. It was also shown that diffusion of sucrose through the porin in reconstituted liposomes was slow (Senaratne *et al.*, 1998)

The TaqMan procedure, whilst being incredibly sensitive and specific, is also prone to erroneous results due to pipetting errors which become exponentially amplified during the PCR procedure. Moreover, the analysis of the results uses a standard curve that plots the logarithm of the standards used, and thus actual results obtained are the results of an inverse log. This means that small differences in the cycle threshold (the point at which fluorescence reaches a threshold level) can be magnified by the mathematical manipulation (in order to relate the template to the standards). Nevertheless the results for the standards are seen to lie accurately on a straight line when $\log_{10}$ of the standards is plotted against average cycle threshold, indicating that the technique is reliable.

Other porins have been identified in *Mycobacterium tuberculosis* (Kartmann *et al.*, 1999), and in BCG (Lichtinger *et al.*, 1999); these may be more important than OmpATb in the response of these organisms to NaCl or to sucrose. It is possible that whilst there is an apparent up regulation of the porin gene *ompATb* in response to NaCl in *Mycobacterium tuberculosis* and a down regulation in BCG, there is a corresponding change in the levels of the other porins identified by Kartmann *et al.*, (1999) or Lichtinger *et al.*, (1999). In this case the results obtained in the study of *ompATb* may well be correct but, in a system analogous to that of *Escherichia coli*, there may be an alteration in the *ratio* of the porin proteins meaning that as one porin is up-regulated another is down regulated dependent upon the size of the pore. More replication and investigation of this phenomenon is necessary in order to determine whether this is true.

The real-time RT-PCR technique has a number of limitations, firstly it is dependent upon very small (pg) levels of input cDNA and thus is open to variability or contamination, and secondly the results are expressed as a ratio to normalising gene *gnd*. Levels of this gene may not remain constant and thus results could be unreliable.

4.5 Gel retardation assays using recombinant Rv0903 protein

The results from this experiment showed two clear regions of binding for the Rv0903 protein to the region around and internal to the *ompATb* gene. The upstream region, lying between bases –1781 and –1331 relative to the transcription start point seems to be a long

way from the gene. However, as already mentioned mycobacterial promoter regions (and associated activator/repressor regions) are often long and complex. The *Escherichia coli* porin *ompF* has a site far upstream (384-354 bp), which is involved in regulation (Huang *et al.*, 1994; Huang and Igo, 1996); this may be analogous to the situation in *Mycobacterium tuberculosis*. It has already been shown (Section 1.13) that the regulation of porins in *Escherichia coli* is a very complex regulatory system which is dependent upon occupation of several binding sites, each of which has a different affinity for the OmpR protein. A similar situation may exist in *Mycobacterium tuberculosis* and the two sites identified may not be the only sites that bind the recombinant Rv0903 protein. Indeed if this protein regulates not only the porin gene but also other genes, as may be postulated due to the lower number of two-component systems in *Mycobacterium tuberculosis*, there may be one or several other binding sites which are unrelated to porin expression.

The presence of a binding site within the *ompATb* gene is unusual but not unique; the *Escherichia coli* Rns protein, a virulence regulator, has binding sites both upstream and downstream of its own promoter region (Munson and Scott, 2000) and the *Escherichia coli* genes *phoA* and *pstS* have internal binding sites for PhoP (Liu *et al.*, 1998). Binding sites internal to the gene that is being regulated are usually, but not always, the targets of proteins that repress rather than activate the gene. Together with the results for the β-galactosidase assays the intriguing possibility is raised that the Rv0903 protein is a transcriptional repressor and not an activator. This hypothesis could be supported by the growth curves of *Mycobacterium smegmatis*, which show an apparent decrease in growth rate when the porin along with the regulatory machinery is expressed; there is a possibility that the *Mycobacterium tuberculosis* Rv0903 protein serves to repress the transcription of native *Mycobacterium smegmatis* genes. The effect that is observed was not due to any effect on the exogenous OmpATb as the strain transformed with only the porin was able to grow as well as the wild type.

DNAseI protection experiments to try and narrow down the binding site to less than the number of base pairs achieved with gel-retardation were performed but were not successful despite several different methods being employed. It was not that the protein did not bind to the fragment used in footprinting, but rather that a satisfactory digest of the fragment could not be obtained or that there was digest of the probe in the absence of

DNase. Attempts to minimise this included preparing a probe from a PCR product (end-labelling with γ^{32}P ATP) or by excision from a vector and a fill-in reaction of a 5' overhang (using α^{32}P dATP). Despite preparing entirely fresh stock solutions it was not possible to obtain a probe that showed no digestion in the absence of DNAse, it was therefore impossible to determine regions of protein binding which protect this probe from DNAse digestion.

The presence of two binding sites raises the possibility that there is a secondary structure of the DNA such that the promoter region becomes inaccessible at times when expression of the porin is not advantageous; one such structure is a hairpin loop as detailed in *Escherichia coli* for regulation of the *ompF* gene (Pratt *et al.*, 1996). If this situation existed then a monomer, or more likely a dimer, of OmpR could contact both sites and pull the DNA together, restricting the promoter region. This is supported by the β-galactosidase assays, particularly if a *Mycobacterium smegmatis* protein is able to fulfil the same role as Rv0903, which show that the longest fragment is not the most active. Removal of one of the binding sites leads to an increase in β-galactosidase activity.

4.5.1 Phosphorylation of Rv0903 protein enhances binding affinity

Phosphorylation of regulator proteins *in vitro* with low molecular weight phospho-donors has been performed before (Deretic *et al.*, 1992; Lukat *et al.*, 1992), and levels of intracellular acetyl phosphate have been implicated in phosphorylation of a number of *Escherichia coli* response regulators (McCleary and Stock, 1994). More recently the pH of the external media has been related to the intracellular acetyl phosphate levels and has been shown to affect the levels of porin gene expression in *Escherichia coli* (Heyde *et al.*, 2000).

It was therefore interesting to observe that phosphorylation of the *Mycobacterium tuberculosis* Rv0903 protein increased its binding affinity to the DNA target. This result supports the theory that the protein may bind as a multimer in a situation homologous to that which occurs in *Escherichia coli*, where multiple binding sites are occupied upon phosphorylation of OmpR. No investigation was performed on the binding affinity of phosphorylated Rv0903 for itself, though this would further support the importance of multiple Rv0903 proteins in gene regulation.

The use of low molecular weight phosphodonors is not the only way in which regulator proteins can become phosphorylated *in vitro*; it is also possible for them to use the kinase component of the 2-component system, or even a kinase from a different 2-component system (Deretic *et al.*, 1992; Via *et al.*, 1996). The importance of this cross-talk between non-cognate sensors and regulators is discussed later in section 4.7.

Using protein alignment software (MegAlign, DNAStar/Lasergene) it is possible to identify a site on the Rv0903 protein that is likely to serve as the phospho-receiver. Alignment with the proteins characterised by Via *et al* (Via *et al.*, 1996) suggests that the aspartate residue which becomes phosphorylated in Rv0903 protein is residue 61 (see figure 11). A similar approach to the sensor protein, Rv0902, reveals that the histidine at position 241 (see figure 12) is the likely residue to donate the phospho- group.

Future experiments could involve a modification of the aspartate at residue 61 such that it was no longer able to receive a phospho-group from Rv0902; this could be achieved by site directed mutagenesis of a cloned copy of Rv0903c in order to change the codon from GAT (coding for asparatate in Rv0903c) to GAA, which would code for glutamate. This could be introduced into the mycobacterial genome by homologous recombination; however it would necessitate the screening and sequencing of many colonies as the difference between the recombinant strain and the wild type is so small. It would not be possible to use Southern blotting to determine a genetic difference between the wild type and the homologous recombinant in this instance.

4.6 Gene knockout of *ompATb* using pRAS5 and pOmpA-aph-rpsL tb

The disruption of genes in *Mycobacterium tuberculosis*, and indeed the fast growing mycobacteria, has always been difficult to achieve (Balasubramanian *et al.*, 1996; Bardarov *et al.*, 1997; Guilhot *et al.*, 1994; Jackson *et al.*, 1999; Kalpana *et al.*, 1991; Pelicic *et al.*, 1997; Pelicic *et al.*, 1996; Rubin *et al.*, 1999). However recent advances in the use of multiple counter-selectable markers have made it easier to select bacterial isolates that have undergone recombination (Hinds *et al.*, 1999; Parish and Stoker, 2000; Parish and Stoker, 2000; Pavelka and Jacobs, 1999; Pelicic *et al.*, 1996; Reyrat *et al.*, 1998).

Using the *sacB* counterselectable marker (Pelicic *et al.*, 1996; Pelicic *et al.*, 1996), and a kanamycin resistance cassette internal to the coding sequence of the gene (plasmid pRAS5) it was not possible to obtain a knockout of *ompATb*. The probable reason for this is that it is not uncommon for spontaneous resistance to kanamycin to occur in the mycobacteria if antibiotic selection pressure is applied, and the *sacB* gene is notoriously unstable in terms of providing sensitivity to sucrose (personal communication, Dr KG Papavinasasundaram, NIMR). Since a one-stage knockout procedure was attempted by plating directly onto sucrose/kanamycin 7H11 agar, the colonies that arose were simply resistant mutants and not homologous recombinants.

A previous attempt to knockout the *ompATb* gene using a very similar system, but with a gentamycin resistance cassette instead of a kanamycin resistance cassette internal to the gene also failed (R. Senaratne, PhD thesis, 1999, NIMR). In this instance the gentamycin resistance cassette failed to provide an adequate selection pressure.

Using a different approach it was attempted to knockout the porin gene in *Mycobacterium tuberculosis* 1424, a point-mutated derivative of *Mycobacterium tuberculosis* H37Rv that is streptomycin resistant. If the mutated native *rpsL* gene is complemented with the wild-type *rpsL* then the streptomycin resistance is lost (Sander *et al.*, 1995). By providing the $rpsL^+$ gene on a plasmid, along with the porin gene disrupted by a kanamycin resistance cassette (plasmid pOmpA-aph-rpsL tb), it was possible to select in a one-step procedure for colonies that grew on 7H11 agar with kanamycin and streptomycin.

Of 8 colonies selected in this way 2 were single crossovers (or random integrants), 2 were wild type and 4 appeared to be mutants by PCR based screening methods. The reason that apparently wild type and single crossover colonies are selected is probably due to the accumulation of kanamycin resistance or failure of the wild-type *rpsL* gene to counterselect properly, perhaps due to mutation. The genotype of the mutant strain was confirmed by Southern blotting, comparing results to the wild type (section 3.7).

Disruption of the porin encoding gene *ompATb* did not seem to have any serious effect on the viability or growth rate of the mutant strain. This may be indicative of the presence of other porins (as identified by Kartmann *et al.*, (1999)) or may simply be due to the fact that the 'correct' selection pressure has not been applied in order to see a dramatic change in phenotype. Sucrose and raffinose were used as examples of a small and a large sugar

respectively which have different permeability coefficients through the OmpATb pore. This has previously been shown by Senaratne *et al.*, (1998) in which diffusion of these sugars into reconsituted liposomes was measured; it was shown that whilst raffinose was not able to diffuse throught the OmpATb pore, sucrose was able to. NaCl was used as it has been previously shown to cause a fall in levels of *ompATb* transcript in BCG.

Porin-deficient mutants of *Escherichia coli* are able to compensate for the loss of the pore forming proteins and remain viable even if both the major porins *ompC* and *ompF* are disrupted (Saint *et al.*, 1993). Furthermore loss of the porins can make the bacteria resistant to compounds which would normally be harmful, such as silver salts (Li *et al.*, 1997). It must be noted that *Escherichia coli* appears to be different to *Mycobacterium tuberculosis* in that there are several characterised outer membrane porins and disruption of the major porins may simply lead to a compensatory increase in the more minor porins; the complexity of adaptation to environmental stress in relation to the porins is discussed in Zhang and Ferenci, (1999). Whilst other porins do appear to have been identified in *Mycobacterium tuberculosis*, they are present at very low levels in the outer membrane and are not as numerous as those of *Escherichia coli*. It was not possible to identify any proteins in *Mycobacterium tuberculosis* that showed homology to OmpF or OmpC of *Escherichia coli* and attempts to purify pore-forming proteins from mycobacterial cell envelopes do not yield high levels of porin, in contrast to *Escherichia coli* in which porins are abundant.

It will be interesting to see how the knockout of the porin gene affects the immunogenicity or pathogenicity of the mutant strain; it has been shown that the outer membrane porin OmpC of *Salmonella typhimurium* is involved in adherence to macrophages (Negm and Pistole, 1999) and that the major outer membrane protein of *Legionella pneumophila* is involved in complement-mediated immunity (Bellinger-Kawahara and Horwitz, 1990). It is therefore possible that the porin-deficient mutant of *Mycobacterium tuberculosis* will be less viable *in vivo*. In a separate line of study by another researcher the porin deficient strain has been inoculated into mice in order to investigate this.

Another study identifies the OmpA porin of *Escherichia coli* as being important in killing of the cells by neutrophil elastase (Belaaouaj *et al.*, 2000), showing that an *ompA* mutant is not degraded after being engulfed by neutrophils and subjected to neutrophil elastase.

The suggestion from this publication is that OmpA is important in maintaining the structural integrity of the cell, and when it is degraded by neutrophil elastase there is a breakdown of the cell membrane; interpretation of this is difficult however, as an ompA deficient strain of *Escherichia coli* remains viable. It is possible that a product released from ompA upon attack by neutrophil elastase is toxic to the *Escherichia coli* cell. The effect of neutrophils on *Mycobacterium tuberculosis* is not certain and is the subject of some debate (Denis, 1991; Jones, Amirault and Andersen, 1990)

The system used to generate the Δ*ompATb* mutant has one major drawback: the use of a streptomycin resistant host strain. This makes any future experiments in animals a greater potential safety risk and also means that any subsequent mutations become limited in the choice of antibiotic resistance cassette, particularly since the original mutation carries kanamycin resistance. The use of an unmarked mutation strategy is more amenable to the creation of further mutations in the same host strain.

4.7 Gene knockout of Rv0903c using pRAS18

Attempts to knock out the function of the regulatory gene were not successful, despite using a strategy that involves at least 3 selectable markers and has been shown to work previously (Parish and Stoker, 2000; Parish and Stoker, 2000). The reason for this is most likely to be that the regulatory gene Rv0903c is essential for the survival of the organism.

The reason that colonies with the correct phenotype for the mutant (kanamycion sensitivity, sucrose resistance and a white colour on X-gal agar) are isolated at the second stage is due to the way in which the knockout is generated. In an attempt to create a gene knockout that was not antibiotic resistant, and thus was more like the wild type, no antibiotic resistance cassette was introduced into the coding sequence of the gene. The knockout depended only on the removal of bases from the gene, and not their replacement with a resistance cassette. While this is beneficial for later experimentation since there is no background antibiotic resistance, it also means that should the second crossover occur within the same flanking region of homology, the result is identical to the wild type host strain.

Further attempts to generate a knockout of this strain should use a marked mutation (gentamycin is a suitable choice), which adds a selection pressure against the second crossover reverting back to wild type. If, as is suspected since nearly 60 colonies with the

correct phenotype were screened, the knockout is lethal then it will be necessary to complement the single crossover (kanamycin resistant, blue colonies on 7H11 agar with X-gal) with an integrated copy of the wild type Rv0903c gene at a site distant from the integration. This has the effect of making the single crossover merodiploid for the gene of interest and allows a second crossover to occur at the target site whilst still maintaining the functionality of the target gene. This technique has been used in mycobacteria to demonstrate the essential nature of *glnE*, a regulator of glutamine synthetase (Parish and Stoker, 2000), *mtrA* (a two-component response regulator) (Zahrt and Deretic, 2000) and *whmD* (a protein required for septum formation) (Gomez and Bishai, 2000). Demonstration that the regulatory *glnE* and *mtrA* genes are essential further supports the theory that the regulatory gene Rv0903c is likely to be necessary for viability, thus far there has been no instance of a regulatory gene in *Mycobacterium tuberculosis* being non-essential. Other genes, such as isocitrate lyase, have been shown to be essential only under certain conditions such as persistence in macrophages (McKinney *et al.*, 2000).

The paucity of two-component systems in *Mycobacterium tuberculosis* may suggest that each of the systems controls more than one target gene or operon. In this case attempts to disrupt the regulatory component are likely to meet with resistance as even if the gene that is studied (in this case the porin) is not essential, as other targets of the same regulatory protein may be.

In *Escherichia coli* the OmpR protein has been identified as a regulator of a gene which is functionally unrelated to the porin genes; a point mutation in the *ompR* gene leads to an ability to form biofilms due to increased curli expression (Vidal *et al.*, 1998), thus even in *Escherichia coli* (where two-component systems are more numerous than in *Mycobacterium tuberculosis*) there is economy in the regulator proteins. There has recently been a suggestion that the lack of his-asp two-component systems in *Mycobacterium tuberculosis* is compensated for by eukaryotic-like serine/threonine protein kinases (Av-Gay and Everett, 2000). This would make the necessity of the conventional two-component systems having more than one target smaller.

Inactivation of the *ompR* gene in *Salmonella typhimurium* by transposon mutagenesis led to the intriguing finding that *ompR* mutants are attenuated *in vivo* in BALB/c mice (Dorman *et al.*, 1989). This is probably due to the effect that such a mutation has on porin

expression and may be related to the same effect reported by Negm and Pistole (1999) who showed that the OmpC porin mediates adherence to macrophages.

Whilst it is unlikely, it is also possible that the Rv0903 protein is involved in alternative signalling pathways to those originating with Rv0902. Such cross-talk of two-component systems has been demonstrated in *Pseudomonas* using CheA to phosphorylate AlgR *in vitro* by Deretic *et al* Deretic *et al.*, (1992) and also *in vivo* (in *Escherichia coli*) showing that EnvZ can phosphorylate ArcB (Matsubara *et al.*, 2000). This raises the possibility that an Rv0903c knockout could have multiple targets and thus the suggestion that it is lethal is a plausible one.

4.8 Microarray analysis of *Mycobacterium tuberculosis*

The aims of the microarray analysis were threefold: to determine if there was still hybridisation of the genomic DNA from the Δ*ompATb* mutant to the PCR product from the porin gene present on the slide, to determine if loss of the porin gene resulted in an upregulation of other (compensatory) genes, and to determine if any genes other than the porin *ompATb* altered their expression upon the addition of NaCl shock.

The genomic DNA from the Δ*ompATb* mutant contained a region of sufficient homology to the wild type H37Rv porin gene that there was no observable difference in binding of the mutant. The 212 bp region of homology (Section 3.11) demonstrates that even though the porin gene is disrupted (as shown by Southern blot and by PCR), a microarray cannot detect a mutant constructed in this way. Similarly the RNA from the mutant must have contained a transcript with sufficient homology to the wild type for binding of the cDNA to occur. This is a limitation of the microarray technique and a pitfall to be aware of when using this technique to evaluate mutants. It was not possible to identify any genes that altered in their expression when the porin gene *ompATb* was knocked out, this may be because there needs to be an addition of stress to the cultures in order to induce a phenotype, or perhaps that the tecnique of microarray analysis is not sensitive enough to detect small changes in expression of other genes which may compensate for the lack of a functional *ompATb*.

The addition of NaCl to *Mycobacterium tuberculosis* cultures had a more observable effect in terms of microarray analysis, with a number of genes obviously upregulated in response to this osmotic shock. Most notable of these was *groEL2*. *gro*ES and *gro*EL1

exist in an operon and are usually upregulated under conditions of stress to the cell, however groEL2, whilst being similar at the genetic level to *gro*EL1, is not part of the same operon (Kong *et al.*, 1993; Rinke de Wit *et al.*, 1992). An upregulation was observed in all three genes (*gro*EL1, *gro*EL2 and *gro*ES) however that in *gro*EL2, the gene which is not part of the operon, was seen to be much more pronounced (see figure 30).

Unfortunately it was not possible to see any difference in the levels of the porin gene transcript using a microarray. However this may be a limitation of the technique rather than an accurate reflection of the situation. Microarrays have an advantage in that it is possible to screen almost every gene in the *Mycobacterium tuberculosis* genome on a single slide; this is in contrast to real-time RT-PCR, which can analyse just one gene at a time. However the sensitivity of real-time RT-PCR is many fold higher than the microarrays which depend on at least a 2-fold difference in gene expression in order to observe any result. In this instance the microarrays agreed with the real-time RT-PCR result and showed no difference in *ompATb* expression; for other genes the limit of resolution of the microarray technique may be more important.

Another difficulty was encountered in the microarray technique as the only slides that were available at the time were poorly printed and did not facilitate examination of every gene. The PCR products on the slide were not all able to be identified; therefore there may be other genes up- or down-regulated upon NaCl shock that were not detected in this experiment.

Microarray analysis is an extremely useful tool, and would have been very interesting to apply to a Rv0903c mutant which would be hypothesised to have a number of changes in gene expression, however in this case for the purposes of examination of the porin gene and the Δ*ompATb* knockout its use is limited. This need not necessarily be the case, as selection of primers to produce a PCR product that would hybridise to the wild type but not to the knockout would certainly be possible.

5. General conclusion and future perspectives

This investigation of the porin OmpATb from *Mycobacterium tuberculosis* has demonstrated several features which were previously unknown, but further study needs to be undertaken in some areas: firstly, it needs to be determined whether Rv0903 protein in *Mycobacterium tuberculosis* does actually serve to regulate the porin. This has not been demonstrated directly, although a number of pieces of evidence point towards it. Foremost in this investigation should be continued attempts to generate a gene knockout of Rv0903c. Should this knockout be obtained then porin levels could easily be investigated by real-time RT-PCR in the presence or absence of the regulatory gene. Alternatively it could be shown by the use of a merodiploid strain that without complementation of the Rv0903c gene in a single crossover strain it is not possible to obtain a double crossover (homologous recombinant).

Whilst it is tempting to draw analogies to *Escherichia coli* it must also be remembered that there are many differences between these two organisms. Levels of homology at the nucleotide level between *Escherichia coli* genes and their *Mycobacterium tuberculosis* counterparts are not strong, yet functional characteristics are often shared. There are important distinctions between the known porin regulation system in *Escherichia coli* and that proposed in *Mycobacterium tuberculosis*, for example little is known about the regulation of OmpA in *Escherichia coli* whereas the regulation of the porins OmpF and OmpC has been the focus of much study. Parallels drawn between the regulation of the mycobacterial *ompATb* and *Escherichia coli ompF* and *ompC* are therefore largely based on the function of the gene products, particularly as *Mycobacterium tuberculosis* is observed to have a very low level of porin in the outer envelope. It is important not to stretch this analogy too far.

The regulatory gene Rv0903c and its protein product are interesting subjects in their own right, regardless of whether regulation of the porin is directly dependent on them, and further study should characterise any other possible targets of this protein. As mentioned previously the site of phosphorylation of both the sensor (Rv0902) and the regulator (Rv0903) could be investigated by site-directed mutagenesis, with the target aspartate and histidine already having being suggested in section 4.5.1.

DNAse protection footprinting to narrow down the binding site of the Rv0903 protein is important, as this may identify a binding motif, which in conjunction with the published

genome could be used to identify likely targets for the regulator. Attempts to use this have so far been unsucessful, however with persistence and careful alteration of reaction conditions it should be possible to demonstrate a specific site of binding. The ability to perform *in silico* analysis of the *Mycobacterium tuberculosis* genome has greatly facilitated study of this organism.

Further study of the Δ*ompATb* knockout is necessary in order to try and identify a phenotype for this strain. The stresses used in this study (sucrose, raffinose and NaCl) did not seem to have an obvious phenotype, however it may be that the mutant bacteria are sensitive to other pressures such as antibiotics, heat shock, oxygen starvation or nutrient limitation. Inoculation of the porin deficient mutant strain into mice is underway and results from this may show that the pathogenicity of the mutant is different to that of the wild type. This could suggest that OmpATb is important in the macrophage or neutrophil elastase mediated response as it is in *Escherichia coli* or *Salmonella*, however there are caveats associated with the neutrophil elastase mediated response as detailed in section 4.6.

It has recently been reported that OmpA from *Escherichia coli* does not have a pore-forming ability (Pautsch and Schulz, 1998); however there are a number of anomalies within this investigation. It has been shown clearly that OmpA from *Escherichia coli* can form pores (Arora *et al.*, 2000; Sugawara and Nikaido, 1994) and it must also be remembered that the protein used by Pautsch and Schulz (1998) was heavily truncated and engineered at several positions in order to obtain crystals. Nevertheless there is controversy as to the exact function of OmpA, and consequently of OmpATb, and it seems likely that the amount of OmpATb present in the mycobacterial envelope in an active form is much less than that which is present in an inactive form. This concept was suggested upon the initial characterisation of OmpATb by Senaratne *et al.*, (1998). The mycobacterial cell envelope is notoriously difficult to extract cleanly using biochemical methods; however study of the recombinant OmpATb protein may be able to further characterise the different states of this porin.

In summary this work represents the beginning of study into the mycobacterial porin and its regulation. Whilst conclusions may be drawn from the work contained in this thesis there are many answers left to find. The mycobacteria are fascinating, yet difficult, subjects of research which present many difficulties and offer many challenges. With patience *ompATb* need not be such an enigma.

Acknowledgements

This work would not have been possible without the tireless and enthusiastic support of the staff of the mycobacterial research division here at the National Institute for Medical Research. In particular I thank my supervisor Dr Philip Draper, who has provided remarkable insight into all matters, mycobacterial and otherwise, throughout my years of study and has certainly been an inspiration to me. I am forever indebted to Dr KG Papavinasasundaram for his help; his knowledge of molecular biology and his willingness to assist at any time were greatly appreciated. I would also like to thank Drs Elaine Davis and Peter Jenner for their assistance and input throughout the project, and for sharing their thoughts and ideas, without which progress would have been far slower. I am also grateful to Dr B Gopal for assistance with the purification of recombinant Rv0903 protein. Of course thanks also to Dr Jo Colston for the opportunity to share a department with some fantastic minds and some marvellous co-workers.

I wish to thank Dr Ryan Senaratne for his patience whilst sharing his lab space during his final PhD year, and also Dr Gili Bachrach for providing an entertaining perspective on molecular biology. I also wish to acknowledge Dr Huw Williams, who started me on the path of mycobacterial research at Imperial College, London.

I would like to thank my friends who have always been around throughout my studies and who have been willing to counsel, advise or listen; when the going got tough, the tough sometimes needed pushing. Thank you in particular Abigail, Christian, Felicity and Rebecca for your valued friendship.

Last but most certainly not least I thank my parents and family for unfailing support and encouragement throughout my time as a PhD student. Their confidence and belief has, I hope, not been misplaced and has been greatly appreciated over the years.

References

Aiba, H., Nakasai, F., Mizushima, S. and Mizuno, T. (1989). Evidence for the physiological importance of the phosphotransfer between the two regulatory components, EnvZ and OmpR, in osmoregulation in *Eschericia coli. Journal of Biological Chemistry* **264**, 14090-14094.

Ames, S. K., Frankema, N. and Kenney, L. J. (1999). C-terminal DNA binding stimulates N-terminal phosphorylation of the outer membrane protein regulator OmpR from *Escherichia coli. Proceedings of the National Academy of Sciences USA* **96**, 11792-11797.

Arora, A., Rinehart, D., Szabo, G. and Tamm, L. K. (2000). Refolded outer membrane protein A of *Escherichia coli* forms ion channels with two conductance states in planar lipid bilayers. *Journal of Biological Chemistry* **275**, 1594-1600.

Av-Gay, Y. and Everett, M. (2000). The eukaryotic-like Ser/Thr protein kinases of *Mycobacterium tuberculosis. Trends in Microbiology* **8**, 238-244.

Balasubramanian, V., Pavelka, M. S., Jr., Bardarov, S. S., Martin, J., Weisbrod, T. R., McAdam, R. A., Bloom, B. R. and Jacobs, W. R., Jr. (1996). Allelic exchange in *Mycobacterium tuberculosis* with long linear recombination substrates. *Journal of Bacteriology* **178**, 273-279.

Bardarov, S., Kriakov, J., Carriere, C., Yu, S., Vaamonde, C., McAdam, R. A., Bloom, B. R., Hatfull, G. F. and Jacobs, W. R., Jr. (1997). Conditionally replicating mycobacteriophages: a system for transposon delivery to *Mycobacterium tuberculosis. Proceedings of the National Academy of Sciences USA* **94**, 10961-10966.

Behr, M. A., Wilson, M. A., Gill, W. P., Salamon, H., Schoolnik, G. K., Rane, S. and Small, P. M. (1999). Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science* **284**, 1520-1523.

Belaaouaj, A., Kim, K. S. and Shapiro, S. D. (2000). Degradation of outer membrane protein A in *Escherichia coli*. Killing by neutrophil elastase. *Science* **289**, 1185-1188.

Bellinger-Kawahara, C. and Horwitz, M. A. (1990). Complement component C3 fixes selectively to the major outer membrane protein (MOMP) of *Legionella pneumophila* and mediates phagocytosis of liposome-MOMP complexes by human monocytes. *Journal of Experimental Medicine* **172**, 1201-1210.

Bergey (1986). *Bergey's Manual of Systematic Bacteriology*. Ed 2. Edited by JG Holt. Baltimore: Williams and Wilkins.

Bloom, B. (1994). Tuberculosis: pathogenesis, protection and control. Washington: American Society for Microbiology Press.

Bloom, B. R. and Murray, C. J. (1992). Tuberculosis: commentary on a reemergent killer. *Science* **257**, 1055-1064.

Brennan, P. and Draper, P. (1994). Ultrastructure of *Mycobacterium tuberculosis*. In *Tuberculosis: Pathogenesis, Protection and Control*, pp. 271-284. Edited by B. Bloom. Washington DC: American Society for Microbiology Press.

Brennan, P. J. and Nikaido, H. (1995). The envelope of mycobacteria. *Annual Reviews in Biochemistry* **64**, 29-63.

Chang, G., Spencer, R. H., Lee, A. T., Barclay, M. T. and Rees, D. C. (1998). Structure of the MscL homolog from *Mycobacterium tuberculosis*: a gated mechanosensitive ion channel. *Science* **282**, 2220-2226.

Chatterjee, D., Lowell, K., Rivoire, B., McNeil, M. R. and Brennan, P. J. (1992). Lipoarabinomannan of *Mycobacterium tuberculosis*. Capping with mannosyl residues in some strains. *Journal of Biological Chemistry* **267**, 6234-6239.

Cirillo, J. D., Barletta, R. G., Bloom, B. R. and Jacobs, W. R., Jr. (1991). A novel transposon trap for mycobacteria: isolation and characterization of IS1096. *Journal of Bacteriology* **173**, 7772-7780.

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., Barrell, B. G. *et al.*, (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537-544.

Collins, D. M. and Stephens, D. M. (1991). Identification of an insertion sequence, IS1081, in *Mycobacterium bovis*. *FEMS Microbiology Letters* **67**, 11-15.

Colston (1997). The Return of the Great White Plague. In *Biologist*, pp. 392-394.

Daffé, M. and Draper, P. (1998). The envelope layers of mycobacteria with reference to their pathogenicity. *Advances in Microbial Physiology* **39**, 132-203.

Daffé, M. and Etienne, G. (1999). The capsule of *Mycobacterium tuberculosis* and its implications for pathogenicity. *Tubercle and Lung Disease* **79**, 153-169.

Daniel, T., Bates, J. and Downes, K. (1994). History of tuberculosis. In *Tuberculosis: pathogenicity, protection and control*, pp. 13-24. Edited by B. Bloom. Washington DC: ASM Press.

Davidson, L. A., Draper, P. and Minnikin, D. E. (1982). Studies on the mycolic acids from the walls of *Mycobacterium microti*. *Journal Of General Microbiology* **128 (4)**, 823-828.

Deretic, V., Leveau, J. H. J., Mohr, C. D. and Hibler, N. S. (1992). *In vitro* Phosphorylation of AlgR, a regulator of mucoidy in *Pseudomonas aeruginosa*, by a histidine kinase and effects of small phospho-donor molecules. *Molecular Microbiology* **6**, 2761-2767.

Denis, M. (1992). Human neutrophils, activated with cytokines or not, do not kill virulent *Mycobacterium tuberculosis*. *Journal of Infectious Diseases* **163**, 919-920.

Dorman, C. J., Chatfield, S., Higgins, C. F., Hayward, C. and Dougan, G. (1989). Characterization of porin and *ompR* mutants of a virulent strain of *Salmonella typhimurium*: *ompR* mutants are attenuated *in vivo*. *Infection and Immunity* **57**, 2136-2140.

Draper, P. (1998). The outer parts of the mycobacterial envelope as permeability barriers. *Frontiers in Bioscience* **3**, 1253-1261.

Fang, Z., Doig, C., Kenna, D. T., Smittipat, N., Palittapongarnpim, P., Watt, B. and Forbes, K. J. (1999). IS6110-mediated deletions of wild-type chromosomes of *Mycobacterium tuberculosis*. *Journal Of Bacteriology* **181**, 1014-1020.

Fine, P. (1994). Immunities in and to tuberculosis: Implications for pathogenesis and vaccination. In *Tuberculosis: Back to the Future*, pp. 53-73. Edited by J. Porter and K. McAdam. Chichester: John Wiley and Sons.

Gomez, J. E. and Bishai, W. R. (2000). *whmD* is an essential mycobacterial gene required for proper septation and cell division. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 8554-8559.

Goodfellow, M. and Magee, J. (1999). Taxonomy of Mycobacteria. In *Mycobacteria I, Basic Aspects*, pp. 1-71. Edited by PRJ Gangadharam and PA Jenkins. New York: International Thomson Publishing / Chapman Hall.

Goren, M. B., D'Arcy Hart, P., Young, M. R. and Armstrong, J. A. (1976). Prevention of phagosome-lysosome fusion in cultured macrophages by sulfatides of *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences USA* **73**, 2510-2514.

Graeme-Cook, K. A. (1991). The regulation of porin expression in *Eschericia coli*: effect of turgor stress. *FEMS Microbiology Letters* **79**, 219-224.

Guilhot, C., Gicquel, B., Davies, J. and Martin, C. (1992). Isolation and analysis of IS6120, a new insertion sequence from *Mycobacterium smegmatis*. *Molecular Microbiology* **6**, 107-113.

Guilhot, C., Jackson, M. and Gicquel, B. (1999). Mobile genetic elements and plasmids: tools for genetic studies. In *Mycobacteria: Molecular Biology and Virulence*, pp. 17-37. Edited by C. Ratledge and J. Dale. Oxford: Blackwell Science.

Guilhot, C., Otal, I., Van Rompaey, I., Martin, C. and Gicquel, B. (1994). Efficient transposition in mycobacteria: construction of *Mycobacterium smegmatis* insertional mutant libraries. *Journal of Bacteriology* **176**, 535-539.

Hall, M. N. and Silhavy, T. J. (1981). Genetic analysis of the *ompB* locus in *Escherichia coli* K-12. *Journal of Molecular Biology* **151**, 1-15.

Haydel, S. E., Dunlap, N. E. and Benjamin, W. H., Jr. (1999). *In vitro* evidence of two-component system phosphorylation between the *Mycobacterium tuberculosis* TrcR/TrcS proteins. *Microbial Pathogenicity* **26**, 195-206.

Head, C. G., Tardy, A. and Kenney, L. J. (1998). Relative Binding Affinities of OmpR and OmpR-phosphate at the *ompF* and *ompC* Regulatory Sites. *Journal of Molecular Biology* **281**, 857-870.

Heyde, M., Laloi, P. and Portalier, R. (2000). Involvement of carbon source and acetyl phosphate in the external-pH- dependent expression of porin genes in *Escherichia coli*. *Journal of Bacteriology* **182**, 198-202.

Hinds, J., Mahenthiralingham, E., Kempsell, K. E., Duncan, K., Stokes, R. W., Parish, T. and Stoker, N. G. (1999). Enhanced gene replacement in mycobacteria. *Microbiology* **145**, 519-527.

Hoch, J. A. (2000). Two-component and phosphorelay signal transduction. *Current Opinion in Microbiology* **3**, 165-170.

Hsing, W. and Silhavy, T. J. (1997). Function of conserved histidine-243 in phosphatase activity of EnvZ, the sensor for porin osmoregulation in *Eschericia coli*. *Journal of Bacteriology* **179**, 3729-3735.

Huang, K. J., Schieberl, J. L. and Igo, M. M. (1994). A distant upstream site involved in the negative regulation of the *Escherichia coli* ompF gene. *Journal of Bacteriology* **176**, 1309-1315.

Huang, K.-J. and Igo, M. M. (1996). Identification of bases in the OmpF regulatory region, which interact with the transcription factor OmpR. *Journal of Molecular Biology* **262**, 615-628.

Huang, K.-J., Lan, C.-Y. and Igo, M. M. (1997). Phosphorylation stimulates the cooperative DNA-binding properties of the transcription factor OmpR. *Proceedings of the National Academy of Sciences USA* **94**, 2828-2832.

Jackson, M., Raynaud, C., Lanéelle, M. A., Guilhot, C., Laurent-Winter, C., Ensergueix, D., Gicquel, B. and Daffé, M. (1999). Inactivation of the antigen 85C gene profoundly affects the mycolate content and alters the permeability of the *Mycobacterium tuberculosis* cell envelope. *Molecular Microbiology* **31**, 1573-1587.

Jap, B. K. and Wallian, P. J. (1996). Structure and functional mechanisms of porins. *Physiological Reviews* **76**, 1073-1088.

Jarlier, V. and Nikaido, H. (1990). Permeability Barrier to Hydrophilic Solutes in *Mycobacterium Chelonei. Journal of Bacteriology* **172**, 1418-1423.

Ji, Y. E., Colston, M. J. and Cox, R. A. (1994). The ribosomal RNA (*rrn*) operons of fast-growing mycobacteria: primary and secondary structures and their relation to *rrn* operons of pathogenic slow-growers. *Microbiology* **140**, 2829-2840.

Jones, G.S., Amirault, H. J. and Andersen, B.R., (1990). Killing of *Mycobacterium tuberculosis* by neutrophils: a nonoxidative process. *Journal of Infectious Diseases* **162.** 700-704.

Kalpana, G. V., Bloom, B. R. and Jacobs, W. R., Jr. (1991). Insertional mutagenesis and illegitimate recombination in mycobacteria. *Proceedings of the National Academy of Sciences USA* **88**, 5433-5437.

Kartmann, B., Stengler, S. and Niederweis, M. (1999). Porins in the cell wall of *Mycobacterium tuberculosis. Journal of Bacteriology* **181**, 6543-6546.

Koebnik, R., Locher, K. P. and Van Gelder, P. (2000). Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Molecular Microbiology* **37**, 239-253.

Kong, T. H., Coates, A. R., Butcher, P. D., Hickman, C. J. and Shinnick, T. M. (1993). *Mycobacterium tuberculosis* expresses two chaperonin-60 homologs. *Proceedings of the National Academy of Sciences USA* **90**, 2608-2612.

Krebs (1999). Agriculture Committee Fifth Report: Badgers and Bovine Tuberculosis. . London: House of Commons.

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

Lan, C.-Y. and Igo, M. M. (1998). Differential expression of the OmpF and OmpC porin proteins in *Eschericia coli* K-12 depends upon the level of active OmpR. *Journal of Bacteriology* **180**, 171-174.

Li, X. Z., Nikaido, H. and Williams, K. E. (1997). Silver-resistant mutants of *Escherichia coli* display active efflux of Ag^+ and are deficient in porins. *Journal of Bacteriology* **179**, 6127-6132.

Lichtinger, T., Burkovski, A., Niederweis, M., Kramer, R. and Benz, R. (1998). Biochemical and biophysical characterization of the cell wall porin of *Corynebacterium glutamicum*: the channel is formed by a low molecular mass polypeptide. *Biochemistry* **37**, 15024-15032.

Lichtinger, T., Heym, B., Maier, E., Eichner, H., Cole, S. T. and Benz, R. (1999). Evidence for a small anion-selective channel in the cell wall of *Mycobacterium bovis* BCG besides a wide cation-selective pore. *FEBS Letters* **454**, 349-355.

Liu, J., Barry, C. and Nikaido, H. (1999). Cell Wall: physical structure and permeability. In *Mycobacteria: Molecular Biology and Virulence*, pp. 220-239. Edited by C. Ratledge and J. Dale. Oxford: Blackwell Science.

Liu, W., Qi, Y. and Hulett, F. M. (1998). Sites internal to the coding regions of *phoA* and *pstS* bind PhoP and are required for full promoter activity. *Molecular Microbiology* **28**, 119-130.

Liu, X. and Ferenci, T. (1998). Regulation of porin-mediated outer membrane permeability by nutrient limitation in *Eschericia coli*. *Journal of Bacteriology* **180**, 3917-3922.

Lukat, G. S., McCleary, W. R., Stock, A. M. and Stock, J. B. (1992). Phosphorylation of bacterial response regulator proteins by low molecular weight phospho-donors. *Proceedings of the National Academy of Sciences USA* **89**, 718-722.

Martin, C., Timm, J., Rauzier, J., Gomez-Lus, R., Davies, J. and Gicquel, B. (1990). Transposition of an antibiotic resistance element in mycobacteria. *Nature* **345**, 739-743.

Matsubara, M., Kitaoka, S. I., Takeda, S. I. and Mizuno, T. (2000). Tuning of the porin expression under anaerobic growth conditions by his-to-Asp cross-phosphorelay through both the EnvZ-osmosensor and ArcB-anaerosensor in *Escherichia coli*. *Genes to Cells* **5**, 555-569.

McCleary, W. R. and Stock, J. B. (1994). Acetyl phosphate and the activation of two-component response regulators. *Journal of Biological Chemistry* **269**, 31567-31572.

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**, 735-738.

McNeil, M. R. and Brennan, P. J. (1991). Structure, function and biogenesis of the cell envelope of mycobacteria in relation to bacterial physiology, pathogenesis and drug resistance; some thoughts and possibilities arising from recent structural information. *Research in Microbiology* **142**, 451-463.

Miller, J. (1972). *Experiments in Molecular Genetics*: Cold Spring Harbour Laboratories.

Minnikin, D. (1982). Lipids: complex lipids, their chemistry, biosynthesis and roles. In *The Biology of the Mycobacteria*, pp. 95-169. Edited by C. Ratledge and J. Stanford. London: Academic Press.

Mizuno, T. and Mizushima, S. (1990). Signal transduction and gene regulation through the phosphorylation of two regulatory components: the molecular basis for the osmotic regulation of the porin genes. *Molecular Microbiology* **4**, 1077-1082.

Morse, D. (1994). Multidrug resistance; the New York experience. In *Tuberculosis Back to the Future*, pp. 225-230. Edited by J. Porter and K. McAdam. Chichester: Wiley and Sons.

Movahedzadeh, F., Colston, M. J. and Davis, E. O. (1997). Determination of DNA sequences required for regulated *Mycobacterium tuberculosis* RecA expression in response to DNA-damaging agents suggests that two modes of regulation exist. *Journal of Bacteriology* **179**, 3509-3518.

Mukhopadhyay, S., Basu, D. and Chakrabarti, P. (1997). Characterisation of a porin from *Mycobacterium smegmatis*. *Journal of Bacteriology* **179**, 6205-6207.

Munson, G. P. and Scott, J. R. (2000). Rns, a virulence regulator within the AraC family, requires binding sites upstream and downstream of its own promoter to function as an activator. *Molecular Microbiology* **36**, 1391-1402.

Nakayama, S.-I. and Watanabe, H. (1998). Identification of cpxR as a positive regulator essential for expression of the *Shigella sonnei* virF gene. *Journal of Bacteriology* **180**, 3522-3528.

Negm, R. S. and Pistole, T. G. (1999). The porin OmpC of *Salmonella typhimurium* mediates adherence to macrophages. *Canadian Journal of Microbiology* **45**, 658-669.

Neidhart, F., Ingraham, J. and Schaechter, M. (1990). *Physiology of the Bacterial Cell, A Molecular Approach.* Sunderland, MA: Sinauer Associates.

Niederweis, M., Ehrt, S., Heinz, C., Klöcker, U., Karosi, S., Swiderek, K. M., Riley, L. W. and Benz, R. (1999). Cloning of the mspA gene encoding a porin from *Mycobacterium smegmatis*. *Molecular Microbiology* **33**, 933-945.

Nikaido, H. (1993). Transport across the bacterial outer membrane. *Journal of Bioenergetics and Biomembranes* **25**, 581-589.

Nikaido, H. (1994). Porins and specific diffusion channels in bacterial outer membranes. *Journal of Biological Chemistry* **269**, 3905-3908.

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

Nikaido, H. (1999). Microdermatology: cell surface in the interaction of microbes with the external world. *Journal of Bacteriology* **181**, 4-8.

Nikaido, H., Kim, S. H. and Rosenberg, E. Y. (1993). Physical organization of lipids in the cell wall of *Mycobacterium chelonae*. *Molecular Microbiology* **8**, 1025-1030.

Norman, E., Dellagostin, O. A., McFadden, J. J. and Dale, J. W. (1995). Gene replacement by homologous recombination in *Mycobacterium bovis* BCG. *Molecular Microbiology* **16**, 755-760.

Ortalo-Magné, A., Dupont, M. A., Lemassu, A., Andersen, A. B., Gounon, P. and Daffé, M. (1995). Molecular composition of the outermost capsular material of the tubercle bacillus. *Microbiology* **141**, 1609-1620.

Ortalo-Magné, A., Lemassu, A., Lanéelle, M. A., Bardou, F., Silve, G., Gounon, P., Marchal, G. and Daffé, M. (1996). Identification of the surface-exposed lipids on the cell envelopes of *Mycobacterium tuberculosis* and other mycobacterial species. *Journal of Bacteriology* **178**, 456-461.

Papa, F., Luquin, M. and David, H. L. (1992). DOT-ELISA for detection of phenolic glycolipid PGL-Tb1 and diacyl-trehalose antigens of *Mycobacterium tuberculosis*. *Research in Microbiology* **143**, 327-331.

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

Parish, T. and Stoker, N. G. (1995). Electroporation of mycobacteria. In *Methods in Molecular Biology*, pp. 237-252. Edited by J. A. Nickoloff. Totowa, NJ: Humana Press Inc.

Parish, T. and Stoker, N. G. (1997). Development and use of a conditional antisense mutagenesis system in mycobacteria. *FEMS Microbiology Letters* **154**, 151-157.

Parish, T. and Stoker, N. G. (2000). *glnE* is an essential gene in *Mycobacterium tuberculosis*. *Journal of Bacteriology* **182**, 5715-5720.

Parish, T. and Stoker, N. G. (2000). Use of a flexible cassette method to generate a double unmarked *Mycobacterium tuberculosis tly*A *plc*ABC mutant by gene replacement *Microbiology* **146**, 1969-1975.

Pautsch, A. and Schulz, G. E. (1998). Structure of the outer membrane protein A transmembrane domain. *Nature Structural Biology* **5**, 1013-1017.

Pavelka, M. S., Jr. and Jacobs, W. R., Jr. (1999). Comparison of the construction of unmarked deletion mutations in *Mycobacterium smegmatis*, *Mycobacterium bovis* bacillus Calmette-Guérin, and *Mycobacterium tuberculosis* H37Rv by allelic exchange. *Journal of Bacteriology* **181**, 4780-4789.

Pelicic, V., Jackson, M., Reyrat, J. M., Jacobs, W. R., Jr., Gicquel, B. and Guilhot, C. (1997). Efficient allelic exchange and transposon mutagenesis in *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences USA* **94**, 10955-10960.

Pelicic, V., Reyrat, J. M. and Gicquel, B. (1996). Expression of the *Bacillus subtilis sacB* gene confers sucrose sensitivity on mycobacteria. *Journal of Bacteriology* **178**, 1197-1199.

Pelicic, V., Reyrat, J. M. and Gicquel, B. (1996). Generation of unmarked directed mutations in mycobacteria, using sucrose counter-selectable suicide vectors. *Molecular Microbiology* **20**, 919-925.

Petrini, B. and Hoffner, S. (1999). Drug-resistant and multidrug-resistant tubercle bacilli. *International Journal of Antimicrobial Agents* **13**, 93-97.

Pratt, L. A., Hsing, W., Gibson, K. E. and Silhavy, T. J. (1996). From acids to *osmZ*: multiple factors influence synthesis of the OmpF and OmpC porins in *Eschericia coli*. *Molecular Microbiology* **20**, 911-917.

Rastogi, N. (1991). Recent observations concerning structure and function relationships in the mycobacterial cell envelope: elaboration of a model in terms of mycobacterial pathogenicity, virulence and drug-resistance. *Research in Microbiology* **142**, 464-476.

Rawling, E. G., Brinkman, F. S. and Hancock, R. E. (1998). Roles of the carboxy-terminal half of *Pseudomonas aeruginosa* major outer membrane protein OprF in cell shape, growth in low-osmolarity medium, and peptidoglycan association. *Journal of Bacteriology* **180**, 3556-3562.

Raynaud, C., Laneelle, M. A., Senaratne, R. H., Draper, P., Lanéelle, G. and Daffé, M. (1999). Mechanisms of pyrazinamide resistance in mycobacteria: importance of lack of uptake in addition to lack of pyrazinamidase activity. *Microbiology* **145**, 1359-1367.

Reyrat, J. M., Pelicic, V., Gicquel, B. and Rappuoli, R. (1998). Counterselectable markers: untapped tools for bacterial genetics and pathogenesis. *Infection and Immunity* **66**, 4011-4017.

Riess, F. G., Lichtinger, T., Cseh, R., Yassin, A. F., Schaal, K. P. and Benz, R. (1998). The cell wall porin of *Nocardia farcinica*: biochemical identification of the channel-forming protein and biophysical characterization of the channel properties. *Molecular Microbiology* **29**, 139-150.

Riess, F. G., Lichtinger, T., Yassin, A. F., Schaal, K. P. and Benz, R. (1999). The cell wall porin of the gram-positive bacterium *Nocardia asteroides* forms cation-selective channels that exhibit asymmetric voltage dependence. *Archives of Microbiology* **171**, 173-182.

Rinke de Wit, T. F., Bekelie, S., Osland, A., Miko, T. L., Hermans, P. W., van Soolingen, D., Drijfhout, J. W., Schoningh, R., Janson, A. A. and Thole, J. E. (1992). Mycobacteria contain two *groEL* genes: the second *Mycobacterium leprae groEL* gene is arranged in an operon with *groES*. *Molecular Microbiology* **6**, 1995-2007.

Rubin, E. J., Akerley, B. J., Novik, V. N., Lampe, D. J., Husson, R. N. and Mekalanos, J. J. (1999). *In vivo* transposition of mariner-based elements in enteric bacteria and mycobacteria. *Proceedings of the National Academy of Sciences USA* **96**, 1645-1650.

Russell, D. G., Sturgill-Koszycki, S., Vanheyningen, T., Collins, H. and Schaible, U. E. (1997). Why intracellular parasitism need not be a degrading experience for mycobacterium. *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences* **352**, 1303-1310.

Saint, N., De, E., Julien, S., Orange, N. and Molle, G. (1993). Ionophore properties of OmpA of *Escherichia coli*. *Biochimica et Biophysica Acta* **1145**, 119-123.

Sander, P., Meier, A. and Böttger, E. C. (1995). *rpsL*$^{+}$: a dominant selectable marker for gene replacement in mycobacteria. *Molecular Microbiology* **16**, 991-1000.

Schirmer, T. (1998). General and specific porins from bacterial outer membranes. *Journal of Structural Biology* **121**, 101-109.

Schulz, G. E. (1993). Bacterial porins: structure and function. *Current Opinion in Cell Biology* **5**, 701-707.

Senaratne, R. H., Mobasheri, H., Papavinasasundaram, K. G., Jenner, P., Lea, E. J. and Draper, P. (1998). Expression of a gene for a porin-like protein of the OmpA family from *Mycobacterium tuberculosis* H37Rv. *Journal of Bacteriology* **180**, 3541-3547.

Slauch, J. M., Russo, F. D. and Silhavy, T. J. (1991). Suppressor mutations in *rpoA* suggest that OmpR controls transcription by direct interaction with the alpha subunit of RNA polymerase. *Journal of Bacteriology* **173**, 7501-7510.

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

Stewart, G. and McFadden, J. (1999). Recombination. In *Mycobacteria: Molecular Biology and Virulence*, pp. 1-16. Edited by C. Ratledge and J. Dale. Oxford: Blakwell Science.

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

Sugawara, E. and Nikaido, H. (1994). OmpA protein *of Escherichia coli* outer membrane occurs in open and closed channel forms. *Journal of Biological Chemistry* **269**, 17981-17987.

Supply, P., Magdalena, J., Himpens, S. and Locht, C. (1997). Identification of novel intergenic repetitive units in mycobacterial two-component system operon. *Molecular Microbiology* **26**, 991-1003.

Sutcliffe, I. C. (1998). Cell envelope composition and organisation in the genus *Rhodococcus. Antonie Van Leeuwenhoek* **74**, 49-58.

Trias, J. and Benz, R. (1993). Characterization of the channel formed by the mycobacterial porin in lipid bilayer membranes. Demonstration of voltage gating and of negative point charges at the channel mouth. *Journal of Biological Chemistry* **268**, 6234-6240.

Trias, J. and Benz, R. (1994). Permeability of the cell wall of *Mycobacterium smegmatis. Molecular Microbiology* **14**, 283-290.

Trias, J., Jarlier, V. and Benz, R. (1992). Porins in the cell wall of mycobacteria. *Science* **258**, 1479-1481.

Tsukamura, M. (1981). Some considerations regarding the classification and identification of mycobacteria. *Reviews of Infectious Diseases* **3**, 829-840.

Via, L. E., Curcic, R., Mudd, M. H., Dhandayuthapani, S., Ulmer, R. J. and Deretic, V. (1996). Elements of signal transduction in *Mycobacterium tuberculosis*: *in vitro* phosphorylation and *in vivo* expression of the response regulator MtrA. *Journal of Bacteriology* **178**, 3314-3321.

Vidal, O., Longin, R., Prigent-Combaret, C., Dorel, C., Hooreman, M. and Lejeune, P. (1998). Isolation of an *Escherichia coli* K-12 mutant strain able to form biofilms on inert surfaces: involvement of a new ompR allele that increases curli expression. *Journal of Bacteriology* **180**, 2442-2449.

Waukau, J. and Forst, S. (1992). Molecular Analysis of the Signalling Pathway between EnvZ and OmpR in *Escherichia coli*. *Journal of Bacteriology* **174**, 1522-1527.

WHO (2000). World Health Report. Annex Table 3. Geneva.

Zahrt, T. C. and Deretic, V. (2000). An essential two-component signal transduction system in *Mycobacterium tuberculosis*. *Journal of Bacteriology* **182**, 3832-3838.

Zhang, E. and Ferenci, T. (1999). OmpF changes and the complexity of *Escherichia coli* adaptation to prolonged lactose limitation. *FEMS Microbiology Letters* **176**, 395-401.

Appendices

1. Tween Glutamate Broth (from Davidson *et al.*, 1992)

L-Monosodiumglutamate	8.0 g/l
Bacto-casitone (Difco)	1.0 g/l
Ferric ammonium citrate	0.1 g/l
Potassium phosphate	1.0 g/l
Disodium hydrogen phosphate	2.5 g/l
Calcium chloride (.2H20)	0.001 g/l
Copper sulphate (.5H2O)	0.0005 g/l
Zinc Sulphate (7H2O)	0.0005 g/l
Tween 80	0.5 ml

Autoclave and add Bovine Serum Albumin at 5 g/l

2. L-Broth

Tryptone	10 g/l
NaCl	10 g/l
Yeast Extract	5 g/l

Dissolve in 950 ml water, adjust pH to 7.5, make up to 1 l and sterilise by autoclaving.

3. Dubos broth

Part A

KH_2PO_4	1 g
$Na_2HPO_4 \cdot 12H_2O$	6.25 g
Na_3citrate	1.25 g
$MgSO_4 \cdot 7H_2O$	0.6 g
Asparagine	2 g
10 % Tween 80	5 ml
Casamino acids	2 g

Dissolve asparagine in hot water (50 ml), add to rest of constituents and make up to 950 ml. Adjust pH to 7.2 with 2 M NaOH and make up to 1 l. Sterilise by autoclaving.

Part B

BSA fraction V	5 g
Distilled water	50 ml

Dissolve without stirring, adjust pH to 7.4. Filter sterilise.

Add 1 part B to 25 parts A before use, and glycerol to 0.2 % if required.

4. Results from TaqMan real-time RT-PCR experiments – spreadsheets presented overleaf.

All spreadsheets with the exception of the last relate to NaCl shock of BCG. The last sheet of data relates to NaCl shock of *Mycobacterium tuberculosis*.

Samples using *gnd* primers/probe

	Ct	Average Ct	log10 amount	Actual amount
NTC	40	40		
NTC	40			
NTC	40			
20 FOLD DILUTION				
0 1 hour	27.02	29.19	1.87	**74.03**
	31.36			
0.2M 1 hour	27.83	28.13	2.14	**138.77**
	28.42			
0.4M 1 hour	27.44	28.74	1.98	**96.54**
	30.04			
0 24 hour	29.06	28.78	1.98	**94.57**
	28.49			
0.2M 24 hour	29.44	28.77	1.98	**95.13**
	28.09			
0.4M 24 hour	28.1	27.83	2.22	**165.64**
	27.55			
50 FOLD DILUTION				
0 1 hour	30.23	30.66	1.49	**31.19**
	31.08			
0.2M 1 hour	28.17	30.64	1.50	**31.56**
	33.1			
0.4M 1 hour	27.53	29.66	1.75	**56.27**
	31.78			
0 24 hour	30.76	30.93	1.42	**26.52**
	31.1			
0.2M 24 hour	31.39	30.16	1.62	**41.90**
	28.92			
0.4M 24 hour	28.66	27.27	2.36	**230.50**
	25.87			

Equation from standard curve: y = -3.9028x + 36.486

Samples using *ompATb* primers/probe

	Ct	Average Ct	log10 amount	Actual amount	Ratio of *ompATb/*
NTC	31.46	31.94			
NTC	32.32				
NTC	32.04				
20 FOLD DILUTION					
0 1 hour	24.45	24.11	1.52	**32.96**	**0.4451**
	23.77				
0.2M 1 hour	24.01	24.39	1.44	**27.42**	**0.1976**
	24.77				
0.4M 1 hour	25.02	24.97	1.27	**18.79**	**0.1946**
	24.91				
0 24 hour	22.38	23.01	1.83	**68.14**	**0.7205**
	23.63				
0.2M 24 hour	23.02	22.80	1.89	**78.22**	**0.8223**
	22.57				
0.4M 24 hour	24.37	24.61	1.38	**23.80**	**0.1437**
	24.84				
50 FOLD DILUTION					
0 1 hour	24.81	25.70	1.07	**11.63**	**0.3727**
	26.58				
0.2M 1 hour	24.42	24.83	1.31	**20.53**	**0.6504**
	25.24				
0.4M 1 hour	25.31	25.56	1.11	**12.75**	**0.2265**
	25.8				
0 24 hour	22.72	22.64	1.94	**86.90**	**3.2766**
	22.55				
0.2M 24 hour	24.59	23.76	1.62	**41.62**	**0.9934**
	22.92				
0.4M 24 hour	26.24	27.44	0.57	**3.69**	**0.0160**
	28.64				

Equation from standard curve: y = -3.5028x + 29.427

Samples using *gnd* primers/probe

	Average Ct	Ct	log10 amount	**Actual amount**
0 NaCl, 1 hour	30.78	30.48	1.778	**59.950**
		29.92		
		31.06		
		31.66		
0.2M NaCl, 1 hour	30.48	29.71	1.960	**91.276**
		31.06		
		30.18		
		30.98		
0.4M NaCl, 1 hour	30.26	30.21	1.842	**69.471**
		31.12		
		29.28		
		30.41		
0 NaCl, 24 hours	29.82	29.29	2.060	**114.799**
		29.7		
		30		
		30.28		
0.2M NaCl, 24 hours	31.68	31.32	1.579	**37.899**
		33.01		
		29.94		
		32.45		
0.4M NaCl, 24 hours	28.61	28.94	2.143	**138.971**
		28.7		
		28.61		
		28.2		

Equation from curve: y = -4.2176x + 37.978

Samples using *ompATb* primers/probe

	Average Ct	Ct	log10 amount	**Actual amount**	***ratio ompATb/gnd***
0 NaCl, 1 hour	24.83	24.76	1.256	**18.020**	**0.301**
		24.49			
		25.28			
		24.8			
0.2M NaCl, 1 hour	24.64	24.77	1.308	**20.332**	**0.223**
		24.64			
		24.41			
		24.73			
0.4M NaCl, 1 hour	25.61	25.46	1.048	**11.170**	**0.161**
		25.95			
		25.4			
		25.61			
0 NaCl, 24 hours	22.65	22.64	1.844	**69.806**	**0.608**
		22.42			
		22.79			
		22.73			
0.2M NaCl, 24 hours	22.79	22.56	1.804	**63.714**	**1.681**
		22.64			
		23			
		23.07			
0.4M NaCl, 24 hours	25.34	25.18	1.121	**13.202**	**0.095**
		25.3			
		25.41			
		25.45			

Equation from standard curve: y = -3.7193x + 29.603

Samples using *ompATb* primers and probe

	Ct	Average Ct	log10 amount	**actual amount**
0 NaCl, 1 hour	23.92	24.3825	1.373725342	**23.644**
	24.08			
	24.41			
	24.49			
	24.73			
	24.38			
	24.46			
	24.59			
0.2M NaCl, 1 hour	24.6	24.805	1.250980506	**17.823**
	25.02			
	25.03			
	25.02			
	24.37			
	25.21			
	24.63			
	24.56			
0.4M NaCl, 1 hour	25.01	24.98125	1.199776299	**15.841**
	25.07			
	24.87			
	24.46			
	25.01			
	25.22			
	25.01			
	25.2			

Equation from standard curve: y = -3.4421x + 29.111

Samples using *gnd* primers and probe

	Ct	Average Ct	log10 amount	**actual amount**	**Ratio of ompATb/gnd**
0 NaCl, 1 hour	29.68	30.06625	1.662719979	**45.996**	**0.51405**
	29.36				
	29.66				
	29.91				
	30.27				
	31.04				
	30.36				
	30.25				
0.2M NaCl, 1 hour	29.56	29.10125	1.91246118	**81.745**	**0.21803154**
	29.65				
	29.54				
	29.14				
	28.48				
	28.86				
	28.43				
	29.15				
0.4M NaCl, 1 hour	28.9	28.7925	1.992365424	**98.257**	**0.16121701**
	29.75				
	29.18				
	29.11				
	29.04				
	28.15				
	28.11				
	28.1				

Equation from standard curve: y = -3.864x + 36.491

Samples using gnd primers/probe

	Average Ct	Ct	log10 amount	**actual amount**
0 NaCl	28.2125	27.54	2.073692675	**118.493**
		28		
		28.17		
		28.61		
		28.02		
		28.9		
		28.38		
		28.08		
0.4M NaCl 1 hour	28.48125	27.61	2.003981635	**100.921**
		40		
		26.29		
		27.04		
		25.72		
		27		
		26.28		
		27.91		
0.4M NaCl 2 hours	25.38125	25.46	2.808090371	**642.821**
		26.18		
		25.59		
		25.45		
		25.24		
		25.49		
		24.99		
		24.65		

equation from curve: y=-3.8552x+36.207

Samples using *ompATb* primers/probe

	average Ct	Ct	log10 amount	**Actual amount**	**Ratio *ompATb/gnd***
0 NaCl	22.7775	22.68	1.90348337	**80.072**	**0.676**
		23			
		22.57			
		22.52			
		22.74			
		22.8			
		23.01			
		22.9			
0.4M NaCl 1 hour	24.1625	23.64	1.514099356	**32.666**	**0.324**
		26.16			
		24.17			
		24.38			
		23.55			
		23.96			
		23.13			
		24.31			
0.4M NaCl 2 hours	23.46875	23.21	1.709142793	**51.185**	**0.080**
		23.44			
		23.44			
		23.38			
		23.42			
		23.47			
		23.94			
		23.45			

equation from curve: y=-3.5569x+29.548

Samples using *ompATb* primers/probe

	Actual amount	log10 amount	Average Ct	Ct
T0	**0.305**	-0.515	23.110	23.12
				23.16
				23.05
T2	**0.549**	-0.261	22.483	22.38
				22.31
				22.76
T6	**0.550**	-0.259	22.480	22.4
				22.39
				22.65
T24	**0.260**	-0.584	23.280	23.45
				23.22
				23.17
T0 + NaCl	**0.327**	-0.485	23.037	23.07
				22.94
				23.1
T2 + NaCl	**0.486**	-0.313	22.613	22.65
				22.64
				22.55
T6 + NaCl	**0.339**	-0.469	22.997	23.03
				22.94
				23.02
T24 + NaCl	**0.298**	-0.526	23.137	23.12
				23.2
				23.09

Equation from standard curve: y = -2.461x + 21.842

Samples using *gnd* primers/probe

	Actual amount	log10 amount	Average Ct	Ct	Ratio *ompATb/gnd*
T0	**1.444**	0.159	21.417	21.31	**0.212**
				21.58	
				21.36	
T2	**0.938**	-0.028	21.870	21.92	**0.585**
				21.91	
				21.78	
T6	**0.969**	-0.014	21.837	21.85	**0.568**
				21.76	
				21.9	
T24	**0.806**	-0.094	22.030	22.03	**0.323**
				22.07	
				21.99	
T0 + NaCl	**1.248**	0.096	21.570	21.23	**0.262**
				21.99	
				21.49	
T2 + NaCl	**1.072**	0.030	21.730	21.74	**0.453**
				21.72	
				21.73	
T6 + NaCl	**0.692**	-0.160	22.190	22.25	**0.490**
				22.01	
				22.31	
T24 + NaCl	**0.856**	-0.068	21.967	22.03	**0.348**
				22.11	
				21.76	

Equation from standard curve: y = -2.4232x + 21.803

SUPPLEMENTARY APPENDIX 1

ADDITIONAL INFORMATION FOR FIGURE LEGENDS

Figure 1	Layers of the envelope are shown as they appear under the transmission electron microscope. The periplasmic space, marked as P, is hypothetical and has not been proven experimentally although it is highly likely to exist by analogy to other bacterial membranes.
Figure 2	The region of the cell envelope referred to as the Cell Wall Skeleton (CWS) is any structure exterior to the periplasmic space up to and not including the capsular layer; this thesis divides the envelope into the inner (cytoplasmic) bilayer membrane, the cell wall, and the capsule. The position of the porin in the diagram of the mycobacterial membrane is a logical one, this has not been experimentally proven yet. The mycobacterial cell envelope shown here is that proposed by McNeil and Brennan, with intercalated mycolic acids and associated lipids. The model of Rastogi does not have these chains intercalated.
Figure 3	This diagram of the mycobacterial wall is representative of the model proposed by McNeil and Brennan.
Figure 4	Regulation of porin levels in *Escherchia coli.* The membrane bound sensor protein, EnvZ, responds to environmental stress. EnvZ is phosphorylated, in the presence of ATP, and under suitable conditions can donate the phospho- group to the regulator protein OmpR. Phosphorylated OmpR then serves as a transcriptional activator or repressor of the genes encoding the major outer membrane porins OmpF and OmpC.
Figure 5	Regions of homology to the mycobacterial genome upstream (1769 bp) and downstream (2083 bp) of the Rv0903c coding sequence are cloned into p2NIL in order to generate pRAS17. 674 bp of the coding sequence of Rv0903c, beginning 8 bp after the translational start codon, have been removed and substituted by 312 bp of p2NIL cloning site-derived DNA. Incorporation of the *PacI* fragment from pGOAL19, containing *lacZ*, *sacB* and *hyg*^{r} genes is cloned into the *PacI* site of pRAS17 to generate pRAS18; this is a suicide delivery vector with 4 selectable markers and a mutated Rv903c allele. pRAS18 is transformed into *Mycobacterium tuberculosis* in order to attempt homologous recombination knockout of Rv0903c function.
Figure 6	Cosmid Y31 (MTCY31) is one of the cosmids used during the sequencing of the *Mycobacterium tuberculosis* genome. This is used as a positive control as it is the cosmid containing the porin gene *ompATb*, as well as the sensor-regulator system Rv0902c/Rv0903c. Bands of amplified DNA (1050 bp) are seen in species that have a gene of sufficient homology to *ompATb* from *Mycobacterium tuberculosis*, the species against which the primers were designed, to give a PCR product. Control amplifications of the *recA* gene verified that DNA had been obtained from the species which gave no *ompATb* product, thus the result is genuine and is not a result of inadequate DNA extraction.

Figure 7	An 895 bp PCR product is observed in species that have a gene sufficiently similar to Rv0903c from *Mycobacterium* tuberculosis to allow amplification. Cosmid Y31 is used as a positive control as this contains the gene Rv0903c from the tuberculosis genome.
Figure 8	An arrow marks the band in a Western blot of an SDS-PAGE of bacterial cell lysate that reacts to a polyclonal rabbit anti-OmpATb antibody. This is of the correct size to be the OmpATb protein. Identical experiments in *Mycobacterium smegmatis* and *Mycobacterium chelonae* failed to identify any reactive band – in agreement with the results obtained by PCR analysis.
Figure 9	The sequence marked 'query' is that of *Escherichia coli* OmpR; matches to the Rv0903c gene (marked 'sbjct') are shown between the two sequences. A + indicates non-identical residues which have similar properties.
Figure 10	A 350 bp PCR product from cDNA of *Mycobacterium bovis* BCG demonstrates that the Rv0903c gene is transcribed in this organism. Controls are reactions that have either no RNA template for the reverse transcription (eliminating contamination of the sample giving a false-positive result) or no reverse transcriptase (eliminating DNA contamination of the RNA template as a source of false-positive results).
Figure 11	A clustal alignment was performed using MegAlign from DNAStar inc. Amino acids boxed in black are those that are identical to their counterparts in Rv0903c from *Mycobacterium tuberculosis*. It can be seen that several motifs appear throughout the regulatory proteins examined.
Figure 12	A clustal alignment was performed using MegAlign from DNAStar inc. Residues that match a consensus sequence, determined by the software, are boxed in black.
Figure 13	The genes around the porin gene, denoted either by Rv number or by their annotation in the *Mycobacterium tuberculosis* genome are shown. Annotations in the *Mycobacterium tuberculosis* genome are: *gltA2* = citrate synthase Rv0897c = possible oxidoreductase Rv0898c = unknown Rv0900 = unknown Rv0901 = unknown *accD3* = acetyl/propionyl CoA carboxylase b subunit *echA6* = enoyl-CoA hydratase/isomerase superfamily Rv0906 = some similarity to romA, probable membrane protein Rv0907 = probable penicillin binding protein *ctpE* = probable cation transporting ATPase
Figure 14	Loading of recombinant Rv0903 protein, in mg, is shown across the top of the gel. A coomassie blue stain identifies a band of the correct size for the recombinant protein, 27 kDa. A smaller breakdown product is also seen.
Figure 15	Fragments of the *Mycobacterium tuberculosis* genome (sizes shown) were amplified by PCR and labelled with radioactive ^{32}P. Each fragment was run through a non-denaturing polyacrylamide gel with (+) and without (-) pre-incubation with recombinant Rv0903 protein. Retardation can be seen in those fragments that bind the protein. The position of the fragments on the genome of *Mycobacterium tuberculosis* can be seen in figure 16.

Figure 16	This diagram is not to scale, it is merely a diagrammatic representation of the fragments used in gel-shift experiments. By making the fragments progressively smaller and discounting the fragments which don't bind protein it is possible to narrow down the site of binding, this has been performed and fragments containing the red areas are those which bind protein.
Figure 17	The 343 bp region that is shown to bind Rv0903 protein lies between bases 530 and 873 of the 981 bp *ompATb* coding sequence. Phosphorylation of this fragment *in vitro* using acetyl phosphate shows that binding affinity is enhanced when the Rv0903 protein is phosphorylated. Approximately a 4-fold lower amount of protein is shown to bind the same DNA fragment.
Figure 18	The Marker lane shows Novex Rainbow protein size standards. Faint bands are seen which react to the anti-OmpATb antibody only in those isolates of *Mycobacterium smegmatis* that have been transformed with plasmid pRAS20. These are expressing the tuberculosis porin protein from its own promoter. Antibody based detection does not identify any proteins which react in the wild-type Mycobacterium smegmatis mc^2155.
Figure 19	This diagram represents a small portion of the growth curve of *Mycobacterium smegmatis* expressing the tuberculosis porin at a high level (from the hsp60 promoter). The initial inoculum was standardised in all cultures and the graph represents the mean value of 3 separate cultures. It can be seen that inclusion of 400 mM NaCl in the media slows down the growth of the mycobacterial cultures considerably, regardless of porin expression.
Figure 20	This graph represents the exponential phase of growth of *Mycobacterium smegmatis* cultures expressing the *Mycobacterium tuberculosis* porin gene (pRAS21) or the porin gene and the sensor/regulator pair Rv0902c/Rv0903c (pRAS20) after an identical inoculum. The average OD_{600} of 3 cultures is plotted. It can be seen that there is little difference between the rates of growth, with pRAS20 (porin + regulatory machinery) transformed *Mycobacterium smegmatis* growing slightly slower than the wild type or porin-alone (pRAS21) transformant.
Figure 21	As figure 20, except growth was measured in the presence of 200 mM NaCl.
Figure 22	Colonies of *Mycobacterium tuberculosis* which were streptomycin resistant and kanamycin resistant, displaying the phenotype which would be expected for a double crossover homologous recombination, had DNA extracted and analysed by PCR. Where 2 bands are seen a genomic wild type copy of the *ompATb* gene and an integrated copy of the disrupted gene from the plasmid are both present; where only one product is seen the strain is either a successful knockout (1360 bp product) or a wild type (512 bp) which has acquired the correct resistance phenotype by other means, such as mutation in the *rpsL* or *aph* gene.
Figure 23	The difference in the size of fragments which hybridise to a 132 bp probe labelled with ^{32}P dCTP is seen. Three different enzymatic digests are used in order to verify the knockout of the *ompATb* coding sequence

Figure 24	A 992 bp product, amplified from genomic DNA from 42 colonies that display the correct phenotype to be an Rv0903c gene knockout, shows that all are in fact still wild type. Cosmid Y31, which has a copy of the wild type *Mycobacterium tuberculosis* Rv0903c gene also shows an amplified fragment of the same size, whereas plasmid pRAS18 (carrying the disrupted allele) shows a fragment of 631 bp.
Figure 25	The level of promoter activity can be measured in Miller Units due to the activity of the downstream *lacZ* gene which encodes a functional ß-galactosidase protein. This is placed under the control of the fragment which is cloned in to the vector pEJ414. It can be seen that promoter activity is strongest in the fragment that includes 452 bp upstream of the *ompATb* translational start codon. Reduction of the size of the fragment to 178 bp leads to a significant loss of promoter function, indicating that the promoter lies between 178 and 452 bases upstream of the *ompATb* coding sequence. An average of 3 separate experiments is presented.
Figure 26	Amplification of a DNA template from 1 well of a 96 well plate is shown; in this case it is well number B3. The fluorescence from the reporter dye (FAM) is measured, as it is proportional to the amount of PCR product produced. The time at which fluorescence passes the threshold value is the cycle threshold (Ct) and is representative of the amount of template DNA included in the reaction. A lower Ct means that more template was present in the PCR reaction. Threshold cycle calculations are the defaults presented by the software (ABI Prism Sequence Detection Systems V1.6.3).
Figure 27	This figure represents the data presented in table 6. It can be seen that the addition of NaCl to the culture medium of *Mycobacterium bovis* BCG affects the level of the *ompATb* transcript as measured by real time RT-PCR. A higher molarity of NaCl has the effect of further reducing the level of *ompATb* transcript. The levels of the normalising gene *gnd* are thought to remain constant (personal communication, Dr E. Davis, NIMR) so the ratio of *ompATb/gnd* reflects only a change in *ompATb* input level.
Figure 28	This figure represents the data presented in table 5. Real-time RT-PCR was used to show that the level of the *ompATb* transcript in *Mycobacterium bovis* BCG fell dramatically in the first hour after the addition of 400 mM NaCl to the culture medium (Tween-glutamate medium) and continued to fall, although more slowly, in the second hour after NaCl addition.
Figure 29	A small section of the microarray slide is shown after hybridisation to Cy3 (red fluorescence) and Cy5 (yellow fluorescence) labelled DNA. The green spots in the corners of each grid are controls in order to align the analysis software correctly. When cDNA from two different samples is labelled, a difference in the transcript (RNA) level can be detected by a predominantly red or yellow spot, depending on which way round the samples were labelled (both orientations should be performed).

Figure 30	Fluorescence of Cy3 (y-axis) against Cy5 (x-axis) is plotted. Each spot on the graph relates to a spot on the microarray slide, thus if fluorescence from both Cy3 and Cy5 is equal then points will lie on a 45° line extending from the origin. Any points which lie outside of this line are interesting as they indicate that one or other of the dye labelled samples is dominant. The points lying up the y-axis all correspond to the green control spots on the microarray slide and as such are not of interest. The spot corresponding to gene *ompATb* is present on the 45° line.
Figure 31	Average OD_{600} of 2 separate cultures is plotted through the exponential growth phase of *Mycobacterium tuberculosis* 1424 wild type, a streptomycin resistant strain, and an *ompATb* (porin gene) knockout of the same strain. Antibiotics were not present in the growth medium. It can be seen that the slopes of the two curves are largely identical, therefore there is no severe detrimental effect on growth of loss of OmpATb function. The initial inoculum was standardised.
Figure 32	NaCl (0.2 M) in the growth medium imposes an osmotic shock on the bacterial cell but growth is not affected by loss of OmpATb function. The initial inoculum was standardised and an average of 2 separate cultures is plotted.
Figure 33	There is no observable difference in the rate of growth between an *ompATb*$^-$ strain and wild type *Mycobacterium tuberculosis* in the presence of 0.1 M raffinose. The initial inoculum was standardised and an average of 2 separate cultures is plotted.
Figure 34	There is no observable difference in the rate of growth between an *ompATb*$^-$ strain and wild type *Mycobacterium tuberculosis* in the presence of 0.2 M sucrose. The initial inoculum was standardised and an average of 2 separate cultures is plotted.

SUPPLEMENTARY APPENDIX 2

SUMMARY OF PLASMIDS GENERATED AND USED IN THIS STUDY

	EXPRESSION OF Rv0903c IN *Escherichia coli.*
pRAS1	Rv0903c gene cloned into pET15b (Invitrogen). This places a 6-Histidine tag upstream of the coding sequence of the Rv0903 gene, which is in-frame with the vector start codon and his- tag. 4 amino acids (12 bases) upstream of the transcriptional start codon of Rv0903c are included in the recombinant protein, as is a thrombin cleavage site (standard for this type of expression vector). The insert extends 124 bases beyond the Rv0903 translational stop codon.
	ATTEMPTED KNOCKOUT OF THE PORIN-ENCODING GENE *ompATb.*
pRAS4	A pRS5 (*SacI* deleted pMCS5, with a 4.7kb *NsiI* fragment of *M. tb* genome which contains the porin gene *ompATb*) derived plasmid for use in homologous recombination. A 1.28 kb Kanamycin resistance cassette from pMV261, flanked by *SacI* sites, is inserted 9 bp into to the coding sequence of the *ompATb* gene. The 5' 9 bp and the 3' 18 bp are flanked by 1704 bp of homology upstream and 2054 bp downstream, 954 bp of the (981 bp) *ompATb* coding sequence are replaced by 1.28 kb of kanamycin resistance gene.
pRAS5	A pRAS4 derived plasmid which has an *XbaI/NheI* flanked *SacB* gene, under the control of the *hsp60* promoter (from pKP140, Dr KG Papavinasasundaram, NIMR) inserted.
	Plasmid **pOmpA-aph-rpsLtb** was constructed by Dr Peter Sander and his team at the Institute for Medical Microbiology, Hannover. A 4.4 kb *EcoRI* fragment of the *M. tb* genome containing *ompATb* was digested with *BsmI* (150 bp after the OmpATb start codon) and *HpaI* (410 bp before the stop codon) and a kanamycin resistance gene (1.27 kb) inserted at these sites. A 1681 bp upstream region of homology and a 1770 bp downstream region of homology to the *M. tb ompATb* gene remain, along with the 5' 150 bp and the 3' 410 bp of the *OmpATb* coding region. The plasmid has a wild-type (dominant) $rpsL^+$ gene, which leads to streptomycin sensitivity if expressed in a streptomycin resistant ($rpsL^-$) strain.

β-GALACTOSIDASE ASSAYS OF *ompATb* PROMOTER ACTIVITY

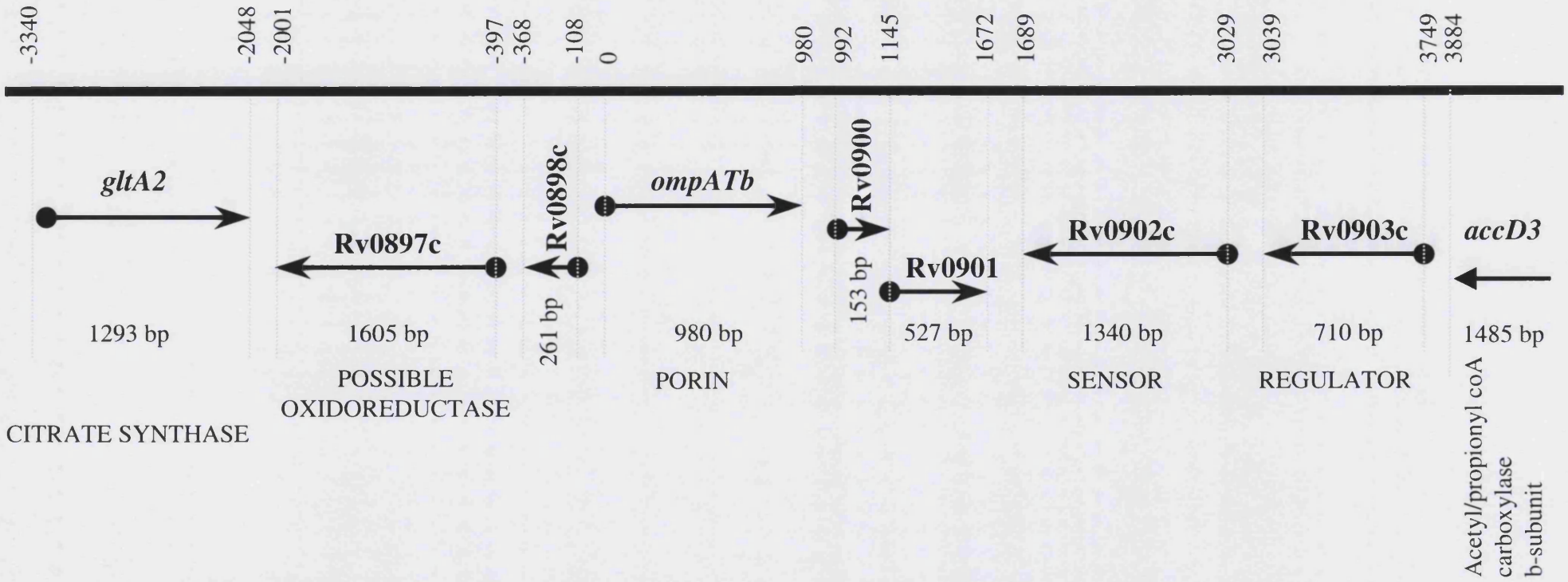
pRAS3	This plasmid contains a region 2.2 kb in length, extending, at the 3' end, 7 bases into the coding sequence of the *M. tb ompATb* gene. A *HindIII* site on the lower primer and an *XbaI* site on the upper primer from which the 2.2 kb fragment was constructed allows insertion of the putative promoter region into pEJ414 (Dr E Davis, NIMR), a mycobacterial integrating vector carrying a promoterless *lacZ* gene downstream of the cloning site.
pRAS6	identical to that above, except the length of the putative promoter region has been truncated, at the 5' end, such that the length of the fragment is 684 bp
pRAS7	The same as above, but truncated at the 5' end so that the fragment from upstream of the *ompATb* gene is 452 bp in length.
pRAS8	The same as above, but truncated at the 5' end so that the promoter fragment is 178 bp upstream of the *ompATb* gene (extending 7 bp into the translational coding sequence)

EXPRESSION OF THE *M. tb* PORIN OmpATb IN *Mycobacterium smegmatis*

pRAS10	A mycobacterial replicating plasmid derived from pMV261. A 1.004 kb fragment containing the porin-encoding gene *ompATb*, flanked by a *BamHI* site and one base before the start codon (in order to keep the sequence in frame) and a *HindIII* site 4 bases beyond the stop codon, is placed under the control of the plasmid-borne hsp60 promoter.
pRAS11	A mycobacterial integrating plasmid derived from pMV306. An *XbaI/NheI* digest of pRAS10 excises the pro-hsp60-*ompATb* construct detailed above, which is cloned into an *XbaI* digest of pMV306. This plasmid expresses the porin gene at a lower level (one copy per cell) than pRAS10, which is replicative.
pRAS19	This plasmid is derived from commercially available pBluescript (KS+). A 6.6 kb *SacI* fragment of the *M. tb* genome containing the porin gene, from 2303 bp upstream of the *ompATb* coding sequence to 3305 bp downstream is incorporated. This fragment contains the porin gene *ompATb*; the sensor component Rv0902c and the regulator component Rv0903c, extending to 534 bp upstream of the Rv0903c start codon (it is on the complimentary strand).
pRAS20	This is a mycobacterial integrating version of pRAS19, the *SacI* fragment (6.6 kb) is excised and blunt-ended, it is then inserted at the *EcoRV* site of pMV306 instead of pBluescript. pMV306 is a commonly used mycobacterial shuttle vector.
pRAS21	In this plasmid a 4.4 kb *EcoRI* sub-digest of the 6.6 kb *SacI* fragment in pRAS19/20 is inserted into the *EcoRI* site of pMV306. The regulatory gene Rv0903c is lost in entirety and only the last 281 bp of gene Rv0902c (the sensor) remain on the complimentary strand. The upstream *EcoRI* site is 1681 bp before the start codon of the *ompATb* gene.

ATTEMPTED KNOCKOUT OF THE REGULATORY GENE Rv0903c

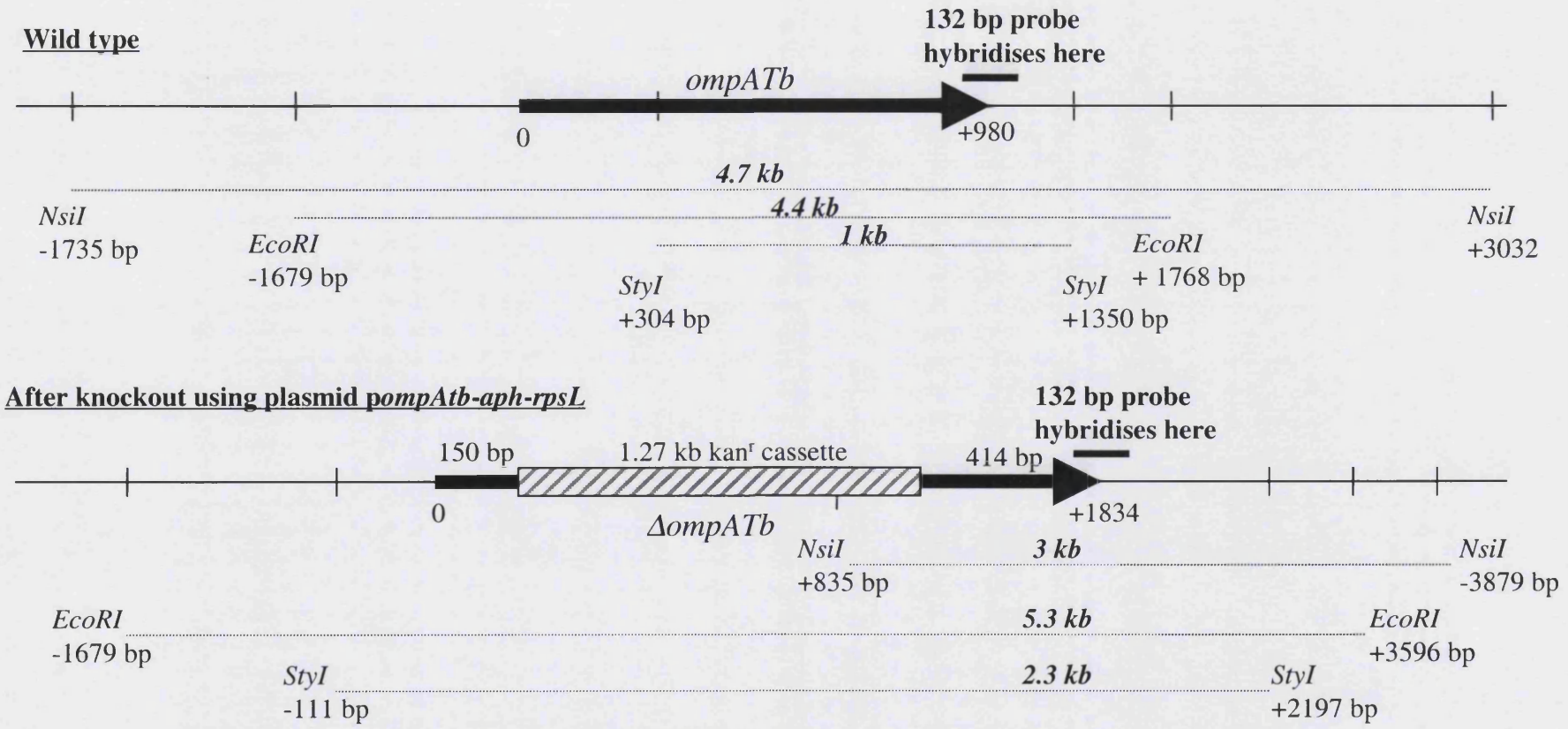
pRAS14	A 1769 bp region, ending 1741 bp after the stop codon of *M. tb* gene Rv0903c, is cloned into a *HindIII/XhoI* digest of plasmid p2NIL (Parish and Stoker, 2000). p2NIL has a kanamycin resistance cassette and a specialised cloning site to allow the later insertion of a *PacI* fragment from pGOAL19 (Parish and Stoker, 2000).
pRAS17	A 2038 bp region 5' to the coding sequence of the Rv0903c gene, including the first 8 bp of coding sequence, is cloned into the *KpnI* site of pRAS14.
pRAS18	A *PacI* digest of pGOAL19 (Parish and Stoker, 2000) releases a *P*Ag85 *lacZ* gene, a *P*hsp60 *sacB* gene and a hygromycin resistance cassette. Cloning this fragment into the *PacI* site of pRAS17 leads to plasmid pRAS18 which has 4 separate selectable markers and contains regions of homology 5' (2030 bp) and 3' (1741 bp) to the Rv0903c coding sequence. The majority of the coding sequence of gene Rv0903c has been removed, only the 5' 8 bases and the 3' 28 base pairs remain, the remainder of the gene (674 bp) has been replaced by vector DNA from between the *SalI* and *KpnI* sites of p2NIL (312 bp).



Supplementary figure I

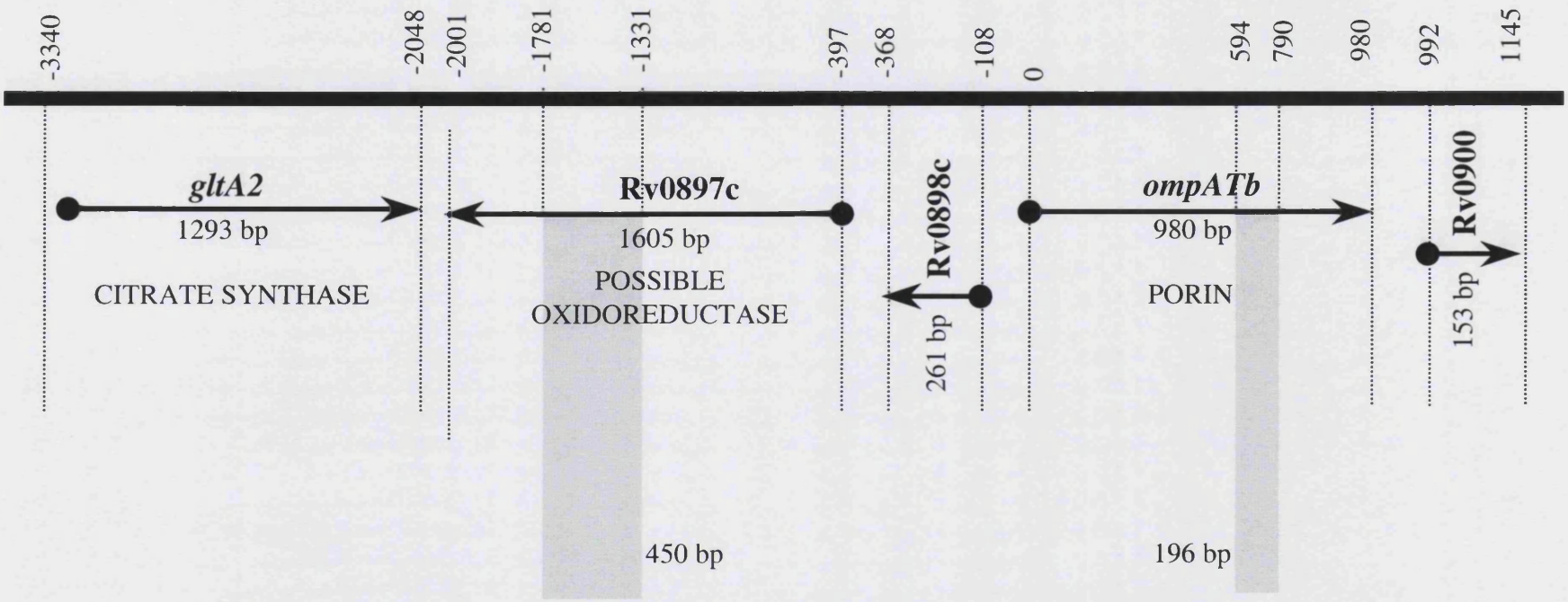
A diagram of the *Mycobacterium tuberculosis* genome around the region coding for the porin gene *ompATb*

Sizes of the genes are given along with their functions where known, and their positions relative to the *ompATb* translational start codon



Supplementary Figure II

The wild type and disrupted *ompATb* gene, showing the insertion of a kanamycin resistance cassette in order to prevent functional OmpATb protein being produced.
The vector used to knock out the gene contained an *EcoRI* fragment of the M. tb genome (see above)
Restriction fragments used in Southern blot verification of the knockout are shown, along with the position of the probe used. Positions of the restriction sites are given relative to the start codon of the *ompATb* gene.

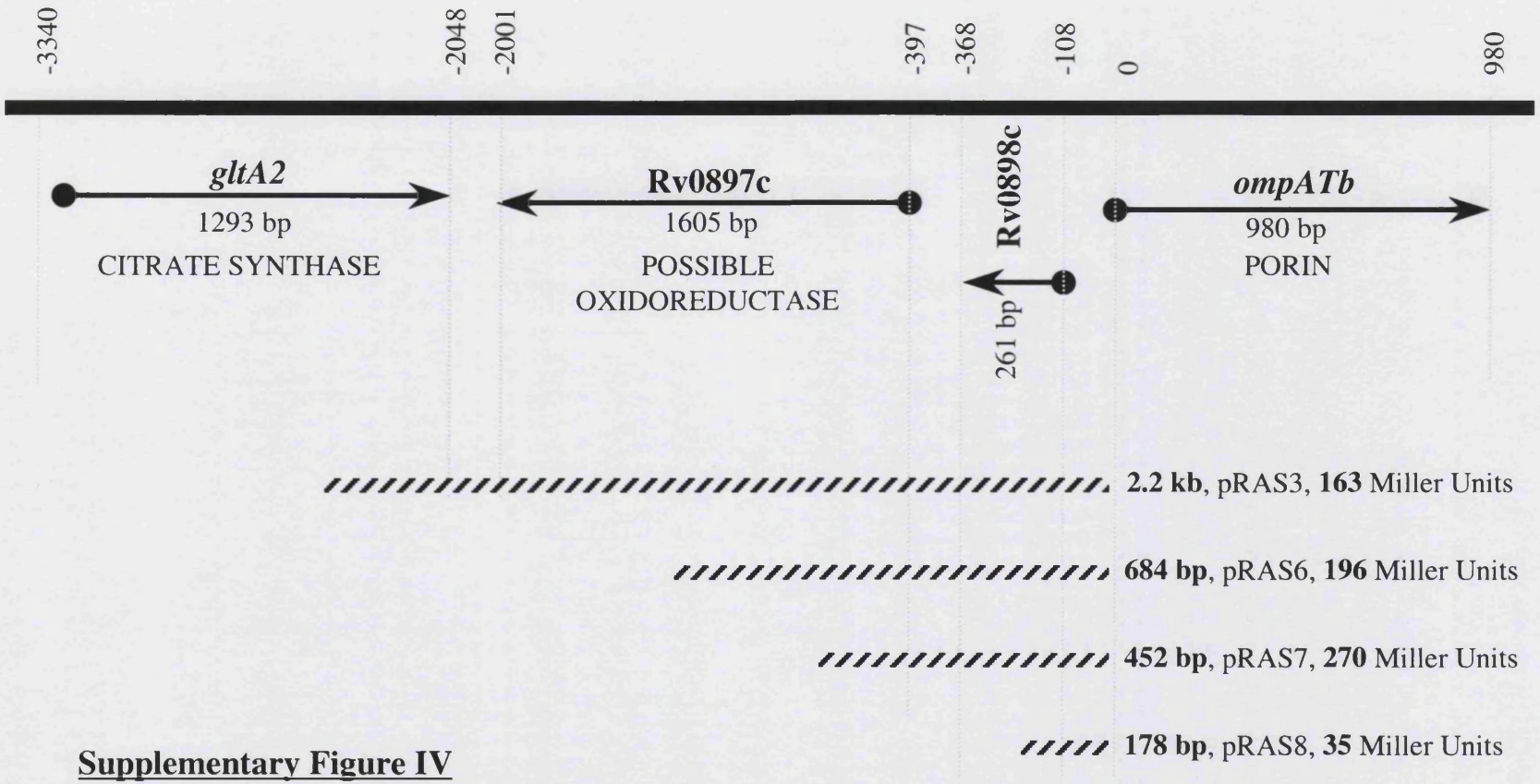


Binding of Rv0903 protein occurs in these shaded regions

Supplementary Figure III

The binding sites of recombinant Rv0903 protein as determined by gel retardation assay are shown in relation to the position of the surrounding genes.

Numbers along the top indicate position relative to the *ompATb* start codon



Supplementary Figure IV

Position of fragments used in β-galactosidase assays to determine promoter strength of the region upstream of *ompATb*.

Length of fragments used, plasmid names and activity of β-galactosidase (Miller Units) is given.

It can be seen that the promoter activity lies largely between 178 and 452 bp upstream of the *ompATb* coding sequence.