STUDIES ON THE CATALASE GENES AND THE ACID TOLERANCE RESPONSE OF MYCOBACTERIA.

Thesis submitted for the degree
of Doctor of Philosophy at the University of Leicester.

by

Stephen Gordon,
Department of Microbiology and Immunology,
University of Leicester.

January 1995

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

For my parents

Statement of Originality

The accompanying thesis submitted for the degree of Ph.D. entitled "Studies of the catalase genes and acid tolerance response of mycobacteria" is based on work conducted by the author in the Department of Microbiology and Immunology of the University of Leicester mainly during the period between October 1991 and October 1994.

All the work recorded in this thesis is original unless otherwise acknowledged in the text or by references.

None of the work has been submitted for another degree in this or any other University.

Signed: Stephen Gardner Date: 1st JUNE 1995

Abstract

Members of the genus *Mycobacterium* are major pathogens of man and animals. Tuberculosis causes almost 3 million deaths per year, currently making it the leading cause of death among bacterial infectious diseases. However, our knowledge of these pathogens is limited. This thesis investigated two questions (i) the role of mycobacterial catalases in isoniazid sensitivity, and (ii) the presence of an acid tolerance response.

The mycobacterial M-catalase is thought to play a role in isoniazid sensitivity. Bacterial catalase enzymes are highly conserved at the amino acid level. An experimental strategy to clone the mycobacterial M-catalase gene sought to exploit this feature. Degenerate oligonucleotide primers were designed, based on previously cloned catalase genes, in an attempt to amplify the mycobacterial catalase gene using the polymerase chain reaction (PCR). Despite repeated attempts, this approach was unsuccessful. An attempt to identify the gene from a mycobacterial library using the catalase gene of *Escherichia coli* (*katE*) as a hybridisation probe also failed.

Resistance to acidic environmental conditions is important for many mycobacterial species. Evidence was sought for the existence of an adaptive acid tolerance mechanism in the mycobacteria. To identify any genes that showed increased transcription in response to a decrease in extracellular pH, a promoter probe library was constructed with bacterial luciferase as the reporter. A clone (pSGA197) was identified that showed increased bioluminescence at the adaptive pH. The mycobacterial DNA insert in this clone was sequenced. Examination of the sequence revealed that the insert's putative acid-regulated promoter was located downstream from an open reading frame that coded for a possible sensor-kinase protein.

Acknowledgements

My sincere thanks to Dr. Peter Andrew for his supervision, encouragement and scintillating wit during the course of my Ph.D.

Thanks to all my eminent colleagues, past and present, in the mycobacteria group, namely Barrie, Lyn, Rachel, Sarah, Seamus, Bob, Carolyn, Jenny and Vincent. Special thanks to Pat for all her help with the catalase project, and for all the Gilson tips I "borrowed" from her.

Many, many thanks to everyone from the department for all their help over the last three years. Although I can't name everyone, I feel I should mention the following individuals: Archie "The Peach" Lovatt, Aras "Whirlwind" Kadioglu, Dave "Geiger-man" Simpson, Fred "Guinness" Esumeth, Janette "Wimp!" Maley, Chris "Footie" Jones, and Esperanza "It's rubbish" Cancio. Also, my utmost gratitude must go to Miguel for his delightful Spanish wines, and in the aftermath to Ben and Anna for their concern in my well-being.

I am indebted to Ireland's Holy Trinity of Eddie Lavelle, Peter Cox and Ger Colleran for their towering intellects and scientific musings. "Go raibh míle maith agaibh!"

Thanks to Dawn and Pete for the Arctic survival course.
PS. You're both crazy.

Extra special thanks to Fiona for all her proof-reading and help during my Ph.D. I owe you a truck-load of chocolate!

Finally, thanks to my family for all their love and support during my studies.

Abbreviations

ADC	Albumin Dextrose Catalase
ATR	Acid tolerance response
BCG	bacille Calmette-Guérin
bp	Base pair
BSA	Bovine serum albumin
CFU	Colony forming units
CIP	Calf intestinal alkaline phosphatase
Da	Dalton
DMSO	Dimethyl sulphoxide
DTT	Dithiothreitol
EDTA	Diaminoethantetra-acetic acid
FMN	Flavin mononucleotide
IPTG	Isopropyl-b-D-thiogalactosylpyranoside
kb	kilobase
kD	kilodaltons
NAD	Nicotinamide adenine dinucleotide
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
pH_i	Cytoplasmic pH
pHo	External pH
RLU	Relative light unit
SDS	Sodium dodecyl sulphate
SOD	Superoxide dismutase
SSC	Saline Sodium Citrate
TE	Tris-EDTA
TEMED	N,N,N,N'-tetramethylethylethylenediamine
X-Gal	5-bromo-4-chloro-3-indoyl-b-D-galactopyranoside
TB	Tuberculosis
UV	Ultraviolet
V	Volts
W	Watts

Table of Contents

Chapter 1.

Introduction

1.1 The genus *Mycobacterium*

The generic name *Mycobacterium*, meaning fungus-bacterium, was first introduced in 1896 by Lehmann and Neumann (Grange, 1988). The name did not imply a relationship to fungi, but rather described the mould-like pellicular growth of the tubercle bacillus on liquid medium. Although the two initial members of the genus were the human pathogens *Mycobacterium tuberculosis* and *Mycobacterium leprae*, it is now recognised that mycobacteria are generally free-living saprophytes. A number of detailed descriptions of mycobacterial classification are available (see Goodfellow and Wayne, 1982; Wayne and Kubica, 1986), so only the salient features of the group shall be dealt with here.

The genus *Mycobacterium* is broadly divided into the slow and fast growers. Slow growers are defined as taking more than 7 days to produce colonies on solid medium visible to the naked eye, while fast growers take less than 7 days (Goodfellow and Wayne, 1982). Comparative analysis of 16S rRNA sequence in the mycobacteria supports this division, with the slow growing species defining a distinct line of evolutionary descent (Stahl and Urbance, 1990). Analysis of the genome structure of mycobacteria has revealed a high guanine plus cytosine (G+C) content, in the range of 64 to 70% (Clark-Curtiss, 1990). Investigations of genome relatedness using DNA hybridisation assays have confirmed groupings based on classical methods of phylogeny (Grosskinsky *et al.*, 1989). Genome size determinations have shown that mycobacteria have larger genomes than most other prokaryotes, in the range of 3.1-4.5 x 10^9 Da (Clark-Curtiss, 1990).

Mycobacteria are resistant to decolouration with weak mineral acids or acid-alcohol after staining with basic dyes, a property defined as acid-fastness. Although acid-fastness is not a unique property of mycobacteria, it is widely used to indicate the presence of mycobacteria in clinical specimens (Heifets and Good, 1994). The chemical basis of acid-fastness is not fully understood, but the mycolic acid constituent of the mycobacterial cell wall appears to be involved (Grange, 1988).

The mycobacterial cell wall is a lipid-rich, highly complex structure. It is responsible for the adjuvant properties of Freund's complete adjuvant, and as such has attracted detailed investigation (Draper, 1982; Brennan and Draper, 1994). A distinctive feature of the cell wall is the presence of long chain mycolic acids, branched fatty acids with a long branch (40-64 carbons) and a shorter branch (22-24 carbons). Two models have been proposed for the structure of the mycobacterial cell wall (Minnikin, 1982; Rastogi, 1991). Both models propose that the mycolic acids are covalently bound to an arabinogalactan layer, in turn bound to the peptidoglycan layer, with their hydrocarbon chains parallel and their methyl ends pointing to the cell exterior. The Minnikin model proposes that the hydrocarbon chains of the mycolic acids are intercalated with the hydrocarbon chains of other cell wall associated lipids. On the other hand, the Rastogi model suggests that the hydrocarbon side chains are arranged in a classic lipid bilayer structure. Currently, experimental evidence favours the Minnikin model, with McNeil and Brennan (1991) including lipoarabinomanam and protein in the structure. It is also worth noting that a related cell wall mycolic acid, cord factor (trehalose dimycolate), is involved in the production of the characteristic "serpentine cords" of tubercle bacilli (Behling *et al.*, 1993).

The lipid rich nature of the mycobacterial cell wall serves the organism well as a shield against the external environment, but creates problems for the uptake of hydrophilic molecules. Indeed, the diffusion of small nutrients across the cell wall of *Mycobacterium chelonae* was greater than 10 000 times slower than through the outer membrane of *E. coli* (Jarlier and Nikaido, 1990). Mycobacteria overcome this limitation by means of protein pores in the cell wall, akin to the porins found in the outer memebranes of Gram negative organisms. Mycobacterial porins were first described in *M. chelonae* (Trias *et al.*, 1992) and later in *M. smegmatis* (Trias and Benz, 1994). These porins were found to be cation specific, with a wide pore diameter (2-3 nm). Crucially, the porins are only minor proteins of the mycobacterial cell wall, unlike the porins of the gram negative outer membrane (Jarlier and Nikaido, 1994). This fact explains the low permeability of the mycobacterial cell wall to hydrophilic compounds.

Metabolically, mycobacteria are essentially the same as most other bacteria. Mycobacterial metabolism has been reviewed in

detail by Ratledge (1982) and Wheeler and Ratledge (1994). Most species have simple nutrient requirements and are able to use a wide variety of carbon and nitrogen sources. A preference for glycerol and asparagine as carbon and nitrogen sources respectively has been observed, but the reason for this is unknown (Ratledge, 1982). An important growth factor is iron, and mycobacteria are unique in synthesising two distinct classes of siderophores, namely exochelins and mycobactins (Ratledge, 1982, 1984). Exochelins are water soluble, low molecular weight peptides, while mycobactins are non-water soluble and are located in the cell wall. Mycobactins are synthesised by all mycobacteria except *M. paratuberculosis* and some strains of *M. avium*, so to cultivate these strains *in vitro* it is necessary to supply mycobactins in the growth medium.

1.2 Tuberculosis

1.2.1 Brief historical overview

In the nineteenth century, tuberculosis was unquestionably the greatest single cause of disease and death in the Western world (Dubos and Dubos, 1952). The rapidly increasing population densities in the cities of Western Europe provided the ideal conditions for person-to-person spread of the pathogen, and it is estimated that at this time tuberculosis was responsible for one-seventh of all deaths in Europe (Daniel, 1982).

Tuberculosis is an ancient disease of man, but before the fifteenth century it was a relatively unimportant disease. The beginning of the TB epidemic can be traced to early fifteenth century England and Western Europe, where TB death rates began to rise sharply. This trend was to continue for the next two centuries, and would earn tuberculosis the title of "The White Plague" (Dubos and Dubos, 1952; Daniel *et al.* 1994). However, during this period the infectious nature of tuberculosis was not appreciated; it took until 1865 for Jean-Antoine Villemin to demonstrate that TB was contagious by transferring sputum and pus from human TB lesions to rabbits that then developed tuberculosis (Dubos and Dubos, 1952).

The isolation of the tubercle bacillus, *Mycobacterium tuberculosis*, by Robert Koch in 1882 and his demonstration that it was

the aetiological agent of tuberculosis in humans heralded a new era of research into TB (Koch, 1932). The culture and staining techniques that Koch and colleagues developed for *M. tuberculosis* were fundamental tools for researchers, while his criteria for proof that the isolated organism was the cause of tuberculosis (Koch's postulates) are now a fundamental concept in the study of infectious disease.

The isolation of *M. tuberculosis* provided the starting point for the development of a vaccine against tuberculosis. Pasteur had discovered that attenuated forms of virulent viruses could be used to vaccinate against viral disease, and this approach was adopted by Calmette and Guérin to produce the BCG (bacille Calmette-Guérin) vaccine (Bloom and Fine, 1994). By repeated subculture of *Mycobacterium bovis* in potato-glycerin-ox bile medium, an attenuated strain was produced which protected animals from challenge with virulent bovine or human tubercle bacilli. BCG was first used as a vaccine in man in 1921, and since then has been administered to over 3 billion people, making it the most widely used vaccine ever (Bloom and Fine, 1994).

The development of chemotherapy for tuberculosis started in the late 1940s and early 1950s with the introduction of streptomycin, p-aminosalicylic acid and isoniazid (Mitchison, 1992). These drugs were mainly given as double-drug combinations for periods of 1 to 2 years. Levels of patient compliance with such prolonged drug regimens were difficult to maintain, so shorter course chemotherapy was needed. The introduction of rifampicin in 1966 allowed the development of "short-course" chemotherapeutic regimens of nine months (Mitchison 1992, Brausch and Bass, 1993). The current six month course of chemotherapy was developed using pyrazinamide, with isoniazid and rifampicin as the other central drugs (Hopewell, 1994). In the UK, the regimens recommended by the Joint Tuberculosis Committee of the British Thoracic Society are an initial phase of two months using pyrazinamide, isoniazid and rifampicin, followed by a continuation phase lasting 4 months using isoniazid and rifampicin (British National Formulary, Number 27, March 1994).

Even before the arrival of a vaccine or drugs for the treatment of tuberculosis, death rates from the disease had been decreasing in the Western world since the middle of the nineteenth century (Dubos and Dubos, 1952). The widespread availability of both

the BCG vaccine and chemotherapeutic drugs from the 1950s onwards, coupled with the decreasing incidence of tuberculosis saw the funding for tuberculosis research dwindle in the 1970s (Bloom and Murray, 1992). Still, tuberculosis remained a major health problem in the developing world, and problems with the protective efficacy of the BCG vaccine were never properly resolved (Fine, 1988). The early 1980s saw an ominous increase in reported cases of TB in the United States. This increase could be due to a number of factors acting in combination, including the HIV epidemic, increasing numbers of homeless people and intravenous drug users in major cities, and the dismantling of healthcare programs (Bloom and Murray, 1992). The emergence of multiple-drug resistant forms of *M. tuberculosis* (MDRTB) in the USA is now a major cause of concern. Death rates from multidrug-resistant tuberculosis in the United States are about 40% for immunocompetent individuals and over 80% for HIV-infected individuals (Bloom, 1994).

The WHO declared tuberculosis a global emergency in April 1993, and current estimates are that the annual death rate from tuberculosis will reach 3.5 million by the year 2000 (WHO press release, WHO/8, 25 January 1994). Still, we rely on a vaccine that shows variable levels of protection and on drug regimens that are relatively long and expensive. In addition, the appearance of multi-drug resistant *M. tuberculosis* now poses a major threat to public health. In this climate, research into the basic biology and pathogenic mechanisms of *M. tuberculosis* and other mycobacteria is of vital importance.

1.2.2 Pathogenesis and immunity

The pathogenesis of tuberculosis and the host's immune response have been detailed in many excellent reviews (Youmans, 1979; Grange, 1988; Dannenburg and Rook, 1994; Rook and Bloom, 1994). This section will therefore only present an outline of the subject, with particular reference to the possible virulence factors of the tubercle bacillus.

Primary infection is used to refer to the initial infection of the host with the tubercle bacillus (Youmans, 1979). Infection usually occurs by inhalation of airborne infectious particles (1-5mm in size) containing 1-3 bacilli. These particles are small enough to reach the terminal alveoli of the mid and lower lung where they are phagocytosed

by the resident alveolar macrophages (Dannenburg and Rook, 1994). Phagocytosis by an incompetent macrophage allows multiplication of the bacilli, a process that eventually leads to lysis of the macrophage and release of bacilli. Ingestion of these bacilli by other alveolar macrophages and polymorphonuclear cells triggers an inflammatory response, the first stages in the formation of a granuloma (Grange 1988). Failure to contain the infection at this stage, usually in children or immunocompromised individuals, leads to dissemination of bacilli from the primary focus by the lymph and blood systems. This can result in tuberculoid meningitis and/or millary TB.

In 90 to 95% of infected individuals the result of primary tuberculosis is recovery (Youmans, 1979). In some individuals however, the bacilli can persist in a state of dormancy, usually in the apices of the lung, resistant to sterilisation by antimicrobials (Rook and Bloom, 1994). Secondary tuberculosis may result from reactivation of these dormant bacilli, or by reinfection with exogenous bacilli. Both mechanisms presumably rely on a weakening or suppression of the immune system, one reason for the association of tuberculosis with HIV positive individuals. Secondary tuberculosis is characterised by a vigorous delayed-type hypersensitivity (DTH) response, with lesions showing marked central necrosis and caseation (Dannenburg and Rook, 1994). Liquefaction of the caseous material and surrounding tissue by macrophage enzymes gives the bacilli an excellent growth medium for extracellular replication. This leads to an escalation of the tissue-damaging DTH response, causing extensive necrosis to surrounding tissue (Dannenburg and Rook, 1994). Rupturing of the bronchial walls leads to cavity formation, releasing the tubercle bacilli into the bronchi and disseminating the bacilli to the rest of the lung and the exterior (Grange, 1988).

1.2.3 Virulence factors

Survival inside macrophages is the key aspect of *M. tuberculosis* virulence. Resistance to and avoidance of the antimicrobial mechanisms of the macrophage allows the bacillus to grow in the phagocytic cell, "hidden" from the rest of the immune system. Characterisation of the factors and strategies employed by the bacillus to effect this survival has greatly increased our understanding of the pathogenesis of tuberculosis. This section shall examine the mechanisms

by which *M. tuberculosis* is thought to avoid some of the key antimicrobial systems of the host.

1.2.3.1 Uptake into the macrophage and avoidance of lysosomes

The mechanism of uptake of *M. tuberculosis* dictates to a large extent how the macrophage can kill the bacillus. *M. tuberculosis* promotes its own uptake by binding to C3bi and fibronectin receptors on the macrophage surface (Schlesinger *et al.*, 1990; Abou-Zeid *et al.*, 1988). Hence, by avoiding phagocytosis involving Fc receptors, *M. tuberculosis* fails to trigger a respiratory burst. The isolation by Arruda and colleagues (1993) of a possible invasion gene from *M. tuberculosis* points to a further possible pathway for internalisation of the bacillus.

Once internalised in a phagosome, avoidance of lysosomes is achieved by inhibition of phagosome-lysosome fusion (Armstrong and Hart, 1971). The ability to inhibit phagosome-lysosome fusion is dependant on living *M. tuberculosis*, as phagosomes containing dead (or opsonised) *M. tuberculosis* fuse with phagosomes (Andrew *et al.*, 1985). Acidification of the phagosome may be an important condition for fusion with the lysosome (Crowle *et al.*, 1991). It has been suggested that virulent *M. tuberculosis* blocks this acidification, possibly by the production of ammonia (Crowle *et al.*, 1991). Also, Sturgill-Koszycki *et al.* (1994) have suggested that *M. avium* blocks acidification of the phagosome by selectively excluding fusion with vesicles containing the proton-ATPase that mediates acidification. Production of cyclic adenosine monophosphate (Lowrie *et al.*, 1975) or the action of cell wall sulpholipids (Goren *et al.*, 1976) have also been proposed as mechanisms involved in inhibition of phagosome-lysosome fusion.

McDonough *et al.* (1993) have suggested that some strains of *M. tuberculosis* may escape from the phagosome into membrane bound vesicles that fail to fuse with secondary vesicles. Controversially, they also reported that virulent, but not attenuated, *M. tuberculosis* can be found free in the cytoplasm of infected macrophages. This is an interesting observation when taken with that of King and colleagues (1993), who showed that virulent *M. tuberculosis* expressed a greater than 3-fold higher level of contact-dependant cytolytic activity than attenuated *M. tuberculosis*. Whether this contact-dependant cytolytic

activity is important in escape from the phagosome, or indeed entry into the macrophage, is not known.

1.2.3.2 Persistence in the macrophage

Persistence of *M. tuberculosis* in the macrophage is an essential step in the disease process. Yet the presence of the bacillus in the macrophage can initiate an immune response, activating the macrophage to kill the bacillus. The bacillus must therefore attempt to block this activation, and/or resist the antimicrobial mechanisms of the macrophage.

Lipoarabinomannan (LAM) is a phosphorylated lipopolysaccharide found in the cell walls of mycobacteria that has been identified as a possible virulence determinant. LAM from *M. tuberculosis* and *M. leprae* is capable of down regulating the effector functions of macrophages by (i) scavenging oxygen free radicals, (ii) blocking the transcriptional activation of gamma interferon (g-IFN) inducible macrophage genes, (iii) inhibiting the activity of protein kinase C, which is involved in macrophage activation and triggering of the respiratory burst (Chan *et al.*, 1989; 1991). Structural differences in LAM from strains of mycobacteria may explain degrees of virulence. For example, purified LAM from avirulent *M. tuberculosis* is over a hundred-fold more potent at inducing TNF-a than LAM from virulent *M. tuberculosis*, due to structural differences at the nonreducing ends (Adams *et al.*, 1993; Roach *et al.*, 1993). Also, LAM from avirulent, but not from virulent, strains of *M. tuberculosis* induces expression of the genes for c-fos, KC and JE, which are involved in macrophage activation (Roach *et al.*, 1993). It appears therefore that LAM from virulent *M. tuberculosis* is particularly suited to allow an *in vivo* lifestyle.

1.2.3.3 Additional putative virulence determinants

The catalase-peroxidase (T-catalase) of *M. tuberculosis* has been proposed as a virulence factor, since catalase-negative, isoniazid-resistant strains have reduced virulence for the guinea pig (Middlebrook and Cohn, 1953). However, as killing of *M. tuberculosis* by activated guinea pig macrophages *in vitro* is not dependent on products of the respiratory burst (O'Brien *et al.*, 1990; O'Brien and Andrew, 1991), the role of catalase is not clear. It may be worth noting that the catalase-

peroxidase is thought to play a role in transport of isoniazid into the bacillus (Wimpenny, 1967). Therefore, assuming that isoniazid is not the only substrate for this transport activity, the role of catalase in virulence may be in transport of essential nutrients or growth factors into the bacillus during intracellular growth (see Section 3.2. for a fuller account of mycobacterial catalase).

Growth rate of *M. tuberculosis in vivo* has also been correlated with virulence. North and Izzo (1993) have established that two virulent strains of *M. tuberculosis*, Erdman and H37Rv, have faster doubling times in mice *in vivo* than attenuated strains. The particular advantage given by a faster growth rate is not obvious, although North and Izzo suggest that it may allow the bacillus to overload the antigen presenting capacity of the macrophage, making the bacillus invisible to the cell-mediated immune system. The recent isolation of a genomic locus from *M. tuberculosis* H37Rv that increases the *in vivo* growth rate of *M. tuberculosis* H37Ra may allow a genetic insight into this phenomenon (Pascopella *et al.*, 1994).

The *recA* genes of *M. tuberculosis* and *M. leprae* have a peculiar structure that seems to be related to virulence. Davis *et al.* (1991; 1992; 1994) have demonstrated the presence of an intragenic spacer region (intein) in the *recA* locus of these pathogens that is spliced from the translated peptide to produce a mature RecA protein and a spacer protein. The fascinating aspect of the inteins is that they are present in different locations in the *recA* genes of *M. tuberculosis* and *M. leprae*, and that the sequence of the inteins shows little homology. Therefore it is reasonable to assume that there has been selection for either the splicing mechanism or the intein in these pathogens (Colston and Davis, 1994). However, the possible role that they may play in the virulence of these pathogens is not clear.

1.3 Molecular Biology of the Mycobacteria

1.3.1 DNA and RNA

1.3.1.1 Base composition and size of genome

Early work on the genetic relatedness of mycobacteria used base ratio analysis and DNA hybridisation assays between the genomes of different mycobacteria (Wayne and Gross, 1968; Gross and Wayne, 1970; Bradley, 1973). Most mycobacterial DNA was found to have a guanine plus cytosine (G+C) content of between 64 and 70%, the exceptions to this being *Mycobacterium leprae* and "*Mycobacterium lufu*" which have G+C contents of 58 and 61% respectively (Clark-Curtiss *et al.*, 1985). DNA hybridisation analysis established or confirmed related groups of mycobacteria by demonstrating amounts of homology among DNA sequences. For example, Grosskinsky and colleagues (1989) revealed that the DNAs of the slow growing mycobacteria *M. tuberculosis*, *Mycobacterium africanum*, *Mycobacterium bovis* and *Mycobacterium microti* were very closely related, while *M. leprae* was found to be only remotely related to any of the mycobacteria tested. Similarly, McFadden *et al.* (1987) found that *M. paratuberculosis* was virtually indistinguishable from *M. avium* on the basis of DNA renaturation.

Measurements of genome size in mycobacteria have revealed that most mycobacteria have large genomes, in the range of 3.1-4.5 x 10^9 Da, in relation to other prokaryotes (Clark-Curtiss, 1990). The exception to this is the pathogenic mycobacteria, whose genome size is smaller: *M. leprae* is estimated to have a genome of 1.3-2.2 x 10^9 Da (Clark-Curtiss *et al.*, 1985; Imeada *et al.*, 1982), while *M. tuberculosis* and *M. bovis* BCG have genome sizes of 2.5 x 10^9 and 2.8 x 10^9 Da respectively (Bradley, 1973).

1.3.1.2 Repetitive DNA elements

The presence of repetitive DNA elements in mycobacteria was first demonstrated by restriction fragment-length polymorphism (RFLP) analysis of mycobacterial DNA. Probes that contained DNA from repetitive elements produced multiple bands on hybridisation to restriction fragments of chromosomal DNA. This led to the identification of IS*900* in *M. paratuberculosis* (Green *et al.*, 1989),

IS*6110* (or IS*986*) in *M. tuberculosis* (Eisenach *et al.*, 1988; Zainuddin and Dale, 1989), and RLEP in *M. leprae* (Grosskinsky *et al.*, 1989; Woods and Cole, 1990). The use of "transposon traps" isolated the insertion sequences IS*1096* (Cirillo *et al.*, 1991) and IS*6120* (Guilhot *et al.*, 1992), both from *M. smegmatis*. The insertion sequence IS*6100*, isolated from *M. fortuitum*, was found to flank a sulphonamide resistance gene and a truncated site-specific integrase, part of a transposon-like element Tn*610* (Martin *et al.*, 1990). Although many repetitive elements present in mycobacterial DNA are insertion sequences, the most abundant repetitive elements in *M. tuberculosis* are the major polymorphic tandem repeats (MPTRs). These are 10 base pair (bp) repeating sequences separated by unique 5bp spacer sequences (Hermans *et al.*, 1992). The MPTRs have been shown to be highly polymorphic in *M. kansasii* and *M. gordonae*, and hence may be useful in DNA fingerprinting analysis of these species (Hermans *et al.*, 1992).

The availability of mycobacterial insertion sequences provides a powerful tool for genetic manipulation of mycobacteria, as well as having diagnostic applications. Guilhot and colleagues (1994) have established that Tn*611*, a derivative of Tn*610* in which sulphonamide resistance is replaced by kanamycin resistance, is efficient at producing insertional mutant libraries in *M. smegmatis*. By introducing the transposon into *M. smegmatis* via a temperature-sensitive shuttle vector, then growing transformants at the non permissive temperature, they were able to show (a) that there was no evidence of a hot-spot for insertion, and (b) that the integration events were stable. The application of this technology to the pathogenic mycobacteria may prove effective in isolating virulence factors, an approach that has been successful with *Salmonella* (Fields *et al.*, 1986; Fields *et al.*, 1989). An alternative use of insertion sequences is the production of integrative vectors. Dellagostin *et al.* (1993) used IS900 to construct an integrative expression vector which was able to drive the expression of foreign proteins in *M. smegmatis*.

The most useful IS-element from the diagnostic point of view is IS*6110*, which is specific for the *M. tuberculosis* complex (Hermans *et al.*, 1990(a). Most *M. tuberculosis* strains have between 5 and 20 copies of IS*6110*, although some strains, mainly from Asia, may have one or no copies (van Soolingen *et al.*, 1993). DNA-fingerprinting analysis using IS*6110* is becoming a standard epidemiological tool, and

has been used to identify clonal transmission between neighbours, prisons, outpatient clinics and hospitals (Small and van Embden, 1994). The establishment of a database containing DNA fingerprints of *M. tuberculosis* isolates from different geographical locations will prove an invaluable tool to the epidemiology of tuberculosis (Small and van Embden, 1994).

1.3.1.3 Mycobacterial plasmids

The presence of plasmids in mycobacteria was first demonstrated in the *M. avium* complex by Crawford and Bates (1979). Subsequent work has shown plasmids to be present in a range of fast and slow growing mycobacteria, with the *M. avium* complex studied in most detail (for reviews see Crawford and Falkinham, 1990; Falkinham and Crawford, 1994). Plasmid encoded functions include a restriction-modification system in *M. avium* LR25 (Crawford *et al.*, 1981), resistance to mercury and copper in *M. scrofulaceum* (Meissner and Falkinham, 1984; Eradi *et al.*, 1987) and control of catalase gene expression in *M. avium* (Pethal and Falkinham, 1989). The observation that a high percentage of clinical isolates of *M. avium* complex, particularly from AIDS-associated isolates, harboured plasmids led to the hypothesis that plasmids may play a role in virulence of these strains (Meissner and Falkinham, 1986). However, this hypothesis has yet to be confirmed (Falkinham and Crawford, 1994).

The most extensively studied mycobacterial plasmid has been pAL5000, a small cryptic plasmid isolated from *M. fortuitum* (Labidi *et al.*, 1985). Snapper and colleagues (1988) constructed a mycobacterium-*E. coli* shuttle vector, pYUB12, based on this plasmid that allowed stable expression of the kanamycin-resistance gene from Tn*5* in both *M. smegmatis* and *M. bovis* BCG. The complete nucleotide sequence of the plasmid was determined by Rauzier *et al.*, (1988), and Stover and co-workers (1991) constructed mycobacterium-*E. coli* shuttle expression vectors based on a minimal 1.8kb *ori* region of pAL5000. By constructing a hybrid plasmid from the *ori* region of pAL5000, the origin of transfer of pRK2, and a kanamycin resistance gene, Lazraq and co-workers (1990) were able to show conjugative transfer of plasmid DNA between *E. coli* and *M. smegmatis*. Plasmids based on temperature sensitive mutants of the pAL5000 *ori* region have been isolated by Guilhot *et al.* (1994), and these have proved to be

useful as vehicles for transposon mutagenesis of *M. smegmatis* (see Section 1.3.2).

1.3.1.4 Mycobacteriophages

Since the first description of a lytic mycobacteriophage by Gardner and Weiser (1947), mycobacteriophages have proved to be useful tools for analysing mycobacteria. Most mycobacteriophages have a broad host range and display the ability to lyse a few strains of different species, but few phages can lyse all strains of the same species. Therefore, phage typing of mycobacteria is used to subdivide known species for epidemiological purposes (reviewed by Grange, 1982).

The construction by Jacobs and colleagues (1987) of the shuttle vector phAE1, consisting of mycobacteriophage TM4 into which the *E. coli* cosmid vector pHC79 had been inserted, was the first description of the introduction of foreign DNA into mycobacteria. Shuttle vectors based on TM4 have also been used to construct luciferase reporter phages (Jacobs *et al.*, 1993). However, as TM4 is lytic for BCG and *M. smegmatis*, it usefulness as a means for stable expression of foreign DNA in mycobacteria is limited. The development of shuttle vectors that allowed stable expression of foreign DNA in mycobacteria was achieved using the temperate mycobacteriophage L1 (Snapper *et al.*, 1988). Similarly, David *et al.* (1992) constructed a *M. smegmatis-E. coli* shuttle vector using the origin of replication from mycobacteriophage D29, and used this construct to express bacterial luciferase in *M. smegmatis*.

The best characterised of the mycobacteriophage is L5 (for review see Hatfull and Jacobs, 1994). The complete nucleotide sequence of this phage has been determined (Hatfull and Sarkis, 1993), and isolation of the phage *att*P site and integrase gene has allowed the production of site-specific integration vectors that are maintained in mycobacteria (Lee *et al.*, 1991; Stover *et al.*, 1991). L5 is now being adapted to function as a lysogenic luciferase reporter phage (Hatfull and Jacobs, 1994).

1.3.1.5 Ribosomal RNA

Analysis of ribosomal RNA (rRNA) copy number in mycobacteria revealed that fast growing mycobacteria had two copies of the rRNA genes per chromosome, while slow growers had one copy per

chromosome (Bercovier *et al.*, 1986; Sela *et al.*, 1989). Since fast-growing prokaryotes have multiple copies of their rRNA genes, e.g. *E. coli* has seven (Brosius *et al.*, 1981), it was suggested that low copy number of rRNA genes in mycobacteria may cause their slow growth (Bercovier *et al.*, 1986). However, rRNA copy number is not the sole determinant of slow growth rate, as the fast growing *Mycobacterium chelonae* and *Mycobacterium abscessus* appear to have only a single 16S rRNA gene (Domenech *et al.*, 1994).

The structure of the rRNA (*rrn*) operon in the slow growing mycobacteria has been found to be identical to that in many other prokaryotes, with genes for 16S, 23S and 5S arranged in that order (Bercovier *et al.*, 1986). Indeed, despite its extremely slow growth, *M. leprae* has a typical mycobacterial *rrn* operon (Ji *et al.*, 1994; Sela *et al.*, 1989). Analysis of the *rrn* operons in slow growing mycobacteria has identified that transcription is started from a single promoter, and that the promoter region sequences are very similar to consensus *E. coli* promoter sequences (Sela and Clark-Curtiss, 1991; Ji *et al.*, 1994), and to the promoter of the *rrnO* operon of *Bacillus subtilus* (Ji *et al.*, 1994).

The sequence of the 16S rRNA gene is widely used for phylogenetic analysis because of its highly conserved nature (Woese, 1987). Sequence analysis of the 16S rRNA region by Boddinghaus *et al.* (1990) established that *M. bovis*, *M. tuberculosis*, *M. microti* and *M. africanum* all had identical 16S rRNA sequences. Frothingham *et al.* (1994) sequenced the 16S to 23S rRNA internal transcribed spacer (ITS) in 13 *M. tuberculosis* complex strains that included *M. tuberculosis* strains with highly variable fingerprints, *M. bovis*, *M. bovis* BCG, *M. microti* and *M. africanum*. The ITS has a high rate of nucleotide substitution, yet in this case the same ITS sequence was found in all 13 strains tested. Under normal phylogenetic criteria, these results would indicate that all four species of the *M. tuberculosis* complex should be considered a single species (Frothingham *et al.*, 1994).

A detailed phylogenetic study of fast and slow growing mycobacteria based on comparative sequence of the 16S rRNAs has been performed by Stahl and Urbance (1990). Grouping of related species on the basis of the 16S rRNA sequences showed a natural division between the fast and slow growers, and clustered the pathogenic mycobacteria (*M. bovis*, *M. kansasii*, MAIS complex and *M.*

paratuberculosis) together. Surprisingly, this study showed that *M. chitae* was distantly related to other mycobacteria, suggesting that it is misclassified.

1.3.2 Cloning and expression of mycobacterial genes.

This section does not attempt to describe the cloning of every mycobacterial gene. Rather, an overview of the strategies and techniques developed for cloning and expressing mycobacterial DNA in *E. coli*, *Streptomyces* and mycobacteria shall be presented.

1.3.2.1 Cloning and expression in *Escherichia coli*

Cloning of mycobacterial DNA was first described by Clark-Curtiss *et al.* (1985) who utilised standard *E. coli* plasmid and cosmid vectors to produce genomic DNA libraries of *M. leprae*, *M. vaccae* and *M. lufu*. However, attempts to complement various *E. coli* auxotrophic mutants with the mycobacterial genomic libraries were unsuccessful. Complementation of an *E. coli* citrate synthase mutant was possible if *M. leprae* DNA was expressed from the strong *asd* gene promoter from *Streptococcus mutans* (Clark-Curtiss *et al.*, 1985; Jacobs *et al.*, 1986). These results suggested that mycobacterial translational signals, but not transcriptional signals, could be recognised by *E. coli*. Expression of a mycobacterial gene from its own promoter in *E. coli* was first reported by Thole and colleagues (1985). By screening lysates from an *M. bovis* BCG lEMBL3 genomic library with anti-*M. bovis* antiserum, they identified a 64kDa antigen that was expressed in *E. coli* from the mycobacterial promoter. Mehra *et al.* (1986) and Shinnick (1987) also demonstrated that the 65 kDa antigens of *M. leprae* and *M. tuberculosis* respectively were expressed from their own promoters in *E. coli*. It therefore appears that only subsets of mycobacterial promoters can be recognised by the *E. coli* RNA polymerase.

An alternative approach to the expression of mycobacterial proteins in *E. coli* that didn't rely on expression from mycobacterial signals was described by Young and co-workers (1985a,b). The *E. coli* vector lgt11 was used to construct expression libraries of *M. tuberculosis* and *M. leprae* that were then screened using monoclonal antibodies. Mycobacterial proteins are expressed as fusion proteins with *E. coli* b-galactosidase, being transcribed and translated from *E. coli*

sequences. This approach was successful in isolating the genes for the 65K, 36K, 28K, 18K and 12K antigens of *M. leprae* (Young *et al.*, 1985b). Screening of the *M. tuberculosis* lgt11 library with four monoclonals identified clones that reacted with three of the monoclonals (Young *et al.*, 1985a), while a further round of screening using rabbit polyclonal serum identified the 65K antigen as an immunodominant antigen shared by *M. tuberculosis* and *M. leprae* (Husson and Young, 1987).

The *M. tuberculosis* lgt11 library was also screened for the ability to complement various *E. coli* auxotrophs as lysogens. This procedure identified the amino acid biosynthesis genes *aroA*, *aroB* and *aroD* (Garbe *et al.*, 1990), and the *lysA* gene (Anderson and Hansen, 1993) of *M. tuberculosis*. Inactivation of the genes for aromatic amino acid biosynthesis in *M. tuberculosis* may attenuate the organism, as is the case in *Salmonella typhimurium* (Hoiseth and Stocker, 1981), and hence may be a route to a new mycobacterial vaccine.

The use of *E. coli* as a surrogate host for the cloning and expression of mycobacterial genes does have some drawbacks. As already noted, many mycobacterial promoters studied so far are non-functional in *E. coli*. Also, *E. coli* may not be the best choice for the expression of secreted mycobacterial proteins. Mycobacterial secreted proteins have to cross the cytoplasmic membrane and the extracellular wall matrix to reach the external environment; in *E. coli*, the outer membrane has to be traversed to reach the exterior. Will mycobacterial proteins become "lost" in the *E. coli* periplasm? As well, *E. coli* possesses many dedicated transport systems for periplasmic, outer membrane and extracellular proteins (Pugsley, 1993). Whether these systems are capable of exporting mycobacterial proteins remains to be seen. Therefore, the expression of mycobacterial genes in hosts other than *E. coli* is fundamental to a proper understanding of mycobacterial genetics.

1.3.2.2 Cloning and expression in *Streptomyces*.

To overcome some of the drawbacks associated with *E. coli* as a surrogate host for mycobacterial genes, and since an efficient method for transformation and stable maintenance of foreign DNA in mycobacteria was still unavailable, *Streptomyces lividans* was proposed as a host for mycobacterial DNA (Keiser *et al.*, 1986). The advantages

of using *S. lividans* included its ability to recognise promoter sequences from a wide range of organisms and the requirement for little complementarity between the 16S rRNA sequence and the ribosome-binding site of the mRNA (Hopwood *et al.*, 1988). In addition, mycobacteria and streptomycetes are related, being members of the high G+C subdivision of the actinomycetes (Hopwood *et al.*, 1988). Transcription and translation of mycobacterial genes from their own expression signals therefore seemed likely.

Keiser and colleagues (1986) investigated the ability of *S. lividans* to use *M. bovis* BCG expression signals. Random fragments (average 5kb) of *M. bovis* BCG DNA were cloned into a *S. lividans* promoter probe vector that contained a promoterless kanamycin resistance gene as the reporter of transcriptional activity. BCG DNA was also cloned into a translational fusion vector to see whether mycobacterial translational signals function in *S. lividans*. These experiments showed that BCG promoters and translational signals function well in *Streptomyces*. However, generation of efficient systems for transformation and expression of DNA in mycobacteria superseded the need for *Streptomyces* as an heterologous host for mycobacterial DNA.

1.3.2.3 Cloning and expression in mycobacteria

The obvious choice as a host for recombinant mycobacterial DNA is a genetically well defined, fast growing mycobacterium. Such a host would be relatively easy to work with, efficiently recognise mycobacterial promoter and translation signals, allow post-translational modification of mycobacterial proteins, and permit studies on genetic regulation of mycobacterial genes to be performed, studies that would be difficult in heterologous hosts. Methods for introducing DNA, construction of cloning vectors, and producing defined mycobacterial mutants are a prerequisite to the use of mycobacteria in molecular biology.

The development of cloning vectors and transformation protocols was the first step towards genetic systems for the mycobacteria. Early work on the introduction of DNA into mycobacteria (via conjugation, transformation, transfection and transduction) is reviewed by Grange (1982). These studies demonstrated that transfer of genetic information to mycobacteria was possible, but

they relied on the transfer of chromosomally encoded markers that are of limited use for genetic studies. Transfection of *M. smegmatis* protoplasts with a shuttle phasmid vector by Jacobs *et al.* (1987) (see Section 1.3.1.4) was the first report of the introduction of foreign DNA into mycobacteria. However, the mycobacteriophage TM4 used to construct the phasmid vector is lytic for both *M. smegmatis* and *M. bovis* BCG, so its usefulness as a cloning vector is limited. Construction of shuttle vectors using the origin of replication from the *M. fortuitum* plasmid pAL5000 (Section 1.3.1.3), and the temperate mycobacteriophage L1 by Snapper *et al.* (1988) allowed the stable maintenance of foreign DNA in mycobacteria. The isolation of efficient plasmid transformation mutants of *M. smegmatis* by Snapper and colleagues (1990) and the development of electroporation protocols to transform mycobacteria (Snapper *et al.*, 1990; Hermans *et al.*, 1990(b); Jacobs *et al.*, 1991) have greatly facilitated genetic analysis of mycobacteria. Transformation of mycobacteria has also been reported using plasmids without an obvious mycobacterial origin of replication, such as the *E. coli* vector pIJ666 (Zainuddin *et al.*, 1989), the cosmid vector pJRD215 (Hermans *et al.*, 1991), and plasmids using a corynebacterial origin of replication (Radford and Hodgson, 1991).

The ability to express the gene products of mycobacteria, and of other pathogens, in surrogate mycobacterial hosts is important for a number of reasons. Expression of proteins from pathogenic mycobacteria in fast growing, well defined, non-pathogenic mycobacteria makes their study much easier, both from a production point of view and as a means to study the genetic systems controlling their regulation. Secondly, expression of protective antigens from pathogenic organisms in BCG opens the possibility of producing a multivalent recombinant BCG vaccine (see Section 1.3.3). Thirdly, post-translational modification of mycobacterial gene products may not be achieved in heterologous hosts. A detailed understanding of expression and post-translational modification of proteins in mycobacteria is therefore central.

Among the first foreign genes to be expressed in mycobacteria were the *aph* and *neo* kanamycin resistance genes from Tn*903* and Tn*5*, both from their own promoters (Snapper *et al.*, 1988). Winter and co-workers (1991) established that the promoter from the *groES/groEL1* operon of *Streptomyces albus* was also capable of

driving the expression of foreign antigens in mycobacteria. However, expression vectors based on mycobacterial transcription and translation signals would be preferable, since these signals may give better expression in mycobacteria, or allow protein expression to be controlled. Mycobacterial expression vectors using the promoter and ribosome binding site of the *M. bovis* BCG *hsp60* (Stover *et al.*, 1991) and *hsp70* (Aldovini and Young, 1991) genes were capable of expressing a range of foreign antigens in BCG. Promoters for heat-shock genes were chosen since these were among the first mycobacterial genes isolated and sequenced, they were known to be recognised in *E. coli*, and they were strong promoters (Burlein *et al.*, 1994). Matsuo and colleagues (1990) demonstrated that expression signals for the *M. kansasii* a-antigen could be used to generate a mycobacterial secretion system for foreign antigens, while Stover *et al.* (1993) used the *M. tuberculosis* 19kD antigen promoter to allow the production of foreign antigens as surface associated lipoproteins in BCG (see section 1.3.3).

Post-translational modification of proteins in mycobacteria has also been established. The antigen 85 complex are secreted mycobacterial proteins possessing signal peptides at the N-terminus that are cleaved before release of proteins to the exterior (Matsou *et al.*, 1990; Content *et al.*; 1991; Wiker and Harboe, 1992). Acetylation of the 19, 26, 27 and 38kD antigens of *M. tuberculosis* has been described by Young and Garbe (1991a). Garbe and colleagues (1993) have also found evidence of post-translational glycosylation of the *M. tuberculosis* 19kD antigen. Worth mentioning here is that efficient expression of the 19kD antigen in *E. coli* could only be achieved when it was produced as a fusion protein, and that glycosylation was only apparent when the *M. tuberculosis* antigen was expressed in *M. smegmatis* and not *E. coli*. This undoubtedly shows the importance of developing mycobacterial expression systems.

The most surprising example of post-translational modification in mycobacteria is evidenced by the *recA* gene products of the *M. tuberculosis* complex and *M. leprae* (Davis *et al.*, 1992; Davis *et al.*, 1994). The *recA* gene product is produced as a precursor protein that undergoes a protein splicing event, excising an internal spacer protein (or "intein") and ligating the N- and C- terminal domains (or "exteins") to produce the mature protein (Colston and Davis, 1994). Protein introns have been identified in a number of unrelated genes in a

phylogenetically diverse group of organisms (Colston and Davis, 1994). The *M. tuberculosis* complex RecA intein (found in *M. microti*, *M. bovis* and *M. tuberculosis*) shows no homology to the *M. leprae* RecA intein, either at the level of sequence or the position of insertion in the *recA* gene (Davis *et al.*, 1994). It therefore appears that the *M. tuberculosis* and *M. leprae* inteins have been acquired independently of each other. Since in mycobacteria the inteins are found only in the major pathogens, Davis and colleagues have suggested that the intein or splicing mechanism may confer a selective advantage to the host (Davis *et al.*, 1994; Colston and Davis, 1994), although what this advantage may be is unknown. An interesting observation is that when expressed in *E. coli*, the *M. tuberculosis recA* gene product is capable of splicing to an active RecA protein, while the *M. leprae* precursor protein shows no evidence of splicing (Davis *et al.*, 1994). Whether this is evidence that splicing in *M. leprae* is environmentally controlled or relies on cellular contents not present in *E. coli* is unknown, but it again shows the importance of using mycobacterial hosts to study mycobacterial genetics.

1.3.3 Recombinant BCG vaccines (rBCG)

The availability of cloning vectors and efficient gene-transfer systems for mycobacteria has opened the possibility of expressing foreign protective antigens in BCG to produce a recombinant BCG multivaccine (Jacobs *et al.*, 1987; Stover *et al.*, 1991; Aldovini and Young, 1991). The advantages of using BCG for such a multivaccine are many. BCG is a highly effective adjuvant, can be given at birth and produces a long-lasting immune response. It is also the most widely used vaccine in the world, and has an excellent record of safe use in man (Jacobs *et al.*, 1987). Expression of bacterial and viral antigens, such as b-galactosidase (Stover *et al.*, 1991; Murray *et al.*, 1992), tetanus toxin (Stover *et al.*, 1991) and a number of HIV antigens (Stover *et al.*, 1991; Winter *et al.*, 1991; Aldovini and Young, 1991) in BCG can elicit antibody and cell-mediated immune responses in model systems. More importantly, protective immune responses against rBCG expressing foreign antigens have been reported. Stover *et al.* (1993) demonstrated that a protective immune response against *Borrelia burgdorferi* (the causative agent of Lyme disease) was elicited by rBCG expressing the OspA surface antigen of *B. burgdorferi* as a lipoprotein, rather than as

a secreted or cytoplasmic protein. Immunisation of mice with rBCG expressing the *Streptococcus pneumoniae* PspA as a lipoprotein also elicited a protective response against *S. pneumoniae* (Burlein *et al.*, 1994). rBCG has also proved successful vaccinating against a protozoan pathogen. Connell and co-workers (1993) constructed rBCG expressing the gp63 proteinase of *Leishmania major* as a cytoplasmic protein. Inoculation of BALB/c and CBA/J mice with this rBCG produced a strong protective response to challenge with promastigotes or amastigotes of *Leishmania mexicana.*

Although rBCG has many advantages as a multivaccine vehicle, certain limitations are apparent. The requirement of some protective antigens to undergo post-translational modifications such as glycosylation or conformational changes may be unachievable in BCG. Also, as Colston has pointed out (1991), people who have received a conventional BCG vaccination may eliminate a second vaccination with rBCG too quickly to permit adequate priming of the immune response to the recombinant antigen. Whether these questions will be a major stumbling block to the development of rBCG as a multivalent vaccine remains to be seen.

1.4 Project aims

The thesis sought to investigate two phenomena: the role of catalase in isoniazid-sensitivity (Part II), and the response of mycobacteria to acid stress (Part III). The aims of each project and the experimental rationale are fully described in Sections 3.3 and 7.6 respectively.

Chapter 2.

Methods

2.1. Chemicals

Unless otherwise stated, chemicals were obtained from Sigma Biochemicals, Fisons or BDH Limited.

2.2 Bacterial strains and plasmids

The bacterial strains and plasmids used in this study are listed in Tables 2.1 and 2.2. *Escherichia coli* strains DH5a and TG2 were grown on Luria agar (L-agar) medium or in Luria broth (L-broth) at 37°C, TG2 being grown in the presence of tetracycline at 15mg/ml. Mycobacteria were normally grown on 7H11 agar medium (Difco) supplemented with Oleic acid-Albumin-Dextrose-Catalase (OADC), or in 7H9 broth (Difco) supplemented with Albumin-Dextrose-Catalase (ADC) and 0.05% (v/v) Tween 80. Mycobacteria harbouring *E. coli*-mycobacteria shuttle vectors were grown in the presence of kanamycin at 20mg/ml. For short-term storage bacteria were maintained on agar plates at 4°C. For long-term storage bacteria were stored at -70°C in media containing 10% (v/v) glycerol as a cryoprotectant.

2.2.1 Bacterial growth media

Media were prepared using distilled water and sterilised by autoclaving at 121°C at 15 psi (pounds per square inch) for 15 minutes.

Luria broth (1L)

10g Peptone
5g Yeast extract
10g NaCl
For Luria-agar, Bacto-agar (Difco) was added to a final concentration of 1.5% (w/v).

When using *E. coli* strains harbouring pUC18, or pUC19, or derivatives thereof, 100mg/ml ampicillin (stock solution 50mg/ml in water), 2ml/ml X-Gal (2% (v/v) solution in dimethylformamide) and 1ml/ml

IPTG (100mM in water) were added to cooled molten L-agar. When using *E. coli* strains harbouring pBR322 or derivatives, 100mg/ml ampicillin and 15mg/ml tetracycline (5mg/ml in 70% (v/v) ethanol) was added to cooled molten L-agar.

7H9 broth (200ml)

0.94g 7H9 base (Difco)
1ml 10% (v/v) Tween 80
Distilled water was added to 180ml, and after autoclaving 20ml of filter-sterilised Albumin-Dextrose-Catalase (ADC; Difco) was added.

7H11 agar (250ml)

5.25g 7H11 agar base (Difco)
1.25ml Glycerol
Distilled water was added to 224ml, and after autoclaving 25ml of filter-sterilised Oleic acid-Albumin-Dextrose-Catalase (OADC; Difco) was added.

SOC medium (1L)

20g Tryptone
5g Yeast Extract
0.5g NaCl
Distilled water was added to 1L, and after autoclaving 20ml of 1M Glucose was added. Before use, 5ml of 2M $MgCl_2$ was added.

2.3 Extraction of DNA

2.3.1 Extraction of *E. coli* plasmid DNA.

Plasmid DNA extraction was performed as described by Sambrook and colleagues (1989) using the following solutions:

Solution I

50mM Glucose
25mM Tris-HCl pH8.0
10mM EDTA pH8.0

Solution II

0.2M NaOH
1% (w/v) SDS

Solution III

5M Potassium acetate	60ml
Glacial Acetic acid	11.5ml
Distilled water	28.5ml

2.3.1.1 Large scale extraction of plasmid DNA

Large scale extraction of plasmid DNA used 500ml stationary phase bacterial cultures. Bacterial cells were harvested by centrifugation at 4,000g and the resulting cell pellet resuspended in 25ml of Solution I containing 10mg of lysozyme. After 10 minutes incubation at room temperature, 50ml of fresh solution II was added and gently mixed to produce a clear lysate. Then 37.5ml of cold Solution III was added and the preparation incubated on ice for 10 minutes. Cell debris was removed by centrifugation at 10,000g for 30 minutes at 4°C. The supernatant was filtered through a 5ml Gilson tip plugged with polymer wool into 0.6 volume of propan-2-ol, incubated at room temperature for 10 minutes, and the precipitated nucleic acid collected by centrifugation at 10,000g for 10 minutes. The pellet was then dissolved in 3ml of TE buffer (10mM Tris-HCl pH 8.0, 1mM EDTA pH 8.0), and 3ml of ice-cold 5M $LiCl_2$ added. The solution was centrifuged at 10,000g for 10 minutes to pellet RNA, and the supernatant added to an equal volume of propan-2-ol and mixed by inversion. Centrifugation at 10,000g for 10 minutes pelleted the DNA. The DNA pellet was dissolved in 500ml of TE, DNAse-free RNAse solution added to 20mg/ml, and then incubated at 65°C for 10 minutes. After the addition of 500ml of 1.6M NaCl containing 13% (w/v) PEG 8000 and incubation on ice for 10 minutes, the preparation was centrifuged for 10 minutes in a MicroCentaur bench-top microfuge at 13,000 rpm. The resulting DNA pellet was dissolved in 400ml of TE at 65°C and extracted with phenol:chloroform:isoamyl alcohol (24:24:1) until no white material was seen at the interface of the organic and aqueous phases after centrifugation at 13,000 rpm for 2 minutes in a MicroCentaur bench-top microfuge. After a chloroform:isoamyl alcohol (24:1) extraction, the DNA was precipitated from the aqueous

phase by adding 0.1 volumes of 3M sodium acetate (pH 5.2) and 2 volumes of absolute ethanol. After incubation at -20°C for 20 minutes, the DNA was recovered by centrifugation at 13,000 rpm for 10 minutes in a MicroCentaur bench-top microfuge, washed in 70% (v/v) ethanol and dissolved in 500ml of TE buffer.

2.3.1.2 Small scale extraction of plasmid DNA

Small scale extraction of plasmid DNA used 3.0ml of stationary phase bacterial cultures. Cells were collected by centrifugation at 6,500 rpm for 1 minute in a MicroCentaur bench-top microfuge and the resulting pellet resuspended in 200ml of Solution I containing lysozyme at 10mg/ml. After incubation at room temperature for 10 minutes, 400ml of fresh Solution II was added and mixed gently by inversion. Incubation on ice for 5 minutes was followed by the addition of 300ml of 7.5M ammonium acetate, gentle mixing, and a further incubation on ice for 5 minutes. Cell debris was removed by centrifugation at 13,000 rpm for 10 minutes in a MicroCentaur microfuge. The supernatant was added to 0.6 volumes of propan-2-ol and incubated at room temperature for 10 minutes. Plasmid DNA was collected by centrifugation at 13,000 rpm in a MicroCentaur microfuge, washed in 70% (v/v) ethanol and the resulting pellet resuspended in 50ml of TE buffer.

2.3.2 Extraction of mycobacterial plasmid DNA.

Plasmid DNA was prepared from *Mycobacterium smegmatis* using a modification of the method of Kado and Liu (1981). *Mycobacterium smegmatis* harbouring recombinant shuttle-plasmids was grown in 7H9 medium containing the appropriate antibiotic for 48 hours at 37°C. Then 5ml of this culture was inoculated into 5ml of fresh 7H9 medium containing the appropriate antibiotic and 2mg/ml D-cycloserine. The culture was then incubated for a further 18-24 hours at 37°C. Cells were collected by centrifugation at 4,000g for 10 minutes, washed in 1ml of PBS (Phosphate Buffered Saline, pH 7.2), and centrifuged in a MicroCentaur bench-top microfuge at 6,000 rpm for 3 minutes. The cell pellet was resuspended in 100ml of E buffer (40mM Tris-acetate pH 7.9, 2mM EDTA) and incubated at room temperature for 10 minutes. 200ml of fresh lysis solution (3% (w/v) sodium dodecyl sulphate, 0.6M NaOH) was added and the preparation incubated at 56°C

for 30 minutes. After two extractions with phenol:chloroform:isoamyl alcohol (24:24:1), and two chloroform:isoamyl alcohol (24:1) extractions, the DNA was precipitated from the final aqueous phase with 0.1 volumes of 3M sodium acetate (pH 5.2) and 2 volumes of absolute ethanol. Following incubation at -20°C for 30 minutes, the DNA was recovered by centrifugation at 13,000 rpm for 10 minutes in a MicroCentaur microfuge. The resulting pellet was washed in 70% (v/v) ethanol and resuspended in 25ml of nanopure water. As the resulting DNA could never be satisfactorily visualised by agarose gel electrophoresis, 5ml of the DNA preparation was used to transform *E. coli* DH5a. Preparation of plasmid DNA from *E. coli* transformants gave greater yield and quality of DNA.

2.3.3 Extraction of mycobacterial chromosomal DNA

Extraction of mycobacterial chromosomal DNA was based on the method of Davis and colleagues (1991). For fast-growing mycobacteria, 1L cultures were allowed to grow for 2 days at 37°C (shaking at 200rpm in the case of *M. smegmatis*), then glycine added to a final concentration of 1% (v/v). After a further 18-24 hour incubation at 37°C, cultures were harvested at 4000g for 10 minutes at 4°C. For slow-growing mycobacteria, cultures were allowed to grow for 2 to 3 weeks, then glycine added to 1% (v/v) and incubated for a further 24 hours before harvesting. Cell pellets were washed twice in wash solution (0.3M sucrose, 50mM Tris-HCl pH 8.0, 10mM EDTA), then resuspended in 1ml of wash solution. Lysozyme (2mg) and lipase VII-S (2mg) were then added and the preparation incubated at 37°C for 12 to 18 hours. Following the addition of 4ml of 6M guanidium chloride-1% (v/v) sarkosyl-20mM EDTA, the preparation was incubated at 37°C for 2 hours, and then extracted with an equal volume of chloroform. DNA was precipitated from the aqueous phase on addition of 0.1 volumes of 3M sodium acetate (pH 5.2) and 2 volumes of absolute ethanol. The DNA was collected by centrifugation at 10,000g for 10 minutes, and the resulting pellet was then washed in 70% (v/v) ethanol and dissolved in 600ml of TE. RNAse T_1 (4,000 Units), Proteinase K (0.5mg) and SDS to a final concentration of 0.5% (w/v) were then added, and the preparation incubated at 37°C for 2 hours. Two phenol:chloroform:isoamyl alcohol (24:24:1) extractions were then performed, followed by a chloroform:isoamyl alcohol (24:1) extraction.

Chromosomal DNA was then precipitated from the aqueous phase by the addition of 0.1 volume 3M sodium acetate and 2 volumes absolute ethanol. Following recovery of DNA by centrifugation at 13,000 rpm in a MicroCentaur microfuge, the DNA pellet was washed in 70% (v/v) ethanol and dissolved in 200ml of TE buffer.

2.4 Transformation of Bacterial Cells

2.4.1 Transformation of *E. coli* using electroporation.

2.4.1.1 Preparation of electro-competent *E. coli*

The preparation of electrocompetent cells was based on the method of Dower and colleagues (1988). LB-broth (1L) was inoculated with a fresh overnight culture and bacteria grown with vigorous aeration to mid-log phase (OD_{600} of 0.5 to 1 unit). The culture was cooled on ice for 30 minutes, then cells harvested by centrifugation in a cold rotor (4°C) at 4,000g for 15 minutes. The resulting cell pellet was washed twice in 1L of ice-cold sterile nanopure water and centrifuged at 6,500g for 20 minutes at 4°C. The cell pellet was resuspended in 25ml ice-cold 10% glycerol, and centrifuged at 5,000 rpm in a Heraceus-Christ centrifuge. The final pellet was resuspended in 1 to 2ml of ice-cold 10% (v/v) glycerol. Aliquots were stored at -70°C for up to 6 months.

2.4.1.2 Electro-transformation with plasmid DNA

Aliquots of electrocompetent cells were thawed on ice and used as soon as possible. Forty microlitres of cell suspension was mixed with 1-2ml of plasmid DNA and placed in a cold 0.2cm electroporation cuvette (Bio-Rad). A Bio-Rad Gene Pulser and Pulse Controller apparatus was used to generate the electric pulse. Settings were 25mF, 2.5kV and 200W, as detailed in Dower *et al.* (1988). The cold electroporation cuvette was placed in the safety chamber of the apparatus, brought into contact with the electrodes, and given a single pulse. Immediately, 1ml of SOC medium (room temperature) was added to the transformed cell suspension, gently mixed, and incubated at 37°C for 60 minutes. After outgrowth, the transformants were selected by plating onto selective agar.

2.4.2 Transformation of mycobacteria using electroporation.

2.4.2.1 Preparation of electro-competent *M. smegmatis* mc²155

Preparation of electro-competent *M. smegmatis* mc²155 was based on the method of Snapper and colleagues (1988). A stationary phase 10ml culture of *M. smegmatis* mc²155 was used to inoculate 100ml of fresh 7H9 medium. The culture was incubated at 37°C on a rotary shaker (200rpm) for 14 to 20 hours, until OD_{600} was approximately 1 unit. The culture was incubated on ice for 60 minutes, then centrifuged at 4,000g for 15 minutes. The cell pellet was washed twice in 100ml of ice-cold 10% (v/v) glycerol and centrifuged at 6,500g for 20 minutes. The cell pellet was then resuspended in 1.2ml of ice-cold 10% (v/v) glycerol and divided into three 400ml aliquots. The aliquots were kept on ice for immediate use.

2.4.2.2 Electro-transformation with plasmid DNA

Aliquots (400ml) of electro-competent *M. smegmatis* mc²155 were mixed with 1-5mg of plasmid DNA in a volume no larger than 5ml. The cell suspension was then placed in an ice-cold 0.2cm electroporation cuvette (Bio-Rad) and placed in the safety chamber of the Bio-Rad Gene Pulser and Pulse Controller apparatus. Settings were 25mF, 2.5kV and resistance at infinity. The transformation mix was then given a single pulse, and immediately added to 4.5ml of 7H9 medium. Transformants were allowed to recover at 37°C for 2 hours before plating onto selective media.

2.5 Routine DNA manipulation techniques

2.5.1 Agarose gel electrophoresis

DNA fragments were separated by agarose gel electrophoresis. Agarose (Seakem) concentration was 0.7% (w/v) for separation of DNA fragments greater than 3kb, or 1% (w/v) for separation of smaller fragments (<3kb). Agarose was solubilised in TAE buffer pH 7.7 (40mM Tris-acetate, 1mM EDTA) containing 0.5mg/ml ethidium bromide, and electrophoresis performed in TAE

buffer containing 0.5mg/ml ethidium bromide. DNA samples were mixed with loading buffer (0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol FF, 15% (w/v) Ficoll Type 400 (Pharmacia) in water) prior to loading into gel wells. Following electrophoresis at constant voltage (<20 volts per cm between electrodes), DNA was visualised using a long wave UV transilluminator. The DNA size markers used were 1kb ladder (GIBCO/BRL), or bacteriophage lambda DNA cut with HindIII (GIBCO/BRL).

2.5.2 DNA restriction digests

Restriction endonuclease cleavage of DNA was performed according to the manufacturer's recommendations. Typically, reactions were performed at 37°C in 20ml volumes using 0.5 to 1 units of enzyme and 500ng of DNA. Restriction endonucleases were obtained from GIBCO/BRL and Promega.

2.5.3 DNA dephosphorylation

Plasmid DNA to be dephosphorylated was first digested to completion using the appropriate restriction endonuclease. DNA (no more than 10pmoles of 5' termini) was then treated with 0.1 units of calf intestinal alkaline phosphatase (CIP) from Promega, using the reaction buffer supplied by the manufacturer, in a 50ml volume.

To dephosphorylate protruding 5' termini, DNA was incubated at 37°C for 30 minutes, then another 0.1 units of CIP was added and incubation continued for 30 minutes at 37°C.

To dephosphorylate blunt ends or recessed 5' termini an incubation for 15 minutes at 37°C and 15 minutes at 56°C was required. Another 0.1 units of CIP was added with two more incubations at the same temperatures. The reaction was stopped by phenol:chloroform extraction, followed by ethanol precipitation.

2.5.4 DNA ligation

T4 ligase (GIBCO/BRL) was used in ligase buffer supplied by the manufacturer. Ratios of vector to insert were determined empirically. Ligations were performed at 14°C for 12 to 18 hours, followed by ethanol precipitation, after which they were ready for transformation.

2.5.5 Phenol:chloroform extraction and ethanol precipitation of DNA

An equal volume of phenol:chloroform:isoamyl alcohol (24:24:1), equilibrated to pH 7.5 with Tris-HCl, was added to the DNA and mixed to form an emulsion. Following centrifugation at 4,000g (or 13,000 rpm in a MicroCentaur bench top microfuge) the aqueous phase was separated and retained. To remove phenol from the aqueous phase, an equal volume of chloroform:isoamyl alcohol (24:1) was added, an emulsion formed by mixing, and the phases separated by centrifugation as before. To precipitate DNA from the aqueous phase, sodium acetate (pH 5.2) was added to 0.3M, followed by 2 volumes of absolute ethanol. The sample was placed at -20°C for 30 minutes, then the DNA recovered by centrifugation at 10,000g. The DNA pellet was then washed twice in 70% (v/v) ethanol to remove residual salt, and dissolved in sterile TE buffer.

2.5.6 DNA extraction from agarose gels

DNA extraction from agarose gels used the Sephaglas Bandprep kit, purchased from Pharmacia, with slight modifications from the manufacturers instructions. After standard agarose gel electrophoresis, the band of interest was excised form the gel and placed in a clean 1.5ml microfuge tube. The agarose slice was then dissolved at 65°C using Gel solubiliser solution (sodium iodide buffered with Tris-HCl), and Sephaglas BP added at 5ml/mg of DNA. The microfuge tube was then placed in a rotary wheel device for 5 minutes to ensure good mixing. The microfuge tube was then centrifuged at 13,000 rpm for 15 seconds in a MicroCentaur microfuge to collect the Sephaglas, and the pellet washed twice in Wash solution (buffered salt solution). The Sephaglas was allowed to air dry before eluting the DNA twice using 20ml of TE buffer (pH 8.0). Eluted DNA was centrifuged at 13,000 rpm for 10 minutes to remove any Sephaglas fines before storage at -20°C.

2.6 DNA Hybridisation

2.6.1 Transfer of DNA to nylon filters by Southern blotting

DNA was transferred from agarose gels to nylon filters by the capillary method as described by Southern (1975). DNA samples were separated by standard agarose gel electrophoresis as described in section 2.5.1, and the gel photographed alongside a linear ruler. The DNA was partially depurinated by soaking the gel in 4 to 5 volumes of 0.25M HCl with gentle rocking for 30 minutes. The gel was then rinsed in distilled water and placed in 4 to 5 volumes of denaturing solution (0.5M NaOH, 1.5M NaCl) for 30 minutes with gentle rocking. The gel was rinsed in distilled water, then placed in 4 to 5 volumes of neutralising solution (0.5M Tris-HCl pH 7.5, 1.5M NaCl, 1mM EDTA) for 15 minutes with gentle rocking. This step was repeated. The gel was rinsed in distilled water, then placed on a "platform" of 10 sheets of pre-wetted (20x SSC) Whatman 3mm paper (20x SSC is 3M NaCl, 0.3M tri-sodium citrate). A pre-wetted (20x SSC) sheet of nylon membrane (Hybond-N+, Amersham International PLC) was placed on the gel, and any air bubbles removed. Three sheets of pre-wetted Whatman 3mm paper were placed over the nylon filter, followed by six dry sheets of 3mm paper and a stack of paper towels. Finally a 500g weight was placed on top, and the capillary transfer allowed to proceed for 6 to 18 hours. After transfer, the blot was dismantled and the nylon filter soaked in 4M NaOH for 30 minutes to fix the DNA to the filter. The filter was then air dried, wrapped in Saran wrap (DuPont) and stored at 4°C until needed.

2.6.2 Generation of a radiolabelled DNA probe.

Probe DNA was radioactively labelled by random priming using the Pharmacia Ready-To-Go labelling kit. The reaction components are provided as a lyophilised pellet containing dATP, dTTP, dGTP, Klenow fragment of DNA polymerase I and random 9-base long oligonucleotides. The lyophilised pellet was dissolved in 20ml of sterile nanopure water. DNA (approximately 50ng) to be labelled was denatured by placing in boiling water for 5 minutes, then rapidly cooling on ice. Denatured DNA was then added to the reconstituted reaction mix, along with 10mCi of [a-^{32}P]dCTP and sterile nanopure water to 50ml (radiolabelled [a-^{32}P]dCTP was obtained from Amersham

International at specific activity of 220 TBq/mmol). The mixture was incubated at 37°C for 30 minutes. Labelled probe could be stored at -20°C for 2 days. Prior to use, labelled probe was denatured by boiling and cooled rapidly on ice.

2.6.3 Hybridisation of probe to target DNA on filter

2.6.3.1 Hybridisation under conditions of high stringency

For prehybridisation, Southern blot filters were incubated for 1 to 3 hours in 25ml of hybridisation solution at 65°C. Hybridisation solution is 5x SSC, 5x Denhardt's solution (50x Denhardt's is 1% (w/v) each of Ficoll Type 400, BSA, polyvinolpyrollidine), 0.5% (w/v) SDS and 100mg/ml denatured salmon sperm DNA. The radiolabelled probe was boiled for 5 minutes to denature the DNA, snap-cooled on ice for 2 minutes, then added to the hybridisation solution. Hybridisation was continued at 65°C for at least 12 hours.

After hybridisation, the filter was washed twice at room temperature in 2x SSC/0.1% (w/v) SDS for 10 minutes, then once in 1x SSC/0.1% (w/v) SDS at 65°C for 15 minutes. Filters were finally washed at 65°C in 0.1x SSC/0.1% (w/v) SDS for 10 minutes, air dried and wrapped in Saran wrap (DuPont) to prevent desiccation. Filters were exposed to Cronex (DuPont) film in cassettes containing intensifying screens for 8-24 hours. Films were developed in an Agfa-Geveart automatic film processor.

2.6.3.2 Hybridisation under conditions of low stringency

For prehybridisation, Southern blot filters were incubated at 53°C for 1 to 3 hours in hybridisation solution (see Section 2.6.3.1). The radioactive probe was denatured by boiling for 5 minutes, snap-cooled on ice for 2 minutes, then added to the hybridisation solution. Hybridisation was allowed to continue at 53°C for at least 12 hours. Following hybridisation, filters were washed twice in 5xSSC/0.1% (w/v) SDS at 53°C for 30 minutes. Filters were then exposed to X-ray film and film developed as described above (Section 2.6.3.1).

2.7 DNA Sequencing

2.7.1 Manual DNA sequencing

2.7.1.1 Sequencing reactions using double stranded DNA

Manual DNA sequencing was accomplished using the chain termination method of Sanger *et al.* (1977). DNA to be sequenced was cloned into pUC18 or pUC19, and plasmid DNA purified using the PEG precipitation method (Section 2.3.1.1). Sequencing reactions were performed using the Sequenase Version 2.0 kit supplied by United States Biochemical Corporation. Primers used for sequencing were either the forward (-47) or reverse sequencing primers (purchased from New England Biolabs), or oligonucleotide primers synthesised on an Applied Biosystems DNA Synthesiser.

Double stranded plasmid DNA had to be denatured prior to performing the sequencing reactions. To achieve this, 4 to 6 mg of DNA (in a total volume of 16ml) was denatured by adding 2ml of fresh 2M NaOH, 2mM EDTA solution and incubating at 37°C for 15 minutes. DNA was then precipitated by adding 2ml of 2M ammonium acetate (pH 4.6), 75ml of absolute ethanol, and incubating at -70°C for 1 to 2 hours. Following centrifugation at 13,000 rpm in a MicroCentaur microfuge and two washes in 70% (v/v) ethanol, the DNA pellet was dissolved in 7ml of sterile nanopure water. Sequencing reactions were then carried out as detailed in the manufacturer's instructions. DNA was radiolabelled by incorporating [a-^{35}S]dATP (370 MBq/ml; Amersham) in the extension reactions, and the radiolabelled fragments separated by gradient gel electrophoresis (Biggin *et al.* 1983).

2.7.1.2 Gradient gel electrophoresis of reaction products

A Sequi-Gen Nucleic Acid sequencing cell (Bio-Rad) was used for gel electrophoresis. Prior to casting, sequencing plates and spacers were thoroughly cleaned with ethanol, and the surface of the notched plate (IPC plate) siliconised. The sequencing cell was then assembled, with 0.4mm spacers separating the plates. Gradient gels (Biggin *et al.* 1983) were prepared from the following solutions:

5x TBE

54g Tris base
27.5g boric acid
20ml 0.5M EDTA pH 8.0
Distilled water to 1L.

0.5x TBE Gel Mix

48g urea
20ml 40% (v/v) Acrylamide Solution
(Accugel, National Diagnostics)
10 ml 5x TBE
Distilled water to 100ml

2.5x TBE Gel Mix

48g urea
20ml 40% (v/v) Acrylamide solution
(Accugel, National Diagnostics)
50ml 5xTBE
5g sucrose
5mg Bromophenol blue
Distilled water to 100ml

To cast the gel, two gel solutions were prepared. Gel solution I contained 40ml of 0.5x TBE gel mix, plus 80ml of 10% (w/v) ammonium persulphate (APS) and 20ml of TEMED (both APS and TEMED were purchased from Bio-Rad Laboratories). Gel solution II contained 15ml of 2.5x TBE gel mix, plus 40ml 10% (w/v) APS and 10ml TEMED. 11ml of gel solution I, followed by 11ml of gel solution II were drawn up into a 25ml pipette. Two air bubbles were allowed into the pipette to produce a gradient, and the gel solution carefully poured between the sequencing plates. The remaining space was filled with gel solution I, and the blunt end of a sharkstooth comb inserted to give the gel-top a straight edge. The gel was allowed to polymerise for 3 to 18 hours before use.

To run the gel, 0.5x TBE was placed in the upper chamber, and 1x TBE in the lower chamber. The gel was pre-run at a constant power of 47W until the gel temperature reached 50°C. Samples were denatured at 80°C for 2 minutes prior to loading onto the gel, and electrophoresis performed at a constant 47W for 3 to 6 hours. Following electrophoresis the gel plates were prised apart and the gel soaked in 4 to 5 volumes of fixing solution (10% (v/v) methanol, 10% (v/v) acetic acid) for 20 minutes. The gel was then briefly rinsed in

distilled water before being transferred to pre-wetted Whatman 3mm paper. The gel was dried onto the paper under vacuum at 80°C for 45 minutes, then exposed to Cronex (DuPont) X-ray film for 12 to 24 hours. Film was processed in an Agfa-Geveart automatic film processor.

2.7.2 Automated DNA sequencing

Automated DNA sequencing used the PRISM DyeDeoxy Cycle Sequencing kit (Applied Biosystems) in conjunction with an Applied Biosystems Model 373A DNA sequencing system. The PRISM DyeDeoxy Cycle sequencing kit contains dye-labelled dideoxynucleotides whose incorporation terminates the extending chain to produce a dye labelled reaction product. The reaction mix also contains the thermostable enzyme AmpliTaq DNA Polymerase (Applied Biosystems) which allows reactions to be performed at high temperatures. The reaction products are then analysed colourimetrically on the 373A sequencing workstation.

To perform the sequencing reactions, 500ng-1mg of double stranded plasmid was mixed with 6.4pmoles of primer and 19ml of Terminator Premix in a final volume of 40ml (Terminator Premix contains the 4 "dyedeoxy" nucleotides, as well as dATP, dCTP, dITP, dTTP and AmpliTaq DNA polymerase). Reactions were always done at least in duplicate. The reactions were then overlaid with 40ml of mineral oil, and placed in a Perkin Elmer Cetus Model 480 thermal cycler (preheated to 98°C). Thermal cycles were then performed as follows

Denaturation	Rapid thermal ramp to 98°C Hold 98°C for 30 seconds
Annealing	Rapid thermal ramp to 50°C Hold 50°C for 15 seconds
Extension	Rapid thermal ramp to 60°C Hold 60°C for 4 minutes

Repeat for 30 cycles

After 30 cycles, the reactions were rapidly brought to 4°C and held at this temperature until removed from the thermal cycler and then stored

at -20°C. When the reaction products were frozen, the mineral oil overlay could be easily pipetted off the frozen reaction.

To clean up the reactions it was necessary to perform a phenol:chloroform step. Briefly, 100ml of sterile nanopure water was added to the reaction mix, followed by 140ml of phenol:water:chloroform (68:18:14, purchased from Applied Biosystems). After vigorous mixing, the sample was centrifuged at 13,000rpm in a MicroCentaur benchtop microfuge to separate the phases. The aqueous phase was removed (110ml) and extracted with phenol:water:chloroform as before. The aqueous phase (80ml) was removed, then 10ml of 2M sodium acetate (pH 4.5) and 2.5 volumes of absolute ethanol added. After incubation at -70°C for 60 minutes, the DNA was collected by centrifugation at 13,000rpm in a MicroCentaur microfuge. The pellet was washed twice in 70% (v/v) ethanol, then allowed to air-dry. The reaction products were analysed on an ABI Model 373A DNA sequencer by Dr. Katherine Lilly at the Protein and Nucleic Acid Sequencing Laboratory, University of Leicester. Sequence data was analysed using the SeqEd program (Version 1.0.3) supplied by Applied Biosystems Inc, and the University of Wisconsin Genetics Computer Group (GCG Version 7.3-UNIX) suite of programs (Genetics Computer Group, 1991).

Table 2.1 Bacterial strains

Strain	Genotype and/or Description	Reference
Escherichia coli		
TG2	*supE hsd*D5 *thi* D(*lac-proAB*) D(*srl-recA*)306::Tn*10*(*tet*^{R}) F'[*traD*36 *proAB*$^{+}$ *lacI*^{q} *lac*ZDM15]	Sambrook *et al.* (1989)
DH5a	*supE*44 DlacU169 (*f*80 *lac*ZDM15) *hsdR*17 *recA*1 *endA*1 *gyrA*96 *thi*-1 *relA*1	Hanahan, (1983)
UM255	Catalase negative *E. coli* mutant	Mulvey *et al.* (1988)
Mycobacteria		
M. smegmatis mc^{2}155	Mutant *M. smegmatis* strain with high transformation phenotype	Snapper *et al.* (1990)
M. smegmatis	Type strain	NCTC 8159
M. terrae	Type strain	NCTC 10856
M. gordonae	Type strain	ATCC 14470

Table 2.2 Plasmids

Plasmid	Characteristics	Reference
pUC18, pUC19	High copy number *E. coli* vectors, ampicillin resistance marker, blue-white selection.	Yanisch-Peron *et al.* (1985)
pBR322	*E. coli* vector, ampicillin and tetracycline resistance markers.	Bolivar *et al.*, (1977)
pMV261	*E. coli*-mycobacterium shuttle vector, kanamycin resistance marker, *hsp60* promoter.	Stover *et al.*, (1991)
pPA3	pMV261 carrying *Vibrio harveyi luxAB* genes expressed from *hsp60* promoter.	Andrew and Roberts (1993)

Chapter 3.

Introduction

3.1 Bacterial antioxidant mechanisms

Active oxygen species can be produced in bacterial cells in a number of ways (reviewed by Cadenas, 1989; Farr and Kogama, 1991; Demple, 1991). For example, the use of oxygen as a terminal electron acceptor during aerobic metabolism entails a four-electron reduction to water. During this reduction oxygen radicals are produced, such as superoxide radicals (O_2^-), hydrogen peroxide (H_2O_2) and hydroxyl radicals ($OH^\bullet$). Other endogenous sources of these radicals include the autoxidation of flavins and ubiquinones, spontaneous dismutation of O_2^-, and radiolysis of water (Farr and Kogama, 1991). These toxic oxygen radicals are capable of damaging a range of cellular constituents such as DNA, RNA, protein and lipids (Demple, 1991). The detoxification of these radicals is therefore of paramount importance to the bacterial cell.

The bulk of protection in bacteria against the harmful effects of oxygen radicals is provided by four enzymes: superoxide dismutase, catalase, glutathione synthetase and glutathione reductase (Farr and Kogama, 1991). The functions of superoxide dismutase (SOD) and catalase shall be dealt with in greater detail below (Sections 3.1.1 and 3.1.2). Glutathione synthetase is responsible for the generation of glutathione (GSH), an antioxidant that is involved in the maintenance of a strong reducing environment in the cell (Loewen, 1979). Glutathione reacts with active oxygen species to form glutathione radicals ($GS^\bullet$) that dimerise to produce GSSG. Glutathione reductase then catalyses the reduction of GSSG using electrons from NADPH, reforming GSH (Cadenas, 1989). Cycling of glutathione between reduced and oxidised forms therefore serves to "mop up" active oxygen species.

3.1.1 Superoxide dismutase

Most bacteria capable of aerobic growth synthesise superoxide dismutase (SOD), an enzyme that catalyses the dismutation of O_2^- to H_2O_2. Two main types of SOD have been characterised in prokaryotes, manganese-containing SOD (MnSOD) and iron-containing SOD (FeSOD) (Farr and Kogoma, 1991). Despite the high degree of homology between FeSOD and MnSOD enzymes, the metal cofactor specificity of these enzymes is strict (Kirby *et al.*, 1980). However, exceptions to this rule have been found in the SOD enzymes of *Streptococcus mutans* and *Bacteroides thetaiotaomicron* (Martin *et al.*, 1986; Pennington and Gregory, 1986). These organisms produce an SOD protein that is active with either manganese or iron as a cofactor.

Escherichia coli produces both MnSOD and FeSOD enzymes, the products of the *sodA* and *sodB* genes respectively, and their expression has been studied in some detail (Fridovich, 1978). FeSOD is expressed under both anaerobic and aerobic conditions, while MnSOD is induced only in the presence of oxygen. This induction is controlled by the products of two positive regulatory genes, *soxR* and *soxS* (Wu and Weiss, 1991). These genes control the SoxRS regulon, which governs the bacterial response to environmental superoxide stress (Demple, 1991). It should be noted that this response is separate to the peroxide stress regulon (Farr and Kogama, 1991).

Escherichia coli mutants unable to produce either form of SOD are incapable of aerobic growth on minimal medium, while single mutants, $sodA^+sodB$ or $sodAsodB^+$, grow normally (Carlioz and Touati, 1986). This suggests that a single functional SOD may be capable of protecting the cell against superoxide. However, studies into the relative importance of the two SOD enzymes in pathogenicity have produced some conflicting data. Papp-Szabo *et al.* (1993) generated isogenic wild-type and *sodB* mutant strains of *E. coli* K-12. These strains were then tested for their survival in human neutrophils, with the results indicating that the *sodB* mutant did not show increased killing compared to the wild-type. Nevertheless, the work of Franzon *et al.* (1990) demonstrated that *Shigella flexneri* lacking FeSOD is extremely sensitive to oxygen stress, despite the presence of a functional *sodA* gene. Considering that *S. flexneri* and *E. coli* are highly related organisms, these results are puzzling. Papp-Szabo *et al.* (1993) suggested the discrepancy is down to differences in the killing assays

employed by the two groups, rather than a reflection on the relative importance of *sodB*.

3.1.2 Catalase enzymes

Heme-hydroperoxidases can be classed into two main groups, catalases and peroxidases. Catalase enzymes (E.C. 1.11.1.6) catalyse the conversion of hydrogen peroxide (H_2O_2) into water and oxygen. The catalase enzymes isolated from animals, plants and micro-organisms resemble each other very closely and as such have been labelled "typical catalases" by Nadler *et al.* (1986). Typical catalase enzymes consist of four subunits of equal size, each subunit containing a ferric heme IX prosthetic group. Common physiochemical characteristics include a molecular weight in the range of 225-270 kDa, wide pH range of activity (pH 5-10.5), inhibition of activity by 3-amino-1,2,4-triazole (aminotriazole), and high Michaelis constants (K_m) for H_2O_2. On the other hand, peroxidases (E.C. 1.11.1.7) catalyse the reduction of H_2O_2 to water in the presence of a reducing agent. Unlike catalases, peroxidases are a heterogeneous group of enzymes consisting mainly of monomers with a wide range of molecular weights (Hochman and Shemesh, 1987).

In addition to catalases and peroxidases a third class of heme containing hydroperoxidases can be defined, namely the "catalase-peroxidases" (Nadler *et al.*, 1986; Hochman and Shemesh, 1987). Catalase-peroxidases differ from typical catalases in a number of characteristics, such as (i) having both catalatic and peroxidatic activities, (ii) displaying a narrow pH range for catalatic activity, (iii) resistance to aminotriazole, and (iv) being relatively heat labile. So far, these enzymes have only been isolated in micro-organisms, such as *M. tuberculosis* (Wayne and Diaz, 1982), *E. coli* (Clairborne and Fridovich, 1979), and *Rhodopseudomonas capsulata* (Hochman and Shemesh, 1987).

Escherichia coli produces both a monofunctional catalase enzyme, HPII, and a catalase-peroxidase enzyme, HPI, the products of the *katE* and *katG* genes respectively (Triggs-Raine *et al.*, 1988; van Ossowski *et al.*, 1991). Although HPII is monofunctional, it shows significant differences from typical catalase enzymes, being a tetramer of 81 kDa subunits and two protoheme IX groups (Clairborne and Fridovich, 1979). Expression of the enzymes is independent of one

another; *katG* is induced by exposure to H_2O_2, while *katE* is induced during stationary phase growth (Loewen *et al.*, 1985). Induction of *katG* is controlled by the *oxyR* gene product, a global regulator that upregulates the synthesis of 8 proteins in response to oxidative stress (Farr and Kogama, 1991). These 8 proteins are a subset of 30 proteins that are induced in response to H_2O_2 stress in *E. coli* (van Bogelan *et al.*, 1987). The expression of the HPII is regulated by the *katF* gene product (Loewen *et al.*, 1985). KatF (also termed RpoS) is a sigma factor that regulates the expression of as many as 32 genes in *E. coli* in response to nutrient limitation (McCann *et al.*, 1991). The KatF sigma factor has also been implicated in virulence gene regulation and the development of acid resistance in *Salmonella typhimurium* (Fang *et al.*, 1992; Li *et al.*, 1994).

Hydrogen peroxide produced by phagocytic cells is believed to be an important antimicrobial mechanism (Andrew *et al.*, 1985; Figure 3.1). Therefore, removal of peroxide by catalase may allow some pathogens to avoid killing by phagocytes. Middlebrook and Cohen (1953) were the first to observe a correlation between catalase activity and virulence, with the finding that catalase-negative *M. tuberculosis* were attenuated for virulence in the guinea pig (the role of catalase as a mycobacterial virulence factor will be dealt with in greater detail in Section 3.2.3). Pathogenicity of *Staphylococcus aureus* also correlated with catalase content, as a mutant strain with low catalase activity was more sensitive than the wild-type to killing by neutrophils (Mandell, 1975). Johnson and colleagues (1993) demonstrated that catalase negative *Neisseria gonorrhoeae* are extremely sensitive to exogenous H_2O_2, and suggested that catalase may be an important virulence factor in this organism. Nevertheless, catalase does not appear to be important for pathogenicity in a number of pathogens. Catalase-negative mutants of *Listeria monocytogenes* and *Shigella flexneri* do not show a significant reduction in virulence (Gaillard *et al.*, 1986; Franzon *et al.*, 1990). Similarly, *S. typhimurium* mutants lacking HPII catalase survive normally in macrophages *in vitro* and are virulent *in vivo* (Fields *et al.*, 1986), while *S. typhimurium* lacking HPI catalase show no defect in survival on exposure to neutrophils (Papp-Szabo *et al.*, 1994). Catalase mutants of *Haemophilus influenzae* were shown to be 2.3-fold less virulent than the wild-type following intraperitoneal inoculation in the infant rat model (Bishai *et al.*, 1994). However, when administered

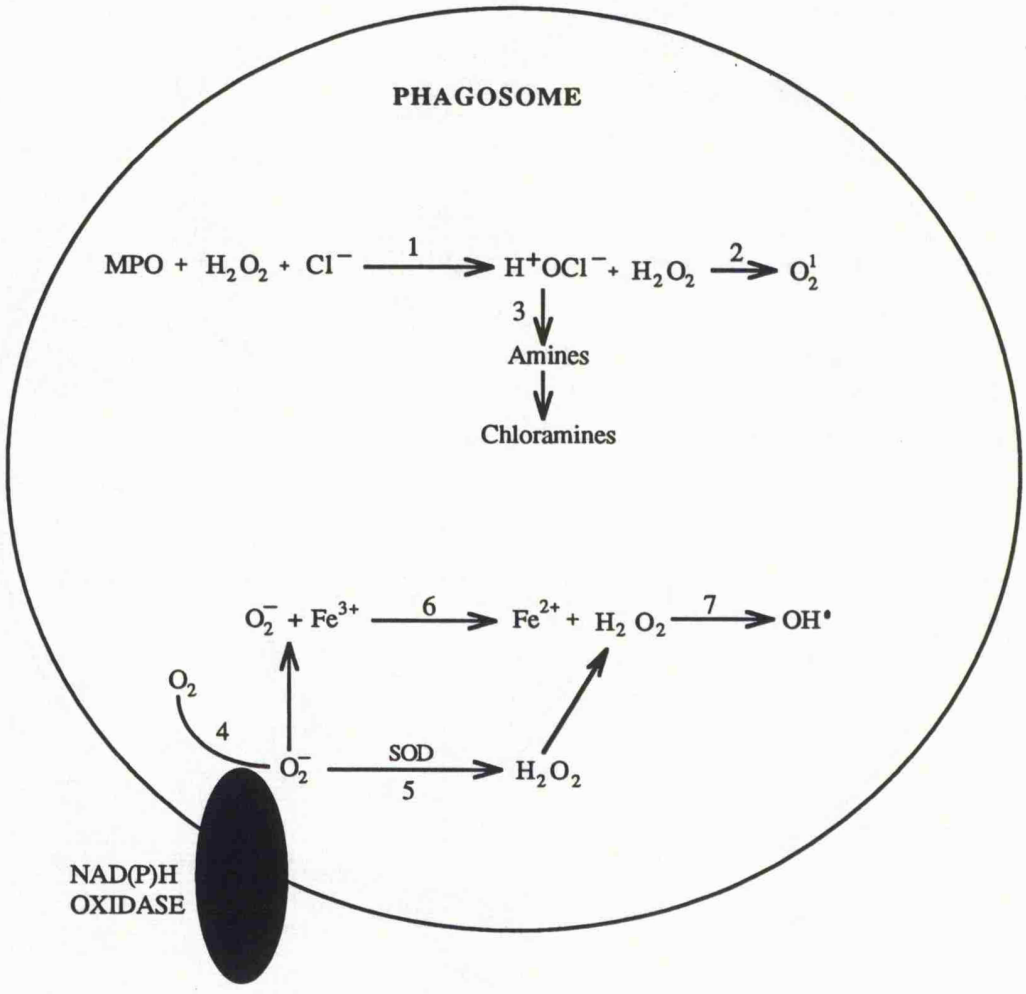


Figure 3.1 Oxygen-dependent antimicrobial systems of the phagosome. Some of the major systems for the production of toxic radicals are shown. Myeloperoxidase (MPO) catalyses the production of (reaction 1) hydrogen perchlorate (H^+OCl^-) and (2) singlet oxygen from hydrogen peroxide (H_2O_2) and chloride ion. (3) Perchlorate can also be used to generate chloramines, compounds that are toxic to microbes. The NAD(P)H oxidase complex reduces molecular oxygen, producing superoxide (4) that can spontaneously dismute to hydrogen peroxide, a reaction that is catalysed by superoxide dismutase (5). Superoxide can also react with ferric iron to generate hydrogen peroxide (6), with the peroxide reacting with ferrous iron to generate hydroxyl radical (7). The central role of hydrogen peroxide in many of these reactions should be noted.

intranasally the capacity of catalase negative mutants to cause lethal sepsis was indistinguishable from the wild-type.

Taken as a whole, these results suggest that the importance of catalase as a virulence factor varies between bacterial pathogens. This may be a reflection of the relative importance of catalase as a protective enzyme between organisms.

3.2 Mycobacterial catalases

3.2.1 Physiochemical properties and distribution

Three classes of catalase enzyme, T, M, and A, have been recognised in the mycobacteria. These enzymes are differentiated on the basis of physiochemical characteristics. Members of the T class are heat labile (inactivated when whole cells are exposed to 68°C for 20 minutes), are resistant to inactivation by aminotriazole, and have K_m values in the range of 3.1 to 6.8 mM H_2O_2 (Wayne, 1984). Members of the M class are heat stable (retain activity after exposure to 68°C for 20 minutes), sensitive to aminotriazole, and have K_m values greater than 140mM H_2O_2 (Wayne, 1984; Wayne and Diaz, 1982). Members of the A class are resistant to inactivation by heat, are sensitive to aminotriazole, and have a K_m of 70mM H_2O_2 (Wayne and Diaz, 1988).

The T class of catalase enzymes are bifunctional, in that they have both catalase and peroxidase activities, and are therefore more correctly classed as "catalase-peroxidase" enzymes (Nadler *et al.*, 1986). The M and A classes lack peroxidase activity and closely resemble the mammalian enzyme, placing them in the "typical catalase" grouping (Nadler *et al.*, 1986). Also, although the M- and A-catalase properties are similar, A-catalase is different in that it (i) is more heat stable, (ii) has a hydrophobic nature, and (iii) fails to cross-react with antibody to M-catalase (Wayne and Diaz, 1988).

Some mycobacteria produce only one class of catalase, while others produce two. For example, *M. tuberculosis*, *M. bovis*, *M. gastri* and *M. marinum* produce only the T-catalase class, while *M. kansasii*, *M. gordonae*, *M. scrofulaceum* and *M. simiae* produce both T and M classes. When both T and M classes are produced together, M-catalase accounts for at least two-thirds of total activity (Wayne and Diaz, 1982). *M. terrae*, *M. nonchromogenicum* and *M. triviale* produce only M-catalase, while *M. avium* and *M. intracellulare* produce both the

T and A class of catalase (Wayne and Diaz, 1982; Wayne and Diaz, 1988). So far the A-catalase class has been found only in the *M. avium-intracellulare* complex, always in conjunction with T-catalase (Wayne and Diaz, 1988). The A-catalase is produced at very low levels relative to T-catalase (Wayne and Diaz, 1982).

3.2.2 Role of catalase in isoniazid sensitivity

Isoniazid (isonicotinic acid hydrazide) was first introduced as an antituberculosis drug in 1952 (Bernstein *et al.*, 1952). It is now established as a central drug in short-course chemotherapeutic regimens, responsible for rapid killing of tubercle bacilli in tuberculous lesions during the first few days of treatment (Mitchison, 1992). The synthesis of the cell wall mycolic acids, essential components of the mycobacterial cell wall, has been proposed as the site of action of isoniazid. A number of lines of evidence point to this, including (i) loss of acid fastness accompanied by isoniazid action, (ii) a decrease in cell wall hydrophobicity of *M. bovis* BCG early in the action of isoniazid, (iii) rapid decrease in synthesis of cell wall material in *M. bovis* BCG exposed to isoniazid (reviewed by Winder, 1982). However, all mycobacteria contain mycolic acids as an integral cell wall component, yet not all mycobacteria are sensitive to isoniazid. Differences between the catalase enzymes of mycobacteria are one reason for the varying sensitivity of mycobacteria to isoniazid.

The role of catalase in isoniazid sensitivity has been widely studied ever since Middlebrook noted that acquisition of isoniazid resistance in *M. tuberculosis* was coincident with loss or reduction of catalase activity (Middlebrook, 1954). It is worth noting at this point that mycobacteria with only T-catalase activity, such as the *M. tuberculosis* complex, are particularly susceptible to isoniazid, and it is this catalase-peroxidase activity that is lost on acquisition of resistance to isoniazid (Winder, 1982). Mycobacteria that produce both T and M classes of catalase tend to be partially or fully resistant to isoniazid, with loss of T-catalase activity conferring complete isoniazid resistance (Wayne *et al.*, 1968). These observations strongly suggest a role for the catalase-peroxidase in isoniazid sensitivity.

A number of possible mechanisms for catalase-peroxidase mediated sensitivity have been suggested, but experimental evidence

favours two in particular (Figure 3.2). First, catalase may play a role in transport of isoniazid into mycobacteria. Susceptible strains of *M. bovis* BCG were shown to take up radiolabelled isoniazid, while resistant strains failed to associate the labelled isoniazid (Wimpenny, 1967). Wimpenny also established that the isoniazid binding mechanism had many of the properties of catalase-peroxidase; i.e., uptake was heat-labile, and sensitive to azide and cyanide. Once inside the cell, isoniazid may then be able to effect its bactericidal mechanism (Wimpenny, 1967).

Second, catalase may be involved in some chemical modification of isoniazid that converts it to a toxic derivative. Evidence for this modification comes from a number of sources. Youatt (1962) showed that *M. bovis* BCG forms yellow pigments when exposed to isoniazid and NAD. Further work demonstrated that cell-free extracts of isoniazid-sensitive, but not isoniazid-resistant, *M. bovis* BCG form pigment precursors in the presence of isoniazid (Youatt and Tham, 1969). This yellow pigment production (the so called Y enzyme reaction, after Youatt) was investigated by Gayatri Devi *et al.* (1975). They purified the peroxidase from *M. tuberculosis* to homogeneity, and showed that it was capable of carrying out the Y enzyme reaction in the presence of isoniazid, NAD and oxygen; Y enzyme activity was therefore attributed to peroxidase. The Y enzyme reaction also suggested two mechanisms for the antimicrobial activity of isoniazid; (i) the pigments are toxic derivatives of isoniazid, (ii) the pigment-forming reaction depletes the cellular pools of NAD. However, neither of these mechanisms appear to be responsible for the action of isoniazid (reviewed in Winder, 1982). For example, Jackett *et al.* (1977) showed that there was insufficient depletion of NAD at low concentrations of isoniazid to have an antibacterial effect, and also that NAD depletion did not correlate well with the antibacterial action of isoniazid. Moreover, pigment production does not appear to occur in *M. smegmatis* or *M. phlei* even though they both show susceptibility to isoniazid (Winder *et al.*, 1967; Youatt, 1969).

Alternative antibacterial mechanisms arising from the interaction of isoniazid and T-catalase have been suggested by other workers. The production of toxic free radicals may be one such mechanism. Shoeb *et al.* (1985a, 1985b) established that horseradish

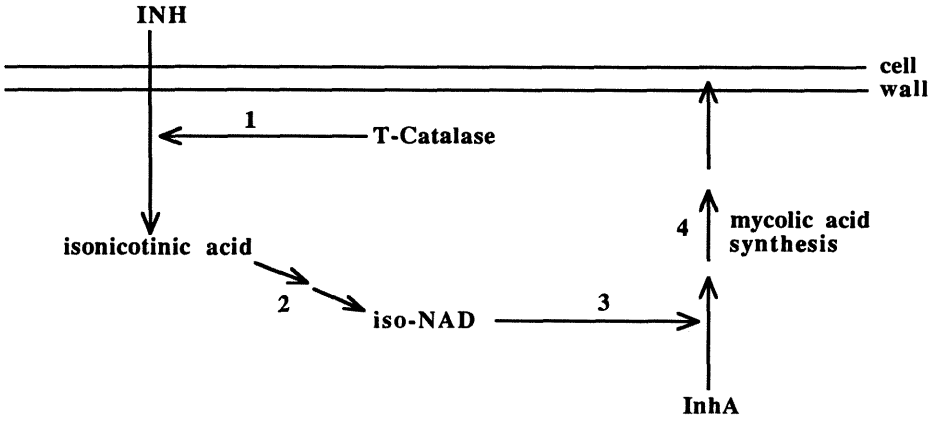


Figure 3.2 Model for the action of isoniazid.
T-catalase has been proposed to act in the uptake of isoniazid (INH) or in the activation of catalase (see text, Section 3.2.2). Both of these proposals can be incorporated in a model where we assume isoniazid is able to cross the mycobacterial cell wall; i.e. a diffusion gradient exists between extracellular and intracellular isoniazid, a gradient that is maintained by the intracellular modification of isoniazid to isonicotinic acid by T-catalase (reaction 1). Isonicotinic acid then poisons the cellular NAD pools by the production of iso-NAD (2). This removal of reducing power would have severe effects on the synthesis of mycolic acids by the InhA protein (3, 4), a protein with homology to enzymes involved in lipid synthesis in other bacteria (Banerjee *et al.*, 1994). This model is based on the theory put forward by Cole (1994).

peroxidase and extracts of isoniazid-sensitive, but not isoniazid-resistant, *M. tuberculosis* can oxidise isoniazid in the presence of H_2O_2 to produce hydroxyl and organic free radicals. Indeed, Winder (1960) was the first to suggest that isoniazid works through the formation of free radicals. A further possibility is that T-catalase converts isoniazid to isonicotinic acid, an analogue of nicotinic acid (Cole, 1994). This molecule is thought to exert an antimetabolic effect by poisoning the nicotinamide biosynthetic pools, leading to the production of iso-NAD (Cole, 1994). This removal of reducing power would have profound effects on many cellular processes, including the synthesis of lipids and hence mycolic acids. This model also accounts for the involvement of the recently described *inhA* gene product in isoniazid resistance (Banerjee *et al.*, 1994). This gene codes for a protein with homology to the *E. coli* enzyme EnvM that is involved in fatty acid biosynthesis, suggesting that InhA may be a mycobacterial analogue of this protein. Mutation of the *inhA* gene from serine to alanine at position 94 produces an isoniazid-resistant phenotype, perhaps by preventing the binding of iso-NAD and thus allowing mycolic acid biosynthesis to continue.

The "activation" of isoniazid by T-catalase suggests a possible mechanism for reduced susceptibility or resistance to isoniazid in mycobacteria with T- and M-catalase classes; the M-catalase may remove H_2O_2 from the internal milieu, preventing the T-catalase from oxidising, and hence "activating", isoniazid. An approach to investigate the validity of this hypothesis would be to inactivate the gene for M-catalase in a *Mycobacterium* species with both T and M classes and see whether the susceptibility to isoniazid increased. Reintroduction of a functional copy of the M-catalase gene should then restore wild-type levels of isoniazid resistance.

3.2.3 Catalase-peroxidase as a virulence determinant.

The antimicrobial mechanisms of phagocytes can be broadly divided into oxygen-dependent and oxygen-independent systems (reviewed by Andrew *et al.*, 1985). A detailed description of these systems is well beyond the scope of this review. Rather, it is sufficient to say that oxygen-dependent mechanisms consist of active oxygen species produced by the respiratory burst, while oxygen-independent systems involve components such as lysosomal enzymes, cationic peptides and tryptophan starvation (Andrew *et al.*, 1985). For the

purposes of this section, the role of catalase in protection against reduced forms of oxygen shall be examined.

Phagocytosis triggers the oxygen-dependent respiratory burst in which the ingested microbe is bombarded with a range of toxic oxygen species, such as superoxide anion, hydroxyl radical and hydrogen peroxide. It would therefore appear to be to the microbe's advantage to be able to detoxify these radicals. Catalase converts hydrogen peroxide to water and oxygen, so this enzyme could be a major defence against oxidative stress (see Section 3.1.2). Middlebrook and Cohen (1953) demonstrated that catalase-negative strains of *M. tuberculosis* have reduced virulence for the guinea pig. This suggested that catalase was indeed an important virulence factor for tubercle bacilli. However, the role of peroxide in killing of mycobacterial pathogens is not clear. Walker and Lowrie (1981) were the first to show that H_2O_2 produced by lymphokine-activated murine macrophages kills *M. microti*. Addition of exogenous catalase protected the bacilli, strongly suggesting that hydrogen peroxide was a critical antimycobacterial agent. Nevertheless, O'Brien and colleagues (1990, 1991) showed that killing of strains of *M. tuberculosis* by guinea-pig alveolar macrophages was independent of the strain's susceptibility to hydrogen peroxide, and that hydrogen peroxide was not involved in the bactericidal mechanism. It has been suggested that the observation of Walker and Lowrie (1981) that exogenously added catalase inhibited the antimycobacterial effect of activated macrophages may in fact be related to the ability of this peroxide scavenger to suppress the production of toxic nitrogen oxides (Chan and Kaufmann, 1994). Parenthetically, when dealing with the area of macrophage killing of *M. tuberculosis* it is important to note the strains and models (guinea-pig or mouse) employed in different studies, as these may well be responsible for many of the apparent conflicts in the literature.

The correlation between loss of catalase-peroxidase activity with avirulence is puzzling. If production of hydrogen peroxide is not a major antimycobacterial mechanism, what is the role of catalase? One possible explanation is that the catalase negative, avirulent mutants of *M. tuberculosis* used in virulence studies may have mutations in loci other than just the catalase gene. Since these strains are not genetically defined, this is a very real possibility. To address this issue, a series of isogenic mutants of *M. tuberculosis*, differing only in catalase activity,

needs to be constructed. This approach has already proved useful in determining the role of catalase and superoxide dismutase in the virulence of *Shigella flexneri* and *Salmonella typhimurium* (Franzon *et al.*, 1990; Papp-Szabo *et al.*, 1994).

A further role of catalase may be in the transport of nutrients into the bacterial cell during *in vivo* growth. Catalase is known to play a role in binding of the antimycobacterial drug isoniazid to the mycobacterial cell (Wimpenny, 1967). Assuming that isoniazid is not the only substrate for this reaction, mycobacteria may use catalase-peroxidase to scavenge nutrients.

Finally, work by Li *et al.* (1992) has raised the possibility that catalase may be involved in protection against toxic reactive nitrogen intermediates (RNI). They found that exogenously added catalase could inhibit the synthesis of nitric oxide and hence the killing of *Leishmania major* by murine macrophages. As nitric oxide is known to be involved in murine macrophage killing of *M. tuberculosis* (Chan *et al.*, 1992; reviewed by Chan and Kaufmann, 1994), it may be that mycobacterial catalase has a protective against killing by RNI. To date however, RNI production *in vitro* has only been demonstrated in murine macrophages; whether RNI have a role to play in antimycobacterial mechanisms of guinea pig and human macrophages remains to be seen.

3.3 Project aims and experimental strategy

The central goal of this project was to clone the genes for the mycobacterial T- and M-catalase. The cloned genes could then be used as starting points for investigations into the role of catalase in isoniazid sensitivity and virulence. For example, an exciting experiment would be to introduce a functional T-catalase gene into isoniazid-resistant, avirulent strains of *M. tuberculosis* and see whether this could restore isoniazid sensitivity and/or virulence.

Previous to this project, an approach to clone the mycobacterial catalase genes that relied on purification of the T- and M-catalase proteins had been pursued in the department. Once purified, the catalase proteins were to be used either to raise antibodies or to obtain N-terminal sequence data; anti-catalase antibodies or oligonucleotides based on the N-terminal sequence could then be used to screen a *M.*

tuberculosis λgt11 library obtained from Dr. R.A. Young, Whithead Institute, USA. However, purification of the proteins to the necessary level of purity proved problematic. As a number of competing groups were also working on the cloning of the mycobacterial catalase genes, a faster approach was needed. This project therefore sought to clone the genes using a molecular genetic approach.

The polymerase chain reaction (PCR) involves the synthesis of multiple copies of a region of DNA from oligonucleotide primers that flank the target sequence on opposite strands. In each cycle, template DNA is denatured, primers are annealed to target sequences, and the primers are extended across the template by a thermostable DNA polymerase (Saiki *et al.*, 1988). Repeating the cycle 20-30 times can result in over a million fold amplification of the target sequence, producing an 'amplicon'. PCR may also be used in situations where the nucleotide sequence of the template DNA is not known. For example, alignment of protein sequences may reveal conserved amino acid regions between related proteins. These conserved regions may then be back-translated to the corresponding coding nucleotide sequence. However, as more than one codon can code for an amino acid, a pool of "degenerate" oligonucleotides needs to be synthesised to cover all possibilities. This pool of primers may then be used to amplify regions of DNA which code for a related protein in other organisms.

The experimental strategy for cloning the catalase genes can be outlined as below:

(i) Identify regions of homology between the catalase enzymes (typical and catalase-peroxidase) of different organisms.
(ii) Use these conserved regions as the basis for PCR oligonucleotide primers.
(iii) Amplify regions of the catalase genes.
(iv) Use these amplicons as hybridisation probes to isolate the full gene from genomic libraries.

It was initially hoped to clone both the T- and M-catalase genes using the PCR strategy. However, it should be stated at this stage that the T-catalase gene from *M. tuberculosis* was cloned by a competing group (Zhang *et al.*, 1992) within the first year of the project. This necessitated a review of the project aims. It was decided that the cloning of the M-catalase gene using PCR should become the

priority as the role of M-catalase in isoniazid sensitivity was still a central question. The results section (Chapter 5) therefore deals largely with the search for the M-catalase gene.

Chapter 4.

Methods

4.1 Application of the Polymerase Chain Reaction (PCR) to amplification of mycobacterial catalase sequences.

To clone the mycobacterial catalase genes, an approach based on the polymerase chain reaction (Saiki *et al.*, 1988) was adopted.

4.1.1 Alignment of catalase protein and gene sequences.

Programs from the University of Wisconsin Genetics Computer Group software version 7-UNIX package (Genetics Computer Group, 1991) were used to search DNA and protein databases for catalase sequences. The CLUSTAL V sequence alignment program (Higgins *et al.*, 1992) was used to align selected catalase sequences (Figures 5.1, 5.2, 5.3).

4.1.2 Design of degenerate oligonucleotide PCR primers.

Conserved protein sequences among the catalase enzymes were back-translated to the corresponding degenerate nucleotide sequences. Mycobacterial codon usage tables (Dale and Patki, 1990) were used to decrease the degeneracy of the primers. All primers were synthesised using an Applied Biosystems model 381A DNA synthesiser at the core facility run by the Department of Biochemistry, University of Leicester. Prior to use, 200µl aliquots of the primer preparations were ethanol precipitated with 0.1 volume of 3M sodium acetate and 2 volumes of 100%(v/v) ethanol. The precipitated primer DNA was collected by centrifugation in a benchtop microfuge at 11,000g for 15 minutes, pellets washed in 70%(v/v) ethanol, then dissolved in sterile nanopure water. The OD_{260} of the primer solution was measured, and the DNA concentration determined by the formula:

$$OD_{260} \times \text{Dilution} \times 20 = [\text{DNA}]\ \mu g/ml$$

Primer solutions were adjusted with sterile nanopure water to give 10µM stock solutions, then stored frozen at -20°C.

4.1.3 PCR Reactions.

In general, PCR reactions were carried out in 100µl volumes, using 100ng of chromosomal template DNA, primers at a final concentration of 1µM and each deoxynucleotide at 200µM. Reaction buffer (x1) was 20mM Tris-HCl (pH 8.3), 25mM KCl, 0.05% (v/v) Tween 80, with $MgCl_2$ added to give a Mg^{2+} concentration of between 1 and 3mM. Reactions were overlaid with 100µl of mineral oil (molecular biology grade, Sigma; oil was exposed to UV light for 15 minutes to destroy any contaminating DNA) and then heated to 95°C for 5 minutes in a Perkin-Elmer model 480 thermal cycler. *Taq* polymerase (Northumbria Biologicals Ltd.) was added to 2.5 units per 100µl reaction, and the reactions briefly spun in a benchtop microfuge. Reaction tubes were then returned to the thermal cycler to undergo 30 cycles of amplification. Thermal cycle parameters were as follows, with annealing temperature decreasing during the cycling in a stepwise manner after the "touchdown" protocol of Don *et al.*, 1991.

Denaturation:	96°C for 60 seconds
Annealing:	42°C for 30 seconds (3 cycles at this temperature, then 3 cycles at 40°C, 3 cycles at 38°C, 6 cycles at 36°C, 6 cycles at 34°C, and 10 cycles at 32°C)
Extension:	72°C for 90 seconds

After a final extension at 72°C for 5 minutes, the block temperature was reduced to 4°C and reactions kept at this temperature until they were removed from the cycler for storage at -20°C. Reaction products were analysed by electrophoresis through a 1% (w/v) agarose gel.

4.1.4 Screening of the PCR products for catalase sequences

PCR products were separated by agarose gel electrophoresis (Section 2.5.1), stained with ethidium bromide and visualised under long wave UV light. Gels were photographed alongside a ruler, then processed for Southern blotting as described in Section 2.6.1. Filters were hybridised under conditions of low stringency (Section 2.6.3.2) against radiolabelled PCR products from the *Listeria seeligeri* and *Escherichia coli katE* genes (Section 4.2).

4.2 PCR amplification of the *E. coli* and *Listeria seeligeri* catalases.

Nucleotide primers based on the known sequence of the *E. coli* (von Ossowski *et al.*, 1991) and *L. seeligeri* (Haas *et al.*, 1991) catalase (*katE*) gene sequences were synthesised using an Applied Biosystems model 381A DNA synthesiser at the Department of Biochemistry, University of Leicester. Primers were purified before use as described above (Section 4.1.2). The *L. seeligeri* primers (Liscat3' and Liscat5') would amplify a 1.3kb fragment containing the entire *L. seeligeri* catalase sequence (Figure 4.1). The *E. coli* primers (ECA3' and ECA5') would amplify a 650bp internal fragment from the *E. coli* catalase gene (Figure 4.2); this internal fragment codes for the section of the *E. coli* catalase enzyme that shows the greatest degree of amino acid homology with other catalases. Sequences of the primers are detailed in Table 4.1.

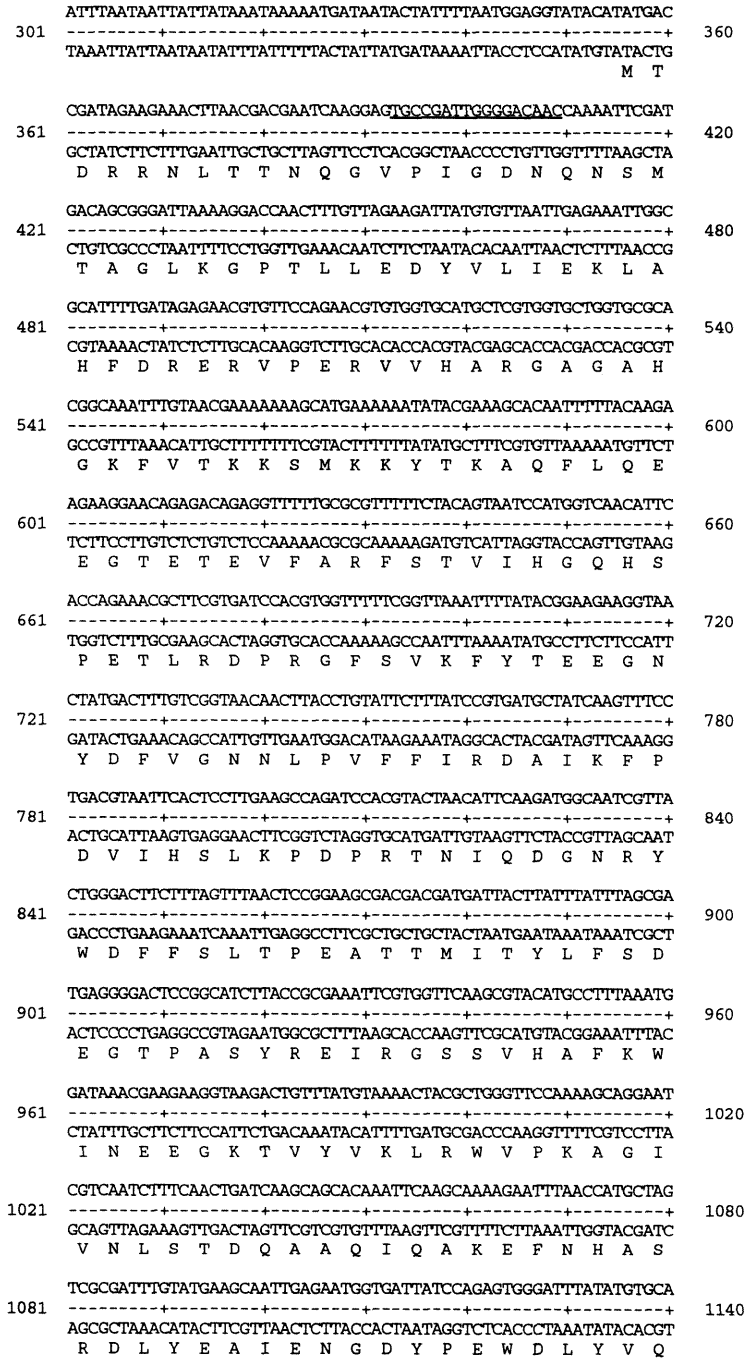
Table 4.1

Primer name	Sequence	Size
Liscat5'	TGCCGATTGGGGACAAC	17mer
Liscat3'	TTCTGCATCTGAGTAACGC	19mer
ECA3'	GTCCCTGCAAACAA	14mer
ECA5'	AAGTTCTATACCGAA	15mer

Reaction mixes and addition of *Taq* polymerase were as described in Section 4.1.3. Thermal cycle parameters (30 cycles) were as follows:

Denaturation: 94°C for 60 seconds
Annealing: 55°C for 2 minutes
Extension: 72°C for 90 seconds

A final extension at 72°C for 5 minutes was performed, and reactions cooled to 4°C. Reactions were stored at -20°C, and analysed by agarose gel electrophoresis.


ATTTAATAATTATTATAAATAAAAATGATAATACTATTTTAATGGAGGTATACATATGAC
301
360
TAAATTATTAATAATATTTATTTTTACTATTATGATAAAATTACCTCCATATGTATACTG
M T
CGATAGAAGAAACTTAACGACGAATCAAGGAGTGCCGATTGGGGACAACCAAAATTCGAT
361
420
GCTATCTTCTTTGAATTGCTGCTTAGTTCCTCACGGCTAACCCCTGTTGGTTTTAAGCTA
D R R N L T T N Q G V P I G D N Q N S M
GACAGCGGGATTAAAAGGACCAACTTTGTTAGAAGATTATGTGTTAATTGAGAAATTGGC
421
480
CTGTCGCCCTAATTTTCCTGGTTGAAACAATCTTCTAATACACAATTAACTCTTTAACCG
T A G L K G P T L L E D Y V L I E K L A
GCATTTTGATAGAGAACGTGTTCCAGAACGTGTGGTGCATGCTCGTGGTGCTGGTGCGCA
481
540
CGTAAAACTATCTCTTGCACAAGGTCTTGCACACCACGTACGAGCACCACGACCACGCGT
H F D R E R V P E R V V H A R G A G A H
CGGCAAATTTGTAACGAAAAAAAGCATGAAAAAATATACGAAAGCACAATTTTTACAAGA
541
600
GCCGTTTAAACATTGCTTTTTTTCGTACTTTTTTATATGCTTTCGTGTTAAAAATGTTCT
G K F V T K K S M K K Y T K A Q F L Q E
AGAAGGAACAGAGACAGAGGTTTTTGCGCGTTTTTCTACAGTAATCCATGGTCAACATTC
601
660
TCTTCCTTGTCTCTGTCTCCAAAAACGCGCAAAAAGATGTCATTAGGTACCAGTTGTAAG
E G T E T E V F A R F S T V I H G Q H S
ACCAGAAACGCTTCGTGATCCACGTGGTTTTTCGGTTAAATTTTATACGGAAGAAGGTAA
661
720
TGGTCTTTGCGAAGCACTAGGTGCACCAAAAAGCCAATTTAAAATATGCCTTCTTCCATT
P E T L R D P R G F S V K F Y T E E G N
CTATGACTTTGTCGGTAACAACTTACCTGTATTCTTTATCCGTGATGCTATCAAGTTTCC
721
780
GATACTGAAACAGCCATTGTTGAATGGACATAAGAAATAGGCACTACGATAGTTCAAAGG
Y D F V G N N L P V F F I R D A I K F P
TGACGTAATTCACTCCTTGAAGCCAGATCCACGTACTAACATTCAAGATGGCAATCGTTA
781
840
ACTGCATTAAGTGAGGAACTTCGGTCTAGGTGCATGATTGTAAGTTCTACCGTTAGCAAT
D V I H S L K P D P R T N I Q D G N R Y
CTGGGACTTCTTTAGTTTAACTCCGGAAGCGACGACGATGATTACTTATTTATTTAGCGA
841
900
GACCCTGAAGAAATCAAATTGAGGCCTTCGCTGCTGCTACTAATGAATAAATAAATCGCT
W D F F S L T P E A T T M I T Y L F S D
TGAGGGGACTCCGGCATCTTACCGCGAAATTCGTGGTTCAAGCGTACATGCCTTTAAATG
901
960
ACTCCCCTGAGGCCGTAGAATGGCGCTTTAAGCACCAAGTTCGCATGTACGGAAATTTAC
E G T P A S Y R E I R G S S V H A F K W
GATAAACGAAGAAGGTAAGACTGTTTATGTAAAACTACGCTGGGTTCCAAAAGCAGGAAT
961
1020
CTATTTGCTTCTTCCATTCTGACAAATACATTTTGATGCGACCCAAGGTTTTCGTCCTTA
I N E E G K T V Y V K L R W V P K A G I
CGTCAATCTTTCAACTGATCAAGCAGCACAAATTCAAGCAAAAGAATTTAACCATGCTAG
1021
1080
GCAGTTAGAAAGTTGACTAGTTCGTCGTGTTTAAGTTCGTTTTCTTAAATTGGTACGATC
V N L S T D Q A A Q I Q A K E F N H A S
TCGCGATTTGTATGAAGCAATTGAGAATGGTGATTATCCAGAGTGGGATTTATATGTGCA
1081
1140
AGCGCTAAACATACTTCGTTAACTCTTACCACTAATAGGTCTCACCCTAAATATACACGT
R D L Y E A I E N G D Y P E W D L Y V Q

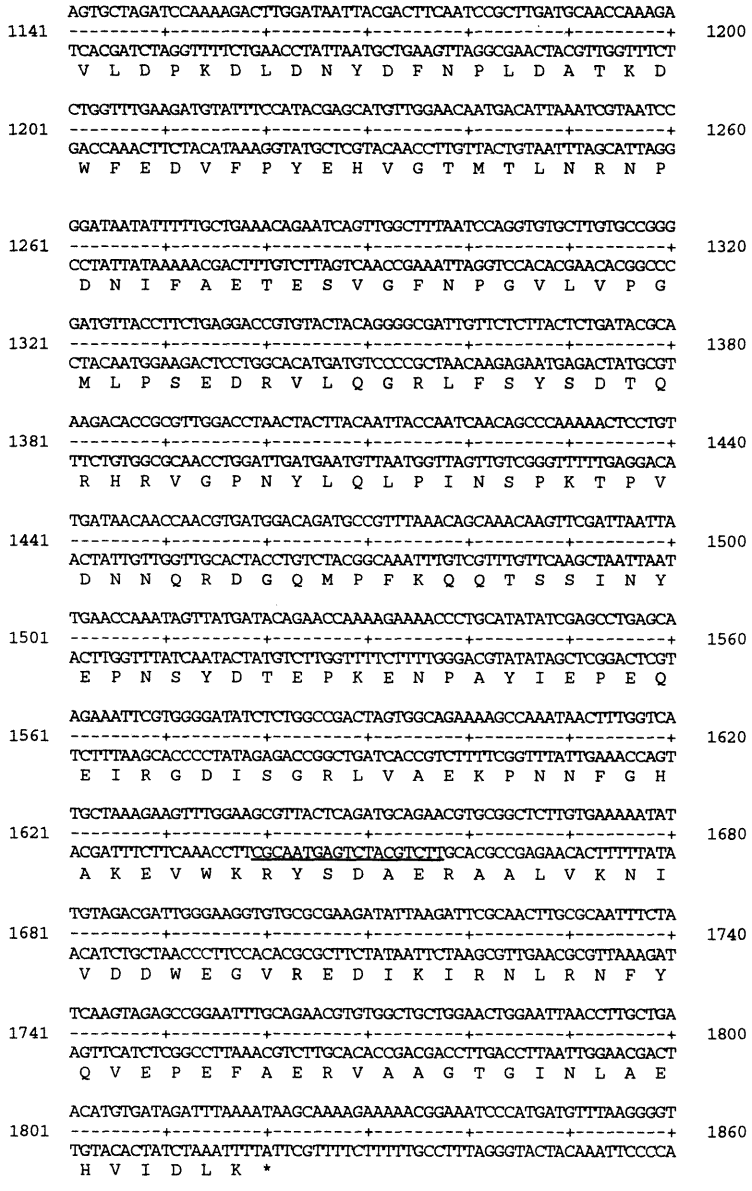
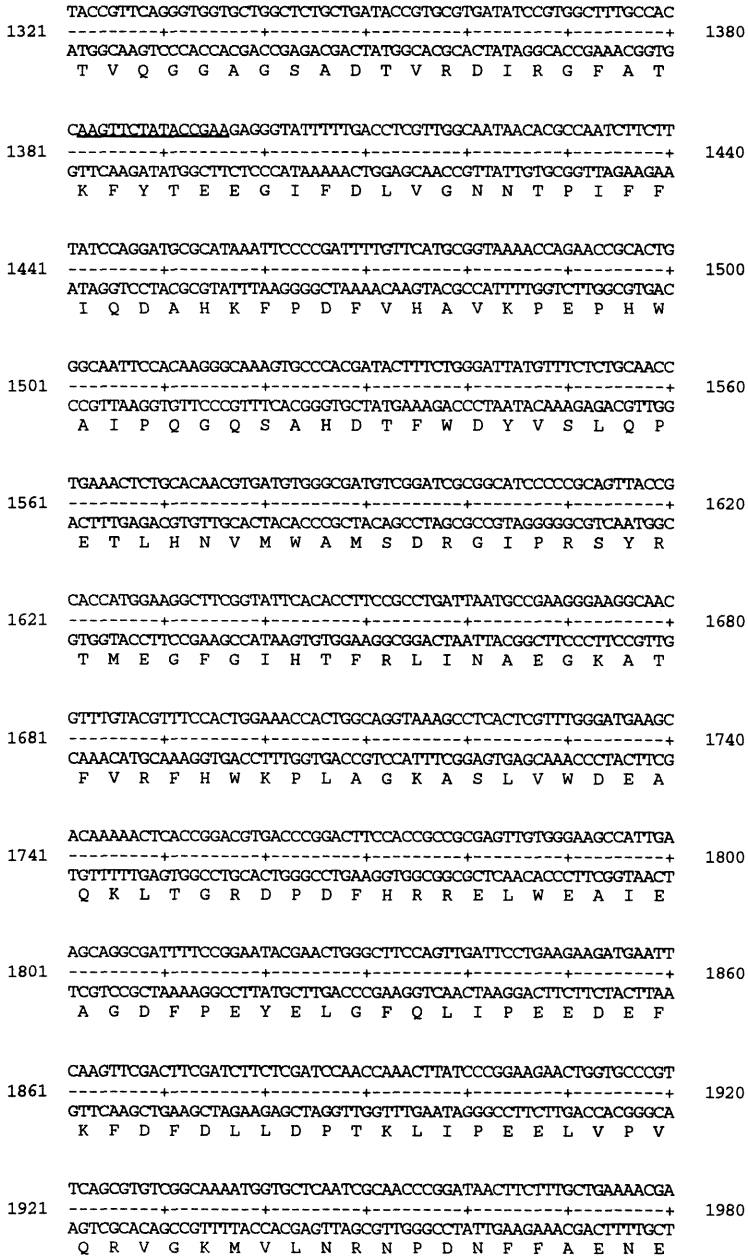


Figure 4.1 Nucleic acid sequence of the *L. seeligeri katE* gene.
Nucleic acid sequence and the translated amino acids of the *L. seeligeri* catalase gene is shown. Sequences used for primers are shown underlined.


TACCGTTCAGGGTGGTGCTGGCTCTGCTGATACCGTGCGTGATATCCGTGGCTTTGCCAC
1321
1380
ATGGCAAGTCCCACCACGACCGAGACGACTATGGCACGCACTATAGGCACCGAAACGGTG
T V Q G G A G S A D T V R D I R G F A T
CAAGTTCTATACCGAAGAGGGTATTTTTTGACCTCGTTGGCAATAACACGCCAATCTTCTT
1381
1440
GTTCAAGATATGGCTTCTCCCATAAAAACTGGAGCAACCGTTATTGTGCGGTTAGAAGAA
K F Y T E E G I F D L V G N N T P I F F
TATCCAGGATGCGCATAAATTCCCCGATTTTGTTCATGCGGTAAAACCAGAACCGCACTG
1441
1500
ATAGGTCCTACGCGTATTTAAGGGGCTAAAACAAGTACGCCATTTTGGTCTTGGCGTGAC
I Q D A H K F P D F V H A V K P E P H W
GGCAATTCCACAAGGGCAAAGTGCCCACGATACTTTCTGGGATTATGTTTCTCTGCAACC
1501
1560
CCGTTAAGGTGTTCCCGTTTCACGGGTGCTATGAAAGACCCTAATACAAAGAGACGTTGG
A I P Q G Q S A H D T F W D Y V S L Q P
TGAAACTCTGCACAACGTGATGTGGGCGATGTCGGATCGCGGCATCCCCCGCAGTTACCG
1561
1620
ACTTTGAGACGTGTTGCACTACACCCGCTACAGCCTAGCGCCGTAGGGGGCGTCAATGGC
E T L H N V M W A M S D R G I P R S Y R
CACCATGGAAGGCTTCGGTATTCACACCTTCCGCCTGATTAATGCCGAAGGGAAGGCAAC
1621
1680
GTGGTACCTTCCGAAGCCATAAGTGTGGAAGGCGGACTAATTACGGCTTCCCTTCCGTTG
T M E G F G I H T F R L I N A E G K A T
GTTTGTACGTTTCCACTGGAAACCACTGGCAGGTAAAGCCTCACTCGTTTGGGATGAAGC
1681
1740
CAAACATGCAAAGGTGACCTTTGGTGACCGTCCATTTCGGAGTGAGCAAACCCTACTTCG
F V R F H W K P L A G K A S L V W D E A
ACAAAAACTCACCGGACGTGACCCGGACTTCCACCGCCGCGAGTTGTGGGAAGCCATTGA
1741
1800
TGTTTTTGAGTGGCCTGCACTGGGCCTGAAGGTGGCGGCGCTCAACACCCTTCGGTAACT
Q K L T G R D P D F H R R E L W E A I E
AGCAGGCGATTTTCCGGAATACGAACTGGGCTTCCAGTTGATTCCTGAAGAAGATGAATT
1801
1860
TCGTCCGCTAAAAGGCCTTATGCTTGACCCGAAGGTCAACTAAGGACTTCTTCTACTTAA
A G D F P E Y E L G F Q L I P E E D E F
CAAGTTCGACTTCGATCTTCTCGATCCAACCAAACTTATCCCGGAAGAACTGGTGCCCGT
1861
1920
GTTCAAGCTGAAGCTAGAAGAGCTAGGTTGGTTTGAATAGGGCCTTCTTGACCACGGGCA
K F D F D L L D P T K L I P E E L V P V
TCAGCGTGTCGGCAAAATGGTGCTCAATCGCAACCCGGATAACTTCTTTGCTGAAAACGA
1921
1980
AGTCGCACAGCCGTTTTACCACGAGTTAGCGTTGGGCCTATTGAAGAAACGACTTTTGCT
Q R V G K M V L N R N P D N F F A E N E

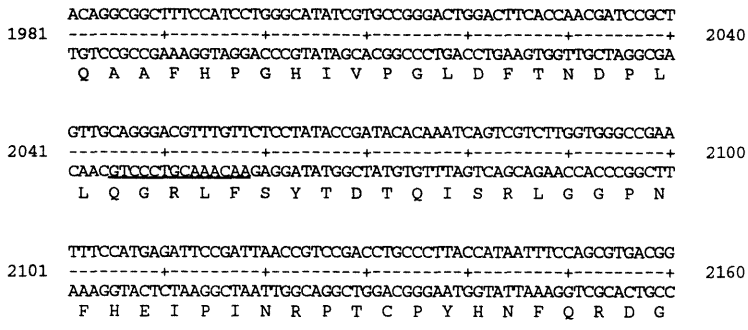


Figure 4.2 Nucleic acid sequence of the *E. coli katE* gene
Partial sequence of the *E. coli* catalase gene is shown. Primer sequences are shown underlined.

4.3 Subcloning PCR products

PCR products were subcloned using the method of Marchuk *et al.* (1990). This method exploits the template-independent terminal transferase activity of *Taq* polymerase, which results in the addition of a single nucleotide, usually adenosine, to the 3' end of the PCR product. The multicopy vector pUC18 was digested to completion with the restriction enzyme *Sma*I to yield a blunt-ended product. The digested vector (1-2 μg) was then added to 1x PCR reaction buffer (Section 4.1.3), 1.5mM magnesium, 2mM dTTP, and 5 units of *Taq* polymerase in a total volume of 100μl. The reaction mix was overlaid with 50μl of mineral oil, and incubated at 72°C for 2 hours. Following two phenol:chloroform extractions, the plasmid DNA was ethanol precipitated and resuspended to 200ng/μl.

To ligate the PCR product to the T-tailed vector, 100ng of vector was mixed with 250-500ng of PCR product, 4μl of 5x ligase buffer (GIBCO/BRL) and 8 units (1μl) of T4 ligase (GIBCO/BRL) in a final volume of 20μl. Ligations were incubated at 14°C for 12-18 hours, then ethanol precipitated and resuspended in 6μl of sterile nanopure water. Two microlitre samples of this solution were then electroporated into electrocompetent *E. coli* DH5α as described previously (Section 2.4.1.2).

4.4 Construction and screening of partial libraries of *Mycobacterium gordonae* and *Mycobacterium terrae*.

4.4.1 Construction of partial genomic libraries of *M. terrae* and *M. gordonae*.

Mycobacterium terrae and *M. gordonae* genomic DNA was isolated using the method described in Section 2.3.3. Genomic DNA (15-20μg) was digested to completion with *Eco*RI, such that a "ladder" of fragments was visible when the digests were analysed by agarose gel electrophoresis. Approximately 5μg of the *M. terrae* and *M. gordonae* digests were electrophoresed through a 0.5% (w/v) preparative agarose gel. Agarose gel slices containing DNA fragments from (a) 4-7 kb and (b) 7-10 kb were removed from the *M. terrae* and *M. gordonae* digests. DNA was purified from these gel slices using the Sephaglas Bandprep kit (Pharmacia) as described in Section 2.5.6.

The plasmid pBR322 was digested to completion with *Eco*RI. The extent of digestion was measured by transforming aliquots of the digest into *E. coli* DH5α and measuring the numbers of transformants; zero or very few transformants indicated the plasmid was completely digested. The plasmid was then dephosphorylated as described in Section 2.5.3, using the protocol for protruding 5' termini. Analysis by self-ligation showed that approximately 90% of the *Eco*RI-cut pBR322 was dephosphorylated.

Size-selected *M. gordonae* and *M. terrae* DNAs were then ligated to the dephosphorylated pBR322 (100-200ng) at an insert:vector ratio of 3:1. Ligations were performed as described in Section 2.5.4. After incubation, ligations were ethanol precipitated and resuspended in 6μl of sterile nanopure water. Two microlitre samples of the ligation were then transformed by electroporation into electrocompetent *E. coli* DH5α as outlined in Section 2.4.1.2. Transformants were selected on L-agar containing 15μg/ml of tetracycline and 100μg/ml of ampicillin.

4.4.2 Screening the partial libraries with the *E. coli katE* PCR product.

As an *E. coli* probe was being used to screen for mycobacterial catalase sequences on low-copy number plasmids in *E. coli* hosts, the library could not be screened using a standard colony-blotting technique as background would be excessively high. Therefore, an alternative approach was adopted. Plasmid DNA was extracted from library transformants using the method described in Section 2.3.1.2. Plasmid DNA was digested with *Eco*RI, and the resulting DNA fragments separated by agarose gel electrophoresis through a 0.7% (w/v) agarose gel. The plasmid inserts were then probed directly in the gel with the 650bp radiolabelled PCR product from the *E. coli katE* gene. This *in situ* gel hybridisation technique was based on the method of Wallace and Miyada (1987). After electrophoresis, gels were stained with ethidium bromide and photographed alongside a ruler. The DNA was denatured by soaking the gel in 0.5M NaOH, 150mM NaCl for 30 minutes at room temperature with gentle rocking. The gel was neutralised by soaking in 0.5M Tris-HCl (pH 8.0), 150mM NaCl for 30 minutes at room temperature. The gel was then placed on 2 sheets of Whatman 3mm filter paper, put into a gel dryer (Bio-Rad), covered with the neoprene sheet (lid left open), and dried at room temperature

under vacuum until the gel was almost flat (30-45 minutes). The heater of the gel dryer was then turned on and set to 60°C. The gel was dried under vacuum for a further 30-60 minutes. After this procedure, the gel was in the form of a thin, translucent membrane attached to the filter paper. For long-term storage, the gel plus filter paper could be wrapped in Saran wrap (DuPont) and placed at 4°C. Prior to hybridisation, the gel was removed from the filter paper by soaking for 1-2 minutes in distilled water; the gel was now treated in the same way as a nylon membrane, except without the need for prehybridisation. The *E. coli katE* PCR product was radiolabelled using the Ready to Go (Pharmacia) probe labelling system (see Section 2.6.2). Hybridisation under conditions of low stringency was performed as detailed in Section 2.6.3.2. After hybridisation, the gel was dried by blotting onto filter paper, then wrapped in Saran wrap (DuPont) and exposed to X-ray film (Cronex, DuPont) at -70°C for 8-24 hours. Films were developed in an Agfa-Geveart automatic film processor according to the manufacturer's instructions.

Chapter 5.

Results

5.1 Alignment of catalase gene and protein sequences

The GenBank and EMBL nucleic acid and protein databases were searched for catalase sequences using programs from the University of Wisconsin's Genetics Computer Group (GCG Version 7-UNIX) software package (Genetics Computer Group, 1991). Sequences identified in this search were divided into two groups, the "typical" catalases and the "catalase-peroxidases". The former group contained many eukaryotic catalase gene sequences, as well as the sequences for the *E. coli* and *L. seeligeri* typical catalase (*katE*) genes. The catalase-peroxidase group consisted of the nucleic acid sequences for the *E. coli*, *Salmonella typhimurium* and *Bacillus stearothermophilus* genes. In cases where protein sequences were not available, gene sequences were translated into the corresponding amino acid sequence using the "translate" program from the GCG software package. Alignment of gene and protein sequences was then achieved using the CLUSTAL V multiple alignment program (Higgins *et al.*, 1992).

Alignment of the nucleic acid sequences of the typical catalase genes, analogous to the mycobacterial M-catalase, revealed no significant homologies. However, alignment of the amino acid sequences revealed three regions in the typical catalase proteins that were conserved across a wide range of organisms (Figure 5.1). The conserved regions were denoted regions 1, 2 and 3. These three regions appeared to be suitable as sites for the production of PCR primers, since (i) they were 5 amino acids (15 nucleotides) in length, long enough for a primer, (ii) they contained a number of codons with low degeneracy.

Alignment of the catalase-peroxidase sequences, analogous to the mycobacterial T-catalase, revealed regions of homology at the nucleic acid (Figure 5.2) and protein level (Figure 5.3). However, as only three sequences were available at the time and since the *E. coli* and *S. typhimurium* gene and amino acid sequences are 78% and 91% homologous respectively, the question arose whether these regions would appear conserved across a wider range of bacterial genera. More catalase-peroxidase sequences would need to be aligned to verify

whether these regions are conserved or just artefacts of aligning 3 sequences.

The amplification of the M-catalase gene was chosen as the starting point for the project. The three conserved regions identified among the typical catalase proteins were suitable for the design of degenerate oligonucleotides, so this appeared to offer a good chance of success. It was hoped that the protocols and experience garnered from the amplification of the M-catalase could then be applied in turn to the T-catalase amplification. Also, more catalase-peroxidase sequences may have been published while the amplification of the M-catalase was ongoing. However, as stated in the project aims (Section 3.3), the cloning of the T-catalase gene from *M. tuberculosis* by Zhang *et al.* (1992) lead to a reappraisal of the project's objectives. The cloning of the M-catalase now became the main goal with a view to deciphering its role in isoniazid sensitivity.

5.2 Design of degenerate oligonucleotide primers

A range of oligonucleotide primers was synthesised that could code for the conserved amino acid residues in the typical catalase enzymes (Table 5.1). Codon usage for the mycobacteria (codon usage tables were those of Dale and Patki, 1990) was taken into account to decrease the complexity of the primer pool. To further decrease complexity, the degenerate nucleotide position at the 3' end of the primer was removed, producing a primer length of 14 rather than 15 nucleotides.

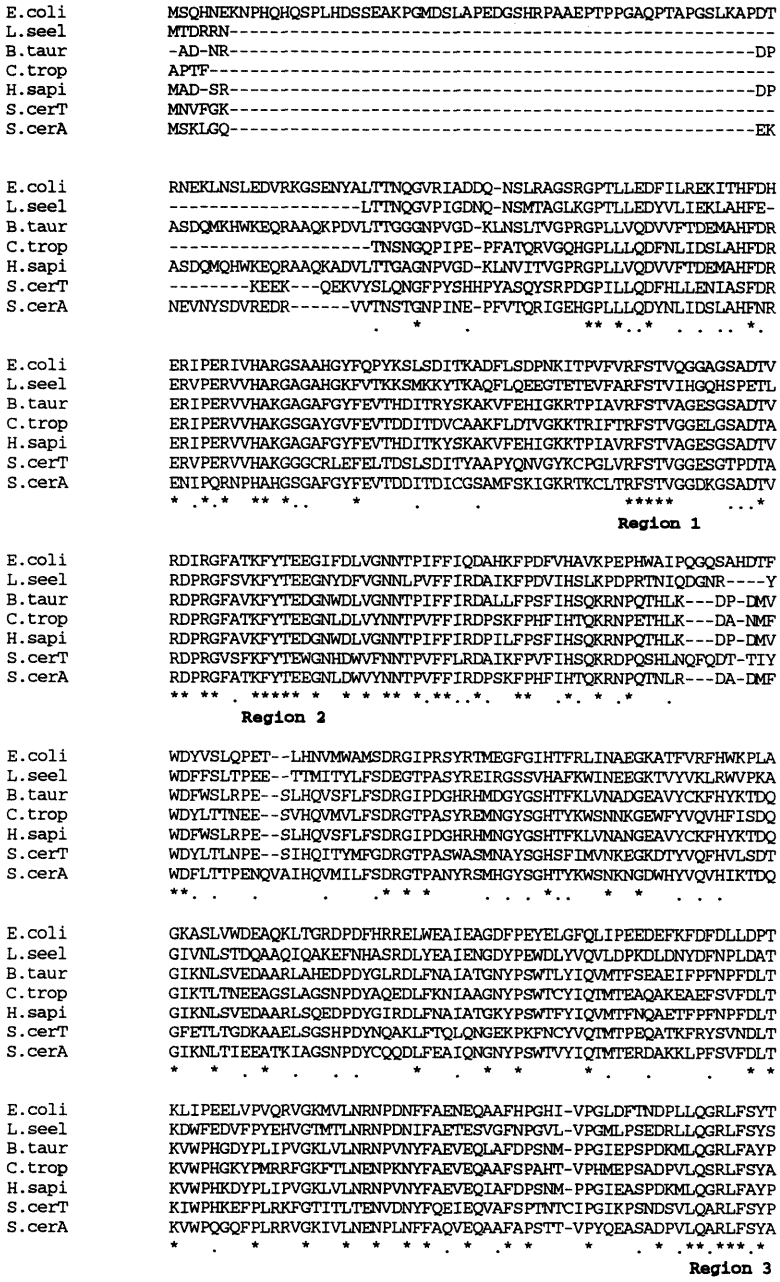
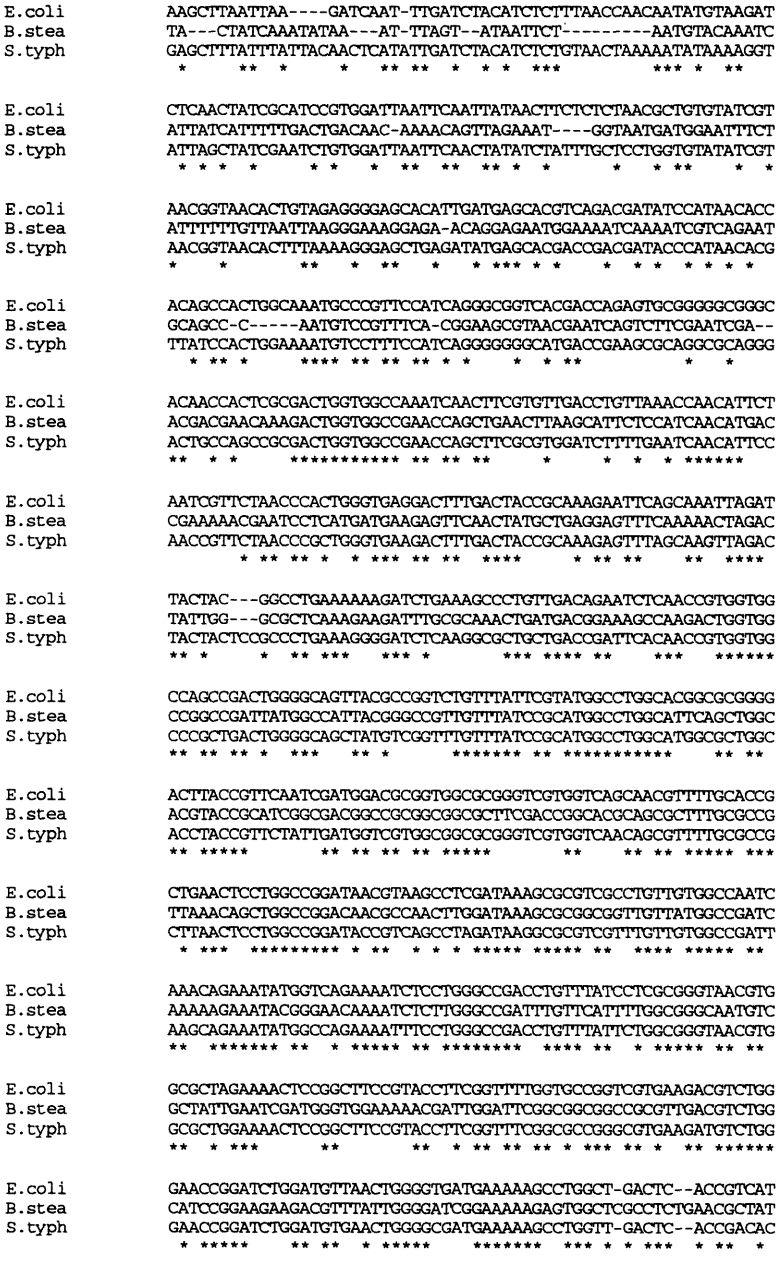

E.coli MSQHNEKNPHQHQSPLHDSSEAKPGMDSLAPEDGSHRPAAEPTPPGAQPTAPGSLKAPDT
L.seel MTDRRN
B.taur -AD-NR DP
C.trop APTF
H.sapi MAD-SR DP
S.cerT MNVFGK
S.cerA MSKLGQ EK
E.coli RNEKLNSLEDVRKGSENYALTTNQGVRIADDQ-NSLRAGSRGPTLLEDFILREKITHFDH
L.seel LTTNQGVPIGDNQ-NSMTAGLKGPTLLEDYVLIEKLAHFE-
B.taur ASDQMKHWKEQRAAQKPDVLTTGGGNPVGD-KLNSLTVGPRGPLLVQDVVFTDEMAHFDR
C.trop TNSNGQPIPE-PFATQRVGQHGPLLLQDFNLIDSLAHFDR
H.sapi ASDQMQHWKEQRAAQKADVLTTGAGNPVGD-KLNVITVGPRGPLLVQDVVFTDEMAHFDR
S.cerT KEEK---QEKVYSLQNGFPYSHHPYASQYSRPDGPILLQDFHLLENIASFDR
S.cerA NEVNYSDVREDR VVTNSTGNPINE-PFVTQRIGEHGPLLLQDYNLIDSLAHFNR
E.coli ERIPERIVHARGSAAHGYFQPYKSLSDITKADFLSDPNKITPVFVRFSTVQGGAGSADTV
L.seel ERVPERVVHARGAGAHGKFVTKKSMKKYTKAQFLQEEGTETEVFARFSTVIHGQHSPETL
B.taur ERIPERVVHAKGAGAFGYFEVTHDITRYSKAKVFEHIGKRTPIAVRFSTVAGESGSADTV
C.trop ERIPERVVHAKGSGAYGVFEVTDDITDVCAAKFLDTVGKKTRIFTRFSTVGGELGSADTA
H.sapi ERIPERVVHAKGAGAFGYFEVTHDITKYSKAKVFEHIGKKTPIAVRFSTVAGESGSADTV
S.cerT ERVPERVVHAKGGGCRLEFELTDSLSDITYAAPYQNVGYKCPGLVRFSTVGGESGTPDTA
S.cerA ENIPQRNPHAHGSGAFGYFEVTDDITDICGSAMFSKIGKRTKCLTRFSTVGGDKGSADTV
Region 1
E.coli RDIRGFATKFYTEEGIFDLVGNNTPIFFIQDAHKFPDFVHAVKPEPHWAIPQGQSAHDTF
L.seel RDPRGFSVKFYTEEGNYDFVGNNLPVFFIRDAIKFPDVIHSLKPDPRTNIQDGNR----Y
B.taur RDPRGFAVKFYTEDGNWDLVGNNTPIFFIRDALLFPSFIHSQKRNPQTHLK---DP-DMV
C.trop RDPRGFATKFYTEEGNLDLVYNNTPVFFIRDPSKFPHFIHTQKRNPETHLK---DA-NMF
H.sapi RDPRGFAVKFYTEDGNWDLVGNNTPIFFIRDPILFPSFIHSQKRNPQTHLK---DP-DMV
S.cerT RDPRGVSFKFYTEWGNHDWVFNNTPVFFLRDAIKFPVFIHSQKRDPQSHLNQFQDT-TIY
S.cerA RDPRGFATKFYTEEGNLDWVYNNTPVFFIRDPSKFPHFIHTQKRNPQTNLR---DA-DMF
Region 2
E.coli WDYVSLQPET--LHNVMWAMSDRGIPRSYRTMEGFGIHTFRLINAEGKATFVRFHWKPLA
L.seel WDFFSLTPEE--TTMITYLFSDEGTPASYREIRGSSVHAFKWINEEGKTVYVKLRWVPKA
B.taur WDFWSLRPE--SLHQVSFLFSDRGIPDGHRHMDGYGSHTFKLVNADGEAVYCKFHYKTDQ
C.trop WDYLTTNEE--SVHQVMVLFSDRGTPASYREMNGYSGHTYKWSNNKGEWFYVQVHFISDQ
H.sapi WDFWSLRPE--SLHQVSFLFSDRGIPDGHRHMNGYGSHTFKLVNANGEAVYCKFHYKTDQ
S.cerT WDYLTLNPE--SIHQITYMFGDRGTPASWASMNAYSGHSFIMVNKEGKDTYVQFHVLSDT
S.cerA WDFLTTPENQVAIHQVMILFSDRGTPANYRSMHGYSGHTYKWSNKNGDWHYVQVHIKTDQ
E.coli GKASLVWDEAQKLTGRDPDFHRRELWEAIEAGDFPEYELGFQLIPEEDEFKFDFDLLDPT
L.seel GIVNLSTDQAAQIQAKEFNHASRDLYEAIENGDYPEWDLYVQVLDPKDLDNYDFNPLDAT
B.taur GIKNLSVEDAARLAHEDPDYGLRDLFNAIATGNYPSWTLYIQVMTFSEAEIFPFNPFDLT
C.trop GIKTLTNEEAGSLAGSNPDYAQEDLFKNIAAGNYPSWTCYIQIMTEAQAKEAEFSVFDLT
H.sapi GIKNLSVEDAARLSQEDPDYGIRDLFNAIATGKYPSWTFYIQVMTFNQAETFPFNPFDLT
S.cerT GFETLTGDKAAELSGSHPDYNQAKLFTQLQNGEKPKFNCYVQIMTPEQATKFRYSVNDLT
S.cerA GIKNLTIEEATKIAGSNPDYCQQDLFEAIQNGNYPSWTVYIQIMTERDAKKLPFSVFDLT
E.coli KLIPEELVPVQRVGKMVLNRNPDNFFAENEQAAFHPGHI-VPGLDFTNDPLLQGRLFSYT
L.seel KDWFEDVFPYEHVGTMTLNRNPDNIFAETESVGFNPGVL-VPGMLPSEDRLLQGRLFSYS
B.taur KVWPHGDYPLIPVGKLVLNRNPVNYFAEVEQLAFDPSNM-PPGIEPSPDKMLQGRLFAYP
C.trop KVWPHGKYPMRRFGKFTLNENPKNYFAEVEQAAFSPAHT-VPHMEPSADPVLQSRLFSYA
H.sapi KVWPHKDYPLIPVGKLVLNRNPVNYFAEVEQIAFDPSNM-PPGIEASPDKMLQGRLFAYP
S.cerT KIWPHKEFPLRKFGTITLTENVDNYFQEIEQVAFSPTNTCIPGIKPSNDSVLQARLFSYP
S.cerA KVWPQGQFPLRRVGKIVLNENPLNFFAQVEQAAFAPSTT-VPYQEASADPVLQARLFSYA
Region 3

Figure 5.1 Alignment of typical catalase protein sequences.

Protein sequences from a range of typical catalase enzymes were aligned using the CLUSTAL V package (Higgins *et al.*, 1992). Catalase sequences of the following were extracted from the Swissprot protein database; *Listeria seeligeri* (L.seel, accession number P24168), *Bos taurus* (B.taur, accession number P00432), *Homo sapiens* (H.sapi, accession number P04040), *Saccharomyces cerevisiae* T catalase (S.cerT, accession number P15202) and *S. cerevisiae* A catalase (S.cerA, accession number P06115). Nucleic acid sequences of the catalase enzymes from *Escherichia coli* (E.coli, accession M55161) and *Candida tropicalis* (C.trop, accession number X06660) were extracted from the EMBL sequence database and translated to the corresponding amino acid sequence using the "translate" program from the Genetics Computer Group package. In the figure, "*" denotes identical residues, "." indicates conservative substitutions.


E.coli AAGCTTAATTAA----GATCAAT-TTGATCTACATCTCTTTAACCAACAATATGTAAGAT
B.stea TA---CTATCAAATATAA---AT-TTAGT--ATAATTCT---------AATGTACAAATC
S.typh GAGCTTTATTTATTACAACTCATATTGATCTACATCTCTGTAACTAAAAATATAAAAGGT
E.coli CTCAACTATCGCATCCGTGGATTAATTCAATTATAACTTCTCTCTAACGCTGTGTATCGT
B.stea ATTATCATTTTTGACTGACAAC-AAAACAGTTAGAAAT----GGTAATGATGGAATTTCT
S.typh ATTAGCTATCGAATCTGTGGATTAATTCAACTATATCTATTTGCTCCTGGTGTATATCGT
E.coli AACGGTAACACTGTAGAGGGGAGCACATTGATGAGCACGTCAGACGATATCCATAACACC
B.stea ATTTTTTGTTAATTAAGGGAAAGGAGA-ACAGGAGAATGGAAAATCAAAATCGTCAGAAT
S.typh AACGGTAACACTTTAAAAGGGAGCTGAGATATGAGCACGACCGACGATACCCATAACACG
E.coli ACAGCCACTGGCAAATGCCCGTTCCATCAGGGCGGTCACGACCAGAGTGCGGGGGCGGGC
B.stea GCAGCC-C-----AATGTCCGTTTCA-CGGAAGCGTAACGAATCAGTCTTCGAATCGA--
S.typh TTATCCACTGGAAAATGTCCTTTCCATCAGGGGGGGCATGACCGAAGCGCAGGCGCAGGG
E.coli ACAACCACTCGCGACTGGTGGCCAAATCAACTTCGTGTTGACCTGTTAAACCAACATTCT
B.stea ACGACGAACAAAGACTGGTGGCCGAACCAGCTGAACTTAAGCATTCTCCATCAACATGAC
S.typh ACTGCCAGCCGCGACTGGTGGCCGAACCAGCTTCGCGTGGATCTTTTGAATCAACATTCC
E.coli AATCGTTCTAACCCACTGGGTGAGGACTTTGACTACCGCAAAGAATTCAGCAAATTAGAT
B.stea CGAAAAACGAATCCTCATGATGAAGAGTTCAACTATGCTGAGGAGTTTCAAAAACTAGAC
S.typh AACCGTTCTAACCCGCTGGGTGAAGACTTTGACTACCGCAAAGAGTTTAGCAAGTTAGAC
E.coli TACTAC---GGCCTGAAAAAAGATCTGAAAGCCCTGTTGACAGAATCTCAACCGTGGTGG
B.stea TATTGG---GCGCTCAAAGAAGATTTGCGCAAACTGATGACGGAAAGCCAAGACTGGTGG
S.typh TACTACTCCGCCCTGAAAGGGGATCTCAAGGCGCTGCTGACCGATTCACAACCGTGGTGG
E.coli CCAGCCGACTGGGGCAGTTACGCCGGTCTGTTTATTCGTATGGCCTGGCACGGCGCGGGG
B.stea CCGGCCGATTATGGCCATTACGGGCCGTTGTTTATCCGCATGGCCTGGCATTCAGCTGGC
S.typh CCCGCTGACTGGGGCAGCTATGTCGGTTTGTTTATCCGCATGGCCTGGCATGGCGCTGGC
E.coli ACTTACCGTTCAATCGATGGACGCGGTGGCGCGGGTCGTGGTCAGCAACGTTTTGCACCG
B.stea ACGTACCGCATCGGCGACGGCCGCGGCGGCGCTTCGACCGGCACGCAGCGCTTTGCGCCG
S.typh ACCTACCGTTCTATTGATGGTCGTGGCGGCGCGGGTCGTGGTCAACAGCGTTTTGCGCCG
E.coli CTGAACTCCTGGCCGGATAACGTAAGCCTCGATAAAGCGCGTCGCCTGTTGTGGCCAATC
B.stea TTAAACAGCTGGCCGGACAACGCCAACTTGGATAAAGCGCGGCGGTTGTTATGGCCGATC
S.typh CTTAACTCCTGGCCGGATACCGTCAGCCTAGATAAGGCGCGTCGTTTGTTGTGGCCGATT
E.coli AAACAGAAATATGGTCAGAAAATCTCCTGGGCCGACCTGTTTATCCTCGCGGGTAACGTG
B.stea AAAAAGAAATACGGGAACAAAATCTCTTGGGCCGATTTGTTCATTTTGGCGGGCAATGTC
S.typh AAGCAGAAATATGGCCAGAAAATTTCCTGGGCCGACCTGTTTATTCTGGCGGGTAACGTG
E.coli GCGCTAGAAAACTCCGGCTTCCGTACCTTCGGTTTTGGTGCCGGTCGTGAAGACGTCTGG
B.stea GCTATTGAATCGATGGGTGGAAAAACGATTGGATTCGGCGGCGGCCGCGTTGACGTCTGG
S.typh GCGCTGGAAAACTCCGGCTTCCGTACCTTCGGTTTCGGCGCCGGGCGTGAAGATGTCTGG
E.coli GAACCGGATCTGGATGTTAACTGGGGTGATGAAAAAGCCTGGCT-GACTC--ACCGTCAT
B.stea CATCCGGAAGAAGACGTTTATTGGGGATCGGAAAAAGAGTGGCTCGCCTCTGAACGCTAT
S.typh GAACCGGATCTGGATGTGAACTGGGGCGATGAAAAAGCCTGGTT-GACTC--ACCGACAC

```
E.coli          CCGGAAG--CGCTGGC--GAAAGCACCGCTGGGTGCAACCGAGATGGGTCTGATTTACGT
B.stea          TCCGGTGATCGCGAGCTCGAAAAC-CCGCTCGCCGCGGTGCAAATGGGGTTAATCTACGT
S.typh          CCTGAAG--CGCTGGC--AAAAGCGCCGCTGGGCGCGACCGAGATGGACCTTATCTACGT
                 * *  *  ***  **   *** * ***** *  **     * ****   * ** *****

E.coli          TAACCCGGAAGGCCCGGATCACAGCGGCGAACCGCTTTCTGCGGCAGCAGCTATCCGCGC
B.stea          CAACCCAGAAGGGCCGGACGGCAAGCCGGATCCAAAAGCAGCGGCGCGCGATATCCGCGA
S.typh          TACACCGGAAGGGCCGAATCACAGCGGCGAACCACTTTCTGCCGCCGCCGCTATTCGCGC
                 *  ** ***** *** *   **     ** **     * ** **    * *** ****

E.coli          GACCTTCGGCAACATGGGCATGAACGACGAAGAAACCGTGGCGCTGATTGCGGGTGGTCA
B.stea          GACGTTCCGCCGCATGGGGATGAACGATGAAGAAACGGTCGCCTTGATCGCCGGCGGTCA
S.typh          TACCTTTGGCAATATGGGGATGAACGACGAAGAGACCGTGGCGTTGATCGCTGGCGGGCA
                 ** **  **   ***** ******** ***** ** ** **  **** ** ** ** **

E.coli          TACGCTGGGTAAAACCCACGGTGCCGGTCCGACATCAAATGTAGGTCCTGATCCAGAAGC
B.stea          TACGTTCGGAAAAGCGCACGGCGCCGGCCCTGCCACGCACGTCGGTCCCGAGCCGGAAGC
S.typh          TACCCTCGGTAAAACCCACGGTCCGGCAGCGGCATCCCATGTAGGGGCCGATCCGGAAGC
                ***  * ** *** * *****  * *   *  *  *  * ** **  * ** ** *****

E.coli          TGCACCGATTGAAGAACAAGGTTTAGGTTGGGCGAGCACTTACGGCAGCGGCGTTGGCGC
B.stea          CGCCCCGATTGAAGCGCAAGGGCTGGGATGGATCAGCTCTTACGGAAAAGGGAAAGGGAG
S.typh          CGCGCCGATTGAAGCGCAGGGCTTAGGTTGGGCCAGCAGCTATGGTAGTGGCGTTGGCGC
                 ** **********  ** **  * ** ***   ***   ** ** *  **    **

E.coli          AGATGCCATTACCTCTGGTCTGGAAGTAGTCTGGACCCAGACGCCGACCCAGTGGAGCAA
B.stea          CGATACGATCACAAGCGGCATTGAAGGCGCTTGGACGCCGACGCCAACCCAGTGGGATAC
S.typh          GGATGCTATCACCTCCGGGCTGGAAGTGGTCTGGACGCAGACGCCGACCCAGTGGAGCAA
                 *** * ** **    **  * ****  *  ***** * ****** *********   *

E.coli          CTATTTCTTCGAGAACCTGTTCAAGTATGAGTGGGTACAGACCCGCAGCCCGGCTGGCGC
B.stea          GTCGTACTTTGACATGCTGTTTGGCTATGACTGGTGGCTGACGAAAAGCCCAGCCGGGGC
S.typh          CTATTTCTTCGAGAACCTGTTCAAATATGAGTGGGTACAAACCCGTAGTCCGGCTGGCGC
                 *  * *** ** *  *****    ***** ***   *  **    ** ** ** ** **

E.coli          AATCCAGTTCGAAGCGGTAGACGCAC-CGGAAA-----TTATCCCGGATCCGTTTGATCC
B.stea          ATGGCAATGGATGGCGGTCGACCCGGATGAAAAAGACTTGGCTCCGGATGCCGAGGATCC
S.typh          TATCCAGTTTGAAGCGGTAGACGCGC-CGGACA-----TCATCCCGGACCCGTTTGATCC
                    ** *     ***** *** *    * * *     *    *****  *    *****

E.coli          GTCGAAGAAACGTAAACCGACAATGCTGGTGACCGACCTGACGCTGCGTTTTGATCCTGA
B.stea          GTCGAAAAAAGTTCCGACGATGATGATGACGACCGATTTGGCGCTTCGGTTTGACCCGGA
S.typh          GTCGAAAAAACGTAAGCCAACCATGCTGGTCACCGACCTGACGCTGCGTTTTGATCCGGA
                ****** ***  *    * *  *** **   *****  ** **** ** ***** ** **

E.coli          GTTCGAGAAGATCTCTCGTCGTTTCCTCAACGATCCGCAGGCGTTCAACGAAGCCTTTGC
B.stea          ATACGAAAAAATCGCCCGCCGATTCCATCAAAATCCAGAAGAATTTGCGGAGGCGTTTGC
S.typh          GTTCGAGAAGATTTCCCGTCGTTTCCTTAACGATCCGCAGGCCTTTAATGAAGCCTTTGC
                 * *** ** **  * ** ** ****   *  ****  * *  **    ** ** *****

E.coli          CCGTGCCTGGTTCAAACTGACGCACAGGGATATGGGGCCGAAATCTCGCTACATCGGGCC
B.stea          CCGGGCATGGTTCAAGCTTACCCATAGGGATATGGGACCGAAAACGAGATATCTCGGCCC
S.typh          TCGTGCCTGGTTCAAACTGACGCACAGAGATATGGGACCAAAAGCGCGTTACATCGGACC
                 ** ** ******** ** ** ** ** ******** ** *** *  * **  **** **

E.coli          GGAAGTGCCGAAAGAAGATCTGATCTGGCAAGATCCGCTGCCGCAGCCGATCTACAACCC
B.stea          GGAAGTTCCGAAAGAAGATTTCATCTGGCAAGATCCGATTCCAGAAGTCGATTACGAATT
S.typh          GGAAGTGCCGAAAGAAGATCTGATCTGGCAGGACCCGTTGCCGCAACCGCTCTATCAGCC
                ****** ************ * ******** ** *** * **  *        **  *

E.coli          GACCGAGCAGGACATTATCGATCTGAAATTCGCGATTGCGGATTCTGGTCTGTCTGTTAG
B.stea          GACAGAAGCGGAAATTGAAGAAATCAAAGCCAAAATTTTGAACTCGGGCCTGACCGTCAG
S.typh          AACGCAGGAAGACATTATCAACCTGAAAGCGGCGATCGCTGCATCCGGGCTTTCTATTAG
                 **  *    ** ***    *  * ***       **        ** ** **  *  * **

E.coli          TGAGCTGGTATCGGTGGCCTGGGCATCTGCTTCTACCTTCCGTGGTGGCGACAAACGCGG
B.stea          TGAACTTGTGAAAACAGCTTGGGCCTCGGCCAGCACGTTCCGCAACTCGGATAAGCGCGG
S.typh          CGAGATGGTTTCGGTTGCCTGGGCATCCGCTTCTACTTTCCGCGGCGGCGATAAGCGTGG
                 **  * **        ** ***** ** **    ** *****        ** ** ** **
```

```
E.coli          TGGTGCCAACGGTGCGCGTCTGGCATTAATGCCGCAGCGCGACTGGGATGTGAACG--CC
B.stea          CGGAGCCAACGGCGCCCGCATTCGTCTCGCGCCGCAAAAAGACTGGGAAGTGAACGAACC
S.typh          CGGCGCTAACGGCGCGCGTCTGGCATTAGCGCCTCAGCGCGACTGGGATGTCAACG--CC
                 ** ** ***** ** **  *      *   *** **    ******** ** ****  **

E.coli          GCAGCCGTTCGT----GCTCTGCCTGTTCTGGAGAAAATCCAG---AAAGAGTCTGGTAA
B.stea          GGAGCGGCTCGCCAAAGTGCTGTCCGTCTACGAGGACATCCAGCGCGAACTGCCGAAAAA
S.typh          GTTGCGGCTCGC----GTTCTGCCGGTATTAGAAGAGATCCAG---AAAACGACGAATAA
                *  ** * ***     *  *** * **    **  * ******    **  * *    **

E.coli          AGCCTCGCTGGCGGATATCATAGTGCTGGCTGGTGTGGTTGGTGTTGAGAAAGCCGCAAG
B.stea          AGTAAGCATCGCCGACTTGATCGTCCTTGGCGGCAGCGCTGCGGTGGAAAAGGCAGCCCG
S.typh          AGCCTCGCTGGCCGATATTATTGTGCTGGCGGGCGTGGTCGGTATCGAGCAGGCGGCCGC
                **       * ** **  * ** ** ** * **    *  *   * **  * ** **

E.coli          CGCCGCAGGTTTGAGCATTCATGTACCGTTTGCGCCGGGTCGCGTTGATGCGCGTCAGGA
B.stea          CGACGCCGGCTTTGATGTCAAAGTGCCATTTTTCCCTGGCCGCGGCGATGCGACACAAGA
S.typh          TGCTGCGCGTGTCAGCATTCACGTACCTTTTCCGCCGGGCCGGGTGGATGCGCGTCATGA
                 *  **  *  *      *  * ** ** ***   ** ** ** *  ******   ** **

E.coli          TCAGACTGACATTGAGATGTTTGAGCTGCTGGAGCCAATTGCTGACGGTTTCCGTAACTA
B.stea          GCAAACCGATGTCGAAAGCTTTGCCGTGTTGGAACCGTTCGCAGATGGCTTCCGCAACTA
S.typh          TCAGACCGACATCGAGATGTTCTCGCTGCTTGAACCGATTGCTGATGGATTCCGTAACTA
                 ** ** **  * ** *  **      ** * ** **  * ** ** ** ***** *****

E.coli          TCGCGCTCGTCTGGACGTTTCCACCACCGAGTCACTGCTGATCGACAAAGCACAGCAACT
B.stea          TCAAAAGCAAGAGTACAGCGTTCCGCCGGAAGAGCTGCTCGTCGACAAAGCCCAGCTCCT
S.typh          TCGTGCGCGTCTGGATGTGTCGACGACCGAATCGCTGTTGATTGATAAAGCGCAGCAGTT
                **      *    * *         *  * **    *** *  * ** ***** ****   *

E.coli          GACGCTGACCGCGCCGGAAATGACTGCGCTGGTGGGCGGCATGCGTGTACTGGGTGGCAA
B.stea          CGGGCTGACCGCCCCAGAAATGACGGTCTTAGTTGGCGGTTTGCGCGTGTTGGGCGCGAA
S.typh          AACGTTGACCGCGCCGGAAATGACGGTACTGGTTGGCGGGATGCGTGTGCTGGGAACCAA
                   * ******* ** ******** *   * ** ***** **** **  ****    **

E.coli          CTTCGATGGCAGCAAAAACGGCGTCTTCACTGACCGCGTTGGCGTATTGAGCAATGACTT
B.stea          CTATCGCGATCTGCCTCATGGCGTCTTCACTGACCGCATCGGGGTGCTGACCAACGACTT
S.typh          CTTTGACGGCAGCCAGAACGGTGTCTTTACCGACAAACCGGGCGTGCTCAGCACTGACTT
                **      *          * ** ***** ** ***       ** **  * * **  *****

E.coli          CTTCGTGAACTTGCTGGATATGCGTTACGAGTGGAAAGCGACCGACGAATCGAAAGAGCT
B.stea          CTTTGTCAACTTGTTGGATATGAACTATGAATGGGTGCCGACAGAC------AGCGGCAT
S.typh          CTTCGCTAATCTGCTGGATATGCGTTACGAGTGGAAGCCCACCGACGACGCTAATGAGCT
                *** *  **  ** ********   ** ** ***    * ** ***       *  *   *

E.coli          GTTCGAAGGCCGTGACCGTGAAACCGGCGAAGTGAAATTTACGGCCAGCCGTGCGGATCT
B.stea          TTATGAAATCCGCGACCGGAAAACGGGCGAAGTGCGGTGGACGGCGACCCGGGTGGATCT
S.typh          GTTCGAAGGCCGGGATCGTCTGACTGGCGAGGTAAAATACACGGCGACCCGCGCCGATCT
                 *  ***  *** ** **    ** ***** **    *  ***** * *** *  *****

E.coli          GGTGTTTGGTTCTAACTCCGTCCTGCGTGCGGTGGCGGAAGTTTACGCCAGTAGCGATGC
B.stea          CATTTTCGGATCCAACTCCATTCTTCGTTCCTACGCGGAATTTTACGCCCAAGACGACAA
S.typh          GGTGTTTGGTTCCAACTCCGTACTGCGCGCGCTGGCGGAAGTTTACGCGTGTAGCGATGC
                  * ** ** ** ****** * ** **  *    ****** *******      ***

E.coli          CCACGAGAAGTTTGTTAAAGACTTCGTGGCGGCATGGGTGAAAGTGATGAACCTCGACCG
B.stea          CCAAGAAAAATTTGTCCGTGATTTCATCAACGCTTGGGTGAAAGTCATGAACGCCGACCG
S.typh          GCACGAGAAGTTTGTGAAGGACTTCGTCGCGGCATGGGTGAAAGTGATGAACCTGGACCG
                 ** ** ** *****    ** *** *    ** ********** ******   *****

E.coli          TTTCGACCTGCTGTAATCTGACCCCGTTCAGCGGCTGCTTGCTGGCAGTCGCTGAACGTT
B.stea          CTTTGATCTGGTGAAA---AA--------AGCAA---------GGGAATCG---------
S.typh          TTTCGATCTGCAATAAT-------------------------------------------
                 ** ** ***    **
```

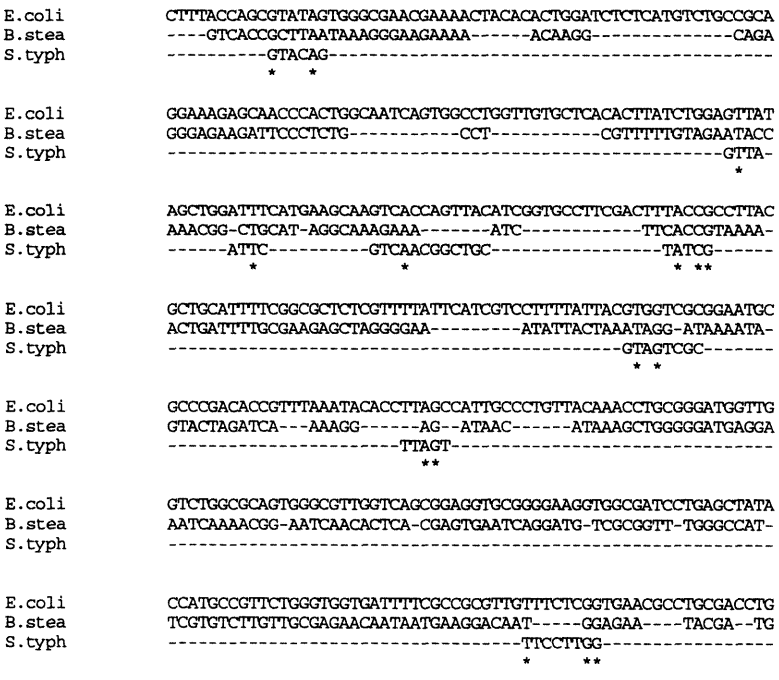


Figure 5.2 Alignment of catalase-peroxidase nucleic acid sequences.
The nucleic acid sequences of the catalase-peroxidase enzymes from *Escherichia coli* (E.coli, EMBL accession number M21516), *Bacillus stearothermophilus* (B.stea, EMBL accession number M29876) and *Salmonella typhimurium* (S.typh, EMBL accession number X53001) were aligned using the CLUSTAL V sequence alignment package (Higgins *et al.*, 1992). In the figure, "*" denotes identical residues.

```
E.coli          MSTSDDIHNTTATGKCPFHQGGHDQSAGAGTTTRDWWPNQLRVDLLNQHSNRSNPLGEDF
B.stea          M----ENQNRQNAAQCPFHESVTNQSSNR-TTNKDWWPNQLNLSILHQHDRKTNPHDEEF
S.typh          MSTTDDTHNTLSTGKCPFHQGGHDRSAGAGTASRDWWPNQLRVDLLNQHSNRSNPLGEDF
                *    . .*   ...****..  ..*.   *. .******* . .*.**  ..** .*.*

E.coli          DYRKEFSKLDYYG-LKKDLKALLTESQPWWPADWGSYAGLFIRMAWHGAGTYRSIDGRGG
B.stea          NYAEEFQKLDYW-ALKEDLRKLMTESQDWWPADYGHYGPLFIRMAWHSAGTYRIGDGRGG
S.typh          DYRKEFSKLDYYSALKGDLKALLTDSQPWWPADWGSYVGLFIRMAWHGAGTYRSIDGRGG
                .*  ** ****   ** **. *.*.** ***** * *  ********.*****  *****

E.coli          AGRGQQRFAPLNSWPDNVSLDKARRLLWPIKQKYGQKIS------WADLFILAGNVA--L
B.stea          ASTGTQRFAPLNSWPDNANLDKARRCYGRSKRNTGTKSLGPICSFWRAMSLLNRWVEKRL
S.typh          AGRGQQRFAPLNSWPDTVSLDKARRLLWPIKQKYGQKIS------WADLFILAGNVA--L
                *. * ***********  .******     *.. * *       *  . .*   *   *

E.coli          ENSGFRTFGFGAGREDVWEPDLDVNWGDEKAWLTHRHPEALAKAPLG-ATEMGLIYVNPE
B.stea          DSAA------GPLTSGIRKKTFIGDRKKSGSPLNAIPVIASSKTRSPRANGVNLRQPRRA
S.typh          ENSGFRTFGFGAGREDVWEPDLDVNWGDEKAWLTHRHPEALAKAPLG-ATEMDLIYVTPE
                ....      *.   ...   . ..    . *      * .*.    *  . *

E.coli          GPDHSGEPLSAAAAIRATFGNMGMNDEETVALIAGGHTLGKTHGAGPTSNVGPDPEAAPI
B.stea          GRQAGSKSRGISAE---TFRRMGMNDEETVALIAGGHTFGKAHRGGPATHVGPEPEAAPI
S.typh          GPNHSGEPLSAAAAIRATFGNMGMNDEETVALIAGGHTLGKTHGPAAASHVGADPEAAPI
                * . .. . . .*    **  *****************.**.*  .....**..******

E.coli          EEQGLGWASTYGSGVGADAITSGLEVVWTQTPTQWSNYFFENLFKYEWVQTRSPAGAIQF
B.stea          EAQGLGWISSYGKGKGSDTITSGIEGAWTPTPTQWDTSYFDMLFGYDWWLTKSPAGAWQW
S.typh          EAQGLGWASSYGSGVGADAITSGLEVVWTQTPTQWSNYFFENLFKYEWVQTRSPAGAIQF
                * ***** *.** * *.*.****.*  ** *****   .*. ** *.*  *.***** *

E.coli          EAVDAPE--IIPDPFDPSKKRKPTMLVTDLTLRFDPEFEKISRRFLNDPQAFNEAFARAW
B.stea          MAVDPDEKDLAPDAEDPSKKVPTMMMTTDLALRFDPEYEKIARRFHQNPEEFAEAFARAW
S.typh          EAVDAPD--IIPDPFDPSKKRKPTMLVTDLTLRFDPEFEKISRRFLNDPQAFNEAFARAW
                 ***. .  . **. *****    *. ***.******.***.*** ..*. * *******

E.coli          FKLTHRDMGPKSRYIGPEVPKEDLIWQDPLPQPIYNPTEQDIIDLKFAIADSGLSVSELV
B.stea          FKLTHRDMGPKTRYLGPEVPKEDFIWQDPIPEVDYELTEAEIEEIKAKILNSGLTVSELV
S.typh          FKLTHRDMGPKARYIGPEVPKEDLIWQDPLPQPLYQPTQEDIINLKAAIAASGLSISEMV
                ***********.**.********.*****.*.  *. *. .* ..*  *  ***..**.*

E.coli          SVAWASASTFRGGDK-RGGANGARLALMPQRDWDVNAAA--VRALPVL-EKIQKESGKAS
B.stea          KTAWASAA--RSATRISAATNGRRIRLAPQKDWEVNEPERLAKVLSVLRGHPARTAEKSK
S.typh          SVAWASASTFRGGDK-RGGANGARLALAPQRDWDVNAVA--ARVLPVL-EEIQKTTNKAS
                  *****.  *.. .  ...** *. * **.**.**      . *.**     . . *.

E.coli          LADIIVLAGVVGVEKAASAAGLSIHVPF---APGRVDARQDQTDIEMFELLEPIADGFRN
B.stea          HRRLDRLGGTLRWKRQPATPALMSKCHFSLAAAMRHKSKPMSKALPCWNRSQMASATIKS
S.typh          LADIIVLAGVVGIEQAAAAARVSIHVPF---PPGRVDARHDQTDIEMFSLLEPIADGFRN
                   .  *.* .   . .... .     *   .. * ..     .      . .  ...

E.coli          YRARLDVSTTESLLIDKAQQLTLTAPEMTALVGGMRVLGANFDGSKNGVFTDRVGVLSND
B.stea          KSTRFRRKSCSSTKPSSSAD----RPRNDGLSWRFARVGPNYRHLPHGVFTDRIGVLTND
S.typh          YRARLDVSTTESLLIDKAQQLTLTAPEMTVLVGGMRVLGTNFDGSQNGVFTDKPGVLSTD
                  .*.   .  *     . .     *    *      .* *.    .*****. ***. *
```

```
E.coli          FFVNLLDMRYEWKATDESKELFEGRDRETGEVKFTASRADLVFGSNSVLRAVAEVYASSD
B.stea          FFVNLLDMNYEWVPTDSG--IYEIRDRKTGEVRWTATRVDLIFGSNSILRSYAEFYAQDD
S.typh          FFANLLDMRYEWKPTDDANELFEGRDRLTGEVKYTATRADLVFGSNSVLRALAEVYACSD
                ** ***** *** .** .  ..* *** ****. **.* **.*****.**. ** **  *

E.coli          AHEKFVKDFVAAWVKVMNLDRFDLL---------
B.stea          NQEKFVRDFINAWVKVMNADRFDLVKKARESVTA
S.typh          AHEKFVKDFVAAWVKVMNLDRFDLQ---------
                 .****.**. ******* *****
```

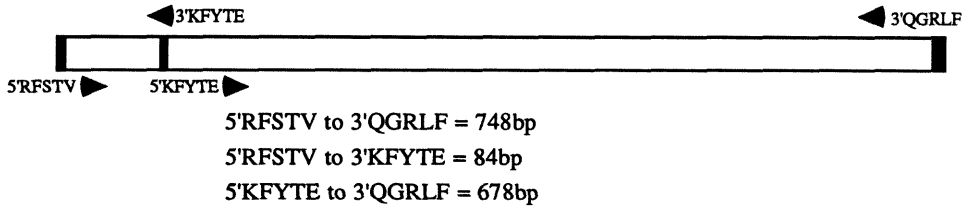
Figure 5.3 Alignment of catalase-peroxidase protein sequences.
Protein sequences of the catalase-peroxidase enzymes from *Escherichia coli* (E.coli, Swissprot accession number P13029), *Bacillus stearothermophilus* (B.stea, Swissprot accession number P14412), and *Salmonella typhimurium* (S.typh, Swissprot accession number P17750) were aligned using the CLUSTAL V sequences alignment package (Higgins *et al.*, 1992). In the figure, "*" denotes identical residues, "·" indicates conservative substitutions.

Table 5.1: Design of oligonucleotide primers.

Region	Motif	Degenerate Primer	Complexity
1	RFSTV	5'CGN/TTY/WSN/ACN/GT3'	256
2	KFYTE	5'AAR/TTY/TAY/ACN/GA3'	32
2'	KFYTE	5'TCN/GTR/TAR/AAY/TT3'	32
3'	QGRLF	5'AA/NAR/NCK/NCC/YTG3'	512

Key: N= A or T or C or G; Y= C or T; W= A or T; S= C or G; R= A or G; K= G or T. Primers for region 1 and 2 are in the sense direction; primers for regions 2' and 3' are in the antisense direction. The single letter amino acid code is used to denote amino acid residues. Complexity is the number of different oligonucleotides in the pool of primers.

The primer sequence for region 2 was made in both the sense and antisense direction, allowing it to act as an upstream primer (5') against the 3'QGRLF primer pool, and as a downstream (3') primer against the 5'RFSTV pool. The degenerate pools of primers were also designed to include the exact nucleotide sequence of regions 1, 2 and 3 from the catalase genes of L. *seeligeri* and *E. coli*. This would act as an internal positive control, since the pool of degenerate oligonucleotides should be able to amplify the *katE* genes of *E. coli* and *L. seeligeri*. On the basis of known catalase gene sequences, the predicted lengths of the PCR amplicons from these primers are as follows:



5'RFSTV to 3'QGRLF = 748bp
5'RFSTV to 3'KFYTE = 84bp
5'KFYTE to 3'QGRLF = 678bp

Figure 5.4 Schematic representation of the amplicons produced by the degenerate oligonucleotide primers. Arrows indicate the direction of extension of the primers by *Taq* polymerase. Sizes of the predicted amplicons are given in base pairs (bp).

5.3 PCR reactions and analysis of amplicons.

5.3.1 Initial amplification reactions

A number of steps were taken to increase the likelihood of amplifying the mycobacterial M-catalase. Firstly, the PCR reaction conditions were optimised to give the maximum number of different amplicons. This was achieved using a "touchdown" thermal-cycling protocol, whereby the initial annealings were performed at the highest melting temperature (Tm) of the primer pair, then the annealing temperature was dropped by two degrees every three cycles to a low annealing temperature (Don *et al.*, 1991). The Tm of an oligonucleotide was calculated empirically using the formula 4(G+C) + 2(A+T) = Tm (Sambrook *et al.*, 1989). This protocol would allow the best primer-template matches to anneal first, selectively amplifying their product in the initial cycles; subsequent cycles would allow the amplification from primer-template hybrids with mismatches. This protocol should also decrease amplicons due to mis-priming, since they are selected against in the early cycles. In addition, the magnesium concentration of the reaction mixes was varied, usually between 1 to 3 mM Mg^{2+}, as this was found to have a major influence on the number of products formed.

As a second step to increase the chances of amplifying the M-catalase, genomic DNA from a wide range of mycobacteria with M-catalase activity was used as the template for the PCR reactions. As even single base mismatches at the 3' end of primers can prevent extension (Sambrook *et al.*, 1989), it was hoped that the combination of a pool of degenerate primers and a range of mycobacterial DNAs would greatly increase the chances of producing the desired amplicon.

Reactions were first performed using the 5'KFYTE and 3'QGRLF pool of primers, since the 5'KFYTE primer pool had lower complexity than the 5'RFSTV pool and hence may have a greater chance of success. Reaction conditions are described in Section 4.1.3. The range of mycobacterial template DNAs used included *Mycobacterium terrae*, *M. gordonae*, *M. avium*, *M. tuberculosis*, *M. phlei*, *M. fortuitum*, *M. smegmatis*, and *M. kansasii*. Reactions using genomic DNA from *L. seeligeri* and *E. coli* were used as positive controls. A typical set of amplicons from these PCR reactions, analysed on a 0.8% (w/v) agarose gel, is shown in Figure 5.5. In general, amplification reactions were repeated three to five times with each source of template DNA for each

set of reaction conditions. With regard to the positive controls, *E. coli* template DNA, but not *L. seeligeri* template, generated an amplicon of the correct size. This failure of the *L. seeligeri* positive control could not be easily explained, and despite numerous attempts no amplicon of the correct size could be produced. Sets of amplicons produced under the standard reaction conditions (Section 4.1.3) from each mycobacterial template DNA proved to be highly reproducible between experiments, with the concentration of template DNA the critical factor in the number of bands produced (e.g. in Figure 5.5, note the decrease in bands in the *M. terrae* amplification reactions when using 100ng [lane 11] rather than 500ng [lane 10] of template DNA). Indeed, the amplicon produced when 100ng of *M. terrae* template DNA was used in the amplification reactions was of approximately the predicted size (i.e. 650bp). This appeared to be the only amplicon produced from mycobacterial template DNA that was of the desired size.

To determine whether an amplicon was indeed a section of the M-catalase gene, two criteria were chosen; (a) the amplicon should be of approximately the predicted size, (b) it should hybridise with other bacterial catalase genes. Therefore, PCR amplicons were Southern blotted onto Hybond-N+ membranes (Amersham) and probed under conditions of low stringency with radiolabelled catalase genes. Two probes were used in the hybridisation experiments, the *E. coli* and *L. seeligeri* typical catalase genes. The entire *L. seeligeri katE* gene was amplified using the primers Liscat5' and Liscat3' that flanked the gene sequence. An internal fragment of the *E. coli katE* gene was used as the other probe. The primers used to amplify the fragment, ECA3' and ECA5', were the *E. coli* gene sequences that code for regions 2 and 3 (KFYTE and QGRLF, respectively). Results of a typical hybridisation experiment are shown in Figure 5.6. The *E. coli katE* probe, but not the *L. seeligeri* probe, hybridised to a number of amplicons. Indeed, despite changing the hybridisation conditions (e.g., lowering the hybridisation temperature, washing for shorter periods), no binding of the *L. seeligeri* probe could be detected. The most notable hybridisation of the *E. coli katE* probe was to an amplicon of approximately 650bp from the *M. terrae* reactions, an amplicon that was also of the desired size; this amplicon therefore fulfilled both criteria for the identification of putative M-catalase sequences and was picked for further analysis.

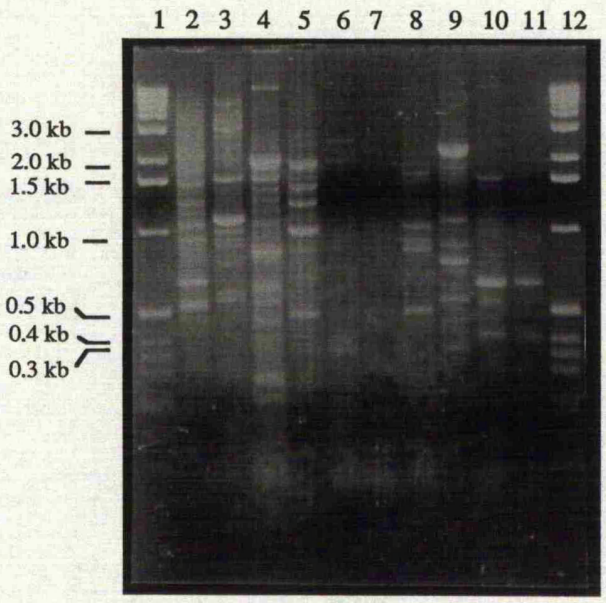


Figure 5.5 Amplification products using the degenerate primer pairs 5'KFYTE and 3'QGRLF.

Amplification reactions were performed using mycobacterial template DNA with the degenerate primer pairs 5'KFYTE and 3'QGRLF. *Escherichia coli* and *L. seeligeri* genomic DNA were used as positive controls. Lane contents are referred to by their template DNA; *E. coli* [2], *L. seeligeri* [3], *M. gordonae* [4], *M. smegmatis* [5], *M. kansasii* [6], *M. avium* [7], *M. phlei* [8], *M. tuberculosis* H37Rv [9], *M. terrae*, 500ng template used in amplification reaction [10], *M. terrae*, 100ng template used in reaction [11]. Lanes 1 and 12 contain 1kb ladder DNA size marker (GIBCO-BRL), with fragment sizes shown in kilobases (kb).

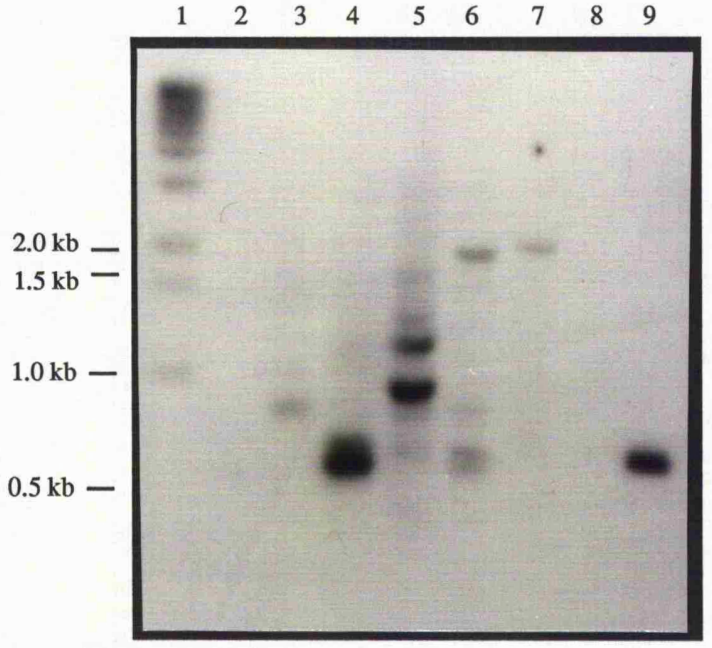


Figure 5.6 Southern blot hybridisation of amplification products.
Amplification products were analysed on a 0.8% (w/v) agarose gel, then blotted to Hybond-N+ nylon membrane. A fragment of the *E. coli katE* gene, amplified by PCR, was labelled with [α-^{32}P]dCTP and used as a hybridisation probe under conditions of low stringency. The membrane was washed twice in 2x SSC, 0.1% SDS at 50°C for 30 minutes. Hybridisation to amplicons from *M. phlei* [3], *M. terrae* [4], *M. smegmatis* [5], *M. gordonae* [6], *M. tuberculosis* [7] is evident. Lane 1 contains 1kb ladder DNA size marker (GIBCO-BRL), with fragment sizes shown in kilobases (kb). Lane 9 is the positive control.

5.3.2 Subcloning and sequencing of the *M. terrae* PCR amplicon.

The 650bp *M. terrae* PCR amplicon was successfully subcloned into pUC18 using the T-vector approach of Marchuk *et al.* (1991). This technique exploits the template independent terminal-transferase activity of *Taq* polymerase, which produces PCR amplicons with a 3' adenosine overhang. Ligation of amplicons to vector molecules that have a complementary thymidine overhang permits the cloning of PCR amplicons. The hybrid pUC18::*M. terrae* PCR amplicon was named pSG1.

DNA sequencing of the *M. terrae* insert in pSG1 was performed using the chain-termination method of Sanger *et al.* (1977). Plasmid DNA was first denatured with a combination of alkali and heat, then primers were annealed and reactions performed using the Sequenase system (see Section 2.6.1.1). Reactions using the forward and reverse sequencing primers produced approximately 200-250 bases of sequence information. An oligonucleotide was synthesised and used in reactions to read between the endpoints of the universal and reverse primer sequence. The whole sequence of the *M. terrae* insert is shown in Figure 5.7. No significant homologies at either the level of nucleic acid or translated amino acid sequence could be found when the sequence was searched against the Swissprot (protein), EMBL (nucleic acid) and GenBank (nucleic acid) databases. When the insert sequences immediately adjacent to the vector sequence were examined, it was found that they both belonged to the 3'QGRLF pool. The 5' end of the insert contains a 12-base oligonucleotide from the 3'QGRLF pool (primer sequence $^{5'}$AAAAGGCGGCCC$^{3'}$), while the 3' end contains a 10-base internal sequence from the same pool (primer sequence $^{3'}$GTTCCGGCGA$^{5'}$). Therefore, the amplicon had been produced from oligonucleotides present only in the 3'QGRLF pool.

5.3.3 PCR Reactions using primers based on Region 3.

The above result is worth considering in greater detail. The degenerate PCR approach depends on the synthesis of an amplicon from primers in the 5'KFYTE pool (acting as "upstream" 5' primers) and the 3'QGRLF pool (acting as "downstream" 3' primers). The *M. terrae* 650bp PCR amplicon was however produced from primers in the same

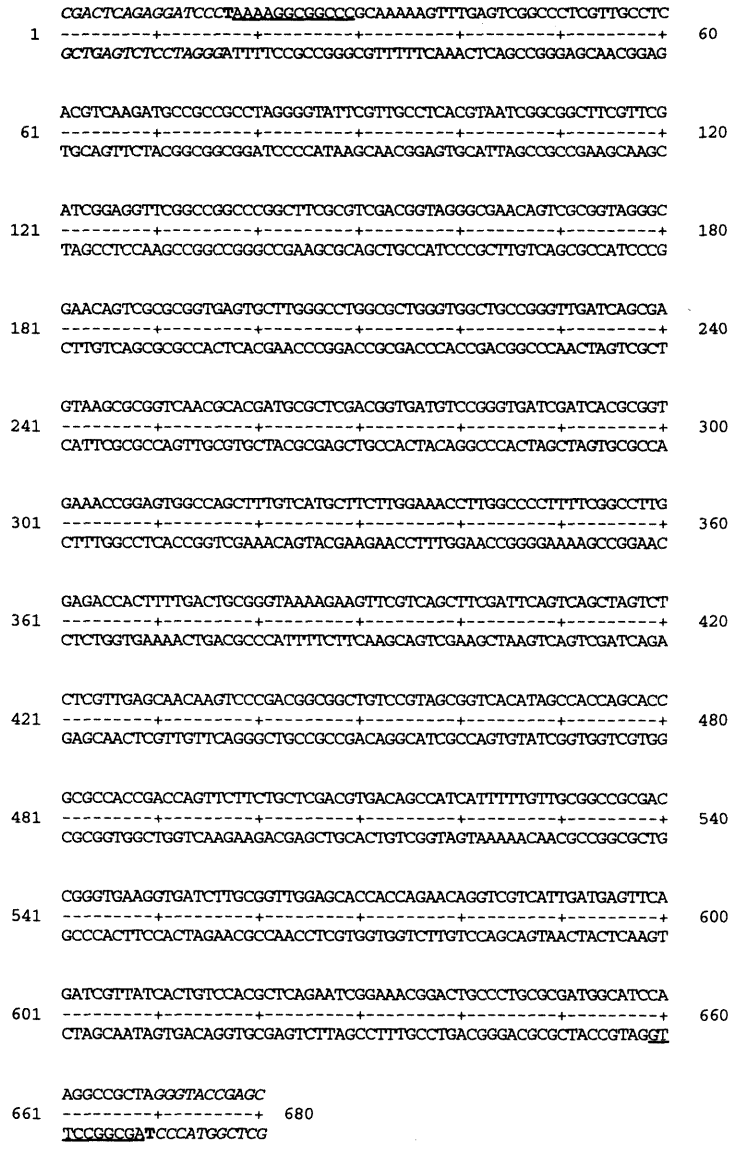


Figure 5.7 Nucleotide sequence of the *M. terrae* 650bp amplicon. Vector sequences are in italics. The T overhangs that were used to subclone the amplicon are in bold type. The primer sequences that generated the amplicon are shown underlined.

pool, producing a nontarget amplicon with no homologies to catalase gene or protein sequences. Hence, the question arose as to whether all of the PCR amplicons were due to mis-priming by the 3'QGRLF degenerate primer pool.

As a control experiment, a set of PCR reactions was performed on *E. coli* genomic template DNA. The two 14-base oligonucleotide primers, ECA5' and ECA3', that code for regions 2 and 3 (KFYTE and QGRLF respectively) in *E. coli* were used as primers with either of the degenerate sets, i.e. ECA5' was used with 3'QGRLF, and ECA3' was used with 5'KFYTE or 5'RFSTV. Since the degenerate primer pools also contained the exact nucleotide sequences that code for the conserved amino acid sequences KFYTE and QGRLF in *E. coli*, these reactions should synthesise amplicons of the correct size. These reactions were performed exactly as with the mycobacterial templates, and the results are shown in Figure 5.8. While reactions using ECA3' and either 5'RFSTV or 5'KFYTE seemed to give the correct size amplicon (approximately 670bp), reactions involving 3'QGRLF with ECA5' gave, amongst others, a shorter amplicon (between 600 and 650bp). From this control it was apparent that the 3'QGRLF degenerate pool had a "dominant" effect at mispriming and producing random amplicons, even in the presence of the correct 5' primer. Indeed, when the 3'QGRLF degenerate pool was the only set of primers added to a PCR reaction with *M. terrae* template DNA, the 650bp amplicon was still produced. Consequently, it was decided to decrease the complexity of the 3'QGRLF pool of primers to reduce the chances of random amplicons being formed. Two new primer sets were synthesised, namely 3'QGRLF(2) and 3'QGRLF(3), that differed in their levels of complexity (Table 5.2).

Table 5.2: New oligonucleotides based on Region 3

Name	Degenerate primer	Size	Complexity
3'QGRLF(2)	GTY/CCV/GCS/RAS/AAG	15mer	48
3'QGRLF(3)	GTC/CCR/GCC/GAS/AA	14mer	4

Key: Y= C or T; V= A or C or G; S= C or G; R= A or G.
The single letter amino acid code is used to denote amino acids. Complexity the number of different oligonucleotides in the pool of primers.

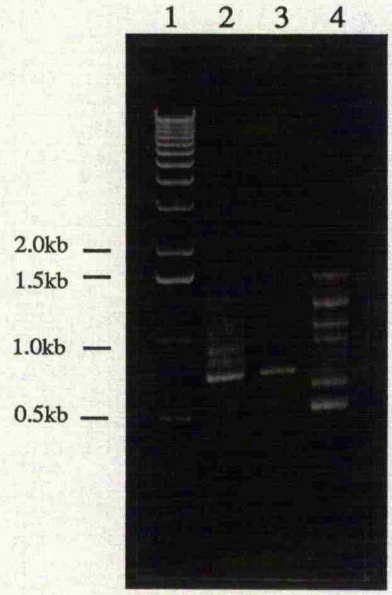


Figure 5.8 Amplicons from *E. coli* template DNA using nondegenerate and degenerate primers.

Lane 1: 1 kb ladder DNA size marker, fragment sizes shown in kilobases (kb)

Lane 2: ECA3' and 5'RFSTV

Lane 3: ECA3' and 5'KFYTE

Lane 4: ECA5' and 3'QGRLF

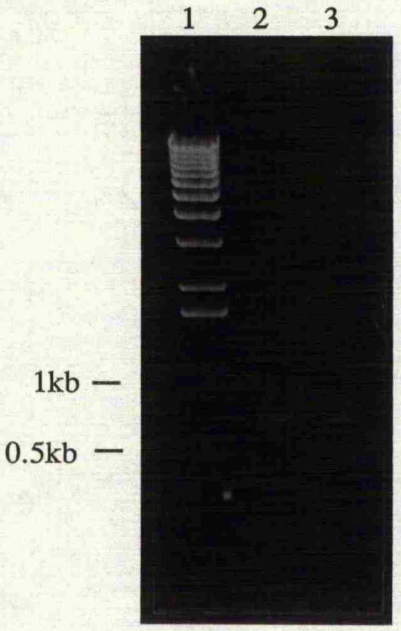


Figure 5.9 Nontarget amplicons generated using the 3'QGRLF degenerate primer.

Amplification reactions were performed using *M. terrae* genomic DNA and the 3'QGRLF degenerate primer. Lanes 2 and 3 contain amplicons produced using 500ng and 250ng of template DNA respectively. Lane 1, 1kb ladder DNA size marker (GIBCO-BRL); sizes shown in kilobases (kb).

When 3'QGRLF(2) was used in PCR reactions with either 5'RFSTV or 5'KFYTE, using *M. terrae* DNA as template, the same 650bp amplicon as before was apparently produced. Indeed, the 650bp amplicon was also produced when only 3'QGRLF(2) was used as the primer (see Figure 5.9). This result was not due to contamination of reaction components with the original 650bp amplicon, as fresh buffer aliquots and dNTP solutions were always used, along with a new preparation of *M. terrae* genomic DNA. In addition, reactions were prepared in a laminar flow cabinet, using a dedicated set of Gilson pipettes with aerosol resistant tips (Northumbria Biologicals Ltd.). Therefore, despite a ten-fold decrease in complexity, and although the primer sequences responsible for synthesis of the sequenced *M. terrae* 650bp PCR amplicon were not present in this new degenerate pool, nontarget amplicons were still produced.

Thermal cycling conditions were also changed in an attempt to produce an amplicon of the desired size. For example, an annealing temperature of 37^{0}C for 30 cycles was tried as this protocol was successfully used by other workers for the amplification of target sequences using degenerate oligonucleotides (J. Ketley, Dept. Genetics, University of Leicester, pers. comm.), but this still led to the production of nontarget amplicons. In addition, a separate set of experiments had shown that an internal fragment of the *E. coli katE* (typical catalase) gene could bind to an 8kb fragment from an *Eco*RI total digest of *M. terrae* DNA (see Section 5.5.1). Sharma and colleagues (1992) had reported that digestion of template genomic DNA with restriction endonucleases can increase the specificity of PCR reactions. Therefore, reactions were set up using total *Eco*RI digests of *M. terrae* DNA with the degenerate primer sets 5'KFYTE, 5'RFSTV and 3'QGRLF(2). However, this approach was unsuccessful, with reactions failing to generate any products.

Other approaches to reduce or remove the nontarget amplicons that were tried included (i) boiling the primer pools then snap-cooling them on ice prior to their addition to the reaction mix to ensure that any primer-dimers or secondary structure in the primer was removed (Kureishi and Bryan, 1992), (ii) adding formamide or DMSO to the reaction mix to permit the complete denaturation of the template DNA (Sarkar *et al.*, 1990), and (iii) making up fresh template DNA and primer stocks in case the solutions being used had somehow become

contaminated. Again, none of these modifications proved successful. It should also be stated here that these modifications to the standard protocol were tried with all the previously synthesised primer sets to see whether they would allow the production of the predicted amplicons; however, no products of the desired size could be generated.

Reactions were then performed using the 3'QGRLF(3) degenerate set of primers. The rationale behind this was that with a complexity of only four, this primer pool shouldn't produce any random amplicons, and may therefore increase the chances of amplifying the desired product. Template DNA was *M. terrae* genomic DNA, while 5'RFSTV and 5'KFYTE were used as the "upstream" pool of primers. Reactions were set up using the primer pairs 5'RFSTV with 3'QGRLF(3), 5'KFYTE with 3'QGRLF, or just 3'QGRLF. Figure 5.10 shows the results of these experiments. Again, despite having a complexity of only four, the 3'QGRLF(3) degenerate primer set was capable of generating a range of amplicons independent of the 5' sets of primers. Despite changing reaction conditions as described above (changing thermal cycling conditions, using digested DNA, adding formamide or DMSO, boiling primers prior to reaction) the nontarget amplicons were still produced. Indeed, the presence of the 5'KFYTE or 5'RFSTV primers appeared to have no effect on the amplicons produced. This result could not be attributed to contamination of the reaction components with previously amplified amplicons. It appeared therefore that some intrinsic property of primers based on region 3 (QGRLF) allowed them to arbitrarily prime throughout the mycobacterial genome, making them useless as primers for degenerate PCR.

5.3.4 PCR reactions using 5'RFSTV and 3'KFYTE

The correct size of the amplicon generated using the degenerate primer pairs 5'RFSTV and 3'KFYTE would be 84bp in length. Again, a range of mycobacterial DNAs were used as templates, with *E. coli* and *L. seeligeri* genomic DNAs acting as the positive controls (since the degenerate primers were synthesised to include the correct *E. coli* and *L. seeligeri* sequences for RFSTV and KFYTE). The standard thermal cycling parameters were altered to take account the smaller reaction product; the extension time was decreased from 90 to 30 seconds.

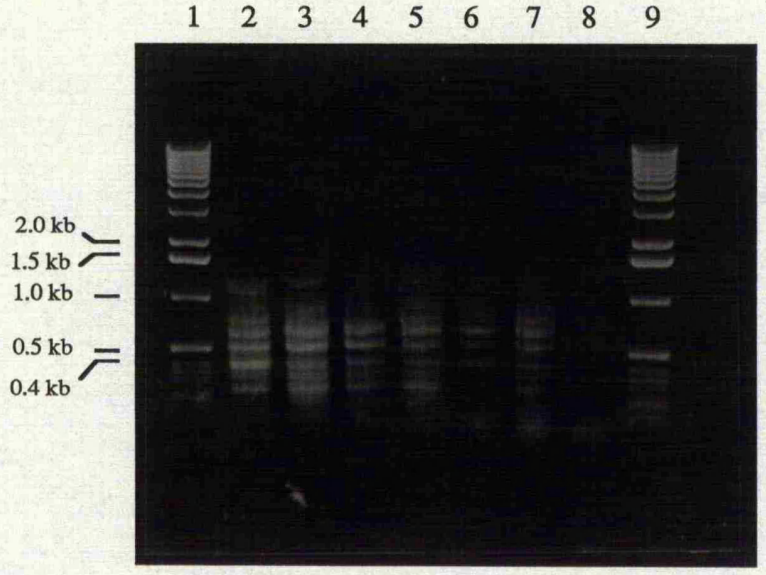


Figure 5.10 Amplification products using the 3'QGRLF(3) degenerate primer.

Mycobacterium terrae genomic DNA (100ng) was the template for amplification reactions. Lanes 2 and 3, 3'QGRLF(3) used as sole primer, reactions performed using 1.5mM or 2.5mM Mg^{2+} respectively. Lanes 4 and 5, 3'QGRLF(3) and 5'KFYTE, reactions using 1.5mM or 2.5mM Mg^{2+} respectively. Lanes 6 and 7, 3'QGRLF(3) and 5'RFSTV, reactions using 1.5mM or 2.5mM Mg^{2+} respectively. Lane 8, negative control (no template added to reactions); lanes 1 and 9, 1kb ladder DNA size marker (GIBCO-BRL), with fragments sizes shown in kilobases (kb).

To visualise the products, reactions were analysed on 3% (w/v) agarose gels (consisting of 2% (w/v) low melting point agarose (Sigma) plus 1% (w/v) "normal" agarose).

No amplicons of the predicted size were synthesised under any conditions. Indeed, even the positive control templates, *E. coli* and *L. seeligeri* genomic DNA, failed to produce amplicons of the correct size. The failure of the positive controls pointed to a technical problem with the either the reaction conditions or the degenerate primers. However, it was felt that investing further effort in this problem was not worthwhile, since (i) even if an amplicon of the desired size was formed, it would be difficult to purify and subclone for sequencing because of its small size, (ii) a lot of time and money had been spent trying to generate the correct size amplicon using the 3'QGRLF degenerate primers, with little success. In addition, a degenerate oligonucleotide, named 5'PROBE, had been synthesised based on region 2. This oligonucleotide "coded" for the amino acid sequence KFYTEEG that is present in all typical bacterial catalase protein sequences elucidated so far. The oligonucleotide had a complexity of four and was designed to include the exact sequence of KFYTEEG coding region from *E. coli*. The 5'PROBE oligonucleotide was end labelled with [γ^{32}P]ATP using T4 polynucleotide kinase, then used to probe *Eco*RI digests of *M. terrae* DNA and *E. coli* DNA that had been blotted to Hybond-N$^+$ nylon membranes. No significant hybridisation of the oligonucleotide to *M. terrae* DNA could be detected, even under conditions of low stringency. The *E. coli* positive control produced a ladder of faint bands between 12kb and 3kb where the probe had bound non-specifically to the digested DNA. This experiment therefore showed that a degenerate oligonucleotide based on the most conserved region of the typical catalase proteins could not identify any homologous sequences in a digest of mycobacterial DNA. This called into question the validity of using the PCR approach to amplify the mycobacterial typical catalase.

Taken as a whole, the results presented above represented a logical, though unsuccessful, approach to amplify the mycobacterial typical catalase gene using an approach based on degenerate oligonucleotide PCR. It was felt that a parallel strategy to isolate the typical catalase gene may prove more fruitful. An approach using the *E. coli* and *L. seeligeri* typical catalase gene sequences as probes was

therefore initiated, using techniques that were already in place (i.e. Southern blotting, probe radiolabelling, hybridisations) rather than embarking in a whole new direction (for example, construction of an expression library and screening for clones that could complement catalase-negative *E. coli* mutants).

5.4 Construction and screening of partial mycobacterial libraries.

5.4.1 Construction of partial libraries of *M. terrae* and *M. gordonae*.

A 670bp radiolabelled fragment of the *E. coli katE* (typical catalase) gene was used as a probe in hybridisations against *Eco*RI digests of mycobacterial genomic DNAs under conditions of low stringency (Section 2.6.3.2). This probe hybridised to single bands in *Eco*RI digests of *M. terrae* and *M. gordonae* genomic DNA in the range of 7-9kb (Figure 5.11). Subsequent attempts to isolate smaller hybridising fragments from double-digests proved to be problematic, with hybridisation signals being much fainter and only visible on autoradiograms after prolonged exposure times. The reason for this reduced hybridisation signal could not be determined, despite repeating the blots many times. It was decided, therefore, to make a partial library of both *M. terrae* and *M. gordonae* based on the original hybridisation result, using size-selected DNA in the range 7-10kb for *M. terrae* and *M. gordonae*.

Mycobacterium terrae and *M. gordonae* genomic DNAs were digested to completion with *Eco*RI and electrophoresed through 0.5% preparative agarose gels. Gel slices containing DNA fragments in the 4-7kb and 7-10kb range were excised, and DNA purified from the agarose using the Pharmacia Bandprep kit. Partial libraries were constructed using the 7-10kb size-selected DNA, as the original binding of radiolabelled *E. coli katE* to *M. terrae* and *M. gordonae* digests was in the 7-9kb range. Size-selected DNA was ligated to *Eco*RI cut, dephosphorylated pBR322 vector DNA, and transformed into *E. coli* DH5α. Low numbers of transformants were obtained, in the region of 20-50 transformants/μg of vector DNA. Since the positive controls

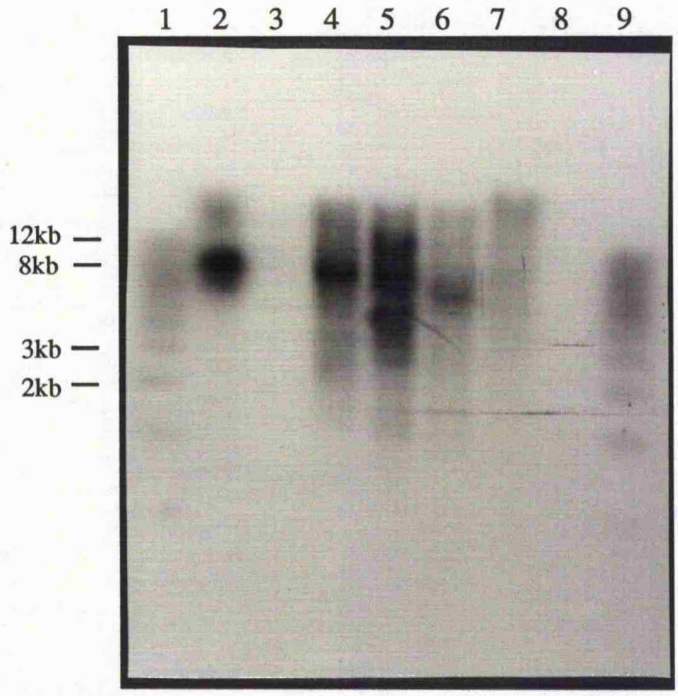


Figure 5.11 Southern hybridisation of mycobacterial chromosomal DNA digests with the *E. coli katE* gene.

Mycobacterial chromosomal DNA was digested with *Eco*RI, analysed on 0.8% (v/v) agarose gels, and blotted to Hybond-N^+ nylon membranes. A fragment of the *E. coli katE* gene, amplified by PCR, was labelled with [α-^{32}P]dCTP and used as a hybridisation probe under conditions of low stringency. An *Eco*RI digest of *E. coli* DNA (Lane 2) is a positive control. Digests of *M. gordonae* (4) and *M. terrae* (6) hybridise at a single position to the *E. coli* probe, while it is harder to discern a specific hybridisation band in the *M. phlei* (7) digest. *M. smegmatis* (3) and *L. seeligeri* (8) show no binding, while a number of *M. tuberculosis* digest fragments (5) appear to bind the probe. Lanes 1 and 9 contain DNA size marker (1kb ladder, GIBCO-BRL), with fragment sizes shown in kilobases (kb).

worked well, these low numbers of transformants may reflect instability of large inserts of G+C rich DNA in plasmid vectors, or the presence of inhibitors in the size selected DNA.

5.4.2 Screening the partial libraries.

The *M. gordonae* partial library was the first to be analysed as this had the most transformants. The 670bp internal PCR fragment of *E. coli katE* catalase was used as the probe. This probe precluded the use of standard colony-blotting techniques to screen the library, since using an *E. coli* gene sequence to screen for low-copy number mycobacterial catalase sequences in *E. coli* hosts would result in high background signals. Instead, plasmid DNA was prepared from transformants, digested with *Eco*RI, and analysed by agarose gel electrophoresis. Insert DNA was present in 80% of transformants tested. Agarose gels were then prepared for direct 'in-gel' hybridisations as described in Section 4.3.2. This technique removed the need for blotting the DNA to nylon membranes, and as such was a considerable saving of time (8-16 hours for gel preparation and capillary transfer reduced to 2-3 hours, plus no requirement to prehybridise before adding radiolabelled probe) and money (no need for nylon membranes). Hybridisations were carried out under the same conditions that allowed hybridisation of the *E. coli katE* gene fragment to genomic digests of *M. terrae* and *M. gordonae* (Section 2.6.3.2)

Thirty-one clones from the *M. gordonae* library were screened, and three clones that hybridised with the *E. coli katE* probe were identified (Figure 5.12). These clones, pAT18, pAT19 and pAT20 contained inserts of approximately 7.5kb, 8kb and 9.5kb respectively. Initially, clone pAT18 was picked for further study as it had the smallest insert size. By performing single- and double-digests on the insert with a range of restriction enzymes and then probing with radiolabelled *E. coli* catalase, it was hoped to isolate a smaller insert fragment that could be sequenced and checked for homology to other catalase gene or protein sequences. A large quantity of pAT18 was prepared and aliquots cut, both as single and double digests, with a range of restriction endonucleases. Digests were analysed by agarose gel electrophoresis, then Southern blotted to nylon membranes. The membranes were probed with the *E. coli katE* gene fragment as before, but no hybridisation signal could be detected. This blot was repeated, with the

same result. The original small scale plasmid preparations of pAT18, pAT19 and pAT20 were analysed again, but this time no binding of *E. coli* catalase could be detected. Again, this blot was repeated, but with the same negative result. A technical reason for the initial hybridisation of the *E. coli katE* gene fragment to these clones, and its subsequent failure to bind, could not be found. The blots were repeated several times, altering the amount of positive control (zero to picogram quantities of the *E. coli katE* probe) on the gel to ensure that the probe was not merely being 'soaked-up' by the positive, changing the hybridisation conditions (on the premise that temperature fluctuations in the oven may have allowed the initial hybridisation to occur), or by blotting the digested clones to nylon membranes and then probing. None of these alterations allowed binding of the *E. coli* catalase gene to the *M. gordonae* sequences. Twenty-one clones from the *M. terrae* partial library were screened using the *E. coli katE* gene fragment as described for the *M. gordonae* library. No clones that hybridised with the *E. coli* typical catalase could be identified.

On the assumption that the original hybridisation result that identified pAT18, pAT19 and pAT20 was correct, these clones were screened for catalase activity. The catalase-negative *E. coli* strain, UM255 (Mulvey *et al.*, 1988), was transformed with the plasmid clones and plated onto selective agar. A 1% (v/v) solution of hydrogen peroxide was then poured over the transformants on the premise that catalase-positive transformants would generate bubbles of oxygen. No catalase-positive transformants were obtained. However, since expression of the mycobacterial catalase gene would have required the catalase promoter to function in *E. coli*, this result may not be surprising.

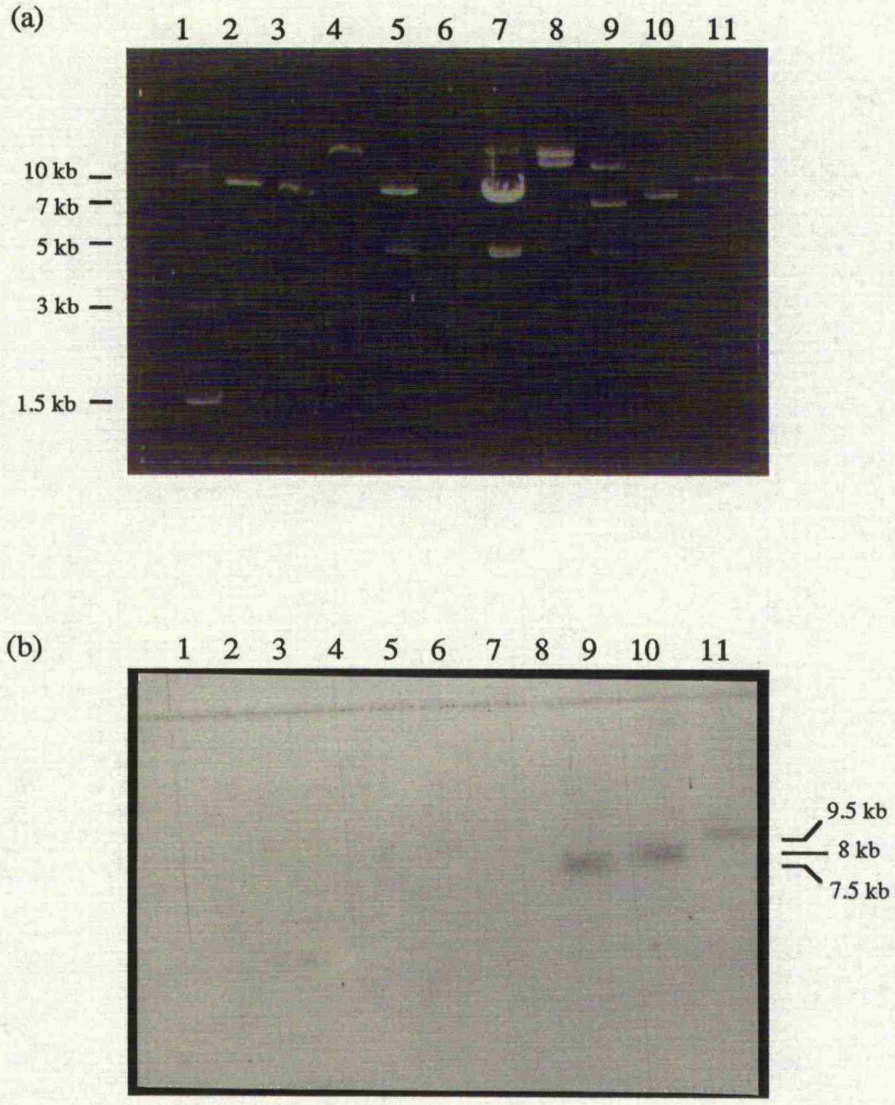


Figure 5.12 Screening the *M.gordonae* partial library.

(a) Plasmid DNA from library clones was prepared, digested with *Eco*RI, and analysed on a 0.8% (w/v) agarose gel. Lane 1, 1kb ladder (GIBCO-BRL). Lanes 2-11, library clones pAT11-20. The gel was treated, dried down, and probed with a radiolabelled internal fragment of the *E. coli katE* gene.

(b) Inserts from clones pAT18 (9), pAT19 (10), and pAT20 (11) hybridised with the *E. coli* probe under conditions of low stringency (hybridisation conditions were 53°C in 5xSSC/0.1%(v/v) SDS overnight; washing was performed twice in 5xSSC/0.1%(v/v) SDS at 53°C for 15 minutes).

Chapter 6.

Discussion

The original project aim was to clone the T- and M-catalase genes from mycobacteria. However, during the initial stages of the project another group cloned the T-catalase gene from *M. tuberculosis* (Zhang *et al.*, 1992). Two strategies to clone the catalase gene were used by this group. In the first instance, a mutant strain of *M. smegmatis* that was highly resistant to isoniazid and showed reduced catalase activity was transformed with a shuttle cosmid library of *M. tuberculosis*. Clones were then screened for sensitivity to isoniazid. The rationale for this was that cosmids that contained a functional T-catalase would render the *M. smegmatis* strain sensitive to isoniazid. Using this approach, a clone was isolated that conferred sensitivity to isoniazid and led to an increase in catalase activity in the *M. smegmatis* strain. The second approach used by this group exploited the conserved nature of the catalase-peroxidase enzymes. The amino-acid sequences of the *E. coli* and *B. stearothermophilus* catalase-peroxidase proteins were aligned to identify any conserved regions, then a nondegenerate oligonucleotide probe was designed to match the amino acid sequence of one such conserved region (amino acids 99-111 in the *E. coli* protein). This oligonucleotide was then used to screen a library of *M. tuberculosis* DNA, and a clone that hybridised with the probe was isolated. The clones isolated by both these strategies were found to contain open reading frames with marked homology to the *E. coli* catalase-peroxidase. The homology to catalase-peroxidases from other organisms, and the ability to confer isoniazid-sensitive phenotype therefore identified the gene as the *M. tuberculosis* T-catalase.

With the T-catalase gene cloned from *M. tuberculosis*, we chose to concentrate on the cloning of the mycobacterial M-catalase gene. The role of M-catalase in isoniazid sensitivity was still an important question, since mycobacteria with both T- and M-catalase show intermediate resistance to isoniazid. Is this due to M-catalase blocking the activation of isoniazid by T-catalase? Or perhaps the T-catalase of mycobacteria with both activities is slightly different to the *M. tuberculosis* T-catalase and unable to activate isoniazid? Cloning the

M-catalase would be a valuable step in finding the answers to these questions.

It should be stated at this stage that a considerable effort had gone into the amplification of the M-catalase using the PCR approach by the time the T-catalase was cloned. Since Zhang and colleagues (1992) had shown that oligonucleotides directed against conserved regions of the catalase-peroxidase could be successfully used to identify homologous sequences in mycobacteria, it was felt that continuing with the PCR approach offered a good chance of success. Along with the PCR approach, a parallel strategy to isolate the M-catalase was also initiated that sought to use previously cloned bacterial typical catalase genes as probes to screen mycobacterial libraries for the M-catalase. This followed the work of Zhang *et al.* (1992) who had used this as one approach to clone the T-catalase.

Alignment of the deduced amino acid sequences of typical catalases from a range of prokaryotes and eukaryotes revealed three conserved regions that appeared suitable as targets for degenerate PCR primers. Designated regions 1, 2 and 3, they coded for the amino acid sequences RFSTV, KFYTE and QGRLF respectively. These regions were conserved in catalase genes from bacteria to humans, so it was reasonable to assume that they would also be present in the mycobacterial M-catalase. During the course of the project, other groups reported the cloning and sequencing of the typical catalase genes of *Lactobacillus skate* (Knauf *et al.*, 1992) and *Bacillus firmus* (unpublished data; EMBL accession number M74194), and these three regions were also conserved in their predicted amino acid sequence. The selection of these regions as sites for degenerate oligonucleotide primers therefore seemed to be a good choice.

Two criteria were used to distinguish the "correct" catalase amplicons from other products; the amplicon should be of approximately the correct molecular size, and it should hybridise to other bacterial catalase genes. However, one product that fulfilled both these criteria, the 650bp *M. terrae* amplicon, was sequenced and found to have no homology to other catalases, either at the level of nucleic acid or protein. This result cast doubt on the method we were using to screen the amplicons produced from the reactions. In addition, the degenerate primers failed to generate the predicted amplicons from the *L. seeligeri* or *E. coli* typical catalase genes, despite the precise *L. seeligeri* and *E.*

coli sequences for regions 1, 2 and 3 being present in the primer pools. This failure of the positive controls pointed to a problem with either the degenerate oligonucleotide primers or the PCR reactions. Nevertheless, despite numerous attempts to get the positive controls working, no amplicons of the correct size were obtained. Hence, we decided to investigate the reasons for the production of nontarget amplicons so that we could block their production and increase the chances of generating the correct amplicon.

Control reactions where the degenerate pools of primers were used in conjunction with the *E. coli* catalase-specific primers ECA5' and ECA3' pointed to a problem with primers based on region 3 (QGRLF). Primers from this pool seemed to have a "dominant' effect on the PCR reactions, mispriming to generate nontarget amplicons, and inhibiting the production of the correct amplicon. Decreasing the degeneracy of primers based on region 3, or altering the reaction conditions, failed to block these unfavourable effects. Since these problems with primers based on region 3 could not be overcome, their use was discontinued.

Amplification of target sequences using degenerate oligonucleotide primers has been successfully used by a number of workers. For example, Dybvig and colleagues (1992) amplified portions of the *recA* genes from a range of Gram-positive bacteria and mycoplasmas using degenerate primers based on conserved regions in *recA* genes. In this work, the RecA amino acid sequences from *Bacillus subtilus*, *Acholeplasma laidlawii* and *E. coli* were aligned to identify conserved regions that would serve as sites for the design of PCR primers. The two primers used had complexities of 192 and 144 respectively, and they amplified fragments of the *recA* genes from *Enterococcus faecalis*, *Streptococcus pyogenes*, *Streptococcus pneumoniae*, *Streptococcus mutans*, *Staphylococcus aureus*, *Mycoplasma pulmonis* and *Lactobacillus lactis*. These results show the power of PCR when using the same set of degenerate oligonucleotides to amplify gene fragments from a wide range of species.

In relation to mycobacteria, PCR using degenerate primers was successfully used by Miller *et al.* (1994) to clone the *M. tuberculosis rpoB* gene that codes for the β subunit of RNA polymerase. Degenerate oligonucleotides were designed from conserved regions of the *E. coli rpoB* gene, with mycobacterial codon usage taken into account to reduce

the complexity of the primer pools. In the resulting primers, one primer had a complexity of 2 while the other primer contained inosine at its only degenerate position to allow a single primer to code for all possibilities (inosine can form hydrogen bonds with all four nucleotides (Sambrook *et al.*, 1989). These primers were then used to amplify regions of the *rpoB* genes from *M. tuberculosis* H37Rv, *M. tuberculosis* H37Ra, *M. bovis* BCG and *M. smegmatis*. The amplicon produced from *M. tuberculosis* H37Rv was used to screen an *M. tuberculosis* H37Rv DNA library for homologous sequences. Using this approach, a clone was isolated from the library and its insert was sequenced. The clone showed considerable homology to the *rpoB* gene from a range of organisms and so was identified as the *M. tuberculosis rpoB* gene.

The two reports described above demonstrate that PCR using degenerate oligonucleotides can be successfully applied to a range of bacteria. Do either of these reports throw any light onto the reasons for the failure of the same approach to amplify the M-catalase gene? To amplify the *M. tuberculosis rpoB* gene Miller *et al.* (1994) used oligonucleotides with a low level of complexity, using both codon usage tables and inosine to reduce the complexity. Perhaps the inclusion of inosine to reduce the complexity of the catalase-primer pools should have been tried. However, as the primer set 3'QGRLF(3) had a complexity of only four and was still capable of producing many nontarget amplicons, the inclusion of inosine may have had little effect. Indeed, considering the success of Dybvig in amplifying *recA* amplicons from a wide range of bacteria despite using primer sets with relatively high complexity (192 and 144 respectively), primer sets with high complexity may have little bearing on the success of the degenerate oligonucleotide PCR approach.

The failure of the PCR strategy required the development of a second approach to isolating the mycobacterial M-catalase. Amplified fragments of the *E. coli* and *L. seeligeri* typical catalase genes had been used to screen the PCR reaction products for amplicons with homology to catalase. A logical step was to investigate whether these catalase probes could hybridise to fragments in genomic digests of mycobacterial DNA. This would then give us a probe to screen mycobacterial libraries for catalase sequences. Restriction endonuclease digests of mycobacterial DNAs were blotted to nylon membranes and than probed with the amplified *E. coli* and *L. seeligeri* catalase genes.

The *E. coli* probe, but not the *L. seeligeri* probe, hybridised to distinct bands in the 7-9kb range in *Eco*RI digests of *M. terrae* and *M. gordonae* genomic DNA. Partial libraries of *M. terrae* and *M. gordonae*, constructed using size-selected DNA in the 7-10kb range, were screened using the *E. coli* probe and three hybridising clones from the *M. gordonae* library were isolated. However, subsequent attempts to repeat this hybridisation result met with failure. Worth noting at this point is recent work by van Ossowski and colleagues (1993) on the evolutionary relationships of prokaryotic and eukaryotic catalase sequences. This work has revealed that the *E. coli* and *L. seeligeri* typical catalases show the least similarity to all other catalases. Their use as hybridisation probes for screening the PCR amplicons and the mycobacterial partial libraries therefore seems an unfortunate choice in retrospect.

An approach to isolating the M-catalase that was not pursued in the course of the project was the construction of an expression library. The rationale in this case would be to construct an expression library from a mycobacterium with M-catalase activity only (e.g. *M. terrae*) to avoid any complications with "false positives" due to T-catalase expression. Library clones could then be screened for phenotypic complementation of a catalase-negative *E. coli* strain such as UM255 (Mulvey *et al.*, 1988). This approach was successfully used to identify catalase genes from *L. sake* (Knauf *et al.*, 1992), *L. seeligeri* (Haas *et al.*, 1991) and *Haemophilus influenzae* (Bishai *et al.*, 1994). For example, in the case of *L. sake* a library was constructed in pBR328 and transformed into the catalase-negative *E. coli* strain UM2. Library clones were then screened by pouring 0.87M hydrogen peroxide over the colonies and immediately picking out with a toothpick any colonies that produced bubbles. Using this method a clone carrying a 3.7kb insert of *L. sake* DNA was isolated that contained the gene for the *L. sake* typical catalase.

This method for isolating catalase genes by complementation of catalase-negative *E. coli* mutants appears simple and straight forward. Perhaps this approach would have offered the best chance for cloning the M-catalase. However, this option was not pursued for two main reasons. First, a considerable amount of effort and expense had been expended on the PCR approach in the first year of the project. It was felt that the experience and techniques gained from this work (such as Southern blotting, nucleic acid hybridisations, etc.) may

be best used in an approach based on the *E. coli* and *L. seeligeri* typical catalase genes as hybridisation probe against genomic libraries of mycobacterial DNA. Second, when the hybridisation approach proved to be unsuccessful, the emphasis of the project turned away from the mycobacterial catalase genes and towards a new project that dealt with the response of mycobacteria to acid stress. This meant that the time to pursue the expression library approach was simply not available.

If time had allowed us, the following experiment into the role of M-catalase in isoniazid resistance would also have been performed. Some workers suspect that mycobacteria with both T- and M-catalase activities owe their reduced isoniazid sensitivity not to the action of M-catalase, but rather to the inability of their T-catalase enzyme to "activate" isoniazid (S.T. Cole, Institute Pasteur, Paris; pers. comm.). The obvious test of this hypothesis would be to inactivate the M-catalase gene from a species with both T- and M-catalase activities that has a low sensitivity to isoniazid (such as *M. smegmatis*). If the resulting mutant had the same minimum inhibitory concentration (MIC) for isoniazid as the wild type, the hypothesis would be proved; i.e. these strains cannot activate isoniazid. To specifically mutate just the M-catalase, the method of choice would be to re-introduce the cloned M-catalase gene using a suicide vector, and select for recombination events that produced M-catalase "knock-out" mutants. However, this option is unavailable at the present since no M-catalase gene has been cloned. Therefore, as an alternative approach to elucidating the role of M-catalase, we proposed to introduce the gene for a typical catalase (i.e. from *E. coli* or *L. seeligeri*) into an isoniazid-senstitive, T-catalase positive mycobacterial strain (i.e. *M. bovis* BCG). The MIC for isoniazid of the transformed *M. bovis* BCG would then be determined; an increase in isoniazid-resistance would suggest that typical catalase activity (i.e. M-catalase) does play an active role in isoniazid-resistance. This experiment would be relatively easy to perform. For example, the typical catalase (HPII) from *E. coli* could be amplified using PCR, then cloned into the *E. coli-Mycobacterium* shuttle vector pMV261, such that the gene was expressed from the *M. bovis* BCG *hsp60* promoter contained in the vector (Stover *et al.*, 1991). The construct would then be transformed into *M. bovis* BCG, selecting for isoniazid-resistant transformants (since if the typical catalase blocked isoniazid activation by the T-catalase, an isoniazid-resistant phenotype would result). Such

an experiment may reveal a lot about the role of M-catalase in isoniazid sensitivity, and is still worth performing.

In summary, both the degenerate oligonucleotide PCR strategy, and the approach based on previously cloned typical catalase genes as probes, failed to isolate the gene coding for the mycobacterial M-catalase gene. The approach that holds the greatest promise for cloning the M-catalase appears to be one based on complementation of catalase-negative *E. coli* mutants. A future project may pursue this goal.

Chapter 7.

Introduction

7.1 Environmental acidic pH as a bacterial stress

7.1.1 Bacterial response to environmental pH alterations.

A dynamic interaction exists between the bacterial cell and its external environment. Alterations in environmental conditions are offset by physiological and genetic readjustments of the cell, allowing it to survive or exploit new conditions. Such adaptive responses have been described for a variety of environmental stresses, such as nutrient limitation, osmotic shock and oxidative stress (Roszak and Colwell, 1987; Farr and Kogama, 1991; Csonka and Hanson, 1991). Analysis of these adaptive systems provides us with an understanding of how bacteria sense and react to their surroundings.

The ability to regulate cytoplasmic pH (pH_i) in response to alterations in external pH (pH_o) is essential for bacterial viability (Booth, 1985; Padan and Schuldiner, 1986). Wide fluctuations in pH_i would have profound effects on such basic cellular processes as ATP generation and enzyme function, and so must be avoided. To maintain the pH_i within a range suitable for viability, the cell employs a range of pH homeostasis mechanisms. These mechanisms provide both constitutive and inducible levels of protection from shifts in the external pH.

Constitutive systems, such as the natural buffering capacity of the cytoplasm and proton transport mechanisms, can be seen as the "day to day" mediators of pH homeostasis (Booth, 1985). However, in times of extreme acid-stress these systems may be overwhelmed, causing the acidification of the cytoplasm and death of the cell. A second, inducible level of acid resistance has been identified that enhances survival at these toxic acid conditions. This "acid-tolerance response" (ATR) is a global response, induced by exposure to mildly acidic pH, and has been identified in a number of organisms such as *Salmonella typhimurium* (Foster and Hall, 1990), *Escherichia coli* (Goodson and Rowbury, 1989), *Listeria monocytogenes* (Kroll and Patchett, 1992), *Streptococcus mutans* and *Enterococcus hirae* (Belli and Marquis,

1991), *Aeromonas hydrophila* (Karem *et al.*, 1994), and root nodule bacteria (O'Hara and Glenn, 1994). The ATR of *S. typhimurium* has been studied in some detail by Foster and colleagues, and will be described in greater detail in Section 7.2 to illustrate the principal aspects of the ATR model.

7.1.2 Elements of bacterial pH homeostasis.

Regulation of pH_i is achieved through three main mechanisms; cytoplasmic buffering, proton transport, and production of acids and bases (Booth, 1985). These mechanisms shall be investigated in turn, with particular note to their roles in maintaining pH_i when the pH_0 is acidic.

The role that cytoplasmic buffering plays in the maintenance of pH_i has been the source of some debate. Zychlinsky and Matin (1983) investigated the capacity of the acidophile *Thiobacillus acidophilus* to maintain pH_i when cells were exposed to ionophores or respiratory chain inhibitors, conditions that cause an influx of protons. Despite a large influx of protons, the cytoplasmic pH remained relatively stable, suggesting that the cytoplasm could buffer against the proton influx. Further work (Goulbourne *et al.*, 1986) with the acidophilic bacterium PW2 established that these bacteria can also maintain a large pH gradient without the action of proton pumps, supporting the earlier conclusions of Zychlinsky and Matin (1983). It could be argued, therefore, that a large cytoplasmic buffering capacity allows an organism to grow over a wide pH range. However, work by Krulwich *et al.* (1985) suggests that this is not always the case. In this work, the cytoplasmic buffering capacity of a range of bacteria was measured. The obligate alkalophiles, *Bacillus firmus* and *Bacillus alcalophilus*, had much higher buffering capacities at high external pH than any of the other bacteria tested. Nevertheless, this buffering capacity was not sufficient to maintain pH_i when the Na^+/H^+ antiporter, the major system for internal acidification, is inactivated. Also, *Bacillus subtilus* had a much greater buffering capacity in the acidic pH range than *E. coli*, despite both being neutralophiles. Therefore, it appears that cytoplasmic buffering capacity may not always be a good indicator of pH_i maintenance capability.

The second major mechanism for pH_i regulation is the movement of protons across the membrane. Three of the major proton

transport systems in the bacterial cell are (i) electron transport chains, (ii) H^+-ATPases, and (iii) antiport systems; these systems shall be described in turn. In the process of oxidative phosphorylation, terminal electron carriers, such as NADH, are oxidised and their electrons transferred between carriers of increasing reduction potential. Movement of electrons down the electron transport chain is coupled to the translocation of protons across the cytoplasmic membrane, generating a pH gradient and electrical potential across the membrane. This potential energy is utilised as a proton motive force; protons flow into the cytoplasm through membrane protein complexes named F_0F_1 ATPases, catalysing the synthesis of ATP (Moat and Foster, 1988). Alterations in factors such as external or internal pH will therefore have important implications for the proton motive force. Indeed, Hicks *et al.* (1991) have shown that different levels of cytochromes are expressed in the facultative alkalophile *Bacillus firmus* in response to the external pH. For example, under alkaline growth conditions the expression of cytochromes caa_3 and *o*, but particularly cytochrome caa_3, is higher than at external pHs approaching neutrality. This may be an attempt to pump more protons out of the cell, as the increased external pH will translate to a decrease in the proton motive force. However, it should be noted that the cycling of protons across the cytoplasmic membrane, which in effect is what is happening during oxidative phophorylation, would have little effect on the proton gradient.

The proton-translocating ATPases of streptococci and related bacteria are essential for regulation of pH_i (Kobayashi *et al.*, 1984; Kobayashi, 1985; Sturr and Marquis, 1992). The H^+-ATPases of these organisms do not function to synthesise ATP from a coupled electron transport chain, but rather act to remove protons from the internal milieu. Hence the expression of the H^+-ATPase of *Streptococcus faecalis* (later renamed *Enterococcus hirae*) was found to increase in response to a decrease in pH_i, triggering alkalisation of the cytoplasm (Kobayashi *et al.*, 1984). Indeed, work by Kobayashi (1985) has suggested that *S. faecalis* relies entirely on the activity of the H^+-ATPase to regulate pH_i by alkalisation, and that it lacks any system for internal acidification. Mutants that lacked H^+-ATPase possessed a pH_i 0.4-0.5 pH units lower than the pH_o when exposed to a pH range from 5.5 to 9.0. Over the same pH_o range, the pH_i of the wild-type was always higher than the mutant at $pH_o < 8.0$, but was the same as the mutant at a pH_o

>8.0. The wild-type sought to maintain its pH_i between 7.7-8.0, but as the medium became more alkaline it lost its ability to do so. No dedicated system for cytoplasmic acidification appeared to be present. Hence Kobayashi (1985) concluded that pH_i in *S. faecalis* is regulated only by cytoplasmic alkalisation through the action of the H^+-ATPase.

Investigations by Marquis and colleagues also suggest a central role for membrane H^+-ATPases in pH_i regulation in oral bacteria. For example *Lactobacillus casei*, which is more acid tolerant than *Actinomyces viscous*, has 3.29 U of ATPase/mg of protein compared to 0.06 U ATPase activity/mg of protein in *A. viscous* (Bender and Marquis, 1987). The ATPase of *L. casei* also has a lower pH optimum (pH 5) than the *A. viscous* ATPase (pH 7). Further work by Sturr and Marquis (1992) developed this point, establishing that bacteria that are more acid tolerant have H^+-ATPases with lower pH optima for activity. Despite the adaptations of these bacterial H^+-ATPases to function as pH_i regulators, they are typical F_0F_1-ATPases. This is reflected both in their ability to synthesise ATP if a proton motive force is artificially imposed (van der Drift *et al.*, 1978), and in the gene structure of the *E. hirae* H^+-ATPase operon (Shibata *et al.*, 1992).

Na^+/H^+ and K^+/H^+ antiporters have been proposed as central mechanisms in the maintenance of bacterial pH_i (Booth, 1985; Padan and Schulinder, 1986; Krulwich and Guffanti, 1986). In the acidic pH_o range, K^+ transport systems sustain an alkaline pH_i by coupling K^+ uptake with H^+ extrusion (Kroll and Booth, 1981; Padan and Schulinder, 1986). In *E. coli*, K^+ uptake is the responsibility of the constitutive, low affinity Trk system and the high affinity, inducible Kdp system (Booth, 1985). The central point of K^+ uptake is that when linked to H^+ efflux it allows a positive charge to enter that is not H^+, hence making the system electroneutral (Kroll and Booth, 1981). Nakamura *et al.* (1988) studied the role of K^+/H^+ antiport in the regulation of cytoplasmic pH in *Vibrio alginolyticus*. They found that with an acidic external pH, the addition of KCl to the medium was necessary for alkalisation of the cytoplasm to a steady value of 7.8 (entry of K^+ would allow H^+ to be expelled via the proton pumps, hence causing alkalisation of the cytoplasm). Worth noting is the role they described for the K^+/H^+ antiporter in intracellular acidification. Cells maintained at pH 8.6, 9.0 and 9.6 require a ΔpK^+ gradient directed

outwards before internal acidification can occur. Hence, the outwardly directed K^+ movement generates an internal acidic pH by H^+ influx.

Similarly, when bacteria are exposed to alkaline pH_o conditions, the action of the Na^+/H^+ antiporter in exchanging intracellular Na^+ for extracellular H^+ permits acidification of the cytoplasm (Padan and Schuldiner, 1992). Na^+/H^+ antiporter activity is also involved in resistance to high Na^+ and Li^+ concentrations (Padan and Schuldiner, 1992). Two Na^+/H^+ antiporters are known to exist in *E. coli* and are termed *nhaA* and *nhaB* (Goldberg *et al.*, 1987; Padan *et al.*, 1989; Pinner *et al.*, 1992). Control of the antiporter encoded by *nhaA* occurs at both the protein and gene level. Analysis of the purified NhaA protein has revealed that its activity in proteoliposomes is increased 2000-fold over a pH_o range of 6.5 to 8.5 (Taglicht *et al.*, 1991). Structural analysis of NhaA has revealed that the histidine residue at position 226 is important for the pH sensing activity of the protein (Gerchman *et al.*, 1993). Histidine-226 was mutated to arginine (H226R) by site directed mutagenesis and the activity of the protein examined. While activation of the wild-type NhaA occurs maximally between pH_o 7 and 8, remaining almost fully active up to pH_o 8.5, the H226R version shows peak activity at pH_o 7.5 and shows only 10-20% activity at pH_o 8.5. These results suggest that histidine-226, while not essential for NhaA activity, is involved in the protein's ability to sense shifts in pH. Karpel *et al.* (1991) constructed fusions of *nhaA* with *lacZ* (*nhaA'-lacZ'*) and found that a change in the pH_o from 6.5 to 8.5 does not by itself increase the expression of the *nhaA* gene. Instead, alterations in pH_o increase the sensitivity of the *nhaA* expression system to changes in extracellular Na^+ concentration. For example, exposure of cells at pH 7.5 to NaCl concentrations from 60-100mM increased the expression of the β-galactosidase reporter by 5-10 fold. However, at pH 8.6, 10mM NaCl could induce the same level of expression of β-galactosidase from *nhaA'-lacZ'*.

The third major system for pH_i regulation is the production of acids and bases. The reactions of degradative amino acid decarboxylases and deaminases produce basic and acidic end products respectively, so these enzymes may play important roles in pH_i regulation. The genes for arginine decarboxylase (*adi*) and lysine decarboxylase (*cadA*) genes are known to be induced in *E. coli* in response to a decrease in the external pH in the presence of their

respective substrates (Auger *et al.*, 1989). Analysis of the lysine decarboxylase *cadA* gene has revealed that is part of an operon with *cadB* whose expression is controlled from the P*cad* promoter (Watson *et al.*, 1992). The P*cad* promoter is under the control of the CadC protein, a positive regulator of transcription whose amino terminus shows homology to the DNA-binding carboxy terminus of proteins such as OmpR and ToxR (Watson *et al.*, 1992). This suggested a model for CadC induction of P*cad*, where the carboxy terminus of CadC would sense alterations in the periplasmic pH while the amino terminus would bind to the P*cad* promoter to induce expression. Meng and Bennet (1992) have also put forward a model for the role of the *cadBA* operon in pH_i regulation. This proposes CadB as a putative lysine-cadaverine antiporter, importing excess lysine and H^+ from the exterior milieu. Once inside, lysine decarboxylase then converts the lysine and protons to cadaverine and CO_2. The cadaverine is then expelled through the CadB antiporter in exchange for lysine and protons, completing the cycle. Therefore, the combined models of Watson *et al.* (1992) and Meng and Bennet (1992) suggest that CadC senses a decrease in the external pH, inducing the expression of the *cadBA* operon that attempts to increase external pH via proton consumption through the action of lysine decarboxylase.

However, Takayama *et al.* (1994) have suggested that this model is flawed, in that it doesn't account for the induction of *cadBA* expression under O_2-limiting conditions. Expression of the *cad* operon was monitored using *cadB-lacZ* fusions, and although expression was induced by a decrease in culture pH (peaking when pH_o 5.4), maximum expression was observed at pH 6.3 under O_2-limiting conditions. Expression was found to be repressed in the presence of carbonate and by the end-products of ornithine decarboxylase reactions (i.e. CO_2). Since *E. coli* is able to grow between pH_o values of 6-8, they suggested expression of the *cad* operon at pH 6.3 would be unnecessary if its major function was to increase pH_o. These observations led Takayama *et al.* (1994) to propose that the primary function of *cadBA* is as a supplier of CO_2 to the cell. Under O_2-limiting conditions, they suggest that the cellular levels of CO_2 are decreased due to a slow down in the Krebs cycle; this would trigger induction of CO_2 supply systems. However, no method for CO_2 regulation of the operon has been proposed by these

workers. These observations suggest that the role of decarboxylases at low extracellular pHs may need to be re-examined.

This section has sought to introduce the various components of pH homeostasis in bacteria. These components operate at both constitutive (i.e. no requirement for *de novo* protein synthesis) and inducible levels to provide protection to the organism. However, whether expression of these components is regulated as a global response to pH stress has not been addressed. Section 7.2 will summarise our current understanding of the global acid tolerance response of *S. typhimurium* to acid stress, the ATR, as this system has been characterised in the most detail.

7.2 Acid tolerance in *S. typhimurium*.

Salmonella typhimurium is a facultative intracellular enteric pathogen that can encounter a range of low pH environments. In the host, the organism is exposed to acidic conditions in the stomach and the macrophage phagosome (Mims, 1987; Garcia-del Portillo *et al.*, 1992). During its free-living phase, the bacteria are aquatic organisms and as such will be exposed to fluctuations in external pH (Roszak and Colwell, 1987). The ability to tolerate exposure to acidic pH is therefore central to the basic biology of *S. typhimurium*.

7.2.1 The acid tolerance response (ATR) model.

The current model for inducible acid resistance in *S. typhimurium* consists of three mechanisms that can be separated into *rpoS*-independent and *rpoS*-dependent responses (Lee *et al.*, 1994). The *rpoS* gene encodes a sigma factor that regulates the expression of a number of growth-phase dependent genes (Matin, 1991).

Examining *rpoS*-independent acid tolerance first, the responses are divided into the log-phase ATR and the stationary-phase ATR (Lee *et al.*, 1994). The log-phase ATR was first described by Foster and Hall (Foster and Hall, 1990; 1991) who established that *S. typhimurium* exposed to a sub-lethal acid stress (adapted cells) could survive exposure to a potentially lethal acid stress better than unadapted cells. The adaptation method involved growing cells at a slightly alkaline pH (pH 7.6), then shifting the culture pH to pH 5.8 (adaptive pH) and allowing for one doubling of the bacterial population before

exposure to the lethal acid stress of pH 3.3. The adaptive pH was chosen on the basis of experiments that showed (i) *Salmonella* genes that are responsive to external acid pH are induced at pH 5.5-6, (ii) *S. typhimurium* tolerates exposure to pH 5.8, with only a slightly slower growth rate than bacteria grown at pH 7.4 (Foster and Hall, 1990). The presence of chloramphenicol, an inhibitor of protein synthesis, in the adapting culture prevented the production of a protective ATR, indicating that protein synthesis is necessary (Foster and Hall, 1991). Indeed, two-dimensional PAGE analysis of cellular proteins labelled with [^{35}S] methionine demonstrated that the expression of 18 proteins was different in adapted cells compared with unadapted cells, 6 being repressed and 12 induced.

The original model of the log-phase ATR was later refined by Foster (Foster, 1991; 1993). The new model divided adaptation into two stages, pre-acid shock and post-acid shock. Pre-acid shock (preshock) entails adapting the cells at pH 5.8 to permit expression of an inducible pH homeostasis mechanism before exposure to the lethal pH. Post-acid shock (postshock) refers to an adaptation step at pH 4.3 that triggers the expression of a set of acid shock proteins (ASPs). Expression of these proteins is regulated in response to a pH_o shift from 7.5 to 4.3, and they comprise approximately 52 proteins as identified by two-dimensional PAGE (Foster, 1991). Apart from 4 shared proteins, preshock and postshock proteins are distinct. Foster established that *S. typhimurium* cells grown originally at pH 7.5, then shifted to pH 4.3 for 15-20 minutes adaptation were better able to survive exposure to pH 3.3 than unadapted cells (Foster, 1993). Increasing the exposure time for postshock adaptation beyond 15-20 minutes caused cells to lose their pH 3.3 tolerance, demonstrating that transient expression of these ASPs is necessary for postshock adaptation (Foster, 1993). An important aspect of the model is that both preshock and postshock stages of adaptation are not essential for an ATR; either step will confer some level of protection against the lethal pH. However, protection is higher and lasts longer if cells undergo both stages of the ATR before exposure to the lethal pH (Foster, 1993). The key point in relation to preshock adaptation is that it allows the cells to synthesise protective ASPs while exposed to normally lethal pH_o values by maintaining the pH_i near pH 6. Protein synthesis at the lethal pH is therefore necessary for preshocked cells to survive (Foster, 1991). Hence, in a batch of *S. typhimurium*

cells grown at pH 7.7, with samples removed and exposed to adaptive and lethal pHs, the following responses occur: (i) direct exposure to pH 3.3 kills the cells [pH_i = 4.5]; (ii) exposure to pH 4.3 for 15-20 minutes before pH 3.3 allows survival of the cells [pH_i at pH_o 3.3 = 4.7]; (iii) exposure to pH 5.8 for one doubling before pH 3.3 allows survival [pH_i at pH_o 3.3 = 5.1]; (iv) exposure of cells to pH 5.8 for one doubling, then to pH 4.3 for 15-20 minutes, permits survival at pH 3.3 [pH_i at pH_o 3.3 = 5.1] (Foster, 1993).

Therefore, a central feature of log-phase ATR in *S. typhimurium* is the synthesis of protective ASPs. In preshock adapted cells, these proteins are synthesised at pH 3.0; in postshock adapted cells they are synthesised at pH 4.3. Thus, preshock adapted cells in effect undergo postshock adaptation at lower external pH conditions.

Stationary-phase ATR is the second low-pH inducible ATR operating in *S. typhimurium* (Lee *et al.*, 1994). Cells taken from stationary-phase cultures at pH 4.3 were found to survive a 4 hour exposure to pH 3.0 at least a 1000-fold better than stationary-phase cells from pH 7.3 cultures. In contrast to the log-phase postshock ATR, where the adaptive pH is also 4.3, long-term rather than transient synthesis of ASPs was necessary for the development of an ATR. Cells from stationary-phase cultures at pH 8.0 were exposed to the adaptive pH 4.3 for various time periods, then challenged at pH 3.0. Resistance to the acid challenge increased in direct proportion to the time at pH 4.3, with maximum protection produced after 2 hours (Lee *et al.*, 1994). A second difference between log- and stationary-phase ATR is that the latter requires continual protein synthesis at the lethal pH to mount a maximal protective response. Addition of chloramphenicol to adapted stationary-phase cells at pH 3.0 caused a 10- to 50-fold decrease in the level of protection (Lee *et al.*, 1994). This contrasts with the post-acid shock component of log-phase ATR (adaptive pH is also 4.3) where cells do not require protein synthesis at the lethal pH (Foster, 1991). Foster suggests that this requirement for continual protein synthesis in stationary-phase ATR may reflect the increased level of protein turnover in stationary-phase cells (Lee *et al.*, 1994).

Bacteria in the stationary-phase of growth are more resistant to a number of stress conditions, such as starvation and low external pH (Gauthier and Clement, 1994; Matin, 1991). This general-stress resistance is mediated by the *rpoS* gene (formerly *katF*) that

encodes a sigma factor (Matin, 1991; Siegele and Kolter, 1992). RpoS co-ordinates the expression of a number of stationary-phase regulated genes, such as the *E. coli* catalase gene *katE* and the gene for exonuclease III (Lange and Hengge-Aronis, 1991; Sak *et al.*, 1989). Lee *et al.* (1994) demonstrated that stationary-phase cells of *S. typhimurium* were more acid tolerant than log-phase cells, and that this mechanism was *rpoS* dependent. This appears similar to the situation in *Shigella flexneri*, where Gorden and Small noted that acid resistance does not develop until the late exponential phase (Gorden and Small, 1993). This "general" acid tolerance displayed by stationary phase cells is distinct from the pH-dependent, *rpoS*-independent stationary-phase ATR described by Lee *et al.* (1994). Nevertheless, it is probable that these general and specific systems work in tandem in *S. typhimurium* to protect against low pH_o during stationary-phase growth.

7.2.2 Genetic loci involved in acid tolerance.

A number of genetic loci have been identified in *S. typhimurium* whose products contribute to the *rpoS*-independent ATR. Since a central feature of the log-phase ATR is the synthesis of ASPs, the model devised by Foster and colleagues predicts two classes of *atr* mutants; those that are defective in preshock adaptation and those defective in postshock adaptation (Foster, 1993). For example, the *unc* operon (subsequently renamed *atp*) codes for the Mg^{2+} dependent F_0F_1 ATPase, a proton translocating mechanism. Foster and Hall (1990) determined that *unc*::Tn*10* mutants were unable to mount an observable log-phase preshock ATR, and were 100- to 1000-fold more acid-sensitive compared to the wild-type. However, further work established that *atp* mutants could mount a successful postshock ATR (Foster, 1993). Therefore, *atp* mutants have a defect in preshock ATR, preventing ASP synthesis at pH 3.3; synthesis of ASPs at pH 4.4 compensates for this defect.

The *fur* gene codes for a regulatory protein that binds available Fe^{2+} and represses the expression of iron uptake genes (Bagg and Neilands, 1987). Mutants deficient in Fur activity are defective in both pre- and post-acid shock ATR, with mutants failing to produce six of the transient ASPs that are necessary for log-phase ATR (Foster, 1993). Although the precise role of Fur in the ATR is unknown, its influence on the expression of a number of pH-regulated genes may

provide a clue to its function (Foster and Hall, 1991). The isolation of another two mutations (*atrD* and *atrF*) that effect iron metabolism and that are defective in pre- and post-shock adaptation hints to a central role for iron in the ATR (Foster and Bearson, 1994).

A pool of Mu*dJ* insertion mutants was screened using a strategy that relied on sensitivity to the protonophore dinitrophenol (DNP) (Foster and Bearson, 1994). The strategy was to plate the mutants onto solid medium containing DNP; mutants that couldn't grow on the DNP medium were thus identified. DNP-sensitive mutants were then tested for a log-phase ATR. From a total of 25,000 mutants screened, seven mutants were isolated that showed a defective ATR on further characterisation. Two of these, *atrD* and *atrF*, were described above. Of the remainder, *atrB* and *atrC* were unable to induce preshock ATR but were capable of a postshock adaptive phase, while *atrG* was incapable of either response. Further characterisation of *atrC* revealed that its map position on the *S. typhimurium* chromosome was in the region of the *polA* gene that codes for DNA polymerase I. Complementation of the *atrC* mutants with fragments of *polA* established that the *atrC* locus was indeed *polA*, and that complementation relied on the Klenow fragment being active. Since one of the major effects of low pH_i will be DNA damage, it is not surprising that the gene for a DNA polymerase is induced by acid stress. The *atrB* mutant is interesting because transposon mapping of the *atrB* locus revealed a gene, *atbR* that negatively regulates *atrB* expression (Foster and Bearson, 1994). Transposon insertions into the *atbR* locus in *atrB* mutants caused over-expression of the *atrB-lac* fusion (produced from the Mu*dJ* insertion), and created a constitutive acid tolerant phenotype. Mutants in the AtbR regulon can therefore produce both acid-sensitive (*atrB*) and acid-resistant (*atbR*) phenotypes. *atbR* mutants also restore acid-tolerance to *polA* and *fur* mutants, indicating that *atbR* controls several genes that contribute to acid resistance.

7.3 Acid tolerance in other bacteria:

7.3.1 *Escherichia coli.*

The adaptive acid tolerance response of *E. coli* is referred to as habituation (Goodson and Rowbury, 1989). Habituation bears

many of the standard features of the ATR in *S. typhimurium*, in that *E. coli* adapted at a sub-lethal acidic pH are better able to survive exposure to a lethal pH than non-adapted cells. Specifically, Goodson and Rowbury (1989) demonstrated that *E. coli* grown at pH 7.0 and adapted at pH 5.0 could survive exposure to the challenge pH (pH 3.0 or 3.5) better than cells maintained at pH 7 before the challenge. Induction at pH 5.0 took approximately 5 minutes in complex broth and 30-60 minutes in minimal medium, and relied on *de novo* protein and RNA synthesis (Raja *et al.*, 1991). Interestingly, *E. coli* grown in glucose minimal medium at pH 7.0 showed a higher level of innate acid resistance than cells grown in rich broth at the same pH (Raja *et al.*, 1991). This probably reflects the general stress resistance shown by starved cells.

Although similar, habituation does show some major differences from the *S. typhimurium* ATR. Habituation does not present pre-acid shock and post-acid shock stages like the log-phase ATR, nor does it require protein synthesis at the challenge pH once habituated, unlike the situation for preshock adapted *S. typhimurium* (Rowbury and Goodson, 1993; Foster, 1993). Also the products of the *atp* operon or the *fur* gene were found to be unnecessary for habituation, while they play central roles in the ATR of *S. typhimurium* (Rowbury and Goodson, 1993).

The presence of phosphate (or polyphosphate) in the medium has two important effects on habituation; (i) it blocks habituation and (ii) it prevents killing by acid challenge (Rowbury *et al.*, 1992; Rowbury and Goodson, 1993). These effects were postulated to be caused by phosphate competing with H^+ for the PhoE pore, an outer membrane channel induced in response to phosphate starvation. The presence of phosphate impeded movement of H^+ through the PhoE pore, preventing entry to the periplasm. The greater activity of polyphosphate is in accordance with this hypothesis, as polyphosphates very efficiently stop entry of other compounds through PhoE (Rowbury and Goodson, 1993). Phosphate at 300 mmol/L almost fully protected *E. coli* grown at pH 7 from the lethal effects of pH 3.5, while polyphosphate was more potent, needing only 10 mmol/L to have the same effect. Further evidence for the role of PhoE came from *phoE* mutant strains that could not be protected from acid killing by the presence of phosphate (Rowbury and Goodson, 1993).

Interestingly, Rowbury and Goodson (1993) suggested that this protective action of phosphate may have influenced the results of Foster and Hall (1990) concerning the constitutive acid resistance of *S. typhimurium*. Foster and Hall (1990) used a medium containing 73 mmol/L phosphate in their work, a concentration that may have a significant protective effect at the challenge pH. However, whether phosphate produces a similar protective effect in *S. typhimurium* remains to be seen.

7.3.2 *Aeromonas hydrophila.*

Aeromonas hydrophila, like *S. typhimurium* and *E. coli*, exists as both a free-living aquatic organism and as an enteric pathogen of humans (Schubert, 1991). As such, it must tolerate a range of environmental pH values. Karem *et al.* (1994) have identified an acid tolerance response in this organism that protects against lethal acid pH. Bacteria grown at pH 7.2 to log-phase, then exposed to pH 5.0 for 25 minutes, showed at least a 5-log higher survival, 1.5 hours after exposure to pH 3.5, than cells without an intermediate adaptation step. This ATR response required protein synthesis, since the presence of chloramphenicol during the adaptive phase blocked the development of tolerance. Proteins produced after a shift in pH_0 from 7.2 to 5.0 were labelled using [^{35}S] methionine, then analysed by two-dimensional PAGE. At least 28 proteins showed an increase in synthesis after the pH shift, while the production of 10 proteins was decreased. The ATR response of *A. hydrophila* therefore appears very similar to the *S. typhimurium* system.

7.3.3 *Listeria monocytogenes.*

The capacity of *L. monocytogenes* to withstand acid stress may be important to its role as a pathogen, both in survival in acid foods and after ingestion by macrophages (Jones, 1990). An investigation by Kroll and Patchett (1992) established that *L. monocytogenes* is capable of mounting an adaptive acid tolerance response. Cells incubated at pH 5.0 for 2.5 hours showed significantly better survival at the challenge pH (pH 3.0) than cells grown at pH 7.0. However, no evaluation of the role of protein synthesis in adaptation was performed, so whether adaptation merely increased activity of the

constitutive homeostasis systems or was the result of an induced global response to acid can not be said.

7.3.4 Root nodule bacteria.

Acidic soil conditions have an adverse effect on the ability of legumes to fix nitrogen, mainly because of poor nodulation by nitrogen-fixing bacteria (Glenn and Dilworth, 1994). An understanding of acid-sensitivity in root nodule bacteria may therefore allow the development of strains with greater acid resistance and hence a greater nodulation capacity. Towards this aim, O'Hara and Graham (1994) demonstrated the existence of an adaptive acid tolerance response in four strains of *Rhizobium* and two strains of *Bradyrhizobium*. Bacteria grown at pH 5.0 were able to survive exposure to pH 3.0 significantly better than bacteria grown at pH 7.0. Adaptation was inhibited in the presence of chloramphenicol and so required new protein synthesis. Also, bacteria grown to stationary-phase were more acid resistant than exponentially growing cells, probably reflecting the general stress resistance of stationary-phase cells. Concurrent work by Graham *et al.* (1994) suggests that not all *Rhizobium* strains are capable of mounting an ATR, as *Rhizobium etli* exposed to pH 5.8 was still very sensitive to pH 3.3. However, the adaptive pH of 5.8 may have been too high to elicit a proper protective response.

7.3.5 *Enterococcus hirae* and *Streptococcus mutans*.

The presence of an adaptive acid tolerance mechanism in streptococci has been demonstrated by Belli and Marquis (1991). To observe adaptation, bacteria were grown in continuous culture at pH 7.0, 6.0 and 5.0. Cells were then harvested, and resuspended in media containing excess glucose to assay for the minimum pH at which glycolysis could still be performed. Cells that were grown at pH 5 were able to continue glycolysis at lower pH values than cells grown at pH 7 or 6; for example, cells grown at pH 5 had a glycolysis pH minimum of 3.45, while pH 7 grown cells had a minimum of pH 3.75. Growth at pH 5 also protected cells against killing by pH 2.5. A further characteristic of pH 5-grown cells was greater ATPase activity than alkaline-grown cells.

7.4 Acid stress and mycobacteria.

7.4.1 Relevance of acidic pH to mycobacteria.

Mycobacteria can encounter a range of acidic environments in nature. Most mycobacteria are free-living saprophytes, commonly isolated from soil and aquatic environments where exposure to acidic conditions, due to acid rain, pollution, or decaying vegetation, will be unavoidable. For instance, Kazda *et al.* (1979) demonstrated the presence of a range of mycobacterial species in sphagnum bogs in south Sweden and coastal Norway, with *M. sphagni*, *M. gordonae* and *M. chelonei* the major isolates. The prevalence of mycobacteria in water has also been well documented, with mycobacteria isolated from fresh and salt waters, as well as in chlorinated domestic water supplies (reviewed by Collins *et al.*, 1984).

Pathogenic mycobacteria will also experience acidic conditions in the host. The action of the antimycobacterial drug pyrazinamide in decreasing extracellular numbers of tubercle bacilli shows this. Pyrazinamide is only active at an acidic pH (Mitchison, 1992), thus suggesting that extracellular bacilli are in an acidic environment (Mitchison, 1992). The necrotic, cheese-like (caseous) interior of the granuloma may also be important. Although solid caseum has an approximately neutral pH, acid conditions may develop during the early stages of caseum development. Dubos (1950, 1955) suggested that acidic conditions may develop in the early stages on inflammation due to restriction of the blood flow and the gathering of mononuclear cells with predominately glycolytic metabolism at the site. Reports that suggested fatty acids accumulate in the caseum due to the lipase activity were also cited by Dubos as evidence for an acidic pH in necrotic areas (Dubos, 1950). The existence of skin and abdominal forms of non-pulmonary tuberculosis also indicate that the bacillus is not acutely acid sensitive, as the skin and abdomen have pHs of approximately 5.5 and 2.0 respectively (Grange, 1988).

Turning to the situation within the macrophage, initial work pointed to an acidic pH in macrophage vacuoles containing tubercle bacilli (detailed in Dubos, 1955). These studies relied on the action of indicator dyes to disclose the pH of the vacuole, but could not differentiate vesicles containing live or dead bacilli. Investigations by Crowle and colleagues (1991) implied that living *M. tuberculosis* and

M. avium did not reside in acidic vesicles. Using the weak bases chloroquine and ammonium chloride to determine the pH of human macrophage vesicles, they showed that dead bacilli, but not living bacilli, were associated with acidic vesicles. Subsequent work by Sturgill-Koszycki *et al.* (1994) showed *M. avium* phagosomes to be only mildly acidic, with a pH of 6.3-6.5. The authors postulated that *M. avium* selectively inhibits fusion of the phagosome with vesicles containing the proton-ATPase, hence blocking acidification.

7.4.2 Acidic pH and mycobacterial growth.

The early observations that tubercle bacilli reside in acidic intracellular and extracellular environments provided the impetus for research into the role of acid as an antimycobacterial agent. Investigating the effect of fatty acids on the *in vitro* growth of *M. tuberculosis*, Dubos established that short chain fatty acids (e.g. lactic, acetic, and capric acids) exhibited growth-promoting properties at low concentrations (near neutrality), but inhibited growth at higher concentrations (Dubos, 1950). Indeed, the bacteriostatic effect of short chain fatty-acids could be enhanced by acidifying the cultures with a few drops of HCl. Further work (Dubos, 1953) indicated that under conditions of oxygen limitation the presence of sodium lactate could be bactericidal as well as bacteriostatic. Dubos therefore postulated that organic acids may be present at the site of infection at concentrations inhibitory for growth, and that this may be a mechanism for controlling extracellular numbers of bacilli. However, this is not to say that resistance to acid is a possible virulence factor; Jackett *et al.* (1978) could not establish any link between virulence and resistance to acid.

An examination by Chapman and Bernard (1962) into the growth pHs of a range of mycobacteria produced some interesting observations. Using Dubos liquid medium adjusted to acidic or alkaline pH with 1M HCl or 1M NaOH, *M. tuberculosis* was found to have a narrow range of pH for growth (pH 5.0 to 8.4), with an optimum growth between pH 6.2 and 6.4. On the other hand, *M. smegmatis* was capable of optimum growth over a very wide pH range form pH 3.5 to 9.5. However, one flaw of the study was that pH measurements were only made before the addition of the inoculum; measurements were not carried out during, or at the end of, the study. As mycobacteria are able

to increase the pH of their growth medium by the deamination of asparagine with the concomitant release of ammonia (Drea *et al.*, 1953), the pH of the medium may have altered significantly from the original value. Hence, of the two cultures that were checked after 6 weeks growth, one had altered the initial pH from pH 4 to pH 5.3, and the other from pH 5 to pH 5.9.

A systematic study of the growth of mycobacteria in relation to the pH was carried out by Portaels and Pattyn (1982). This investigation employed Dubos oleic acid albumin agar and Lowenstein-Jensen medium adjusted to a range of pH with Sorenson's citrate and McIlvaine's citrate buffer. Mycobacterial strains were inoculated onto each medium at each pH, and incubated at the appropriate temperature. In general, the results showed that the slow growers grew optimally over a narrower pH range than the fast growers. The slow growers could be divided into two groups, one having a narrow pH range for growth (<1.1 pH units), the other having a somewhat wider pH range (>1.5 pH units). The group with the narrow pH range contained many of the slow growing pathogens, such as *M. tuberculosis*, *M. microti* and *M. bovis*; members of the group with a wider pH range included *M. gordonae*, *M. terrae* and *M. ulcerans*. Among the fast growers, *M. smegmatis*, *M. nonchromogenicum* and *M. fortuitum* showed the widest pH range, with *M. smegmatis* capable of optimum growth between pH 5.0 and 7.4 (pH 7.4 was the most alkaline pH tested). These results were in general agreement with those of Chapman and Bernard (1962) who showed that saprophytic mycobacteria can grow over a wider pH range than the pathogenic mycobacteria. Growth of mycobacteria was also independent of the buffers and media used.

Nakamura (1970) reported that growth of *M. smegmatis* in an acidic medium (pH 3.8-4.6) caused a state of competency for transformation to develop in the bacterial population. A minimal medium (standard acid medium, SAM) containing 0.1% $(NH_4)_2SO_4$, 0.15% KH_2PO_4, 0.0001% $CaCl_2$, 0.2% glycerin and 0.06% Tween 80 was used in the experiments, with pH adjustments made using 0.1M HCl or 1M NaOH. Overnight cultures of *M. smegmatis* grown in SAM were inoculated into fresh SAM, with this point taken as zero-time. To analyse the time-course of transfection, 0.5 ml samples of the culture were removed at defined time points and mixed with 1µg of phage B1 DNA. After a 30 minute incubation at 37°C, the mixture was plated

onto Trypto-soya agar with the indicator bacteria and incubated at 37°C for 24 hours to allow the plaques to develop. *Mycobacterium smegmatis* grown in SAM between pH 3.8 and 4.6 was capable of being transfected with phage DNA, with competence developing after a 6 hour lag period and reaching a plateau after 15 hours of incubation. The mechanism for the development of competence is unknown, but the lag period would suggest some biochemical alteration of the cell. Perhaps the role of pH in the development of mycobacterial competence has parallels with *Streptococcus pneumoniae* competence, where DNA uptake by competent cells relies on a pH_i of 8.3 ± 0.2 (Clave and Trombe, 1989). However, no further descriptions of mycobacterial competence could be found in the literature.

From these studies, we can draw the conclusion that mycobacteria can grow over a range of pH, with saprophytes generally displaying a wider pH range for growth than pathogenic species. The regulation of pH_i by cytoplasmic buffering, proton transport, or production of acids and bases has not been investigated in these organisms. Whether mycobacteria are capable of an adaptive response to external pH is not known.

7.5 Bacterial luciferase and its use as a reporter gene.

7.5.1 Bioluminescence and the luciferase enzyme.

The generation of light by biological processes has been recognised for centuries, with the Irish scientist Robert Boyle noting in 1667 the requirement of air for light emission from rotting wood and dead fish (cited in Meighen, 1988). Bioluminescence can be produced by two enzymes, eukaryotic luciferase or bacterial luciferase. The best example of the eukaryotic enzyme is firefly luciferase, an enzyme that catalyses the production of light from ATP and the luciferin substrate. Bacterial bioluminescence is the product of the luciferase enzyme, a dimer (approximately 80 kDa) consisting of an α and β subunit (Meighen, 1988; 1991). Luciferase catalyses the oxidation of long chain fatty acids (>7 carbons) and reduced riboflavin, generating blue-green light (λ=490nm) in the process (Stewart and Williams, 1992). The reaction can be written as follows:

$$FMNH_2 + RCHO + O_2 \longrightarrow FMN + H_2O + RCOOH + Light$$

The mechanism of light production from this reaction has been studied extensively, with the work of Hastings and Gibson (1963) providing a good reaction model. In their system, luciferase first binds with $FMNH_2$ to form intermediate I; this rapidly reacts with oxygen to form intermediate II. Reaction of II with aldehyde triggers the emission of light. Decomposition of the ternary luciferase-FMN-RCOOH compound is slow and rate-limiting, a feature that may explain why luciferase is one of the "slowest" enzymes known (Hastings, 1978; Li and Meighen, 1994).

Almost all marine luminescent bacteria are classified into three genera, *Vibrio*, *Photobacterium* and *Xenorhabdus* (Meighen, 1991). The genes responsible for the synthesis of luciferase (*luxA* and *B*) and the fatty acid substrate (*luxC*, *D* and *E*) are arranged in the same order in all operons, namely *luxCDABE* (Meighen, 1988). A set of regulatory genes, *luxI* and *luxR* have also been identified in *Vibrio fischeri* and *Vibrio harveyi* (Meighen, 1991). *luxI* is involved in the synthesis of a freely diffusible autoinducer, N-(3-oxohexanoyl)-L-homoserine lactone (HSL). The HSL autoinducer interacts with the

positive regulator produced by *luxR* to activate transcription of the *lux* operon, increasing production of the autoinducer and hence creating a positive feedback loop (Meighen, 1991). As the autoinducer is freely diffusible, the induction of one cell leads to the induction of surrounding cells, allowing a co-ordinated response from a population of cells.

The genes coding for the α and β subunits of luciferase, *luxA* and *luxB*, have been cloned from a number of species belonging to these groups, such as *Vibrio harveyi*, *Vibrio fischeri* and *Xenorhabdus luminescens* (Stewart and Williams, 1992). This has allowed the exploitation of luciferase as a reporter gene.

7.5.2. Bacterial luciferase as a reporter gene.

Reporter genes code for a product that is easy to assay. This property is particularly useful when studying the expression of a gene whose product is difficult to assay. By linking a gene promoter to a reporter gene, the activity of the promoter can be measured as a function of the reporter's expression. Hence, the conditions under which the gene would show differential expression can be easily ascertained. Commonly used reporter genes include the genes for β-galactosidase and chloramphenicol acetyltransferase (CAT) (Brosius and Lupski, 1987). The bacterial luciferase genes, with their phenotype of light production, represent ideal candidates for a reporter system (Meighen, 1991). The advantages of a luciferase-based reporter system are many; (i) the ease of light detection (at the most basic level, the naked eye in a dark room will usually suffice); (ii) the nondestructive nature of the assay, since the aldehyde substrate readily diffuses through the bacterial cell wall, removing any requirement for cell lysis; (iii) the possibility of real time measurement of promoter activity; (iv) luciferase is not present in the majority of bacteria, hence removing the possibility of background from endogenous sources (Hill *et al.*, 1993; Stewart and Williams, 1992; Hastings *et al.*, 1978).

The basic experimental design is therefore to clone the genes for the luciferase α and β subunits (*lux A* and *B*) downstream from the promoter of interest; the addition of exogenous aldehyde substrate then initiates light production. Of course, reduced riboflavin and oxygen are assumed to be present in the cell.

The utility of bacterial luciferase as a reporter of gene expression has been demonstrated by a number of groups. In one of the first such studies, Engebrecht *et al.* (1985) used the luciferase structural genes from *Vibrio fischeri* to construct a mini-Mu *lux* transposon. Using this construct, expression from the *E. coli* lactose and arabinose operon promoters, and from the *Vibrio parahaemolyticus* lateral flagella promoter, was quantified. Luciferase proved to be a sensitive and simple indicator of gene expression. This Mu*dlux* system was later employed by Francis and Gallagher (1993) to identify genetic loci in *S. typhimurium* that were induced in response to oxidative stress. Clones with *lux*-tagged loci were then allowed to interact with *in vitro* grown macrophages to see whether these loci may be induced in an *in vivo* situation. The use of a *lux* system for such a study allows gene expression in the microbe to be visualised with minimal disturbance to the macrophage. A similar approach to the above, using *lux*-bearing transposons to tag genetic loci, was utilised by Wolk *et al.* (1991) to identify genes regulated by nitrogen limitation in the cyanobacterium *Anabaena.*

A number of promoter probe vectors have also been developed using promoter-less copies of the *V. fischeri* luciferase structural genes. Carmi *et al.* (1987) developed a luciferase promoter probe vector for use in *Bacillus* species, and successfully used this system to identify promoters that are active during sporulation. The regulation of virulence genes in *L. monocytogenes* was also examined using a promoter probe vector (Park *et al.*, 1992). The *luxAB* genes were placed under the transcriptional control of the promoters for two virulence genes, *hlyA* and *plcA*. Light production was found to increase in response to a heat shock of 49°C for 30 minutes, hence permitting the induction of the virulence genes to be quantitatively measured. Park and co-workers (1989) used a promoter probe approach to identify genes that were induced in response to osmotic shock in *E. coli.* Using lac-α and tetracycline resistance as the initial reporters, a regulatory region of the *proU* operon was identified that responded to osmotic stress. This locus was further characterised (i.e., the kinetics of induction determined) using luciferase as a reporter, as the real-time, nondestructive nature of the luciferase assay makes it a better choice for these types of measurements.

One problem in the above studies was that luciferase assays had to be performed at <30°C as the *V. fischeri* enzyme is heat labile (Mackey *et al.*, 1994). The decreased growth rate of mesophiles at this temperature would therefore make the *V. fischeri* enzyme unsuitable for use with many bacteria. A more "user friendly" enzyme is the *Vibrio harveyi* luciferase, since this is a more thermostable enzyme (half-life at 45°C of 5 minutes), allowing luciferase assays to be performed at 37°C (Meighen, 1991). Work by David *et al.* (1992) and by Andrew and Roberts (1993) on the generation of bioluminescent mycobacteria established that the *V. harveyi* enzyme is functional in *M. smegmatis*. Indeed, Andrew and Roberts (1993) demonstrated that bacterial luciferase can be employed as a reporter of mycobacterial viability, making the luciferase system an ideal basis for high-throughput screening of antimycobacterial agents. In relation to bioluminescence and mycobacteria, the work of Jacobs and colleagues should be mentioned here. Rather than the bacterial luciferase genes, these workers have used the firefly luciferase gene to construct phage and plasmid based systems to screen for antimycobacterial agents (Jacobs *et al.*, 1993; Cooksey *et al.*, 1993). As mentioned previously, firefly luciferase is different from the bacterial enzyme in that it requires ATP and the substrate luciferin to produce light. The requirement for ATP in the light reaction is exploited in the screen for drug sensitivity, as drug-sensitive mycobacteria will be damaged or killed by a drug, blocking ATP synthesis and hence stopping light production. The advantage in using bacterial luciferase is the greater sensitivity of this system over the firefly system (Dr. Karen Padilla, Becton Dickinson, USA; pers comm.).

Despite the ease of use of the luciferase system, two reports have thrown some doubt on the validity of bacterial luciferase as a reporter of gene expression. Firstly, Forsberg *et al.* (1994) noted an inconsistency in the measured activity of two promoters when using the *V. harveyi* luciferase, compared to *lacZ*, to report promoter activity. This study used a mutant form of the leucine operon promoter, *leuP*, from *S. typhimurium* that is termed *leu-500*. This promoter is usually inactive but can have its activity restored if placed in a strain with a mutation in the DNA topoisomerase I gene, a mutation that causes increased negative supercoiling of DNA. Expression from the *leu-500* promoter is therefore sensitive to DNA supercoiling. However, when

this promoter was placed upstream from the luciferase genes in a wild-type genetic background, the promoter was activated and produced light. Repeating the experiment using β-galactosidase as the reporter revealed the promoter to be inactive.

Similarly, experiments using the *proU* promoter produced conflicting results when β-galactosidase or luciferase was used as the reporter. Expression from the *proU* promoter is normally repressed at low osmolarity and induced at high osmolarity. Osmotic regulation requires a downstream regulatory element that represses the promoter at low osmolarity by affecting local DNA topology (Forsberg *et al.*, 1994). Hence when the promoter, lacking the regulatory element, was cloned upstream of the β-galactosidase reporter, the promoter showed constitutive levels of expression at high or low osmolarity. However, when the same promoter was cloned upstream of *luxAB* expression was regulated in response to osmolarity; the promoter was repressed at low osmolarity and induced at high osmolarity. The presence of *luxAB* downstream of the promoter seemed to reproduce the effect of the downstream regulator.

These results suggest that *luxAB* affects the local topology of the DNA, altering the activity of promoters that are influenced by supercoiling. This does not appear to be a general effect on all topologically sensitive promoters; expression from the *gyrB* promoter was unaffected by *luxAB*.

The second reported problem is that the presence of luciferase may place the cell under oxidative stress (Gonzalez-Flecha and Demple, 1994). The activity of the *E. coli soxS* gene promoter, a locus that is induced in response to intracellular superoxide, was measured using a transcriptional fusion based on the *V. harveyi* luciferase. The promoter was found to be induced not only by superoxide-generating agents (e.g., paraquat), but also when the *lux* genes were present in the cell. The latter situation could be explained by the generation of superoxide as a consequence of the luciferase reaction. This would involve the reaction of the luciferase-$FMNH_2$ intermediate with oxygen, then the transfer of electrons from $FMNH_2$ to oxygen to produce superoxide. This "dark" reaction may be especially prevalent in the absence of the aldehyde substrate. Hence, bacterial luciferase may be unsuitable for the identification or study of promoters regulated by oxidative stress.

Bacterial luciferase therefore allows real-time, nondestructive measurements of promoter activity to be made, giving it clear advantages over other reporter systems. Nevertheless, some problems may arise when studying the expression of certain promoters, and these problems should be kept in mind when using luciferase as a reporter. For in-depth analysis of a promoter, it appears a combination of reporter systems may provide the best picture of the promoter's activity.

7.6 Project aims and experimental strategy

The aim of this work was to investigate the responses of mycobacteria to acid stress. *Mycobacterium smegmatis* mc^{2}155 (Snapper *et al.*, 1990) was the model organism for these studies. The choice of this organism was based on a number of criteria, including (i) it is non-pathogenic, (ii) it has a relatively rapid growth rate compared to other mycobacteria, and (iii) it is a mutant strain that is easier to transform than other mycobacteria, facilitating genetic studies. Three questions were addressed:

(i) Can mycobacteria adapt at sub-lethal acid conditions to survive exposure to lethal acid conditions?

(ii) Does adaptation involve any change in protein synthesis?

(iii) Is there a genetic response to acid stress?

The answers to these questions would be sought using the following strategies:

(i) Determine the lethal and adaptive pHs for *M. smegmatis* mc^{2}155. Determine whether exposure to the adaptive pH elicits protection against the lethal pH.

(ii) Expose *M. smegmatis* mc^{2}155 to the adaptive pH in the presence of [^{35}S] methionine. Analyse protein profiles of adapted cells on SDS-PAGE gels to see whether any proteins are induced or repressed.

(iii) Construct a promoter probe library from *M. smegmatis* mc^{2}155 genomic DNA. Isolate clones that show increased expression of the reporter at the adaptive pH. Sequence the insert DNA present in the

clones as a means to identify the function of these loci in *M. smegmatis* mc^2155.

Using these approaches it was hoped to generate a model of the responses to acid stress in *M. smegmatis* mc^2155 that may be generally applicable to all mycobacteria.

Chapter 8

Methods

8.1 Demonstration of an acid tolerance response (ATR) in *M. smegmatis*.

Methods in Section 8.1 were devised by Lyn O'Brien (Department of Microbiology and Immunology, University of Leicester) in collaboration with Stephen Gordon.

8.1.1 Survival of *M. smegmatis* over a pH range.

Mycobacterium smegmatis mc^{2}155 (Snapper *et al.*, 1990) was the model organism for these studies. This strain has a high transformation efficiency, making genetic analysis much easier, and so is a useful starting point to investigate the response to environmental acid stress at both the genetic and physiological level.

Mycobacterium smegmatis mc^{2}155 was grown overnight in 50ml of 7H9 medium (pH 7.6) at 37°C with shaking (200 rpm). A 1ml sample of the culture was sonicated with three 5 second bursts of 40W (Ultrasonic Engineering, UK) to disrupt clumps of bacilli. Numbers of bacilli per millilitre were estimated by counting dilutions of the culture microscopically using a Thoma chamber, then using the formula:

$$\text{bacteria/ml} = \frac{\text{numbers of bacilli x dilution x } 2\text{x}10^7}{\text{number of squares counted}}$$

Samples of non-sonicated culture containing approximately $1\text{x}10^6$ bacteria/ml were then added to 1ml volumes (in duplicate) of 7H9 medium at pH 6.0, 5.0, 4.5, 4.0, 3.5, 3.0, 2.5, and 2.0 (the 7H9 medium had been acidified using either 2M H_3PO_4 or 2M HCl). Cultures were then incubated at 37°C. At zero, two and four hours a 1ml culture at each pH was removed, sonicated (as above), and 50µl samples from serial dilutions plated onto 7H11 agar to determine culture viability.

8.1.2 Measurement of acid tolerance response (ATR).

The procedure for measurement of the mycobacterial acid tolerance response was based on the methods of Foster and Hall (1990; 1991). *Mycobacterium smegmatis* mc^{2}155 (Snapper *et al.*, 1990) was used throughout the adaptation experiments. Two to three isolated

colonies from a stock 7H11 agar plate of *M. smegmatis* were used to inoculate 20ml of 7H9 medium plus ADC. Cultures were incubated at 37°C overnight with shaking (200rpm). A 1ml sample of the culture was sonicated and bacterial numbers estimated as described previously (Section 8.1.1). A sample of non-sonicated culture was then added to two 50ml volumes of 7H9 plus ADC (pH 7.6) in 500ml vessels, such that the final cell concentration was 10^6 bacteria/ml. To produce 7H9 growth medium with a pH of 7.6, 0.94g of 7H9 powder (Difco) was dissolved in 180ml of distilled water, then ADC supplement (20ml ADC per 180 ml of 7H9) and 1ml of 10% (v/v) Tween 80 added. The pH of the medium was adjusted to pH 7.6 by dropwise addition of 2M NaOH, then filter sterilised. Cultures were incubated at 37°C with shaking (200rpm) until the cell concentration had attained 10^8 bacteria/ml (approximately 12 to 14 hours).

After 12-14 hours incubation, the pH of one of the cultures was dropped to pH 5.0 with filter sterilised 2M H_3PO_4, then reincubated at 37°C to allow for cell numbers to reach $2x10^8$ bacteria/ml (2 hours for cultures at pH7.6, 4 hours for cultures at pH 5.0). The pH of both cultures was then dropped to pH 3.5 with filter sterilised 2M H_3PO_4. A 1ml fraction of each culture was taken directly after the drop to pH 3.5 and subsequently at one hour intervals for 4 hours, all fractions being immediately placed on ice. The culture samples were sonicated as before and 50µl volumes from serial dilutions plated onto 7H11 agar. Agar plates were incubated at 37°C for 36 hours and bacteria/ml determined. Viability was defined as the numbers of bacilli remaining as a percentage of the starting count

8.1.3 The effect of chloramphenicol on adaptation at pH 5.0.

The first step was to determine the minimum inhibitory concentration (MIC) of chloramphenicol for *M. smegmatis* mc^2155 at pH 5.0. A culture of *M. smegmatis* mc^2155 was grown overnight at 37°C with shaking. A sample of culture was removed, then sonicated, and bacterial numbers estimated using a Thoma chamber as described in Section 8.1.1. Non-sonicated *M. smegmatis* ($1x10^9$ bacteria/ml) were then inoculated into 10ml volumes of 7H9 medium at pH 5 (acidified using 2M H_3PO_4) containing chloramphenicol ranging in concentrations from 0-100µg/ml. The cultures were incubated at 37°C with shaking (200 rpm) for 16 hours, then viable counts determined (Section 8.1.1).

The MIC of chloramphenicol was found to be 20μg/ml. The ATR experiment (Section 8.1.2) was then repeated, except that 30 minutes before the pH was reduced to 5.0, 200μg/ml chloramphenicol was added to two cultures, while an identical volume of 100% ethanol (the solvent for chloramphenicol) was added to two control cultures.

8.2 SDS-PAGE analysis of *M. smegmatis* protein profile after acid stress.

8.2.1 Radiolabelling of *M. smegmatis* proteins synthesised after acid stress.

The strategy to determine whether any proteins were induced or repressed at the adaptive pH was to radiolabel proteins synthesised after the pH shift from 7.6 to 5.0, then analyse them using SDS-PAGE. Radiolabelling of proteins from cultures maintained at pH 7.6 would serve as the control.

An overnight culture of *M. smegmatis* mc^2155 was used to inoculate 10ml of fresh 7H9 medium plus supplements (1:20 dilution). Following incubation at 37°C overnight with vigorous aeration (200 rpm), a sample of culture was sonicated (Ultrasonic Engineering, UK) with three 5 second bursts of 40W to disrupt clumps of bacilli. Bacterial numbers were then estimated by counting bacilli microscopically in a Thoma chamber, as described in Section 8.1. A volume of the culture was then added to 50ml of 7H9 medium (pH 7.6) to give a final concentration of 10^6 bacteria/ml. Cultures were incubated at 37°C with vigorous aeration until cell concentrations of 10^8 bacteria/ml had been attained (approximately 14 hours). One of the 50ml pH 7.6 cultures was then removed from the incubator and maintained at 37°C in a water bath. Working quickly, a 1ml sample of the culture was removed to a 1.5ml sterile eppendorf containing 20μCi of [^{35}S]methionine (>37 TBq/mmol, >1000Ci/mmol; Amersham) and incubated at 37°C. The pH of the remaining culture was then dropped to pH 5.0 with approximately 500μl of 2M H_3PO_4 or 2M HCl (thermoequilibrated to 37°C). A 1ml sample of the adjusted culture was then removed to a 1.5ml eppendorf containing 20μCi of [^{35}S]methionine and incubated at 37°C (pH adjustment took approximately 1 minute, and was to within 0.1 pH unit). Samples (200μl) of the pH 7.6 and pH 5.0 labelled cultures

were removed at specific time points and added to 1ml of ice-cold 0.1M phosphate buffer (pH 7.3) in 1.5ml eppendorfs and kept on ice. After labelling, samples were stored frozen at -20°C until required.

8.2.2 PAGE analysis of labelled proteins.

Separating Gel (12%):

3.35ml	Distilled water
2.5ml	1.5m Tris-HCl (pH 8.8)
100µl	10% (w/v) SDS
4.0ml	Acrylamide/Bis 30% (v/v) (Protogel, National Diagnostics)
50µl	10% (w/v) ammonium persulphate (fresh)
5µl	TEMED

Stacking Gel (4%)

6.1ml	Distilled water
2.5ml	0.5M Tris-HCl (pH 6.8)
100µl	10% (w/v) SDS
1.3ml	Acrylamide/Bis 30%(v/v) (Protogel, National Diagnostics)
50µl	10% (w/v) ammonium persulphate (fresh)
10µl	TEMED

5x Running buffer

9g	Tris base
43.2g	Glycine
3g	SDS
Distilled water to 600ml	

Sample Loading buffer

4.0ml	Distilled water
1.0ml	0.5M Tris-HCl (pH 6.8)
0.8ml	Glycerol
1.6ml	10% (w/v) SDS
400µl	β-mercaptoethanol
200µl	0.05% (w/v) bromophenol blue

Labelled proteins were separated and visualised after electrophoresis through denaturing polyacrylamide gels using the method of Laemmli (1970). A Bio-Rad mini-PROTEAN II electrophoresis system was used in accordance with the manufacturer's instructions. Glass plates were cleaned rigorously with ethanol prior to assembly. The separating gel mix was poured between the gel plates and overlaid with 1-2ml of water-saturated n-butanol. The gel was allowed to polymerise for 45-60 minutes before the n-butanol was washed off with distilled water and the gel surface blotted dry. The stacking gel solution was poured onto the stacking gel, and a comb inserted. After polymerisation of the stacking gel, the comb was carefully removed and the wells washed out with distilled water.

Radiolabelled *M. smegmatis* samples were thawed to room temperature, then centrifuged at 13,000 rpm in a MicroCentaur bench-top centrifuge for 3 minutes. The resulting cell pellet was resuspended in 50µl of sample loading buffer. Samples of the cell suspension were heated to 100°C in a boiling water bath for 10 minutes, then cooled to room temperature prior to loading onto the gel. Gels were subjected to a constant 200V for 45-60 minutes, until the blue dye-front had reached the bottom of the gel. The gel was recovered from the electrophoresis assembly, then stained and fixed in 3-5 volumes of Comassie blue stain (0.25% (w/v) Comassie blue R250 made up in SDS-PAGE destain) for 15-30 minutes. Gels were destained in 5-10 volumes of SDS-PAGE destain (300ml methanol, 100ml glacial acetic acid, 600ml distilled water) to visualise the loaded protein. The amount of protein loaded in each well could therefore be seen, hence ensuring that equal amounts of protein had been loaded into each well and preventing any false results due to loading different quantities. Gels were then soaked in 2-3 volumes of the fluorographic enhancer Amplify (Amersham) for 20 minutes, and dried onto Whatman 3mm filter paper for 90 minutes at 80°C. Dried gels were exposed to Cronex X-ray film (DuPont) at -70°C, then film was processed in an Agfa-Geveart automatic film processor in accordance with the manufacturer's instructions.

8.3 Construction of a promoter probe library.

To identify promoters that showed increased activity in response to acid stress a promoter probe library was constructed. This entailed cloning random fragments of *M. smegmatis* DNA upstream of a reporter gene and screening for increased expression of the reporter. Bacterial luciferase was the chosen reporter system.

Mycobacterium smegmatis mc^2155 genomic DNA was isolated using the method described in Section 2.3.3. DNA was partially digested with *Sau*3AI using serial dilutions of enzyme, as described in Sambrook *et al.* (1989). Briefly, 20µg of *M. smegmatis* genomic DNA was added to 1x *Sau*3AI enzyme buffer in a final volume of 90µl, then left at 4°C for approximately 60 minutes to allow good mixing to occur. The DNA solution was then dispensed into six 10µl fractions and one 20µl fraction (all fractions were kept on ice). Five units of *Sau*3AI were added to the 20µl fraction and mixed thoroughly. This gave a final concentration of 0.25U *Sau*3AI/µl. Ten microlitres of this solution was removed and mixed with a 10µl fraction of the original DNA solution. These two-fold serial dilutions were repeated until a final concentration of 0.07U *Sau*3AI/10µl fraction was reached. As a negative control, one fraction had no enzyme added. Reaction tubes were then incubated at 37°C for 60 minutes exactly. Restriction of the DNA was stopped by adding EDTA (pH 8.0) to a final concentration of 40mM. Digests were analysed by electrophoresis through 0.8% (w/v) agarose gels. Restriction enzyme concentrations that gave fragments in the desired range (2-6kb) were repeated, scaling up the digests to use 10-15µg of genomic DNA. Partial digests were electrophoresed through 0.5% (w/v) preparative agarose gels and gel slices containing the desired fragment size excised. DNA was purified from the agarose slices using the Pharmacia Bandprep kit as described in Section 2.5.6.

The plasmid pSG10 was digested to completion with *Bam*HI. The extent of digestion was measured by transforming fractions of the digest into *E. coli* DH5α and measuring the numbers of transformants; zero or very few transformants indicated the plasmid was completely digested. The plasmid was then dephosphorylated as described in Section 2.5.3, using the protocol for protruding 5' termini. Analysis by self-ligation showed that approximately 90% of the *Bam*HI digested pSG10 was dephosphorylated.

Size-selected *M. smegmatis* DNA was then ligated to the dephosphorylated pSG10 (100ng-250ng) at an insert:vector ratio of 3:1. Ligations were performed as described in Section 2.5.4. Ligations were ethanol precipitated, resuspended in 6µl of sterile nanopure water, and transformed by electroporation into electrocompetent *E. coli* DH5α as outlined in Section 2.4.1.2. Transformants were selected on L-agar containing kanamycin at 20µg/ml. Transformants were washed off the agar plates using 2-3ml of sterile PBS, then plate washings were pooled (approximately 10,000 transformants per pool) and used to inoculate 1L of LB broth containing kanamycin at 20µg/ml. The culture was allowed to grow for 1 hour at 37°C, then plasmid DNA was harvested from the culture as described in Section 2.3.1.1. Fractions (2-5µl) of this amplified plasmid DNA library were then electroporated into electrocompetent *M. smegmatis* mc²155 as described in Section 2.4.2. Transformants were selected for growth on L-agar plates containing 20µg/ml kanamycin, and these master plates stored at 4°C.

8.4 Luciferase assays.

8.4.1 Luciferase activity of *M. smegmatis* mc²155 [pSG15] cultures.

Mycobacterium smegmatis mc²155 harbouring the plasmid pSG15 was grown to mid-log phase in 7H9 growth medium plus supplements and containing 20µg/ml kanamycin. To induce the acetamidase promoter, cultures were then diluted 1:250 into minimal medium (Draper, 1967; see below) containing 20µg/ml kanamycin, 0.2% (v/v) trace element solution (Hopwood and Wright, 1978), 0.05% (v/v) Tween 80 and 0.2% (w/v) acetamide (from a 20% (w/v) stock solution) as the carbon source. Cultures of *M. smegmatis* mc²155 [pSG15], grown in this medium but with 0.2% (w/v) succinate (from a 20% (w/v) stock solution) as the carbon source, were used as controls. The minimal medium (Draper, 1967) was made as follows:

Minimal medium:

1g	K_2HPO_4
1g	KH_2PO_4
1g	NH_4Cl
2ml	Trace elements solution*
5ml	10% (v/v) Tween 80
	Distilled water to 1L

After sterilisation by autoclaving at 121°C (15 psi), 5ml of 10%(w/v) $MgSO_4$ (filter sterilised) was added.

**Trace elements solution:*

40mg	$ZnCl_2$
200mg	$FeCl_3.6H_2O$
10mg	$CuCl_2.2H_2O$
10mg	$MnCl_2.4H_2O$
10mg	$Na_2B_4O_7.H_2O$
10mg	$(NH_4)_6Mo_7O_{24}.4H_20$
	Distilled water to 1L

Sterilised by autoclaving

Decanal (Sigma) was the chosen substrate for luciferase (Hill *et al.*, 1991). A Luminoskan luminometer (Labsystems), thermoequilibrated to 37°C, was used to measure light production. Samples (200µl) of the acetamide and succinate grown cultures were dispensed into black 96 well fluoroplates (Labsystems) and placed in the sample tray of the luminometer. The luminometer was programmed to deliver 10µl of a 1% (v/v) decanal solution in ethanol to 200µl samples of culture, mix for 1 second, then integrate light production over 15 seconds at 37°C.

A sample of culture (approximately 2ml) was removed to determine the viable bacterial count. The culture sample was sonicated with three 5 second bursts of 40W (Ultrasonic Engineering, UK) to disrupt mycobacterial clumps, then serially diluted in sterile nanopure water to 10^{-5}. Fifty microlitre fractions of the dilutions were then dispensed onto dried L-agar plates (containing 20mg/ml kanamycin) and incubated at 37°C for 48 hours.

8.4.2 Screening the promoter probe library for clones with acid-inducible luciferase activity.

7H9 culture medium was made at both pH 7.6 and 5.0. This was achieved by dissolving 0.94g of 7H9 pre-made mix (Difco) in 180ml of distilled water, then adding 20ml of ADC enrichment. The pH of the medium (pH 6.6-6.8 as standard) was then adjusted to pH 7.6 using 2M NaOH, or to pH 5.0 using 2M H_3PO_4, filter sterilised and stored at 4°C. All pH adjustments were to within 0.1 pH unit of pH 7.6 or 5.0 respectively. Library clones were picked from master plates and (i) inoculated into 200µl of media at pH 7.6 containing 20µg/ml kanamycin in 96 well fluoroplates (Labsystems), then (ii) streaked onto a fresh L-agar plate containing 20µg/ml kanamycin. The L-agar plates were incubated at 37°C for 48 hours, while fluoroplates were wrapped in Saran Wrap (DuPont) and placed on wet towels in an air-tight box (to prevent evaporation) at 37°C for 48 hours. Twenty microlitres of the 48 hour old pH 7.6 culture was then added, in duplicate, to 200µl of fresh media at pH 7.6 and pH 5.0 (thermoequilibrated to 37°C) in black fluoroplates. The fluoroplates were wrapped in Saran-wrap and placed at 37°C for 4 hours exactly. After incubation, fluoroplates were placed in the Luminoskan luminometer sample tray (thermoequilibrated to 37°C). Ten microlitres of 1% decanal (v/v) solution was added to each 220µl culture, and after a 1 second mixing step light production over 15 seconds was integrated.

8.4.3 Time course of luciferase induction in *M. smegmatis* [pSGA197] and *E. coli* [pSGA197].

Growth medium was adjusted to pH 7.6 using 2M NaOH, or to pH 5.0 using 2M H_3PO_4 (Section 8.4.2), then filter sterilised and stored at 4°C. A 10ml volume of the pH 7.6 medium was inoculated with 2-3 colonies of *M. smegmatis* [pSGA197] from a fresh agar plate and incubated at 37°C with shaking (200rpm) for 36-48 hours. A 2ml sample was then used to inoculate 50ml of 7H9 media (pre-warmed to 37°C) at pH 7.6 and pH 5.0, and 2-3ml of culture removed before incubating the culture at 37°C. The 2-3ml that was removed (i) had its pH measured using a Mettler Delta 340 pH meter, and (ii) 4x 200µl samples were dispensed into a fluoroplate in a Labsystems luminometer (thermoequilibrated to 37°C) and light production measured over 15 seconds after the addition of 10µl of 1% (v/v) decanal solution in

ethanol. The remainder of the culture sample was then briefly sonicated with 3x 5 second bursts of 40W (Ultrasonic Engineering, UK) to disrupt clumps of bacilli, diluted to 10^{-5} in sterile nanopure water, and 4x 50µl fractions of the 10^{-3} to 10^{-5} dilutions dispensed onto dried L-agar plates containing 20µg/ml kanamycin. This procedure was repeated for each culture at pH 7.6 and 5.0, every 60 minutes.

To measure luciferase activity in *E. coli* [pSGA197], the experimental details were identical to the above, except that *E. coli* [pSGA197] "starter culture" was allowed to grow for 12-14 hours, then 500µl used to inoculate 50ml of L broth at pH 7.6 and 5.0. Culture samples of *E. coli* did not have to be sonicated prior to dilution in PBS.

Chapter 9

Results

9.1 The acid tolerance response (ATR) of *M. smegmatis* mc²155.

Experiments described in Sections 9.1.1, 9.1.2, 9.1.3 and 9.1.4 were performed by Lyn O'Brien (Department of Microbiology and Immunology, University of Leicester) in collaboration with Stephen Gordon.

9.1.1 Selecting the adaptive and lethal pH values for *M. smegmatis*.

The demonstration of an acid tolerance response (ATR) in *Salmonella typhimurium* by Foster and Hall (1990) relied on exposure of bacteria to a mildly acidic "adaptive" pH before exposing to a normally lethal pH, at which they showed significantly greater survival than unadapted cells. The adaptive pH was chosen on the basis of experiments that showed: (i) *Salmonella* genes that are responsive to external acid pH are induced at pH 5.5-6, (ii) *S. typhimurium* tolerates exposure to pH 5.8, with only a slightly slower growth rate than bacteria grown at pH 7.4 (Foster and Hall, 1990). Therefore, as a starting point for the investigation of acid tolerance in *M. smegmatis*, the adaptive and lethal pH values for this organism had to be determined. For this purpose, the lethal pH was defined as the external pH that caused a significant decrease in the viability of *M. smegmatis* cultures. The adaptive pH was defined as the lowest pH at which *M. smegmatis* still replicated; it was assumed that at this pH the organism would undergo fundamental biochemical changes that would allow survival at normally lethal acidic pH values.

Cultures of *M. smegmatis* mc²155 were grown in 7H9 medium plus supplements at pH 7.6, then inoculated into medium at pH values ranging from pH 6 to 2, using either 2M H_3PO_4 or 2M HCl to acidify the media. Cultures were checked for viability over a 4 hour period, with the results calculated as the mean $\log_{10}$ CFU/ml at each time point (Table 9.1). The results suggest that pH 5 should be used as the adaptive pH when using HCl to acidify the growth medium, since

after 4 hours at this pH *M. smegmatis* was still capable of replication, while cultures at pH 4.5 showed a decrease in viability. Similarly, pH 3 was picked as the lethal pH for HCl acidified cultures, as the next lowest pH tested, pH 2.5, rapidly killed the bacteria and so might prove toxic even to adapted bacteria. The same criteria were used to select pH 5 and pH 2.5 as the adaptive and lethal pH respectively when using H_3PO_4 to acidify the growth media.

Table 9.1. Viability of *M. smegmatis* mc²155 over a range of pH.

	mean $\log_{10}$ CFU/ml[a]			mean $\log_{10}$ CFU/ml[b]		
pH	**t=0**	**t=2**	**t=4**	**t=0**	**t=2**	**t=4**
6	6.04	6.05	6.25	6.12	6.46	6.84
5	6.04	6.06	6.28	6.10	6.19	6.45
4.5	6.03	6.03	6.09	6.12	6.05	6.07
4	6.01	5.98	6.00	6.19	6.04	5.61
3.5	5.99	5.95	5.91	6.12	5.84	5.10
3	5.98	5.78	5.79	6.11	5.99	4.24
2.5	5.96	0	0	6.15	5.86	4.22
2	5.89	0	0	6.06	0	0

a 2M HCl was used to acidify the medium.
b 2M H_3PO_4 was used to acidify the medium.
"t" denotes time of sample taking in hours.
All $\log_{10}$ CFU/ml values are the mean of two counts taken at the appropriate time point.

9.1.2 Demonstration of an acid tolerance response (ATR) in *M. smegmatis* mc²155.

The working hypothesis for demonstrating an acid tolerance response in *M. smegmatis* was that bacteria grown at the adaptive pH should show significantly greater survival at the lethal pH than unadapted cells. This hypothesis was tested using 2M HCl or 2M H_3PO_4 to acidify the 7H9 growth medium to the adaptive and lethal pH. Duplicate cultures of *M. smegmatis* mc²155 were grown initially at pH 7.6 for 14 hours (to ensure that inducible acid-resistance mechanisms were "off"), then the pH of one culture dropped to the lethal pH (pH 3.0 when using 2M HCl, or pH 2.5 with 2M H_3PO_4; see Table 9.1) while the other culture was brought to the adaptive pH (pH 5 when using HCl or H_3PO_4) and maintained at this pH for 4 hours (one doubling) before being dropped to the lethal pH. Viable counts of the adapted and unadapted cultures were then performed for four hours after the drop to the lethal pH, and percentage viability calculated. Results from these experiments are shown as paired observations in Tables 9.2a and 9.2b (HCl experiments), and Table 9.3 (H_3PO_4 experiments). Percentage viability was determined for each culture at 1 hour intervals for 4 hours. Using 2M HCl to effect the drop in culture pH and taking experiment 1 as an example, 4 hours after the drop to the lethal pH the adapted culture showed a percentage viability of 1%, while the unadapted culture had a percentage viability of 0.3% (Tables 9.2a and 9.2b). However, when 2M H_3PO_4 was used to acidify the culture to the adaptive and lethal pH (pH 5 and 2.5 respectively) no ATR was evident, with adapted and unadapted cultures showing a similar decrease in viability in response to the lethal pH (Table 9.3).

As an ATR had been demonstrated using HCl with the same adaptive pH as with H_3PO_4 (pH 5) but with a lethal pH of 3.0 rather than pH 2.5, the question arose as to whether pH 2.5 may have been too toxic, overcoming any protection provided by the adaptation step. Therefore, the H_3PO_4 experiment was repeated exactly as before, but this time with pH 3.5 as the lethal pH. Tables 9.4a and 9.4b show the results of 7 such experiments. Taking experiment 1 as an example, the adapted culture showed a greater percentage viability after 1 hour's exposure (91%) than the culture sample exposed directly to pH 3.5 from pH 7.6 (57%). After 4 hours at the lethal pH, viability of the adapted culture was 52% compared to 0.2% in the unadapted culture.

Taken together with the results from the HCl ATR experiments, these results suggested that exposure of *M. smegmatis* to a mild acid stress triggered a biochemical realignment that allowed the organism to survive a normally lethal acidic pH.

9.1.3 Time course of adaptation at pH 5.0.

The length of time that cultures were left at the adaptation pH was based on the work of Foster and Hall (1990), who allowed *S. typhimurium* cultures to undergo one doubling at the adaptive pH before exposure to the lethal pH. In the case of *M. smegmatis* cultures, one doubling at pH 5.0 was achieved in 4 hours, and this was sufficient to allow expression of an ATR. To determine if a shorter exposure to the adaptive pH would still allow the development of an ATR, cultures of *M. smegmatis* were held at pH 5.0 for 15, 60 and 120 minutes before dropping the culture pH to 3.5. The results of these experiments are shown in Tables 9.5a and 9.5b. Cultures held at the adaptive pH for 15 minutes showed no evidence of an ATR, with a similar decrease in viability of "adapted" bacteria at the lethal pH as unadapted bacteria. However, an ATR was apparent in cultures that had been maintained at pH 5.0 for 60 or 120 minutes, indicating that 60 minutes exposure to the adaptive pH is sufficient for the development of an ATR in *M. smegmatis.*

9.1.4 Effect of chloramphenicol on development of ATR.

The development of acid tolerance in *M. smegmatis* could be achieved by a genetic process that requires gene induction and the synthesis of new protein, or by a physiological alteration of the cell by the action of pre-existing molecules, such as an increase in the buffering capacity of the cytoplasm. To determine whether acid tolerance relies on the synthesis of new proteins, the ATR experiments were repeated as described in Section 8.1.3 with the addition of the protein synthesis inhibitor chloramphenicol during adaptation. If *de novo* protein synthesis is required for adaptation, inclusion of chloramphenicol during the pH 5 adaptation step would prevent the development of tolerance and render "adapted" bacteria sensitive to the lethal pH. The results of these experiments are shown in Table 9.6, and they support a role for protein synthesis in the *M. smegmatis* ATR. Cultures that were adapted at pH 5.0 showed a greater percentage viability at pH 3.5 than

unadapted cultures, with 12% viability for adapted cultures compared to 1% for unadapted cultures. However, cultures that had chloramphenicol (200 μg/ml) added 30 minutes before exposure to the adaptive pH showed lower levels of viability (4% and 3% viability) than cultures adapted in the absence of chloramphenicol.

9.1.5 Protein profile produced after acid stress.

Protein synthesis in *M. smegmatis* mc^2155 was monitored after a shift in external culture pH from 7.6 to 5 by incubating culture samples in the presence of [^{35}S] methionine (method described in Section 8.2.1). Radiolabelled proteins were then separated by electrophoresis through denaturing polyacrylamide gels and visualised by autoradiography (Section 8.2.2). A typical autoradiogram from these experiments is shown in Figure 9.1, where samples were removed from pH 7.6 and pH 5 cultures at 5, 15, 30, 60 and 90 minutes. The profile of protein synthesis at pH 7.6 was as might be expected, with a gradual increase in the quantity of labelled protein over the 90 minute sampling period. However, the protein profile produced by bacteria exposed to a pH shift from 7.6 to 5 showed a cessation of protein synthesis after the drop to pH 5, with radiolabelled proteins becoming visible again after approximately 60 minutes. This dramatic effect on protein synthesis was evident when using either 2M HCl or 2M H_3PO_4 to effect the pH drop. The other feature of the protein profile was that there appeared to be no increase in the expression of any "shock" proteins, such as Hsp60 and Hsp70, even when sampling was continued for up to 135 minutes.

Table 9.2a. Viability of adapted *M. smegmatis* mc²155 cultures at the lethal pH, using 2M HCl to acidify the medium.

Time (hours)	% Viability[a] Expt. 1	Expt. 2	Expt. 3	Expt. 4
0	100	100	100	100
1	55	83	78	95
2	12	78	71	74
3	4	78	59	60
4	1	66	58	46

[a] The culture was adapted at pH5.0 for four hours before exposure to pH 3.0. Percentage viability was calculated from the mean of two viable counts taken at the appropriate time point after the pH was dropped to 3.0 with 2M HCl.

Table 9.2b. Viability of unadapted *M. smegmatis* mc²155 cultures at the lethal pH, using 2M HCl to acidify the medium.

Time (hours)	% Viability[a] Expt. 1	Expt. 2	Expt. 3	Expt. 4
0	100	100	100	100
1	34	69	68	93
2	7	63	59	60
3	0.5	48	50	18
4	0.3	14	15	2.5

[a] Percentage viability was calculated from the mean of two viable counts taken at the appropriate time point after the pH was dropped to 3.0 with 2M HCl.

Table 9.3. Viability of adapted and unadapted *M. smegmatis* mc²155 cultures at the lethal pH (pH 2.5), using 2M H_3PO_4 to acidify the medium.

	% Viability[a]	
Time (hours)	**Adapted[b]**	**Unadapted**
0	100	100
1	9	11
2	8	9
3	5	4
4	5	2

[a] Percentage viability was calculated from the mean of two viable counts after the pH was dropped to 2.5 using 2M H_3PO_4.

[b] The culture was adapted at pH 5.0 for four hours before exposure to pH 2.5.

Table 9.4a. Viability of adapted *M. smegmatis* mc²155 cultures at the lethal pH, using H_3PO_4 to acidify the medium.

Time (hours)	% Viability[a]						
	Expt. 1	Expt. 2	Expt. 3	Expt. 4	Expt. 5	Expt. 6	Expt. 7
0	100	100	100	100	100	100	100
1	91	95	76	77	88	98	58
2	78	95	43	59	43	52	22
3	76	71	17	42	23	21	2
4	52	52	4	27	21	10	1

[a] The culture was adapted at pH5.0 for four hours. Percentage viability was calculated from two viable counts taken at the appropriate time point after the pH was dropped to 3.5 with 2M H_3PO_4.

Table 9.4b. Viability of unadapted *M. smegmatis* mc²155 cultures at the lethal pH, using H_3PO_4 to acidify the medium.

Time (hours)	% Viability[a]						
	Expt. 1	Expt. 2	Expt. 3	Expt. 4	Expt. 5	Expt. 6	Expt. 7
0	100	100	100	100	100	100	100
1	57	69	60	35	28	50	32
2	13	28	38	13	12	24	16
3	1	2	12	8	11	18	4
4	0.2	0.4	4	6	8	12	2

[a] Percentage viability was calculated from two viable counts taken at the appropriate time point after the pH was dropped to 3.5 with 2M H_3PO_4.

Table 9.5a. Viability of adapted *M. smegmatis* mc²155 cultures.

	% Viability at the lethal pH[a]					
	t=15		t=60		t=120	
Time (hrs)	Expt. 1	Expt. 2	Expt. 1	Expt. 2	Expt. 1	Expt. 2
0	100	100	100	100	100	100
1	48	43	98	93	89	83
2	22	29	67	63	71	62
3	18	13	56	59	43	45
4	7	3	48	41	36	38

[a] Cultures were adapted at pH5.0 for 15, 60 and 120 minutes ("t" values in table). Percentage viability was calculated from the mean of two viable counts taken at the appropriate time point after the pH was dropped to 3.5 with H_3PO_4.

Table 9.5b. Viability of unadapted *M. smegmatis* mc²155 cultures.

	% Viability at the lethal pH[a]					
	t=15		t=60		t=120	
Time (hrs)	Expt. 1	Expt. 2	Expt. 1	Expt. 2	Expt. 1	Expt. 2
0	100	100	100	100	100	100
1	49	43	47	53	51	49
2	21	21	39	35	43	36
3	19	11	15	12	12	4
4	10	1	10	8	8	2

[a] Cultures were maintained at pH 7.6 for 15, 60 or 120 minutes ("t" values) as controls. Percentage viability was calculated from the mean of two viable taken at the appropriate time point after the pH was dropped to 3.5 with H_3PO_4.

Table 9.6. The effect of chloramphenicol on the ATR of *M. smegmatis* mc²155.

	% Viability at the lethal pH[a]					
	Adapted[b]		Adapted + Cm[c]		Unadapted	
Time (hrs)	Expt.1	Expt.2	Expt.1	Expt.2	Expt.1	Expt.2
0	100	100	100	100	100	100
1	71	68	27	26	16	17
2	57	56	15	17	9	12
3	40	48	17	12	6	2
4	12	12	4	3	1	1

[a] Percentage viability was calculated from one viable count taken at the appropriate time point after the pH was dropped to 3.5 with H_3PO_4.

[b] The culture was adapted at pH 5.0 for four hours before exposure to pH 3.5.

[c] 200µg/ml chloramphenicol was added thirty minutes before the pH was dropped to 5.0. The culture was then adapted at pH 5.0 for four hours before exposure to pH 3.5.

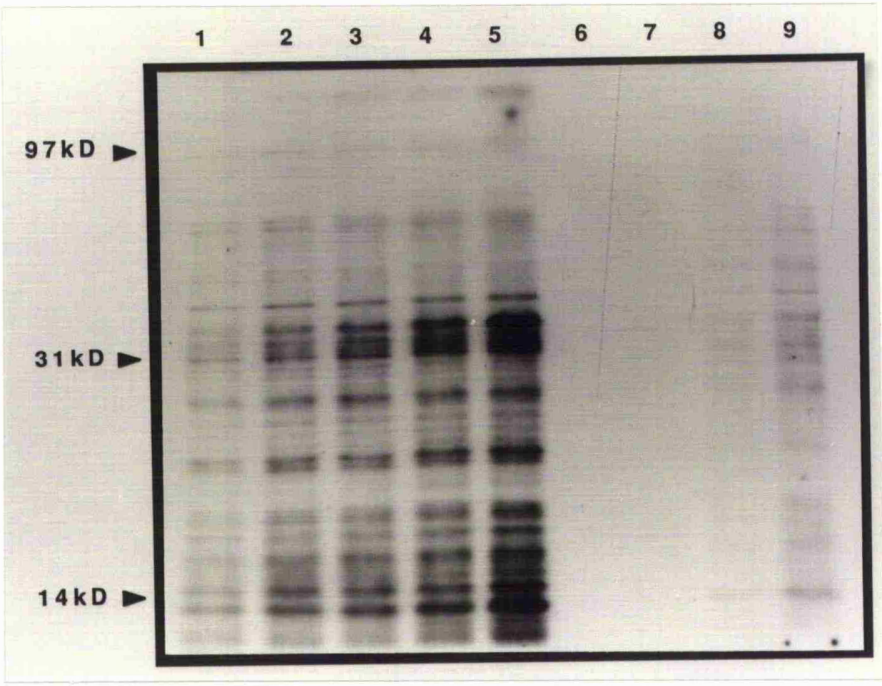


Figure 9.1 Protein profile of *M. smegmatis* at pH 7.6 and pH 5.0.
The protein profile of *M. smegmatis* mc^{2}155 was monitored using [^{35}S]methionine after a shift in external pH from 7.6 to 5.0. Lanes 1, 2, 3, 4, and 5 show proteins labelled 5, 15, 30, 60 and 90 minutes after addition of [^{35}S]methionine to pH 7.6 cultures (positive control), while lanes 6, 7, 8, and 9 show proteins produced 15, 30, 60, and 90 minutes after a shift from pH 7.6 to pH 5.0. Protein molecular weights are displayed in kilodaltons (kD) in the left hand margin.

NB: During the Ph.D. viva examination it was pointed out that flaws in the design of this experiment rendered any conclusions unsound. Two critical points were made: (i) The radiolabelling was carried out in an eppendorf, a vessel that would not allow sufficient aeration. Hence cells would experience both anaerobic and acidic stress. (ii) A control experiment to measure the uptake of [^{35}S] methionine at pH 5 was not performed. Therefore it could be argued that the lack of labelled proteins after the acid shift to pH 5 was due to a decrease/defect in the uptake of the radiolabel rather than a decrease in protein synthesis. Conclusions from the experiment are therefore void and have been removed from the discussion section.

9.2 Construction of the promoter probe shuttle vector pSG10.

The starting point for construction of the promoter probe vector pSG10 was the plasmid pPA3 (Andrew and Roberts, 1993). The backbone of pPA3 was the *E. coli-Mycobacterium* shuttle vector pMV261 (Stover *et al.*, 1991), into which the *Vibrio harveyi luxAB* gene fusion (Hill *et al.*, 1991) had been cloned. Expression of the *V. harveyi luxAB* fusion was under the control of the *M. bovis* BCG *hsp60* gene promoter that allows transcription in both *E. coli* and *Mycobacterium* hosts. To construct the promoter probe vector, the *hsp60* promoter was excised from pPA3 by a *Bam*HI-*Xba*I double digest, and the overhanging termini end-filled using the Klenow fragment of DNA polymerase I (Figure 9.2). Blunt-ended vector molecules were then recircularised by ligating at low DNA concentrations (1-10ng vector DNA/50µl ligation reaction) to produce the vector pPA51. As the removal of the *hsp60* promoter also removed many useful restriction enzyme sites, the *Eco*RI-*Hinc*II fragment of the pUC multiple cloning site was excised from pUC18 by a double digest, then added to ligations containing pPA51 digested with *Eco*RI and *Pvu*II (Figure 9.2). Ligations were transformed into *E. coli* DH5α, then plasmid DNA extracted from transformants and digested with *Bam*HI to identify transformants with the pUC polylinker. One clone, termed pSG10, in which the polylinker fragment was present, was chosen as the promoter probe shuttle vector.

9.3 Expression of luciferase by the acetamidase gene promoter.

9.3.1 Construction of pSG15

Before a promoter probe library in pSG10 could be constructed and examined for environmentally regulated promoters, it was necessary to investigate the efficacy of bacterial luciferase as a reporter of mycobacterial promoter activity. Although the *hsp60* promoter of pPA3 would appear to be an ideal candidate for this, Stover and colleagues (1991) had shown that this promoter was not inducible by environmental stress when present on an extrachromosomal vector. The acetamidase gene of *Mycobacterium smegmatis* is known to be highly inducible when grown using simple

amides, such as acetamide, as sole carbon source. Parish (1993) determined that the chloramphenicol acetyl transferase gene (*cat*) could be used as a reporter of the acetamidase promoter's activity, and that the promoter was inducible when present on an extrachromosomal vector. The acetamidase gene promoter was therefore picked as a model environmentally regulated mycobacterial promoter.

The plasmid pAMI1 (Mahenthiralingam *et al.*, 1993; a gift from Dr. M.J. Colston, NIMR, Mill Hill, London) contains the *M. smegmatis* acetamidase gene and upstream regulatory regions on a 4.3kb *Bam*HI restriction fragment. *Bam*HI digested pAMI1 was electrophoresed through a 0.5% (w/v) preparative agarose gel and the restriction fragment containing the acetamidase gene and regulatory regions excised in an agarose slice. The DNA was recovered from the agarose slice using the Pharmacia Bandprep kit as described in Section 2.5.6. This *Bam*HI restriction fragment was then ligated into the promoter probe vector pSG10 that had been digested with *Bam*HI and dephosphorylated (as described in Section 2.5.3). Ligations were transformed into *E. coli* TG2, selecting for kanamycin resistant transformants. Plasmid DNA was extracted from transformants and analysed by restriction enzyme digestion and agarose gel electrophoresis. One of the clones, termed pSG15, in which the acetamidase gene was present in the correct orientation relative to the luciferase genes, was used for further studies. The construction of pSG15 is detailed in Figure 9.3. Plasmid DNA of the pSG15 construct was electroporated into *M. smegmatis* mc^2155, and kanamycin resistant transformants selected on 7H11 agar containing 20μg/ml kanamycin.

9.3.2 Luciferase activity in pSG15 cultures.

Cultures of *M. smegmatis* mc^2155 [pSG15] were grown in minimal medium with acetamide or succinate as sole carbon source. Samples from each culture were removed at various time points over 7 hours and (i) assayed for bioluminescence as described in Section 8.4.1, and (ii) serially diluted and plated on 7H11 agar containing 20μg/ml kanamycin to determine numbers of viable bacteria in the sample. The results of one such experiment, representative of four, are shown in Figure 9.4. Acetamide-grown cultures produced a high level of light due to read-through from the induced acetamidase gene promoter into the *lux* genes. Induction of luciferase expression began after 2 hours of

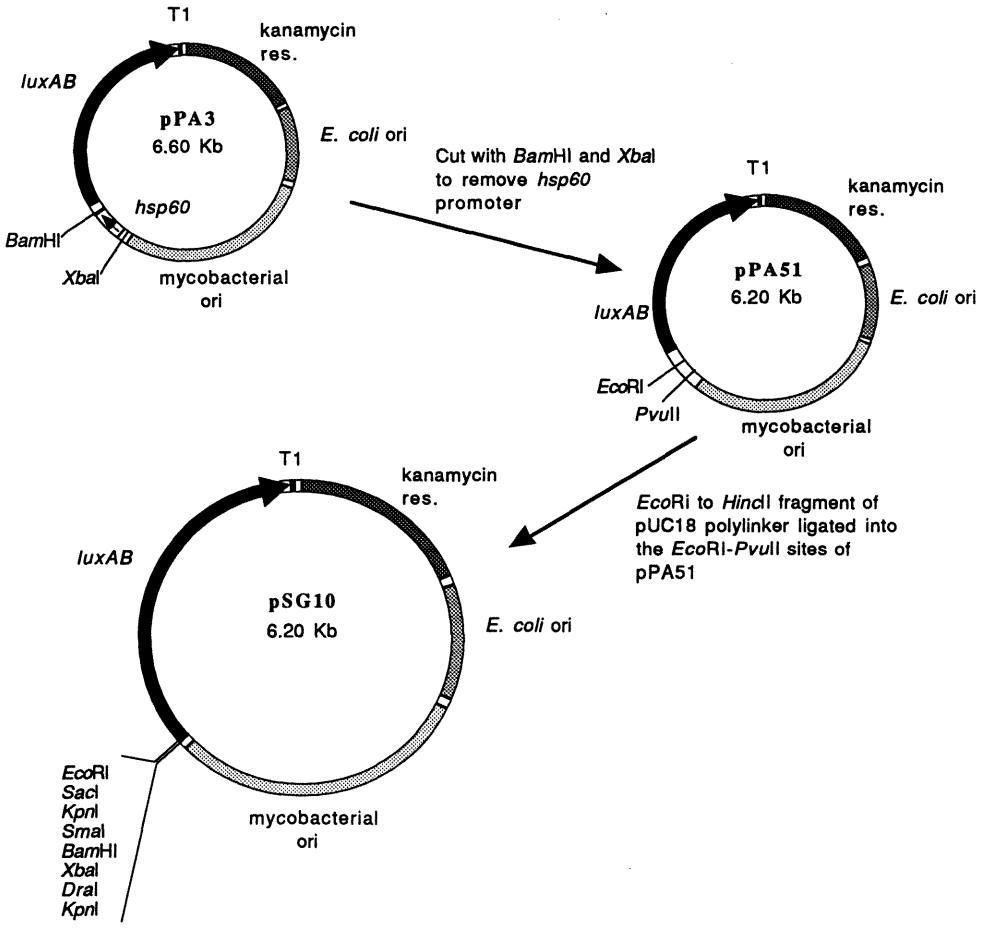


Figure 9.2 Construction of the promoter probe shuttle vector pSG10.
The *hsp60* gene promoter region of pPA3 (Andrew and Roberts, 1993) was removed to create pPA51. The *Eco*RI-*Hinc*II region of the pUC18 polylinker was then ligated into the *Eco*RI-*Pvu*II site of pPA51 to produce pSG10. Abbreviations are kanamycin resistance (kanamycin res.), *Escherichia coli* and mycobacterial origins of replication (*E. coli* and mycobacterial ori respectively), *hsp60* gene promoter (*hsp60*), transcriptional terminator (T1).

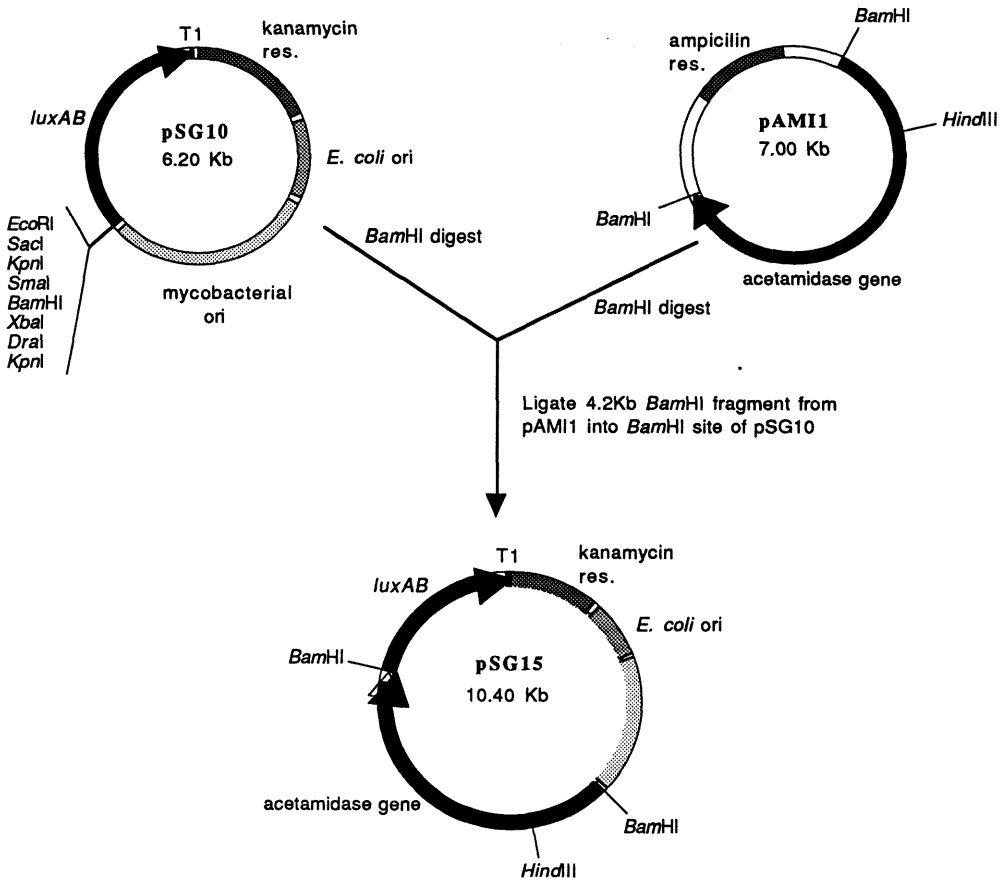


Figure 9.3 Construction of pSG15.

Plasmid pAMI1 (Mahenthiralingam *et al.*, 1993) was digested with *Bam*HI to excise the acetamidase gene and regulatory regions. This 4.2 kb *Bam*HI fragment was ligated into the *Bam*HI site of pSG10 to create pSG15. Abbreviations are as in Figure 9.2.

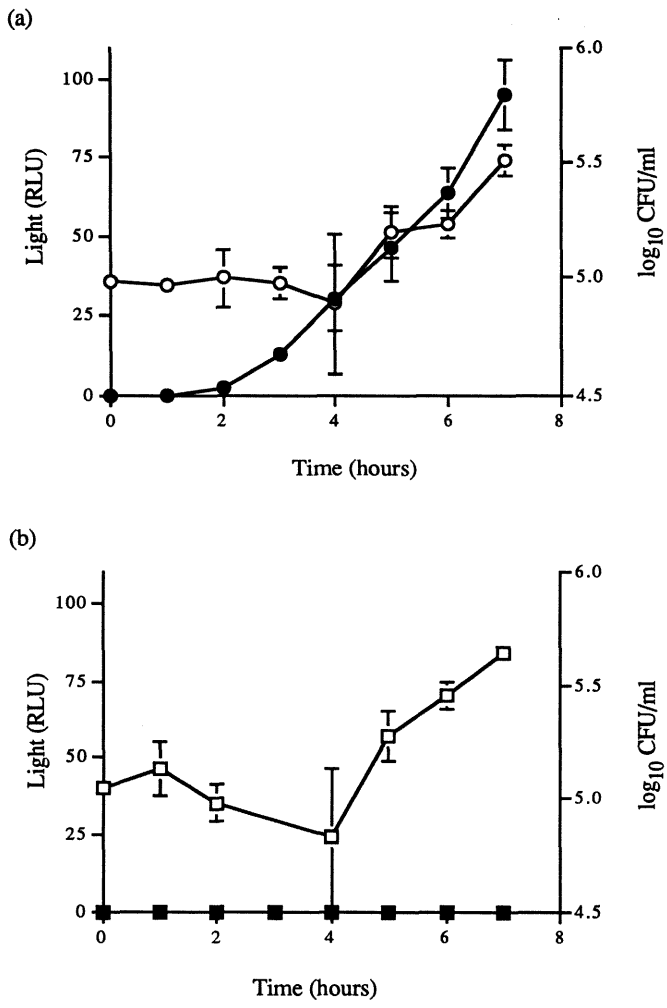


Figure 9.4 Bioluminescence of *M. smegmatis* mc^{2}155 [pSG15] cultures. Cultures were grown in minimal media with (a) acetamide or (b) succinate as the sole carbon source. Cell numbers of both the acetamide- (O) and succinate-(□) grown cultures are given as $\log_{10}$ CFU/ml. Light produced by the acetamide- (●) and succinate- (■) grown cultures is given in relative light units (RLU). Data are the mean ± standard deviation of eight readings from one experiment, representative of four experiments.

culture and after 4 hours light levels were 1000-fold higher from acetamide-grown cultures relative to succinate-grown cultures (30 and 0.03 relative light units respectively). The initial increase in light from acetamide-grown cultures was independent of an increase in cell numbers. By 7 hours, acetamide-cultured cells produced over 1500-fold more light relative to succinate-grown cells (95 and 0.06 relative light units respectively). Differences in light production were not a consequence of differences in growth rates between the cultures, or of total cell numbers. Cell numbers in both succinate and acetamide cultures were approximately equal throughout the 7 hour sampling period.

9.4 Construction and screening of a promoter probe library for acidic pH-regulated promoters.

9.4.1 Construction of a promoter probe library in pSG10.

The starting point for identification of genetic elements responsive to changes in external pH was the construction of a promoter probe library. The experimental strategy was to place the promoterless luciferase genes in pSG10 under the transcriptional control of random sections of *M. smegmatis* genomic DNA to form a promoter probe library, then to transform the library into *M. smegmatis* mc^2155. Library clones would then be exposed to acidic external pH and monitored for bioluminescence. Clones that produced light in response to acid would be expressing luciferase from an acid-inducible promoter contained in the upstream mycobacterial DNA. These clones could then be isolated and the mycobacterial insert analysed to identify the promoter and gene that was driving luciferase expression.

The size-selected DNA that was used for library construction was in the 2-3kb range; this size of insert was chosen for two reasons. First, Parish (1993) had found that for the acetamidase gene to show full induction (on a plasmid) in response to acetamide, 3kb of DNA upstream from the start codon is needed; with less of this upstream DNA present, the gene showed lower levels of induction. This requirement for upstream DNA may be unique to the acetamidase promoter, and could be due to a number of factors such as the presence of upstream enhancer elements. However, since so little is known about

mycobacterial promoters and due to the reported difficulties that other investigators have had in inducing mycobacterial promoters while on plasmids (Stover *et al.*, 1991; J.J. McFadden, University of Surrey, pers. comm.), we felt this observation with acetamidase may be significant. The presence of control regions upstream from mycobacterial promoters that are necessary for maximum promoter induction may be a general feature of mycobacterial gene regulation. Therefore, relatively large fragments of mycobacterial DNA (2-3kb) were used in library construction compared to other mycobacterial promoter probe libraries (i.e. Gupta *et al.* (1993) used fragment sizes of 50-800bp). The second reason for using a large insert size was to increase the chances of isolating the gene usually linked to the promoter from any positive inserts. The sequence of any gene fragment attached to the promoter could then be translated into protein and used to search the protein sequence databases, possibly revealing why the organism switches the gene on in response to acid stress.

Genomic DNA from *M. smegmatis* mc^2155 was partially digested with the restriction endonuclease *Sau*3AI and digests analysed by agarose gel electrophoresis. Agarose gel slices, containing restriction fragments in the 2-3kb range, were excised, then the DNA recovered using the Pharmacia Bandprep kit and ligated to *Bam*HI digested, dephosphorylated pSG10. Electroporation of ligations into *M. smegmatis* mc^2155 gave very low transformation efficiencies, approximately 10-20 transformants/μg DNA. This efficiency of transformation was too low, preventing the direct transformation of the library into *M. smegmatis*. To circumvent this, a transformation protocol using *E. coli* as an intermediate host was employed. Ligations were transformed into *E. coli* DH5α, selecting for kanamycin resistance. Transformation efficiencies in *E. coli* DH5α were approximately 10^6/μg of DNA. Plasmid DNA was purified from 10 transformants, digested with the restriction enzyme *Kpn*I (*Kpn*I sites flank the *Bam*HI site in pSG10) and analysed by agarose gel electrophoresis. Figure 9.5 shows the results of this analysis, which suggested that the library was approximately 70% recombinant. Although 70% recombinant appears low for a useful library, it was felt sufficient to work with since the aim was not to identify every pH regulated promoter, but rather to investigate the efficacy of luciferase

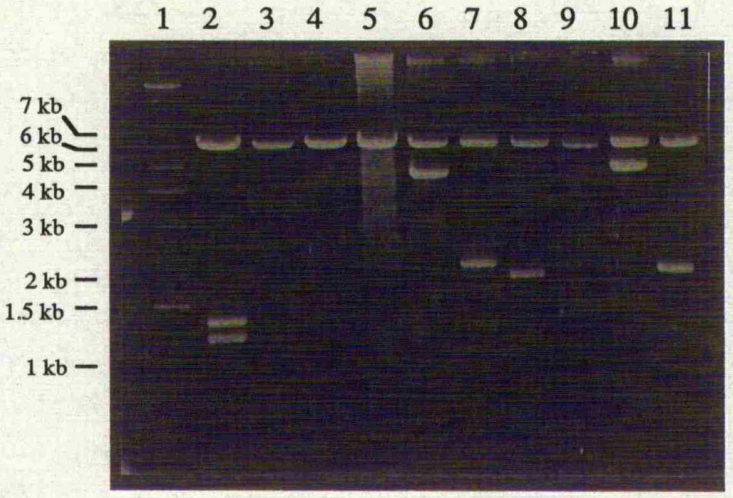


Figure 9.5 Restriction enzyme digests of promoter-library clones.
Plasmid DNA was prepared from 10 random promoter probe library clones and digested with the restriction enzyme *Kpn*I (*Kpn*I sites flank the *Bam*HI site used for library construction, so KpnI digestion will excise any insert). Lanes 2-11, *Kpn*I digested plasmid DNA from library clones. Lane 1, 1kb DNA size marker (GIBCO-BRL), fragment sizes are shown in the left hand column in kilobases (kb).

as a model system for identifying mycobacterial acid-regulated promoters.

Transformants on the original *E. coli* DH5α library transformation plates were washed off with 2-3ml of sterile PBS. Washings from 8-12 agar plates were pooled such that each pool contained about 10,000 transformants. A pool of transformants was then used to inoculate 1L of L-broth containing 20μg/ml kanamycin. After an incubation period of 1 hour at 37°C, plasmid DNA was purified from the culture. Aliquots of this stock of library DNA were then used to transform *M. smegmatis* mc²155, selecting for kanamycin resistance. These transformations generally gave efficiencies of 10^3 transformants/μg of DNA. Transformants of *M. smegmatis* mc²155 were stored on L-agar, containing 20μg/ml kanamycin, at 4°C.

9.4.2 Screening of the promoter probe library for acid-regulated promoters.

Initial attempts to screen the library in *M. smegmatis* mc²155 for clones that were luminescent at acidic pH were based on exposure of library transformants to decanal vapour then placing the agar plates on X-ray film (Cronex, DuPont). To achieve this, transformants were replica-plated onto L-agar at pH 7.6 (adjusted using 2M NaOH) and pH 5 (adjusted using 2M H_3PO_4), then incubated at 37°C for 48 hours. A few drops of a 1% (v/v) solution of decanal in ethanol were added to the lids of the Petri plates, plates left at room temperature for 5 minutes, then placed on X-ray film for 2-5 minutes. The exposed film was then developed, with black dots revealing the positions of bioluminescent colonies. It was hoped that this method would identify colonies that were luminescent at pH 5 but were dark at pH 7.6, indicating that luciferase expression in these clones was induced at acidic pH.

Despite screening approximately 700 clones by this method, no clones that showed reproducible induction at acidic pH could be identified. One of the pitfalls of this screening method was that levels of bioluminescence produced by colonies was dependant on colony size and time elapsed since exposure to decanal, rather than external pH. Indeed, colonies at pH 5 generally showed a lower level of bioluminescence compared to replica colonies at pH 7.6. Secondly, the microenvironment of a colony is highly complex, with bacteria exposed to a wide range of

stresses such as oxygen limitation, nutrient availability and pH fluctuations. It would therefore be difficult to attribute the luciferase activity of a colony exclusively to the pH of the growth medium.

A screening protocol that would allow the pH of the growth medium to be altered while all other variables remained relatively constant was needed. The parameters that needed to be considered were the method of administering the drop in pH, the most suitable length of time to leave the cultures to adapt at pH 5 before measuring luminescence, and the choice of medium (i.e. 7H9 or L-broth) in which to grow the library clones. The initial approach to screening the promoter probe library was to grow the clones (4 replicates) in 200µl of 7H9 medium at pH 7.6 in 96-well microtitre plates for 48 hours, then dispense 50µl of 0.1M H_3PO_4 to half of the replicates to drop the pH to 5 (the method for dropping the pH to 5 was determined by adding varying amounts of 0.1M H_3PO_4 to 200µl of culture and then measuring the resulting culture pH with pH paper). The plates were incubated at 37°C for 60 minutes, and bioluminescence measured for 15 seconds in a Luminoskan luminometer following the addition of 10µl of a 1% (v/v) decanal solution in ethanol. Using this method 100 clones were screened, but the results were disappointing. All cultures that had been dropped to pH 5 produced very little or zero luminescence, despite some of the replica cultures that were maintained at pH 7.6 generating very high levels of light. This seemed to point to a fundamental flaw in the screening procedure. The most probable fault was the procedure for altering the culture pH, in that high local concentrations of acid may be produced after the addition of the 50µl volume of 0.1M H_3PO_4 to 200µl of culture, killing or at least damaging the bacilli. A method for uniformly exposing the culture to pH 5 was needed.

The screening method that was used was developed with *M. smegmatis* mc^2155 [pPA3] as the model organism. Media used for cultivating *M. smegmatis* mc^2155 [pPA3] were L-broth and 7H9, since they are known to support the growth of *M. smegmatis* and are commonly used in the laboratory. The pH of the growth medium was adjusted to pH 6.6 (optimum for mycobacterial growth) and 5; pH 6.6, not pH 7.6, was chosen as the growth pH for these experiments since the purpose was to determine the effect of an acidic external pH on the activity of luciferase, rather than of an acid-shock. Cultures of *M. smegmatis* mc^2155 [pPA3] were grown in 100ml of L-broth or 7H9

broth (both containing 0.05% (v/v) Tween 80 and 20μg/ml kanamycin) at pH 6.6 for 48 hours. Cultures were then split into two 50ml fractions, and the pH of one fraction dropped to pH 5 using filter-sterilised 2M H_3PO_4. Cultures were incubated at 37°C (with shaking at 200 rpm) and assayed for bioluminescence and colony forming units over 6 hours (Figures 9.6 and 9.7). In the L-broth cultures, light production by cultures exposed to pH 5 was initially lower than cultures maintained at pH 6.6. After a period of 2 hours, the pH 5 cultures appeared to adapt to the new growth conditions and increased their generation of light. The 7H9 cultures exposed to pH 5 produced more light during the course of the experiment than the pH 6.6 culture, with an increase in light in both cultures after 4 hours. Bacterial numbers in pH 5 cultures remaining roughly equivalent to pH 6.6 cultures during the initial 4 hours of measurement in both L-broth and 7H9 media. These results suggested that bioluminescence should be determined 4 hours after the exposure to pH 5.0, as this would allow maximum time for adaptation of the culture at pH 5.0 without a considerable change in bacterial numbers. The medium of choice was 7H9, as it was felt that the phosphate buffer present in this medium would allow for greater buffering of culture pH compared to L-broth.

Library clones were grown in 7H9 medium (containing 20μg/ml kanamycin) at pH 7.6 for 48 hours. These cultures were then used to inoculate (in duplicate) 200μl of fresh 7H9 medium at both pH 7.6 and pH 5.0 in black 96-well fluoroplates (thermoequilibrated to 37°C). Cultures were incubated at 37°C for 4 hours, after which time bioluminescence was measured for 15 seconds following the addition of 10μl of 1% (v/v) decanal solution. This screening technique had the advantages over an agar-based method that (i) a quantitative value of luciferase induction at pH 5 over expression at pH 7.6 could be obtained, (ii) culture pH could be accurately determined during the assay period, (iii) bioluminescence in broth cultures could be adjusted to account for bacterial numbers.

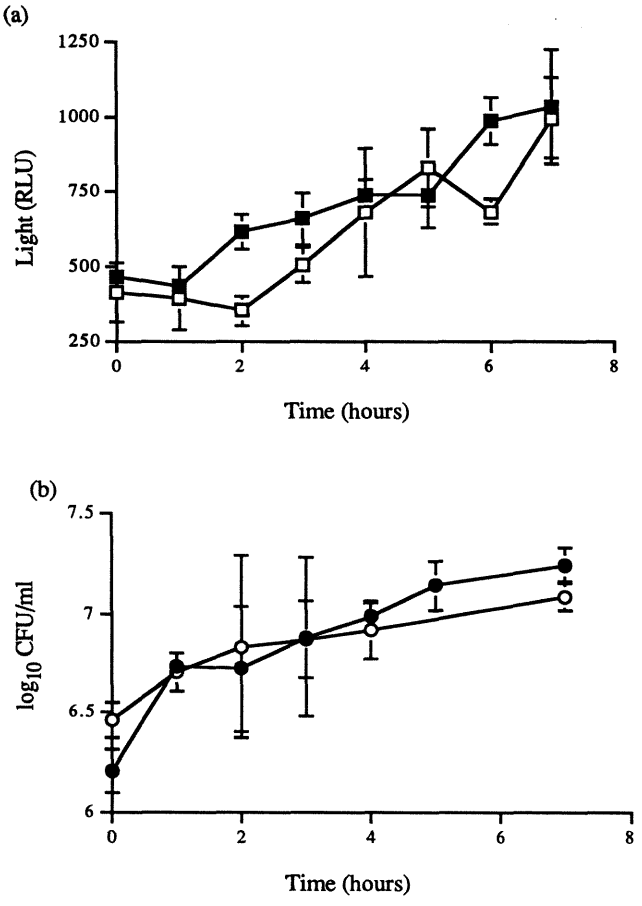


Figure 9.6 Light production by *M. smegmatis* mc²155 [pPA3] grown in L-broth.

The graphs show the effect of pH 5.0 on bioluminescence and viability of *M. smegmatis* mc²155[pPA3]. Cultures of *M. smegmatis* mc²155 [pPA3] were grown in L-broth at pH 6.6 for 48 hours, then divided into two aliquots, with the pH of one aliquot reduced to 5.0 using 2M H_3PO_4. Graph (a) depicts light production from cultures grown at pH 6.6 (■) and pH 5.0 (□) over 7 hours. Graph (b) presents $\log_{10}$ CFU/ml of pH 6.6 (●) and pH 5.0 (○) cultures over the same period. Bioluminescence is given in relative light units (RLU). Bacterial numbers are given as CFU/ml. Data are the means ± standard deviation of four readings from one experiment.

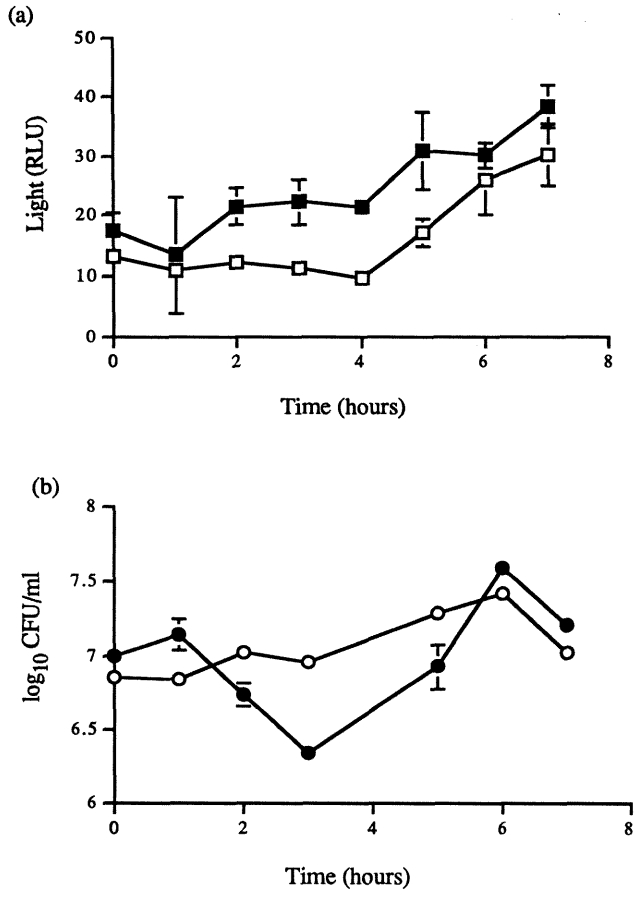


Figure 9.7 Light production by *M. smegmatis* mc^{2}155 [pPA3] grown in 7H9 broth.

The graphs show the effect of pH 5.0 on bioluminescence and viability of *M. smegmatis* mc^{2}155[pPA3]. Cultures of *M. smegmatis* mc^{2}155 [pPA3], grown in 7H9 medium (plus ADC) at pH 6.6 for 48 hours, were divided into two aliquots. The pH of one aliquot was then reduced to 5.0 using 2M H_3PO_4. Graph (a) depicts light production from cultures grown at pH 6.6 (□) and pH 5.0 (■) over 7 hours. Graph (b) presents $\log_{10}$ CFU/ml of pH 6.6 (○) and pH 5.0 (●) cultures over the same period. Bioluminescence is given in relative light units (RLU). Data are means ± standard deviation of four readings from one experiment.

Using the broth culture screening method, 451 clones were analysed. An example of a typical set of results obtained from a round of screening is represented in Figure 9.8. Clones were considered "induced" if luciferase levels at pH 5.0 were at least twice those at pH 7.6. A number of clones that fitted this criterion were identified in the first round of screening, but subsequent rounds of repeat screening revealed them to be false positives. One clone that showed a reproducible increase in bioluminescence at pH 5.0 relative to pH 7.6 was identified. This clone, termed pSGA197, produced 15-20 fold more light at pH 5.0 than at pH 7.6 (Figure 9.8). As time was limited, it was decided to concentrate on the pSGA197 clone rather than screening the library for more acid-inducible clones.

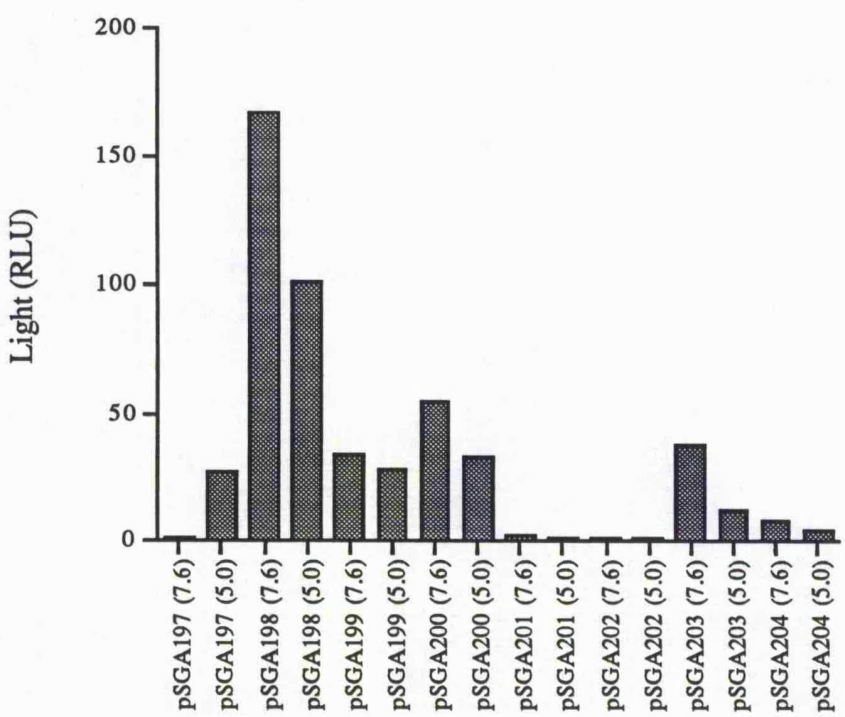


Figure 9.8 Screening of promoter library clones for increased bioluminescence in response to acidic pH of medium.

Library clones were grown in 7H9 medium (plus ADC) at pH 7.6, then inoculated in duplicate into 7H9 medium (plus ADC) at pH 7.6 and pH 5.0. Clones were allowed to adapt at 37°C for 4 hours, then bioluminescence determined in a luminometer after the addition of 10µl of a 1% (v/v) decanal solution in ethanol. Bioluminescence was measured in relative light units (RLU) and data presented are the mean of two readings.

9.5 Analysis of pSGA197.

9.5.1 Expression of bioluminescence by *M. smegmatis* [pSGA197].

9.5.1.1 Time course of luminescence induction.

Cultures (10ml) of *M. smegmatis* mc^2155 [pSGA197] were grown for 36-48 hours (mid to late log phase) at 37°C in 7H9 medium (pH 7.6) containing 20μg/ml kanamycin. Two millilitres of this culture was used to inoculate 50ml of 7H9 medium (plus ADC and 20μg/ml kanamycin) at pH 7.6 and 5.0 (thermoequilibrated to 37°C), then cultures incubated at 37°C with shaking. Samples of the pH 7.6 and pH 5.0 cultures were removed at timed intervals to assay for bioluminescence, pH and bacterial numbers. Figure 9.9 presents the results of three such assays. Graphs (a) and (b) show light production from *M. smegmatis* mc^2155 [pSGA197] cultures grown at pH 7.6 for 36 hours, then inoculated into fresh 7H9 medium at both 7.6 and 5.0; graph (c) shows light production from a culture that was grown for 48 hours before inoculating into fresh medium. Induction of luciferase activity was rapid, with light levels generally 5 to 10-fold higher in pH 5.0 cultures after 60 minutes. After 4 hours, light levels were 15- to 20-fold higher in cultures exposed to pH 5.0 than cultures maintained at pH 7.6. The pH of the culture did not alter significantly from the initial pH, with culture pH readings staying within 0.2 pH units of the original medium pH during the 4 hour assay period. Differences in bioluminescence were not a result of differences in bacterial numbers, with cell numbers remaining approximately equal during the 4 hour assay period.

9.5.1.2 Response to HCl, Heat and Peroxide.

During the screening procedure, the pH of the culture medium was dropped using 2M H_3PO_4. To determine whether induction of luciferase by *M. smegmatis* [pSGA197] was a response to phosphoric acid rather than to acidic pH *per se*, a batch of 7H9 medium was prepared and 50ml samples were acidified using 2M H_3PO_4 or 2M HCl. Samples (200μl) of 7H9 medium at pH 7.6, 5.0 (HCl) and 5.0 (H_3PO_4) were dispensed into black 96-well fluoroplates, thermoequilibrated to 37°C, then inoculated with 20μl of mid-log cultures of *M. smegmatis*

[pSGA197]. Fluoroplates were incubated at 37°C for 4 hours, then bioluminescence determined as before (Section 7.4). The level of luciferase induction was found to be independent of the acid used (Figure 9.10), with pH 5.0 cultures producing approximately 20-fold more light than pH 7.6 cultures.

To determine whether luciferase expression could be induced by other environmental stresses, *M. smegmatis* [pSGA197] was exposed to heat and peroxide (Figure 9.11). A culture of *M. smegmatis* [pSGA197], grown at pH 7.6 and inoculated into 7H9 medium at pH 5.0 as before, was used as a positive control. Heat-shock was performed on 36 hour old cultures of *M. smegmatis* [pSGA197] by shifting the culture to 45°C for 30 minutes, then returning the cultures to 37°C (method used by Stover *et al.* (1991) to induce expression of the *M. bovis* BCG *hsp60* gene promoter). Prior to the heat shock, a sample of culture was briefly sonicated, serially diluted and plated onto L-agar containing 20μg/ml kanamycin; light production from 200μl culture samples was also measured. Bioluminescence was determined directly after the return to 37°C from 45°C (t=0), and again 3 hours later (t=3 hours). At t=0, light levels were effectively zero probably due to destruction of the luciferase enzyme by exposure to 45°C (Meighen, 1991). Heat shock did not induce the expression of luciferase, with light production returning to pre-shock levels after 3 hours (Figure 9.11). Response to oxidative stress was investigated by measuring light production from cultures of *M. smegmatis* [pSGA197] after addition of H_2O_2 to a final concentration of 0.02% (v/v) (method used by Stover *et al.* (1991) to induce expression of the *M. bovis* BCG *hsp60* gene promoter). Viable bacterial counts were performed and light levels of cultures measured prior to addition of H_2O_2. Bioluminescence was determined 3 hours after the addition of H_2O_2, and was found to be the same as pre-shock light levels (Figure 9.11). Therefore, as with heat shock, there was no evidence of induction of luciferase in response to oxidative stress.

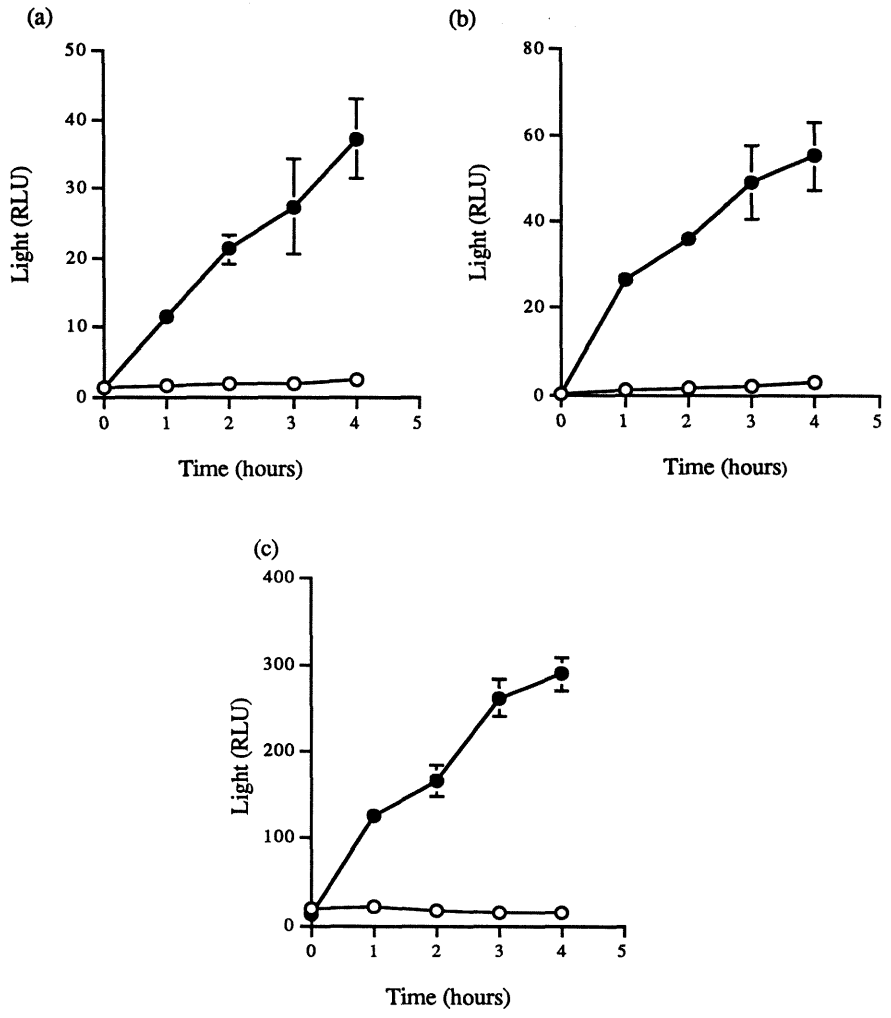


Figure 9.9 Time course of luciferase induction by *M. smegmatis* [pSGA197]. *M. smegmatis* mc^{2}155 [pSGA197] was grown in 10ml volumes of 7H9 medium (containing ADC and 20μg/ml kanamycin) at pH 7.6 for 36 hours (Graphs a and b) or 48 hours (Graph c). A 2ml fraction of culture was used to inoculate 50ml of fresh 7H9 medium at pH 7.6 and pH 5.0. Samples were removed from the cultures at one hour intervals, and light production measured in a luminometer over 15 seconds after the addition of 10μl of 1% (v/v) decanal. Light production at pH 7.6 (O) and pH 5.0 (●) is presented in relative light units (RLU). Data points in each graph are the means ± standard deviation of 4 readings.

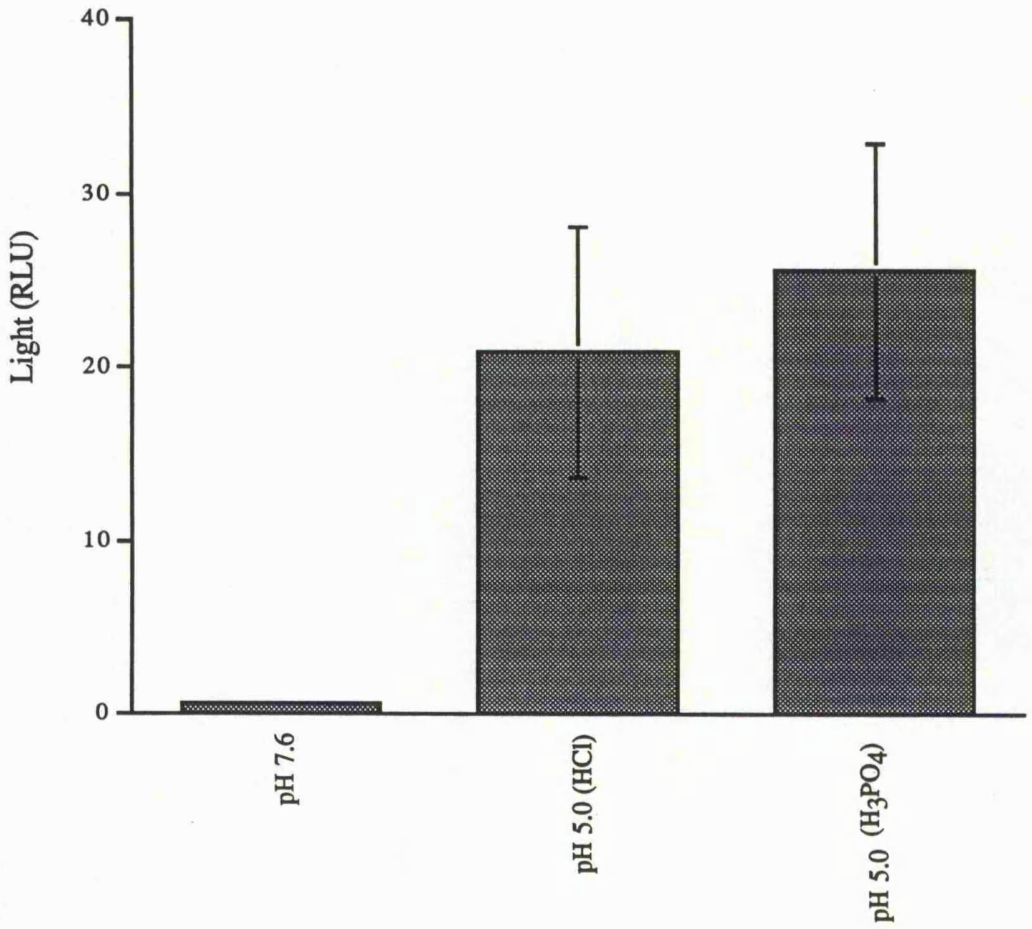


Figure 9.10 Light production by *M. smegmatis* [pSGA197] in response to medium acidified using 2M H_3PO_4 or 2M HCl.

7H9 cultures of *M. smegmatis* [pSGA197], grown at pH 7.6 for 36 hours, were used to inoculate 7H9 medium at pH 7.6 and pH 5.0 (acidified using either 2M HCl or 2M H_3PO_4). Cultures were allowed to adapt for 4 hours, then bioluminescence measured in a luminometer following the addition of 10µl of 1% (v/v) decanal. Light production is given in relative light units (RLU). Data are the means ± standard deviation of 4 readings from one experiment.

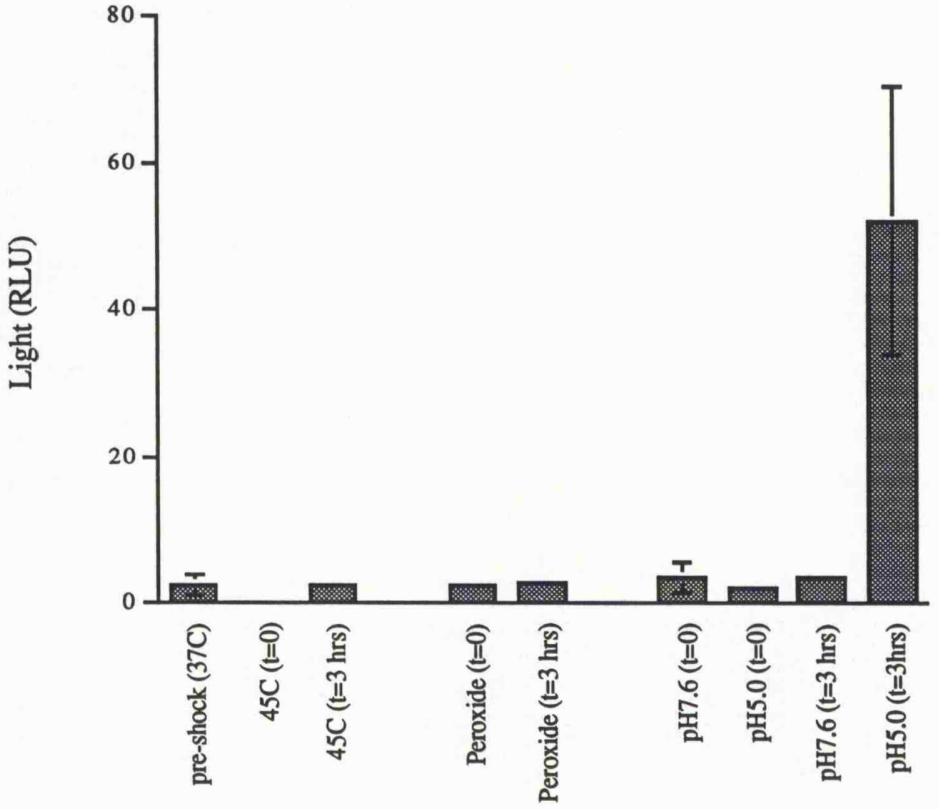


Figure 9.11 Response of *M. smegmatis* [pSGA197] to oxidative and heat stress.

M. smegmatis mc^2155 [pSGA197] was cultured in 7H9 medium (plus ADC and 20μg/ml kanamycin) at pH 7.6 for 36 hours. To perform a heat shock, cultures were shifted to 45°C for 30 minutes, then returned to 37°C. Light production and CFU/ml were determined immediately before the heat shock, after the return to 37°C, and again after 3 hours. To produce an oxidative stress, 3μl of 3% (v/v) H_2O_2 solution (peroxide) was added to 5ml of culture, then light production and CFU/ml determined immediately and again after 3 hours. Cultures exposed to pH 5.0 were used as positive controls. Bioluminescence is presented as relative light units (RLU)/10^5 bacteria. Data are the means ± standard deviation of three experiments.

9.5.1.3 Induction of luciferase over a range of pH values.

7H9 medium was adjusted to pH values between 7.6 and 3.0, using 2M H_3PO_4 to drop the pH and 2M NaOH to raise the pH, in 0.5 pH unit increments. Samples (200µl) of the media were dispensed into fluoroplates, equilibrated to 37°C, and inoculated with 20µl of *M. smegmatis* [pSGA197] culture that had been grown at pH 7.6 for 48 hours. Fluoroplates were incubated at 37°C for 4 hours, after which time bioluminescence was determined by addition of 10µl of 1% (v/v) decanal solution in ethanol and measuring light production over 15 seconds (Section 8.4). This experiment was repeated twice, and the results are presented in Figure 9.12. Figure 9.12(a) shows light production using approximately $7x10^3$ bacilli per well, while Figure 9.12(b) was generated using approximately $5x10^4$ bacilli per well. Light production is given as relative light units (RLU), with data representing the means ± standard deviation of 8 readings. Bioluminescence was directly related to the acidity of the medium, with light increasing as the pH dropped. At a pH value of approximately 3.5 light production decreased, probably as the drop in external pH from 7.6 to 3.5 was too great for the organism to maintain an internal pH at which luciferase could maximally function.

9.5.2 Expression of luciferase by *E. coli* [pSGA197]

Some mycobacterial promoters fail to function in *E. coli*, suggesting that they may be unrecognised by the *E coli* transcription machinery (see Section 1.3.2.1). Analysis of mycobacterial promoter sequences may reveal why this is so. As the promoter probe library had been constructed in a *Mycobacterium-E. coli* shuttle vector, we could determine whether (a) the mycobacterial promoter locus of pSGA197 was recognised by *E. coli*, and (b) whether luciferase expression was induced in *E. coli* in response to external pH changes. Plasmid DNA was first purified from *M. smegmatis* [pSGA197] using the method described in Section 2.3.2. A sample of the plasmid preparation was then used to transform *E. coli* DH5α, selecting for kanamycin-resistant transformants (Section 2.4.1.2). Plasmid DNA was extracted from transformants, then analysed by restriction enzyme digestion and agarose gel electrophoresis. Transformants contained a plasmid of approximately 8.5kb in size that released a 2.2kb fragment on digestion with *Kpn*I (*Kpn*I sites flank the *Bam*HI restriction enzyme site that was

originally used for library construction, Section 9.3.1). Plasmid DNA purified from *E. coli* [pSGA197] was then used to transform *M. smegmatis* mc^2155, selecting for kanamycin resistance. Mycobacterial transformants were grown in 7H9 broth for 36 hours, then inoculated into 7H9 medium at pH 7.6 and 5.0. Transformants induced luciferase expression in response to the decrease in external pH as before, ensuring that the plasmid recovered from the *E. coli* transformants was functional and had not undergone any genetic re-arrangements.

L-broth was adjusted to pH 7.6 with 2M NaOH, or to pH 5.0 using 2M H_3PO_4. A 500µl sample of a 12 hour culture of *E. coli* [pSGA197] was used to inoculate 50ml of L-broth at pH 7.6 and 5.0 (pre-warmed to 37°C) with the inoculation point taken as t=0, then the cultures incubated at 37°C with shaking. Culture samples (2-3 ml) were removed at 60 minute intervals and assayed for bioluminescence, pH and bacterial numbers (Section 8.4.3). The results shown in Figure 9.13 are the mean ± standard deviation of three such experiments. Although the *E. coli* [pSGA197] cultures were capable of generating light, the constitutive levels of bioluminescence were lower than *M. smegmatis* [pSGA197] cultures (e.g. immediately after the inoculation of fresh medium at pH 7.6, cultures of *E. coli* [pSGA197] produced <0.1 RLU/ $2x10^6$ CFU, while pH 7.6 cultures *M. smegmatis* [pSGA197] produced approximately 1.5-2 RLU/10^6 CFU). This may reflect inefficient transcription of the luciferase genes from the mycobacterial promoter in *E. coli*. However, light production did increase in cultures exposed to pH 5 relative to cultures maintained at pH 7.6 over the first 2 hours of measurement (0.06 ± 0.1 RLU in pH 7.6 cultures compared to 0.38 ± 0.04 RLU in pH 5.0 cultures). This would suggest that the *E. coli* transcriptional machinery was capable of increasing transcription from the mycobacterial promoter locus in response to an acidic external pH. The level of light produced by the pH 7.6 cultures increased significantly after 2 hours, rising to a level comparable with that produced by cultures exposed to pH 5.0 (1.09 ± 0.13 RLU and 1.21 ± 0.11 RLU respectively). As cell numbers and rate of growth were equal in both the pH 7.6 and pH 5.0 cultures, one explanation for the increase in light production by pH 7.6 cultures was the drop in external culture pH from 7.6 to approximately 6.9. This decrease in external pH may have triggered an increase in transcription from the mycobacterial promoter locus, leading to a rise in light production.

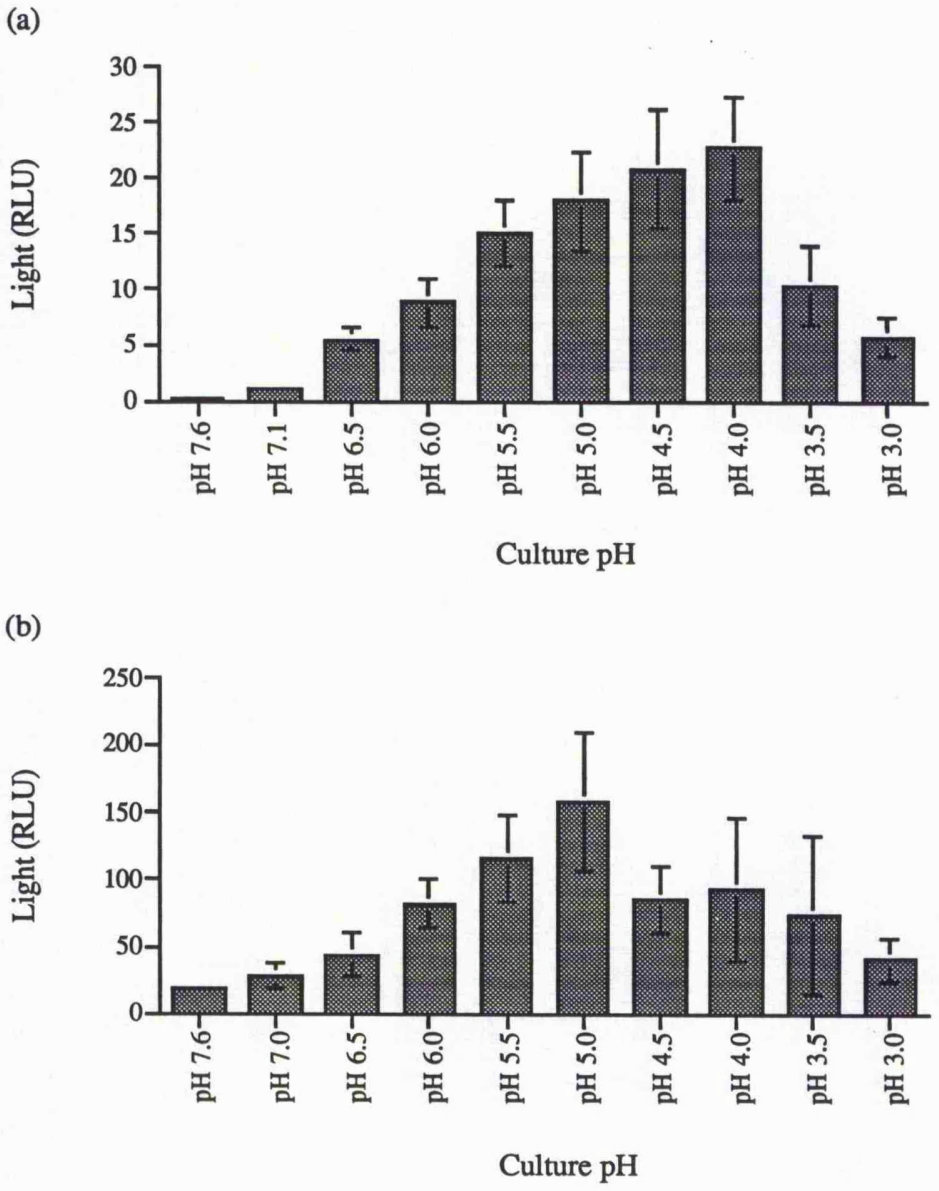


Figure 9.12 Light production by *M. smegmatis* [pSGA197] over a pH range.

Graphs (a) and (b) show light production by *M. smegmatis* [pSGA197] over a pH range from 7.6 to 3.0. Cultures of *M. smegmatis* [pSGA197], grown in 7H9 medium (plus ADC and 20µg/ml kanamycin) at pH 7.6 for 36 hours, were used to inoculate 7H9 medium at pHs from 7.6 to 3.0, left to adapt for 4 hours at 37°C, then bioluminescence measured in a luminometer after adding 10µl of a 1% (v/v) decanal solution in ethanol. Graph (a) shows light production using approximately $7x10^3$ bacilli per well, while graph (b) was generated using approximately $5x10^4$ bacilli per well. Data are the means ± standard deviation of 8 readings.

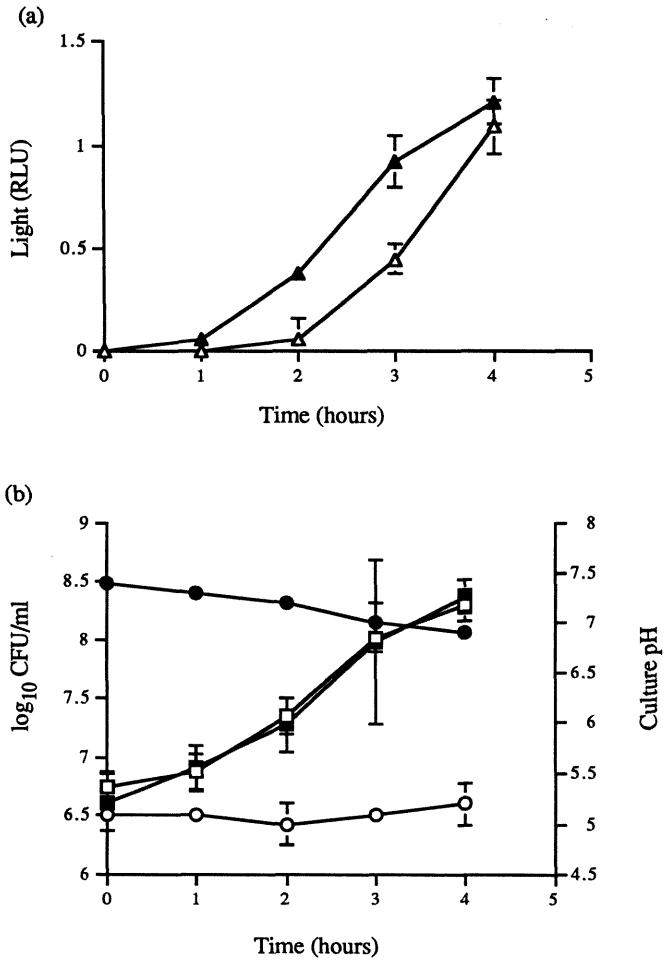


Figure 9.13 Light production by *E. coli* [pSGA197].
Cultures of *E. coli* DH5α [pSGA197] were examined for their ability to induce bioluminescence in response to external acidic pH. An overnight culture of *E. coli* [pSGA197] was used to inoculate L-broth at pH 7.6 and 5.0, then cell numbers, culture pH and bioluminescence measured over a 4 hour period. Graph (a) shows light production by cultures exposed to pH 7.6 (△) or pH 5.0 (▲), with light measured in relative light units (RLU). Graph (b) shows the alterations in (i) culture pH in pH 7.6 (●) and pH 5.0 cultures (○) and (ii) bacterial numbers in $\log_{10}$ CFU/ml from pH 7.6 (■) and pH 5.0 (□) cultures. Data are the means ± standard deviation of 3 experiments, with four readings in each experiment.

9.6 Analysis of the pSGA197 DNA insert.

9.6.1 Subcloning and mapping the pSGA197 insert.

To facilitate the sequencing of the *M. smegmatis* insert present from pSGA197 it was necessary to subclone the fragment into pUC19. *Escherichia coli* DH5α [pSGA197] had been generated previously (Section 9.5.2) and plasmid DNA was extracted by the usual method (Section 2.3.1.1). Small scale digests (using 0.5μg DNA) of the purified pSGA197 DNA, performed with *Eco*RI and *Xba*I, released a fragment of approximately 2.2kb. The digestions were repeated, this time using approximately 10μg of pSGA197 DNA, then electrophoresed through 0.7% (w/v) agarose gels. Gel slices containing the 2.2kb fragment were excised, and DNA purified from the agarose using the Pharmacia Sephaglas Bandprep kit (Section 2.5.6).

pUC19 plasmid DNA was digested with *Eco*RI and *Xba*I. The efficiency of digestion was checked by performing self-ligations with the digested DNA, then transforming into *E. coli* DH5α. Low numbers of transformants (approximately 20-50 transformants/μg plasmid DNA) indicated that the plasmid was sufficiently digested. Ligations were performed as detailed in Section 2.5.4, using 50-100ng of pUC19-digested DNA with 200-400ng of the 2.2kb *M. smegmatis* fragment isolated from pSGA197. Ligation reactions were incubated at 14°C for 12-16 hours, then ethanol precipitated, resuspended in nanopure water, and transformed by electroporation into *E. coli* DH5α, selecting for ampicillin resistance. Plasmid DNA was extracted from white transformants and analysed by restriction enzyme digestion and agarose gel electrophoresis. One clone that had the 2.2kb fragment present, termed pOP-1, was picked for further analysis.

To generate a restriction map of the 2.2kb insert, a range of restriction enzymes were used in single and double digests of pOP-1. The map deduced from these digests is detailed in Figure 9.14. Sites for the restriction enzymes *Eco*RI, *Hind*III, *Kpn*I, *Pst*I, and *Xba*I were not found, while *Sma*I, *Sst*I and *Sph*I had unique recognition sites, and *Hinc*II multiple sites, in the insert. This information was used to generate a set of subclones. Digests of pOP-1 were performed using *Hinc*II, *Sma*I, *Sph*I and *Sst*I, and the products electrophoresed through 0.7% (w/v) agarose gels. Restriction fragments consisting of the pUC19 "backbone" attached to the remaining insert DNA were excised in

agarose slices from each digest lane and the DNA extracted using the Pharmacia Sephaglas Bandprep kit. The extracted plasmid DNA was then self-ligated at low DNA concentrations to re-circularise the vector, then transformed into competent *E. coli* DH5α. Using this strategy, a set of nested deletions of the pOP-1 insert was constructed (Figure 9.14). These clones were termed pOP-2 (generated by *Hinc*II digestion and religation), pOP-3 (*Sma*I), pOP-4 (*Sph*I) and pOP-5 (*Sst*I).

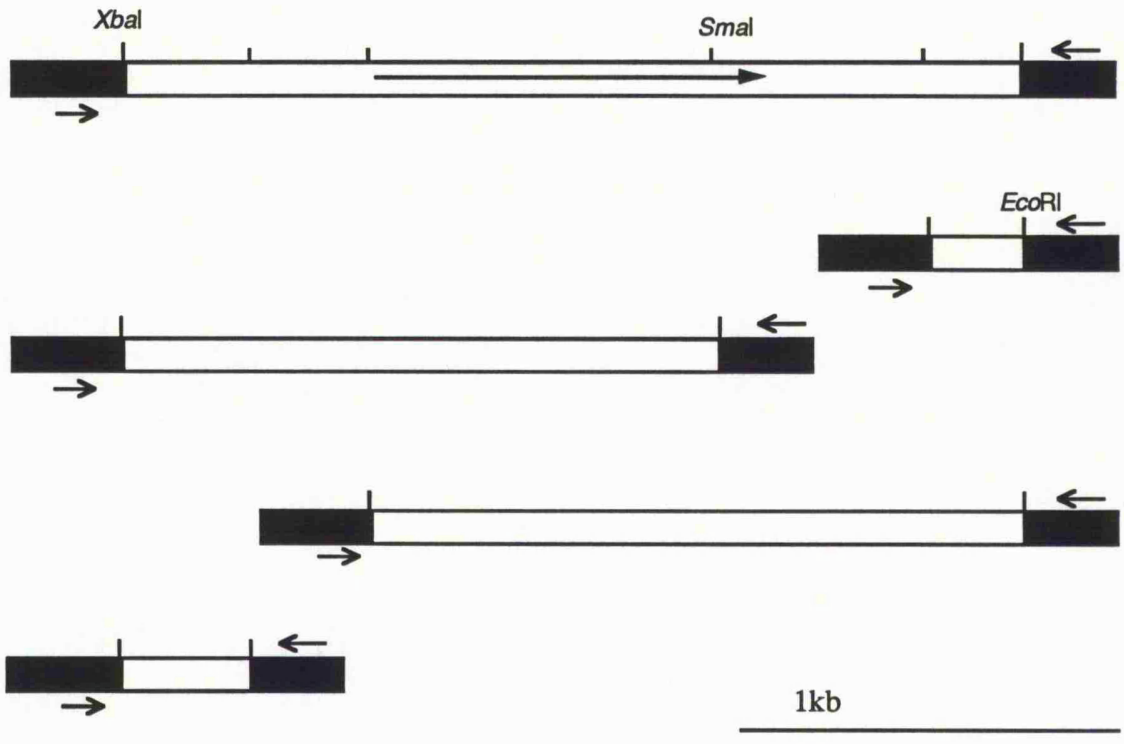


Figure 9.14 Restriction map and pOP subclones of the pSGA197 insert. The entire 2.2kb insert from pSGA197 was excised by an *Eco*RI-*Xba*I double digest and cloned into the *Eco*RI-*Xba*I site of pUC19 to produce pOP-1. Direction of transcription from the insert promoter is shown by the arrow within the pOP-1 figure. A restriction enzyme map of the pOP-1 insert was determined, and using this information a set of subclones, pOP-2 to pOP-5, was generated. Filled areas indicate pUC19 DNA, while open areas represent the mycobacterial DNA insert. Arrows above and below the pUC19 regions represent the universal and reverse sequencing primers respectively.

9.6.2 DNA sequencing of the pOP subclones.

9.6.2.1 Choice of sequencing method.

Initial attempts to sequence the pOP range of subclones used the Sequenase Version 2.0 kit supplied by United States Biochemical Corporation (Section 2.7.1). Plasmid DNA was denatured using heat and alkali to produce single stranded template that could be used in the Sequenase reactions. Reaction products were analysed by electrophoresis through gradient buffer polyacrylamide gels and subsequent autoradiography (Section 2.7.1). Sequence data produced by this method was generally very poor, with compressions, pile-ups and other artefacts evident throughout the autoradiograms. The poor quality of sequence obtained was attributed to a probable high percentage of guanine and cytosine nucleotides in the sequence. This would promote the formation of hairpin loops in the single stranded template molecule, causing the Sequenase enzyme to stall or non-specifically terminate during the strand synthesis reactions, and altering the mobility of the chain terminated products during electrophoresis through polyacrylamide gels. To inhibit the formation of hairpin loops during electrophoresis, dITP rather than dGTP was used in the extension reactions. However, this modification had little or no effect on the quality of sequence produced.

In an attempt to overcome these problems, sequencing reactions were performed using a thermostable DNA polymerase. The CircumVent Cycle Sequencing kit manufactured by New England Biolabs utilises a highly thermostable DNA polymerase, $Vent_R(exo^-)$, that allows sequencing reactions to be performed at 72°C rather than the 37°C needed for Sequenase based reactions. This higher reaction temperature should prevent hairpin loop formation, giving better sequence data from G+C rich templates. There is also no requirement to convert the plasmid to single stranded DNA as the reaction protocol includes a denaturation step of 98°C for 20 seconds, sufficient to denature the plasmid. Sequencing reactions using pOP-1 plasmid DNA and the universal forward primer were performed with the CircumVent Cycle Sequencing kit according to the manufacturer's instructions. Reaction products were analysed by gradient polyacrylamide gel electrophoresis followed by autoradiography (Section 2.7.1). Despite repeated attempts, the quality of sequence data obtained using this kit

was also poor, with common problems including extensive smearing and shadow bands in the gel lanes, and failure of one of the reactions (i.e. nothing visible in the lane of the autoradiogram).

The sequencing method that gave the most satisfactory results was "automated" sequencing using the Prism Ready Reaction DyeDeoxy Terminator cycle sequencing kit (Applied Biosystems) in conjunction with an Applied Biosystems Model 373A DNA sequencing system. This system relies on the incorporation of fluorescent dye-labelled dideoxynucleotides by the thermostable enzyme Ampli*Taq* DNA polymerase to terminate the extension of the primer. This sequencing strategy has two main advantages; (i) the use of dye-labelled ddNTPs reduces signals due to false stops since the false stops won't be labelled, (ii) reaction results are analysed by laser densitometry and computer analysis, a process that allows long reads of sequence to be obtained from a single run.

9.6.2.2 DNA sequencing strategy and sequence determination

The first step in the sequencing strategy was to determine the sequence of the pOP subclones using the universal forward and reverse primers. On the basis of the sequence obtained from these subclones, oligonucleotide primers (18-24 bases in length) were synthesised to complete the sequence. A list of the primers used is shown in Table 9.7. Figure 9.15 depicts the sequencing strategy for the pOP-1 subclone, with primer numbers corresponding to the primers in Table 9.7.

The primary goal of the sequencing of pOP-1 was to see if the cloned mycobacterial DNA insert had any homologies to sequences in the EMBL and GenBank nucleic acid databases, or the Swissprot protein database. For example, if any open reading frames (ORFs) present in the clone had homologies to genes that were known to be regulated by acid, this would be a good start to deducing the function of the mycobacterial locus. Also, the location of ORFs in the insert would give a good indication of where the acid-responsive promoter might be located. Therefore, it was felt that putting a lot of effort into determining the sequence of the insert on both strands was unnecessary at this stage. However, this is not to say that the sequence contains a lot

of ambiguities; each base was determined on the basis of at least duplicate reactions (with triplicate reactions performed in many cases).

The sequence of the insert in pOP-1 is presented in Figure 9.16. The figure was produced using the 'map' program from the Genetics Computer Group (GCG) software package (Genetics Computer Group, 1991). The sequence is shown 5' to 3' in the direction of transcription from the insert's putative acid-regulated promoter. The translation products (a, b and c) from the three forward reading frames are shown under the sequence, with amino acids represented by the single letter code. Positions where a base could not be determined with certainty are denoted "N" to show that the base could be A, T, C, or G. Codons containing an "N" base are translated to "?" by the program.

Table 9.7 Primers used in sequencing the pOP-1 DNA insert

Primer number	Sequence (5'-3')
1 (Universal)	CGCCAGGGTTTTCCCAGTCACGA
2 (Reverse)	AGCGGATAACAATTTCACACAGGA
3	AAACCGTTGTAGCCCAGCGCC
4	CGCGAACCGTTCGAACACGT
5	GCTGACGGTGTCGACGACCA
6	CAACTCGGTGGTCGACGACA
7	GCGGGAGGTCGCCGATCTCG
8	GAGACGCGCGTGCGCCAGTTC
9	GAGGATGACTTGGTGACACT
10	CTGTTCGGGTCGAGCGACCCG
11	GTGCCGTGCTGGCGCTCACACCCG

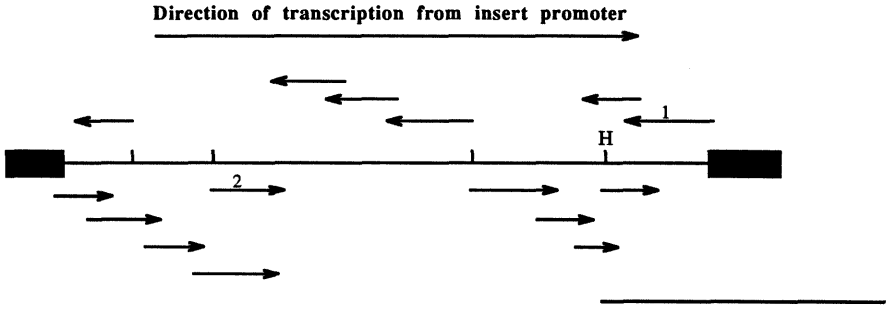
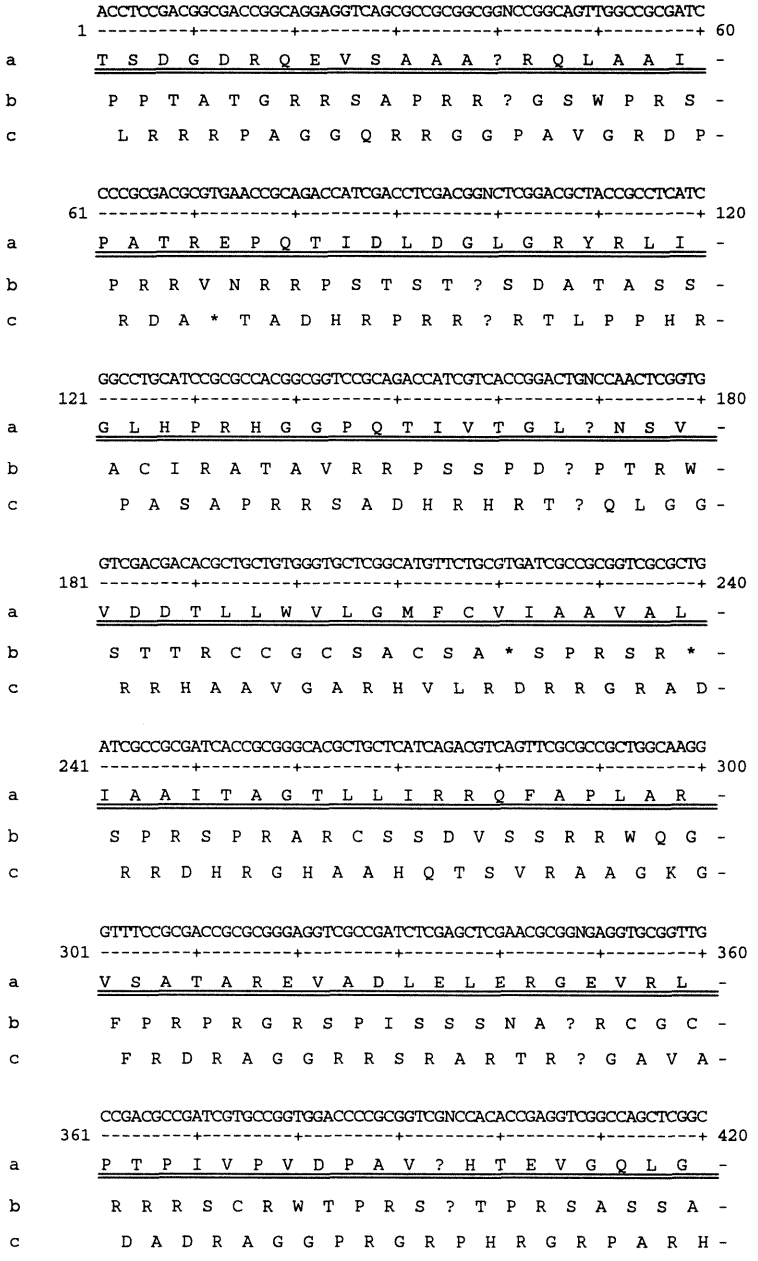
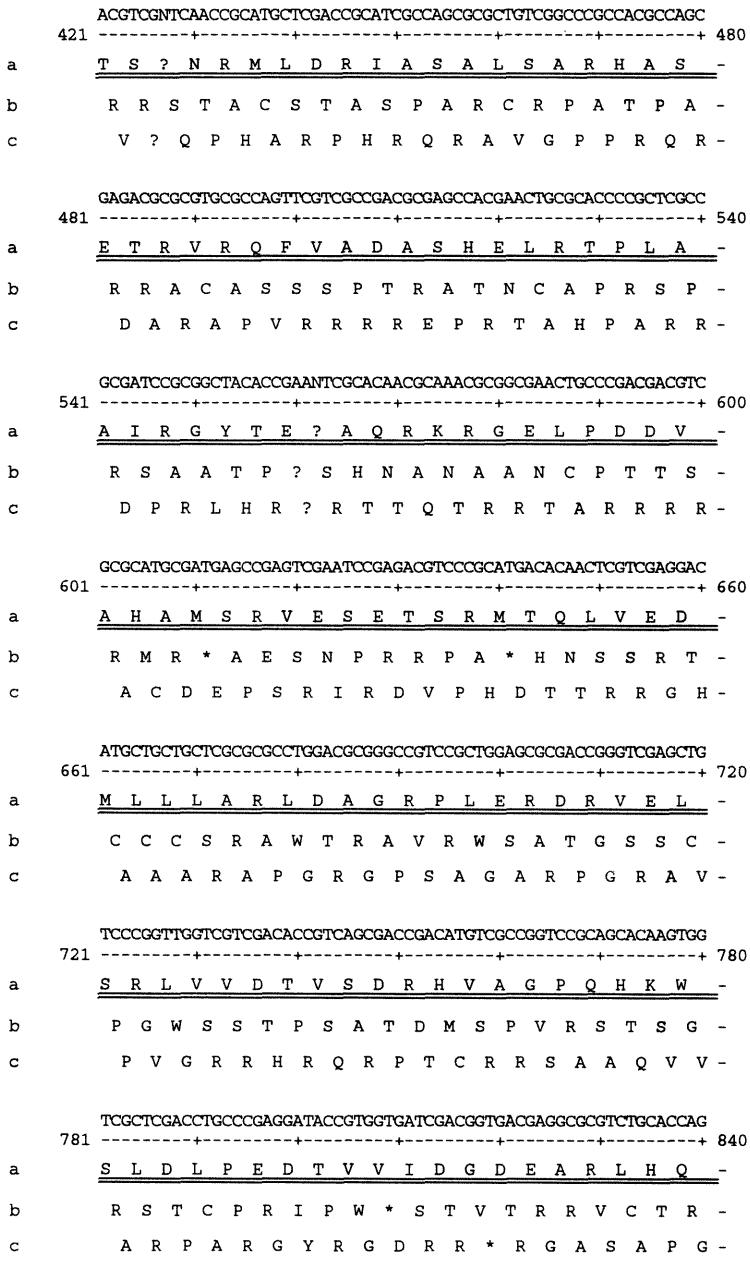
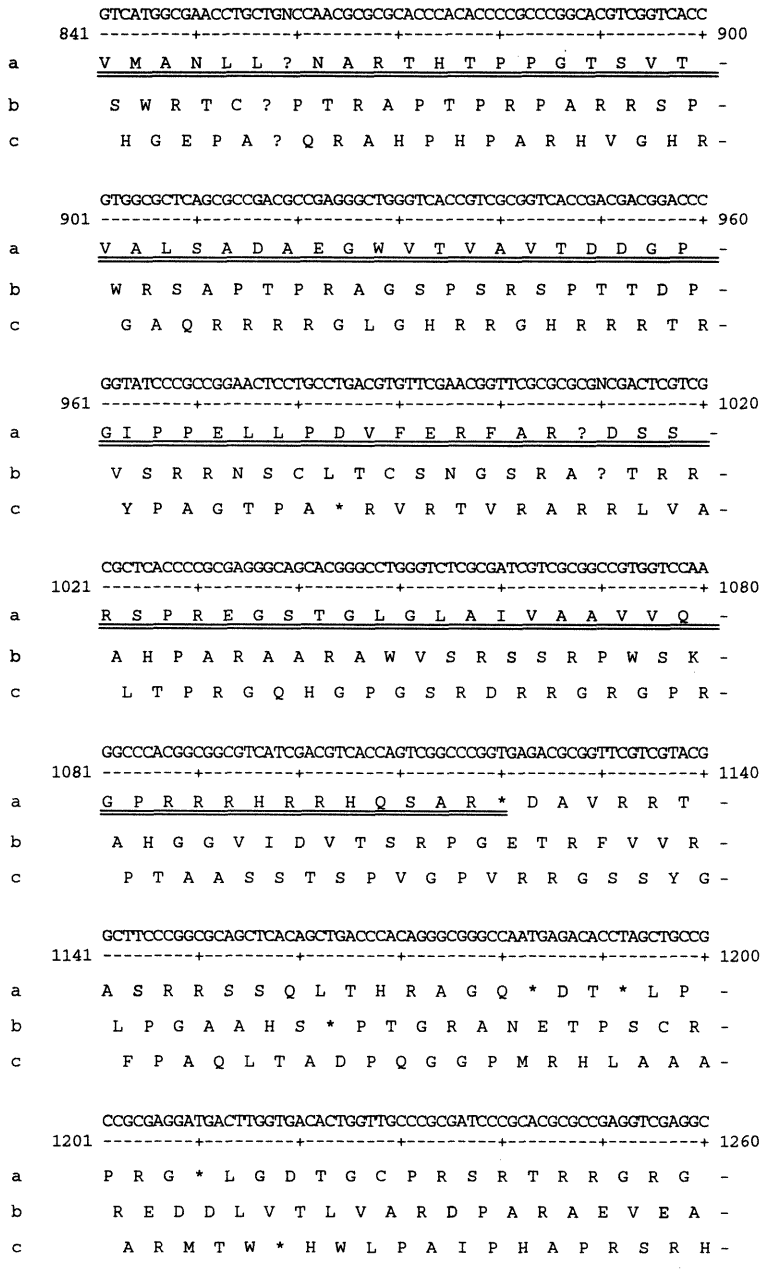


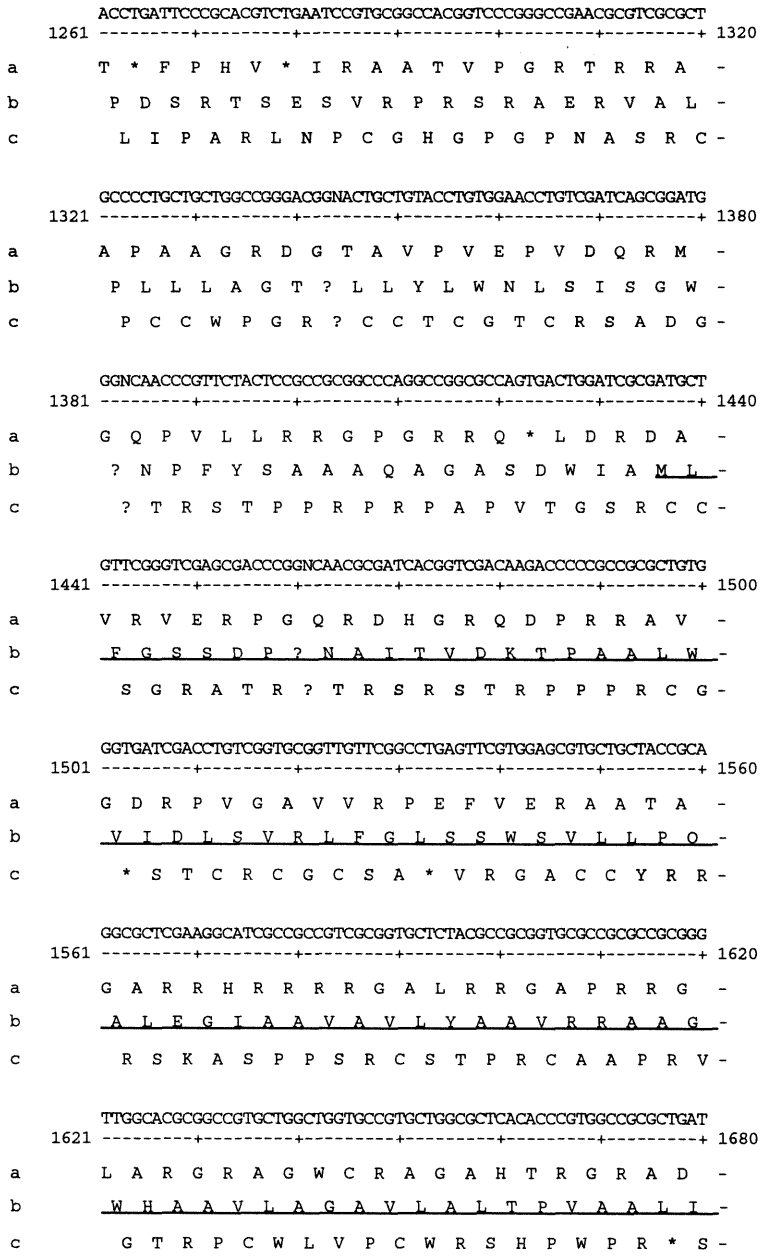
Figure 9.15 Sequencing strategy of pOP-1 to pOP-5

In the diagram, solid rectangles represent pUC19 DNA, while the pOP-1 insert is represented by a line. Arrows indicate both the length of fragment sequenced from the oligonucleotide primer and the direction of primer extension. Numbers over the arrows indicate the oligonucleotide primer used as detailed in Table 9.7. Sequence produced from the universal ("1") or reverse ("2") primers indicates positions where the sequence was obtained from the pOP subclones. Restriction sites are as follows: H, *Hind*III; Sm, *Sma*I; Sp, *Sph*I; Ss, *Sst*I.


ACCTCCGACGGCGACCGGCAGGAGGTCAGCGCCGCGGCGGNCCGGCAGTTGGCCGCGATC
1 ---------+---------+---------+---------+---------+---------+ 60
a T S D G D R Q E V S A A A ? R Q L A A I -
b P P T A T G R R S A P R R ? G S W P R S -
c L R R R P A G G Q R R G G P A V G R D P -
CCCGCGACGCGTGAACCGCAGACCATCGACCTCGACGGNCTCGGACGCTACCGCCTCATC
61 ---------+---------+---------+---------+---------+---------+ 120
a P A T R E P Q T I D L D G L G R Y R L I -
b P R R V N R R P S T S T ? S D A T A S S -
c R D A * T A D H R P R R ? R T L P P H R -
GGCCTGCATCCGCGCCACGGCGGTCCGCAGACCATCGTCACCGGACTGNCCAACTCGGTG
121 ---------+---------+---------+---------+---------+---------+ 180
a G L H P R H G G P Q T I V T G L ? N S V -
b A C I R A T A V R R P S S P D ? P T R W -
c P A S A P R R S A D H R H R T ? Q L G G -
GTCGACGACACGCTGCTGTGGGTGCTCGGCATGTTCTGCGTGATCGCCGCGGTCGCGCTG
181 ---------+---------+---------+---------+---------+---------+ 240
a V D D T L L W V L G M F C V I A A V A L -
b S T T R C C G C S A C S A * S P R S R * -
c R R H A A V G A R H V L R D R R G R A D -
ATCGCCGCGATCACCGCGGGCACGCTGCTCATCAGACGTCAGTTCGCGCCGCTGGCAAGG
241 ---------+---------+---------+---------+---------+---------+ 300
a I A A I T A G T L L I R R Q F A P L A R -
b S P R S P R A R C S S D V S S R R W Q G -
c R R D H R G H A A H Q T S V R A A G K G -
GTTTCCGCGACCGCGCGGGAGGTCGCCGATCTCGAGCTCGAACGCGGNGAGGTGCGGTTG
301 ---------+---------+---------+---------+---------+---------+ 360
a V S A T A R E V A D L E L E R G E V R L -
b F P R P R G R S P I S S S N A ? R C G C -
c F R D R A G G R R S R A R T R ? G A V A -
CCGACGCCGATCGTGCCGGTGGACCCCGCGGTCGNCCACACCGAGGTCGGCCAGCTCGGC
361 ---------+---------+---------+---------+---------+---------+ 420
a P T P I V P V D P A V ? H T E V G Q L G -
b R R R S C R W T P R S ? T P R S A S S A -
c D A D R A G G P R G R P H R G R P A R H -


ACGTCGNTCAACCGCATGCTCGACCGCATCGCCAGCGCGCTGTCGGCCCGCCACGCCAGC
421 480
a T S ? N R M L D R I A S A L S A R H A S -
b R R S T A C S T A S P A R C R P A T P A -
c V ? Q P H A R P H R Q R A V G P P R Q R -
GAGACGCGCGTGCGCCAGTTCGTCGCCGACGCGAGCCACGAACTGCGCACCCCGCTCGCC
481 540
a E T R V R Q F V A D A S H E L R T P L A -
b R R A C A S S S P T R A T N C A P R S P -
c D A R A P V R R R R E P R T A H P A R R -
GCGATCCGCGGCTACACCGAANTCGCACAACGCAAACGCGGCGAACTGCCCGACGACGTC
541 600
a A I R G Y T E ? A Q R K R G E L P D D V -
b R S A A T P ? S H N A N A A N C P T T S -
c D P R L H R ? R T T Q T R R T A R R R R -
GCGCATGCGATGAGCCGAGTCGAATCCGAGACGTCCCGCATGACACAACTCGTCGAGGAC
601 660
a A H A M S R V E S E T S R M T Q L V E D -
b R M R * A E S N P R R P A * H N S S R T -
c A C D E P S R I R D V P H D T T R R G H -
ATGCTGCTGCTCGCGCGCCTGGACGCGGGCCGTCCGCTGGAGCGCGACCGGGTCGAGCTG
661 720
a M L L L A R L D A G R P L E R D R V E L -
b C C C S R A W T R A V R W S A T G S S C -
c A A A R A P G R G P S A G A R P G R A V -
TCCCGGTTGGTCGTCGACACCGTCAGCGACCGACATGTCGCCGGTCCGCAGCACAAGTGG
721 780
a S R L V V D T V S D R H V A G P Q H K W -
b P G W S S T P S A T D M S P V R S T S G -
c P V G R R H R Q R P T C R R S A A Q V V -
TCGCTCGACCTGCCCGAGGATACCGTGGTGATCGACGGTGACGAGGCGCGTCTGCACCAG
781 840
a S L D L P E D T V V I D G D E A R L H Q -
b R S T C P R I P W * S T V T R R V C T R -
c A R P A R G Y R G D R R * R G A S A P G -


GTCATGGCGAACCTGCTGNCCAACGCGCGCACCCACACCCCGCCCGGCACGTCGGTCACC
841 900
a V M A N L L ? N A R T H T P P G T S V T -
b S W R T C ? P T R A P T P R P A R R S P -
c H G E P A ? Q R A H P H P A R H V G H R -
GTGGCGCTCAGCGCCGACGCCGAGGGCTGGGTCACCGTCGCGGTCACCGACGACGGACCC
901 960
a V A L S A D A E G W V T V A V T D D G P -
b W R S A P T P R A G S P S R S P T T D P -
c G A Q R R R R G L G H R R G H R R R T R -
GGTATCCCGCCGGAACTCCTGCCTGACGTGTTCGAACGGTTCGCGCGCGNCGACTCGTCG
961 1020
a G I P P E L L P D V F E R F A R ? D S S -
b V S R R N S C L T C S N G S R A ? T R R -
c Y P A G T P A * R V R T V R A R R L V A -
CGCTCACCCCGCGAGGGCAGCACGGGCCTGGGTCTCGCGATCGTCGCGGCCGTGGTCCAA
1021 1080
a R S P R E G S T G L G L A I V A A V V Q -
b A H P A R A A R A W V S R S S R P W S K -
c L T P R G Q H G P G S R D R R G R G P R -
GGCCCACGGCGGCGTCATCGACGTCACCAGTCGGCCCGGTGAGACGCGGTTCGTCGTACG
1081 1140
a G P R R R H R R H Q S A R * D A V R R T -
b A H G G V I D V T S R P G E T R F V V R -
c P T A A S S T S P V G P V R R G S S Y G -
GCTTCCCGGCGCAGCTCACAGCTGACCCACAGGGCGGGCCAATGAGACACCTAGCTGCCG
1141 1200
a A S R R S S Q L T H R A G Q * D T * L P -
b L P G A A H S * P T G R A N E T P S C R -
c F P A Q L T A D P Q G G P M R H L A A A -
CCGCGAGGATGACTTGGTGACACTGGTTGCCCGCGATCCCGCACGCGCCGAGGTCGAGGC
1201 1260
a P R G * L G D T G C P R S R T R R G R G -
b R E D D L V T L V A R D P A R A E V E A -
c A R M T W * H W L P A I P H A P R S R H -


ACCTGATTCCCGCACGTCTGAATCCGTGCGGCCACGGTCCCGGGCCGAACGCGTCGCGCT
1261 1320
a T * F P H V * I R A A T V P G R T R R A -
b P D S R T S E S V R P R S R A E R V A L -
c L I P A R L N P C G H G P G P N A S R C -
GCCCCTGCTGCTGGCCGGGACGGNACTGCTGTACCTGTGGAACCTGTCGATCAGCGGATG
1321 1380
a A P A A G R D G T A V P V E P V D Q R M -
b P L L L A G T ? L L Y L W N L S I S G W -
c P C C W P G R ? C C T C G T C R S A D G -
GGNCAACCCGTTCTACTCCGCCGCGGCCCAGGCCGGCGCCAGTGACTGGATCGCGATGCT
1381 1440
a G Q P V L L R R G P G R R Q * L D R D A -
b ? N P F Y S A A A Q A G A S D W I A M L -
c ? T R S T P P R P R P A P V T G S R C C -
GTTCGGGTCGAGCGACCCGGNCAACGCGATCACGGTCGACAAGACCCCCGCCGCGCTGTG
1441 1500
a V R V E R P G Q R D H G R Q D P R R A V -
b F G S S D P ? N A I T V D K T P A A L W -
c S G R A T R ? T R S R S T R P P P R C G -
GGTGATCGACCTGTCGGTGCGGTTGTTCGGCCTGAGTTCGTGGAGCGTGCTGCTACCGCA
1501 1560
a G D R P V G A V V R P E F V E R A A T A -
b V I D L S V R L F G L S S W S V L L P Q -
c * S T C R C G C S A * V R G A C C Y R R -
GGCGCTCGAAGGCATCGCCGCCGTCGCGGTGCTCTACGCCGCGGTGCGCCGCGCCGCGGG
1561 1620
a G A R R H R R R R G A L R R G A P R R G -
b A L E G I A A V A V L Y A A V R R A A G -
c R S K A S P P S R C S T P R C A A P R V -
TTGGCACGCGGCCGTGCTGGCTGGTGCCGTGCTGGCGCTCACACCCGTGGCCGCGCTGAT
1621 1680
a L A R G R A G W C R A G A H T R G R A D -
b W H A A V L A G A V L A L T P V A A L I -
c G T R P C W L V P C W R S H P W P R * S -

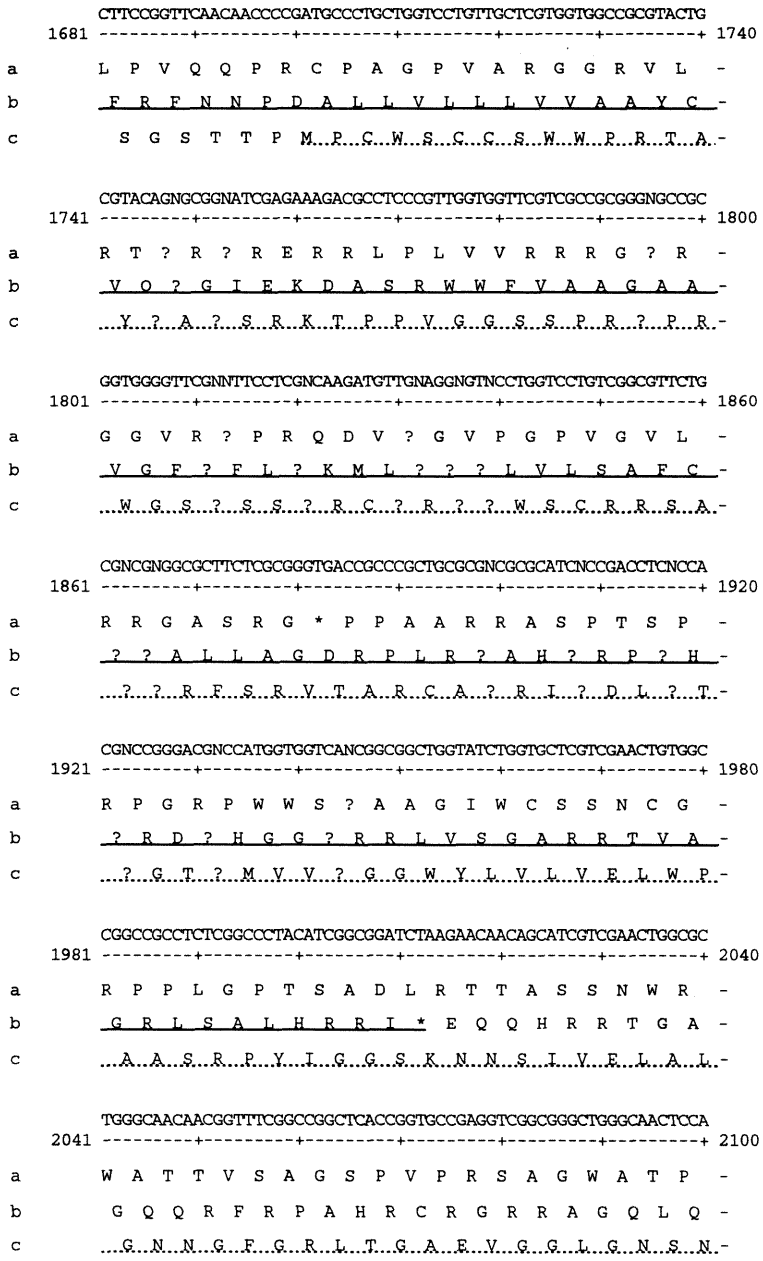

CTTCCGGTTCAACAACCCCGATGCCCTGCTGGTCCTGTTGCTCGTGGTGGCCGCGTACTG
1681 1740
a L P V Q Q P R C P A G P V A R G G R V L -
b F R F N N P D A L L V L L L V V A A Y C -
c S G S T T P M P C W S C C S W W P R T A -
CGTACAGNGCGGNATCGAGAAAGACGCCTCCCGTTGGTGGTTCGTCGCCGCGGGNGCCGC
1741 1800
a R T ? R ? R E R R L P L V V R R R G ? R -
b V Q ? G I E K D A S R W W F V A A G A A -
c Y ? A ? S R K T P P V G G S S P R ? P R -
GGTGGGGTTCGNNTTCCTCGNCAAGATGTTGNAGGNGTNCCTGGTCCTGTCGGCGTTCTG
1801 1860
a G G V R ? P R Q D V ? G V P G P V G V L -
b V G F ? F L ? K M L ? ? ? L V L S A F C -
c W G S ? S S ? R C ? R ? ? W S C R R S A -
CGNCGNGGCGCTTCTCGCGGGTGACCGCCCGCTGCGCGNCGCGCATCNCCGACCTCNCCA
1861 1920
a R R G A S R G * P P A A R R A S P T S P -
b ? ? A L L A G D R P L R ? A H ? R P ? H -
c ? ? R F S R V T A R C A ? R I ? D L ? T -
CGNCCGGGACGNCCATGGTGGTCANCGGCGGCTGGTATCTGGTGCTCGTCGAACTGTGGC
1921 1980
a R P G R P W W S ? A A G I W C S S N C G -
b ? R D ? H G G ? R R L V S G A R R T V A -
c ? G T ? M V V ? G G W Y L V L V E L W P -
CGGCCGCCTCTCGGCCCTACATCGGCGGATCTAAGAACAACAGCATCGTCGAACTGGCGC
1981 2040
a R P P L G P T S A D L R T T A S S N W R -
b G R L S A L H R R I * E Q Q H R R T G A -
c A A S R P Y I G G S K N N S I V E L A L -
TGGGCAACAACGGTTTCGGCCGGCTCACCGGTGCCGAGGTCGGCGGGCTGGGCAACTCCA
2041 2100
a W A T T V S A G S P V P R S A G W A T P -
b G Q Q R F R P A H R C R G R R A G Q L Q -
c G N N G F G R L T G A E V G G L G N S N -

Figure 9.16 Sequence of the pOP-1 insert

Nucleic acid sequence of the pOP-1 mycobacterial DNA insert is shown. The sequence is shown 5' to 3' in the direction of transcription from the insert's acid-regulated promoter. The translation products (a, b and c) from the three forward reading frames are shown under the sequence, with amino acids represented by the single letter code. Refer to Figure 9.17 for open reading frames (ORF). ORF1 is shown with a double underline; ORF5 is shown with a single underline; ORF6 is shown with each character underlined; ORF9 is shown with a dotted underline. Positions where a base could not be called with certainty were denoted "N" to show that the base could be A, T, C, or G. In the translated products, codons containing a "N" base are translated to "?". "*" denotes a translational stop.

9.6.3 Sequence analysis of the pOP-1 insert.

DNA sequence analysis was performed using programs from the University of Wisconsin's Genetics Computer Group (GCG) software package (Genetics Computer Group, 1991). As a starting point, the sequence was searched for the presence of open reading frames (ORFs) using the 'frames' program. The result of this search is presented in Figure 9.17. The three forward reading frames (5' to 3', in the direction of transcription from the acid-regulated promoter) are denoted a, b, and c, with possible ORFs labelled 1 to 9. The amino acid translation products from these ORFs are shown in Figure 9.16 (cross-referencing the labels a, b and c).

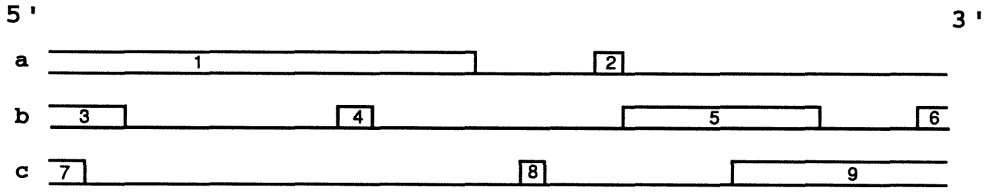


Figure 9.17 ORFs in the pOP-1 sequence

The sequence from pOP-1 was searched for ORFs using the 'frames' program of the GCG software package. The three forward reading frames, in the direction of transcription from the acid-regulated promoter, are shown (a, b and c), with the ORFs labelled 1-9.

The next step was to see whether any of the products of these ORFs had homology with any amino acid sequences in the databases. To accomplish this, the entire sequence was translated using 'pepdata', a program that translates nucleic acid sequences in all six reading frames. The resulting translation product was then searched against the Swissprot protein database for homologous sequences using

the 'FASTA' program (word size = 2). The top ten scores from this search are shown in Table 9.8 (two scores have been left out for the purposes of clarity; a "hypothetical protein" from *B. subtilus* (P37483) showed 34% identity over 194 amino acids but this was with a section of the insert that was not associated with any ORF; a mouse proline-rich protein showed 19.2% identity over 182 amino acids to a section translated in the opposite direction to that of the acid-regulated promoter).

Table 9.8 Results of the FASTA search of the Swissprot database

ORF	Description	Acc.*	% Identity
1	Sensor protein PhoR, *B. subtilus*	P23545	33.8% in 204 aa
1	Hypothetical 66.8kD protein, *B. subtilus*	P35164	32.8% in 201 aa
1	Sensor protein CreC, *E. coli*	P08401	27.2% in 327 aa
1	Sensor protein PhoR, *E. coli*	P08400	30.6% in 248 aa
1	Sensor protein KdpD, *E. coli*	P21865	24.7% in 243 aa
1	Sensor protein Afsq2, *S. coelicolor*	Q04943	30.4% in 336 aa
1	Sensor protein CpxA, *E. coli*	P08336	28.2% in 213 aa
1	Sensor kinase SarS, *Synechococcus* 7942	Q06904	25.7% in 249 aa
1	Sensor protein PmrB, *S. typhimurium*	P36557	27.2% in 283 aa
1	Nitrogen regulator NtrB, *V. alginolyticus*	P19906	22.4% in 277 aa

* Acc.= Accession number for the Swissprot database

The ORF1 translation product was found to have significant homology to histidine-sensor-kinase proteins. These proteins are integral constituents of the so-called two component regulator systems (Gross *et al.*, 1989). Taking the *B. subtilus* PhoR protein as an example, positions 172-175aa in the ORF1 protein showed identity to amino acids 359-362 (..SHEL..) in the PhoR protein (Figure 9.18). The histidine residue contained in this sequence is conserved in all histidine-sensor-kinases, and is the site of phosphorylation (Roberts *et al.*, 1994). Other motifs that are conserved in histidine-sensor kinases are also found in the ORF1 product, such as GLGLA (amino acids 349-353) and Asn-288 (N) (Roberts *et al.*, 1994; Bayles, 1994).

Of the other ORFs, particular attention was given to the analysis of ORF5, ORF6 and ORF9 as they could be transcribed from a promoter located downstream from ORF1 (i.e., the putative acid-regulated promoter). ORFs 2, 3, 4, 7 and 8 were considered too small to warrant an extensive examination. Homology searches of the translation products from ORFs 5, 6 and 9 were performed against the Swissprot protein database in an effort to reveal homologies that may have been "hidden" when the whole insert was used. However, none of these ORFs showed any significant homologies with other proteins in the Swissprot database.

A search at the nucleic acid level was performed using the whole insert sequence to see if the DNA databases contained any homologous sequences. Searches based on DNA sequence are generally less sensitive than protein-based searches, so care was taken in the conclusions drawn from the search results. The most interesting result from the search of the EMBL and GenBank bacterial databases was the homology shown between the pOP-1 insert and a section of the *M. leprae* cosmid B2168 (EMBL accession U00018). This cosmid contains approximately 43 kb of *M. leprae* DNA and was constructed as part of the *M. leprae* genome sequencing project (Eiglmeier *et al.*, 1993). The region of the cosmid that showed homology (52.2% over 851bp) to the DNA sequence of pOP-1 was a region that codes for a putative *M. leprae* PhoR protein. This was the same region of DNA that coded for ORF1. At the amino acid level, the ORF1 protein product had 25.7% identity with this putative *M. leprae* PhoR protein (Figure 9.19). These results suggest that ORF1 codes for part of an *M. smegmatis* kinase protein.

(a) *B. subtilus* PhoR sensor-kinase

```
            140       150       160       170       180       190
ORF1    VXHTEVGQLGTSXNRMLDRIASALSARHASETRVRQFVADASHELRTPLAAIRGYTEXAQ
                                          | : ::|||::||||:||:::|:|:||:
PhoR    EVDGVPIMGPDDEWKGIVLVFHDMTETKKLEQMRKDFVANVSHELKTPITSIKGFTETLL
        320       330       340       350       360       370

            200       210       220       230       240
ORF1    RKRGELPDDVAHAMSRVESETSRMTQLVEDMLLLARLDAGR-PLERDRVELSRLV--VDT
          |  ::::: :| : :|::|:  ||:|:| |::::: : :|: :  | ::::  ::|
PhoR    DGAMEDKEALSEFLSIILKESERLQSLVQDLLDLSKIEQQNFTLSIETFEPAKMLGEIET
        380       390       400       410       420       430

        250       260       270       280       290       300
ORF1    VSDRHVAGPQHKWSLDLPEDTVVIDGDEARLHQVMANLLXNARTHTPPGTSVTVALSADA
        : ::::::   :  |::|:|:  ::||  ||:||: ||::|| |:|| |:||::::::  :
PhoR    LLKHKADEKGISLHLNVPKDPQYVSGDPYRLKQVFLNLVNNALTYTPEGGSVAINVKPRE
        440       450       460       470       480       490

        310       320       330       340       350       360
ORF1    EGWVTVAVTDDGPGIPPELLPDVFERFARXDSSRSPREGSTGLGLAIVAAVVQGPRRRHR
        :: : ::|:|:| ||: | :| :|||| |:|::||:::|:|||||||  :::::
PhoR    KD-IQIEVADSGIGIQKEEIPRIFERFYRVDKDRSRNSGGTGLGLAIVKHLIEAHEGKID
         500       510       520       530       540       550
```

(b) *B. subtilus* hypothetical 66.8kD protein

```
          140       150       160       170       180       190
ORF1    VGQLGTSXNRMLDRIASALSARHASETRVRQFVADASHELRTPLAAIRGYTEXAQRKRGE
                                        :|:|::|||||||:: ::|||:|:  : ::
66.8kD SPLYAESHVRGAVAVLRDMTEERRLDKLREDFIANVSHELRTPISMLQGYSEAIVDDIAS
         340       350       360       370       380       390

          200       210       220       230       240       250
ORF1    LPDDVAHAMSRVESETSRMTQLVEDMLLLARLDAGRP-LERDRVELSRLVVDTVSDRHVA
          :|  :    : :|: ||:||:|:|:| |||:::|:: |: :::::: :: :::  |: :
66.8kD SEEDRKEIAQIIYDESLRMGRLVNDLLDLARMESGHTGLHYEKINVNEFLEKII--RKFS
         400       410       420       430       440       450

            260       270           280       290       300       310
ORF1    GPQHKWSLDLPEDTVVIDG----DEARLHQVMANLLXNARTHTPPGTSVTVALSADAEGW
        | ::: :::| :|: ::::    ||:::||::||::||  ||::|:||:::: :  :|
66.8kD GVAKEKNIALDHDISLTEEEFMFDEDKMEQVFTNLIDNALRHTSAGGSVSISVHSVKDG-
           460       470       480       490       500       510

              320       330       340       350       360       370
ORF1    VTVAVTDDGPGIPPELLPDVFERFARXDSSRSPREGSTGLGLAIVAAVVQGPRRRHRRHQ
        ::::::|:|:||| | || :|||| ::|::|::  ::|||||||| ::|:::
66.8kD LKIDIKDSGSGIPEEDLPFIFERFYKADKARTRGRAGTGLGLAIVKNIVEAHNGSITVHS
            520       530       540       550       560       570
```

Figure 9.18 Homology of ORF1 to sensor-kinase proteins

Two high homologies from the FASTA Swissprot database search are shown. Both proteins are *B. subtilus* sensor-kinase proteins and share homology with the translation product from ORF1 at critical residues (shown in bold type).

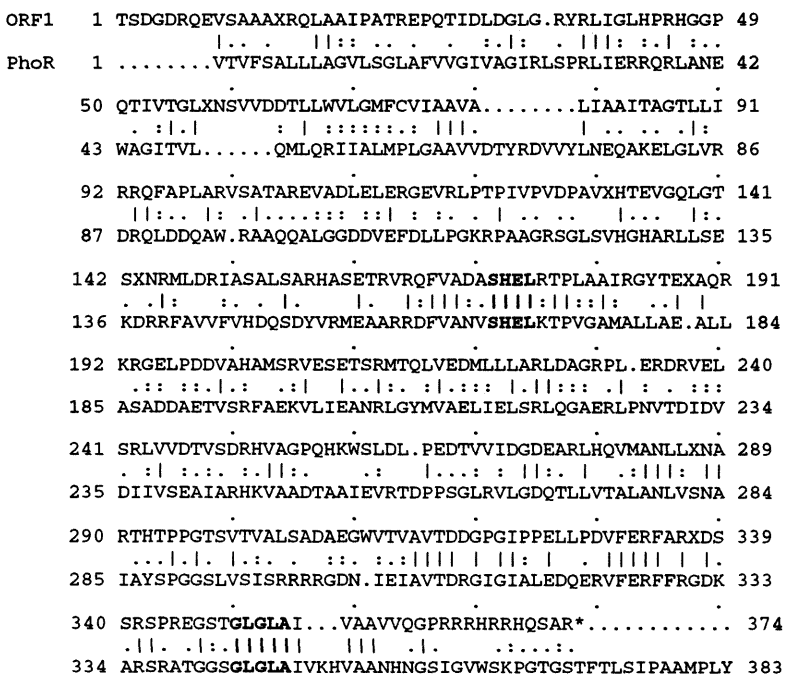


Figure 9.19 Homology of ORF1 protein to the *M. leprae* PhoR protein. The putative *M. leprae* PhoR protein (from cosmid B2168; EMBL accession no. U00018) was aligned against the ORF1 protein using the GCG program 'gap' (settings were gap weight 3.0, length weight 0.1). The proteins showed 46.2% similarity and 25.7% identity. Regions showing homology with two component regulators are shown in bold.

Chapter 10.

Discussion

Adaptation to changes in the external environment is central to bacterial survival. Systems that sense extracellular conditions and react to them must therefore exist in the bacterial cell. These survival systems have been characterised in many bacteria for a range of stress conditions. This project set out to investigate the previously uncharacterised response to acid stress in mycobacteria.

The answers to three questions were sought:
(i) can mycobacteria adapt to external acid? If so, then
(ii) does adaptation require the synthesis of new proteins?
(iii) could we identify any genes involved in adaptation?
The approaches used to investigate these questions and the conclusions reached shall be addressed in turn.

10.1 Evidence for an adaptive tolerance response.

The extensive work of Foster and colleagues on the acid tolerance response (ATR) of *S. typhimurium* provides us with a model for bacterial adaptation to acidic conditions (Foster, 1993; Lee *et al.*, 1994). In this model, the central function of adaptation is to allow the synthesis of a set of acid shock proteins (ASPs) that enhance survival at normally lethal acid conditions. In the log-phase ATR model, these ASPs can be synthesised in two ways: (i) bacteria that are exposed to an adaptive pH of 5.8 (pre-acid shock) after growth at pH 7.6 induce a pH homeostasis system that permits ASP synthesis at pH_o values that normally block protein synthesis (i.e. pH 3.3); (ii) bacteria grown at pH 7.5 and then shifted to pH 4.3 (post-acid shock) induce ASPs that allow survival at pH 3.3 (Foster, 1993).

As a first step in describing the response to acid stress in mycobacteria, the presence of a log-phase ATR in *M. smegmatis* mc^2155 was investigated. Specifically, the work of Foster and Hall (1990) in describing the pre-acid shock ATR of *S. typhimurium* was repeated using *M. smegmatis*. To begin with, values for the adaptive and lethal pHs had to be determined. Cultures of *M. smegmatis* were grown at pH

7.6, then the culture pH dropped to values from pH 6.5 to pH 2.0 using either HCl or H_3PO_4. The adaptive pH was defined as the lowest pH at which *M. smegmatis* still replicated (after Foster and Hall, 1990), and this was found to be pH 5.0. The lethal pH was defined as the pH that decreased cell viability over the four hour assay period, with pH 3.0 chosen when HCl was used to acidify the medium, or pH 3.5 when H_3PO_4 was used.

Mycobacterium smegmatis grown at pH 7.6, then shifted to the adaptive pH of 5.0, showed a greater percentage viability at the lethal pH than unadapted cells. This adaptive response was independent of the acid used to acidify the medium and required *de novo* protein synthesis. One hour at the adaptive pH was sufficient to produce a protective response. The magnitude of protection was generally a 2-3 fold increase in percentage viability after 4 hours exposure at the lethal pH, but this varied between experiments. Compared to studies with other organisms, this level of protection is low. For example, in *S. typhimurium*, adapted cells showed 2-3 log greater percentage survival after 90 minutes at the lethal pH (pH 3.3) than unadapted cells (Foster and Hall, 1990); in *A. hydrophila* adaptation produced a >5 log higher percentage survival at pH 3.5 after 90 minutes (Karem *et al.*, 1994); in *S. mutans*, adaptation produced around a 2 log increase in percentage viability after 30 minutes at the lethal pH (pH 2.5) (Belli and Marquis, 1991).

Mycobacterium smegmatis mc^2155 therefore shows a relatively low level of induced protection against acid stress. Perhaps this is to be expected, considering that *M. smegmatis* is reported to be one of the most acid-resistant mycobacteria. In their study of the growth of mycobacteria in relation to the pH of the medium, Portaels and Pattyn (1982) showed that *M. smegmatis* was capable of growth over one of the widest pH ranges of any mycobacterium, with optimum growth between pH 5.0 to 7.4 and partial growth at pH 4.6 (the lowest pH tested). This is compared to *M. tuberculosis* and *M. bovis* that showed optimum growth between pH 5.8 and 6.5, and were incapable of growth below pH 5.4. Chapman and Bernard (1962) had seen a similar picture, with *M. smegmatis* showing growth between pH 3.5 and 9.5, while *M. tuberculosis* grew between pH 5.0 and 8.4. These studies were performed with no obvious adaptive step; inocula consisted of organisms (grown between pH 6.6 and 7.0) suspended in saline, with the inoculum

added to liquid (Chapman and Bernard, 1962) or solid (Portaels and Pattyn, 1982) medium at a range of pH. Hence, after inoculation, survival at the new pH would depend on the constitutive systems that had existed at the original growth pH (i.e. pH 6.6-7.0).

Indeed, work done in this thesis agrees with the view that *M. smegmatis* is a relatively acid-resistant mycobacterium. The $\log_{10}$ CFU/ml of *M. smegmatis* cultures only decreased from 5.99 to 5.79 over a 4 hour period after shifting the culture pH from 7.6 to 3.0 (using HCl to acidify the medium); this decrease was from 6.11 to 4.24 when H_3PO_4 was used (Table 9.1).Thus, it appears that *M. smegmatis* can tolerate a 4 log increase in acidity without being rapidly killed, suggesting that the organism possesses efficient constitutive pH_i homeostasis systems.

It could be argued therefore that *M. smegmatis* is not a good organism to investigate mycobacterial adaptive acid tolerance since it possesses powerful constitutive pH_i homeostasis mechanisms that diminish the requirement for high levels of inducible acid-protection. This argument was appreciated at the start of the project, but equally one could argue the opposite case; i.e. since *M. smegmatis* shows a high degree of acid-resistance, this suggests that acid-resistance is important to the biology of this organism and hence it may also possess an inducible acid-resistance system that augments the constitutive systems further. There were also a number of advantages in using *M. smegmatis* over other mycobacteria as the model organism for this work. For example, it is a nonpathogen, so special precautions in its handling (such as level 3 containment, necessary for *M. tuberculosis*) are avoided; it is a rapidly growing mycobacterium (doubling time of 1-2 hours under optimum conditions), allowing a faster "turn-around" time for experiments; a number of genetic systems have been developed for it, making genetic analysis much easier than with other mycobacteria. Also worth noting is the strain of *M. smegmatis* used in this study; *M. smegmatis* mc^2155 is a mutant strain with an uncharacterised mutation that confers a high-transformation phenotype (Snapper *et al.*, 1990). This feature allows mycobacterial acid tolerance to be investigated at both the physiological and genetic level in the same organism. Therefore, *M. smegmatis* was chosen as the model mycobacterium for this study; it would serve to investigate the principal of adaptive acid tolerance in this genus.

Returning to the reasons for the relatively low level of induced protection seen, perhaps the adaptive pH used was not sufficient to induce a complete adaptive response. The adaptive pH was defined as the lowest pH at which the bacteria still replicated, a definition that was used by Foster and Hall (1990) to choose the adaptive pH for *S. typhimurium*. In *S. typhimurium*, this pH triggers the expression of an inducible pH homeostasis system that allows expression of protective ASPs at the lethal pH. However, this may not be the situation in *M. smegmatis*, where ASPs may need to be synthesised before exposure to the lethal pH (akin to the *S. typhimurium* post-acid shock ATR; Foster, 1993).

Indeed, how the *M. smegmatis* acid tolerance response described here fits in with Foster's *S. typhimurium* ATR model is not clear. For example, does the *M. smegmatis* ATR present pre-acid shock and post-acid shock phases? Or does it more closely resemble the habituation response in *E. coli*, where preshock and postshock phases are not seen (Rowbury and Goodson, 1993). On the basis of the results obtained, it is impossible to answer such questions. The hypothesis we set out to test (i.e. does exposure to sub-lethal acid conditions protect against lethal acid conditions) was based on Foster and Hall's original description of the *S. typhimurium* ATR (1990; 1991), a model that did not contain the more complex elements of the *S. typhimurium* ATR such as pre- and post-acid shock (Foster, 1993). To reach a better understanding of mycobacterial adaptation to acid conditions the later experiments of Foster and colleagues (Foster, 1993; Lee *et al.*, 1994) need to be repeated with *M. smegmatis* mc^2155. Specifically, we need to see whether protein synthesis is required for survival at the lethal pH after adaptation at pH 5.0. This could be investigated by adding chloramphenicol to adapted and unadapted cultures at pH 3.0, then measuring viability over time. If adapted and unadapted bacteria died at a similar rate, the adaptive response we see with *M. smegmatis* at pH 5.0 is a preshock event; i.e. there is a requirement for protective ASPs synthesis at the lethal pH. Hence, this would suggest that *M. smegmatis* possess an inducible pH homeostasis system that is switched "on" at pH 5.0 after growth at pH 7.6. On the other hand, survival of adapted bacteria at the lethal pH in the presence of chloramphenicol would suggest that pH 5 induces a postshock ATR; i.e. protective ASPs are synthesised at the adaptive pH.

Linked to this, a survey of different pH_o values and their ability to elicit protection at a lethal pH should be performed. This would take the form of a "checker board" experiment, where *M. smegmatis* would be grown at pH 7.6, then exposed to a range of "adaptive" pHs (e.g. from pH 7.0 to 2.0) for one generation before exposure to a range of "lethal" pHs (e.g. from pH 5.0 to pH 1.0). Using this strategy, combined with the use of chloramphenicol at the lethal pH, we may be able to identify pH values capable of eliciting preshock or postshock ATRs. We could then see whether consecutive exposure to both the preshock and postshock pHs increased the magnitude of protection against the lethal pH, as is the case in *S. typhimurium* (Foster, 1993). These experiments will help to produce a much more refined model for the acid tolerance response in *M. smegmatis*. These experiments are now being undertaken.

10.2 Protein synthesis at the adaptive pH

During the Ph.D. viva examination it was pointed out that critical controls were not performed during the course of the radiolabelling experiments. Discussion of the results obtained is therefore unsound and has been deleted.

10.3 Genetic response to acid stress.

Adaptation appeared to require *de novo* protein synthesis at the adaptive pH, suggesting that induction of genes coding for new "protective" proteins was central to the adaptive response. Two questions arose from this observation: can we identify any genes induced at the adaptive pH, and what role do they play in the response to acid stress? To answer these questions a promoter probe library was constructed and screened for promoters that were induced in response to acid stress.

Promoter probe libraries are constructed in vectors containing a reporter gene, a gene that codes for a product that is easy to assay. This permits the activity of a promoter to be measured as a function of the reporter's expression. Promoters can therefore be identified by screening the library for clones with increased expression of the reporter. A number of reporter genes have been described, including β-galactosidase (*lacZ*), chloramphenicol acetyltransferase (CAT) and bacterial luciferase (Brosius and Lupski, 1987; Stewart and

Williams, 1992). The advantages of luciferase over other reporters include the ease of light detection, the nondestructive nature of the assay, and the possibility of real time measurement of promoter activity (Hill *et al.*, 1993; Stewart and Williams, 1992). These advantages, particularly in relation to the ease in rapidly screening library clones for increased light production in response to acid stress, led to luciferase being chosen as the reporter system for this project.

Bacterial luciferase had not previously been used as a reporter gene in mycobacteria. Andrew and Roberts (1993) had shown that *V. harveyi* luciferase could be expressed and produce a bioluminescent phenotype in *M. smegmatis*. However, a number of control experiments had to be performed to determine the efficacy of luciferase as a reporter of mycobacterial promoter activity. Specifically, light production from luciferase should show a direct relationship to promoter activity in the mycobacterial host. To investigate the suitability of the luciferase reporter system, the acetamidase gene promoter from *M. smegmatis* was cloned upstream of the luciferase genes in the mycobacterial promoter probe shuttle vector pSG10. Using CAT as the reporter, Parish (1993) had shown that the acetamidase gene was induced approximately 10-15 fold when acetamide was used as the sole carbon source. Repeating these experiments with luciferase as the reporter showed that the acetamidase gene was induced greater than 1000-fold in response to acetamide. These results demonstrated the sensitivity of luciferase relative to CAT as a reporter of mycobacterial promoter activity. Also, luciferase assays are much easier to perform than CAT-based assays (T. Parish, pers. comm.), adding to the attraction of luciferase.

The promoter probe library was screened for clones that showed an increase in light production at pH 5 relative to pH 7.6. Of the 451 clones analysed to date, one clone termed pSGA197 has been identified that shows a reproducible 15-20 fold increase in light production at pH 5.0. The magnitude of induction was related to the pH shift; i.e., a decrease in pH form 7.6 to 4.0 caused a greater increase in light production than a shift from pH 7.6 to 6.0 (Section 9.5.1.3). Light production was not increased in response to heat shock or oxidative stress, indicating that the promoter is not a general stress responsive element (Section 9.5.1.2).

Recent reports suggest that bacterial luciferase can alter expression from promoters that are sensitive to oxidative stress and local DNA topology. Therefore, can we be sure that induction from pSGA197 is a real response to acid and not just an artefact? The fact that the level of light induction was dependent on the decrease in pH_o seems to discount this possibility (Section 9.5.1.3). Further, pSGA197 showed acid-induced light production when transformed into *E. coli*, demonstrating that induction was not host dependent and a real event.

Sequence analysis of the pSGA197 DNA insert revealed a number of opening reading frames (ORFs). One of these, denoted ORF1, produced a translation product that showed significant levels of homology with histidine-sensor-kinase proteins. These proteins are constituents of "two component regulators", a family of proteins involved in the response to environmental stimuli (Gros *et al.*, 1989; Parkinson, 1993). Specifically, the sensor-kinase transmits information (via phosphorylation) to a receiver protein that is capable of regulating gene expression. The organisation of the genes for the two components is interesting, in that the regulator and sensor are often transcribed from the same promoter. Do we find the same organisation with ORF1? For example, perhaps ORF1 is part of an operon where the regulator gene is transcribed first, followed by the kinase gene. This operon structure is found in a number of cases, such as *phoP* and *phoR* in *B. subtilus* (Sorokin *et al.*, 1993), and *bvgA* and *bvgS* in *Bordetella pertussis* (Melton and Weiss, 1993). Alternatively, the sensor (ORF1) may be transcribed first followed by the regulator, an organisation that is found in *E. coli* with *baeS* and *baeR* (Nagasawa *et al.*, 1993), and *kdpD* and *kdpE* (Walderhaug *et al.*, 1992).

The starting point of ORF1 appeared to be beyond the 5' end of the pOP-1 insert sequence. Hence, the promoter for this ORF is not the acid-regulated promoter that was driving luciferase expression in pSGA197. Since ORF1 showed a significant level of homology to *M. leprae phoR*, *B. subtilus phoR* and *E. coli phoR* (see Table 9.8, Figure 9.19), is ORF1 part of the *M. smegmatis phoR* gene? In addition, does this give us any clues to the identity of the gene controlled by the insert's acid-regulated promoter?

In the *M. leprae* chromosome, *phoR* is preceded by a *phoP* homologue, reflecting the gene organisation seen in *B. subtilus* (Sorokin *et al.*, 1993). This also appears to be the organisation seen with ORF1,

part of a possible *phoR* gene. Therefore, perhaps the next gene downstream from the *M. leprae phoR* is also downstream from ORF1, and hence is the gene controlled by the acid-regulated promoter. Examining the open reading frames downstream from the *M. leprae phoR* gene in cosmid B2168, the gene immediately downstream codes for a putative phosphoglycerate mutase, followed by a number of uncharacterised ORFs. Is a phosphoglycerate mutase gene downstream from ORF1? No significant homologies to any phosphoglycerate mutase genes with any region downstream from ORF1 have been found. However, the genome organisation seen in *M. leprae* may not reflect that seen in *M. smegmatis*. Alternatively, perhaps ORF1 codes for a sensor-kinase that has yet to be identified in the mycobacteria.

Two ORFs (ORF6 and ORF9) continue beyond the 3' end of the insert. These ORFs are the most likely candidates to be controlled by the insert's acid-regulated promoter. However, these ORFs show no significant homologies at either the protein or nucleic acid level with other sequences in the database. Therefore, to identify the gene controlled by the acid-regulated promoter, the section of DNA immediately downstream from the pOP-1 insert will have to be cloned and characterised. This work is now being undertaken.

10.4 Conclusions

This project sought to investigate the response of *M. smegmatis* to acid stress. Evidence was found for an adaptive acid tolerance mechanism that enhanced protection to normally lethal acidic conditions. A genomic locus that responds to acid stress by increasing transcription of a downstream gene has been identified. Analysis into these physiological and genetic systems that allow the response to acid are being investigated further.

Now that we have obtained some feel for the systems operating in *M. smegmatis*, a further step will be to investigate the acid response in *M. tuberculosis*. Adaptation to acidic environments in the host may be an important feature of this organism's virulence (section 7.4.1). Blocking this adaptive response to acid may prevent the bacillus setting up a successful infection. Therefore, this adaptive mechanism may prove an ideal drug target.

The presence of an acid tolerance response in *M. tuberculosis* may also have implications for vaccine development. Work by Hone *et al.* (1994) has shown that the *S. typhi* live oral vaccine strain Ty21a has a defective ATR response. This strain was produced by non-specific chemical mutagenesis, and the genetic basis for its attenuation is poorly characterised. Hone *et al.* (1994) suggested that this defective ATR may contribute to the attenuation of this organism. Extrapolating this work to *M. tuberculosis,* it may be possible to select for *M. tuberculosis* mutants defective in their response to acid. As a lot of work at the moment is focusing on the creation of a new generation of TB vaccines that will replace BCG, these strains may be useful as candidate vaccine strains.

REFERENCES

Abou-Zeid, C., Ratliff, T. L., Wiker, A. G., Harbe, M., Bennedson, J., and Rook, G. A. W., 1988. Characterisation of fibronectin-binding antigens released by *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG. Infection and Immunity, **56**:3046-3051.

Adams, L. B., Fukutomi, Y., and Krahenbuhl, J. L., 1993. Regulation of murine macrophage effector functions by lipoarabinomannan from mycobacterial strains with different degrees of virulence. Infection and Immunity, **61**:4173-4181.

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

Andersen, A. B., and Hansen, E. B., 1993. Cloning of the *lysA* gene from *Mycobacterium tuberculosis*. Gene, **124**:105-109.

Andrew, P. W., Jackett, P. S., and Lowrie, D. B., 1985. Killing and degradation of microorganisms by macrophages. In: Dean, R. T. and Jessup, W. (Editors), Mononuclear Phagocytes: Physiology and Pathology. Elsevier Science, Amsterdam, pp. 311-335.

Andrew, P. W., and Roberts, I. S., 1993. Construction of a bioluminescent *Mycobacterium* and its use for assay of antimycobacterial agents. Journal of Clinical Microbiology, **31**:2251-2254.

Armstrong, J. A., and Hart, P. D., 1971. Response of cultured macrophages to *Mycobacterium tuberculosis*, with observations on fusion of lysosomes with phagosomes. Journal of Experimental Medicine, **134**:713-740.

Arruda, S., Bomlim, G., Knights, R., Humia-Byron, T., and Riley, L. W., 1993. Cloning of a *Mycobacterium tuberculosis* DNA fragment associated with entry and survival inside cells. Science, **261**:1454-1457.

Auger, E. A., K.E., R., Plumb, T., Childs, L. C., Meng, S. Y., and Bennett, G. N., 1989. Construction of *lac* fusions to the inducible arginine- and lysine decarboxylase genes of *Escherichia coli* K12. Molecular Microbiology, **3**:609-620.

Bagg, A., and Neilands, J. B., 1987. Molecular mechanism of regulation of siderophore-mediated iron assimilation. Microbiological Reviews, **51**:509-518.

Banerjee, A., Dubnau, E., Quemard, A., Balasubramanian, V., Sun Um, K., Wilson, T., Collins, D., de Lisle, G., and Jacobs, W.R. Jr., 1994. *inhA*, a gene encoding a target for isoniazid and ethionamide in Mycobacterium tuberculosis. Science, **263**:227-230.

Bayles, K. W., 1994. Investigating environmental sensing mechanisms of *Staphylococcus aureus*. ASM news, **60**:597-599.

Behling, C. A., Bennett, B., Takayama, K., and Hunter, R. L., 1993. Development of a trehalose 6,6'-dimycolate model which explains cord formation by *Mycobacterium tuberculosis*. Infection and Immunity, **61**:2296-2303.

Belli, W. A., and Marquis, R. E., 1991. Adaptation of *Streptococcus mutans* and *Enterococcus hirae* to acid stress in continuous culture. Applied and Environmental Microbiology, **57**:1134-1138.

Bender, G. R., and Marquis, R. E., 1987. Membrane ATPases and acid tolerance of *Actinomyces viscous* and *Lactobacillus casei*. Applied and Environmental Microbiology, **53**:2124-2128.

Bercovier, H., Kafri, O., and Sela, S., 1986. Mycobacteria posses a surprisingly small number of ribosomal RNA genes in relation to the size of their genome. Biochemical and Biophysical Research Communications, **136**:1135-1141.

Bermudez, L. E., Young, L. S., Martinelli, J., and Petrofsky, M., 1993. Exposure to ethanol up-regulates the expression of *Mycobacterium avium* complex proteins associated with bacterial virulence. Journal of Infectious Diseases, **168**:961-968.

Bernstein, J., Jambor, W. P., Lott, W. A., Pansy, F., Steinberg, B. A., and Yale, H. L., 1952. Chemotherapy of experimental tuberculosis. V. Isonicotinic acid hydrazide and related compounds. American Review of Tuberculosis, **65**:357-364.

Biggin, M. D., Gibson, T. J., and Hong, G. F., 1983. Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. Proceedings of the National Academy of Sciences (USA), **80**:3963-3965.

Bishai, W. R., Smith, H. O., and Barcak, G. J., 1994. A peroxide/ascorbate-inducible catalase from *Haemophilus influenza* is homologous to the *Escherichia coli katE* gene product. Journal of Bacteriology, **176**:2914-2921.

Bloom, B. R., and Murray, C. J. L., 1992. Tuberculosis: commentary on a re-emergent killer. Science, **257**:1055-1064.

Bloom, B. R., and Fine, P. M., 1994. The BCG experience: Implications for Future Vaccines against Tuberculosis. In: Bloom, B. R. (Editor), Tuberculosis: Pathogenesis, Protection and Control. American Society for Microbiology, Washington, D.C., pp. 531-538.

Bloom, B. R., 1994. Preface. In: Bloom, B. R. (Editor), Tuberculosis: Pathogenesis, Protection, and Control. American Society for Microbiology Press, Washington D.C., pp. xiii-xiv.

Boddinghaus, B., Rogall, T., Flohr, T., Blocker, H., and Bottger, E. C., 1990. Detection and identification of mycobacteria by amplification of rRNA. Journal of Clinical Microbiology, **28**:1751-1759.

Bolivar, F., Roderigeuz, R. L., Greene, P. J., Betlach, M. C., Heyenker, H. L., Boyer, H. W., Crosa, J. H., and Falkow, S., 1977. Construction and characterisation of new cloning vehicles. II. A multipurpose cloning system. Gene, **2**: 95-113.

Booth, I. R., 1985. Regulation of cytoplasmic pH in bacteria. Microbiological Reviews, **49**:359-378.

Bradley, S. G., 1973. Relationships among mycobacteria and norcardiae based on deoxyribonucleic acid reassociation. Journal of Bacteriology, **113**:645-651.

Brausch, L. M., and Bass Jr, J. B., 1993. The Treatment of Tuberculosis. In: Bass Jr, J. R. (Editor), The Medical Clinics of North America - Tuberculosis. W.B. Saunders, 77, pp. 1277-1288.

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

Brosius, J., Ullrich, A., Raker, M. A., Gray, A., Dull, T. J., Gutell, R. R., and Noller, H. F., 1981. Construction and fine mapping of recombinant plasmids containing the *rrnB* ribosomal RNA operon of *Escherichia coli*. Plasmid, **6**:112-118.

Brosius, J., and Lupski, J. R., 1987. Plasmids for the selection and analysis of prokaryotic promoters. Methods in Enzymology, **153**:54-68.

Burlein, J. E., Stover, C. K., Offutt, S., and Hanson, M. S., 1994. Expression of foreign genes in mycobacteria. In: Bloom, B. R. (Editor), Tuberculosis: Pathogenesis, Protection and Control. American Society for Microbiology Press, Washington, D.C., pp. 239-252.

Cadenas, E., 1989. Biochemistry of oxygen toxicity. Annual Review of Biochemistry, **58**:79-110.

Carlioz, A., and Touati, D., 1986. Isolation of superoxide dismutase mutants in *Escherichia coli*: Is superoxide dismutase necessary for aerobic growth? EMBO Journal, **5**:623-630.

Carmi, O. A., Stewart, G. S. A. B., Ulitzer, S., and Kuhn, J., 1987. Use of bacterial luciferase to establish a promoter probe vehicle capable of nondestructive real-time analysis of gene expression in *Bacillus* spp. Journal of Bacteriology, **169**:2165-2170.

Chan, J., Fujiwara, T., Brennan, P. J., McNeil, M., Turco, S. J., Sibille, J.-C., Snapper, M., Aisen, P., and Bloom, B. R., 1989. Microbial glycolipids: Possible virulence factors that scavenge oxygen radicals. Proceedings of the National Academy of Sciences (USA), **86**:2453-2457.

Chan, J., Fan, X., Hunter, S. W., Brennan, P. J., and Bloom, B. R., 1991. Lipoarabinomannan, a possible virulence factor involved in persistence of *Mycobacterium tuberculosis* within macrophages. Infection and Immunity, 59:1755-1761.

Chan, J., Xing, Y., Magliozzo, R. S., and Bloom, B. R., 1992. Killing of virulent *Mycobacterium tuberculosis* by reactive nitrogen intermediates produced by activated murine macrophages. Journal of Experimental Medicine, **175**:1111-1122.

Chan, J., and Kaufmann, S. H. E., 1994. Immune mechanisms of protection. In: Bloom, B. R. (Editor), Tuberculosis: Pathogenesis, Protection and Control. American Society for Microbiology Press, Washington D.C., pp. 389-415.

Chapman, J. S., and Bernard, J. S., 1962. The tolerances of unclassified mycobacteria. American Review of Respiratory Disease, **86**:582-583.

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

Clairborne, A., and Fridovich, I., 1979. Purification of the o-dianisidine peroxidase from *Escherichia coli* B. Journal of Biological Chemistry, **254**:4245-4252.

Clark-Curtiss, J. E., Jacobs Jr, W. R., Docherty, M. A., Ritchie, L. R., and Curtiss III, R., 1985. Molecular analysis of DNA and construction of genomic libraries of *Mycobacterium leprae*. Journal of Bacteriology, **161**:1093-1102.

Clark-Curtiss, J. E., 1990. Genome structure of mycobacteria. In: McFadden, J. J. (Editor), Molecular Biology of the Mycobacteria. Surrey University Press, London, pp. 77-96.

Clave, C., and Trombe, M. C., 1989. DNA uptake in competent *Streptococcus pneumoniae* requires ATP and is regulated by cytoplasmic pH. FEMS Microbiology Letters, **65**:113-118.

Cole, S. T., 1994. *Mycobacterium tuberculosis*: drug-resistance mechanisms. Trends in Microbiology, **2**:411-415.

Collins, C. H., Grange, J. M., and Yates, M. D., 1984. Mycobacteria in water. Journal of Applied Bacteriology, **57**:193-211.

Colston, M. J., 1991. Rebirth of a star performer. Nature, **351**:442-443.

Colston, M. J., and Davis, E. O., 1994. The ins and outs of protein splicing elements. Molecular Microbiology, **12**:359-363.

Connell, N. D., Medina-Acosta, E., McMaster, W. R., Bloom, B. R., and Russell, D. G., 1993. Effective immunization against cutaneous leishmaniasis with recombinant BCG expressing the *Leishmania* surface proteinase (gp63). Proceedings of the National Academy of Sciences (USA), **90**:11473-11477.

Content, J., de la Cuvellerie, A., De Wit, L., Vincent-Levy-Frebault, V., Ooms, J., and De Bruyn, J., 1991. The genes for the antigen 85 complex of *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG are members of a gene family: Cloning, sequence determination, and genomic organisation of the gene encoding for antigen 85-C of *M. tuberculosis*. Infection and Immunity, **59**:3205-3212.

Cooksey, R. C., Crawford, J. T., Jacobs, W. R., and Shinnick, T. M., 1993. A rapid method for screening antimicrobial agents for activities against a strain of *Mycobacterium tuberculosis* expressing firefly luciferase. Antimicrobial Agents and Chemotherapy, **37**:1348-1352.

Crawford, J. T., and Bates, J. H., 1979. Isolation of plasmids from mycobacteria. Infection and Immunity, **24**:979-981.

Crawford, J. T., Cave, M. D., and Falkinham, J. O., 1981. Evidence for plasmid-mediated restriction-modification in *Mycobacterium-avium-intracellulare*. Journal of General Microbiology, **127**:333-338.

Crawford, J. T., and Falkinham III, J. O., 1986. Analysis of plasmids in *Mycobacterium avium-intracellulare* isolates from persons with acquired immunodeficiency syndrome. American Review of Respiratory Disease, **134**:659-661.

Crawford, J. T., and Falkinham III, J. O., 1990. Plasmids of the *Mycobacterium avium* complex. In: McFadden, J. J. (Editor), The Molecular Biology of the Mycobacteria. Surrey University Press, London, pp. 97-119.

Crowle, A. J., Dahl, R., Ross, E., and May, M. H., 1991. Evidence that vesicles containing living, virulent *Mycobacterium tuberculosis* or *Mycobacterium avium* in cultured human macrophages are not acidic. Infection and Immunity, **59**:1823-1831.

Csonka, L. N., and Hanson, A. D., 1991. Prokaryotic osmoregulation: genetics and physiology. Annual Review of Microbiology, **45**:569-606.

Dale, J. W., and Patki, A., 1990. Mycobacterial gene expression and regulation. In: McFadden, J. J. (Editor), The Molecular Biology of the Mycobacteria. Academic Press, London, pp. 173-198.

Daniel, T. M., 1982. Robert Koch, tuberculosis, and the subsequent history of medicine. American Review of Respiratory Disease, **125**(pt.2):1-3.

Daniel, T. M., Bates, J. H., and Downes, K. A., 1994. History of Tuberculosis. In: Bloom, B. R. (Editor), Tuberculosis: Pathogenesis, Protection and Control. American Society for Microbiology, Washington D.C., pp. 13-24.

Dannenberg, A. M., and Rook, G. A. W., 1994. Pathogenesis of Pulmonary Tuberculosis: an interplay of tissue-damaging and macrophage-activating immune responses - dual mechanisms that control bacillary multiplication. In: Bloom, B. R. (Editor), Tuberculosis: Pathogenesis, Protection and Control. American Society for Microbiology Press, Washington, D.C., pp. 459-483.

David, M., Lubinsky-Mink, S., Ben-Svi, A., Ulitzur, S., Kuhn, J., and Suissa, M., 1992. A stable *Escheerichia coli-Mycobacterium smegmatis* shuttle vector containing the mycobacteriophage D29 origin. Plasmid, **28**:267-271.

Davis, E. O., Sedgwick, S. G., and Colston, M. J., 1991. Novel structure of the *recA* locus of *Mycobacterium tuberculosis* implies processing of the gene product. Journal of Bacteriology, **173**:5653-5662.

Davis, E. O., Jenner, P. J., Brooks, P. C., Colston, M. J., and Sedgwick, S. G., 1992. Protein splicing in the maturation of *Mycobacterium tuberculosis* RecA protein: A mechanism for tolerating a novel class of intervening sequence. Cell, **71**:201-210.

Davis, E. O., Thangaraj, H. S., Brooks, P. C., and Colston, M. J., 1994. Evidence of selection for protein introns in the RecAs of pathogenic mycobacteria. EMBO Journal, **13**:699-703.

Dellagostin, O. A., Wall, S., Norman, E., O'Shaughnessey, T., Dale, J. W., and McFadden, J. J., 1993. Construction and use of integrative vectors to express foreign genes in mycobacteria. Molecular Microbiology, **10**:983-993.

Demple, B., 1991. Regulation of bacterial oxidative stress genes. Annual Review of Genetics, **25**:315-337.

Domenech, P., Menendez, M. C., and Garcia, M. J., 1994. Restriction fragment length polymorphisms of 16S rRNA genes in the differentiation of fast-growing mycobacterial species. FEMS Microbiology Letters, **116**:19-24.

Don, R. H., Cox, P. T., Wainwright, B. J., Baker, K., and Mattick, J. S., 1991. 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nucleic Acids Research, **19**:4008.

Dower, W. J., Miller, J. F., and Ragsdale, C. W., 1988. High-efficiency transformation of *Escherichia coli* by high-voltage electroporation. Nucleic Acids Research, **16**:6127-6145.

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

Draper, P., 1981. Mycobacterial inhibition of intracellular killing. In: O'Grady, F. and Smith, H. (Editors), Microbial Perturbation of Host Defenses. Academic Press, London, pp. 143-164.

Draper, P., 1982. The Anatomy of Mycobacteria. In: Ratledge, C. and Stanford, J. (Editors), The Biology of the Mycobacteria. Academic Press, London, 1, pp. 9-52.

Drea, W. F. and Andrejew, A., 1953. Nutrition and growth: nitrogenous nutrition. In: The metabolism of the tubercle bacillus. Charles C. Thomas, Springfield, Illinois., pp. 40-49.

Dubos, R. J., 1950. The effect of organic acids on mammalian tubercle bacilli. Journal of Experimental Medicine, **92**:319-332.

Dubos, R. J. and Dubos, J., 1952. The White Plague: Little, Brown and Company, Boston, 277 pp.

Dubos, R. J., 1953. Effect of the composition of the gaseous and aqueous environments on the survival of tubercle bacilli *in vitro*. Journal of Experimental Medicine, **97**:357-366.

Dubos, R. J., 1955. The micro-environment of inflammation, or Metchnikoff revisited. Lancet, **269**:1-5.

Dybvig, K., Hollinhgshead, S. K., Heath, D. G., Clewell, D. B., Sun, F., and Woodward, A., 1992. Degenerate oligonucleotide primers for enzymatic amplification of *recA* sequences form Gram-positive bacteria and mycoplasmas. Journal of Bacteriology, **174**:2729-2732.

Eiglmeier, K., Honore, N., Woods, S. A., Caudron, B., and Cole, S. T., 1993. Use of an ordered cosmid library to deduce the genomic organisation of *Mycobacterium leprae*. Molecular Microbiology, **7**:197-206.

Eisenach, K. D., Crawdord, J. T., and Bates, J. H., 1988. Repetitive DNA sequences as probes for *Mycobacterium tuberculosis*. Journal of Clinical Microbiology, **26**:2240-2245.

Engebrecht, J., Simon, M., and Silverman, M., 1985. Measuring gene expression with light. Science, **227**:1345-1347.

Eradi, F. X., Failla, M. L., and Falkinham III, J. O., 1987. Plasmid-encoded copper resistance and precipitation by *Mycobacterium scrofulaceum*. Applied and Environmental Microbiology, **53**:1951-1954.

Falkinham III, J. O., and Crawford, J. T., 1994. Plasmids. In: Bloom, B. R. (Editor), Tuberculosis: Pathogenesis, Protection and Control. American Society for Microbiology Press, Washington, D.C., pp. 185-198.

Fang, F. C., Libby, S. J., Buchmeimer, N. A., Loewen, P. C., Switala, J., Harwood, J., and Guiney, D. G., 1992. The alternative sigma factor KatF (RpoS) regulates *Salmonella* virulence. Proceedings of the National Academy of Sciences (USA), **89**:11978-11982.

Farber, J., and Peterkin, P., 1991. *Listeria monocytogenes*, a food-borne pathogen. Microbiological Reviews, **55**:476-511.

Farr, S. B., and Kogoma, T., 1991. Oxidative stress responses in *Escherichia coli* and *Salmonella typhimurium*. Microbiological Reviews, **55**:561-585.

Fields, P. I., Swanson, R. V., Haidaris, C. G., and Heffron, F., 1986. Mutants of *Salmonella typhimurium* that cannot survive within the macrophage are avirulent. Proceedings of the National Academy of Sciences (USA), **83**:5189-5193.

Fields, P. I., Groisman, E. A., and Heffron, F., 1989. A *Salmonella* locus that controls resistance to microbiocidal proteins from phagocytic cells. Science, **243**:1059-1062.

Fine, P. E. M., 1988. BCG vaccination against tuberculosis and leprosy. British Medical Journal, **44**:691-703.

Flesch, I. E. A., and Kaufmann, S. H. E., 1988. Attempts to characterise the mechanisms involved in mycobacterial growth inhibition by gamma-interferon-activated bone marrow macrophages. Infection and Immunity, **56**:1464.

Forsberg, A. J., Pavitt, G. D., and Higgins, C. F., 1994. Use of transcriptional fusions to monitor gene expression: a cautionary tale. Journal of Bacteriology, **176**:2128-2132.

Foster, J. W., and Hall, H. K., 1990. Adaptive acidification tolerance response of *Salmonella typhimurium*. Journal of Bacteriology, **172**:771-778.

Foster, J. W., and Hall, H. K., 1991. Inducible pH homeostasis and the acid tolerance response of *Salmonella typhimurium*. Journal of Bacteriology, **173**:5129-5135.

Foster, J. W., 1991. *Salmonella* acid shock proteins are required for the adaptive acid tolerance response. Journal of Bacteriology, **173**:6896-6902.

Foster, J. W., and Hall, H. K., 1992. Effect of *Salmonella typhimurium* ferric uptake regulator (*fur*) mutations on iron- and pH-regulated protein synthesis. Journal of Bacteriology, **174**:4317-4323.

Foster, J. W., 1993. The acid tolerance response of *Salmonella typhimurium* involves transient synthesis of key acid shock proteins. Journal of Bacteriology, **175**:1981-1987.

Foster, J. W., and Bearson, B., 1994. Acid-sensitive mutants of *Salmonella typhimurium* identified through a dinitrophenol lethal screening strategy. Journal of Bacteriology, **176**:2596-2602.

Francis, K. P., and Gallagher, M. P., 1993. Light emission form a Mudlux transcriptional fusion in *Salmonella typhimurium* is stimulated by hydrogen peroxide and by interaction with the mouse macrophage cell line J774.2. Infection and Immunity, **61**:640-649.

Franzon, V. L., Arondel, J., and Sansonetti, P. J., 1990. Contribution of superoxide dismutase and catalase activities to *Shigella flexneri* pathogenesis. Infection and Immunity, **58**:529-535.

Fridovich, I., 1978. The biology of oxygen radicals. Science, **201**:875-880.

Frothingham, R., Hills, H. G., and Wilson, K. H., 1994. Extensive DNA sequence conservation throughout the *Mycobacterium tuberculosis* complex. Journal of Clinical Microbiology, **32**:1639-1643.

Gaillard, J. L., Berche, P., and Sansonetti, P., 1986. Transposon mutagenesis as a tool to study the role of hemolysin in the virulence of *Listeria monocytogenes*. Infection and Immunity, **52**:50-55.

Garbe, T., Jones, C., Charles, I., Dougan, G., and Young, D., 1990. Cloning and characterisation of the *aroA* gene from *Mycobacterium tuberculosis*. Journal of Bacteriology, **172**:6774-6782.

Garbe, T., Harris, D., Vordermeier, M., Lathigra, R., Ivanyi, J., and Young, D., 1993. Expression of the *Mycobacterium tuberculosis* 19-kilodalton antigen in *Mycobacterium smegmatis*: Immunological analysis and evidence of glycosylation. Infection and Immunity, **61**:260-267.

Garcia-del Portillo, F., Foster, J. W., Maguire, M. E., and Finlay, B. B., 1992. Characterisation of the micro-environment of *Salmonella typhimurium*-containing vacuoles within MDCK epithelial cells. Molecular Microbiology, **6**:3289-3297.

Gardner, G. M., and Weiser, R. S., 1947. A bacteriophage from *Mycobacterium smegmatis*. Proceedings of the Society for Experimental Biology and Medicine, **66**:205-206.

Gauthier, M. J., and Clement, R. L., 1994. Effect of a short period of starvation in oligotrophic waters on the resistance of enteric bacterial pathogens to gastric pH conditions. FEMS Microbiology Ecology, **14**:275-284.

Gayatri Devi, B., Shaila, M. S., Ramakrishna, T., and Gopinathan, K. P., 1975. The purification and properties of peroxidase in *Mycobacterium tuberculosis* H37Rv and its possible role in the mechanism of action of isonicotinic acid hydrazide. Biochemical Journal, **149**:187-197.

Gerchman, Y., Olami, Y., Rimon, A., Taglicht, D., Schulinder, S., and Padan, E., 1993. Histidine-226 is part of the pH sensor of NhaA, a Na^+/H^+ antiporter in *Escherichia coli*. Proceedings of the National Academy of Sciences (USA), **90**:1212-1216.

Glenn, A. R., and Gilworth, M. J., 1994. The life of root nodule bacteria in the acidic underground. FEMS Microbiology Letters, **123**:1-10.

Goldberg, E. B., Arbel, T., Chen, J., Karpel, R., Mackie, G. A., Schulinder, S., and Padan, E., 1987. Characterisation of a Na+/H+ antiporter gene of *Escherichia coli*. Proceedings of the National Academy of Sciences (USA), **84**:2615-2619.

Gonzalez-Flecha, B., and Demple, B., 1994. Intracellular generation of superoxide as a by-product of *Vibrio harveyi* luciferase expressed in *Escherichia coli*. Journal of Bacteriology, **176**:2293-2299.

Goodfellow, M., and Wayne, L. G., 1982. Taxonomy and Nomenclature. In: Ratledge, C. and Stanford, J. (Editors), The Biology of the Mycobacteria. Academic Press, London, 1, pp. 471-521.

Goodson, M., and Rowbury, R. J., 1989. Habituation to normally lethal acidity by prior growth of *Escherichia coli* at a sub-lethal acid pH value. Letters in Applied Microbiology, **8**:77-79.

Gorden, J., and Small, P. L. C., 1993. Acid resistance in enteric bacteria. Infection and Immunity, **61**:364-367.

Goren, M. B., Hart, P. D., 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.

Goulbourne, E., Matin, M., Zychlinsky, E., and Matin, A., 1986. Mechanisms of pH maintenance in active and inactive cells of an obligately acidophilic bacterium. Journal of Bacteriology, **166**:59-65.

Graham, P. H., Draeger, K. J., Ferrey, M. L., Conroy, M. J., Hammer, B. E., Martinez, E., Aarons, S. R., and Quinto, C., 1994. Acid pH tolerance in strains of *Rhizobium* and *Bradyrhizobium*, and initial studies on the basis for acid tolerance of *Rhizobium tropici* UMR1899. Canadian Journal of Microbiology, **40**:198-207.

Grange, J. M., 1982. The Genetics of Mycobacteria and Mycobacteriophages. In: Ratledge, C. and Stanford, J. (Editors), The Biology of the Mycobacteria. Academic Press, London, 1, pp. 309-351.

Grange, J. M., 1988. Mycobacteria and Human Disease: Edward Arnold, London

Green, E. P., Tizard, M. L. V., Moss, M. T., Thompson, J., Winterbourne, D. J., McFadden, J. J., and Herman-Taylor, J., 1989. Sequence and characteristics of IS900, an insertion element identified in a human Crohn's disease isolate of *Mycobacterium paratuberculosis*. Nucleic Acids Research, **17**:9063-9072.

Gross, W. M., and Wayne, L. G., 1970. Nucleic Acid Homology in the Genus *Mycobacterium*. Journal of Bacteriology, **104**:630-634.

Gross, R., Arico, B., and Rappuoli, R., 1989. Families of bacterial signal-transducing proteins. Molecular Microbiology, **3**:1661-1667.

Grosskinsky, C. M., Jacobs Jr, W. R., Clark-Curtiss, J. E., and Bloom, B. R., 1989. Genetic Relationships among *Mycobacterium leprae*, *Mycobacterium tuberculosis*, and candidate Leprosy vaccine strains determined by DNA hybridisation: Identification of an *M. leprae*-specific repetitive sequence. Infection and Immunity, **57**:1535-1541.

Genetics Computer Group, 1991. Program manual for GCG package, Version 7-UNIX. Madison, Wisconsin, USA.

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

Gupta, S. K. D., Bashyam, M. D., and Tyagi, A. K., 1993. Cloning and assessment of mycobacterial promoters by using a plasmid shuttle vector. Journal of Bacteriology, **175**:5186-5192.

Haas, A., Brehm, K., Kreft, J., and Goebel, W., 1991. Cloning, characterisation and expression in *Escherichia coli* of a gene encoding *Listeria seeligeri* catalase, a bacterial enzyme highly homologous to mammalian catalases. Journal of Bacteriology, **173**:5159-5167.

Hastings, J. W., 1978. Bacterial Bioluminescence: an overview. Methods in Enzymology, **57**:125-135.

Hastings, J. W., Baldwin, T. O., and Nicoli, M. Z., 1978. Bacterial luciferase: assay, purification, and properties. Methods in Enzymology, **57**:135-152.

Hastings, J. W. and Gibson, Q. H., 1963. Intermediates in the bioluminescent oxidation of reduced flavin mononucleotide. Journal of Biological Chemistry, **238**: 2537-2554.

Hatfull, G. F., and Sarkis, G. J., 1993. DNA sequence, structure and gene expression of mycobacteriophage L5: a phage system for mycobacterial genetics. Molecular Microbiology, **7**:395-405.

Hatfull, G. F., and Jacobs Jr, W. R., 1994. Mycobacteriophages: Cornerstones of mycobacterial research. In: Bloom, B. R. (Editor), Tuberculosis: Pathogenesis, Protection and Control. American Society for Microbiology Press, Washington, D.C., pp. 165-183.

Heifets, L. B., and Good, R. C., 1994. Current laboratory methods for the diagnosis of tuberculosis. In: Bloom, B. R. (Editor), Tuberculosis: Pathogenesis, Protection and Control. American Society for Microbiology Press, Washington, D.C., pp. 85-110.

Hermans, P. W. M., van Soolingen, D., Dale, J. W., Schuitema, A. R. J., McAdam, R. A., Catty, D., and van Embden, J. D. A., 1990(a). Insertion element IS986 from *Mycobacterium tuberculosis*: a useful tool for diagnosis and epidemiology of tuberculosis. Journal of Clinical Microbiology, **28**:2051-2058.

Hermans, J., Boschloo, J. G., and de Bont, J. A. M., 1990(b). Transformation of *Mycobacterium aurum* by electroporation: the use of glycine, lysozyme and isonicotinic acid hydrazide in enhancing transformation efficiency. FEMS Microbiology Letters, **72**:221-224.

Hermans, J., Martin, C., Huijberts, G. N. M., Goosen, T., and deBont, J. A. M., 1991. Transformation of *Mycobacterium aurum* and *Mycobacterium smegmatis* with the broad host-range Gram-negative cosmid vector pJRD215. Molecular Microbiology, **5**:1561-1566.

Hermans, P. W. M., van Soolingen, D., and Embden, J. D. A., 1992. Characterisation of a major polymorphic tandem repeat in *Mycobacterium tuberculosis* and its potential use in the epidemiology of *Mycobacterium kansasii* and *Mycobacterium gordonae*. Journal of Bacteriology, **174**:4157-4165.

Heyde, M., and Portalier, R., 1990. Acid shock proteins of *Escherichia coli*. FEMS Microbiology Letters, **69**:19-26.

Hickey, E. W., and Hirshfield, I. N., 1990. Low-pH-induced effects on patterns of protein synthesis and on internal pH in *Escherichia coli* and *Salmonella typhimurium*. Applied and Environmental Microbiology, **56**:1038-1045.

Hicks, D. B., Plass, R. J., and Quirk, P. G., 1991. Evidence for multiple terminal oxidases, including cytochrome d, in facultatively alkaliphilic *Bacillus firmus* OF4. Journal of Bacteriology, **173**:5010-5016.

Higgins, D. G., Bleasby, A. J., and Fuchs, R., 1992. CLUSTAL V: improved software for multiple sequence alignment. Computer Applications in the Biosciences, **8**:189-191.

Hill, P. J., Swift, S., and Stewart, G. S. A. B., 1991. PCR based gene engineering of the *Vibrio harveyi lux* operon and the *Escherichia coli trp* operon provide for biochemically functional native and fused gene products. Molecular and General Genetics, 226:41-48

Hill, P. J., Rees, C. E. D., Winson, M. K., and Stewart, G. S. A. B., 1993. The application of *lux* genes. Biotechnology and Applied Biochemistry, **17**:3-14.

Hochman, A., and Shemesh, A., 1987. Purification and characterisation of a catalase-peroxidase from the photosynthetic bacterium *Rhodopseudomonas capsulata*. The Journal of Biological Chemistry, **262**:6871-6876.

Hoiseth, S. K., and Stocker, B. A. D., 1981. Aromatic-dependent *Salmonella typhimurium* are non-virulent and are effective live vaccines. Nature, **241**:238-239.

Hone, D. M., Harris, A. M., and Levine, M. M., 1994. Adaptive acid tolerance response by *Salmonella typhi* and candidate live oral typhoid strains. Vaccine, **12**:895-898.

Hopewell, P. C., 1994. "The cure": organisation and administration of therapy for tuberculosis. In: Porter, J. D. H. and. McAdam, K.P.W.J. (Editors), Tuberculosis: Back to the Future. John Wiley and Sons, Chichester, England, pp. 99-121.

Hopwood, D. A., and Wright, H. M., 1978. Bacterial protoplast fusion: recombination in fused protoplasts of *Streptomyces coelicolor*. Molecular and General Genetics, **162**:307-317.

Hopwood, D. A., Keiser, T., Colston, M. J., and Lamb, F. I., 1988. Molecular biology of mycobacteria. British Medical Bulletin, **44**:528-546.

Husson, R. N., and Young, R. A., 1987. Genes for the major protein antigens of *Mycobacterium tuberculosis*: The aetiological agents of tuberculosis and leprosy share an immunodominant antigen. Proceedings of the National Academy of Sciences (USA), **84**:1679-1683.

Imeada, T., Kirchheimer, W. F., and Barksdale, L., 1982. DNA isolated from *Mycobacterium leprae*: Genome size, base ratio, and homology with other related bacteria as determined by optical DNA-DNA reassociation. Journal of Bacteriology, **150**:414-417.

Ishizuka, H., Horinouchi, S., Keiser, H. M., Hopwood, D. A., and Beppu, T., 1992. A putative two-component regulatory system involved in secondary metabolism in *Streptomyces* spp. Journal of Bacteriology, **174**:7585-7594.

Jackett, P. S., Aber, V. R., and Lowrie, D. B., 1978. Virulence and resistance to superoxide, low pH and hydrogen peroxide among strains of *Mycobacterium tuberculosis*. Journal of General Microbiology, **104**:37-45.

Jackett, P. S., Aber, V. R., and Mitchison, D. A., 1977. Relationship between antibacterial activity of isoniazid and nicotinamide adenine dinucleotide concentration in Mycobacterium tuberculosis. American Review of Respiratory Disease, **115**:601-607.

Jacobs Jr, W. R., Docherty, M. A., Curtiss III, R., and Clark-Curtiss, J. E., 1986. Expression of *Mycobacterium leprae* genes from a *Streptococcus mutans* promoter in *Escherichia coli* K-12. Proceedings of the National Academy of Sciences (USA), **83**:1926-1930.

Jacobs Jr, W. R., Tuckman, M., and Bloom, B. R., 1987. Introduction of foreign DNA into mycobacteria using a shuttle phasmid. Nature, **327**:532-535.

Jacobs Jr, W. R., Kalpana, G. V., Cirillo, J. D., Pascopella, L., Snapper, S. B., Udani, R. A., Jones, W., Barletta, R. G., and Bloom, B. R., 1991. Genetic Systems for Mycobacteria. Methods in Enzymology, **204**:537-555.

Jacobs Jr, W. R., Barletta, R. G., Udani, R., Chan, J., Kalkut, G., Sosne, G., Keiser, T., Sarkis, G. J., Hatfull, G. F., and Bloom, B. R., 1993. Rapid assessment of drug susceptibilities of *Mycobacterium tuberculosis* by means of Luciferase Reporter Phages. Science, **260**:819-822.

Jarlier, V., and Nikaido, H., 1990. Permeability barrier to hydrophilic solutes in Mycobacterium chelonae. Journal of Bacteriology, **172**:1418-1423.

Jarlier, V., and Nikaido, H., 1994. Mycobacterial cell wall: structure and role in natural resistance to antibiotics. FEMS Microbiological Reviews, **123**:11-18.

Ji, Y., Colston, M. J., and Cox, R. A., 1994. Nucleotide sequence and secondary structures of precursor 16s rRNAs of slow-growing mycobacteria. Microbiology, **140**:123-132.

Johnson, S. R., Steiner, B. M., Cruce, D. D., Perkins, G. H., and Arko, R. J., 1993. Characterisation of a catalase-deficient strain of *Neisseria gonorrhoeae*: Evidence for the significance of catalase in the biology of *N. gonorrhoeae*. Infection and Immunity, **61**:1232-1238.

Jones, D., 1990. Foodborne listeriosis. Lancet, **336**:1171-1174.

Kado, C. I., and Liu, S. T., 1981. Rapid procedure for detection and isolation of large and small plasmids. Journal of Bacteriology, **145**:1365-1373.

Karem, K. L., Foster, J. W., and Bej, A. K., 1994. Adaptive acid tolerance response (ATR) in *Aeromonas hydrophila*. Microbiology, **140**:1731-1736.

Karpel, R., Alon, T., Glaser, G., Schulinder, S., and Padan, E., 1991. Expression of a sodium proton antiporter (NhaA) in *Escherichia coli* is induced by Na^+ and Li^+. Journal of Biological Chemistry, **266**:21753-21759.

Kazda, J., Muller, K., and Irgens, L. M., 1979. Cultivable mycobacteria in sphagnum vegetation of moors in south Sweden and coastal Norway. Acta Pathologica et Microbiologica Scandinavica, **87B**:97-101.

Keiser, T., Moss, M. T., Dale, J. W., and Hopwood, D. A., 1986. Cloning and expression of *Mycobacterium bovis* BCG DNA in *Streptomyces lividans*. Journal of Bacteriology, **168**:72-80.

King, C. H., Mundayoor, S., Crawford, J. T., and Shinnick, T. M., 1993. Expression of a contact-dependant cytolytic activity of *Mycobacterium tuberculosis* and isolation of the locus that encodes the activity. Infection and Immunity, **61**:2708-2712.

Kirby, T., Blum, J., Kahane, I., and Fridovich, I., 1980. Distinguishing between Mn-containing and Fe-containing superoxide dismutases in the crude extracts of cells. Archives of Biochemistry and Biophysics, **210**:551-555.

Knauf, H. J., Vogel, R. F., and Hammes, W. P., 1992. Cloning, sequence and phenotypic expression of *katA*, which encodes the catalase of *Lactobacillus sake* LTH677. Applied and Environmental Microbiology, **58**:832-839.

Kobayashi, H., Suzuki, T., Kinoshita, N., and Unemoto, T., 1984. Amplification of the *Streptococcus faecalis* proton-translocating ATPase by a decrease in cytoplasmic pH. Journal of Bacteriology, **158**:1157-1160.

Kobayashi, H., 1985. A proton-translocating ATPase regulates pH of the bacterial cytoplasm. Journal of Biological Chemistry, **260**:72-76.

Koch, R., 1932. Die Aetiologie der Tuberculose (translated from the original by B. Pinner and M. Pinner). American Review of Tuberculosis and Pulmonary Disease, **25**:285-323.

Krishna Murti, C. R., 1975. Isonicotinic acid hydrazide. In: Corcoran, V. W. and Hahn, F. E. (Editors), Antibiotics III. Springer Verlag, Berlin, pp. 621-652.

Kroll, R. G., and Booth, I. R., 1981. The role of potassium transport in the generation of a pH gradient in *Escherichia coli*. Biochemical Journal, **198**:691-698.

Kroll, R. G., and Patchett, R. A., 1992. Induced acid tolerance in *Listeria monocytogenes*. Letters in Applied Microbiology, **14**:224-227.

Krulwich, T. A., Agus, R., Schneier, M., and Guffanti, A. A., 1985. Buffering capacity of bacilli that grow at different pH ranges. Journal of Bacteriology, **162**:768-772.

Krulwich, T. A., and Guffanti, A. A., 1986. Regulation of internal pH in acidophilic and alkalophilic bacteria. Methods in Enzymology, **125**:352-365.

Kureishi, A., and Bryan, L. E., 1991. Pre-boiling high GC content, mixed primers with 3' complementation allows the successful PCR amplification of *Pseudomonas aeruginosa* DNA. Nucleic Acids Research, **20**:1155.

Labidi, A., David, H. L., and Roulland-Dussoix, D., 1985. Cloning and expression of mycobacterial plasmid DNA in *Escherichia coli*. FEMS Microbiology Letters, **30**:221-225.

Laemmli, U. K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature, **227**:680-685.

Lange, R., and Hengge-Aronis, R., 1991. Identification of a central regulator of stationary phase gene expression in *Escherichia coli*. Molecular Microbiology, **5**:49-59.

Lazraq, R., Clavel-Seres, S., David, H. L., and Roulland-Dussoix, D., 1990. Conjugative transfer of a shuttle plasmid for *Escherichia coli* to *Mycobacterium smegmatis*. FEMS Microbiology Letters, **69**:135-138.

Lee, M. H., Pascopella, L., Jacobs Jr, W. R., and Hatfull, G. F., 1991. Site-specific integration of mycobacteriophage L5: Integration-proficient vectors for *Mycobacterium smegmatis*, *Mycobacterium tuberculosis*, and bacille Calmette-Guerin. Proceedings of the National Academy of Sciences (USA), **88**:311-3115.

Lee, I. S., Slonczewski, J. L., and Foster, J. W., 1994. A low-pH-inducible, stationary-phase acid tolerance response in *Salmonella typhimurium*. Journal of Bacteriology, **176**:1422-1426.

Li, Y., Severn, A., Rogers, M. V., Palmer, R. M. J., Moncada, S., and Liew, F. Y., 1992. Catalase inhibits nitric oxide synthesis and the killing of intracellular *Leishmania major* in murine macrophages. European Journal of Immunology, **22**:441-446.

Li, Z., and Meighen, E. A., 1994. The turnover of bacterial luciferase is limited by a slow decomposition of the ternary enzyme-product complex of luciferase, FMN and fatty acid. Journal of Biological Chemistry, **269**:6640-6644.

Loewen, P. C., 1979. Levels of glutathione in *Escherichia coli*. Canadian Journal of Biochemistry, **57**:107-111.

Loewen, P. C., Switala, J., and Triggs-Raine, B. L., 1985. Catalases HPI and HPII in *Escherichia coli* are induced independently. Archives of Biochemistry and Biophysics, **243**:144-149.

Lowrie, D. B., Jackett, P. S., and Ratcliffe, N. A., 1975. *Mycobacterium microti* may protect itself from intracellular destruction by releasing cyclic AMP into phagosomes. Nature, **254**:600-602.

Mackey, B. M., Cross, D., and Park, S. F., 1994. Thermostability of bacterial luciferase expressed in different microbes. Journal of Applied Microbiology, **77**:149-154.

Mahenthiralingam, E., Draper, P., Davis, E. O., and Colston, M. J., 1993. Cloning and sequencing of the gene which encodes the highly inducible acetamidase of *Mycobacterium smegmatis*. Journal of General Microbiology, **139**:575-583.

Mandell, G. L., 1975. Catalase, superoxide dismutase and virulence of *Staphylococcus aureus*: *in vitro* and *in vivo* studies with emphasis on staphylococcal-leukocyte interaction. Journal of Clinical Investigation, **55**:561-566.

Marchuk, D., Drumm, M., Saulino, A., and Collins, F. S., 1990. Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. Nucleic acids Research, **5**:1154.

Martin, M. E., Rowe-Byers, B., Olson, M. O. J., Salin, M. L., Arceneaux, J. E. L., and Tolbert, C., 1986. A *Streptococcus mutans* superoxide dismutase that is active with either manganese or iron as a cofactor. The Journal of Biological Chemistry, **261**:9361-9367.

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

Matin, A., 1991. The molecular basis of carbon-starvation-induced general resistance in *Escherichia coli*. Molecular Microbiology, **5**:3-10.

Matsuo, K., Yamaguchi, R., Yamazaki, A., Tasaka, T., Terasaka, K., Totsuka, M., Kobayashi, K., Yukitake, H., and Yamada, T., 1990. Establishment of a foreign secretion system in mycobacteria. Infection and Immunity, **58**:4049-4054.

McCann, M. P., Kidwell, J. P., and Matin, A., 1991. The putative sigma factor KatF has a central role in development of starvation-mediated resistance in *Escherichia coli*. Journal of Bacteriology, **173**:4188-4194.

McDonough, K. A., Kress, Y., and Bloom, B. R., 1993. Pathogenesis of tuberculosis: interaction of *Mycobacterium tuberculosis* with macrophages. Infection and Immunity, **61**:2763-2773.

McFadden, J. J., Butcher, P. D., Chiodini, R. J., and Hermon-Taylor, J., 1987. Determination of genome size and DNA homology between and unclassified *Mycobacterium* species isolated from patients with Crohn's disease and other mycobacteria. Journal of General Microbiology, **113**:211-214.

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.

Mehra, V., Sweetser, D., and Young, R. A., 1986. Efficient mapping of protein antigenic determinants. Proceedings of the National Academy of Sciences (USA), **83**:7013-7017.

Meighen, E. A., 1988. Enzymes and genes from the *lux* operons of bioluminescent bacteria. Annual Review of Microbiology, **42**:151-176.

Meighen, E. A., 1991. Molecular biology of bacterial bioluminescence. Microbiological Reviews, **55**:123-142.

Meissner, P. S., and Falkinham III, J. O., 1984. Plasmid-encoded mercuric reductase in *Mycobacterium scrofulaceum*. Journal of Bacteriology, **157**:669-672.

Meissner, P. S., and Falkinham III, J. O., 1986. Plasmid DNA profiles as epidemiological markers for clinical and environmental isolates of *Mycobacterium avium, Mycobacterium intracellulare,* and *Mycobacterium scrofulaceum.* Journal of Infectious Disease, **153**:325-331.

Melton, A. R., and Weiss, A. A., 1993. Characterisation of environmental regulators of *Bordetella pertussis.* Infection and Immunity, **61**:807-815.

Meng, S. Y., and Bennett, G. N., 1992. Nucleotide sequence of the *Escherichia coli cad* operon: a system for neutralisation of low extracellular pH. Journal of Bacteriology, **174**:2659-2669.

Middlebrook, G., and Cohn, M. L., 1953. Some observations on the pathogenicity of isoniazid resistant variants of tubercle bacilli. Science, **118**:297-299.

Middlebrook, G., 1954. Isoniazid-resistance and catalase activity of tubercle bacilli. American Review of Tuberculosis, **69**:471-472.

Miller, L. P., Crawford, J. T., and Shinnick, T. M., 1994. The *rpoB* gene of *Mycobacterium tuberculosis.* Antimicrobial Agents and Chemotherapy, **38**:805-811.

Mims, C. A., 1987. The Pathogenesis of Infectious Disease: Academic Press, London, 342 pp.

Minnikin, D. E., 1982. Lipids: Complex lipids, their chemistry, biosynthesis and roles. In: Ratledge, C. and Stanford, J. (Editors), The Biology of the Mycobacteria. Academic Press, London, 1, pp. 95-184.

Mitchison, D. A., 1992. Understanding the Chemotherapy of tuberculosis - current problems. Journal of Antimicrobial Chemotherapy, **29**:477-493.

Moat, A. G., and Foster, J. W., 1988. Carbohydrate metabolism and energy metabolism. In: Moat, A. G. and Foster, J. W. (Editors), Microbial Physiology. Wiley-Interscience, New York, pp. 118-220.

Mulvey, M. R., Sorby, P. A., Triggs-Raine, B. L., and Loewen, P. C., 1988. Cloning and physical characterisation of *katE* and *katF* required for catalase HPII expression in *Escherichia coli.* Gene, **73**:337-345.

Murray, A., Winter, N., Lagranderie, M., Hill, D. F., Rauzier, J., Timm, J., Leclerc, C., Moriarty, K. M., Gheorghiu, M., and Gicquel, B., 1992. Expression of *Escherichia coli* b-galactosidase in *Mycobacterium bovis* BCG using and expression system isolated from *Mycobacterium paratuberculosis* which induced humoral and cellular immune responses. Molecular Microbiology, **6**:3331-3342.

Nadler, V., Goldberg, I., and Hochman, A., 1986. Comparative study of bacterial catalases. Biochemical and Biophysical Acta, **882**:234-241.

Nagasawa, S., Ishige, K., and Mizuno, T., 1993. Novel members of the two-component signal transduction genes in *Escherichia coli.* Journal of Biochemistry, **114**:350-357.

Nagaya, M., Aiba, H., and Mizuno, T., 1993. Cloning of a sensory-kinase-encoding gene that belongs to the two-component regulatory family from the cyanobacterium *Synechococcus* sp. PCC7942. Gene, **131**:119-124.

Nakamura, R. M., 1970. Transfection of *Mycobacterium smegmatis* in an acidic medium. In: Juhasz, S. E. and Plummer, G. (Editors), Host-virus relationships in *Mycobacterium*, *Norcardia* and *Actinomyces*. Charles C. Thomas, Springfield, Illinois., pp. 166-178.

Nakamura, T., Tokuda, H., and Unemoto, T. 1984. K^+/H^+ antiporter functions as a regulator of cytoplasmic pH in the marine bacterium, *Vibrio alginolyticus*. Biochimica et Biophysica Acta, **776**:330-336.

North, R. J., and Izzo, A. A., 1993. Mycobacterial virulence. Virulent strains of *Mycobacterium tuberculosis* have faster *in vivo* doubling times and are better equipped to resist growth-inhibiting functions of macrophages in the presence and absence of specific immunity. Journal of Experimental Medicine, **177**:1723-1733.

O'Brien, S., Jackett, P. S., Lowrie, D. B., and Andrew, P. W., 1990. Guinea pig alveolar macrophages kill *Mycobacterium tuberculosis in vitro*, but killing is independent of susceptibility to hydrogen peroxide or triggering of the respiratory burst. Microbial Pathogenesis, **10**:199-207.

O'Brien, S., and Andrew, P. W., 1991. Guinea-pig alveolar macrophage killing of *Mycobacterium tuberculosis, in vitro*, does not require hydrogen peroxide or hydroxyl radical. Microbial Pathogenesis, **11**:229-236.

O'Farrell, P. H., 1975. High resolution two-dimensional electrophoresis of proteins. Journal of Biological Chemistry, **250**:4007-4021.

O'Hara, G. W., and Glenn, A. R., 1994. The adaptive acid tolerance response in root nodule bacteria and *Escherichia coli*. Archives of Microbiology, **161**:286-292.

Olsen, E. R., 1993. Influence of pH on bacterial gene expression. Molecular Microbiology, **8**:5-14.

Padan, E., and Schulinder, S., 1986. Intracellular pH regulation in bacterial cells. Methods in Enzymology, **125**:337-352.

Padan, E., Maisler, N., Taglicht, D., Karpel, R., and Schulinder, S., 1989. Deletion of *ant* in *Escherichia coli* reveals its function in adaptation to high salinity and an alternative Na+/H+ antiporter systems(s). Journal of Biological Chemistry, **264**:20297-20302.

Padan, E., and Schulinder, S., 1992. Na+ transport systems in prokaryotes. In: Bakker, E. (Editor), Alkali Cation Transport Systems in Prokaryotes. CRC, Boca Raton, Florida.

Papp-Szabo, E., Sutherland, C. L., and Josephy, P. D., 1993. Superoxide dismutase and the resistance of *Escherichia coli* to phagocytic killing by human neutrophils. Infection and Immunity, **61**:1442-1446.

Papp-Szabo, E., Firtel, M., and Josephy, P. D., 1994. Comparison of the sensitivities of *Salmonella typhimurium oxyR* and *katG* mutants to killing by human neutrophils. Infection and Immunity, **62**:2662-2668.

Parish, T., 1993. The nature and activity of regulators of mycobacterial genes. [PhD thesis]. The Open University, UK.

Park, S. F., Stirling, D. A., Hulton, C. S. J., Booth, I. R., Higgins, C. F., and Stewart, G. S. A. B., 1989. A novel, non-invasive promoter probe vector: cloning of the osmoregulated *proU* promoter of *Escherichia coli* K12. Molecular Microbiology, **3**:1011-1023.

Park, S. F., Stewart, G. S. A. B., and Kroll, R. G., 1992. The use of bacterial luciferase for monitoring the environmental regulation of expression of genes encoding virulence factors in *Listeria monocytogenes*. Journal of General Microbiology, **138**:2619-2627.

Parkinson, J. S., 1993. Signal transduction schemes of bacteria. Cell, **73**:857-871.

Pascopella, L., Collins, F. M., Martin, J. M., Lee, M. H., Hatfull, G. F., Stover, C. K., Bloom, B. R., and Jacobs Jr, W. R., 1994. Use of *in vivo* complementation in *Mycobacterium tuberculosis* to identify a genomic fragment associated with virulence. Infection and Immunity, **62**:1313-1319.

Pennington, C. D., and Gregory, E. M., 1986. Isolation and reconstitution of iron- and manganese-containing superoxide dismutases from *Bacteroides thetaiotaomicron*. Journal of Bacteriology, **166**:528-532.

Pethal, M. L., and Falkinham III, J. O., 1989. Plasmid-influenced changes in *Mycobacterium avium* catalase activity. Infection and Immunity, **57**:1714-1718.

Pinner, E., Padan, E., and Schulinder, S., 1992. Cloning, sequencing, and expression of the **nhaB** gene, encoding a Na+/H+ antiporter in *Escherichia coli*. Journal of Biological Chemistry, **267**:11064-11068.

Portaels, F., and Pattyn, S. R., 1982. Growth of mycobacteria in relation to the pH of the medium. Annals de Institute Pasteur, **133B**:213-221.

Pugsley, A.P., 1993. The complete general secretory pathway in Gram-negative bacteria. Microbiological Reviews, **57**:50-108.

Radford, A. J., and Hodgson, A. L. M., 1991. Construction and characterisation of a *Mycobacterium-Escherichia coli* shuttle vector. Plasmid, **25**:149-153.

Raja, N., Goodson, M., Chul, W. C. M., Smith, D. G., and Rowbury, R. J., 1991. Habituation to acid in *Escherichia coli*: conditions for habituation and its effects on plasmid transfer. Journal of Applied Bacteriology, **70**:59-65.

Rastogi, N., 1991. Recent observations concerning structure and function relationships in the mycobacterial cell envelope. Research in Microbiology, **142**:464-476.

Ratledge, C., 1982. Nutrition, Growth and Metabolism. In: Ratledge, C. and Stanford, J. (Editors), The Biology of the Mycobacteria. Academic Press, London, 1, pp. 186-271.

Ratledge, C., 1984. Metabolism of iron and other metals by mycobacteria. In: Kubica, G. P. and Wayne, L. G. (Editors), The Mycobacteria: a sourcebook. Marcel Dekker Inc., New York, Part A, pp. 603-627.

Rauzier, J., Moniz-Perrira, J., and Giquel-Sanzey, B., 1988. Complete nucleotide sequence of pAL5000, a plasmid from *Mycobacterium fortuitum*. Gene, **71**:315-321.

Roach, T. I. A., Barton, C. H., Chatterjee, D., and Blackwell, J. M., 1993. Macrophage activation: Lipoarabinomannan from avirulent and virulent strains of *Mycobacterium tuberculosis* differentially induces the early genes c-fos, KC, JE, and tumor necrosis factor-α. The Journal of Immunology, **150**:1886-1896.

Roberts, D. L., Bennett, D. W., and Forst, S. A., 1994. Identification of the site of phosphorylation on the osmosensor, EnvZ, of *Escherichia coli*. Journal of Biological Chemistry, **269**:8728-8733.

Roland, K. L., Martin, L. E., Esther, C. R., and Spitznagel, J. K., 1993. Spontaneous *pmrA* mutants of *Salmonella typhimurium* LT2 define a new two-component regulatory system with a possible role in virulence. Journal of Bacteriology, **175**:4154-4164.

Rook, G. A. W., and Bloom, B. R., 1994. Mechanisms of pathogenesis in tuberculosis. In: Bloom, B. R. (Editor), Tuberculosis: Pathogenesis, Protection and Control. American Society for Microbiology Press, Washington, D.C., pp. 485-501.

Roszak, D. B., and Colwell, R. R., 1987. Survival strategies of bacteria in the natural environment. Microbiological Reviews, **51**:365-379.

Rowbury, R. J., Goodson, M., and Wallace, A. D., 1992. The PhoE porin and transmission of the chemical stimulus for induction of acid resistance (acid habituation) in *Escherichia coli*. Journal of Applied Bacteriology, **72**:233-243.

Rowbury, R. J., and Goodson, M., 1993. PhoE porin of *Escherichia coli* and phosphate reversal of acid damage and killing and of acid induction of the *cadA* gene product. Journal of Applied Bacteriology, **74**:652-661.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A., 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science, **239**:487-491.

Sak, B. D., Eisenstark, A., and Touati, D., 1989. Exonuclease III and the hydoperoxidase II in *Escherichia coli* are both regulated by the *katF* product. Proceedings of the National Academy of Sciences (USA), **86**:3271-3275.

Sambrook, J., Fritsch, E. F., and Maniatis, T., 1989. Molecular Cloning: a Laboratory Manual, 2nd ed: Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Sarkar, G., Kapelner, S., and Sommer, S. S., 1990. Formamide can dramatically improve the specificity of PCR. Nucleic Acids Research, **18**:7465.

Schlesinger, L. S., Bellinger-Kawahara, C. G., Payne, N. R., and Horwitz, M. A., 1990. Phagocytosis of *Mycobacterium tuberculosis* is mediated by human monocyte complement receptors and complements component C3. Journal of Immunology, **144**:2771-2780.

Schubert, R. H. W., 1991. Aeromonads and their significance as potential pathogens in water. Journal of Applied Bacteriology, Symposium Supplement, **70**:131S-135S.

Sela, S., Clark-Curtiss, J. E., and Bercovier, H., 1989. Characterisation and taxonomic implications of the rRNA genes of *Mycobacterium leprae*. Journal of bacteriology, **171**:70-73.

Sela, S., and Clark-Curtiss, J. E., 1991. Cloning and characterisation of the *Mycobacterium leprae* putative ribosomal RNA promoter in *Escherichia coli*. Gene, **98**:123-127.

Sharma, J. K., Gopalkrishna, V., and Das, B. C., 1992. A simple method for elimination of unspecific amplifications in polymerase chain reaction. Nucleic Acids Research, **20**:6117-6118.

Shibata, C., Ehara, T., Tomura, K., Igarashi, K., and Kobayashi, H., 1992. Gene structure of *Enterococcus hirae* (*Streptococcus faecalis*) F1F0 ATPase, which functions as a regulator of cytoplasmic pH. Journal of Bacteriology, **174**:6117-6124.

Shinnick, T. M., 1987. The 65-kilodalton antigen of *Mycobacterium tuberculosis*. Journal of Bacteriology, **169**:1080-1088.

Shoeb, H. A., Bowman, B. U., Ottolenghi, A. C., and Merola, A. J., 1985(a). Peroxidase-mediated oxidation of isoniazid. Antimicrobial Agents and Chemotherapy, **27**:399-403.

Shoeb, H. A., Bowman, B. U., Ottolenghi, A. C., and Merola, A. J., 1985(b). Evidence for the generation of active oxygen by isoniazid treatment of extracts of *Mycobacterium tuberculosis* H37Ra. Antimicrobial Agents and Chemotherapy, **27**:404-407.

Siegele, D. A., and Kolter, R., 1992. Life after log. Journal of Bacteriology, **174**:345-348.

Small, P. M., and van Embden, J. D. A., 1994. Molecular epidemiology of tuberculosis. In: Bloom, B. R. (Editor), Tuberculosis: Pathogenesis, Protection and Control. American Society for Microbiology Press, Washington, D.C., pp. 569-582.

Snapper, S. B., Lugosi, L., Jekkel, A., Melton, R. E., Kieser, T., Bloom, B. R., and Jacobs Jr, W. R., 1988. Lysogeny and transformation in mycobacteria: stable expression of foreign genes. Proceedings of the National Academy of Sciences (USA), **85**:6987-6991.

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

Snider, D. E., 1994. Tuberculosis: the world situation. In: Porter, J. D. H. and. McAdam, K. P. W. J,. (Editors), Tuberculosis: Back to the Future. John Wiley and Sons, Chichester, pp. 13-33.

Sorokin, A., Zumstein, E., Azevedo, V., Ehrlich, S. D., and Serror, P., 1993. The organisation of the *Bacillus subtilus* 168 chromosome region between the *spoVA* and *serA* genetic loci, based on sequence data. Molecular Microbiology, **10**:385-395.

Southern, E. M., 1975. Detection of specific sequences among DNA fragments separated by gel eletrophoresis. Journal of Molecular Biology, **98**:503-517.

Spector, M. P., Aliabadi, Z., Gonzalez, T., and Foster, J. W., 1986. Global control in *Salmonella typhimurium*: two-dimensional electrophoretic analysis of starvation-, anaerobiosis-, and heat shock-inducible proteins. Journal of Bacteriology, **168**:420-424.

Stahl, D. A., and Urbance, J. W., 1990. The division between fast- and slow-growing species corresponds to natural relationships among the mycobacteria. Journal of Bacteriology, **172**:116-124.

Stewart, G. S. A. B., and Williams, P., 1992. *lux* genes and the applications of bacterial bioluminescence. Journal of General Microbiology, **138**:1289-1300.

Stover, C. K., de la Cruz, V. F., Fuerst, T. R., J.E., B., Benson, L. A., Bennett, L. T., Bansal, G. P., Young, J. F., Lee, M. H., Hatfull, G. F., Snapper, S. B., Barletta, R. G., Jacobs Jr, W. R., and Bloom, B. R., 1991. New use of BCG for recombinant vaccines. Nature, **351**:456-460.

Stover, C. K., Bansal, G. P., Hanson, M. S., Burlein, J. E., Palaszynski, S. R., Young, J. F., Koenig, S., Young, D. B., Sadziene, A., and Barbour, A., 1993. Protective immunity elicited by a recombinant bacille Calmette-Guerin (BCG) expressing outer surface protein A (OspA) lipoprotein: a candidate Lyme disease vaccine. Journal of Experimental Medicine, **178**:197-209.

Sturgill-Koszycki, S., Schlesinger, P. H., Chakraborty, P., Haddix, P. L., Collins, H. L., Fok, A. K., Allen, R. D., Gluck, S. L., Heuser, J., and Russell, D. G., 1994. Lack of acidification in *Mycobacterium* phagosomes produced by exclusion of the vesicular proton-ATPase. Science, **263**:678-681.

Sturr, M. G., and Marquis, R. E., 1992. Comparative acid tolerance and inhibitor sensitivities of isolated F-ATPases of oral lactic acid bacteria. Applied and Environmental Microbiology, 58:2287-2291.

Taglicht, D., Padan, E., and Schulinder, S., 1991. Overproduction and purification of a functional Na+/H+ antiporter coded by *nhaA* (*antA*) from *Escherichia coli*. Journal of Biological Chemistry, **266**:11289-11294.

Takayama, M., Ohyama, T., Igarashi, K., and Kobayashi, H., 1994. *Escherichia coli cad* operon functions as a supplier of carbon dioxide. Molecular Microbiology, **11**:913-918.

Thole, J. E. R., Dauwerse, H. G., Das, P. K., Groothius, D. G., Schouls, L. M., and van Embden, J. D. A., 1985. Cloning of *Mycobacterium bovis* BCG DNA and expression of antigens in *Escherichia coli*. Infection and Immunity, **50**:800-806.

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.

Triggs-Raine, B. L., Doble, B. W., Mulvey, M. R., Sorby, P. A., and Loewen, P. C., 1988. Nucleotide sequence of *katG*, encoding catalase HPI of *Escherichia coli*. Journal of Bacteriology, **170**:4415-4419.

van Bogelan, R. A., Kelley, P. M., and Neidhardt, F. C., 1987. Differential induction of heat shock, SOS, and oxidative stress regulons and accumulation of nucleotides in *Escherichia coli*. Journal of Bacteriology, **169**:26-32.

van der Drift, C., Jannessen, D. B., and van Wezenbeek, P. M. G. F., 1978. Hydrolysis and synthesis of ATP by membrane-bound ATPase from a motile *Streptococcus*. Archives of Microbiology, **119**:31-36.

van Ossowski, I., Hausner, G., and Loewen, P. C., 1993. Molecular evolutionary analysis based on the amino acid sequence of catalase. Journal of Molecular Evolution, **31**:71-76.

van Soolingen, D., Hermans, P. W. M., de Haas, P. E. W., Soll, D. R., and van Embden, J. D. A., 1993. Occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains: Evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. Journal of Clinical Microbiology, **29**:2578-2586.

von Ossowski, I., Mulvey, M. R., Leco, P. A., Borys, A., and Loewen, P. C., 1991. Nucleotide sequence of *Escherichia coli katE*, which encodes catalase HPII. Journal of Bacteriology, **173**:514-520.

Walderhaug, M. O., Polarek, J. W., Voelkner, P., Daniel, J. M., Hesse, J. E., Altendorf, K., and Epstein, W., 1992. KdpD and KdpE, proteins that control expression of the *kdpABC* operon, are members of the two-component sensor-effector class of regulators. Journal of Bacteriology, **174**:2152-2159.

Walker, L., and Lowrie, D. B., 1981. Killing of *Mycobacterium microti* by immunologically activated macrophages. Nature, **293**:69-70.

Watson, N., Dunyak, D. S., Rosey, E. L., Slonczewski, J. L., and Olsen, E. R., 1992. Identification of elements involved in transcriptional regulation of the *Escherichia coli cad* operon by external pH. Journal of Bacteriology, **174**:530-540.

Wayne, L. G., and Gross, W. M., 1968. Base composition of deoxyribonucleic acid isolated from mycobacteria. Journal of Bacteriology, **96**:1915-1919.

Wayne, L. G., Diaz, G. A., and Doubek, J. R., 1968. Acquired isoniazid resistance and catalase activity of mycobacteria. American Review of Respiratory Disease, **97**:909-913.

Wayne, L. G., and Diaz, G. A., 1982. Serological, taxonomic, and kinetic studies of the T and M classes of mycobacterial catalase. International Journal of Systematic Bacteriology, **32**:296-304.

Wayne, L. G., 1984. Mycobacterial enzymes involved in generation and destruction of hydrogen peroxide. In: Kubica, G. P. and Wayne, L. G. (Editors), The Mycobacteria: a sourcebook. Marcel Dekker Inc., New York, Part A, pp. 595-601.

Wayne, L. G., and Kubica, G. P., 1986. Genus *Mycobacterium*. In: Sneath, P. H. A., Mair, N. S., Sharpe, E. M. and Holt, J. G. (Editors), Bergey's Manual of Systematic Bacteriology. Williams and Wilkins, London, 2, pp. 1436-1457.

Wayne, L. G., and Diaz, G. A., 1988. Detection of a novel catalase in extracts of *Mycobacterium avium* and *Mycobacterium intracellulare*. Infection and Immunity, 56:936-941.

Wheeler, P. R., and Ratledge, C., 1994. Metabolism of *Mycobacterium tuberculosis*. In: Bloom, B. R. (Editor), Tuberculosis: Pathogenesis, Protection and Control. American Society for Microbiology Press, Washington, D.C., pp. 353-385.

Wiker, H. G., and Harboe, M., 1992. The antigen 85 complex: a major secretion product of *Mycobacterium tuberculosis*. Microbiological Reviews, **56**:648-661.

Wimpenny, J. W. T., 1967. The uptake and fate of isoniazid in *Mycobacterium tuberculosis* var. bovis BCG. Journal of General Microbiology, **47**:389-403.

Winder, F. G., 1960. Catalase and peroxidase in mycobacteria: possible relationship to the mode of action of isoniazid. American Review of Respiratory Disease, **81**:68-78.

Winder, F. G., 1982. Mode of action of the antimycobacterial agents. In: Ratledge, C. and Stanford, J. (Editors), The Biology of the Mycobacteria. Academic Press, London, 1, pp. 353-438.

Winter, N., Lagranderie, M., Rauzier, J., Timm, J., Leclerc, C., Guy, B., Kieny, M. P., Gheorghiu, M., and Gicquel, B., 1991. Expression of heterologous genes in *Mycobacterium bovis* BCG: induction of a cellular response against HIV-1 Nef protein. Gene, **109**:47-54.

Woese, C. R., 1987. Bacterial Evolution. Microbiological Reviews, **51**:221-271.

Wolk, C. P., Cai, Y., and Panoff, J. M., 1991. Use of a transposon with luciferase as a reporter to identify environmentally responsive genes in a cyanobacterium. Proceedings of the National Academy of Sciences (USA), **88**:5355-5359.

Woods, S. A., and Cole, S. T., 1990. A family of dispersed repeats in *Mycobacterium leprae*. Molecular Microbiology, **4**:1745-1751.

Wu, J., and Weiss, B., 1991. Two divergently transcribed genes, *soxR* and *soxS*, control a superoxide response regulon in *Escherichia coli*. Journal of Bacteriology, **173**:2864-2871.

Yanisch-Peron, C., Vieira, J., and Messing, J., 1985. Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. Gene, **33**:103-119.

Youatt, J., 1962. Pigments produced by mycobacteria on exposure to isoniazid. II. Australian Journal of Experimental Biology and Medical Science, **40**:197-200.

Youatt, J., 1969. A review of the action of isoniazid. American Review of Respiratory Disease, **99**:729-749.

Youatt, J., and Tham, H. S., 1969. An enzyme system of *Mycobacterium tuberculosis* that reacts specifically with isoniazid. II. Correlation of this reaction with binding and metabolism of isoniazid. American Review of Respiratory Disease, **100**:31-37.

Youmans, G. P., 1979. Tuberculosis: W.B. Saunders Company, Philadelphia, 511 pp.

Young, R. A., Bloom, B. R., Grosskinsky, C. M., Ivanyi, J., Thomas, D., and Davis, R. W., 1985a. Dissection of *Mycobacterium tuberculosis* antigens using recombinant DNA. Proceedings of the National Academy of Sciences (USA), **82**:2583-2587.

Young, R. A., Mehra, V., Sweetser, D., Buchanan, T., Clark-Curtiss, J. E., Davis, R. W., and Bloom, B. R., 1985b. Genes for the major protein antigens of the leprosy parasite *Mycobacterium leprae*. Nature, **316**:450-452.

Young, D. B., and Garbe, T. R., 1991a. Lipoprotein antigens of *Mycobacterium tuberculosis*. Research in Microbiology, **142**:55-65.

Young, D. B., and Garbe, T. R., 1991b. Heat shock proteins and antigens of *Mycobacterium tuberculosis*. Infection and Immunity, **59**:3086-3093.

Zainuddin, Z. F., and Dale, J. W., 1989. Polymorphic repetitive DNA sequences in *Mycobacterium tuberculosis* detected with a gene probe from a *Mycobacterium fortuitum* plasmid. Journal of General Microbiology, **135**:2347-2355.

Zainuddin, Z. F., Kunze, Z. M., and Dale, J. W., 1989. Transformation of *Mycobacterium smegmatis* with *Escherichia coli* plasmids carrying a selectable resistance marker. Molecular Microbiology, **3**:29-34.

Zhang, Y., Heym, B., Allen, B., Young, D., and Cole, S. T., 1992. The catalase-peroxidase and isoniazid resistance of *Mycobacterium tuberculosis*. Nature, **358**:591-593.

Zychlinsky, E., and Matin, A., 1983. Cytoplasmic pH homeostasis in an acidophilic bacterium, *Thiobacillus acidophilus*. Journal of Bacteriology, 156:1352-1355.