THE *MYCOBACTERIUM TUBERCULOSIS KATG* GENE: IDENTIFICATION OF A NOVEL FUNCTION AND ANALYSIS OF THE REGULATION OF EXPRESSION

Michelle Anne Mulder

Thesis Presented for the Degree of

DOCTOR OF PHILOSOPHY

In the Department of Medical Microbiology

Faculty of Medicine

UNIVERSITY OF CAPE TOWN

June 1998

CERTIFICATION OF SUPERVISORS

In terms of paragraph 9 of "General regulations for the degree of Doctor of Philosophy," University of Cape Town, we, as supervisors of the candidate Michelle Anne Mulder, certify that we approve of the incorporation in this thesis of material that has been submitted for publication.

Signature Removed

Supervisor
Associate Professor L. M. Steyn
Dept. Medical Microbiology
University of Cape Town

17/2/98

Signature Removed

Co-supervisor
Dr H. Zappe
Dept. Medical Microbiology
University of Cape Town

17/2/98

BRIEF CONTENTS

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor Lafras Steyn and co-supervisor Harold Zappe for their support and help in the writing of my thesis, and Val Abratt for all her help and advice on DNA repair.

A special thanks to my colleagues and friends in the laboratory for their continual enthusiasm, support and advice. A big thank you to my family and to Gregg for their love, support and patience throughout my PhD, and especially to Susi and Nicky for all their friendship, help and encouragement.

I gratefully acknowledge the financial assistance received from Glaxo-Wellcome and the Foundation for Research and Development, without which this study would not have been possible.

ABBREVIATIONS

A	Adenosine
A	Alanine
ACP	Acyl carrier protein
ADC	Albumin dextrose catalase
AIDS	Acquired Immunodeficiency Syndrome
Ala	Alanine
AMP	Adenosine monophosphate
AP	Apurinic/apyrimidinic
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
ATP	Adenosine triphosphate
BCG	Bacille Calmette-Guérin
bp	Base pairs
BSA	Bovine serum albumin
°C	Degrees Centigrade
C	Cysteine
C	Cytidine
C-	Carboxy-
CAP	Catabolite activator protein
CAT	Chloramphenicol acetyl transferase
CCP	Cytochrome *c* peroxidase
Ci	Curie
Cl	Chloride
cm	Centimetre
Cu	Copper
Cys	Cysteine
D	Aspartic acid
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytidine triphosphate
DEPC	Diethyl pyrocarbonate
dGTP	Deoxyguanosine triphosphate
dITP	Deoxyinosine triphosphate
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
dRpase	DNA deoxyribophosphodiesterase
dsDNA	Double stranded deoxyribonucleic acid
DTT	Dithiothreitol
dTTP	Deoxythymidine triphosphate
E	Glutamic acid
EDTA	Ethylenediaminetetraacetic acid (disodium salt)
ETH	Ethionamide
F	Phenylalanine
$FADH_2$	1,5-dihydroflavin adenine dinucleotide
Fe	Iron

FMN	Flavin mononucleotide
$FMNH_2$	Reduced flavin mononucleotide
g	Grams
G	Glycine
G	Guanosine
GFP	Green fluorescent protein
Gln	Glutamine
Glu	Glutamic acid
Gly	Glycine
GSH	Glutathione
H	Histidine
H_2O_2	Hydrogen peroxide
His	Histidine
HIV	Human Immunodeficiency Virus
HOO·	Hydroperoxyl radical
HPI	Hydroperoxidase I
HPII	Hydroperoxidase II
I	Isoleucine
IFA	Incomplete Freund's adjuvant
IgG	Immunoglobin G
Ile	Isoleucine
INH	Isoniazid
IPTG	Isopropyl-β-D-Thiogalactoside
J	Joules
K	Lysine
KAS	β-ketoacyl synthase
kb	Kilobases
kDa	Kilodalton(s)
K_m	Michaelis constant
kV	Kilovolts
l	Litre
L	Leucine
L/L	Luciferase/luciferin
LAM	Lipoarabinomannan
Leu	Leucine
Lys	Lysine
m	Metre
M	Methionine
M	Molar
mA	Milliampere
MAC	*Mycobacterium avium* complex
MBP	Maltose binding protein
MCP	*M. tuberculosis* catalase-peroxidase
MDR	Multidrug resistant
Met	Methionine
mg	Milligram
MIC	Minimum inhibitory concentration
ml	Millilitre
mM	Millimolar

mm	Millimetre
Mn	Manganese
MOPS	Morpholinepropanesulfonic acid
MPTR	Major polymorphic tandem repeat
mRNA	Messenger ribonucleic acid
MTHF	5,10-methenyl tetrahydrofolyl
Mw	Molecular weight
μF	MicroFaraday
μg	Microgram
μl	Microlitre
μM	Micromolar
N	Asparagine
N-	Amino-
NAD	Nicotinamide adenine dinucleotide
NADH	Reduced nicotinamide adenine dinucleotide
NADP	Nicotinamide adenine dinucleotide phosphate
NADPH	Reduced nicotinamide adenine dinucleotide phosphate
ng	Nanogram
nm	Nanometre
nmol	Nanomoles
O_2	Molecular oxygen
O_2^-	Superoxide radical
$OH^·$	Hydroxyl radical
OD_{600}	Optical density at 600 nm
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
pg	Picogram
Phen	Phenylalanine
POD	Peroxidase
Pro	Proline
R	Arginine
R.B.S.	Ribosome binding site
RFLP	Restriction fragment length polymorphism
RIF	Rifampin
RLU	Relative light units
RNA	Ribonucleic acid
rRNA	Ribosomal ribonucleic acid
RNase	Ribonuclease
rpm	Revolutions per minute
s	Second
S	Serine
S	Sulphur
SDS	Sodium dodecyl sulfate
Ser	Serine
SOD	Superoxide dismutase
SSCP	Single strand conformation polymorphism
SSBP	Single strand DNA binding protein

ssDNA	Single stranded deoxyribonucleic acid
T	Threonine
T	Thymidine
TAE	Tris-acetate EDTA buffer
TB	Tuberculosis
TCA	Trichloroacetic acid
TEMED	N,N,N',N'-tetramethyl-ethylenediamine
Thr	Threonine
TNF	Tumour necrosis factor
tRNA	Transfer ribonucleic acid
Tris	Tris(hydroxymethyl)aminomethane
Trp	Tryptophan
T.S.C.	Translation start codon
tss	Transcription start site
UAR	Upstream activator region
UP	Upstream
UV	Ultraviolet
V	Valine
v/v	Volume/volume
V_{max}	Maximal velocity of reaction
W	Watts
W	Tryptophan
WHO	World Health Organization
w/v	Weight/volume
X-GAL	5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside
Y	Tyrosine
YT	Yeast tryptone
Zn	Zinc

ABSTRACT

A clone containing the terminal third of the *Mycobacterium tuberculosis katG* gene was previously shown to confer resistance to ethyl methane sulfonate on DNA repair-deficient *Escherichia coli* cells. The first aim of this study, therefore, was to examine the role played by the *M. tuberculosis katG* gene in DNA repair. The strategy used was overexpression of different regions of the gene in DNA repair-deficient mutants of *E. coli*, and examination of the sensitivities of the transformants to DNA damaging agents. Overexpression of the gene resulted in an increase in the survival of *recA* mutants exposed to ultraviolet (UV) light irradiation (254 nm) and hydrogen peroxide, and *uvr* mutants exposed to mitomycin C. Both the 5' and 3' regions of the *M. tuberculosis* KatG protein conferred the above effects, and this was independent of the catalase or peroxidase activity of the enzyme. The results suggest that the *M. tuberculosis katG* gene may encode a novel function related to the repair of DNA damage, and this may have implications for the survival of *M. tuberculosis* in the presence of DNA damaging agents, for example, in the macrophage. UV sensitivity tests on *M. intracellulare* and *M. tuberculosis* strains mutant in *katG* revealed that the *katG* gene product does not play a demonstrable role in the survival of repair-competent mycobacterial cells after exposure to UV irradiation.

The second aim of this study was to examine the regulation of expression of the *M. tuberculosis katG* gene. An *E. coli*-mycobacterial shuttle vector, pJCluc, containing the luciferase reporter gene, was constructed and used to examine the *katG* promoter sequences. The region required for optimal expression in *M. smegmatis* was localized to a 559 bp fragment immediately upstream of the gene. Two transcription start sites were mapped and putative -10 and -35 promoter sequences identified. It was demonstrated that expression from the promoter peaks during late exponential phase, and declines during stationary phase, and that the promoter is induced by ascorbic acid, and is repressed by oxygen limitation and growth at elevated temperatures. An upstream element that increased expression from the *M. tuberculosis katG* and the *M. paratuberculosis* P_{AN} promoters was identified, and shown to bind to one or more *M. smegmatis* proteins. Similar results were obtained in *M. bovis* BCG. Understanding the regulation of gene expression in mycobacteria is essential for determining the processes that govern interaction with the host. This study provides information on both the mycobacterial transcription signals and gene regulatory mechanisms.

SUMMARY

The *Mycobacterium tuberculosis katG* gene, encoding a catalase-peroxidase, plays an important role in the susceptibility of the organism to the antituberculosis drug, isoniazid, and in the virulence of the organism. The ability of *M. tuberculosis* to repair damage to its DNA, caused by mutagenic agents, may also play a significant role in the survival of the organism in the macrophage. A clone, designated pD3, containing the terminal third of the *Mycobacterium tuberculosis katG* gene, in the vector pEcoR252, was previously shown to confer resistance to ethyl methane sulfonate on DNA repair-deficient *E. coli* cells. The first aim of this study, therefore, was to examine the role played by the *M. tuberculosis katG* gene in DNA repair.

The *M. tuberculosis* H37Rv *katG* gene was amplified by PCR. The whole *katG* gene, the proximal two thirds, and the terminal third were cloned in the expression vector, pMAL™-c2, in-frame with the maltose binding protein. The resulting fusion proteins were overexpressed in, and purified from *E. coli*, and used in catalase and peroxidase assays. Only the fusion protein containing the whole KatG enzyme exhibited enzyme activity. Constructs containing the terminal third of the *katG* gene in pMAL™-c2 in the two other reading frames, with respect to the *malE* gene, were made for use as controls in DNA repair experiments. The pMAL™-c2-*katG* constructs were introduced into 5 DNA repair-deficient strains of *E. coli*. The resulting transformants were tested for their ability to survive exposure to three DNA damaging agents, short wavelength ultraviolet (UV) light irradiation, mitomycin C, and hydrogen peroxide. Overexpression of the proteins resulted in an increase in the survival of *recA* mutants exposed to UV irradiation and hydrogen peroxide, and in the survival of excision repair mutants exposed to mitomycin C. UV sensitivity experiments with *E. coli recA* mutants harbouring derivatives of the pD3 insert in pMAL™-c2, revealed that increased resistance to UV requires expression of the *katG* gene, and is independent of sequences downstream of the gene. Both domains of the *M. tuberculosis* KatG protein conferred the above effects, and this was independent of the catalase or peroxidase activity of the enzyme. The results suggest that the *M. tuberculosis katG* gene may encode a novel function related to the repair of DNA damage, and this may have implications for the survival of *M. tuberculosis* in the presence of DNA damaging agents, for example, in the macrophage.

UV sensitivity tests on a selection of mycobacterial species revealed that the *M. tuberculosis* complex organisms are more sensitive than the 'atypical mycobacteria.' UV sensitivity tests on *M. intracellulare* and *M. tuberculosis* strains mutant in *katG* revealed no significant differences in the sensitivity of the cells in the absence of this gene. This suggests that the *katG* gene product does not play an additive role in the survival of mycobacterial cells after exposure to short wavelength UV irradiation. This function may, however, be important in the presence of mutations in essential DNA repair genes.

Understanding the regulation of gene expression in mycobacteria is essential for determining the processes involved in interaction with the host, and hence in virulence. The second aim of this study, therefore, was to examine the regulation of expression of the *M. tuberculosis katG* gene, in order to provide insight into the factors which mediate cell survival within the macrophage and susceptibility to antituberculosis agents, particularly isoniazid, and to provide more information on the mycobacterial transcription signals and regulatory mechanisms.

An *E. coli*-mycobacterial shuttle vector, pJCluc, containing the promoterless firefly luciferase gene as a reporter, was constructed for the examination of mycobacterial promoter sequences. The vector was used to characterize the *M. tuberculosis katG* promoter. A 1.9 kb region immediately upstream of the *katG* gene was amplified by PCR and cloned into pJCluc, and promoted expression of the luciferase gene in *E. coli* and *M. smegmatis*. A smaller promoter fragment (559 bp) promoted expression with equal efficiency, and was used in all further studies. Two transcription start sites were mapped by primer extension analysis, and used to identify putative -10 and -35 regions. These show strong similarity to previously identified mycobacterial promoters. *M. smegmatis* recombinants harbouring the promoter construct were used to determine under what conditions the *katG* promoter is expressed. Luciferase activities revealed that expression from the promoter peaks during late exponential phase, and declines during stationary phase. The promoter is induced by ascorbic acid, and is repressed by oxygen limitation and growth at elevated temperatures.

Other regions of the 559 bp *M. tuberculosis katG* promoter fragment that are essential for promoter activity were examined. Deletions in the fragment were made using Bal31 nuclease and restriction endonucleases. The luciferase activities of the resulting constructs revealed that

a region between 559 and 448 bp upstream of the translation initiation codon, which forms part of an upstream activator region (UAR), is essential for promoter activity in *E. coli*, and is required for optimal activity in *M. smegmatis*. Primer extension analysis on selected deletions revealed that an alternative promoter is used in the absence of this region, and in the absence of the previously identified putative *katG* promoters. Cloning of a fragment containing the *katG* UAR upstream of the *M. paratuberculosis* P_{AN} promoter, in pJCluc, resulted in a 14.7-fold increase in luciferase activity in *E. coli*, and an 11.6-fold increase in *M. smegmatis*, relative to the P_{AN} promoter alone. Mobility shift assays revealed that one or more *M. smegmatis* proteins bind to the *katG* UAR. This region, therefore, acts as a transcriptional enhancer, possibly by binding to one or more regulatory proteins. The luciferase activities of a selection of promoter constructs were tested in *M. bovis* BCG. These constructs exhibited similar activities in this host as they did in *M. smegmatis*, suggesting that the recognition and utilization of promoters is similar in the two mycobacterial species.

The *M. tuberculosis katG* gene has been extensively studied with respect to its catalase and peroxidase activities, its role in virulence, and its role in susceptibility to isoniazid, however, this study provides the first report on the involvement of the gene in DNA repair processes. In-depth analysis of the regulation of expression of the gene has also not been previously reported. This study provides a foundation for the characterization of DNA sequences involved in both the expression and regulation of the gene, and the proteins or cellular factors required for this.

CHAPTER 1

GENERAL INTRODUCTION

CONTENTS Pg.

CHAPTER 1

GENERAL INTRODUCTION

1.1. OVERVIEW OF TUBERCULOSIS

Tuberculosis (TB) is a debilitating disease that accounts for approximately 7% of deaths in developing countries and 26% of all avoidable adult deaths worldwide (Murray *et al.*, 1990). Approximately 8-10 million people are infected with TB each year and, of these, 2-3 million die. Approximately 95% of these TB cases occur in developing countries (WHO, 1989). In South Africa it is estimated that 6-10 million people are infected with TB (Glatthaar, 1982), resulting in 6000 registered deaths per year (Department of National Health and Population Development, 1990). Of all the infectious diseases known to man, tuberculosis is the most prevalent and widespread over an extensive time period and has recently been declared a global emergency by the World Health Organization (WHO, 1994).

1.1.1. The history of tuberculosis

Available evidence indicates that people have been infected with TB since before written records. It has been suggested that *Mycobacterium tuberculosis,* the etiological agent of TB in humans, evolved from a mutant of *Mycobacterium bovis,* which causes a similar disease in cattle, i.e. that TB was first spread to humans by domestic cattle (Morse, 1961; Hare, 1967). The reason for this is that *M. bovis* has a broad host range (Grange and Collins, 1987), while *M. tuberculosis* is non-pathogenic for cattle (Clarke *et al.*, 1987). The earliest evidence of TB in humans comes from skeletal remains from approximately 5000 BC, which contained deformities suggestive of TB (Sager *et al.*, 1972), and from studies on Egyptian mummies. These date TB infection back to about 3400 BC (Zimmerman, 1979). Recently, Salo and co-workers (1994) reported on the recovery of DNA unique to *M. tuberculosis* from a lung lesion of a spontaneously mummified 1000 year old adult female body in Southern Peru, which supports the idea of the pre-Columbian presence of TB in the New World.

The incidence of TB and mortality rates from it are only known through written records, which have been best kept by the British. There are accounts of the increase in TB in Britain before the start of the industrial revolution in about 1780 (Chalke, 1962). This may have been due to enlargement of urban centres caused by population growth. Industrialization resulted in a number of conditions which were conducive to the spread of TB, such as large-scale urbanization, the development of industrial slums, poorly ventilated houses, long working hours, overcrowded factories and poor diet (Galdston, 1942). In Europe in the nineteenth century, pulmonary TB accounted for more deaths than war or cholera or smallpox, measles, whooping cough, scarlet fever, and typhus combined, and yet, it was given less attention than these other infectious diseases (Kochi, 1991). The course of respiratory TB death rates in Britain is illustrated in figure 1.1. Tuberculosis mortality rates showed a decline from 1838, possibly due to sanitary and labour reform, which occurred around that time. In the twentieth century, these mortality rates continued to decline and levels were relatively low by the 1950s, when effective TB chemotherapy was introduced. Similar declines in TB mortality were noted in other developed countries such as the United States, and these have been attributed to improved living conditions (Young and Cole, 1993).

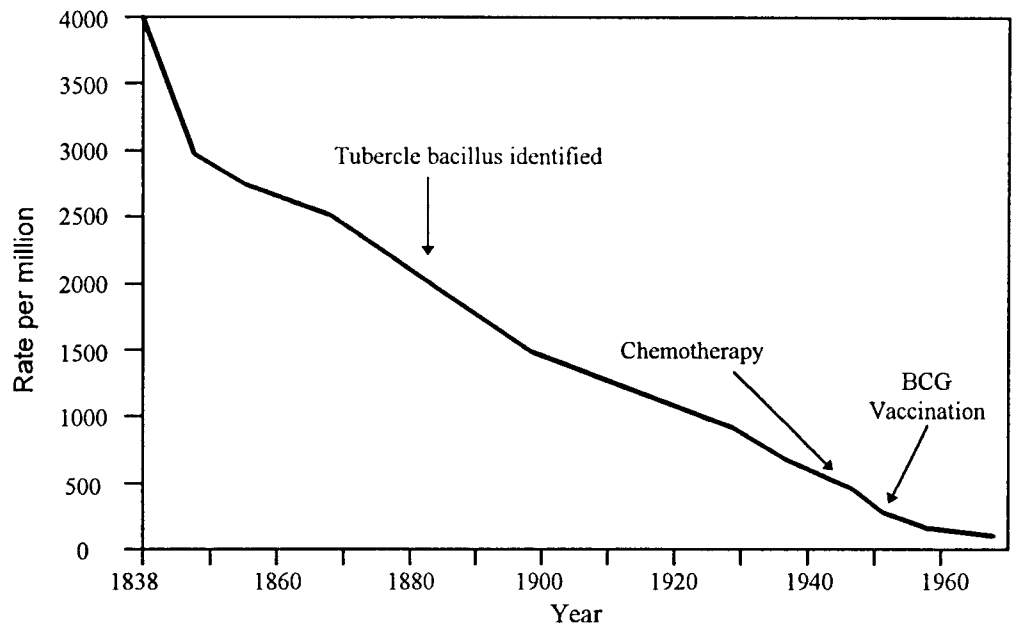


Figure 1.1 The decline in respiratory tuberculosis death rates in England and Wales since 1838. Standardized to the 1901 population. (Source: McKeown, 1979).

Little is known about the incidence of TB in South Africa before this century due to a lack of written accounts. Cummins (1920) and Lichenstein (1928) suggest that TB was virtually unknown in the interior of sub-Saharan Africa until the early part of the twentieth century. The spread of TB among the native people appears to be related to the degree of European contact, and the disease may have been introduced by European colonization (McVicar, 1932). The incidence of TB in the black population increased in the early part of the century due to urbanization, and, in particular, due to poor conditions on the mines. The introduction of effective chemotherapy in the 1950s resulted in a sharp decline in TB mortality rates, but there has been a continued increase in the incidence rates. TB therefore remains a problem in developing countries (Kochi, 1991).

Due to the steady decline in TB mortality in the developing world, and since the introduction of chemotherapy, little attention has been paid to the disease, and few attempts have been made to understand the pathogenesis of TB or the physiology and genetics of the causative organism. A new interest in TB has, however, been aroused by the recent re-emergence of the disease in the developed world (Bloom and Murray, 1992). In 1986 a resurgence of TB was noted in the United States, Britain and Europe, which coincided with the spread of HIV during this time. AIDS causes an increased susceptibility to and accelerated progression of TB. Another problem is the emergence of multidrug resistant (MDR) strains of *M. tuberculosis*, which, together with HIV infection, result in an early mortality in TB patients (Young and Wormser, 1994). These strains are difficult to eradicate, and therefore also contribute to the spread of the disease. Other proposed reasons for the re-emergence of TB in developed countries include immigration of individuals with TB from high-prevalence areas, and the increasing incidence of TB in institutions such as homeless shelters, schools and prisons (Ordway *et al.*, 1995). New solutions to the disease are evidently required in order to bring the epidemic under control, particularly in developing countries. Molecular biological techniques are now being used to obtain a better understanding of the pathogenesis of mycobacteria and interaction with their mammalian hosts. These techniques were initially applied to the study of *Mycobacterium leprae*, the causative agent of leprosy, and are now being extended to the study of *M. tuberculosis* (Young and Cole, 1993).

1.1.2. The control of tuberculosis

1.1.2.1. Tuberculosis chemotherapy

Although tuberculosis has infected people for thousands of years, it was only in this century that chemotherapy was introduced for the treatment of the disease. Streptomycin was first used against TB in 1944, only eleven months after its isolation (Schatz *et al.*, 1944; Pfuetze *et al.*, 1955), and isoniazid (INH) was first used in 1952 (Robitzek and Selikoff, 1952). In the same year, pyrazinamide, a synthetic derivative of nicotinamide, was introduced as a secondary drug (McDermott and Tompsett, 1954), and it was at this time that the potential of using more than one drug in combination against TB was realized. Rifampicin was introduced in 1966 (Verbist and Gyselen, 1968) and first used to treat TB in the early 1970s, and ethambutol was introduced in 1968 (Grumbach, 1969). Today, these five drugs are used as first-line antituberculosis agents. Ethionamide (ETH) is a structural analogue of INH which is sometimes used as a second-line antituberculosis drug (Hok, 1964; Riddell *et al.*, 1960). The successful treatment of TB relies on the eradication of all viable tubercle bacilli within the patient. There are several different populations of these organisms at the site of infection. These include: actively multiplying bacilli at neutral pH in material lining the cavity walls; bacilli within the macrophages at an acidic pH, and with a slower growth rate; bacilli within solid caseous areas with low oxygen tension, and therefore extremely slow growth rates; and a population of dormant bacilli (discussed in a later section). INH and streptomycin act on the large, actively multiplying population and may eliminate all extracellular bacilli, except for those that have developed resistance. INH penetrates most tissues and macrophages (Mackaness, 1956), and also acts on the slower growing bacilli to eliminate the "persisters." Rifampicin and pyrazinamide are able to kill the semi-dormant organisms which are not affected by INH (Grosset, 1980). Pyrazinamide, in particular, acts optimally at acid pH, and is therefore effective against intracellular organisms (Cole, 1994).

Immediately after the introduction of the various antituberculosis drugs, there were reports on strains which had developed resistance. Acquired resistance to drugs during or after treatment is usually due to non-adherence to the drug regimine. Primary resistance occurs in patients who have never received prior treatment and are infected with drug-resistant strains. Primary drug resistance is an indicator of the efficacy of TB control programs (Weyer and Kleeberg, 1992).

The recent resurgence of TB in developed countries has been accompanied by an increase in the incidence of strains which are resistant to one or more drugs. In April of 1991, for example, 19% of all prevalent strains in New York were resistant to two or more drugs (Frieden *et al.*, 1993). In the United States as a whole, 13% of all new TB cases are resistant to at least one of the first-line drugs, and 3.2% are resistant to both INH and rifampin (Bloch *et al.*, 1994). Understanding the genetic basis for resistance to drugs is expected to aid in the development of more rapid ways of detecting resistant strains, and in the rational design of new antimycobacterial agents (Musser, 1995). The modes of action of the antituberculosis drugs and mechanisms of resistance to them are discussed in a later section

For approximately 30 years after the introduction of ethambutol, no new anti-TB drugs were reported, however, the search has been reinitiated. Some possibilties for new drugs are listed by Reynolds (1989). These include the fluoroquinolones such as ciprofloxacin and ofloxacin; the rifamycin derivatives rifapentine and rifabutin; macrolides such as roxithromycin; and phenazines such as clofazimine. There is a need to develop a new TB treatment regimine which involves more powerful drugs which can be used in much shorter courses of therapy. The problem with new, powerful antimicrobial drugs, however, is that, in many cases, for example immunocompromised patients, these drugs are not effective. Thus a move is being made to rather stimulate the immune system to deal with the infection itself. A variety of cytokines and growth factors involved in the regulation of the immune response have been isolated. These factors, for example, interleukin-1 and tumour necrosis factor (TNF)-alpha, when administered to experimental animals, protect the animals against a number of infections (Kullberg *et al.*, 1994), and are under investigation as alternatives to drugs for the treatment of TB.

1.1.2.2. BCG vaccine

One of the most important control measures for prevention of the spread of tuberculosis is the BCG (Bacille Calmette-Guérin) vaccine. It has long been known that people who develop TB and recover are more resistant to reinfection with a new strain. For this reason, attempts were made to generate an attenuated organism which could be used as a vaccine against TB. In 1908, two students of Louis Pasteur, Calmette and Guérin, began to attenuate a strain of *M. bovis* by serial passages on potato-bile broth. After 13 years and 230 passages, they had a

strain which was avirulent in guinea pigs, rabbits, cattle and horses. The strain was given to the first human in 1921. Today more than 80% of children world-wide receive the BCG vaccine. It is inexpensive, safe, single dose, can be administered at birth, and confers long lasting immunity. The efficacy of the vaccine is, however, controversial. Roche *et al.* (1995) review the past disappointments and future hopes of the BCG vaccine. Protective efficacy is defined as the incidence of TB in the unvaccinated population minus that in the vaccinated population divided by the incidence in the unvaccinated population. Large scale trials on the efficacy of the BCG vaccination revealed that in countries with a high incidence of TB, BCG induced a lower degree of protection (Bloom and Fine, 1994). The vaccine is effective against childhood and disseminated TB but is less effective against adult pulmonary TB (Colditz *et al.*, 1994). The data from various trials on the efficacy of the vaccine are difficult to interpret and consolidate due to large variations in the methodologies used. Various reasons have, however, been put forward for the variability in the efficacy of BCG in different populations and in different geographical locations. One major problem is that BCG has been subcultured for years in various laboratories using a variety of culture conditions, and it is evident that BCG from different localities has evolved into strains with some distinct antigenic differences. BCG is given as a viable vaccine but there is large variation (5-45%) in the proportion of viable bacilli in different vaccine preparations (Gheorghiu, 1990). Induction of protective immunity depends on active secretion of proteins from viable bacilli (Andersen, 1994), therefore the potency of the vaccine depends on the proportion of viable bacilli. The optimal vaccine dose has not yet been established. The method of administration of the vaccine differs in different parts of the world, and recent work suggests that the route of administration has a profound effect on the type of immunity induced (Barnes *et al.*, 1992). Evidence also suggests that the immunity induced by BCG is not life long (Medical Research Council, 1972). In developing countries, for example, immunity may be waning at the age of peak exposure to TB. It is not known as yet whether boosting enhances efficacy.

There are also various host factors that affect the efficacy of the vaccine. Nutrition strongly affects immunity and immune responses to vaccines, however no correlation has been found between protective efficacy and nutritional status. Efficacy may be affected by concurrent infections with other pathogens such as HIV, and by certain host genetic factors. The most likely explanation for the variation in the protective efficacy of the BCG vaccine, however,

appears to be exposure to environmental mycobacteria (Bloom and Fine, 1994), which may induce a cross-reactive immunity that protects against TB almost as effectively as BCG does (Palmer and Long, 1966). This protection is not improved by the BCG vaccine and, therefore, lowers its apparent efficacy. Considerable work is underway to develop new and improved vaccines against tuberculosis. Approaches being used include new attenuated strains of *M. tuberculosis* in which virulence genes have been deleted; subunit vaccines composed of immunologically dominant mycobacterial antigens in an adjuvant; vaccines using environmental mycobacteria such as *M. vaccae* (Stanford *et al.*, 1994); recombinant BCG strains expressing heterologous antigens or various cytokines (O'Donnell *et al.*, 1994), and DNA vaccines (Lowrie *et al.*, 1994).

1.1.3. The pathogenesis of tuberculosis

1.1.3.1. The genus *Mycobacterium*

M. tuberculosis, the etiological agent of tuberculosis, is a member of the genus *Mycobacterium*, which is one of the oldest and best known bacterial genera. The two most well known mycobacteria are *M. tuberculosis* and *M. leprae*, because of their ability to cause disease in man. However, most mycobacteria form part of natural microbial communities in habitats such as the soil. Mycobacteria are aerobic actinomycetes which are morphologically intermediate between corynebacteria and proactinomycetes. They do not form mycelia but form slightly curved or straight rods which occassionally branch. They are non-motile, gram-positive and acid fast (i.e. cannot be decolourized by acid after staining with aniline dyes), due to the high levels of mycolic acid in their cell walls, which makes them wax-like and strongly hydrophobic (Schlegel, 1986). Mycobacterial cell walls are composed of a variety of soluble proteins, carbohydrates and lipids, and three insoluble macromolecules, arabinogalactan, peptidoglycan, and mycolic acid which, together, constitute the mycoylarabinogalactan peptidylglycan core of the cell wall. Mycolic acids are 2-branched, 3-hydroxycarboxylic acids (McNeil and Brennan, 1991). Although the cell walls of all the mycobacteria are thought to be very similar, different species and strains differ in their degree of acid-fastness, which can be related to differences in their cell wall structure or thickness. There is a relationship between the degree of acid-fastness and both the degree of parasitism and the speed at which the organisms grow. Slower growing

mycobacteria, such as *M. tuberculosis*, have a higher degree of acid-fastness than the faster growing saprophytic mycobacteria, and are more pathogenic. The order of acid-fastness is as follows: *M. leprae* > *M. paratuberculosis* > *M. tuberculosis* (virulent) > *M. tuberculosis* (avirulent or attenuated) > *M. avium* > saprophytic mycobacteria (non-pathogenic) (Murohashi and Yoshida, 1968). It is not clear what genetic or physiological differences are responsible for the differences in the growth rates of the mycobacteria.

The genus *Mycobacterium* consists of over seventy recognized species and includes intracellular bacterial parasites from the obligate intracellular pathogens *M. leprae*, *M. tuberculosis*, *M. bovis*, and *M. avium*, to environmental species such as *M. gordonae*, *M. fortuitum* etc., which are seldom pathogenic. Between these are the opportunistic pathogens such as *M. intracellulare*, *M. scrofulaceum*, *M. simiae*, and *M. szulgai*. Mycobacteria are classified into 4 or 5 broad taxonomic groups according to pathogenicity for humans and animals, rate of growth at optimum temperature, and effect of visible light on pigment production, corroborated by comparative 16S rRNA sequence studies. The *M. tuberculosis* complex consists of *M. tuberculosis*, *M. bovis*, *M. africanum*, and *M. microti*, all of which have conserved 16S rDNA sequences (Frothingham *et al.*, 1994). "Atypical" mycobacteria are those that do not cause tuberculosis. Runyon *et al.* (1974) classified atypical mycobacteria into four groups based on their laboratory culture characteristics. Group I (photochromogens) consists of mycobacteria that form deep yellow to orange or brick-red pigments when exposed to light but produce little or no pigments if grown in the dark. Group II (scotochromogens) form bright yellow or orange pigments when grown in light or dark. Group III (nonphotochromogens) are slow growers and produce little or no pigment, and Group IV grow very rapidly and are generally saprophytes, except for *M. fortuitum*, which can be pathogenic in man (Runyon *et al.*, 1974). This classification of atypical mycobacteria is now being replaced by molecular classification. *Mycobacterium avium* complex (MAC) organisms are slowly growing bacilli which may produce yellow pigment in the absence of light. The complex is composed of opportunistic pathogens which are ubiquitous in nature and can be isolated from natural sources of water, pools, soils, plants and bedding material. They are found in both fresh and salt water, especially with relatively high acidity, and in aerosols above fresh water. *M. avium* causes diseases in poultry and swine but direct transmission to humans is rare, except in HIV-infected individuals who are immunodeficient (Inderlied *et al.*, 1993).

The mycobacterial genome consists of a single length of DNA in a closed loop which is tightly packed but has no nuclear membrane. Their genomes are 2.8-4.5 x 10^9 bp in size, and most have a G+C content of 64-70%. *M. tuberculosis* has a genome of approximately 4.4 x 10^9 bp and a G+C content of 63-65% (Clarke-Curtiss *et al.*, 1985; Imaeda *et al.*, 1982). Attempts are underway to characterize mycobacterial genome structures and to sequence the entire genome of *M. leprae* and *M. tuberculosis*, with the aim of identifying immunogenic and biochemical features of the organism. A number of observations have been made with respect to the sequence of the *M. tuberculosis* genome obtained to date. The genome contains a large number of insertion sequences except in the area around the *oriC* and the *rpoB* gene, possibly because these are essential for replication. There are also a number of genes encoding protein kinases and phosphatases which are involved in signalling. Other interesting features include large amounts of phage sequence, and many genes encoding anaerobic functions (Cole, ASM Presentation, 1997).

1.1.3.2. Mechanisms of pathogenesis

In 1882 Robert Koch described the tubercle bacillus, *Mycobacterium tuberculosis*, and demonstrated that it was the cause of tuberculosis (Koch, 1932; Sakula, 1982; Grange and Bishop, 1982). TB is transmitted by the airborne spread of 3-10 bacilli suspended in droplet nuclei, produced when people with active pulmonary TB cough or speak (Wells, 1955; Riley *et al.*, 1959). The droplet nucleus is very stable and remains suspended in the air for long periods of time. Humans are highly susceptible to infection with *M. tuberculosis*, and inhalation of only 1 or 2 viable bacilli can cause an infection, although the risk of developing disease after one exposure is low (3-5%), and many who are exposed never develop clinical disease (Collins, 1989). In general, there are three possible outcomes to infection with *M. tuberculosis*: 1) primary TB which results from acute infection shortly after infection, 2) reactivation TB, in which active disease appears many years after infection, 3) establishment of a chronic latent infection. Only 5-15% of people infected by the tubercle bacilli develop active disease at some stage in their lives. This may be determined by a number of factors such as the immune response to the infection, the genetic make-up of the infected individual, and various characteristics of the invading organism. Today much is known about the immunology of TB

(Reviewed in Bloom, 1994), however, little is known about the pathogenesis of mycobacterial infection and, as yet, a molecular basis for virulence in *M. tuberculosis* has not been established.

M. tuberculosis is an intracellular pathogen which resides within the macrophage. The danger of the pathogen lies in its ability to persist in tissues and activate an immune response, rather than causing direct damage to host cells. Different *M. tuberculosis* strains differ in their ability to cause disease, and the virulent strains presumably contain a unique set of features, referred to as virulence determinants, which allow them to enter macrophages and to persist and multiply in this environment. Macrophages have developed a number of defences against intracellular pathogens which are reviewed by Beaman and Beaman (1984). These include the production of reactive oxygen intermediates, such as the superoxide anion (Babior *et al.*, 1973), singlet oxygen ($^1O_2$) (Allen *et al.*, 1974), and hydroxyl radicals (OH^-) (Tauber and Babior, 1978), in the host respiratory burst. This is accompanied by an increased oxygen consumption, and induction of a number of enzymes such as myeloperoxidase, acid hydrolases, and lysozyme (Bainton *et al.*, 1971), which degrade macromolecules. The latter enzymes are contained within lysosomes, which fuse with the plasma membrane of phagosomes and release the lysosomal contents in a process known as degranulation (Cohn, 1963). Human macrophages are relatively efficient at controlling most *M. tuberculosis* infections, but some organisms have adapted to overcome their defence mechanisms (Riley, 1995). Virulent *M. tuberculosis* strains, for example, can interfere with normal lysosome function. The mechanisms involved are not clear but it may be due to the production of a large amount of ammonia, which also results in the raising of the intralysosomal pH (Gordon *et al.*, 1980). Some components of the mycobacterial cell wall can also inhibit some of the lysosomal enzymes. The environment inside phagosomes is usually acidic, which potentiates the toxicity of other factors, however, acidification may be inhibited by the ammonia released by *M. tuberculosis*. More recent evidence suggests that *M. tuberculosis* arrests the maturation of its phagosome at a stage at which the phagosome interacts with early and late endosomes, but not with lysosomes (Clemens and Horwitz, 1996; Sturgill-Koszycki *et al.*, 1996), and that the lack of acidification of the vacuoles may be due to the absence of the vesicular proton-adenosine triphosphatase (Sturgill-Koszycki *et al.*, 1994).

It appears, therefore, that *M. tuberculosis* can prevent inhibition by the acidic conditions and lysosomal enzymes of the phagosome, however, within these cells the organism is still exposed

to the toxic oxygen and nitrogen intermediates. Walker and Lowrie (1981) identified hydrogen peroxide, which is generated by the host respiratory burst (Iyer *et al.*, 1961; Klebanoff, 1980), as the molecule that mediates the mycobactericidal effects of mononuclear phagocytes, and there is substantial evidence to suggest that hydrogen peroxide and associated reactive oxygen intermediates are responsible for much of the microbicidal powers of the macrophage (Andrew *et al.*, 1985; Babior, 1978; Jackett *et al.*, 1981a; 1981b; Lowrie, 1983). Chan *et al.* (1989) suggest that mycobacteria possess inherent mechanisms of resistance to these radicals by virtue of their cell walls. *M. tuberculosis* and other pathogenic and slow growing mycobacteria have an apparently unique major modification to their mycolic acids, i.e. cyclopropanation (Minnikin, 1982). Fast growing, saprophytic mycobacteria do not cyclopropanate their mycolic acids, but produce a large amount of unsaturated mycolic acids. Yuan *et al.* (1995) suggest that cyclopropanation may be an adaptation to the hostile intracellular environment in which reactive oxygen species would cross-link and degrade mycolic acids containing surface-exposed double bonds. The authors isolated the gene responsible for cyclopropanation from *M. tuberculosis* and demonstrated that overexpression of the gene in *M. smegmatis* resulted in a significant increase in resistance to killing by hydrogen peroxide (Yuan *et al.*, 1995). The importance of the cell wall in virulence is also illustrated by the results of Murohashi and Yoshida (1965), who showed that the duration of ultraviolet irradiation required for the loss of acid-fastness was longer for virulent strains than for avirulent or attenuated strains, again suggesting a link between the degree of virulence and the cell wall structure. Differences in the cell wall of virulent and avirulent strains could, therefore, be responsible for the ability of the organisms to survive at initial entrance into host cells during infection. Resistance to toxic radical species produced by macrophages is also mediated by the major cell wall-associated phosphatidylinositonyl-anchored complex lipopolysaccharide (LAM), which is produced in large amounts by *M. tuberculosis,* and has been shown to scavenge reactive oxygen intermediates and to downregulate the oxidative burst (Chan *et al.*, 1991). Although the mycobacterial cell wall provides a formidable barrier against reactive oxygen species, a reasonable proportion of these are probably still able to enter the cell where they can cause considerable damage to cellular components (reviewed in Chapter 2). This is where the radical scavenging enzymes such as the catalases and peroxidases come into play. These enzymes probably provide the second line of defence to these radicals, and therefore play a vital role in the survival of *M. tuberculosis* in

macrophages. One of these enzymes, the catalase-peroxidase, is the focus of this study and is discussed in detail in section 1.2.

1.1.3.3. Dormancy

Another factor which is important for the virulence of the organism and its ability to withstand drug treatment, is its ability to remain in a dormant state for years within the host. Dormant organisms are those that are not actively replicating and are metabolically inactive. As such they are resistant to most antimicrobial factors, and may be more accurately referred to as "non-replicating persisters." Nothing is known about how *M. tuberculosis* remains dormant or how it is activated out of latency. Under optimal conditions, *M. tuberculosis* takes 16-18 hours to undergo one cycle of replication (Wayne, 1977), and this slow growth may confer a selective advantage. It is possible that priority is given to cell survival rather than cell division to cause little perturbation of the host cells. It has been suggested that the natural physiological state of the cells resembles that of *Escherichia coli* in stationary phase, in which a set of proteins are induced that allow increased survival under poor external conditions. Dormant mycobacteria within the host may then represent an extreme form of stationary phase in which no replication occurs (Young and Cole, 1993). Very abrupt exposure to anaerobic conditions is lethal to actively growing *M. tuberculosis* cultures, however, gradual depletion of oxygen causes *M. tuberculosis* cultures to shift from active replication to dormancy. These dormant bacilli are partially or completely resistant to INH and rifampin but are sensitive to metranidazole, a drug specific for anaerobes and which is normally inactive against actively growing aerobic cultures (Wayne and Sramek, 1994). It is interesting to note that *M. tuberculosis* contains many enzymes required for anaerobic metabolism (Segal and Edwards, 1984), and these may play an important role in dormancy.

1.2. THE RADICAL SCAVENGING ENZYMES

Metabolism in an oxygen-rich environment may result in the generation of reactive oxygen species such as superoxide anions, hydroxyl radicals, and hydrogen peroxide (Fridovich, 1986). These radicals can oxidize membrane fatty acids, which initiates lipid peroxidation, oxidize

proteins, and damage DNA. Organisms in these environments therefore possess enzymes which neutralize these molecules e.g. superoxide dismutase, catalase and peroxidase, as well as DNA repair enzymes, e.g. exonuclease III, DNA polymerase I, the RecBC nuclease and RecA, which will be discussed in more detail in Chapter 2. The superoxide dismutases (SODs) are a family of metalloenzymes containing either iron (Fe), manganese (Mn), or both copper and zinc (CuZn) at the active site. Bacteria generally contain both Mn- and FeSODs in the cytoplasm (Hassan, 1989). The CuZnSOD is usually found in eukaryotes but is found in a small number of gram-negative species, in which it resides in the periplasm. SODs convert the superoxide anion into oxygen and hydrogen peroxide (McCord and Fridovich, 1988), which is then removed by the catalases and peroxidases. Catalases and peroxidases are both involved in the decomposition of hydrogen peroxide in a two-electron transfer between two substrates, one of which is reduced and one of which is oxidized. The difference between the two enzyme activities lies in the nature of the donor and acceptor. Catalases catalyze the dismutation of hydrogen peroxide to form oxygen and water, while peroxidases use a variety of organic and inorganic electron donors to reduce hydrogen peroxide (Deisseroth and Dounce, 1970; Dunford and Stillman, 1976). Hydroperoxidases exhibit both catalase and peroxidase activities.

1.2.1. Catalases

Catalases from various higher organisms share a number of similarities. Their molecular weight (Mw) ranges from 225 000 to 270 000 and they consist of four subunits of equal size, each of which contains one ferric heme prosthetic group (protoporphyrin IX). All show a broad pH range (5-10.5), and are specifically inhibited by 3-amino-1,2,4,-triazole (Nadler *et al.*, 1986; Schonbaum and Chance, 1976). The structure of the beef liver catalase has been examined in detail (Fita and Rossmann, 1985). Each subunit of the tetramer consists of four domains: 1) the 70 residue amino terminal, which forms an arm extending from the globular region of each subunit and interacts with the neighbouring subunit; 2) residues 76-320, which form a large anti-parallel β-barrel; 3) residues 321-436, which lacks secondary structure from residues 366-420, but includes the essential α-helix and is known as the "wrapping domain"; 4) residues 437-506, consisting of 4 α-helices on the external part of the molecule. Protohemin is essential for activity and specificity, and is thus a major structural component of the active site. Each subunit contains one protoporphyrin IX moiety containing an Fe^{3+} ion. In catalase, the heme group is

the organic radical, while in peroxidase the radical is situated on the amino acid side chain or on a sulfur atom. The heme group is accessible by a main funnel-shaped channel in the molecule. Both the amino- and carboxy-terminal sections form part of the lining of the entrance of this channel. Another minor channel leading to the heme group has a narrow entrance between Lys-236 and His-304, which is associated with the NADP binding site.

The first step in the catalytic reaction is the transfer of two electrons from the heme iron, Fe[III], to hydrogen peroxide to form Compound I (Fe[V]) (Schonbaum and Chance, 1976; Fita and Rossmann, 1985). This then reacts with a second hydrogen peroxide molecule to regenerate the native enzyme. Compound I can also react with H_2O_2 to form Compound II, which can then react with H_2O_2 in two different ways; 1) a slow reduction to native catalase, and 2) irreversible oxidation to Compound III, which is then slowly reduced to native catalase (see Fig. 1.2) (Schonbaum and Chance, 1976). Since these reactions are slow, the regeneration of native catalase for the catalytic cycle is low, resulting in an apparent decrease in enzyme turnover.

1.2.2. Peroxidases

Peroxidases are monomeric proteins of differing molecular weight. Their heme ions are reducible by dithionite (Dunford and Stillman, 1976). The cytochrome *c* peroxidase found in yeast mitochondria is the most widely studied. It consists of a single chain 294-residue heme enzyme which catalyzes the oxidation of ferrocytochrome *c* by peroxides in the following reaction:

$$2H^+ + ROOH + 2 \text{ cyt } c\ (Fe^{2+}) \rightarrow 2 \text{ cyt } c\ (Fe^{3+}) + ROH + H_2O$$

The cytochrome *c* peroxidase is a relatively small monomer, and serves as a simplified model for multisubunit cytochrome *c* oxidases. Finzel *et al.* (1984) used high resolution crystallographic refinement to obtain a structural model for the enzyme. Cytochrome *c* peroxidase is folded into two clearly defined domains, and the overall shape resembles a prolate ellipsoid. It consists of 10 α-helical segments, 2 antiparallel-β-pairs, and one small β-sheet (Poulos and Kraut, 1980). The amino- and carboxy-termini are situated in the same domain. A large channel at the interface between the two domains probably allows transport of the substrates to the active site (Finzel *et al.*, 1984).

The first step in the peroxidatic reaction also involves the oxidation of the enzyme by hydrogen peroxide to form Compound I, known as Compound ES for cytochrome *c* peroxidase (Yonetani, 1976). In this state, one of the oxidizing equivalents is stored as the oxy-ferryl (Fe[IV]=O) state of the heme iron (Lang *et al.*, 1976), while the other exists as a reversibly oxidized amino acid residue. Sivaraja *et al.* (1989) identified the primary site of the cytochrome *c* peroxidase Compound ES radical with tryptophan, most likely Trp^{191}. After the formation of Compound ES, two sequential one-electron transfers occur, in which ferrocytochrome *c* is oxidized (Yonetani, 1976), and the native protein is regenerated. Compound I of peroxidases also reacts with an electron donor, AH_2, to form Compound II, which then reacts with another electron donor to regenerate the native enzyme (see Fig. 1.2).

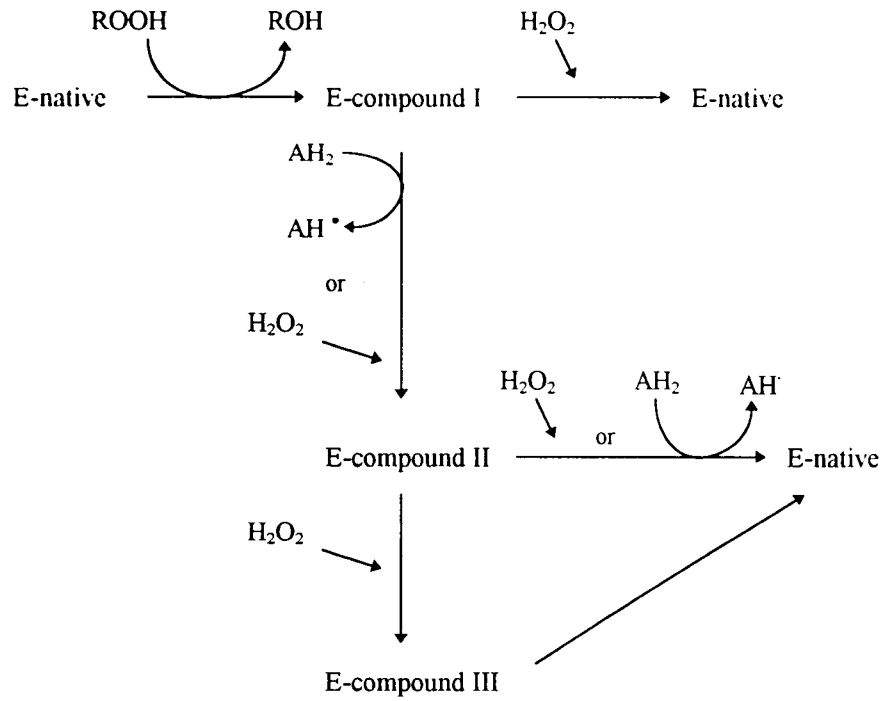


Figure 1.2 The oxidation and reduction reactions of catalases and peroxidases. E represents the enzyme; ROOH and ROH represent the oxidized and reduced forms of the peroxide, respectively; and AH_2 and AH· represent the reduced and oxidized forms of the electron donor, respectively.

1.2.3. Catalase-peroxidases

Hydroperoxidases (HPI) or catalase-peroxidases are members of the superfamily of plant, fungal and bacterial peroxidases (Welinder, 1992). They are mostly multimers of identical subunits of approximately 80 kDa each. Most form tetramers of Mw 240 000, with two protoheme IX groups which can be reduced by dithionite. These enzymes are different to normal catalases in their narrow pH ranges, sensitivity to temperature, hydrogen peroxide, and organic solvents, and resistance to 3-amino-1,2,4,-triazole (Goldberg and Hochman, 1989; Hochman *et al.*, 1992) So far, these types of catalase-peroxidases have been found in *E. coli* (Claiborne and Fridovich, 1979), *Rhodopseudomonas capsulata* (Hochman and Shemesh, 1987), *Klebsiella pneumoniae* (Hochman and Goldberg, 1991), *Chromatium vinosum* (Nadler *et al.*, 1986), *Rhodobacter capsulatus* (Hochman *et al.*, 1992), *Salmonella typhimurium* (Loewen and Stauffer, 1990), *Citrobacter freundii*, *Edwardsiella tarda*, *Enterobacter aerogenes* (Switala *et al.*, 1990), and the alkalophilic Bacillus strain YN 2000 (Yumoto *et al.*, 1990). These enzymes are highly homologous and consist of two related domains, each with strong similarity to the yeast cytochrome *c* peroxidase. The catalase-peroxidase activity is predicted to lie in the N-terminal domain and, as yet, no function has been assigned to the C-terminal domain, although it may play a role in substrate binding or subunit interaction. In bacterial catalase-peroxidases, the first 50 residues have no counterpart in the yeast enzyme, and the N-terminal halves are more related to the yeast enzyme than the C-terminal halves. (Welinder, 1991; 1992).

The catalytic mechanisms of catalase-peroxidases are thought to occur in a similar manner to those described above for the individual enzymes. Hochman and Shemesh (1987) suggest that, for both catalases and catalase-peroxidases, Compound I is formed at the same rate while the rate of transition to Compound II is significantly higher for catalase-peroxidases. This results in a lower steady state concentration of Compound I. This explains why, in the presence of low and steady state concentrations of H_2O_2, activity is lost, since there is an increased rate of conversion of Compound I to Compound II, which then takes a long time to regenerate the enzyme. The inhibitor 3-amino-1,2,4,-triazole reacts with Compound I. Since the conversion to Compound II is very rapid for catalase-peroxidases, they appear to be more resistant to inhibition. When there is a substrate for peroxidase, the steady state level of Compound I will probably be lower than when only H_2O_2 is present (Hochman and Shemesh, 1987).

The catalase-peroxidase enzyme, as mentioned previously, occurs in a variety of bacterial species. However, it has been studied most extensively in *E. coli*. This organism possesses two catalase enzymes, HPI, which has associated peroxidase activity, and HPII, which possesses only catalase activity. HPII, which is located in the cytoplasm (Heimberger and Eisenstark, 1988), is encoded by the *katE* gene and is a hexamer of 93 kDa subunits and six heme d-like groups (Loewen and Switala, 1986). HPI, the catalase-peroxidase, is encoded by the *katG* gene. The *E. coli katG* promoter contains a -35 region with poor homology to the consensus sequence and requires the OxyR protein for optimal expression. The 2.181 kb gene encodes a monomer of 80.109 kDa (Triggs-Raine *et al.*, 1988). The protein is active as a tetramer with two protoheme IX groups (Claiborne and Fridovich, 1979), and is found in the periplasmic and cytoplasmic membrane fractions (Heimberger and Eisenstark, 1988). To establish which domains of the enzyme were important for activity, Loewen *et al.* (1990) generated mutants in HPI in *E. coli* using nitrosoguanidine and cloned and sequenced their *katG* genes. The mutants contained one each of the following substitutions, a Gly → Asp change in codons 119 and 314, and a Leu → Phen change in codon 139. The changes at positions 139, 119, and 314 resulted in 13, < 1, and 18% of the wild type catalase activity and 43, 4, and 45% of the wild type peroxidase activity, respectively. All three bound less protoheme IX than the wild type enzyme. The mutant enzymes were more sensitive to high temperature and β-mercaptoethanol than the wild type, but the pH profiles and sensitivities to inhibitors and activators were unchanged (Loewen *et al.*, 1990).

Another well characterized catalase-peroxidase is that from *S. typhimurium*. The *katG* gene from this organism was cloned and sequenced by Loewen and Stauffer (1990). The gene is 81.9% similar to that of *E. coli* at a nucleotide level and 90% similar at an amino acid level. The -10 region is identical to that of *E. coli*, while the -35 region differs in 2 of the 6 positions, and also requires OxyR for optimal expression. The gene encodes a protein of 727 amino acids with a predicted molecular weight of 80.042 kDa. Loprasert *et al.* (1988) isolated a catalase-peroxidase from *Bacillus stearothermophilus*. It is a thermostable dimeric enzyme with one molecule of protoheme XI per molecule and an optimum temperature of 70°C. The same authors cloned and characterized the gene encoding the enzyme (Loprasert *et al.*, 1989). The 2.193 kb gene encodes a polypeptide of 731 amino acids with a predicted molecular weight of 82.963 kDa. The *B. stearothermophilus* catalase-peroxidase shows 48% identity to that of

E. coli at the amino acid level (Loprasert *et al.*, 1989). Brown-Peterson and Salin (1993) purified and characterized a hydroperoxidase from *Halobacterium halobium*. It exhibited similar properties to other hydroperoxidases but had some unique features, including a greater amount of peroxidase activity in relation to catalase activity, a diminished heme content, inability to be reduced by dithionite, a requirement for salt for activity, and stability at higher temperatures. Another salt-dependent catalase-peroxidase was purified from the extreme halophile *Haloarcula marismortui* (Cendrin *et al.*, 1994). The catalase activity of this enzyme was enhanced by sodium chloride or ammonium sulphate, but the peroxidase activity decreased with increasing concentration of salt. Other catalase-peroxidases which have been characterized include one from *Caulobacter crescentus*, which is also located in both the cytoplasm and periplasm, associated with the outer surface of the inner membrane (Schnell and Steinman, 1995), and one from the cyanobacterium *Synechococcus* PCC 7942 (Mutsuda *et al.*, 1996). The latter enzyme has a molecular mass of 150 kDa, and is composed of two identical 79 kDa subunits with one protoheme molecule. It exhibits high catalase activity and moderate peroxidase activity with *o*-dianisidine and pyrogallol. Catalase activity is inhibited by potassium cyanide and sodium azide, but not by 3-amino-1,2,4-triazole (Mutsuda *et al.*, 1996).

1.2.4. The physiological role of catalase-peroxidases

A number of studies have therefore been carried out on the structure of the catalase-peroxidase enzyme and on the gene encoding it. Hochman *et al.* (1992), however, were the first to attempt to elucidate the physiological function of the enzyme. They studied the catalase-peroxidase from *Rhodobacter capsulatus*, a photosynthetic, facultatively anaerobic bacterium capable of fixing molecular nitrogen. *R. capsulatus* has a catalase-peroxidase and a peroxidase, as shown by activity gels. The authors used transposon mutagenesis to isolate a catalase-peroxidase mutant, AH18. The mutant reacted to molecular oxygen in a similar way to the wild type with respect to generation time under aerobic conditions, pigment content, mutation rate, and *in vivo* nitrogen fixation capacity, but was affected by hydrogen peroxide with respect to viability and inhibition of nitrogen fixation. Under aerobic conditions, the catalase-peroxidase mutant had a much higher rate of cell death in the stationary phase than the wild type. These results indicate, as has been demonstrated for other catalases, that they do not have a crucial function in actively growing aerobic cultures, unless the cells are exposed to externally added hydrogen peroxide.

Catalase mutants of *E. coli*, for example, exhibit a similar growth rate to wild type cells, but are more sensitive to the lethal effects of hydrogen peroxide (Levine, 1977; Meir and Yagil, 1984). In the absence of catalase, these cells can switch to utilizing glutathione as an antioxidant (Alonso-Moraga *et al.*, 1987). There are only a few reports in the literature which suggest a distinct function for catalase-peroxidases. In *Vitreoscilla* sp., the catalase-peroxidase is thought to play a specific role in scavenging peroxide generated by its bacterial hemoglobin (Abrams and Webster, 1990), while in *E. coli*, HPI is thought to play a general antioxidant role in exponentially growing cells, with a specific effect on membrane structure and function, and possibly active transport (Farr *et al.*, 1988; Leven *et al.*, 1990). In the latter organism, HPI is responsible for a very small part of the total catalase activity, however, it is the most important cellular determinant for hydrogen peroxide sensitivity (Schellhorn, 1994). Hochman *et al.* (1992) suggest that the *R. capsulatus* catalase-peroxidase has a minor function in exponential phase cells, but is crucial in stationary phase when the cells are starved or stressed. They suggest that, in actively growing cells, the peroxidase or damage repair systems protect against oxidative damage. Starved bacteria are subjected to oxidative stress, therefore catalase-peroxidases are crucial for survival under these conditions. The peroxidase activity of this enzyme may also catalyze biosynthesis or degradation activities for cellular adaptation to starvation or oxidize NADH, NADPH, or cytochrome *c*, therefore lowering the cell's reducing activity (Hochman *et al.*, 1992). In pathogenic bacteria, there is strong evidence to suggest that the catalase-peroxidase plays an important role in virulence (discussed in a later section).

1.2.5. Catalases in the mycobacteria

The catalase contents of various mycobacterial species have been studied extensively. In 1958, Schweiger and co-workers showed that virulent, avirulent and saprophytic mycobacteria could be differentiated by the thermal stability of their catalase enzyme, and, in 1960, Kubica and Pool introduced the 68°C, pH 7.0 catalase test for the subgrouping of mycobacteria based on this property. Wayne (1962) subsequently developed the semi-quantitative catalase assay, which entails the addition of 10% Tween 80 and 30% hydrogen peroxide to a culture and measuring the height of the column of bubbles produced. He used this to demonstrate that different variants of *M. kansasii* have differences in their catalase activity. Kubica *et al.* (1966) tested 326 cultures of slowly growing mycobacteria and 204 strains of rapidly growing mycobacteria

for catalase activity using both the semi-quantitative and pH 7/ 68°C tests. The tests revealed two natural groupings, one composed of strains of *M. tuberculosis*, *M. bovis*, clinically insignificant strains of *M. kansasii*, and other mycobacteria producing 40 mm or less of bubbles in the semi-quantitative test and showing a negative pH 7/ 68°C test, and the other containing the rest of the cultures which produced > 50 mm of bubbles and a positive pH 7/ 68°C test. A small group of strains showing no apparent catalase activity consisted primarily of tubercle bacilli resistant to varying levels of INH. Mycobacteria, therefore contain two types of catalases, the M-catalases which retain activity after heating to 68°C for 5 minutes, but are inactivated by aminotriazole, and the T-catalases which are inactivated by heat but are resistant to aminotriazole. Other studies since have demonstrated the presence of both types of catalase in *M. smegmatis* and only the T-catalase in *M. gastri* (Bartholomew, 1968). These results were independent of the age of the culture and were, therefore, not a result of the differences in the growth rates of the two mycobacterial species. Wayne and Diaz (1982; 1988) have also tested the catalase contents of several mycobacterial species and found the following; *M. terrae*, *M. triviale*, and *M. nonchromogenicum* have no T-catalase; *M. tuberculosis*, *M. bovis*, *M. intracellulare*, *M. avium*, *M. gastri*, *M. marinum*, and *M. xenopi* have a T-catalase but no M-catalase; *M. szulgai*, *M. simiae*, *M. kansasii*, *M. gordonae*, *M. scrofulaceum*, and *M. asiaticum* have both T- and M-catalases; and *M. avium* and *M. intracellulare* have a novel type of catalase, A-catalase, which is heat stable, sensitive to inhibition by 3-amino-1,2,4-triazole, has no peroxidase activity, and does not react with antibody to the M-catalase from *M. scrofulaceum*. Apart from the catalases and catalase-peroxidases, mycobacteria produce two types of SODs, Mn- and Fe-containing. *M. tuberculosis* produces and secretes an FeSOD which also plays a role in protection against reactive oxygen species (Zhang *et al.*, 1991).

1.3. THE *M. TUBERCULOSIS* CATALASE-PEROXIDASE

As mentioned above, certain mycobacterial species possess only a T-catalase. This catalase has associated peroxidase activity, and is therefore classified as a catalase-peroxidase or hydroperoxidase. The *katG* gene encoding the *M. tuberculosis* catalase-peroxidase enzyme was cloned and sequenced by Zhang *et al.* (1992), and characterized by Heym *et al.* (1993). The gene shows 53.3% homology with the enterobacterial HPI enzymes and 45.7% homology with that of *B. stearothermophilus*. Upstream of the gene are two copies of a 700 bp direct repeat

which are 68% identical. Probing of an *M. tuberculosis* cosmid library with a region of this repeat revealed that the repetitive sequence is not dispersed in the genome. The markers linked to *katG* on the chromosome are LL105, encoding an anonymous antigen, and *fbpB*, encoding the putative fibronectin-binding protein or α-antigen. *M. leprae* and *M. avium* contain sequences homologous to the *M. tuberculosis katG* gene, while *M. gordonae* and *M. szulgai* do not (Zhang *et al.*, 1992).

The *katG* gene encodes a protein of 735 amino acids with a predicted molecular weight of 80.029 kDa. The *M. tuberculosis* KatG, like all catalase-peroxidases, consists of two related domains, each with strong similarity to the yeast cytochrome *c* peroxidase. The region spanning residues 55 to 422, and that between residues 423 and 735, are weakly related (30% identity at the amino acid level), although many of the catalytic amino acid residues are absent from the C-terminal region (see Fig. 1.3), and it may play a role in substrate binding or subunit interaction. Zhang *et al.* (1992) demonstrated that deletion of a *Bam*HI-*Kpn*I fragment from the 3' end of the gene results in a loss of enzyme activity, and ability to confer sensitivity to INH. Within the N-terminal region, the putative active site regions (residues 94-154 and 257-276) are 45% identical, and contain the putative catalytic or substrate binding residues, Arg-104, Trp-107, His-108, Asn-138, His-269, Thr-274, His-275, Trp-320, and Asp-380, which are conserved (Heym *et al.*, 1993).

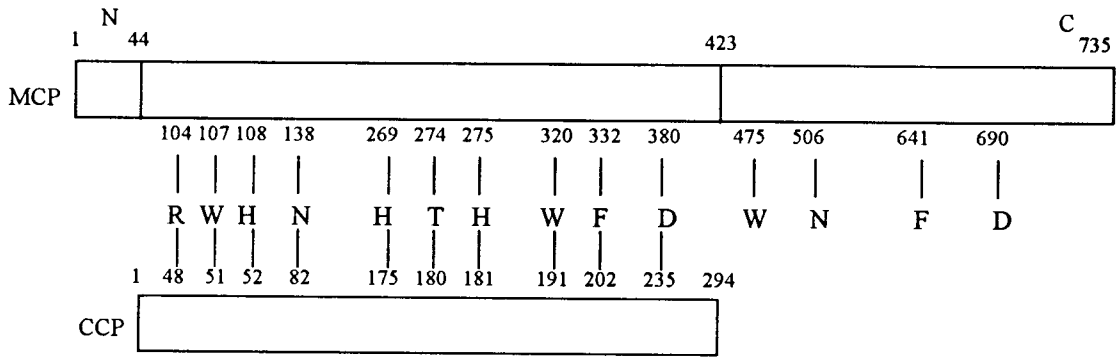


Figure 1.3 Schematic representation of the gene duplication observed in the *M. tuberculosis* catalase-peroxidase (MCP). The residues corresponding to the conserved active site residues of the yeast cytochrome *c* peroxidase (CCP) are shown.

The consensus sequences for the active sites of typical peroxidases are: peroxidase 1, [DET]-[LIVMT]-X_2-[LIVM]-[LIVMSTAG]-[SAG]-[LIVMSTAG]-H-[STA]-[LIVMFY] and peroxidase 2, [SGAT]-X_3-[LIVMA]-R-[LIVMA]-X-[FW]-H-X-[SAC] (Bairoch *et al.*, 1992). Only the first is found in the *M. tuberculosis* KatG, since there are two deviations around His-269 where the second motif should be. A possible ATP-binding motif (G-X-X-X-X-G-K-T) was also detected, although this overlaps with the active site and may not be functional.

The *M. tuberculosis* catalase-peroxidase was first purified by Devi *et al.* (1975). The purified protein was reddish-brown and had a molecular weight of 240 000. The enzyme activity was in the supernatant fraction, therefore it is not localized in the membrane in this organism. Johnsson *et al.* (1997) recently purified the protein from catalase mutants of *E. coli* expressing the *M. tuberculosis katG* gene on a plasmid, and studied its spectroscopic and kinetic properties in detail. They determined the molecular weight of the protein to be approximately 161-175 kDa, suggesting that the protein forms a dimer in solution. This agrees with previous reports (Diaz and Wayne, 1974; Nagy *et al.*, 1995), but not with that of Devi and co-workers (1975). The purified enzyme possessed high catalase activity which showed a significant dependence on pH, with a maximum activity at pH 7.0. The specific catalase activity of the recombinant KatG was approximately 6-fold higher than the KatG purified from *M. tuberculosis* cell extracts (Diaz and Wayne, 1974), and approximately 150-fold higher than that purified from *M. tuberculosis* H37Rv (Devi *et al.*, 1975). This may be due to difficulties in purifying the enzyme from the native hosts and differences in the purification protocol. The peroxidase activity was approximately 8-fold lower than that of horseradish peroxidase (Hochman and Goldberg, 1991). This may be due to direct competition between the catalase and peroxidase activities of the KatG enzyme for hydrogen peroxide. The enzyme oxidizes both NADH and NADPH, and this may provide a regulatory role for the enzyme in controlling metabolic rates (Hochman and Goldberg, 1991; Marcinkeviciene *et al.*, 1995).

Other mycobacterial catalases which have been characterized further are the *M. intracellulare, M. smegmatis,* and *M. leprae* KatG, and the *M. avium* KatE. The *M. leprae katG* gene was recently isolated and sequenced by Nakata *et al.* (1997). The sequence was approximately 70% identical to that of the *M. tuberculosis katG*, however, the *M. leprae* gene had several mutations. The *M. tuberculosis katG* GTG initiation codon was replaced by GTA in *M. leprae*

and the latter gene had two deletions of approximately 100 and 110 bp at the 5' end of the coding region, as well as several minor deletions, insertions, and termination codons. The same lesions were detected in 7 other *M. leprae* strains. The authors were unable to detect sequences homologous to the *M. avium katE* gene. It appears, therefore, that *M. leprae* produces no functional catalases, and this may explain both the natural resistance of the organism to INH, and its very slow growth *in vivo* due to its lack of defence against oxygen radicals (Nakata *et al.*, 1997). The *M. intracellulare katG* gene, which is 60% identical to that of *M. tuberculosis* (Heym *et al.*, 1993), has a G+C content of 70%, and encodes a protein of 745 amino acids with a predicted Mw of 83.5 kDa (Morris *et al.*, 1992). The protein is 60, 59, and 47% identical to the corresponding genes from *E. coli*, *S. typhimurium*, and *B. stearothermophilus*, respectively. The C-terminal region, in particular, shows a relatively high identity to the above three proteins. The *M. intracellulare katG* gene hybridized to homologous sequences in *M. asiaticum*, *M. avium*, *M. bovis*, *M. bovis* BCG, *M. fortuitum*, *M. kansasii*, *M. smegmatis*, *M. tuberculosis* H37Ra, *M. tuberculosis* H37Rv, *M. phlei*, *M. scrofulaceum*, and *M. terrae* (Morris *et al.*, 1992).

The *M. smegmatis* KatG was purified by Marcinkeviciene *et al.* (1995). The 330.0 kDa enzyme was composed of four 81.89 kDa subunits, and had a heme content of 2 molecules per tetramer. The peroxidase activity of the enzyme was competitively inhibited by isoniazid and nicotinic acid hydrazide, to the same extent as that from *M. tuberculosis*. Interestingly, the *M. smegmatis* catalase-peroxidase exhibited 10-fold greater peroxidase activity than the *M. tuberculosis* enzyme. The *M. avium katE* gene encodes a protein of 706 amino acids with a predicted molecular weight of 91 kDa. The gene shows 81.4; 80.8; 76.2; 76.4; and 83% similarity with the bovine liver, human kidney, *Saccharomyces cerevisiae*, *Lactobacillus sake* and *E. coli* catalases, respectively. The three active site residues of bovine liver catalase and the seven interacting with the heme group are conserved, as are 2 of the 4 residues involved in NADPH binding. Homologous *katE* sequences were identified in *M. aurum*, *M. kansasii*, and *M. marinum*, but not in *M. tuberculosis* (Milano *et al.*, 1996). The authors were also able to amplify the gene from different *M. avium* clinical isolates and *M. simiae*, *M. xenopi*, *M. terrae*, *M. kansasii*, and *M. gordonae*, and a minor band in *M. smegmatis* and *M. marinum*. No bands were amplified in *M. tuberculosis*, *M. bovis* BCG, *M. flavescens*, *M. fortuitum*, *M. abcessus*, or *M. cholenae*. Some of these results contradict those of Wayne and Diaz (1982; 1988) with

respect to the catalase contents of certain species. Milano *et al.* (1996) suggest that their *M. avium* A-catalase is the mycobacterial M-catalase, or HPII. The A-catalase is more similar to the M-catalase, but does not react with antibodies to the M-catalase from *M. scrofulaceum* (Wayne and Diaz, 1988).

1.3.1. Association of KatG with virulence

As mentioned previously, as an intracellular pathogen, *M. tuberculosis* has to overcome the host respiratory burst in order to survive. This respiratory burst involves the release of hydrogen peroxide and other toxic radicals, and it follows, therefore, that the removal of these molecules by radical scavenging enzymes would play an important role in the persistence of the organism in this environment, and hence in virulence. There are a number of lines of evidence to show that catalase is a virulence factor in the mycobacteria. Wayne (1962) showed that clinically significant strains of *M. kansasii* all had high levels of catalase activity and exhibited the greatest virulence for the guinea pig. INH-resistant strains of *M. tuberculosis* which have low catalase activity show a reduced virulence for guinea pigs (Middlebrook and Cohn, 1953; Mitchison *et al.*, 1960). A significant correlation was also found between virulence and resistance to hydrogen peroxide in INH-sensitive tubercle bacilli freshly isolated from patients (Mitchison *et al.*, 1963; Narayanan Nair *et al.*, 1964). Jackett *et al.* (1978) confirmed the association of hydrogen peroxide susceptibility with low virulence in native and INH-resistant tubercle bacilli, however, they showed that this is not the sole determinant of virulence, and that possession of catalase is only one of the features which confers resistance to hydrogen peroxide. Walker and Lowrie (1981) demonstrated that exogenous catalase protects *M. microti* from killing by lymphokine-activated murine macrophages. Direct evidence for the association between catalase and virulence is provided by Wilson *et al.* (1995), who showed that integration of a functional *M. bovis* BCG *katG* gene into an avirulent, catalase-negative *M. bovis* strain resulted in full virulence. These authors point out, however, that some clinical *M. tuberculosis* isolates that are resistant to INH and have lost catalase activity remain virulent. They suggest that these organisms have another method of coping with the oxidative burst from macrophages. More recently, Heym *et al.* (1997) showed clearly that *M. tuberculosis* strains which produce catalase-peroxidase have a competitive advantage in immunocompetent hosts, once the macrophages have been activated to release large amounts of reactive oxygen species. More

evidence is provided by Kelley *et al.* (ASM Abstract U-96, 1997), who showed that a *katG*-deleted, INH-resistant mutant of *M. tuberculosis* H37Rv was avirulent for mice and guinea pigs relative to the parent, but that transformation of the mutant with the *M. tuberculosis* wild type *katG*, one with an Arg to Leu change at codon 463, and the *M. avium* complex *katG* restored full catalase-peroxidase activity and ability to survive *in vivo*. The *katG* gene alone, however, could not restore virulence, although this may have been due to other mutations in this strain. These observations all suggest an important role for catalase in the survival and persistence of mycobacteria in macrophages, however, a number of recent studies contradict this. It was recently shown, for example, that in *M. intracellulare*, resistance to peroxide does not correlate with catalase content and, in *M. tuberculosis*, susceptibility to killing by activated macrophages is not related to peroxide susceptibility (Gangadharam and Pratt, 1984; O'Brien *et al.*, 1991). Further evidence was reported by Marklund and co-workers (ASM Abstract U-95, 1997), who found no differences in the virulence of two strains of *M. intracellulare* with mutated *katG* genes, although they showed elevated resistance to INH and sensitivity to hydrogen peroxide relative to the wild type parent strains. The role of catalase in the virulence of mycobacteria is therefore not clear. Tubercle bacilli appear to possess at least two independent mechanisms of resisting the lethal effects of hydrogen peroxide, and loss of virulence evidently can occur without loss of resistance to hydrogen peroxide (Jackett *et al.*, 1978). There are probably a number of factors involved in virulence which are interdependent. SOD has also been associated with virulence in some pathogenic bacteria (Babior, 1978; Beaman *et al.*, 1983; Franzon *et al.*, 1990; Welch *et al.*, 1979), but not in others (Mandell, 1975; Bloch *et al.*, 1989). No correlation was found between the SOD content and virulence of *M. tuberculosis* (Jackett *et al.*, 1978).

Catalase has been shown to be a potential virulence factor in a number of other pathogens such as *Haemophilus influenzae* (Bishai *et al.*, 1994), *Rhodobacter capsulatus* (Hochman *et al.*, 1992), *Staphylococcus aureus* (Mandell, 1975), *Nocardia asteroides* (Filice, 1983), *Leishmania donovani* (Beaman and Beaman, 1984), and *Neisseria gonorrhae* (Zheng *et al.*, 1992). Kagaya *et al.* (1992), for example, showed that purified stress-induced catalase (catalase-peroxidase), rather than constitutive catalase induces an increased protection to *S. typhimurium* infection in mice. Their results suggest that stress proteins play a major role in the induction of cell mediated immunity, and that the hydrogen peroxide-inducible catalase is one such effective antigen. The association of catalase with virulence in these organisms has also been suggested

to be due to its ability to protect from the oxidative burst of macrophages and other phagocytic cells (Beaman and Beaman, 1984). There are also some reports, however, which contradict the above observations. Genetically defined catalase mutants of *Listeria monocytogenes*, *S. typhimurium*, and *Shigella flexneri*, for example, have been shown to be fully virulent in animal models (Fields *et al.*, 1986; Franzon *et al.*, 1990; Gaillard *et al.*, 1986). Bishai *et al.* (1994) showed that *hktE* (catalase) mutants of *H. influenzae* exhibited a 2.3-fold reduction in their ability to produce lethal sepsis in infant rats, but were still able to produce persistent bacteremia. The virulence mechanisms in these organisms are evidently also complex and not fully understood. There appear to be multiple determinants related to pathogenicity and no component can be singled out.

1.3.2. Association of KatG with INH susceptibility

As mentioned in a previous section, isoniazid (INH) was first described in 1912 (Meyer and Malley, 1912) and was reported to be effective against *M. tuberculosis* and *M. bovis* in 1952 (Bernstein *et al.*, 1952; Fox, 1952; Offe *et al.*, 1952). It has been used both as a prophylactic agent and as a component of the drug regime to treat TB for nearly 40 years. Mycobacterial species show a range of sensitivities to INH. *M. tuberculosis* is highly susceptible (minimum inhibitory concentration (MIC) 0.02-0.05 μg/ml) while the saprophytic species, such as *M. smegmatis,* are the most resistant (MIC 32 μg/ml) (Zhang *et al.*, 1992). INH is chemically very similar to the nicotinamide in NAD and NADP (see Fig. 1.4) and is a very powerful chelator of metals (reviewed by Krishna-Murti, 1975).

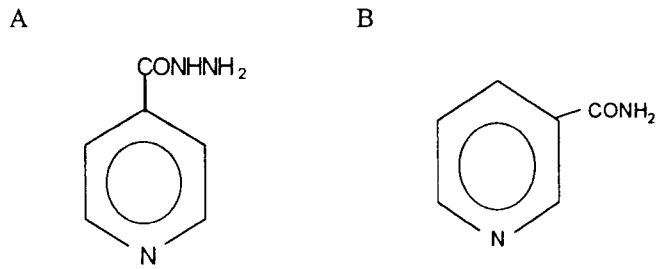


Figure 1.4 Chemical structures of A) INH, and B) nicotinamide.

1.3.2.1. Effect of INH on susceptible cells

A number of observations have been made over the years with regard to the effect of INH on the cells, although it is only recently that the mode of action and cellular target of the drug are becoming apparent. Nicotinamide adenine nucleotide co-enzymes play an important role in the regulation of metabolism. The concentrations of the reduced and oxidized forms are dependent on the composition of the medium. Bekierkunst (1966) reported that the exposure of *M. tuberculosis* H37Rv and *M. bovis* BCG (Bekierkunst and Bricker, 1967) to INH results in a decrease in the content of oxidized nicotinamide nucleotides. Their results were confirmed by Winder and Collins (1969), who noted that the concentrations of both NAD^+ and $NADP^+$ were higher in BCG than in *M. tuberculosis* H37Ra, and decreased to a smaller extent in the former. In human and bovine bacteria, nicotinamide or nicotinic acid is not incorporated directly into NAD, instead nicotinamide nucleotides are synthesized from quinolinic acid (reviewed by Krishna-Murti, 1975). *M. tuberculosis* has a higher capacity to produce the precursor of NAD than *M. bovis*, but has lower levels of NAD (Konno *et al.*, 1966; Dudley and Willet, 1966). It has been suggested that NAD^+ glycohydrolase, which cleaves NAD to degrade it, regulates the concentration of NAD in *M. tuberculosis*, and that INH antagonizes the action of an inhibitor of the enzyme, since the activity of NAD^+ glycohydrolase appears to be activated in the presence of INH (Bekierkunst and Bricker, 1967). In *M. bovis* BCG, no evidence exists for high levels of activity of this enzyme, therefore, the reduction of NAD concentration in the presence of INH would be smaller in this organism (Winder and Collins, 1968). Another suggestion is that INH is converted to isonicotinic acid which replaces the nicotinamide of NAD to form nicotinic acid analogues, which would disturb the NAD and NADP dehydrogenases (Krüger-Theimer, 1958). It was recently reported that sensitivity to INH could be restored to INH-resistant mutants of *M. smegmatis* by introduction of the *M. smegmatis ndh* gene encoding NADH dehydrogenase (Miesel *et al.*, ASM Abstract U-4, 1997). The authors suggest that the resistant mutants have elevated levels of NADH which may interfere with the interaction of INH with its target. Alternatively, the drug requires the oxidized form, NAD, for activity.

It has been shown that incubation with INH results in a loss of acid-fastness in approximately 90% of tubercle bacilli (Middlebrook, 1952), and blocks mycolic acid synthesis in *M. tuberculosis* (Winder, 1982; Winder and Collins, 1970) and in cell free extracts of mycobacteria (Quemard *et al.*, 1991), by inhibiting synthesis of unsaturated long-chain fatty

acids, which are precursors of mycolic acids. Winder and Rooney (1968), for example, showed that exposure of *M. bovis* BCG cells to INH resulted in a decrease in the triglyceride content of the wax A fraction. However, the triglyceride content of the chloroform-soluble and bound lipid fractions increased. Further experiments revealed a marked decrease in the content of fatty acids of chain length > C_{16} relative to those of length C_{12}-C_{16} in the triglycerides of all the lipid fractions of cells treated with INH. The authors suggest that INH may either inhibit the synthesis or incorporation of a fatty acid into triglyceride. Winder and Brennan (1965) also noted a reduced lipid content and altered cell wall permeability caused by INH. Takayama and co-workers (Takayama *et al.*, 1972; 1975) demonstrated clearly that INH inhibits the synthesis of unsaturated fatty acids of chain length ≥ 26 in *M. tuberculosis*, and that treatment with INH results in the accumulation of saturated fatty acids of the corresponding length.

Other changes which have been noted in *M. tuberculosis* or *M. bovis* BCG cells exposed to INH include altered pigment formation (Youatt, 1961b), reduced nucleic acid synthesis (Gangadharam *et al.*, 1963; McClatchy, 1971), and reduced protein synthesis (Tsukamura and Tsukamura, 1963), some of which may be due to inhibition of growth. DNA synthesis requires the enzyme DNA ligase which, in turn, depends on NAD as a co-factor (Olivera, 1971). The reduced nucleic acid synthesis could therefore be explained by the decrease in the levels of NAD caused by INH. INH also acts as an antimetabolite of pyridoxal. Further effects of INH have been recently reported. Kumar *et al.* (1994) observed that INH increases the activity of gamma-glutanyl transferase and decreases the activities of glutathione (GSH) biosynthetic enzymes in *M. smegmatis*, thus resulting in depletion of GSH. Garbe *et al.* (1996), on the other hand, found that exposure of *M. tuberculosis* H37Rv to INH resulted in induction and expression of several secreted proteins. Two of the prominent proteins induced are members of the antigen 85 complex, which have mycolyl transferase activity. The authors postulate that INH inhibits an intermediate which is involved in negatively regulating genes involved in mycolic acid metabolism (Garbe *et al.*, 1996). The effect of INH on susceptible cells is, therefore, varied and multifaceted, and may involve complex interactions between a number of cellular components. The mechanisms of action and cellular targets of the drug have become more apparent by the isolation and characterization of mutants which are resistant to it.

1.3.2.2. Resistance to INH - association with catalase activity

As early as 1954, Middlebrook reported on virulent tubercle bacilli which had developed resistance to INH. He noted that these bacilli exhibited a reduction in catalase activity, which was subsequently shown to correlate with the attenuation of their virulence for guinea pigs (Cohn *et al.*, 1954; Wolinsky *et al.*, 1956). Murohashi and Yoshida (1968) suggest that the effect of ultraviolet (UV) irradiation on the acid-fastness of tubercle bacilli, i.e. the strength of the cell wall, is related to their virulence, and used this to study strains which were resistant to > 5 μg/ml INH, of which > 50% had greatly attenuated virulence for guinea pigs. All of these strains lost their acid-fastness in a shorter period of UV irradiation than both virulent strains, and streptomycin- and PAS-resistant mutants which were not attenuated in virulence. The authors therefore showed a high correlation among the degree of resistance to INH, attenuation of virulence and lost of acid-fastness by exposure to UV irradiation. These findings suggest that an increase in resistance to INH in tubercle bacilli results in changes in the cell wall structure, while no marked changes in the wall occur upon the development of resistance to either streptomycin or PAS.

As mentioned above, the development of resistance to INH correlates with a reduction in catalase activity. Mycobacteria possess two different types of catalase which differ in their susceptibility to inactivation by heat. Winder (1960) proposed that the heat-labile catalase found in *M. tuberculosis* catalyzes the oxidation of INH to produce toxic free radicals which then cause inhibition of these organisms. This is supported by the results of Wayne and co-workers (1968), who noted that *M. tuberculosis*, *M. bovis*, and *M. gastri* strains were susceptible to ≤ 1 μg/ml INH and lost all catalase activity after heating to 68°C. However, they also showed that low-catalase *M. kansasii* strains which also contained only the heat-labile catalase exhibited identical susceptibility to the high-catalase strains. *M. flavescens* was the only rapidly growing mycobacteria studied to exhibit a susceptibility to INH as low as 1 μg/ml, and yet it retained some catalase activity after heating. Mycobacterial cultures possessing only a thermolabile catalase lost all catalase activity on the development of resistance to INH, while those possessing both a thermolabile and thermostable catalase retained some catalase activity, although this activity was reduced with increasing resistance to INH. The authors showed that *M. tuberculosis*, *M. bovis*, *M. gastri*, and the low-catalase *M. kansasii* cultures have a high affinity for INH, but lose the ability either to take up the drug, or to degrade it to release a

^{14}C-carbonyl group when they acquire resistance to INH. Up until 1975, loss of peroxidase activity had also been reported in all INH-resistant mutants (Devi *et al.*, 1975). This agrees with the fact that some atypical mycobacteria, which are naturally resistant to INH, also lack peroxidase activity (Winder, 1960).

1.3.2.3. Reaction of INH with peroxidase

When INH susceptible strains of *M. tuberculosis* are exposed to ^{14}C-isoniazid, a considerable amount of radioactivity is taken up by the cells. The maximal uptake depends on the concentration of INH up to 200 μg/ml, and is not determined by equilibration between the internal and external environments, since accumulation of INH in the cells is observed at low concentrations. Two non-inhibitory chemicals, isonicotinic acid and 4-pyridylmethanol have been identified inside the cells after exposure to INH (Youatt and Tham, 1969b). In 1961 Youatt noted that extracellular pigments are produced when washed whole cell suspensions of *M. tuberculosis* BCG are treated, in the presence of oxygen, with INH, its 2-methyl derivative or, to a lesser extent, its 2-ethyl derivative. These pigments were only produced by INH-sensitive strains and only in the presence of INH, i.e. not with other antituberculosis drugs or the isomers of isoniazid, nicotinic and picolinic acid hydrazides. Youatt and Tham (1969a) subsequently described the occurrence of a pigment precursor upon reaction of cell-free extracts of BCG with INH. This precursor was readily converted by dilute acid to a pigment with similar spectral properties to the pigment produced by the whole cells, suggesting that this may be the same product which has been metabolized further in the whole cells. INH was not lethal under these conditions, unless nutrients were added to the system. The reaction again occurred only with INH and its two derivatives, and the amount of pigment produced correlated with the bacteriostatic activity of the agent. No reaction occurred with any other hydrazides tested, but these, including benzoic acid, and *m*- and *p*-amino benzoic acid, competitively inhibited the reaction with INH and antagonized its bactericidal activity. The rate at which the cell extracts reacted with INH depended on a number of factors, including protein content, pH, temperature, aeration (the reaction does not occur under anaerobic conditions), INH concentration up to 200 μg/ml, NAD concentration, and other unidentified co-factors. Production of the pigment precursor occurred only in extracts from INH sensitive strains. Addition of 4-pyridylmethanol, 4-pyridine aldehyde, H_2O_2, hydrazine and isonicotinic acid, the known and possible products of

INH metabolism, inhibited the reaction. This suggests that these may be later products than those resulting from the above reaction. Further experiments showed that the reaction of INH with the cell-free extracts was not affected by the addition of various growth substrates to the medium, but that production of the precursor increased with the age of the culture to the end of the stationary phase, and then declined. The binding of INH to cells, and pigment production, are separable from the lethal reaction, which requires the addition of nutrients and occurs later. From their results, the authors proposed the following theory on the primary action of INH:

HYPOTHESIS: INH enters both sensitive and resistant cells freely (there was no indication that the rate of metabolism of INH by sensitive cells was controlled by the rate of entry into the cells). Once in the cell, INH, or a product thereof, binds to an enzyme system present in sensitive strains and displaces a pigment precursor molecule. If metabolic substrates are added, the situation becomes lethal, either due to modification of the enzyme system, or the synthesis of toxic molecules from the pigment precursor. Binding of INH is thought to occur through both the ring nitrogen and the hydrazide group, since loss of these results in a loss of bacteriostatic activity. The 2-methyl and 2-ethyl derivatives of INH have more bulky ring substituents, thus lowering the efficiency of binding, and hence the bacteriostatic activity. Other hydrazides are thought to bind through their hydrazide group only, and therefore do not displace the pigment precursor, but compete with INH for the hydrazide binding site (Youatt and Tham, 1969a).

In 1975, Devi *et al.* (1975) purified the enzyme involved in the the so-called Y-reaction, i.e. the pigment production reported by Youatt. Catalase, peroxidase and Y-enzyme activities were all associated with the same protein peak, and the enzyme activity increased with increasing concentration of INH and showed a biphasic response. The K_m obtained at lower INH concentrations was 7.69×10^{-5} M, while that at higher INH concentrations was 13.9×10^{-5} M, i.e. at lower concentrations of INH there is one binding site, while at higher concentrations there may be more than one site, suggesting a negative co-operativity. A similar effect was seen with NAD. The reaction was dependent on the presence of oxygen, and was inhibited by an analogue of INH, picolinic acid hydrazide; reducing agents such as GSH, β-mercaptoethanol, dithiothreitol, L-cysteine and sodium dithionite; thiol inhibitors such as *p*-chloromercuribenzoate and N-ethyl-maleimide; metal chelating agents such as *o*-phenanthroline, diethyl dithiocarbamate, and 8-hydroxyquinoline; and other inhibitors such as cyanide and azide. The

authors were also able to demonstrate catalysis of the Y-reaction by horseradish peroxidase and lactoperoxidase, but not by any catalase preparations. The above peroxidases were of a different size and had different temperature optimums for the reaction, but showed similar pH requirements and sensitivities to various inhibitors. This suggests that the Y-enzyme reaction may be a peroxidatic-type reaction. Shoeb *et al.* (1985a; 1985b) also demonstrated that both eukaryotic and mycobacterial peroxidases can catalyze oxidation of INH by hydrogen peroxide to yield toxic free radicals.

It had been demonstrated that the *M. tuberculosis* catalase-peroxidase reacts with INH in the presence of oxygen to produce pigments, however, this had not been related to its bactericidal activity. Shoeb *et al.* (1985), therefore, examined the possibility that oxidation of INH by exogenous peroxidase results in the production of active oxygen species in susceptible mycobacteria. Intermediates of the oxidation of INH by peroxidase include peroxy, carbonyl, and isonicotinyl radicals, while side reactions may yield hydroxyl radicals and superoxide anion. The authors found that extracts of a susceptible strain could mediate hydroxylation of phenol when INH and hydrogen peroxide were present, while extracts from a resistant strain could not. The hydroxylation rate was decreased by the addition of hydroxyl radical scavengers such as mannitol and dimethyl sulfoxide. The reaction between INH and peroxidase was enhanced at an alkaline pH and in the presence of SOD, which indicates that the superoxide anion is also produced. van Zyl and van der Walt (1994) recently demonstrated aromatic hydroxylation and the formation of thiobarbituric acid-reactive substances when INH and catalase were mixed, and this was stimulated by phytic acid, a potent metal chelator. This suggests that the reaction does not involve transition metal-catalyzed hydroxyl radical generation. Hydroxylation also occured with INH and phytic acid in the absence of catalase, although to a lesser extent. These authors suggest that peroxynitrite or a similar species is formed during INH oxidation. Johnsson and Schultz (1994) also identified isonicotinic acid, isonicotinamide, and pyridine-4-carboxaldehyde as products of INH oxidation by KatG.

INH and its meta isomer, nicotinic acid hydrazide, inhibit peroxidase enzyme activity equally, but the latter is not mycobactericidal, therefore, both isomers can bind through the hydrazine group to the enzyme and undergo initial oxidation, but oxidation intermediates may differ, depending on the pyrimidine ring. The differences in activities between INH and its isomer may,

therefore, be due to differences in the oxidation pathway. In the rate-determining step of the reaction, the INH carbanion species, which is the precursor of the carbonyl radical, is stabilized by a resonance structure in which two electrons are accommodated by the strong electronegative nitrogen of the ring. In the isomer, all resonance structures have the two electrons accommodated by the weak electronegative carbons of the ring. The INH intermediate is the most stable, and has a low transition energy and is, therefore, more effective than the *m*-isomer in the reaction.

It is evident from the above studies that a number of reactive species are produced by the oxidation of INH, and it is thought that these are direct or indirect effectors in the mechanism of action of INH, i.e. they are responsible for its mycobactericidal activity. The activation of INH by peroxidase requires the formation of metabolic hydrogen peroxide, therefore the effect of INH on mycobacteria decreases with reductions in metabolic rate, for example with carbon source depletion. The uptake and metabolism of INH also requires oxygen, therefore the MIC of INH against mycobacteria decreases with increasing oxygen concentration.

1.3.2.4. The genetics of resistance to INH

Zhang and co-workers (1992) were the first to study the genetics of INH susceptibility in *M. tuberculosis*. They transformed a highly resistant (MIC_{INH} 500 μg/ml), low catalase *M. smegmatis* mutant strain with an *M. tuberculosis* H37Rv genomic library and screened transformants on plates containing 32 μg/ml INH. They isolated a clone, designated pBH4, which conferred hypersensitivity to INH (MIC 8μg/ml), although it did not alter the MICs for other antituberculosis drugs, and caused a 2-3-fold overproduction of the catalase enzyme. Using oligonucleotide probes based on the *E. coli* and *B. stearothermophilus* hydroperoxidase I enzymes and enriched cloning, they isolated the gene encoding the *M. tuberculosis* heat-labile catalase enzyme. This region of DNA was found to be present in the pBH4 clone and could confer sensitivity to INH on the resistant *M. smegmatis* mutant. Wild type *E. coli* is naturally completely resistant to INH, even when transformed with constructs expressing the *M. tuberculosis katG* gene. Overexpression of this gene in a catalase negative strain, UM2 (*katG*, *katE*) (Mulvey *et al.*, 1988), however, rendered the cells susceptible to high concentrations of INH (50-500 μg/ml). The authors, therefore, demonstrated that the *katG*

gene alone can confer sensitivity to INH and must play an important role in the action of this drug. When they probed genomic DNA from INH-resistant and INH-sensitive strains of *M. tuberculosis* with the *katG* gene, however, they found the gene to be absent from only two highly resistant isolates (MIC > 50 μg/ml), suggesting that deletion of the gene is not a common mechanism of resistance. Zhang and Young (1994) later examined the *katG* region of the genomes of a panel of INH-sensitive and INH-resistant isolates of *M. tuberculosis* and found it to be highly variable and unstable in all. They isolated a 2 kb *Kpn*I fragment located 10 kb upstream of the *katG* gene in *M. tuberculosis* H37Rv which was specific to *M. tuberculosis* complex members and contained three copies of an internal 75 bp repeat flanked by multiple copies of repeated sequences related to the major polymorphic tandem repeat (MPTR) sequence reported by Hermans *et al.* (1992). PCR amplification of the three repeats from different *M. tuberculosis* isolates resulted in products of varying size due to differences in the number of copies of the repeat. The presence of these repeats may stimulate recombination or other DNA rearrangements, and may contribute to the instability of the *katG* region of the genome. Goyal *et al.* (1994) report the use of the above PCR amplification to differentiate *M. tuberculosis* complex strains. They were able to identify seven different subtypes based on the product size, but found no correlation between the pattern of subtypes and the geographical location, HIV status of the patient, or the INH susceptibility of the organism.

M. aurum is another mycobacterial species that is extremely sensitive to INH. Heym and Cole (1992) isolated and characterized INH-resistant mutants of this species and of *M. smegmatis*. Resistance to INH occurred at a frequency of 10^{-6} when these strains were plated on medium containing INH at 6x MIC. Two mutants with similar resistance levels, BH0 (*M. aurum*) and BH1 (*M. smegmatis*), were studied further. These mutants had unaltered MICs for other drugs i.e. resistance was probably not due to an increase in cell wall permeability, but grew more slowly than their parental strains, and the *M. aurum* mutants had altered colony morphology. Enzymological studies showed that the *M. smegmatis* mutant, BH1, produced two-fold less catalase activity and showed 14-fold less peroxidase activity than its parent, while BH0 produced barely detectable levels of catalase activity and showed a greater than 18-fold reduction in peroxidase activity. The authors extracted the mycolic acids from the INH-resistant mutants and found no difference in the pattern or levels of mycolate production when compared with the parent strains. Southern blot analysis of chromosomal DNA isolated from

the mutants revealed no alterations in the banding patterns when probed with the *M. tuberculosis katG* gene. The fact that the two INH-resistant mutants studied showed similar phenotypes to the corresponding mutants of *M. tuberculosis*, suggests a common drug target and mechanism of resistance in most mycobacteria. The authors suggest that the residual catalase activity seen in the mutants may be due to low levels of expression of *katG*, missense mutations in the gene, or a second catalase gene which has no peroxidase activity, but has the same electrophoretic mobility as KatG.

Since the cloning and sequencing of the *M. tuberculosis katG* gene, and the presentation of direct evidence for its role in INH susceptibility, a number of studies have been carried out in order to elucidate the genetic basis for INH resistance in naturally occurring and laboratory strains of *M. tuberculosis*. Stoeckle *et al.* (1993) examined the presence of the *katG* gene in 80 strains of *M. tuberculosis* isolated from patients in New York City, using PCR. They observed that 31/41 INH-resistant strains were PCR-positive, and 32 exhibited catalase activity, while 35/39 INH-sensitive strains were PCR-positive, and 30 exhibited catalase activity. Many of the PCR-negative strains had catalase activity and many PCR-positive strains had no catalase activity. These discrepancies may be due to mutations in PCR primer binding sites, technical factors related to PCR, and point mutations either in a regulatory region or within the structural *katG* gene, outside of the binding region of the primers used here. Altamirano and co-workers (1994) were able to amplify a fragment near the 5' end of the *katG* gene from 8/9 INH-resistant and one INH-sensitive isolate of *M. tuberculosis* from Vancouver, however the products from the resistant strains contained randomly distributed mutations including point mutations, deletions of one base, and insertion of up to three bases. Seven of the resistant isolates had reduced catalase activity, and one had no catalase activity at all. Ferrazoli *et al.* (1995) studied 53 INH-resistant and 49 INH-sensitive *M. tuberculosis* isolates from patients in Sao Paulo, Brazil. They found that 43 of the INH-resistant isolates expressed catalase activity, and 48 were *katG* PCR-positive. Only one isolate with an MIC > 50 μg/ml had a complete deletion of the *katG* gene. The first comprehensive analysis of the *katG* gene in *M. tuberculosis* isolates with a broad range of susceptibilities to INH was reported by Cockerill *et al.* (1995). These authors sequenced the entire *katG* gene from 15 isolates and used these to generate a consensus nucleotide sequence for use in assessing mutations. Only two strains, H37Rv and an isolate with an MIC of 8.0 μg/ml INH were identical to the consensus sequence. In 5/9 strains with

MICs ≥ 1.0 μg/ml, there was a conversion of arginine at codon 463 to leucine. This base alteration (G → T) also results in the loss of an *Nci*I-*Msp*I restriction endonuclease site. *Nci*I-digested *katG* fragments were, therefore, used in restriction fragment length polymorphism (RFLP) studies with 75 *M. tuberculosis* strains. Results revealed that 19/43 INH-resistant mutants had lost the restriction site. The authors found no consistent correlation between loss of catalase activity and resistance to INH (Cockerill *et al.*, 1995). Pretorius *et al.* (1995) studied the *katG* genes of 39 INH-resistant isolates of *M. tuberculosis* from different geographical locations by PCR-SSCP (single-strand conformation polymorphism). None of them had a complete deletion in *katG*, however 6 had point mutations, which included mutations in Thr-275 and Arg-463, and one had a small deletion. The authors later identified a mutation in codon 315 in 52% of the INH-resistant South African isolates, using a new primer set, which elevated the overall incidence of *katG* mutations in these strains to approximately 64% (Victor *et al.*, 1996). Recently, Haas *et al.* (1997) reported on the characterization of the *katG* genes of 124 INH-resistant *M. tuberculosis* complex isolates from south, central, and west Africa. They identified point mutations in codon 315 in 64% of these. As noted by others, none had complete *katG* deletions. Goto *et al.* (1995) analyzed 22 INH-resistant isolates of *M. tuberculosis* from Japan and Yemen for *katG* by PCR and catalase assays. Only 6 (27%) were PCR-negative and catalase-negative, 8 (36%) were PCR-positive and catalase-negative, and 8 were PCR-positive and catalase-positive. Rouse *et al.* (1995) observed perturbations in the *katG* genes of INH-resistant clinical isolates of *M. tuberculosis* from the US and South Korea. These included frameshift mutations in 3 highly resistant strains, a single base insertion at codon 521, small deletions, one of 8 bp and one of 22 bp, and the following amino acid substitutions, Thr-262 to Arg, Ala-350 to Ser, Gly-629 to Ser, Ala-717 to Pro, His-108 to Gln, Asp-63 to Glu, Ser-315 to Thr, Arg-463 to Leu, and Ile-335 to Thr. These resulted in a decreased catalase activity and increased resistance to INH. Of the 9 missense mutations, 6 occurred within functionally important residues (conserved in several bacterial hydroperoxidases), normally Asp-63, His-108, Thr-262, Ala-350, Ser-315, and Gly-629, while 5 of them, His-108, Thr-262, Ser-315, Ile-335, and Ala-350 may affect the structure of the protein (Welinder, 1992; 1991). Heym *et al.* (1995) also identified distinct mutations in the *katG* genes of a panel of INH-resistant clinical isolates of *M. tuberculosis*. These included; KatG1, Arg-463 → Leu; KatG2, deletion of residues 120-123; KatG3, insertion of isoleucine between residues 125 and 126; KatG4, Thr-275 → Pro; KatG5, Ser-315 → Thr; KatG6, novel

N-terminus; and KatG7, Leu-587 → Met. Sequencing of the *M. bovis* BCG and BCG Pasteur *katG* genes revealed three base differences from that of *M. tuberculosis* H37RV, codons 29 and 203, which were silent changes, and codon 463 which resulted in an Arg → Leu change, as seen in some INH-resistant *M. tuberculosis* strains. The KatG1 and KatG7 proteins displayed normal catalase and peroxidase activities and heat stability, whereas the activities of all the rest were reduced. The KatG2, 3, and 5 proteins showed a > 10-fold decrease in catalase activity and reduced heat stability, and the KatG4 protein showed no detectable activity. Five of the seven mutations identified lie in the N-terminal domain of KatG, which is predicted to contain the catalase and peroxidase active sites (Welinder, 1991; 1992). Four of the amino acid substitutions occur in positions analogous to those found in the *E. coli* HPI (Loewen *et al.*, 1990), which correlate with modification of the active site.

Because most of the INH-resistant strains studied thus far are clinical strains, they do not necessarily have the same genetic background, therefore it is not possible to directly correlate specific mutations with INH-resistance. For this reason, Rouse and Morris (1995) isolated and characterized INH-resistant mutants of *M. tuberculosis* H37Rv. They found that substitution of a proline for threonine at residue 275 resulted in loss of catalase activity and decreased susceptibility to INH. The authors also examined the molecular mechanisms of INH-resistance in 6 *in vitro* mutants of *M. tuberculosis* complex strains using DNA hybridization, SSCP, sequence analysis, immunoblotting and catalase assays. They performed immunoblots using a polyclonal antibody against the 80 kDa catalase-peroxidase to assay for the presence of the enzyme in 10 *M. tuberculosis* complex strains, including 4 INH-sensitive parental strains (H37Ra, H37Rv, and 2 *M. bovis* BCG strains) and 6 INH-resistant mutants (4 *M. tuberculosis* and 2 *M. bovis* BCG). Of the INH-resistant strains, 5 were resistant to at least 50 μg/ml INH, while the last was resistant to 5 μg/ml INH. The catalase-peroxidase was not detectable in 2 of the INH-resistant *M. tuberculosis* strains and the 2 INH-resistant *M. bovis* BCG strains. These strains also had no catalase activity, and neither did one of the strains which synthesizes an apparent full-length enzyme. A closer look at the genetics revealed that 5 of the highly INH-resistant *M. tuberculosis* complex strains had significant alterations in the *katG* gene, with 3 having deletions, one of which was a complete deletion of the gene, and 2 having point mutations. As a further means of correlating mutations in *katG* with resistance to INH, Rouse *et al.* (1996) used site-directed mutagenesis to alter the wild type *M. tuberculosis katG* gene at codons shown to be mutated in clinical isolates, and introduced the altered genes into INH-

resistant strains of *M. smegmatis* and *M. bovis* BCG. An Arg → Leu change at codon 104 resulted in a 100-fold reduction in peroxidase activity in *M. smegmatis*, while mutations in codons 108, 138, 148, 270, 275, 321, and 381 resulted in a 10-100-fold decrease in activity. Mutations in codons 463 and 315 resulted in a 50% decrease in activity, while mutations in codons 587 and 350 had no significant effect on peroxidase activity. All the genes containing mutations which reduced the peroxidase activity were unable to confer sensitivity to INH on the *M. smegmatis* cells, while the wild type enzyme could. Similar results were observed in BCG, except for the enzyme carrying the mutation in codon 463, which exhibited wild type peroxidase activity and ability to confer sensitivity to INH. Immunoblot studies indicated that, for some of the transformants, the decrease in peroxidase activity was due to a decrease in the levels of the KatG protein. Therefore, some of the mutations primarily affect the intracellular concentration of KatG, possibly by affecting KatG stability, some have a more direct effect on enzyme activity, while others may affect both the intracellular levels and the enzyme activity. Interestingly, the authors noted no differences in the *katG* mRNA levels of the transformants, suggesting that the regulation of the protein may occur at a post-transcriptional level (Rouse *et al.*, 1996). These results confirm that residues 104 and 108, which form the catalytic sites, and residues 270, 275, and 315, which may be involved in heme binding, are important for both catalase and peroxidase activity and the ability to confer sensitivity to INH. Residue 321 may be involved in electron transfer, while residue 381 is critical for the formation of a hydrogen-bond with His-270. Residues 138 and 148 are also important for activity, however their role is undefined. Mutations in the C-terminal residues, which have no counterpart in the yeast cytochrome *c* peroxidase, do not significantly alter activity. Interestingly there were no mutations which affected only one activity, which suggests that the catalase and peroxidase activities share catalytic sites (Rouse *et al.*, 1996).

All of the above studies provide evidence that the *katG* gene product plays a significant role in the susceptibility of certain mycobacteria to INH. In general, reduction in catalase activity could be correlated either with mutations in important catalytic residues, frameshift mutations, or small deletions. So far at least 20 different mutations have been identified in the *katG* gene. Musser (1995) reviews all the studies carried out on INH-resistant strains, and suggests that a restricted subset of missense changes in *katG* are repeatedly associated with resistance. The mutations are summarized in table 1.1. Many of these affect either the binding sites, active sites, or conformation of the enzyme. Two of the frequent mutations affect a glycine-rich loop which

holds together 2 α-helices, one of which contains essential catalytic residues. The mutations change the length of the loop and, therefore, possibly the position of the catalytic residues. A Ser → Thr conversion distorts the plane of the heme group in the active site. In general, mutations in the C-terminal region of *katG* do not affect enzyme activity and result in very low levels of resistance to INH (0.1 μg/ml). The most frequent mutation in INH-resistant *M. tuberculosis* isolates (20-40%) is an arginine to leucine substitution at codon 463, which lies in the C-terminal repeat at the end of helix A, and not in the vicinity of the peroxidase active site. This may represent a substrate-binding site within a domain that interacts with INH. The mutation appears to confer low level resistance to INH, but does not significantly affect enzyme activity. Interestingly, this substitution occurs in all BCG strains tested, and is also found in the KatG-like proteins of *B. stearothermophilus* (Loprasert *et al.*, 1988) and *M. intracellulare* (Morris *et al.*, 1992). BCG and *M. intracellulare* are both naturally more resistant to INH than *M. tuberculosis* (*M. tuberculosis* H37Rv is inhibited by 0.05 μg/ml INH while BCG Pasteur grows normally on 0.2 μg/ml INH, and the MIC for INH of *M. intracellulare* is approximately 5-10 μg/ml). More recent evidence, however, suggests that the substitution has no effect on INH-resistance, and is simply a polymorphism. For example, the introduction of an Arg → Leu substitution at codon 463 of the *katG* gene of *M. tuberculosis* H37Rv by site-directed mutagenesis, and introduction of the altered gene into the catalase-negative, INH-resistant *M. smegmatis* strain, BH1, resulted in at least a 20-fold increase in the susceptibility to INH, as noted for the wild type gene. Catalase and peroxidase activity was essentially the same for both, which suggests that the mutation does not significantly alter activity (Rouse *et al.*, 1996). These results were confirmed recently by Johnsson *et al.* (1997), who showed that a purified KatG enzyme containing the mutation displayed identical spectroscopic properties, catalase and peroxidase activities, and ability to mediate inactivation of InhA by INH, to the wild type enzyme. These authors point out, however, that the experiments were performed *in vitro*, and may differ from the *in vivo* situation. They suggest that, *in vivo*, the mutant enzyme may have a higher activity against another substrate which may compete with INH. However, other evidence suggests that codon 463 mutations are independent of INH-resistance and are linked to the geographical origin of the strains (Haas *et al.*, 1997; Shim, 1997).

Table 1.1 Summary of mutations identified in the *katG* gene of INH-resistant strains of *M. tuberculosis*.

katG mutation	Catalase activity	Reference
Asp-63	-	Rouse *et al.*, 1995
Arg-104	-	Rouse *et al.*, 1996
His-108	-	Rouse *et al.*, 1995
Thr-262	-	Rouse *et al.*, 1995
His-270	-	Rouse *et al.*, 1996
Thr-275	-	Rouse *et al.*, 1995; Pretorius *et al.*, 1995; Rouse & Morris, 1995
Ser-315	-	Rouse *et al.*, 1995; Pretorius *et al.*, 1995; Haas *et al.*, 1997
Ile-335	-	Rouse *et al.*, 1995
Ala-349	-	Rouse *et al.*, 1995
Ala-350	-	Rouse *et al.*, 1995
Arg-463	+	Rouse *et al.*, 1995; Pretorius *et al.*, 1995; Cockerill *et al.*, 1995
Leu-587	+	Rouse *et al.*, 1995
Gly-629	-	Rouse *et al.*, 1995
Ala-717	-	Rouse *et al.*, 1995
deletion of 120 - 123	-	Rouse *et al.*, 1995
insertion of Ile-126	-	Rouse *et al.*, 1995
insertion at codon 521	-	Rouse *et al.*, 1995
complete deletion	-	Ferrazoli *et al.*, 1995; Rouse & Morris, 1995; Zhang *et al.*, 1992
frame-shift mutations	-	Rouse *et al.*, 1995

Deletion of the entire *katG* gene is apparently not a common mechanism of resistance to INH, although it results in the highest levels of resistance. It is probably not favourable for the growth of the organism, particularly in macrophages. In many cases, the resistant strains produce an altered enzyme which still has catalase activity although it is reduced. It is possible that the maintenance of some catalase-peroxidase activity is required in order for the organism to withstand the hostile environment or to grow more rapidly. The reaction between the altered enzymes and INH that results in activation of the drug possibly no longer occurs or occurs at a reduced rate in INH-resistant strains, thus resulting in a decreased bactericidal effect (Heym *et al.*, 1995). Apart from its role in INH oxidation, it has been suggested that KatG is involved in the transport of INH into the cell. Studies in *E. coli*, *M. bovis* BCG, and *M. smegmatis* have demonstrated that KatG is probably surface-associated, and therefore may be involved in transport processes (Heimburger and Eisenstack, 1988; Winder, 1960; Youatt and Tham, 1969). Many INH-resistant, catalase-negative *M. tuberculosis* strains have a reduced capacity

to take up INH (Youatt, 1969), and expression of *katG* in *E. coli* allows rapid recovery of transport processes inhibited by oxidative stress (Farr *et al.*, 1988). However, in *M. tuberculosis* the enzyme is not localized on the membrane, therefore its involvement in active transport in this organism is not clear.

1.3.2.5 The cellular target/s of INH

It is evident that the *M. tuberculosis* KatG catalyzes the hydrogen peroxide-dependent peroxidation of INH to an activated product which then acts on the target, however, no indication has been given as to what the target might be. Earlier observations on the effect of INH on mycolic acid synthesis (Takayama *et al.*, 1975) suggest that the target must be an enzyme involved in the production of unsaturated 24-carbon fatty acids. In 1994, Banerjee and co-workers identified the *inhA* gene as a possible target of INH. These authors constructed a genomic library from a spontaneous INH-ETH-resistant mutant of *M. smegmatis* mc^2651 and an INH-resistant strain of *M. bovis*, and used these to transform wild type *M. smegmatis* strains and screened for resistant colonies. They identified fragments that conferred resistance to INH and ETH on wild type INH-ETH-sensitive strains of *M. smegmatis*, *M. tuberculosis*, *M. bovis*, *M. bovis* BCG and *M. avium*, and isolated the corresponding fragments from wild type *M. smegmatis*, *M. bovis*, and *M. tuberculosis*, and INH-resistant *M. smegmatis* and *M. bovis*. Each had two open reading frames (ORFs), one encoding a protein of 29 kDa, and the other 32 kDa, the latter of which, from *M. smegmatis*, was alone able to confer resistance to INH, and was renamed the *inhA* gene. The *M. tuberculosis* and *M. bovis inhA* genes lie on an operon with the first ORF and have a much shorter noncoding region between the two. The InhA protein shows 32% identity with enoyl-acyl carrier protein (ACP) reductase of *Brassica napus*, and 40% identity over 203 amino acids to the EnvM protein of *S. typhimurium* and *E. coli*, which catalyzes the reduction of crotonoyl-ACP (Bergler *et al.*, 1994), and is the target of the diazaborine group of antibacterial compounds (Bergler *et al.*, 1992; Grassberger *et al.*, 1984; Turnowsky *et al.*, 1989). The *M. tuberculosis* H37Rv and *M. bovis inhA* genes are identical, and have > 95% identity with the *M. smegmatis inhA* gene. The first ORF also shows similarity to other fatty acid biosynthesis proteins, particularly the *E. coli* 3-ketoacyl-acyl carrier protein reductase. The InhA protein has a putative binding site for nicotinamide or flavin nucleotides, and may use these as substrates or co-factors (Banerjee *et al.*, 1994). The difference between

the InhA of the mutant strains and that of the wild type was a single substitution of serine to alanine at position 94. It is interesting to note that resistance of the *E. coli* EnvM to the diazaborine group of antibacterial compounds is mediated by a Gly → Ser mutation at codon 93, which lies in a similar position to codon 94 of InhA (Grassberger *et al.*, 1984). Further experiments proved that this mutation mediates resistance to INH in *M. smegmatis*.

Dessen *et al.* and Quemard *et al.* (1995) further characterized the *M. tuberculosis* InhA by looking at its biochemical properties and chemical structure. They showed that the protein, like other ACP reductases, catalyzes reduction of 2-*trans*-octenoyl-ACP. It functions as a component of the type-II fatty acid synthesis system, responsible for the final reduction step in chain elongation to form conventional fatty acids (Dessen *et al.*, 1995; Johnsson *et al.*, 1995). The enzyme preferentially reduces long chain substrates (12-24 carbons), and therefore may be involved in mycolic acid synthesis. The authors were, however, unable to demonstrate binding of INH, its KatG-generated metabolites, or ETH to the enzyme or its complex with NADH. The authors also examined the mutated form and showed that, for the same reaction, the K_m and V_{max} values were the same, but the affinity for NADH was 5-fold lower. Therefore, drug resistance may be related to the specific interactions between InhA and the co-factor NADH. Possibly activated INH binds to and inactivates InhA when bound to NADH, therefore the mutant enzyme, which has a lower affinity for NADH, would be less inhibited. In *M. tuberculosis* H37Rv, the *in vivo* levels of NADH are low (Winder and Collins, 1969), therefore small changes in NADH affinity may have dramatic effects on the MIC of INH. The authors suggest that binding of NADH causes a conformational change in the active site which affects the substrate binding site (Dessen *et al.*, 1995). Alternatively, the activity of InhA may be inhibited by incorporation of an NAD analogue containing isonicotinic acid, which would block mycolic acid synthesis and cause a loss of acid-fastness. Johnsson *et al.* (1995) report that oxidation of INH by KatG leads to inactivation of InhA. This inactivation is inhibited by the addition of the enzyme substrate, and requires the presence of NADH or NAD^+. The inactivation of InhA also requires oxygen, the manganese ion, and INH, but occurs 2.7 times faster in the presence of KatG. The KatG-independent inactivation is probably mediated by superoxide, since it can be prevented by the addition of SOD (Zabinski and Blanchard, 1997). The activated forms of ETH and INH react with Cys-243 of InhA, leading to irreversible inactivation (Johnsson *et al.*, 1995).

There have been several reports of INH-resistant strains with mutations in the *inhA* locus, and at least six different mutations have been identified (Heym *et al.*, 1995; Rouse *et al.*, 1995). Reported perturbations include mutations resulting in the following substitutions, Ser-94 → Ala and Ile-16 → Thr (the latter mutation changes the configuration of the active site) (Heym *et al.*, 1994; Banerjee *et al.*, 1994; Kapur *et al.*, 1995); and mutations near the translation start codon of *orf1*, which may be a regulatory region of the *inhA* gene. For example, a T → G transversion 8 bp upstream of the *orf1* initiation codon, adjacent to a predicted ribosome binding site, and a T → G transversion 24 bp upstream of the initiation codon. These mutations may cause overexpression of *inhA*, therefore increasing levels of the INH target, and resulting in increased resistance. Wilson *et al.* (1995) demonstrated that loss of catalase and mutation of *inhA* have a cumulative effect in INH-resistance, but that the *katG* mutation contributes significantly more to the resistance than the *inhA* mutation.

In total, of all the INH-resistant strains studied, 50-64% had mutations in *katG*, 14% had mutations in *inhA*, some of which also had *katG* mutations, and 22% were $katG^+$ but had no detectable enzyme (Musser, 1995). Therefore several INH-resistant clinical isolates have no detectable mutations in *katG* or *inhA*, suggesting either that other mechanisms of resistance to INH may exist, or that *inhA* or *katG* have certain regulatory defects in these strains. In the strains which had no mutations in *katG* and yet had no detectable catalase or peroxidase activity, expression of the *katG* gene may have been impaired, possibly by mutation of an unlinked regulatory gene. Alternatively, KatG may be subject to post-transcriptional regulation. It is also likely that the mode of action of INH and mechanism of resistance to it are far more complex and involve interactions between a number of cellular components. Dubnau *et al.* (1996), for example, showed that expression of ribosomal protein S13 from either *M. bovis* BCG or *M. smegmatis* on a multicopy plasmid confers sensitivity to INH and a small colony phenotype on *M. smegmatis*, but not on BCG. This increased sensitivity occurs due to an increased KatG production. It may be that overproduction of S13 induces a stress response which then induces *katG*, and that this response is absent in BCG.

It is evident from the studies on INH-resistant clinical and laboratory strains that resistance to INH is, in most cases, not mediated by mutations in *inhA*. This suggests that it may not be the target of INH in these strains, and this is supported by the results of Mdluli *et al.* (1996). These

authors showed that treatment of *M. tuberculosis* with INH resulted in an immediate accumulation of 24- and 26-carbon fatty acids. In *M. smegmatis*, however, there was no such accumulation at any INH concentration tested. Treatment with INH resulted only in inhibition of short chain fatty acid biosynthesis, possibly by inhibition of InhA. These authors confirmed that overexpression of the *M. tuberculosis inhA* gene causes resistance to INH in *M. smegmatis*, but not in *M. tuberculosis*. InhA is, therefore, not the most significant target of INH in the *M. tuberculosis* complex organisms. Both the wild type and mutant InhA enzymes could confer resistance in *M. smegmatis*, therefore resistance is determined more by the level of expression than by the presence of a particular mutation. The observations suggest that the target of INH in *M. tuberculosis* is much further along in mycolic acid biosynthesis.

Barry and co-workers (Personal communication) have recently identified another possible target of INH in *M. tuberculosis*. They found that treatment of *M. tuberculosis* with INH, at a concentration known to induce the accumulation of hexacosanoic acid (C_{26}), resulted in the induction of the 12 kDa AcpM protein, a small acyl carrier protein (ACP), as well as a larger form of the same protein (80 kDa), which incorporates radio-labelled INH. ACPs are small proteins, containing a phosphopantetheine attachment, that transport the growing fatty acyl chain between component enzymes of type-II fatty acid synthase systems (Prescott and Vagelos, 1972). AcpM from INH-treated cells was shown to carry predominantly hexacosanoic acid, while AcpM from untreated cells carries a broader range of fatty acids (up to 50 carbon atoms). This suggests that AcpM is involved in the type-II fatty acid synthase system which produces the meromycolate branch of mycolic acids. Further analysis of the larger form of the AcpM protein revealed that it is a complex of AcpM-INH-KasA, where KasA is a β-ketoacyl synthase, which plays a role in fatty acid chain extension by adding 2 carbons to an ACP-bound saturated fatty acid. Inhibition of KasA would, therefore, be expected to stop fatty acid elongation, resulting in the accumulation of precursors, as observed upon treatment with INH. The gene encoding KasA is located immediately upstream of that encoding AcpM on the genome. Expression of the former gene in *M. tuberculosis* conferred resistance to approximately 2 μg/ml INH, while overexpression of the latter was toxic to cells. The upregulation of the *acpM* gene by INH may, therefore, be partly responsible for killing the cells. These results strongly suggest that AcpM and KasA are the cellular targets of INH in *M. tuberculosis* (Barry, Personal communication).

1.3.2.6. Why is *M. tuberculosis* uniquely sensitive to INH?

A great deal of information is now available on the mechanism of activation of INH and its cellular targets in the mycobacteria, however, the reasons for the unique sensitivity of the *M. tuberculosis* complex organisms to INH remain unclear. The fact that other mycobacteria, including *M. leprae,* are far less susceptible to INH suggests that *M. tuberculosis* has some unique features which are targeted by the drug. Barry (ASM Presentation, 1997) points out that, in *M. tuberculosis*, the MIC of INH correlates directly with mycolic acid synthesis, and that loss of mycolic acids results in cell death. In *M. avium*, on the other hand, mycolic acid synthesis shuts off at a much lower concentration of INH than the MIC, which suggests that the organisms can survive without mycolic acids. The loss of the permeability barrier in the presence of INH, however, lowers the MIC of other drugs. All mycobacteria have similar permeability barriers, therefore differences in the uptake of INH are unlikely, and the mycolic acid biosynthetic pathway is present in all mycobacteria. The *katG* and *inhA* genes, which are known to be involved in INH susceptibility, are both highly conserved and are present in almost all mycobacteria, including those that are naturally resistant to INH. Cole (1994) suggests that the KatG of *M. tuberculosis* has unique features that allow it to use INH as a substrate. However, peroxidases from widely different sources can react with INH, and both the *E. coli* and *M. intracellulare* KatG can partially restore INH-sensitivity to INH-resistant mutants of *M. tuberculosis* (Zhang and Young, 1993; Rouse *et al.*, 1996). Another possibility is that the differences in INH-susceptibility may be related the levels of hydrogen peroxide in the cell and may, therefore, depend on the presence of a second, heat-stable catalase. In the presence of two catalases, one of which is not inhibited by INH, destruction of hydrogen peroxide by the second catalase should inhibit hydrogen peroxide-dependent, peroxidase-mediated oxidation of INH (Shoeb *et al.*, 1985a; 1985b; Johnsson and Schultz, 1994). Milano *et al.* (1996), however, demonstrated that introduction of the *M. avium katE* into *M. tuberculosis* did not alter the MIC for INH and, therefore, did not antagonize activation of INH. These results suggest that the reasons for the exquisite sensitivity of *M. tuberculosis* to INH lie elsewhere.

1.3.2.7. The role of the oxidative stress response in susceptibility to INH

The catalase-peroxidase, apart from its role in INH oxidation, also plays an important role in protection of the organism against excess hydrogen peroxide. Deretic *et al.* (1995) suggest the possibility that other enzymes which are involved in the protection of *M. tuberculosis* against oxidative stress may also play a role in the sensitivity of the organism to INH. In *E. coli* and *S. typhimurium*, under conditions of oxidative stress, i.e. in the presence of excess hydrogen peroxide, the *oxyR* gene induces expression of *katG* as well as a number of other genes including *ahpC* (small subunit of alkyl hydroperoxide reductase), *dps* (19 kDa DNA binding protein from starved cells), *oxyS* (divergently transcribed from *oxyR* in *E. coli*), and *gorA* (glutathione reductase) (Christman *et al.*, 1985; Tartaglia *et al.*, 1989; Altuvia *et al.*, 1994; Toledano *et al.*, 1994; Storz *et al.*, 1990; Farr and Kogoma, 1991). The oxidative stress response is discussed in more detail in Chapter 5. Mutations in either *oxyR* or *ahpC* in *E. coli* result in sensitivity of the organism to INH (Rosner, 1993). Rosner and Storz (1994) demonstrated that transcription of the *oxyR*-regulated *oxyS* gene was not induced by INH in wild type cells, therefore INH does not activate *oxyR*. However, a *katG ahpCF* mutant had constituvely high levels of *oxyR* transcription. The authors suggest that, in these mutants, endogenous oxidants accumulate which leads to constitutive *oxyR* activation and INH susceptibility. Consistent with this, was the finding that hydrogen peroxide and cumene hydroperoxide potentiated the INH susceptibility of wild type cells, while the antioxidant, ascorbic acid, protected the susceptible mutant strain from INH. Superoxide radicals generated by paraquat did not enhance the INH susceptibility of wild type cells. Hydrogen peroxide also potentiated the INH susceptibility of susceptible and resistant (*katG*$^-$) *M. smegmatis* strains. The authors suggest that INH is converted by reaction with peroxides to a more active drug, and that the INH susceptibilities of enterobacteria and mycobacteria are mechanistically related (Rosner and Storz, 1994). Sherman *et al.* (1996) report that treatment of *M. bovis* BCG with sub-inhibitory concentrations of both hydrogen peroxide and INH resulted in a 100-fold reduction in cell viability after 72 hours, while the same concentrations of the two agents added separately had no effect, therefore they act synergistically. This agrees with the suggestion by Rosner and Storz (1994) that increasing the oxidative burden on the cells should improve the efficacy of INH. This effect was, however, not seen in a strain in which *katG* had been deleted.

To test whether the oxidative stress proteins, AhpC and OxyR, play a role in INH susceptibility in the mycobacteria, Deretic *et al.* (1995) studied the corresponding genes in *M. tuberculosis*. In *M. leprae*, the *ahpC* gene is linked to and divergently transcribed from the *oxyR* gene (Bergh and Cole, 1994). The authors isolated and sequenced the *ahpC* gene from *M. tuberculosis* and found it to have > 90% homology with that of *M. avium* and *M. leprae*, and to be very similar to other members of the *ahpC* family. The *oxyR-ahpC* intergenic region was also very similar to that from *M. leprae*, however, the *oxyR* gene of *M. tuberculosis* was shorter and carried multiple lesions, which prevent it from forming a single, contiguous translation product. Similar results were reported by Sherman and co-workers (1995) (discussed in Chapter 5). To test whether these lesions in *oxyR* affected INH susceptibility, the authors introduced two cosmids containing the *ahpC-oxyR* region of *M. leprae* into *M. tuberculosis* H37Rv, and found the transformants to be slightly more resistant to INH. They suggest that the activated products of INH may be detoxified by AhpC and possibly other OxyR-controlled functions, which may be absent or present at low levels in *M. tuberculosis* (Deretic *et al.*, 1995).

Several studies have confirmed an indirect association between INH susceptibility and the AhpC protein. Sherman *et al.* (1996), for example, found an increase in the expression of AhpC in INH-resistant, *katG*⁻ mutant *M. tuberculosis* strains, relative to the INH-sensitive *M. tuberculosis* H37Rv strain. However, selection for INH-resistant mutants revealed no upregulation of AhpC, even though *katG* had been mutated. The mutation resulting in increased AhpC expression, therefore, requires a second selection event, either during infection or by further passage in culture. Sequencing of the 5' regions of *ahpC* in *katG* mutants in which *ahpC* is upregulated revealed at least one mutation in the region from 39-81 bp upstream of the *ahpC* start codon in each isolate. Using the *lux* gene as a reporter, it was shown that these mutations positively affected promoter activity. Heym *et al.* (1997) recently reported results which confirm these observations. They detected 7 different mutations in the regulatory region of *ahpC* in 10 out of 77 INH-resistant clinical *M. tuberculosis* isolates (MIC 0.2 to > 10 μg/ml). All strains with complete *katG* deletions had altered *ahpC* regulatory regions. They also showed, using *lacZ* fusions and immunoblot analysis, that some of the mutations resulted in increased transcription of *ahpC*. Wilson and Collins (1996) identified a point mutation (C → T) at position 72 in the promoter region of the *ahpC* gene from a catalase-negative, INH-resistant mutant of *M. bovis*, as well as three other INH-resistant *M. tuberculosis* complex strains. This

mutation resulted in a 6-fold increase in expression. Dhandayuthapani *et al.* (1996) also identified mutations or polymorphisms in the 105 bp intergenic region between *ahpC* and *oxyR* in catalase-negative, INH-resistant *M. tuberculosis* strains. Zhang *et al.* (1996) mapped the transcription start site of the *M. bovis* BCG *ahpC* to 42 bp upstream of the translation start codon. This means that most of the identified mutations lie within the promoter region. These results show that INH-resistant strains mutant in *katG* frequently contain additional mutations which contribute to the overall level of sensitivity to INH. It may be that mutations which increase expression of AhpC compensate for the lack of catalase activity, thus allowing greater survival in the presence of reactive oxygen species (Sherman *et al.*, 1996). AhpC is involved in the conversion of potentially harmful hydroperoxide derivatives of lipids and nucleic acids into the corresponding alcohols. The enzyme may, therefore, be involved in the detoxification of the products of INH activation, and may reduce the consequences of increased oxidative damage to cellular molecules. In contrast, Sreevatsan *et al.* (1997) recently showed that most of the INH-resistant *M. tuberculosis* complex clinical isolates which they studied, which had reduced catalase and peroxidase activity, had no alterations in *ahpC* or the intergenic region. They were, however, able to detect 2 polymorphic sites in the *ahpC* gene and 9 in the intergenic region of some INH-resistant organisms. These authors suggest that the second selection event for the *ahpC* mutation had not occured as yet in their clinical isolates.

Despite the presence of upregulatory mutations in the *ahpC* of many INH-resistant isolates, there was no correlation between overexpression of either the wild type or mutated *ahpC* genes from *M. leprae, M. avium,* and *M. tuberculosis,* and resistance to INH in *M. tuberculosis* H37Rv, *M. bovis* BCG, or *M. smegmatis,* in the absence of a *katG* mutation (Wilson and Collins, 1996; Sherman *et al.*, 1996; Heym *et al.*, 1997). AhpC, therefore, probably does not play a direct role in the detoxification of INH. Zhang *et al.* (1996) showed that insertional inactivation of the *ahpC* gene in *M. smegmatis* resulted in increased susceptibility to INH. In all INH-sensitive strains of *M. tuberculosis* and *M. bovis*, the AhpC protein is below detectable levels, possibly due to the defects in OxyR. *M. aurum*, another strain which is naturally sensitive to INH, also has no detectable AhpC (Dhandayuthapani *et al.*, 1996). It may be these very low levels of AhpC which prevent the removal of excess reactive oxygen species, and result in the high sensitivity to INH. In *M. tuberculosis* complex organisms, the *oxyR* gene is naturally mutated resulting in overproduction of *katG*. Therefore, there is a large amount of

enzyme per cell converting INH to its activated form. In INH-resistant mycobacterial strains, such as *M. avium*, the same mechanisms of resistance to INH occur, however, there are lower basal levels of *katG*, and therefore less activation of INH. This suggests that the defects in the oxidative stress response are a major contributing factor to the exceptionally high sensitivity of *M. tuberculosis* to INH. Other putative members of the *oxyR* regulon may also be involved in INH-susceptibility and the effect may be cumulative (Zhang *et al.*, 1996).

1.3.2.8. Mechanisms of resistance to other anti-tuberculosis drugs

Mycobacteria show a high degree of primary resistance to most antibiotics and chemotherapeutic agents, probably due to their low cell wall permeability. Hydrophilic and lipophilic agents are slowed down by the abnormal thickness and unusually low fluidity of the lipid bilayer (Jarlier and Nikaido, 1994). A number of studies have been carried out on the mechanisms of resistance of *M. tuberculosis* and other mycobacteria to various antimycobacterial drugs, using the resistance mechanisms of *E. coli* as analogies. Streptomycin, an aminocyclitol glycoside antibiotic, causes misreading of the genetic code, inhibition of the initiation of mRNA translation, and aberrant proof-reading (Gale *et al.*, 1981; Moazed and Noller, 1987). Resistance to streptomycin can be caused by aminoglycoside-modifying enzymes and mutations in the ribosomal protein S12 gene, *rpsL* (Böttger, 1994). Most *M. tuberculosis* strains are highly susceptible to streptomycin, and only one molecule of streptomycin is required to inhibit one ribosome *in vitro*. Mutations in the *M. tuberculosis rpsL* gene (Kenney and Churchward, 1994; Finken *et al.*, 1993; Honoré and Cole, 1994), and in the 16S rDNA sequence (Douglass and Steyn, 1993; Meier *et al.*, 1994; Honoré and Cole, 1994), have been associated with resistance to this drug. In *M. smegmatis*, however, no mutations have been identified in the 16S rDNA, only in *rpsL* (Kenney and Churchward, 1994), and this may be due to the presence of two copies of the gene in this organism. The slow growing mycobacteria possess only one copy of the 16S rRNA gene, therefore single nucleotide changes can result in antibiotic resistance (Musser, 1995). Approximately 30% of resistant *M. tuberculosis* isolates studied had wild type *rpsL* and 16S rRNA genes, suggesting that other mechanisms of resistance occur, for example, alterations in cellular permeability, which may be responsible for low levels of resistance (Musser, 1995).

Rifampin (RIF) inhibits *E. coli* by binding to the β subunit of the RNA polymerase and preventing transcription (McClure and Cech, 1978). The same mechanism appears to function in *M. smegmatis* (Levin and Hatfull, 1993). Resistance to rifampin in *E. coli* is mediated by a number of different mutations and deletions in the *rpoB* gene, encoding the β subunit (Jin and Gross, 1988), and these have been used for studying the resistance mechanisms in mycobacteria (Honoré and Cole, 1993). Miller *et al.* (1994) cloned and sequenced the *M. tuberculosis* H37Rv *rpoB* gene, whose product showed 57% amino acid identity with that of *E. coli*. Introduction and expression of the gene in a highly rifampin-resistant strain of *M. smegmatis* LR223 (MIC > 200 μg/ml) resulted in an increased sensitivity to rifampin. Examination of the *rpoB* gene of rifampin-resistant strains of *M. tuberculosis* revealed mutations in an 81 bp region of the gene in > 90% of them (Heym *et al.*, 1994; Bodmer *et al.*, 1995; Kapur *et al.*, 1994; 1995). Similar mutations have been found in the *rpoB* gene of RIF-resistant isolates of *M. leprae* (Honoré and Cole, 1993), *M. africanum*, and *M. avium*. In the small proportion of strains with wild type *rpoB* genes, resistance may be due to changes in cell permeability or mutations in other RNA polymerase subunits.

The mechanism of action of pyrazinamide has not been elucidated, however it is known that susceptible strains possess a pyrazinamidase, which converts pyrazinamide into pyrazinoic acid and nicotinamide into nicotinic acid (Konno *et al.*, 1967). The activated drug is only active at a low pH (McDermott and Tompsett, 1954) or inside macrophages (Mackaness, 1956). Pyrazinamide-resistant strains of *M. tuberculosis* and strains of *M. bovis*, which are naturally resistant to pyrazinamide, lack this enzyme activity (Konno *et al.*, 1959; Konno *et al.*, 1967; Butler and Kilburn, 1983). The gene encoding pyrazinamidase, *pncA*, was recently cloned and characterized by Scorpio and Zhang (1996). The amino acid sequence is 35.5% identical to that of the *E. coli* nicotinamidase. The authors detected the gene in 3 *M. bovis* strains and 3 *M. bovis* BCG substrains, however, all had the same point mutation, C → G at nucleotide 169, resulting in a His-57 → Asp substitution, which presumably renders the enzyme inactive. A direct correlation was also observed between resistance to pyrazinamide and the presence of point mutations in the *pncA* gene of *M. tuberculosis* strains. In addition, transformation of *M. bovis* BCG and a pyrazinamide-resistant *M. tuberculosis* strain with a plasmid carrying the gene resulted in susceptibility to the drug (Scorpio and Zhang, 1996). Scorpio and co-workers (1997) analyzed a panel of pyrazinamide-resistant isolates of *M. tuberculosis*. They detected

mutations, including nucleotide substitutions, insertions, and small deletions, in 33/38 clinical isolates and 8 *in vitro* mutants. Those isolates with no detectable mutations were shown to be "falsely resistant" to pyrazinamide. Altogether, the authors identified 26 types of mutations dispersed along the gene, although some conservation of mutations was observed, for example, Asp-12 → Ala or Asn, Leu-85 → Pro, Gly-132 → Ser, and Thr-142 → Lys or Met. No mutations were observed in susceptible strains. These results suggest that mutations in *pncA* are indicative of resistance to pyrazinamide, and that these are the major mechanism of resistance to the drug (Scorpio *et al.*, 1997). Sreevatsan *et al.* (1997) report similar results. They detected mutations in the *pncA* gene of 72% of pyrazinamide-resistant *M. tuberculosis* isolates from diverse geographic localities, and no mutations in the susceptible isolates. These results suggest that at least one other gene is involved in resistance. As yet, no resistant strains have been isolated which have mutations in the target, i.e. strains which are resistant to pyrazinoic acid, which suggests that these types of mutations may be lethal (Scorpio *et al.*, 1997).

DNA gyrase is the main target of the fluoroquinolones in bacteria. The A and B subunits of the enzyme are encoded by the *gyrA* and *gyrB* genes, respectively. Point mutations have been identified in the *gyrA* genes of ofloxacin-resistant mutants of *M. smegmatis* (Revel *et al.*, 1994), and ciprofloxacin-resistant mutants of *M. tuberculosis* (Takiff *et al.*, 1994; Kapur *et al.*, 1995). Viomycin and capreomycin interfere with ribosome translocation (Inderlied, 1991). High level resistance to these antibiotics appears to be conferred by ribosomal mutations (Böttger, 1994). Ethionamide, a structural analogue of INH used as a second-line antituberculosis drug, also blocks mycolic acid synthesis in mycobacteria and, in some cases, low level resistance to INH involving mutations in *inhA* correlates with resistance to ETH (Canetti, 1965; Hok, 1964). Ethambutol inhibits the polymerization of cell wall arabinan, possibly by interfering with the transfer of arabinose to the cell wall acceptor. Resistance to this drug results from overexpression of the Emb proteins, encoded by *embCAB*, structural mutation of EmbB, or both (for references see Telenti *et al.*, 1997). Mycobacteria exhibit a range of sensitivities to β-lactam antibiotics and this is due differences in their β-lactamase activities. The *M. tuberculosis* complex organisms possess strong β-lactamase activity, the rapid growers possess strong to intermediate activity, and *M. avium* possesses no activity (Kwon *et al.*, 1995). Although a number of mutations have been identified in drug-resistant mycobacterial isolates, the exact

mechanisms of action of, and mechanisms of resistance to, the many of the antituberculosis drugs remain unclear. It is evident that both are very complex and, in most cases, more than one cellular component is involved. A common element in all the drug resistant mechanisms identified thus far in *M. tuberculosis* is that resistance is due to mutations in chromosomal genes rather than the acquisition of new genetic elements encoding antibiotic-altering enzymes. *M. tuberculosis* is an intracellular pathogen and, as such, is not exposed to the pool of drug determinants available in naturally occurring microbial communities. Multidrug-resistance appears to arise from the accumulation of sequential mutations which confer resistance to single agents (Musser, 1995).

1.4. PROJECT AIMS

Outlined above is a brief review of the history, control, and pathogenesis of tuberculosis, with particular reference to the *M. tuberculosis* catalase-peroxidase enzyme, and its role in both virulence and susceptibility to INH. The enzyme evidently plays an important role in these organisms, and has been studied by a large number of researchers. Our interest in the *M. tuberculosis* KatG was generated fortuitously during another study. As mentioned in the review, the virulence of the organism relies, to an extent, on the removal of reactive oxygen species by the radical scavenging enzymes. Those radicals which may escape degradation, as well as other toxic molecules within the macrophage can cause considerable damage to the DNA of the organism. The ability of the organism to repair damage to its DNA is, therefore, also expected to be an important factor involved in the survival of mycobacteria in macrophages. For this reason, our laboratory initiated an investigation of DNA repair enzymes and mechanisms in *M. tuberculosis*. A genomic library of *M. tuberculosis* H37Rv was screened for DNA repair genes by transformation of a DNA repair-deficient *E. coli* strain with the library and selection on the DNA damaging agent, ethyl methane sulfonate. One of the clones which conferred resistance to this agent contained a fragment of the *M. tuberculosis katG* gene. One of the aims of this thesis, therefore, was to determine whether resistance was indeed due to the *katG* gene, and to examine the role of this gene in DNA repair. The results reported suggest that the gene product increases the resistance of *E. coli* DNA repair mutants to short wavelength ultraviolet light irradiation, mitomycin C, and hydrogen peroxide. This observation,

further emphasizes the importance of the *M. tuberculosis katG* gene in the survival of the organism in the macrophage, and suggests that it may encode a novel activity.

Obtaining a better understanding of the *katG* gene and its expression within the cells, was the second aim of this study. This section of work deals with the characterization of the *M. tuberculosis katG* promoter and the manner in which its expression is regulated. The absence of a functional *oxyR* gene in this organism suggests that the regulation of *katG* may differ from that in other bacteria. Examination of expression from the promoter under different conditions revealed that it has the same growth phase-dependence as other catalases, but that its response to various stresses differs from that in other bacteria. It was demonstrated that the *katG* promoter requires an upstream activator sequence for optimal activity, and that this region may bind an activator protein. The regulation of the gene in this organism may, therefore be complex, and may require a number of cellular factors.

PART A

IDENTIFICATION OF A NOVEL ACTIVITY OF THE *M. TUBERCULOSIS* KATG IN REPAIR OF DNA DAMAGE

CHAPTER 2

INTRODUCTION

CONTENTS

CHAPTER 2

INTRODUCTION

The ability of *Mycobacterium tuberculosis* to repair damage to its DNA caused by reactive oxygen species, and other mutagenic agents within the macrophage, probably contributes significantly to its survival in this environment, and hence to its virulence. DNA repair has been extensively studied in a number of organisms, particularly *E. coli*. A wealth of information is available on the pathways and mechanisms involved in DNA repair in this host. The subject is extensively reviewed by Friedberg *et al.* (1995), and will not be covered in detail in this review. Only the pathways and mechanisms relevant to this study are outlined below.

2.1. DNA REPAIR IN *E. COLI*

2.1.1. DNA damage caused by mutagenic agents

Ionizing radiation can cause DNA damage either directly, if energy is transferred without intermediates to ionize or excite components of DNA, or indirectly, through the generation of partially reduced oxygen species. Both forms of damage occur, therefore agents which act by scavenging toxic products probably only provide protection from indirect processes. Short wavelength ultraviolet (UV) light also causes both direct and indirect damage to DNA. The most common structural defects resulting from irradiation are pyrimidine dimers, formed due to the direct absorption of photons by the bases in the DNA, resulting in covalent linkage of adjacent pyrimidines (Setlow, 1966a; 1966b; 1968). These pyrimidine dimers change the conformation of the DNA and prevent base pairing and replication. Dimerization is reversible, and, if *E. coli* is continually irradiated, a dynamic equilibrium is established in which the rates of dimer formation and reversal are equal (Setlow and Carrier, 1966). Intermolecular DNA cross-links are also a minor product of damage due to short wavelength UV irradiation (Love *et al.*, 1986). Mitomycin C causes covalent cross-linking of complementary DNA strands which prevents strand separation and causes inhibition of DNA replication and transcription (Iyer and Szybalski, 1963).

Hydrogen peroxide is lethal to many bacteria, including *E. coli*. Reactive oxygen species attack and damage DNA in a number of ways, for example: damage to sugar residues, which can cause fragmentation, base loss and strand breaks; modification of bases; and cleavage of DNA. The main source of damage caused by hydrogen peroxide is the hydroxyl radical, generated by the reaction of hydrogen peroxide with iron-complexed DNA (Imlay *et al.*, 1988; Imlay and Linn, 1988). The hydroxyl radical can also attack the double bond of thymine to form thymine glycol, which blocks replication, or 8-hydroxyguanine, which causes mispairing (Demple and Linn, 1982). Strand breaks and other lesions that block DNA replication are likely to contribute more to the lethal effects of hydrogen peroxide than base damage (Farr and Kogoma, 1991).

2.1.2. The SOS response and RecA

Central to the survival of *E. coli* cells after exposure to DNA damaging agents is the SOS response. The SOS response in *E. coli* involves the induction of more than 20 genes in response to DNA damage, including *recA*, *uvrA*, *uvrB*, *uvrD*, *polB*, and *umuDC*. The response is regulated by the LexA repressor and the RecA protein. In uninduced cells, LexA binds to operator sequences, or SOS boxes, upstream of the genes induced by the SOS response, including *recA* and *lexA*. When DNA replication is inhibited, for example due to single stranded DNA (ssDNA) regions, the RecA protein binds to these single stranded regions, in the presence of a nucleoside triphosphate, and is reversibly converted to an activated form. This then interacts with LexA and facilitates proteolytic cleavage of the latter protein, thus inactivating it. Transcription of the SOS genes is, therefore, induced. As the DNA damage is repaired and the levels of ssDNA decrease, the RecA molecules return to their inactivated state, resulting in an increase in the intracellular pools of LexA, and renewed repression of the SOS genes (Little and Mount, 1982; Walker, 1984; 1985).

In addition to its role in the SOS response, the RecA protein is involved in recombination, and DNA repair and damage tolerance processes. These include the repair of daughter strand gaps generated during the replication of a damaged DNA template; repair of double strand breaks; and trans-lesion replicative bypass (Friedberg *et al.*, 1995). The substrate for RecA-mediated strand exchange is ssDNA, bound by ssDNA binding proteins (SSBPs). These proteins remove hairpin formations from the ssDNA. RecA binds co-operatively to gapped duplex DNA. The

RecA-DNA complex then binds to the duplex DNA molecule, and RecA promotes unidirectional branch migration (Cox and Lehman, 1981).

2.1.3. The repair of UV-induced damage

As mentioned above, the most common structural defects resulting from UV irradiation are pyrimidine dimers. Cellular responses to UV damage involve either the direct reversal of DNA damage or its excision. An example of direct reversal of damage is a mechanism known as photoreactivation. Photoreactivation of pyrimidine dimers in DNA is a light-dependent process, which is catalyzed by DNA photolyase. The enzyme is widely distributed in nature, but has been best studied in *E. coli*. In the absence of light at wavelengths between 300 and 500 nm, DNA photolyase binds specifically to DNA containing pyrimidine dimers in the so-called dark reaction. The enzyme probably binds to the dimer-containing strand in duplex DNA, and makes close contacts with the phosphodiester backbone and the cyclobutane ring of the dimer (Husain *et al.*, 1987). The light reaction involves the absorption of photoreactivating light by chromophores associated with the enzyme. One of the chromophores, 5,10-methenyl tetrahydrofolyl (MTHF) polyglutamate, captures light and transfers energy to the second chromophore, 1,5-dihydroflavin adenine dinucleotide ($FADH_2$), which then initiates electron transfers that destabilize the pyrimidine dimers in the DNA and result in monomerization (Jorns *et al.*, 1984; Sancar and Sancar, 1987; 1984; Sancar, 1994). DNA photolyase bound to the damaged DNA also stimulates the Uvr-dependent nucleotide excision repair pathway discussed below (Sancar *et al.*, 1984).

There are three ways in which pyrimidine dimers are removed from DNA, base excision repair, nucleotide excision repair, and recombinational repair. Base excision repair involves the removal of chemically modified moieties from the deoxyribose-phosphate backbone, as free bases, by DNA glycosylases (Duncan *et al.*, 1976). DNA glycosylases have limited substrate specificity, and generally recognize specific forms of damage, such as pyrimidine dimers (Friedberg and King, 1969), thymine glycol (Breimer and Lindahl, 1984), and uracil (Duncan and Weiss, 1982). The removal of modified bases results in apurinic or apyrimidinic (AP) sites, which are then cleaved on either side by specific AP endonucleases. This is followed by gap repair synthesis and ligation. The repair of AP sites in *E. coli* DNA requires initial incisions in

the DNA catalyzed by specific AP endonucleases (Lindahl, 1979). The *xthA* gene product, exonuclease III (Yajko and Weiss, 1975), which is regulated by KatF (Sak *et al.*, 1989), represents approximately 90% of the AP endonuclease activity of *E. coli*. Bacteria which are mutant in *xthA*, however, are not noticeably sensitive to alkylating agents such as methyl methane sulfonate (Ljungquist *et al.*, 1976), suggesting that other AP endonuclease activities in the cell are sufficient to initiate DNA repair of these base-free sites (Milcarek and Weiss, 1972). The 5' terminal deoxyribose-phosphate residue resulting from the AP endonucleases is removed by exonucleases or DNA deoxyribophosphodiesterase (dRpase), to leave a single nucleotide gap in the DNA duplex (Kornberg and Baker, 1992). Gap repair synthesis requires the DNA polymerases, of which there are three in *E. coli*. DNA polymerase I, encoded by *polA*, contains the catalytic functions for DNA polymerization, pyrophosphorolysis exchange, 3'-5' exonucleolytic degradation, and 5'-3' exonucleolytic degradation. It can promote DNA replication at a nick, unaided by other proteins. DNA polymerase II, encoded by *polB*, has no 5'-3' exonuclease activity. Its optimal substrate is duplex DNA with short gaps of < 100 nucleotides, and is, therefore, ideal for repair synthesis of excision gaps generated during base or nucleotide excision repair. DNA polymerase III has 3'-5' and 5'-3' exonuclease activity, but requires a single stranded substrate (Kornberg and Baker, 1992). The DNA glycosylase/ AP endonuclease-mediated incision of DNA containing pyrimidine dimers is a highly specialized form of base excision repair observed only in *E. coli* cells infected with bacteriophage T4 (Friedberg *et al.*, 1995).

Nucleotide excision repair is the most common repair process, and involves the *uvrA*, *uvrB*, and *uvrC* products. The UvrABC complex recognizes and removes many different types of DNA damage by catalyzing excision of oligonucleotides containing the altered bases from the DNA duplex. Resultant gaps are then filled in by gap repair synthesis, and covalently sealed by DNA ligase. Nucleotide excision repair is biochemically more complex than base excision repair, and involves many more gene products. The *uvrA*, *uvrB*, and *uvrC* gene products are produced constitutively at approximately 25, 250, and 20 molecules per cell, respectively, in *E. coli*, however, the *uvrA* and *uvrB* gene products are induced by RecA in the SOS response (Selby and Sancar, 1990; Van Houten, 1990; Fogliano and Schendel, 1981). UvrA is a DNA-independent ATPase and a DNA-binding protein (Van Houten, 1990). In its dimeric form, UvrA binds to DNA containing various forms of damage (Selby and Sancar, 1990; Van Houten

and Snowden, 1993). These complexes are, however, short-lived, and dissociate rapidly, unless bound to UvrB. UvrB interacts specifically with UvrA to form protein-protein and protein-DNA complexes, which are important in damage-specific recognition. ATP hydrolysis is critical for the formation of the $(UvrA)_2(UvrB)_1$ complex (Sancar and Sancar, 1988). It has been suggested that the only role played by UvrA is in delivering the UvrB protein to the damage sites, after which it rapidly dissociates to leave a highly stable UvrB-DNA complex (Orren and Sancar, 1989). The binding of UvrB to DNA results in a conformational change in the protein and local unwinding of the DNA (Oh and Grossman, 1986). The working model is that two molecules of UvrA associate with one of UvrB to form a complex which binds non-specifically to DNA, but is able to unwind DNA using energy from ATP hydrolysis, and it translocates unidirectionally (Grossman and Thiagalingam, 1993). When base damage is encountered, the complex unwinds and kinks the DNA , UvrA dissociates, and a stable UvrB-DNA complex is formed (Sancar and Hearst, 1993; Van Houten, 1990). UvrC binds to specific UvrB-[damaged-DNA] complexes and allows incision of the DNA, mediated by both UvrB and UvrC (Lin and Sancar, 1992). Incisions are made on either side of the damaged DNA, typically 8 nucleotides upstream and 4 or 5 nucleotides downstream (Orren and Sancar, 1989). The *uvrD* gene product, DNA helicase II, is also induced by the SOS response (Kushner *et al.*, 1983). This protein, together with DNA polymerase I, is responsible for effecting the turnover of UvrB and UvrC (Caron *et al.*, 1985). DNA helicase II is required for the release of the oligonucleotide fragment and UvrC from the post-incision complex, while DNA polymerase I binds to the 3' OH terminus generated at the 5' incision, and displaces the UvrB protein during repair synthesis (Orren *et al.*, 1992). All three DNA polymerases of *E. coli* are able to perform repair synthesis of UV-damaged DNA (Masker *et al.*, 1973), however DNA polymerase I is probably the most important. It has been suggested that a number of other cellular functions may also be involved in nucleotide excision repair (Friedberg *et al.*, 1995).

Recombinational repair is more of a tolerance mechanism than a repair system. If altered bases have not been removed before the replication fork reaches them, the replication machinery may leave a gap and then continue synthesis after the defective base/s. The gap that is formed is then filled in by recombination with an intact daughter strand. The altered base is, therefore, not removed, but is eventually titrated out after successive rounds of replication. Recombinational repair requires the RecA protein as well as the RecBCD complex, and only occurs during

replication (Howard-Flanders *et al.*, 1984; West *et al.*, 1981). Therefore, during stationary phase when the cells are not replicating, the predominant form of repair will be excision repair. It has been proposed that the RecBCD enzyme enters into the DNA at a double stranded break, and then travels along the DNA, unwinding and rewinding it until it encounters a specific recombination, or chi site. It then cuts one strand of the DNA to form a single stranded tail, which invades the homologous intact duplex in a reaction mediated by RecA (Dixon and Kowalczykowski, 1991). Recombinational repair is responsible for the repair of single stranded gaps and single and double stranded DNA breaks. RecF binds ssDNA and is required for daughter strand gap repair (Walker, 1987), while RecBCD is required for sister duplex recombination (Wang and Smith, 1983). RecBCD is also important for the maintenance of cell viability during normal growth. It contains ATP-dependent exonuclease and helicase activity on double stranded DNA (dsDNA), and exonuclease and ATP-independent endonuclease activity on ssDNA. RecE is an exonuclease which uses double stranded breaks to produce gaps with a single stranded overhang, which then undergo homologous recombination with an undamaged sister strand. Apart from its role in recombinational repair, RecA is essential for the survival of cells exposed to UV irradiation due to its role in induction of the SOS response.

2.1.4. The repair of damage caused by cross-linking

Mitomycin C causes covalent cross-linking of complementary DNA strands which results in inhibition of DNA replication and transcription (Iyer and Szybalski, 1963). These DNA lesions are removed in a similar manner to those caused by UV, i.e. excision repair, except that both strands must be removed, leaving a gap in the DNA which must be repaired by recombination, involving RecA. An incisional-recombinational model has been proposed for the repair of this type of damage. In this model, the UvrABC complex effects incision of one strand, generating an oligonucleotide covalently linked to DNA through the cross-link. Displacement of the oligonucleotide generates a gap which can be repaired by recombinational repair. Subsequent incision of the other strand results in release of the cross-linked oligonucleotide, leaving a second gap, which can be repaired by repair synthesis and DNA ligation (Cole, 1973). Repair of cross-links in DNA has been shown to occur in *uvr* mutants of *E. coli* (Bridges, 1983; 1984; Bridges and Stannard, 1982), suggesting the presence of Uvr protein-independent pathways for

the repair of cross-links in DNA. In addition, a novel enzyme which exclusively incizes DNA containing cross-links, has been described by Sladek *et al.* (1988)

2.1.5. The repair of oxidative damage

The exact mechanisms involved in repair of oxidative damage to DNA are not known, but may involve both specific and non-specific enzyme systems. Bacteria have two main mechanisms of protection against oxidative stress: 1) enzymes which scavenge activated oxygen species, such as catalases (Loewen, 1984), peroxidases (Saunders *et al.*, 1964), and superoxide dismutase (Carlioz and Touati, 1986); and 2) enzymes which repair damage to DNA resulting from oxidative stress, such as exonuclease III (Demple *et al.*, 1983), thymine glycol DNA glycosylase (Cunningham and Weiss, 1985), DNA polymerase I (Ananthaswamy and Eisenstark, 1977), DNA polymerase III (Hagensee and Moses, 1989), RecBC nuclease (Imlay and Linn, 1987), and RecA (Yoakum and Eisenstark, 1972). When bacteria are pretreated with low doses of hydrogen peroxide, they become resistant to subsequent doses that would otherwise be lethal, due to the induction of at least 30 proteins, some of which include DNA repair enzymes (Demple and Halbrook, 1983). Among these, 9, including catalase, alkyl hydroperoxide reductase, and OxyR, are regulated by OxyR (Loewen, 1984; Morgan *et al.*, 1986). The oxidative stress response is discussed in more detail in Chapter 5. Treatment with low concentrations of hydrogen peroxide also results in the induction of *recA* (Greenberg and Demple, 1989), and Goerlich *et al.* (1989) provide direct evidence that, in *E. coli*, hydrogen peroxide induces the SOS response. They demonstrated this using a *lacZ* operon fusion to *sfiA*, one of the SOS response genes. The authors concluded that hydrogen peroxide makes DNA lesions, possibly strand breaks, which cause SOS induction, and that some of these can be repaired by the *uvr*ABC-dependent excision repair pathway. *E. coli* cells pretreated with hydrogen peroxide also exhibit an increased resistance to ionizing radiation, such as gamma rays, but not to alkylating agents or UV irradiation (Demple and Halbrook, 1983). The superoxide stress response also includes an increased DNA repair capacity (Farr *et al.*, 1985).

It has been shown by Carlsson and Carpenter (1980) and Ananthaswamy and Eisenstark (1977), that the radical-removing enzymes, such as the catalases, peroxidases, and superoxide dismutase, play a surprisingly small role in protection of *E. coli* from killing by hydrogen

peroxide. Carlsson and Carpenter (1980) examined the hydrogen peroxide sensitivities of a variety of DNA repair-deficient strains of *E. coli* K12. They found no correlation between the death rate of the strains and the ability to decompose hydrogen peroxide by intact cells, or the level of catalase and superoxide dismutase in cell-free extracts. The actual involvement of catalases in this protection is unclear, due to a number of conflicting observations. It has been shown that, at relatively high cell densities, wild type cells are less sensitive to hydrogen peroxide than HPI mutants, while at low cell densities, wild type and mutant strains are equally sensitive. This is, presumably, due to the efficient removal of hydrogen peroxide through the mass action of the cells (Ma and Eaton, 1992). Others have demonstrated that catalase mutants are hypersensitive to killing by hydrogen peroxide (Demple, 1991; Farr and Kogoma, 1991), but that, in wild type strains, there is no correlation between the levels of catalase and hydrogen peroxide killing (Hassett *et al.*, 1990). The role of catalases in protection against hydrogen peroxide toxicity also differs depending on the hydrogen peroxide concentration. Catalase mutants are hypersensitive to millimolar concentrations (Loewen, 1984), however, in experiments with micromolar concentrations, the catalases provided only weak protection (Goerlich *et al.*, 1989). Thus survival of cells exposed to hydrogen peroxide must rely on repair of the damage after it has occured. Hagensee and Moses (1989) have shown that there are at least two pathways for repair of hydrogen peroxide damage, both of which require DNA polymerase I. One of these pathways also utilizes exonuclease III and DNA polymerase III. Demple *et al.* (1983) have shown that *E. coli* cells mutant in *xthA*, encoding exonuclease III, are 10-20-fold more sensitive to hydrogen peroxide than wild type cells, and are more sensitive than cells mutant in *recA*. This suggests that these enzymes play an important role in repair of oxidative damage. Demple *et al.* (1983) hypothesize that the free radicals formed during the decomposition of hydrogen peroxide may form nicks containing 3'-phosphate groups in the DNA, which in turn block DNA polymerase. Since > 99% of the 3'-phosphatase activity in *E. coli* is due to exonuclease III (Milcarek and Weiss, 1972), this enzyme may be vital in removing these groups to allow DNA replication to continue. Another possibility is that hydrogen peroxide causes base damages which are cleaved by endonuclease III (Demple and Linn, 1982). This enzyme generates 3' termini that are inefficiently used by DNA polymerase I, thus requiring exonuclease III for the initiation of repair synthesis (Demple *et al.*, 1983).

A major form of damage caused by hydrogen peroxide is base modification, especially thymine glycol (Demple and Linn, 1980). N-glycosylases for thymine glycol and urea are known to act specifically on oxidized DNA (Demple and Halbrook, 1983). However, a mutation in *nth*, encoding thymine glycol DNA glycosylase, had no effect on the induction of the SOS response by hydrogen peroxide (Goerlich *et al.*, 1989), and Cunningham and Weiss (1985) have observed that an *nth* mutant is not hypersensitive to killing by hydrogen peroxide. Excision repair mutants are also far more resistant to hydrogen peroxide than *xthA* mutants (Demple *et al.*, 1983), which suggests that excision repair is not essential either for repair of oxidative damage. The UvrABC complex is capable of removing thymine glycols and AP sites from oxidatively damaged DNA (Lin and Sancar, 1989), however, and probably does participate, to an extent, in repair of oxidative damage to DNA.

2.1.6. Mutagenesis

Some lesions caused by DNA damaging agents escape repair, and can lead to the introduction of mutations. Most mutagenesis is induced due to trans-lesion DNA synthesis. Mutagenesis by UV and many chemicals is not a passive process in *E. coli*, and is brought about by error-prone repair or SOS repair (Radman, 1974; Witkin, 1976). In the trans-lesion synthesis model, a polymerase inserts one or more nucleotides directly opposite the lesion and extends the chain (Bridges and Woodgate, 1985). The process probably involves DNA polymerase III, and requires the *umuD* and *umuC* gene products.

2.2. DNA REPAIR IN THE MYCOBACTERIA

As mentioned in Chapter 1, *M. tuberculosis* is exposed to hydrogen peroxide as part of the macrophage defence mechanism (Lowrie, 1983). Jackett *et al.* (1978) have shown an association between hydrogen peroxide-susceptibility and low virulence in native and INH-resistant tubercle bacilli, but that this is not the sole determinant of virulence. They found that resistance to hydrogen peroxide depends on catalase, but also on other unknown features. Virulent *M. tuberculosis*, therefore, seem to possess two independent means of resistance to hydrogen peroxide, but can become attenuated without losing this resistance (Jackett *et al.*,

1978). The cell wall probably plays an important role in resistance. The gene encoding the inducer of the oxidative stress response, *oxyR*, of *M. tuberculosis* and other *M. tuberculosis* complex members is naturally mutated (Sherman *et al.*, 1995; Deretic *et al.*, 1995; Dhandayathupani *et al.*, 1996) (Discussed in Chapter 5). This suggests that this organism may use alternative pathways for survival in the presence of reactive oxygen species, and possibly other DNA damaging agents.

2.2.1. Sensitivity of mycobacteria to UV irradiation

Different mycobacterial species exhibit different sensitivities to UV light irradiation (254 nm). David (1973) investigated the properties of mycobacteria which give rise to their sensitivity to UV light irradiation. All the mycobacteria tested showed greater survival than *E. coli* grown in the same medium, and the order of UV sensitivity was as follows: *M. tuberculosis* > *M. fortuitum* > *M. avium-intracellulare* > *M. phlei* > *M. marinum* > *M. kansasii* > *M. smegmatis* > *M. flavescens*. The UV sensitivity depends partly on the G+C content, but this value has a narrow range in these mycobacteria (63-70%), therefore the differences in sensitivity could not be attributed to the base composition of their DNA. It is also evident that the UV sensitivities did not correlate with generation time or clumping of the cells. However, the UV sensitivities of six of the mycobacterial species did correlate with the sizes of their genomes. This suggests that the smaller the DNA molecule, the higher the probability that a lesion will be lethal. The mycobacteriophage Minetti reproduces lytically in both *M. smegmatis* and *M. fortuitum* and was used by the author as a means of determining the repair efficiency of each host. *M. smegmatis*, which is the less UV sensitive of the two, was shown to confer higher survival to the mycobacteriophage than *M. fortuitum*, suggesting that the differences in UV sensitivity were directly due to differences in their repair efficiencies. Unfortunately, no phages have been identified that reproduce lytically in every species of mycobacteria. Different mycobacterial species produce different amounts of pigment (carotenoids) under different conditions. In some, for example *M. marinum* and *M. kansasii*, the synthesis of carotenoids is light inducible (Batra, 1967), while in others, for example *M. flavescens*, pigments are synthesized constitutively. David (1973) demonstrated that the production of pigments masks the UV sensitivity in mycobacteria. This may be due to the absorption of radiation by intermediates in the production pathway (David, 1973). The author demonstrated

photoreactivation in *M. avium-intracellulare*, and both photoreactivation and dark repair in *M. kansasii* to the same degree as that reported for other mycobacteria (David *et al.*, 1971; Sellers *et al.*, 1970).

2.2.2. Sensitivity of mycobacteria to hydrogen peroxide

Due to the lack of a functional OxyR, *M. tuberculosis* probably has no inducible oxidative stress response, and, therefore, relies on the constitutive defence offered by its inert and largely impermeable cell wall. Only KatG is peroxide-inducible, but this induction does not protect against hydrogen peroxide challenge. Sherman *et al.* (1995; 1996) demonstrated that sub-lethal hydrogen peroxide concentrations did not protect *M. tuberculosis* against subsequent lethal hydrogen peroxide doses, as in *E. coli*, *S typhimurium* and *M. smegmatis* (Christman *et al.*, 1985; Rosner and Storz, 1994; Sherman *et al.*, 1995), even though the expression of *katG* was induced. Induction of *katG* did, however, allow a 35-fold increase in survival of *M. bovis* BCG exposed to cumene hydroperoxide. A *katG*-deleted strain showed a much higher sensitivity to oxidative stress, and no protection against organic peroxides could be induced by pre-exposure to sub-inhibitory hydrogen peroxide concentrations. Overexpression of the *ahpC* gene from *M. tuberculosis* Erdman in strain H37Rv resulted in a significant increase in resistance to cumene peroxide. Similar results were obtained for BCG. Increased expression of AhpC therefore appears to compensate for the loss of *katG* in detoxifying the organic peroxides (Sherman *et al.*, 1996). These experiments do not take into consideration the possible effects of DNA repair systems in the recovery from oxidative damage, and little is known about this subject.

2.2.3. The *M. tuberculosis* RecA

The *M. tuberculosis recA* gene has been cloned and sequenced (Davis *et al.*, 1991). It encodes a protein with high homology to the *E. coli* RecA over the first 254 and last 96 amino acids, and the protein cross-reacts with antibodies to the *E. coli* RecA. The *M. tuberculosis* gene, however, also encodes a centrally located 440 amino acid non-homologous region, which represents an intein. This intein must be spliced out after translation to produce the mature protein (Davis *et al.*, 1991). It is only in the *M. tuberculosis* complex organisms, and in

M. leprae that unusual *recA* genes exist (Davis *et al.*, 1994). The inteins from *M. tuberculosis* and *M. leprae* are, however, different in size, sequence and location in the coding sequence. Therefore acquisition of the inteins must have occurred independently in the two species. Only the spliced form of RecA is active, and is able to bind ssDNA in the presence of ATP, promote pairing between ssDNA and homologous dsDNA, and possesses ATPase activity (Kumar *et al.*, 1996). This may provide a novel means of regulating RecA activity. In *M. tuberculosis* complex organisms, integration of plasmids into the chromosome by illegitimate recombination occurs at a frequency of $1x10^{-3}$, therefore illegitimate recombination is a frequent event in these organisms. This, together with the presence of a specific protein splicing mechanism, suggests that the *M. tuberculosis* RecA possesses some unique features, and it is possible that these may play a significant role in the virulence of the organism. Protein splicing has also been demonstrated for the *gyrA* genes of *M. leprae*, *M. flavescens*, *M. gordonae*, and *M. kansasii* (Fsihi *et al.*, 1996).

Davis *et al.* (1991) cloned the *M. tuberculosis recA* gene under control of the *lacZ* promoter, and were able to demonstrate complementation of an *E. coli recA* mutant with respect to recombinational activity, and partial complementation of the repair and regulatory responses required for recovery from DNA damage. The *M. tuberculosis recA* clone could restore mutability to a *recA*13 mutant, but did not complement for cleavage of the λ repressor. The protein is defective in displacing SSB from ssDNA, and has a reduced affinity for ATP and a reduced ability to hydrolyze it, which may explain the partial complementation of the *E. coli recA* mutants. It shows significant differences in its requirements for binding to DNA, and in the formation of joint molecules. Studies with *sfiA-lacZ* fusions have shown that it can bring about LexA cleavage and SOS induction.

The *M. tuberculosis recA* promoter region contains no *E. coli*-like LexA-binding site or promoter, but it does have a sequence resembling heat shock promoters 100 bp upstream of the initiation codon, and a sequence identical to a consensus $GAACN_4$-GTTC sequence from various DNA damage-inducible promoters in *B. subtilis* (Cheo *et al.*, 1991). The promoter region and its interaction with the *M. tuberculosis* LexA homologue have been examined further. Movahedzadeh *et al.* (1997) examined the expression from and induction of the promoter in *M. smegmatis*, using a fusion to the chloramphenicol acetyl transferase reporter

gene. They demonstrated induction of the promoter in response to ofloxacin, mitomycin, and UV irradiation. The purified *M. tuberculosis* LexA protein bound to the Cheo box, 121 bp upstream of the translation initiation codon. Induction of the promoter occurred to the same extent in *M. tuberculosis*, but was delayed in this organism. In the *M. tuberculosis* LexA homologue, amino acids required for autocatalytic cleavage are conserved, while those required for specific DNA binding are not. The region upstream of the *lexA* gene also contains a Cheo box which binds to the LexA (Movahedzadeh *et al.*, 1997).

2.3. AIMS OF THIS STUDY

It is evident that some elements of the SOS response of *M. tuberculosis* are similar to those in other bacteria. However, other important elements of DNA repair in this organism have yet to be examined. This, together with the possible importance of DNA repair pathways in the virulence of the organism, prompted a search for DNA repair genes from *M. tuberculosis*. Shamila Nair (PhD Thesis, University of Cape Town, 1993) initiated this investigation by constructing a *Pst*I library of *M. tuberculosis* H37Rv genomic DNA in the suicide vector pEcoR252, and selecting for clones which conferred resistance to ethyl methane sulfonate on *recA*-deficient *E. coli* cells. One of the clones which consistently conferred resistance to this agent was found to contain a portion of the *M. tuberculosis katG* gene. The aim of this study was, therefore, to study the effect of the *M. tuberculosis katG* gene on the survival of DNA repair-deficient mutants of *E. coli* exposed to various DNA damaging agents, and to extrapolate these observations to the mycobacteria.

CHAPTER 3

CLONING OF THE *M. TUBERCULOSIS KATG* GENE AND OVEREXPRESSION IN *E.COLI*

CONTENTS Pg.

CHAPTER 3

CLONING OF THE *M. TUBERCULOSIS KATG* GENE AND OVEREXPRESSION IN *E.COLI*

3.1. SUMMARY

This chapter describes the sequencing and analysis of a clone that conferred resistance to ethyl methane sulfonate on DNA repair-deficient *E. coli* cells. Also described is the PCR amplification of the *M. tuberculosis* H37Rv *katG* gene and cloning of the whole gene, the proximal two thirds, and the terminal third of the gene in the expression vector pMAL™-c2. The resulting fusion proteins were overexpressed in, and purified from *E. coli*, and used in catalase and peroxidase assays. Only the fusion protein containing the whole KatG enzyme exhibited enzyme activity. Constructs containing the terminal third of the *katG* gene in pMAL™-c2 in the two other reading frames, with respect to the *malE* gene, were made for use as controls in the DNA repair experiments described in Chapter 4.

3.2. INTRODUCTION

The role played by DNA repair processes in the survival of *M. tuberculosis* in the macrophage is outlined in the introduction presented in Chapter 2. Little is known about DNA repair in the mycobacteria. The *recA* genes from a number of mycobacteria have been cloned and sequenced (Davis *et al.*, 1991). These genes have an unusual structure in the pathogenic mycobacteria, *M. leprae* and *M. tuberculosis*, in that they encode an intein which must be spliced out after translation to produce the mature, functional protein (Davis *et al.*, 1991; 1994). Another important determinant involved in the survival of cells exposed to agents which cause oxidative damage is the *oxyR* gene. In *M. tuberculosis*, this gene is naturally mutated (Sherman *et al.*, 1995; Deretic *et al.*, 1995; Dhandayathupani *et al.*, 1996). This, together with the unusual nature of the *recA* gene, suggests that the pathways and mechanisms of DNA repair in this organism may differ from those in other, well-characterized bacteria. These unusual features may also contribute to the virulence of *M. tuberculosis*.

In an attempt to improve our understanding of the mechanisms of DNA repair in the mycobacteria, a search for DNA repair genes in *M. tuberculosis* was initiated by Shamila Nair (PhD thesis, University of Cape Town, 1993). In this study, a *Pst*I genomic library of *M. tuberculosis* H37Rv in the suicide vector pEcoR252, was screened for DNA repair genes by selection of *E. coli* DK1 (*recA*) transformants on 0.1% (w/v) ethyl methane sulfonate. Two clones were isolated that were able to confer resistance to this agent, one of which was described earlier (Nair and Steyn, 1991). The other, designated pD3, is the subject of this study. pD3 contained an insert of 1747 bp (GenBank Accession number L14268), which included the terminal third of the *M. tuberculosis katG* gene. The aim of this study, therefore, was to determine whether the increased survival of the DNA repair-deficient *E. coli* cells observed was due to expression of the 3' portion of the *M. tuberculosis katG* gene. The strategy used for this was overexpression of different portions of the *M. tuberculosis katG* gene as fusions with the maltose binding protein (MBP) in the expression vector, pMAL^{TM}-c2 (New England Biolabs), in different DNA repair-deficient *E. coli* hosts. In pMAL^{TM}-c2, the *malE* gene, encoding the MBP (Guan *et al.*, 1987; Maina *et al.*, 1988), is under control of the *tac* promoter (Amann and Brosius, 1985), which is repressed under normal conditions by the repressor encoded by the *lacI*^{q} gene, which is also present on the vector. High levels of expression from the promoter can be obtained by induction of cells harbouring the vector with an analogue of lactose, IPTG, resulting in a fusion that constitutes 20-40% of the total cellular protein. The advantage of the pMAL^{TM}-c2 expression vector is that the resultant fusion protein can be purified by affinity chromatography, using an amylose resin (Kellerman and Ferenci, 1982), and the protein of interest can be cleaved off from the MBP with the protease, factor Xa (Nagai and Thogersen, 1984; 1987), whose cleavage recognition sequence lies immediately upstream of the multiple cloning site of the vector. This chapter describes the analysis of the *M. tuberculosis katG* sequence, the construction of the various translational fusions, and the overexpression, purification, and assaying of the fusions in *E. coli*.

3.3. MATERIALS AND METHODS

The bacterial strains, plasmids, and culture conditions used are described in Appendix A. Details of the commonly used techniques and solutions are given in Appendices B and C, respectively.

3.3.1. Isolation of pD3 and cloning into pUC19

The isolation and sequencing of the clone, pD3, has been described elsewhere (Shamila Nair, PhD thesis, University of Cape Town, 1993). Briefly, a *Pst*I genomic library of *M. tuberculosis* H37Rv was constructed in the suicide vector pEcoR252, a derivative of pEcoR251 (Zappe *et al.*, 1986). The library was used to transform *E. coli* DK1 (*recA*) cells, and transformants were selected on 0.1% (w/v) ethyl methane sulfonate. One of the two clones that conferred resistance to this agent on the DNA repair-deficient *E. coli* cells, pD3, was analyzed further. The 1747 bp insert of pD3 was cloned into the *Pst*I restriction endonuclease site of pUC19, to form pSNS100. The insert of this construct was sequenced using the pUC19 forward and reverse sequencing primers (Promega), and the T7 Sequencing kit (Pharmacia) (GenBank Accession number L14268).

3.3.2. Sequence determination and analysis

In this study, the sequence of the 1747 bp insert of pD3 and pSNS100 was analyzed using the computer packages described in Appendix B. Sequencing of various regions of the insert in pSNS100 was performed using the pUC19 forward and reverse sequencing primers (Promega), a reverse primer designed to bind to sequences 251 bp downstream of the *Pst*I restriction endonuclease site of the insert, and the T7 Sequencing kit (Pharmacia). The 18-mer *katG* sequencing primer had the sequence: 5'-GCTCCAGCACGGCAAAGG-3'. Band compressions in the sequencing ladder were resolved using the Sequenase™ Version 2.1 DNA Sequencing kit (USB, Amersham Life Science) dITP reagents, and a protocol modified for GC-rich templates (see Appendix B).

3.3.3. Cloning of the 1747 bp insert in pMAL™-c2

The 1747 bp *Pst*I fragment from pD3 was cloned into the *Pst*I restriction endonuclease site of pMAL™-c2 (New England Biolabs, Inc., Beverley U.S.A.), to form pSMK3'. Use of this restriction endonuclease results in an in-frame fusion of the terminal third of the *M. tuberculosis* *katG* gene with the MBP. Cloning of an insert in the multiple cloning site of pMAL™-c2 results in inactivation of the β-galactosidase α-fragment activity of the MBP-LacZ fusion. Therefore,

recombinants form white colonies on X-GAL plates in an α-complementing *E. coli* host (Yanisch-Perron *et al.*, 1985). During cloning, a construct was isolated which contained the 1747 bp insert in the opposite orientation to that in pSMK3'. This construct was designated pSMK3'$_{opp}$. The reading frame of the *M. tuberculosis katG* gene, with respect to the *malE* gene, was altered using the following strategy: pSMK3' was digested to completion with *Bam*HI, which restricts the construct upstream of the insert. The resulting cohesive ends were removed by digestion with Mung Bean nuclease (Boehringer Mannheim) (see Appendix B), and the blunt-ended, linear fragment was religated to form the construct pSMK3'a, in which the *katG* sequences are translated in frame 2. The construct pSMK3'b was made by digestion of pSMK3' with *Sal*I, which digests the construct upstream of the insert, filling in of the resulting cohesive ends using T4 DNA polymerase (Amersham DNA Blunting kit) (see Appendix B), and religation of the blunt-ended, linear fragment. In this construct, the *katG* sequences are translated in frame 3. Another construct, pSMKd, was made by digestion of the 1747 bp fragment with *Alu*I and *Pst*I (see Fig. 3.3.b), to generate a 961 bp fragment, containing the sequences downstream of the *M. tuberculosis katG* gene. The fragment was cloned into the *Xmn*I and *Pst*I restriction endonuclease sites of pMAL™-c2 to yield pSMKd. The orientation and frame of the inserts of these constructs were confirmed by restriction analysis and sequencing over the junctions using the *malE* primer (New England Biolabs), which initiates sequence near the 3' end of the *malE* gene, and the pUC19 forward primer (Promega).

3.3.4. PCR amplification of the *M. tuberculosis katG* and cloning into pMAL™-c2

Oligonucleotides were designed for the amplification of the whole *M. tuberculosis katG* gene. The primers were based on the published sequence (GenBank/EMBL Accession number X68081), and were designed to bind to the beginning and end of the gene. Mismatches were introduced near the 5' end of each primer to generate convenient restriction endonuclease recognition sites. The forward primer was a 24-mer with the sequence: 5'-AATGaatTcCCCGAGCAACACCCA-3', and the reverse primer was a 24-mer with the sequence: 5'-GGCAGtctaGATCAACCCGAATCA-3'. The lower case nucleotides indicate the mismatches, and the underlined nucleotides represent the restriction endonuclease sites for *Eco*RI and *Xba*I, respectively. The modifications in the forward primer resulted in the replacement of the GTG initiation codon of the *katG* gene with TTC, which encodes

phenylalanine. The complement of the TGA termination codon has been included at the 3' end of the reverse primer. The primers were designed such that cloning of the resulting PCR product into the *Eco*RI and *Xba*I restriction endonuclease sites of pMAL™-c2 would result in an in-frame fusion of the *M. tuberculosis* KatG with the MBP. The primers were predicted to amplify a 2250 bp fragment. PCR amplification was performed as described in Appendix B, using *M. tuberculosis* H37Rv genomic DNA (0.1 µg) as the template, and an annealing temperature of 58°C. Control reactions included those with no template DNA or primers, or those with single primers only, to ensure that the PCR products were not the result of non-specific binding of primers, or due to contaminating DNA. PCR products were separated by agarose gel electrophoresis, stained with ethidium bromide, and viewed under 310 nm UV transillumination. The 2250 bp PCR product was excised from an agarose gel, purified by the phenol squeeze method described in Appendix B, and digested to completion with the restriction endonucleases *Eco*RI and *Xba*I. The resulting fragment was cloned into the corresponding restriction endonuclease sites of pMAL™-c2, to yield pSMKA. The 3' end of the *katG* gene was deleted from pSMKA by digestion of the construct with *Pst*I to release a 756 bp fragment. The restriction endonuclease, *Pst*I, digests the construct at a recognition site within the *M. tuberculosis katG* gene (see Fig. 3.7), and one within the multiple cloning site of pMAL™-c2, on the 3' end of the insert. The 8132 bp linear fragment, containing pMAL™-c2 and the 5' two thirds of the *katG* gene, was religated to form the construct pSMK5'. The 756 bp *Pst*I fragment described above was cloned into the corresponding site of pMAL™-c2, with the *katG* sequences in-frame with the MBP, to form pSMK3'*. The sequence and frame of the inserts were verified by sequencing with the pUC19 forward (Promega), the *katG* (see section 3.3.2.), and the *malE* (New England Biolabs) sequencing primers, and the *katG* forward and reverse PCR primers described above.

3.3.5. SDS-PAGE and immunoblot analysis of proteins

Stationary phase cultures of *E. coli* UM255 (*katG*, *katE*) transformants harbouring pSMKA, pSMK5', pSMK3', and pMAL™-c2 were diluted 100-fold into 5ml of fresh 2x YT medium, containing ampicillin, and incubated at 37°C with shaking. When the cultures reached an OD_{600nm} of 0.4, a 1 ml sample was removed, and the cells harvested by centrifugation. The resulting pellet was resuspended in 50 µl of SDS-PAGE sample buffer (see Appendix C), and

represents the uninduced control. IPTG (Melford Laboratories, England) was added to the remainder of the culture, to a final concentration of 0.3 mM, and the cells were incubated for a further two hours at 37°C. The cells (1 ml) were harvested by centrifugation, and the cell pellet resuspended in 100 µl of SDS-PAGE sample buffer. All samples were heated to boiling for 5 minutes, loaded on a 12.5% (w/v) SDS-polyacrylamide gel, and separated by electrophoresis at a constant current of 40 mA (see Appendix B). Proteins were visualized by staining with Coomassie® Brilliant Blue R-250 (Bio-Rad Laboratories). For immunoblots, the proteins in unstained gels were transferred onto an Amersham C-super nitrocellulose membrane (see Appendix B). Immunoblots were performed using the Chemiluminescence Blotting Substrate (POD) (Boehringer Mannheim), according to the manufacturers' instructions. Blots were probed with the anti-MBP rabbit serum supplied by New England Biolabs, and anti-rabbit IgG conjugated to peroxidase (POD) (Boehringer Mannheim).

3.3.6. *In vitro* expression analysis of proteins encoded by the pMAL™-c2 constructs

The synthesis of proteins from the inserts of the pMAL™-c2 constructs was investigated using the *E. coli* S30 Extract System for Circular DNA (Promega, Madison U.S.A.). For the analysis, 1 µg of DNA was incubated for 2 hours at 37°C with 1 µl of Amino Acid Mixture Minus Methionine, 4 µl of Premix Without Amino Acids, 0.5 µl of ^{35}S-methionine (Amersham, UK), and 3 µl of S30 extract, in a 10 µl reaction. The resulting radiolabelled proteins were precipitated with 40 µl of acetone, collected by centrifugation, and resuspended in 20 µl of SDS-PAGE sample buffer. The samples were boiled for 5 minutes, loaded on a 17.5% SDS-polyacrylamide gel, and separated by electrophoresis at a constant current of 40 mA. The gel was stained with Coomassie® Brilliant Blue R-250 (Bio-Rad Laboratories) for 30 minutes, and destained overnight in destaining solution, in order to visualize the molecular weight markers (see Appendix B). The gel was dried and the ^{35}S-labelled bands visualized by autoradiography.

3.3.7. Purification of the fusion proteins by affinity chromatography

Overnight cultures of *E. coli* UM255 transformants containing the recombinant plasmids, pSMKA, pSMK5', and pSMK3', were diluted 100-fold into fresh 2x YT medium (500 ml final volume) containing ampicillin. The cultures were grown to an OD_{600nm} of 0.5 (approximately 2×10^8 cells/ml). One millilitre of culture was centrifuged to collect the cells, as an uninduced control. The pellet was resuspended in 50 µl of SDS-PAGE sample buffer. IPTG was added to the remainder of the culture to a final concentration of 0.3 mM, to induce expression of the fusion proteins from the *tac* promoter, and the culture was incubated at 37°C for a further 3 hours, with shaking. The cells were harvested by centrifugation, and 1 ml of supernatant was retained as a control. The pellet was resuspended in 30 ml column buffer (see Appendix C), and frozen at -20°C overnight. The cells were lyzed by passing the cell extract twice through an Aminco® French Pressure Cell (SLM Instruments, Inc., IL, U.S.A.), and the cell debris removed by centrifugation. An aliquot of the supernatant and pellet were retained for SDS-PAGE analysis. The fusion protein was purified from the whole cell extract by affinity chromatography, as recommended in the New England Biolabs Protein Fusion and Purification manual. The amylose resin (New England Biolabs) was poured into a 2.5 x 10 cm column, and washed with 8 bed volumes of column buffer. The crude extract was diluted to a protein concentration of 2.5 mg/ml and loaded onto the column. The flow rate of the column was adjusted to approximately 1 ml/minute. The column was washed with 10 volumes of column buffer, and the fusion protein eluted with column buffer containing 10 mM maltose (BDH Chemicals Ltd., Poole England). Twenty 3 ml fractions were collected and the optical density at 280 nm measured. Fractions with the highest OD_{280nm} values were pooled, and the protein concentrated using an Amicon Centriprep-10 (Amicon Inc., Beverley U.S.A.). The entire purification procedure was performed at 4°C. The concentration of the purified fusion proteins was determined using the Bio-Rad Protein Assay kit (see Appendix B). Digestion of the purified fusion proteins with the restriction protease, factor Xa (Boehringer Mannheim), was carried out at room temperature with the protease at a concentration of 200 µg/ml.

3.3.8. Preparation of cell extracts and enzyme assays

E. coli UM255 (*katG, katE*) cells harbouring the pMAL™-c2 constructs were grown and induced with IPTG, as described in section 3.3.5. The cells were collected by centrifugation, and resuspended in 500 µl of potassium phosphate buffer (pH 7.0) (see Appendix C) and 500 µl of lysozyme (4 mg/ml) (Sigma Chemicals, St. Louis U.S.A.). The cell suspensions were incubated at 37°C for 60 minutes and frozen at -20°C overnight. The extracts were thawed in ice water, and the cellular debris removed by centrifugation. The supernatants were recovered, and the pellets were resuspended in 1 ml of potassium phosphate buffer. The protein concentrations of the supernatant and pellet fractions were determined using the Bio-Rad Protein Assay kit. Gel-staining catalase and peroxidase activity assays were performed by the method of Clare *et al.* (1984) and Wayne and Diaz (1986). Briefly, 70 µg each of the supernatant and pellet fractions were separated on 7.5% (w/v) non-denaturing polyacrylamide gels. For catalase stains, the gels were immersed first in 5 mM hydrogen peroxide (Saarchem, Krugersdorp S.A.) for 20 minutes, and then in a 50:50 solution of 2% ferric chloride (Sigma) and 2% potassium ferricyanide (Sigma). Peroxide stains were performed by immersing the gels in a solution containing 5 mM hydrogen peroxide and 0.5 mg/ml 3,3-diaminobenzidine (Sigma). Catalase activity bands were visualized as clear bands on a dark green background, and peroxidase activity bands were visualized as brown bands on a clear background. The semi-quantitative catalase test of Wayne (1962) was performed using 80 µl of 10% (v/v) Tween 80 (Sigma), 80 µl of 30% hydrogen peroxide and 50 µg of purified fusion protein. Catalase activity was assessed by monitoring the production of bubbles from the reaction mixture. Quantitative catalase activity assays were performed by the method of Beers and Sizer (1951). A total of 42.5 ng of purified fusion protein was mixed with 5 mM hydrogen peroxide in 50 mM potassium phosphate buffer (pH 7.0), and the decomposition of hydrogen peroxide was monitored by measuring changes in OD_{240nm} for 60 seconds. One unit of catalase activity was defined as the amount of enzyme that catalyzes the decomposition of 1 µmol of hydrogen peroxide per minute at 25°C.

3.4. RESULTS

3.4.1. Confirmation of the sequence of the pD3 insert

The insert of the clone, pD3, which conferred resistance to ethyl methane sulfonate on DNA repair-deficient *E. coli* cells, was sequenced in both orientations (Shamila Nair, PhD thesis, University of Cape Town, 1993). The 1747 bp insert contains the 724 bp (241 amino acid) terminal portion of the *M. tuberculosis katG* gene, as well as approximately 1 kb of downstream sequence Since the above study, the sequencing of the entire *katG* genes from *M. tuberculosis* H37Rv and INH-resistant clinical isolates of *M. tuberculosis* has been reported. One such sequence from the GenBank database (Accession number X68081) was aligned with the sequence obtained from the pD3 clone, using CLUSTAL W (see Appendix B). The sequences were identical except in the two positions indicated in figure 3.1.

```
                                    ↓       ↓
ctgcagccacaagtcgggtgggaggtcaacgaccccgacggggatctgcgcaaggtcatt  60
cgcaccctggaagagatccaggagtcattcaactccgcggcgccggggaacatcaaagtg 120
tccttcgccgacctcgtcgtgctcggtggctgtgccgccatagagaaagcagcaaaggcg 180
gctggccacaacatcacggtgcccttcaccccgggccgcacggatgcgtcgcaggaacaa 240
accgacgtggaatcctttgccgtgctggagcccaaggcagatggcttccgaaactacctc 300
ggaaagggcaacccgttgccggccgagtacatgctgctcgacaaggcgaacctgcttacg 360
ctcagtgcccctgagatgacggtgctggtaggtggcctgcgcgtcctcggcgcaaactac 420
aagcgcttaccgctgggcgtgttcaccgaggcctccgagtcactgaccaacgacttcttc 480
gtgaacctgctcgacatgggtatcacctgggagccctcgccagcagatgacgggacctac 540
cagggcaaggatggcagtggcaaggtgaagtggaccggcagccgcgtggacctggtcttc 600
gggtccaactcggagttgcgggcgcttgtcgaggtctatggcgccgatgacgcgcagcca 660
aagttcgtgcaggacttcgtcgctgcctgggacaaggtgatgaacctcgacaggttcgac 720
gtgcgctgattcgggttgatcggccctgcccgccgatcaacacaacccgccgcagcaccc 780
cgcgagctgaccggctccggggctgtgtttgcccggcgcgatttgtcagaccccgcgtgc 840
atggtggtcgcaggcacgacgagacggggatgacgagacggggatgaggagaaagggcgc 900
cgaaatgtgctggatgtgcgatcacccggaagccaccgccgaggagtacctcgacgaggt 960
gtacgggataatgctcatgcatggctgggcggtacagcacgtggagtgcgagcgacggcc 1020
atttgcctacacggttggtctaacccggcgcggcttgcccgaactggtggtgactggcct 1080
ctcgccacgacgtgggcagcggttgttgaacatcgccgctcgcagggctctggtcggtga 1140
cttgctgaactcccggtatgcagaccaccctcccagccggccctcttgtcgaaacggtcc 1200
aggttacacatccggacgcgcatttgtattgtgcgatcgccatctttggcgacaaggtga 1260
cggccttgcagttggtgtgggccgaccggtggtcgctggccgtggggcggcggacttcga 1320
cgaaggtcggtacccagccggtgctcgggatgcgagccaccaggaggtcagcctgacgcg 1380
agcgctctaaacgtgttgccggagaacgatttgtgacagcaatgcagtgacagcggtggt 1440
tgctcgacacggaggtggccaacgcttccggacatcctccgatacatgcggtctgggcgg 1500
tgtgctggcggacaaggctgtgactcggtggaagcgttggcgccgtccgcgatttaccgc 1560
atcattgccggacactgcgtcggtggcgcgccgtccgtgcgctcgcaggcgcatatgggg 1620
atcgaacaagcccatcggcggcgggagaacgagtgattgcttccaggccgagtgagttcc 1680
ggataccggtaggccccagtaggctattgtgtgatgcgcttgaagccagccccatctcct 1740
gctgcag                                                       1747
```

Figure 3.1 Nucleotide sequence of the 1747 bp *Pst*I insert of pD3. The C and G residues indicated with arrows were absent from the sequence determined by S. Nair (PhD thesis, University of Cape Town, 1993), but present in the published *katG* sequence (GenBank Accession number X68081). The *katG* coding sequence is underlined, and the *Pst*I restriction endonuclease sites are in bold print.

In order to clear up these discrepancies, these regions of the insert were resequenced using the construct pSNS100, which contains the 1747 bp insert of pD3 in pUC19. Sequencing reactions generated with the Pharmacia T7 sequencing kit resulted in band compressions in these areas. Separation of the reactions on a polyacrylamide gel containing 40% (v/v) formamide, and sequencing in the reverse direction with a custom designed *katG* sequencing primer (see section 3.3.2.) did not resolve these compressions. The sequencing reactions were, therefore, repeated using the Sequenase™ Version 2.1 DNA Sequencing kit, and the base analogue deoxyinosine triphosphate (dITP). The resulting sequencing ladder (Fig. 3.2) confirmed the presence of a fourth C at position 33, and a fourth G at position 40, labelled with respect to the beginning of the *Pst*I restriction endonuclease recognition site (see Fig. 3.1). The sequence of the pD3 insert was, therefore, identical to the published *katG* sequence.

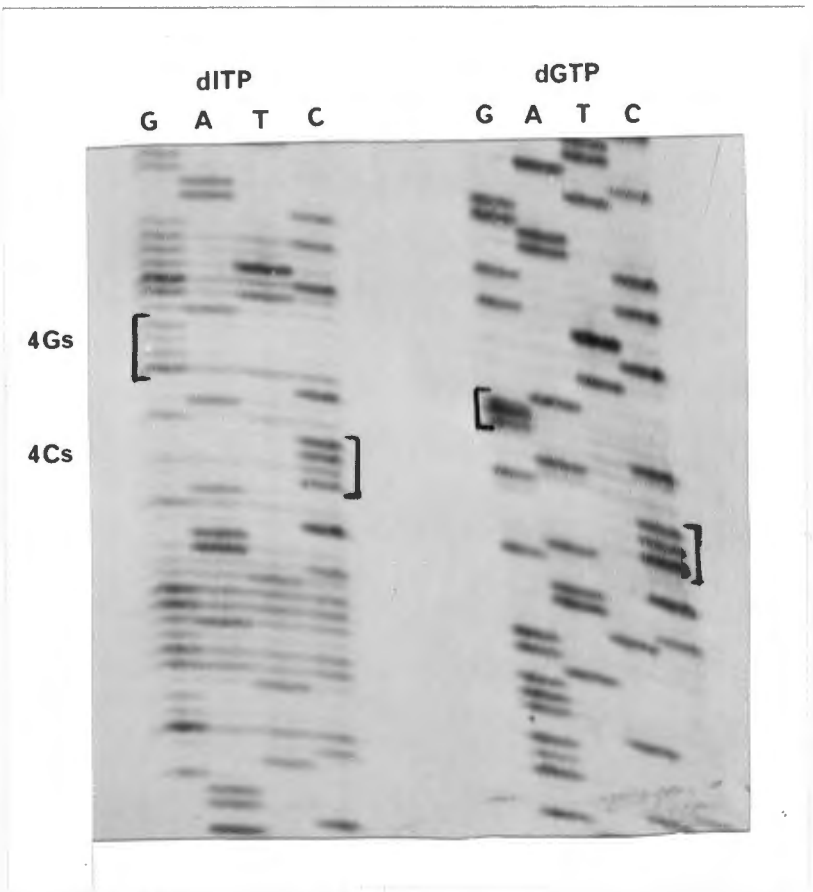
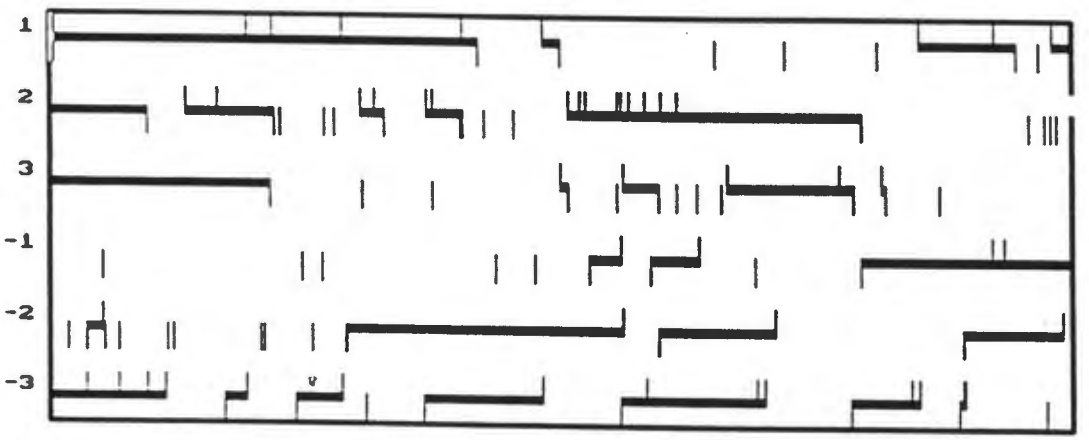


Figure 3.2 Sequencing ladders generated using the Sequenase™ Version 2.1 dITP and dGTP reagents. The positions of the band compression (dGTP) and the corresponding resolved regions (dITP) are indicated.

3.4.2. Computer analysis of the 1747 bp insert sequence

The experiments outlined in the introduction showed that a clone containing 1747 bp of *M. tuberculosis* H37Rv DNA could confer resistance to ethyl methane sulfonate on DNA repair-deficient *E. coli* cells. In order to determine a possible basis for this observation, the DNA sequence of the fragment was analyzed. Analysis of the sequence on GENEPRO™ Version 6.10 revealed three putative ORFs in the forward orientation (Fig. 3.3.a). The truncated *katG* gene is encoded in frame 1. Downstream of this are two putative open reading frames of 168 (ORF168, frame 2) and 56 (ORF56, frame 1) amino acids, encoding putative proteins of approximately 19 and 6 kDa, respectively. BLAST searches of these open reading frames revealed no significant homologies with any database amino acid sequences.

a)



b)

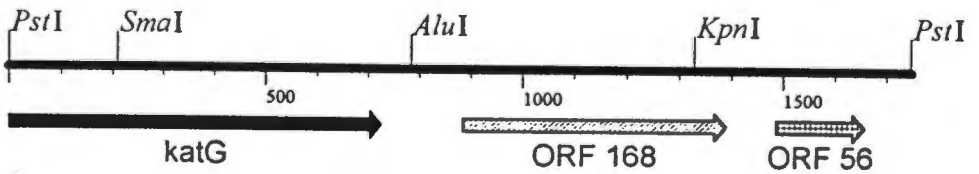
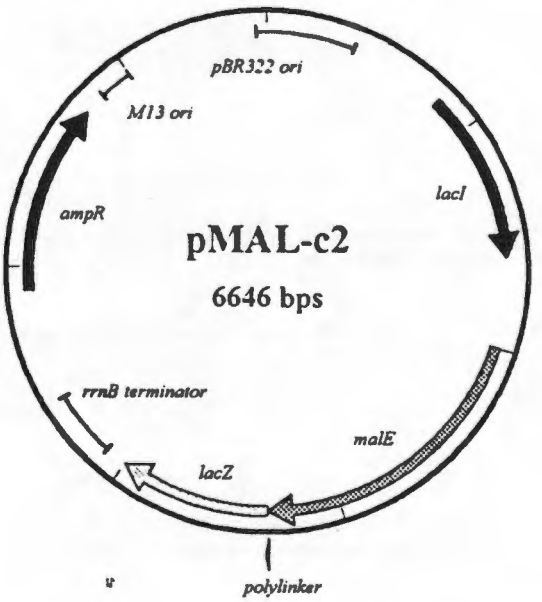


Figure 3.3 a) Open reading frame pattern of the 1747 bp *PstI* insert from pD3, generated using GENEPRO™. b) Linear restriction endonuclease map of the 1747 bp insert, including the more common restriction endonuclease sites and the putative ORFs identified using GENEPRO™.

3.4.3. Cloning of the 1747 bp fragment in pMAL™-c2

In order to assess the protective effects of the insert of pD3 on DNA repair-deficient cells *E. coli*, the insert was cloned into the expression vector pMAL™-c2 (Fig. 3.4). The use of pMAL™-c2 as an expression vector involves cloning the gene of interest downstream of, and in-frame with, the *malE* gene, which encodes the maltose binding protein, to form a translational fusion. The *malE* gene is under control of the *tac* promoter, which is repressed under normal conditions by the repressor encoded by the *lacI*^{q} gene, which is also present on the vector. High levels of expression from the promoter can be obtained by induction of cells harbouring the vector with IPTG.



POLYLINKER

	*Xmn*I	*Eco*RI	*Bam*HI	*Xba*I	*Sal*I	*Pst*I	*Hin*dIII	
malE... ATC GAG GGA AGG ATT TCA GAA TTC GGA TCC TCT AGA GTC GAC CTG CAG GCA AGC TTG ...*lacZ*								

↑
Factor Xa cleavage site

Figure 3.4 Restriction endonuclease map of the expression vector, pMAL™-c2, and sequence of the polylinker region.

The 1747 bp *Pst*I fragment from pD3 was cloned into pMAL™-c2, with the truncated *katG* ORF in-frame with the MBP, to form the construct pSMK3' (see Fig. 3.7). To determine whether the resistance to ethyl methane sulfonate, conferred on the *E. coli* DK1 cells, was due to expression of the truncated *kat*G gene, or whether other regions of the insert were involved, the following constructs were made, as described in the materials and methods; pSMK3'*, pSMK3'a, pSMK3'b, pSMK3'$_{opp}$, and pSMKd. The construct pSMK3'* contains only the *katG* portion of pSMK3', and no downstream sequences (see Fig. 3.7). Both pSMK3'a and pSMK3'b have bases either removed or added to change the frame of the insert with respect to the MBP. In pSMK3'$_{opp}$, the 1747 bp insert is cloned in the opposite orientation to that in pSMK3' The latter three constructs, therefore, contain the *kat*G sequences but do not express the gene product. The construct pSMKd contains the region downstream of the *Alu*I site of the insert (see Fig 3.7), i.e. no *kat*G sequences.

3.4.4. Analysis of expression of proteins from the pSMK3' derivatives

The vector, pMAL™-c2, encodes an MBP-LacZ α-fragment fusion protein of 50.8 kDa, while the *malE* gene alone encodes a protein of 42.7 kDa. In pSMK3', the truncated *katG* ORF encodes a 241 amino acid, or 26 kDa, portion of the KatG enzyme. Thus the construct encodes an MBP-KatG3' fusion protein of 68 kDa. In pSMK3'a, frame 2 of the insert contains a termination codon 160 bp downstream of the *Pst*I site. Thus, the construct should produce a MBP fusion of 49 kDa. Similarly, construct pSMK3'b should produce a fusion protein of 56 kDa, and pSMK3'$_{opp}$ should produce a protein of 52 kDa, due to fusion of the *malE* gene with sequences on the 3' end of the insert. The synthesis of proteins encoded by the constructs was analyzed using the *E. coli* S30 cell-free transcription and translation system (Promega) (Fig. 3.5). Fusion proteins of 50 kDa, 68 kDa, 52 kDa, 49 kDa, and 56 kDa were produced by pMAL™-c2, pSMK3', pSMK3'$_{opp}$, pSMK3'a, and pSMK3'b, respectively, as predicted. The construct pSMKd produced a protein of < 49 kDa, possibly due to termination of translation immediately after the end of the *malE* gene. No differences were observed in the proteins produced by pSMK3' and pSMK3'*, which suggests either that the downstream sequences present in pSMK3' do not encode any proteins, or that they do not contain promoters which function in *E. coli*. The transcription/translation results were confirmed by immunoblot analysis using anti-MBP antibodies (results not shown).

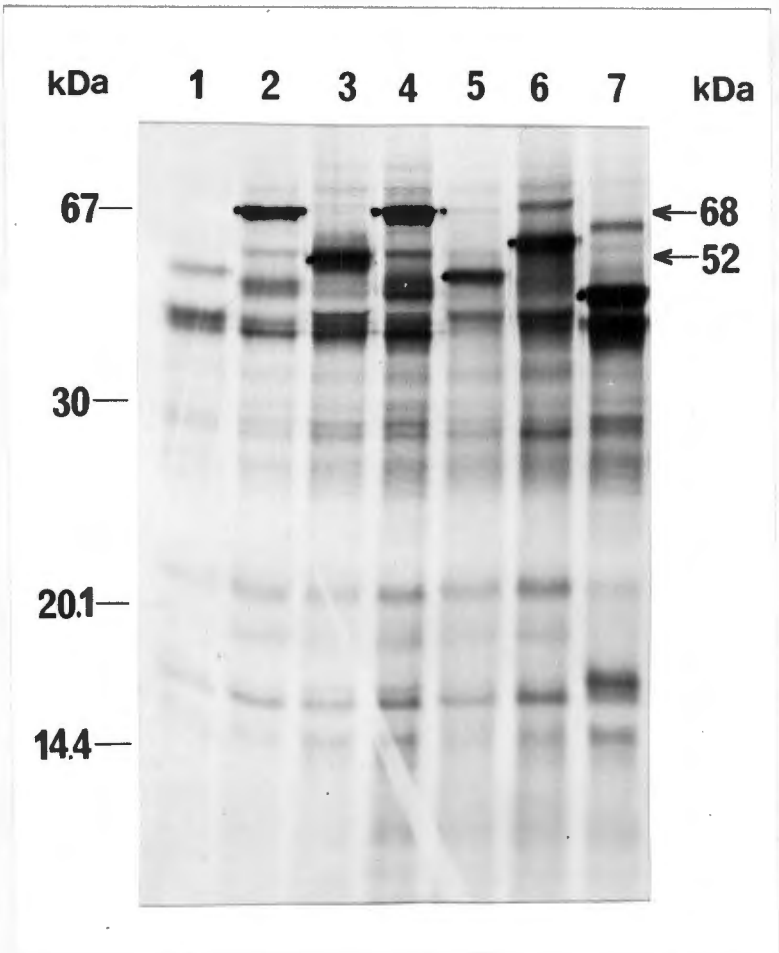


Figure 3.5 Autoradiogram of proteins produced by the pSMK3' derivatives, using the *E. coli* S30 cell-free transcription and translation system (Promega). The lanes were loaded as follows: 1) pMAL™-c2; 2) pSMK3'; 3) pSMK3'$_{opp}$; 4) pSMK3'*; 5) pSMK3'a; 6) pSMK3'b; and 7) pSMKd. The positions of the molecular weight markers (Pharmacia) are indicated. The major protein bands corresponding to the MBP fusion proteins are indicated with black dots.

3.4.5. Amplification and cloning of the whole *katG* gene in pMAL™-c2

Oligonucleotides were designed for the amplification of a 2250 bp fragment containing the *M. tuberculosis katG* gene. The primers were used in a PCR with *M. tuberculosis* H37Rv genomic DNA as the template. Reactions containing no template DNA or no primers yielded no product, while a reaction containing the forward primer yielded two products of approximately 700 and 800 bp. Only a reaction containing both primers at equal concentrations, together with the template DNA, resulted in a product of the predicted size (Fig. 3.6).

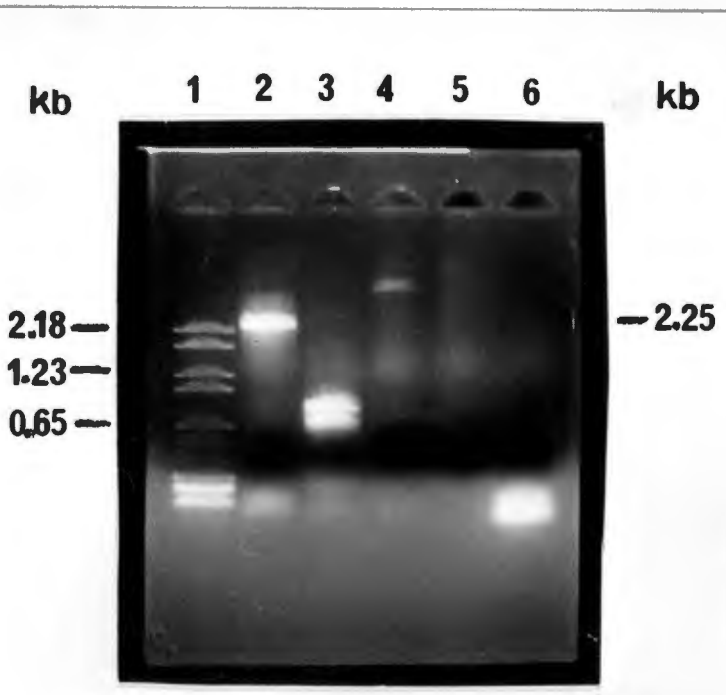


Figure 3.6 Agarose gel of the amplified *katG* gene from *M. tuberculosis* H37Rv and the control PCR reactions. The lanes were loaded as follows: 1) DNA molecular weight marker VI (Boehringer Mannheim); 2) Both primers and *M. tuberculosis* H37Rv template DNA; 3) Forward primer only; 4) Reverse primer only; 5) No primers; and 6) No template DNA.

The resulting 2250 bp fragment was cloned into the *Eco*RI and *Xba*I sites of pMAL™-c2, in-frame with the MBP, to form the construct pSMKA. A total of 1.1 kb of the insert was sequenced, in order to ensure that no errors were introduced during amplification, and found to be identical to the published *M. tuberculosis katG* sequence. Another construct, pSMK5', was made by deleting the 756 bp *Pst*I fragment from the 3' end of the gene in pSMKA (Fig. 3.7).

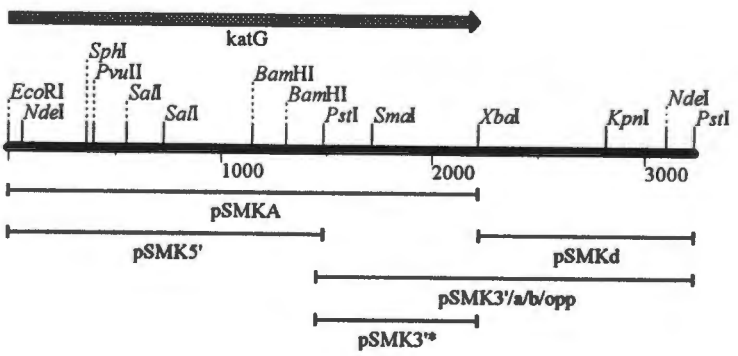


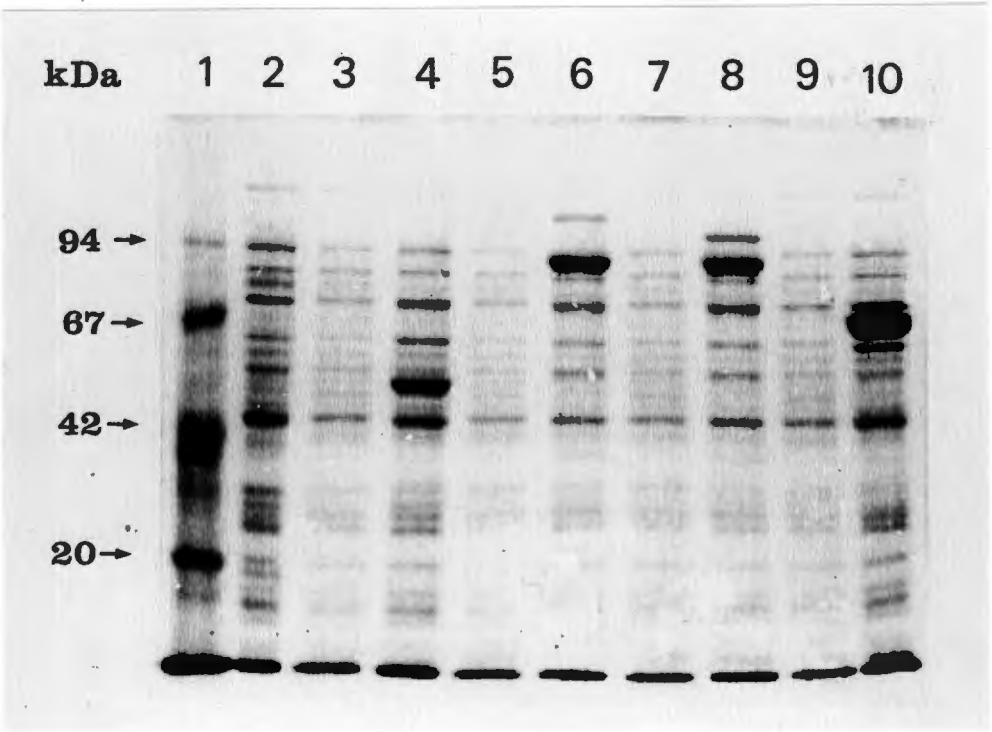
Figure 3.7 Linear restriction endonuclease map of the *M. tuberculosis katG* gene and downstream sequences. The restriction endonuclease sites, *Eco*RI and *Xba*I, were introduced by PCR. The position of the *katG* coding region is indicated. The regions present in the pMAL™-c2 constructs are shown below.

3.4.6. Analysis of expression and activity of the MBP-KatG fusion proteins

Expression of the fusion proteins from pSMKA, pSMK5', and pSMK3' was examined by IPTG-induction of *E. coli* UM255 cells carrying the plasmids, and separation of the resulting proteins by electrophoresis on a 12.5% (w/v) SDS-polyacrylamide gel (Fig. 3.8.a). The pSMKA construct was predicted to produce a fusion protein of approximately 125 kDa, due to fusion of the 82 kDa KatG with the MBP, while the MBP-KatG5' fusion protein was predicted to be 99 kDa. The fusion proteins were only visible after induction with IPTG, due to repression of the *tac* promoter by the $LacI^q$ repressor. The fusion proteins had relative mobilities of approximately 50.8 kDa (MBP-LacZ), 125 kDa (MBP-KatGAll), 99 kDa (MBP-KatG5'), and 68 kDa (MBP-KatG3'), as predicted. The two larger fusion proteins, however, appeared predominantly as cleavage products of approximately 90 kDa, which cross-reacted with anti-MBP antibodies (Fig. 3.8.b), suggesting that the cleavage site lies within the KatG protein, near the C-terminal. Analysis of the bands on the autoradiograph in figure 3.8.b, using a Shimadzu CS-9000 dual wavelength flying spot scanner, revealed that 94.6% of the MBP-KatGAll, and 92.8% of the MBP-KatG5' fusion protein was present as the cleavage product. In an attempt to prevent the formation of the cleavage product, expression of the fusion proteins was tested in a variety of *E. coli* strains, namely LK111, LK112, TB1, DH5α, JM109, and HB101, and by induction at different temperatures, using different amounts of IPTG, and for different time periods. Under all of these conditions, the same pattern of protein synthesis was observed (results not shown). Cloning of the insert in-frame with the maltose binding protein in pMAL™-p2 did not resolve this problem either. In this vector, the *malE* gene contains a signal peptide, which directs fusion proteins to the periplasm. However, fusion protein expressed from pMAL™-p2 constitutes only 5-10% of total cellular protein, and could not be detected on an SDS-polyacrylamide gel (results not shown). The cleavage is probably due to an *E. coli* protease.

Cell extracts from uninduced and IPTG-induced *E. coli* UM255 cells harbouring the pMAL™-c2 constructs were separated by electrophoresis in 7.5% (w/v) non-denaturing polyacrylamide gels. Gel staining catalase and peroxidase assays (Fig. 3.9) revealed that only the fusion containing the whole KatG protein was active. Activity, therefore, requires both domains of the protein, and is not affected by the presence of the MBP. The *E. coli* UM255 cells are mutant in both *katG* and *katE*, and therefore display no endogenous catalase activity.

a



b

Figure 3.8 Expression of the MBP-KatG fusions in *E. coli* UM255. The proteins were separated on 12.5% (w/v) SDS-polyacrylamide gels. Proteins were visualized by a) staining with Coomassie® Brilliant Blue R-250, and b) immunoblotting with anti-MBP serum. The lanes were loaded as follows: 1) Molecular weight standards (Pharmacia); 2) untransformed *E. coli* UM255 induced; 3) and 4) uninduced and induced fractions from *E. coli* (pMAL™-c2); 5) and 6) corresponding fractions from *E. coli* (pSMKA); 7) and 8) corresponding fractions from *E. coli* (pSMK5'); 9) and 10) corresponding fractions from *E. coli* (pSMK3').

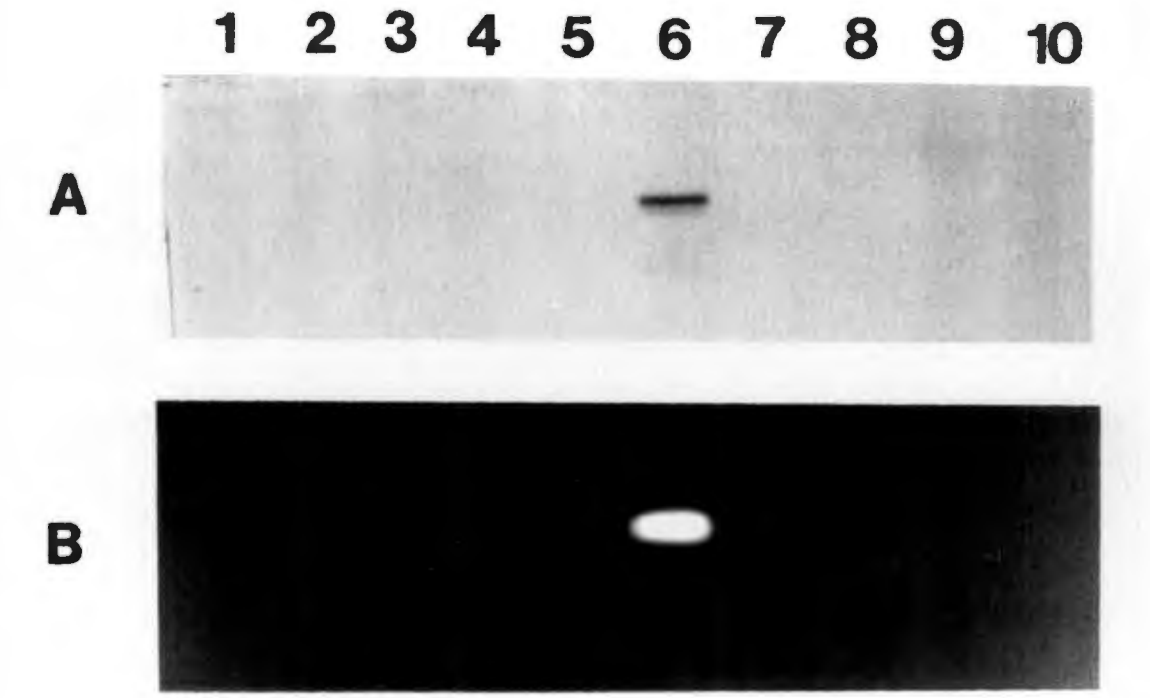


Figure 3.9 Gel-staining catalase and peroxidase activity assays. Approximately 70 μg of total protein isolated from uninduced and induced cultures of *E. coli* UM255 harbouring the pMAL™-c2 constructs were separated by electrophoresis in a 7.5% non-denaturing polyacrylamide gel, and stained for A) peroxidase, and B) catalase activity. The lanes were loaded as follows: 1) untransformed *E. coli* UM255 uninduced 2) untransformed *E. coli* UM255 induced; 3) and 4) uninduced and induced fractions from *E. coli* (pMAL™-c2); 5) and 6) corresponding fractions from *E. coli* (pSMKA); 7) and 8) corresponding fractions from *E. coli* (pSMK5'); 9) and 10) corresponding fractions from *E. coli* (pSMK3').

3.4.7. Purification of the MBP-KatG fusion proteins and analysis of enzyme activity

The MBP-KatG fusion proteins were purified by affinity chromatography, as described in the materials and methods. The purity of the fusion proteins was analyzed by SDS-polyacrylamide gel electrophoresis of samples taken during the purification process (Fig. 3.10). These results indicated that the fusion proteins were present after induction of the cells, in the supernatant after lysis, and in the fractions eluted from the column with column buffer containing maltose. For all three fusion proteins, the protein was eluted from the amylose column in fractions 3-6.

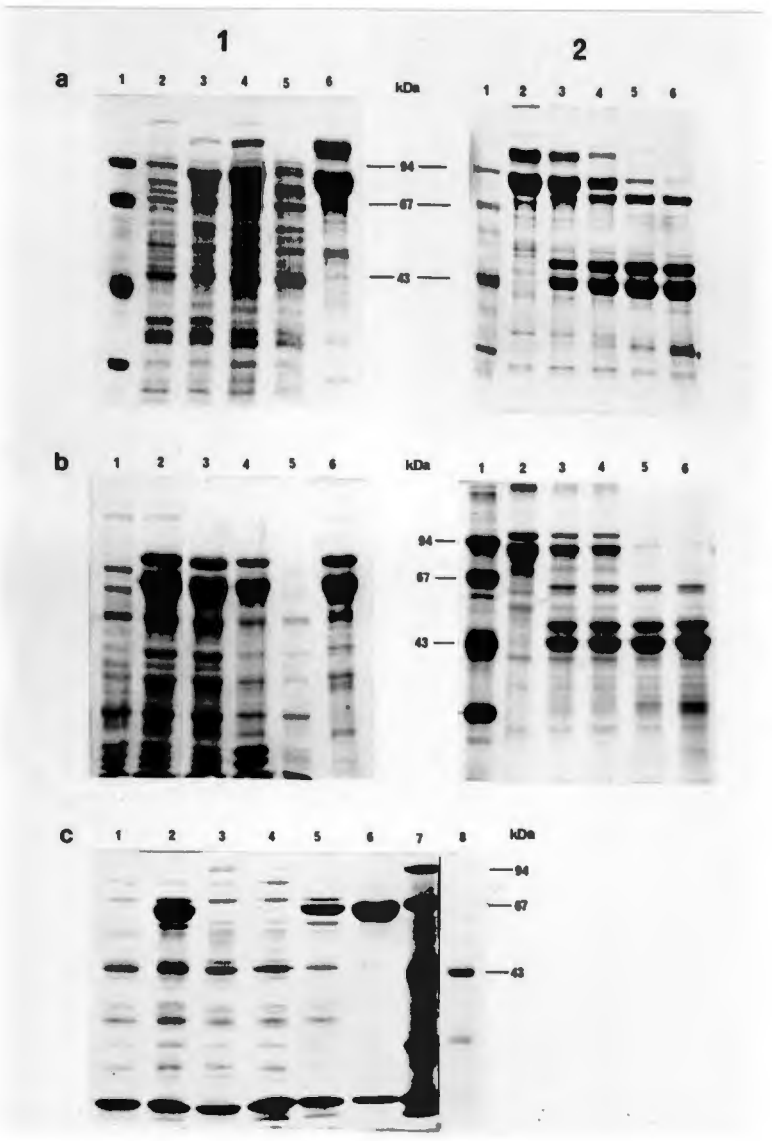


Figure 3.10 Separation of protein purification products on a 12.5% (w/v) SDS-polyacrylamide gel. a) 1) Purification of MBP-KatGAll. The lanes were loaded as follows: 1) Molecular weight standards (Pharmacia); 2) *E. coli* UM255 (pSMKA) uninduced; 3) *E. coli* UM255 (pSMKA) induced; 4) *E. coli* UM255 (pSMKA) cell extract; 5) Cell extract passed through the amylose column; 6) Eluant from the amylose column. 2) Cleavage of the MBP-KatGAll fusion protein with factor Xa: 1) Molecular weight standards; 2) No factor Xa; 3) 200 ng factor Xa digested at 4°C; 4) 200 ng factor Xa digested at room temperature; 5) 400 ng factor Xa digested at room temperature; 6) 1 μg factor Xa digested at room temperature. All of the lanes contain 20 μg of fusion protein. b) Purification of MBP-KatG5'. The lanes were loaded as follows: 1) *E. coli* UM255 (pSMK5') uninduced; 2) *E. coli* UM255 (pSMK5') induced; 3) Cell extract; 4) Pellet fraction; 5) Cell extract passed through the amylose column; 6) Eluant from the amylose column. 2) Cleavage of the fusion protein with factor Xa: 1) Molecular weight standards; 2) No factor Xa; 3) 200 ng factor Xa digested at 4°C; 4) 200 ng factor Xa digested at room temperature; 5) 400 ng factor Xa digested at room temperature; 6) 1 μg Xa digested at room temperature. All of the lanes contain 20 μg of fusion protein. c) Purification of MBP-KatG3' and cleavage with factor Xa. The lanes were loaded as follows: 1) *E. coli* UM255 (pSMK3') uninduced; 2) *E. coli* UM255 (pSMK3') induced; 3) Supernatant after harvesting of the cells; 4) Pellet fraction of the cell extract; 5) Supernatant fraction of the cell extract; 6) Eluant from the amylose column; 7) Molecular weight standards; and 8) MBP- KatG3' cleaved with factor Xa.

The total yield of MBP-KatGAll was 10.2 mg/l of culture, however, the full-size product is only present as 31.5% of the preparation (as calculated from a densitometer scan of Fig. 3.10.a.1, lane 6), therefore, the yield of the full-size MBP-KatGAll protein is approximately 3.57 mg/l of culture. Cleavage with factor Xa was performed using increasing amounts of the protease, at room temperature and at 4°C. The most efficient digestion was obtained at room temperature with 1 μg of factor Xa per 20 μg of fusion protein (Fig. 3.10.a.2). Cleavage of MBP-KatGAll resulted in proteins of 35, 42.7, 47, and 82 kDa. The 42.7 kDa protein is the MBP, and the 82 kDa protein is the full-size *M. tuberculosis* KatG. The smaller proteins, 35 and 47 kDa, represent the cleavage products of the KatG enzyme. The total yield of MBP-KatG5' was 27.6 mg/l of culture. The yield of the full-size fusion protein, which is present as 40.4% (as calculated from a densitometer scan) of the total preparation, is 11.6 mg/l of culture. Cleavage with factor Xa resulted in proteins of approximately 42.7, 47, and 56 kDa, representing the MBP, the KatG cleavage product, and KatG5', respectively. The 56 kDa KatG5' protein is present in lower levels and appears to have a retarded mobility in the polyacrylamide gel (Fig. 3.10.b.2). The yield of MBP-KatG3' was 50 mg/l culture, and the purity was > 95%, as estimated from an SDS-polyacrylamide gel (Fig. 3.10.c). Cleavage of the fusion with protease factor Xa resulted in the release of the 42.7 kDa MBP, and the 26 kDa KatG3' protein. The yields of fusion protein obtained from affinity chromatography were similar to those predicted by the manufacturers (10-40 mg/litre culture) (New England Biolabs Protein Fusion and Purification manual). The variations in the yield and purity of the fusion proteins may have been due to the size differences of the proteins. Alternatively, the low yield may have been due to the retention of significant levels of the larger fusion proteins in the pellet fractions of the cell extracts. The full-size KatG and the 5' two thirds may have formed non-specific associations with other proteins in the cell extract, thus resulting in protein contaminants after chromatography.

A semi-quantitative catalase assay, using 50 μg of purified fusion protein, revealed activity with MBP-KatGAll, but not with the other fusions. Catalase activity was quantified by measuring the decomposition of hydrogen peroxide by the fusion proteins at 240 nm. A total of 42.5 μg of the full-size MBP-KatGAll caused a decrease in the concentration of hydrogen peroxide of 198630 μM/minute. Thus the specific catalase activity of the preparation was calculated as

2336.824 units/mg of protein at 25°C. The fusion proteins containing the 5' or 3' segments of KatG exhibited no detectable catalase activity.

3.5. DISCUSSION

The primary subject of this study was a clone, designated pD3, which conferred resistance to ethyl methane sulfonate on DNA repair-deficient *E. coli* cells. Analysis of the sequence of the 1747 bp insert of pD3 revealed two putative open reading frames downstream of the *M. tuberculosis katG* gene. Analysis of protein synthesis by constructs containing the insert, however, revealed no additional proteins in the presence of these sequences. If the sequences downstream of the *katG* gene were transcribed on a polycistronic message with the *malE-katG* fusion, the level of proteins encoded by downstream sequences would be high, due to the high activity of the IPTG-induced *tac* promoter. Thus the protein bands would be present at approximately equal intensity to those of the fusion proteins on an SDS-polyacrylamide gel, which they clearly were not. Little is known about transcription termination in mycobacteria. Nair *et al.* (1993) identified a 12 bp inverted repeat, 36 bp downstream of the termination codon of the gene encoding MI43 from *M. intracellulare*. The authors suggest that the formation of stable hairpin-loop secondary structures in regions of dyad symmetry may be a common mechanism of terminating transcription in mycobacteria (Collins *et al.*, 1990; Nair *et al.*, 1992; Shinnick *et al.*, 1989). Immediately downstream of the *M. tuberculosis katG* gene is a 7 bp inverted repeat sequence with dyad symmetry, which may be analogous to that described by Nair *et al.* (1993). The lack of detectable protein production by the sequences downstream of the *M. tuberculosis katG* gene suggests that, either no proteins are encoded in this region, or they are not transcribed as a part of the *malE-katG* mRNA message, and do not contain promoters which function in *E. coli*.

Expression of the whole *M. tuberculosis katG* gene, as well as the 5' two-thirds, as a translational fusion with the maltose binding protein resulted in the formation of a cleavage product. As mentioned in section 3.4.5, only half of the amplified *katG* gene was sequenced to check for PCR errors. However, the region sequenced included the area between nucleotides 3180 and 3240, labelled with respect to the published sequence (GenBank Accession number X68081). This is the approximate region of the gene which should contain the cleavage site in

order to produce a fusion protein of 90 kDa. No discrepancies in the sequence were noted in this region, suggesting that the cleavage was not due to an error in the coding sequence introduced during the PCR amplification. This is substantiated by the fact that the full-size MBP-KatGAll fusion protein exhibits both catalase and peroxidase activities. The presence of cleavage products of the *M. tuberculosis* KatG has been noted by Johnsson *et al.* (1997) and Heym *et al.* (1997). The first-mentioned authors observed a cleavage product of approximately 77 kDa, and suggest that this is due to proteolytic degradation of the protein. In addition, proteolytic cleavage has been observed in the catalase isolated from rat liver (Robbi and Lazarow, 1978; Crane *et al.*, 1982). Cleavage of the fusion proteins in *E. coli* cells could not be prevented by using another *E. coli* strain, or by varying the concentration of IPTG, or the time period of induction (results not shown). Only the full-size fusion product exhibited catalase and peroxidase activity. This agrees with the observations of Zhang *et al.* (1992), who demonstrated that deletion of the 3' end of the *M. tuberculosis katG* gene resulted in loss of enzyme activity. Although the catalase and peroxidase activities are predicted to lie in the N-terminal region of the protein (Welinder, 1991; 1992), the C-terminal region may play an important role in substrate binding or subunit interaction. The specific activity of the MBP-KatG fusion was 2336.824 units/mg of protein. This is similar to that recently reported by Johnsson *et al.* (1997) for a recombinant *M. tuberculosis* KatG isolated in *E. coli* (2420 units/mg). The fact that fusion of the KatG protein with the maltose binding protein did not affect the activity of the enzyme, shows that dimerization and correct folding of the enzyme still occur. Similarly, Zhang *et al.* (1992) demonstrated that fusion of the *M. tuberculosis* KatG to a portion of the LacZ protein did not interfere with enzyme activity.

The pMAL^{TM}-c2 vector provides a convenient and efficient means of overproducing a protein of interest in *E. coli* cells. IPTG-induction of the translational fusion with the maltose binding protein results in large quantities of the fusion which can easily be detected by SDS-polyacrylamide gel electrophoresis. The purification of the fusion protein from total cell extracts by affinity chromatography is also rapid and simple. The constructs described in this chapter allow overexpression of different regions of the *M. tuberculosis* KatG gene in *E. coli*, and provide a useful means of assessing the effects of the various regions of the KatG protein on *E. coli* cells exposed to various DNA damaging agents, which is the focus of the next chapter.

CHAPTER 4

COMPLEMENTATION OF DNA REPAIR-DEFICIENT *E. COLI* STRAINS WITH THE *M. TUBERCULOSIS KATG* GENE

CONTENTS Pg.

CHAPTER 4

COMPLEMENTATION OF DNA REPAIR-DEFICIENT *E. COLI* STRAINS WITH THE *M. TUBERCULOSIS KATG* GENE

4.1. SUMMARY

This chapter describes complementation studies with the *M. tuberculosis katG* gene in *E. coli* strains mutant in DNA repair enzymes. The whole KatG, the proximal two thirds, and the terminal third of the enzyme were overexpressed, as translational fusions to the maltose binding protein, in 5 DNA repair-deficient strains of *E. coli*. The transformants were tested for their ability to survive exposure to three DNA damaging agents, short wavelength ultraviolet (UV) light irradiation, mitomycin C, and hydrogen peroxide. Overexpression of the proteins resulted in an increase in the survival of *recA* mutants exposed to UV irradiation and hydrogen peroxide, and in the survival of excision repair mutants exposed to mitomycin C. Both regions of the *M. tuberculosis* KatG protein conferred the above effects, and this was independent of the catalase or peroxidase activity of the enzyme. The results suggest that the *M. tuberculosis katG* gene may encode a novel function related to the repair of DNA damage, and this may have implications for the survival of *M. tuberculosis* in the presence of DNA damaging agents, for example, in the macrophage. UV sensitivity tests on a selection of mycobacterial species revealed that the *M. tuberculosis* complex organisms are more sensitive than the 'atypical mycobacteria.' UV sensitivity tests on *M. intracellulare* and *M. tuberculosis* strains mutant in *katG* revealed no significant differences in the sensitivity of the cells in the absence of this gene. This suggests that the *katG* gene product does not play a significant role in the survival of mycobacterial cells after exposure to short wavelength UV irradiation, in the absence of mutations in essential DNA repair genes.

4.2. INTRODUCTION

The previous chapter describes the construction and testing of systems for the overexpression of different regions of the *M. tuberculosis* KatG enzyme in *E. coli* cells. Fusions of the *katG* gene with the *malE* gene, encoding the maltose binding protein, were constructed in the vector

pMAL™-c2, and shown to express fusion proteins of the predicted size. Only a fusion containing the full size KatG enzyme exhibited catalase and peroxidase activity, as expected from previous reports (Zhang *et al.*, 1992). The effect of these fusions on DNA repair-deficient *E. coli* cells exposed to various DNA damaging agents was tested.

Five DNA repair-deficient *E. coli* strains were selected for the study, based on their sensitivity to short wavelength (254 nm) UV light and mitomycin C. In *E. coli* strain DK1, the *recA* gene is deleted, and the *recE* gene is non-functional due to a point mutation. These cells are, therefore, unable to induce the production of many of the DNA repair genes in the SOS response, in response to DNA damaging agents, and do not have a functional recombinational repair system. Strain CC118 contains a mutation in the N-terminal recombinase domain of RecA, and in the *recE* gene (Manoil and Beckwith, 1985). The cells are, therefore, able to induce the SOS response, but are defective in recombinational repair. The strains AB1884, AB1885, and AB1886 are derived from strain AB1157, in which the *recE* gene is deleted (Howard-Flanders *et al.*, 1966). These strains are mutant in the *uvrC*, *uvrB*, and *uvrA* genes, respectively, which are required for Uvr-dependent nucleotide excision repair. The cells mutant in *uvrA* and *uvrB* are completely defective in excision repair, while cells mutant in *uvrC* exhibit residual nucleotide excision repair activity. This may be due to the sufficient stabilization of damaged DNA by the $(UvrA)_2(UvrB)$ complex to increase the efficiency of other repair systems (Van Houten, 1990). These cells are also mutant in *recE*, but are able to undergo recombinational repair via a RecE-independent pathway. The different strains were used in this study in an attempt to identify which aspects or pathways of DNA repair were complemented by the *M. tuberculosis katG* gene. In addition, three different DNA damaging agents were used. These agents each cause different lesions in the DNA, which are repaired via specific pathways. As reviewed in Chapter 2, short wavelength UV irradiation primarily causes pyrimidine dimers, mitomycin C causes cross-linking of complementary DNA strands, and hydrogen peroxide causes a variety of lesions, including base modification, strand breaks, and base fragmentation.

Until recently, attempts to introduce mutations in or to inactivate specific genes in *M. tuberculosis* complex organisms have been unsuccessful. Experiments have been hampered by the high frequency of illegitimate recombination in these organisms (Kalpana *et al.*, 1991; Aldovini *et al.*, 1993; Colston and Davis, 1994). Thus the availability of appropriate DNA

repair-deficient mutants of the mycobacteria has been limited. The construction of mutants of *M. smegmatis*, in which the *recA* gene has been inactivated by allelic exchange, has recently been reported (Scholtz *et al.*, ASM Abstract U-117, 1997). These mutants exhibited a > 100-fold reduction in viability after exposure to ethyl methane sulfonate, relative to wild type *M. smegmatis* cells. These mutants may be useful for future complementation studies with the *M. tuberculosis katG* gene as described in this chapter with *E. coli* mutants.

In an attempt to examine the role played by the *katG* gene in the survival of mycobacteria exposed to DNA damaging agents, in the absence of DNA repair mutants of mycobacteria, the association between the presence of a functional KatG and sensitivity to UV was examined. Interspecies differences in the sensitivity of mycobacteria to short wavelength UV irradiation have been reported (Collins, 1971; David *et al.*, 1971). The UV sensitivities of a range of mycobacteria were examined by David (1973). The order of sensitivity was found to be: *M. tuberculosis* > *M. fortuitum* > *M. avium-intracellulare* > *M. kansasii* > *M. smegmatis*. The author examined the underlying factors related to the variable responses of these organisms to UV irradiation. It was demonstrated, using phage reactivation experiments, that the different responses of, at least, *M. smegmatis* and *M. fortuitum* were due to differences in the efficiency of their respective repair systems. The UV sensitivity of the mycobacteria was shown to be independent of the growth rate, and G+C content, but correlated with the size of the genome and the cellular concentration of carotenoids.

4.3. MATERIALS AND METHODS

The source and characteristics of the DNA repair-deficient *E. coli* strains and the mycobacterial strains used in this study are given in Appendix A.

4.3.1. Transformation of DNA repair-deficient *E. coli* strains with the pMAL™-c2 constructs

Competent *E. coli* DK1, CC118, AB1884, AB1885, and AB1886 cells were prepared by the method of Chung and Miller (1988), using DMSO. The cells were transformed with the plasmids pMAL™-c2, pSMKA, pSMK5', and pSMK3', and selected on 2x YT agar plates

containing ampicillin. Transformants were verified by analysis of protein production after IPTG-induction, as described in section 3.3.5.

4.3.2. UV sensitivity tests

IPTG-induced cultures of *E. coli* mutants, transformed with the pMAL™-c2 constructs (see section 3.3.5), were diluted 10-fold in water, and 5 ml placed in a glass petri dish. These were irradiated with short wavelength (254 nm) UV light, at a dose of 1 J/m^2, for different time periods, with a Model UVG-11 254 nm single wavelength lamp (UV Products, Upland U.S.A.). UV doses were measured using a Model J-225 short wavelength Blak-Ray UV meter (UV Products, Upland U.S.A.). Samples were withdrawn before irradiation and at the indicated times, diluted in water, and plated on 2x YT agar containing ampicillin. After overnight incubation at 37°C, colonies were counted, and surviving fractions calculated relative to the unirradiated sample. All experiments were performed in the absence of direct light, and the plates were incubated in the dark to avoid photoreactivation (Jones and Woods, 1981; Abratt *et al.*, 1985). Semi-quantitative UV sensitivity tests were performed by streaking 10 µl of stationary phase *E. coli* cells on a 2x YT agar plate, and selective exposure to short wavelength UV irradiation at a dose of 1 J/m^2. Sections of the plate were shielded from irradiation with cardboard for the appropriate time periods (Sambrook *et al.*, 1989). The UV sensitivities of a selection of mycobacterial species were tested using the quantitative method described above. Mycobacterial cells were grown to stationary phase in 5 ml Middlebrook-ADC-Tween (3-4 days for the rapid growers, and 14 days for the slower growers). Irradiated samples were plated on Middlebrook-ADC agar plates containing 0.01% (w/v) cycloheximide (Fluka Biochemicals), and incubated at 37°C for 5 days (rapid growers) or 4-6 weeks (slow growers). Clumping of the mycobacterial cells was minimized by including Tween 80 in the medium, and by vigorous stirring of the cell suspension before diluting and plating, to facilitate accurate colony counts. All experiments were performed in triplicate.

4.3.3. Mitomycin C sensitivity tests

IPTG-induced cultures of the *E. coli* transformants were diluted 10-fold in fresh medium, and a sample was removed for plating. Mitomycin C (Merck, Darmstadt Germany) was added to a

final concentration of 6 μg/ml, and the cells were shaken at 37°C. Samples were removed at different time points, diluted in water and plated on 2x YT agar plates containing ampicillin at a concentration of 50 μg/ml. Colonies were counted after overnight incubation at 37°C, and surviving fractions calculated as described for the UV sensitivity experiments. Mitomycin C was freshly prepared for each set of experiments, and kept on ice between experiments. Semi-quantitative mitomycin C sensitivity tests were performed by plating stationary phase *E. coli* cells on plates containing mitomycin C at a concentration of 1μg/ml, and on control plates without mitomycin C. Sensitivity to mitomycin C was determined by comparison of growth on the mitomycin C plates with those on the control plates.

4.3.4. Hydrogen peroxide sensitivity tests

Hydrogen peroxide sensitivity tests were carried out by the method of Demple *et al.* (1983), as described for mitomycin C, except that hydrogen peroxide (Saarchem) was added to a final concentration of 50 mM. Results were processed in the same way as for the UV and mitomycin C experiments.

4.4. RESULTS

4.4.1. Selection of *E. coli* DNA repair-deficient mutants for this study

The following DNA repair-deficient *E. coli* strains were tested for their ability to survive treatment with short wavelength UV light and mitomycin C, using semi-quantitative methods: DK1(*recA*, *recE*), HB101 (*recA*), JC5519 (*recE*, *recBC*), JC8679 (*recBC*), CC118 (*recA1*, *recE*), AB1157 (*recE*), AB1884 (*uvrC*, *recE*), AB1885 (*uvrB*, *recE*), and AB1886 (*uvrA*, *recE*). Five strains, DK1, CC118, AB1884, AB1885, and AB1886, exhibited significant sensitivity to either or both of the agents (results not shown), and were selected for further experiments. Strain MC1060, the parent strain of DK1, and AB1157, the parent strain of AB1884, AB1885, and AB1886, were included as controls. Pilot experiments were performed to determine the doses of UV irradiation, mitomycin C, and hydrogen peroxide which resulted in a significant difference between the survival curves of parent and mutant strains, but which allowed sufficient survival of mutants such that colony counts remained accurate. Doses of

8 J/m^2/s UV irradiation, 6 μg/ml mitomycin C, and 50 mM hydrogen peroxide were selected for further experiments.

4.4.2. Transformation of the mutant *E. coli* strains with the pMAL^{TM}-c2 constructs

Competent *E. coli* DK1, CC118, AB1884, AB1885, and AB1886 cells were transformed with the pMAL^{TM}-c2-*katG* constructs, pSMKA, pSMK5', and pSMK3', described in Chapter 3, as well as the vector control. In addition, competent *E. coli* DK1 cells were transformed with the pSMK3' derivatives, pSMK3'*, pSMK3'a, pSMK3'b, pSMK3'$_{opp}$, and pSMKd. Transformants were verified by SDS-PAGE analysis of proteins produced after IPTG-induction. The transformants all overproduced the predicted size fusion proteins after induction (results not shown).

4.4.3. The UV sensitivities of the *E. coli* DNA repair-deficient transformants

The transformants described in section 4.4.2. were examined for their ability to survive exposure to short wavelength (254 nm) UV light irradiation. The parent strains, MC1060 and AB1157, were included as controls. A representative selection of survival curves is shown in figure 4.1.a, and the results are summarized in table 4.1. The *recA*$^+$ and *uvr*$^+$ parent strains, MC1060 and AB1157, showed < 1% killing at the doses used, due to the presence of efficient DNA repair systems for the removal of damaged DNA. *E. coli* DK1 (pMAL^{TM}-c2) had a similar UV survival curve to that of untransformed DK1 (not shown), i.e. a shoulderless curve representing exponential killing of the cells, which levels out as the damage sites on the DNA become saturated. When DK1 was transformed with pSMKA or pSMK5', the survival curves levelled out with a higher surviving fraction (Fig. 4.1.a, panel 1), and the cells exhibited a > 200-fold increase in survival at a UV dose of 8 J/m^2/s (Table 4.1). A similar effect was observed with DK1 (pSMK3'), with the increase in survival being 108-fold. The presence of the *katG* constructs, therefore, restored a level of resistance to the cells, although not to wild type levels. The UV survival curves for the CC118 transformants were similar to those of DK1, but this strain is naturally more resistant to UV (Fig. 4.1.a, panel 2). The presence of both pSMKA and pSMK5' increased the survival of the cells 36.7- and 28.2-fold, respectively, at a dose of 6 J/m^2/s (Table 4.1). For CC118 (pSMK3'), survival was approximately 2-fold greater than the

control at 6 J/m^2/s. From the AB1886 survival curves (Fig. 4.1.a, panel 3), and the relative survival values given in table 4.1, it is evident that none of the *katG* constructs conferred any resistance to UV on the excision repair mutants. In fact, in some cases the constructs increased the sensitivity of the cells.

Table 4.1 Analysis of the UV sensitivities of *E. coli* DNA repair mutants transformed with the *katG* constructs.

Strain	Construct	Surv. fraction (6 J/m^2/s)[1]	Rel. surv. x/c[2]
DK1	pMAL™-c2	2.07 (± 0.79) x 10^{-5}	1.00
	pSMKA	4.56 (± 1.01) x 10^{-3}	220.3
	pSMK5'	4.36 (± 1.25) x 10^{-3}	210.6
	pSMK3'	2.25 (± 1.01) x 10^{-3}	108.7
CC118	pMAL™-c2	6.24 (± 2.45) x 10^{-5}	1.00
	pSMKA	2.29 (± 0.63) x 10^{-3}	36.7
	pSMK5'	1.76 (± 0.63) x 10^{-3}	28.2
	pSMK3'	1.52 (± 0.68) x 10^{-4}	2.40
AB1884	pMAL™-c2	1.50 (± 1.45) x 10^{-4}	1.00
	pSMKA	1.06 (± 0.21) x 10^{-5}	0.07
	pSMK5'	1.26 (± 1.47) x 10^{-4}	0.84
	pSMK3'	4.13 (± 3.09) x 10^{-5}	0.28
AB1885	pMAL™-c2	1.13 (± 1.07) x 10^{-5}	1.00
	pSMKA	6.98 (± 0.78) x 10^{-6}	0.62
	pSMK5'	1.11 (± 0.50) x 10^{-5}	0.98
	pSMK3'	1.22 (± 0.62) x 10^{-5}	1.08
AB1886	pMAL™-c2	2.07 (± 0.56) x 10-5	1.00
	pSMKA	1.48 (± 0.46) x 10^{-5}	0.71
	pSMK5'	8.89 (± 3.63) x 10^{-6}	0.43
	pSMK3'	9.57 (± 2.46) x 10^{-6}	0.46

[1] Surviving fraction at a dose of 6 J/m^2/s. The mean of three independent experiments ± the standard error of the mean.
[2] Relative survival i.e. the surviving fraction of the construct at a dose of 6 J/m^2/s divided by the surviving fraction of the control, pMAL™-c2, at 6 J/m^2/s.

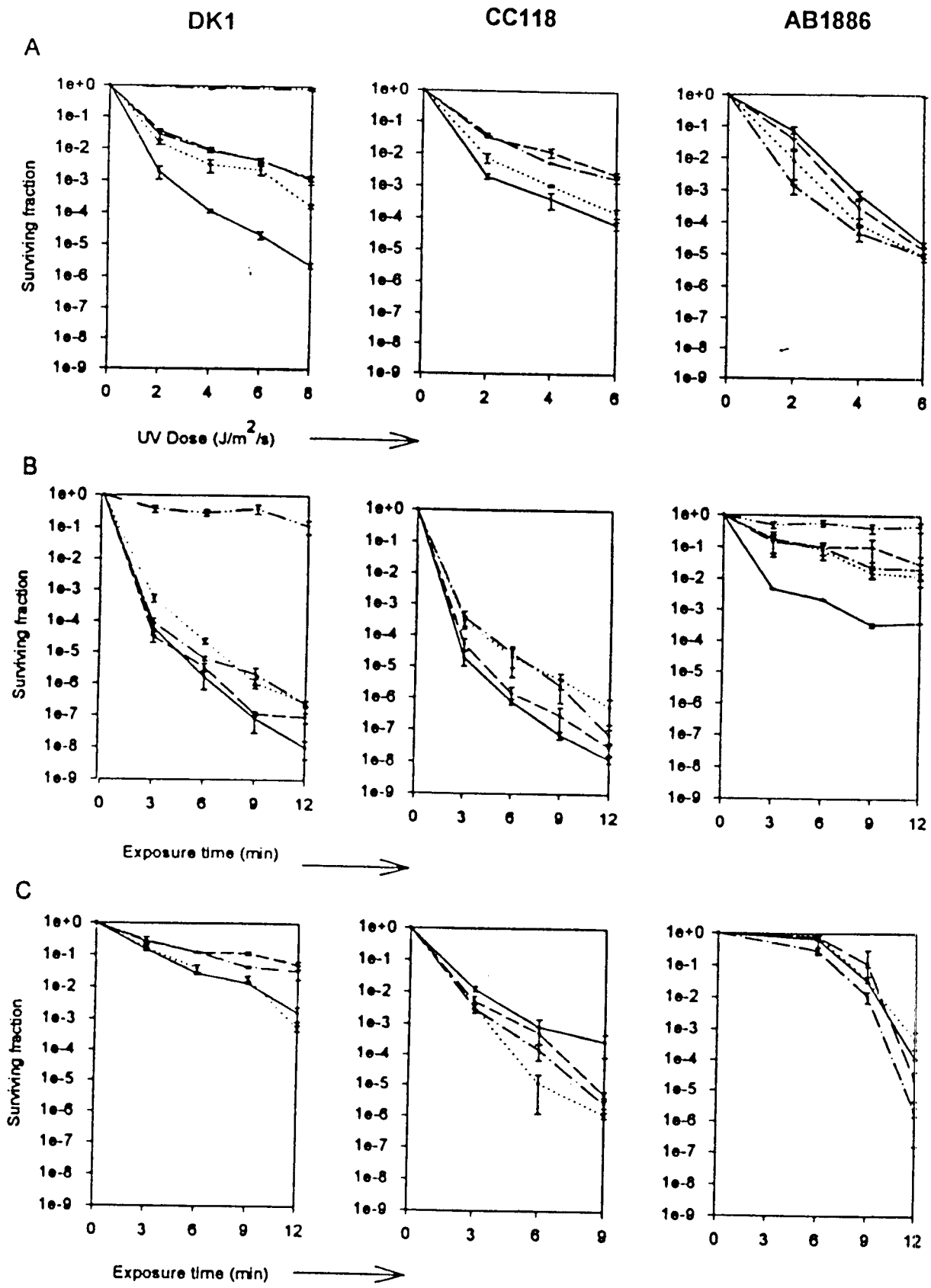


Figure 4.1 Effect of the pMAL™-c2- *katG* constructs on the survival of *E. coli* DNA repair mutants exposed to A) UV, B) mitomycin C, and C) hydrogen peroxide. The UV and hydrogen peroxide survival curves represent the average of 3 experiments, and the mitomycin C survival curves represent the average of two experiments performed on the same day. Error bars were calculated using the standard error of the mean. Curves are shown for DK1 (first column), CC118 (second column), and AB1886 (third column). — pMAL™-c2;– – pSMKA;–·– pSMK5';·····pSMK3'. The survival curves of the AB1884 and AB1885 transformants showed similar trends to those of the AB1886 transformants, and have, therefore, not been included. The survival curves of the parent strains, MC1060 and AB1157 are shown in A and B.

4.4.4. The UV sensitivities of *E. coli* DK1 cells harbouring the pSMK3' derivatives

To determine whether the protective effects observed with the pSMK3' construct were due to the presence of the *M. tuberculosis katG*, or sequences downstream of the gene, *E. coli* DK1 cells harbouring the pSMK3' derivatives, pSMK3'*, pSMK3'a, pSMK3'b, pSMK3'$_{opp}$, and pSMKd, were tested for their ability to survive exposure to short wavelength UV irradiation. The characteristics of these constructs are described in section 3.4.3. Briefly, pSMK3'* contains only the *katG* portion of pSMK3', and no downstream sequences, pSMK3'a and pSMK3'b have bases either removed or added to change the frame of the insert with respect to the maltose binding protein, pSMK3'$_{opp}$, contains the 1747 bp insert in the opposite orientation to that in pSMK3', and pSMKd contains the region downstream of the *katG* gene. Only the constructs pSMK3' and pSMK3'* express the MBP-KatG3' fusion protein. The results of three independent UV sensitivity experiments are represented in figure 4.2. It is evident that only the constructs expressing the MBP-KatG3' fusion, pSMK3' and pSMK3'*, conferred resistance to UV irradiation on the DK1 cells, and that this property is independent of the presence of the downstream sequences.

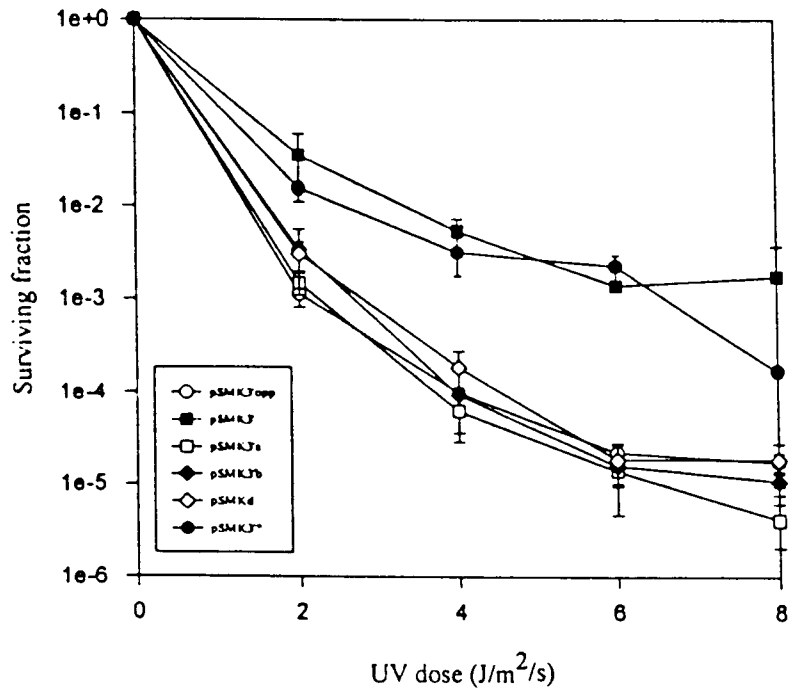


Figure 4.2 UV sensitivity curves of *E. coli* DK1 cells harbouring the pSMK3' derivatives. The curves represent the average of 3 experiments. Error bars were calculated using the standard error of the mean.

4.4.5. Sensitivities of the transformants to mitomycin C

The DK1, CC118, AB1884, AB1885, and AB1886 transformants were tested for their ability to survive exposure to mitomycin C (6 µg/ml) for various time periods. A representative selection of survival curves is shown in figure 4.1.b, and the results are summarized in table 4.2.

Table 4.2 Analysis of the mitomycin C sensitivities of *E. coli* DNA repair mutants transformed with the *katG* constructs.

Strain	Construct	Surv. fraction (12 min)[1]	Rel. surv. x/c[2]
DK1	pMAL™-c2	1.05 (± 0.84) x 10^{-8}	1.00
	pSMKA	9.28 (± 4.82) x 10^{-8}	8.84
	pSMK5'	2.42 (± 0.80) x 10^{-7}	23.05
	pSMK3'	2.61 (± 0.66) x 10^{-7}	24.86
CC118	pMAL™-c2	1.29 (± 0.42) x 10^{-8}	1.00
	pSMKA	3.06 (± 1.71) x 10^{-8}	2.37
	pSMK5'	7.48 (± 5.38) x 10^{-8}	5.84
	pSMK3'	5.74 (± 5.89) x 10^{-7}	44.50
AB1884	pMAL™-c2	1.81 (± 0.01) x 10^{-4}	1.00
	pSMKA	1.07 (± 0.10) x 10^{-2}	59.11
	pSMK5'	3.18 (± 0.59) x 10^{-3}	17.57
	pSMK3'	3.25 (± 0.50) x 10^{-4}	1.80
AB1885	pMAL™-c2	4.53 (± 2.26) x 10^{-4}	1.00
	pSMKA	3.13 (± 0.03) x 10^{-3}	6.91
	pSMK5'	7.38 (± 1.24) x 10^{-3}	16.29
	pSMK3'	4.15 (± 0.19) x 10^{-3}	9.16
AB1886	pMAL™-c2	3.85 (± 0.08) x 10^{-4}	1.00
	pSMKA	2.95 (± 2.37) x 10^{-2}	76.6
	pSMK5'	2.04 (± 1.03) x 10^{-2}	52.99
	pSMK3'	1.22 (± 0.31) x 10^{-2}	31.69

[1] Surviving fraction after 12 minutes of exposure. The mean of three independent experiments ± the standard error of the mean.

[2] Relative survival i.e. the surviving fraction of the construct after 12 minutes of exposure divided by the surviving fraction of the control, pMAL™-c2, after 12 minutes of exposure.

Mitomycin C enters the cell using a specific uptake system, and is metabolically activated within the cell. Once activated, it is very unstable, and rapidly loses its DNA cross-linking ability (Iyer and Szybalski, 1963). As a result, mitomycin C sensitivity experiments showed relatively large

variation, depending on the age of the mitomycin C preparation, and the metabolic state of the cells. There were, therefore, large variations in the surviving fractions obtained in the different experiments, although the shapes of the curves and the relative survival values were similar. The values plotted in the curves have, therefore, been calculated from two independent experiments performed on the same day, to minimize the standard errors, but are representative of the general trend observed in each of three separate experiments.

The growth of repair-competent MC1060 and AB1157 cells was not significantly affected by exposure to mitomycin C at the doses used in these experiments. The DK1 cells harbouring pSMKA, pSMK5' and pSMK3' exhibited 8.8-, 23.0-, and 24.9-fold higher survival, respectively, than cells harbouring the control, pMAL^{TM}-c2, after exposure to mitomycin C for 12 minutes, while the corresponding CC118 transformants exhibited 2.4-, 5.8-, and 44.5-fold increases in survival respectively (Table 4.2). These *E. coli* strains are particularly sensitive to mitomycin C, and it may be that small effects due to the presence of the *katG* constructs are more noticeable on this low background. In general, the AB1884, AB1885, and AB1886 mutants showed a greater natural resistance to mitomycin C than DK1 and CC118 (Fig. 4.1.b, panel 3). This resistance was enhanced by the presence of pSMKA, pSMK5' and, in some cases, pSMK3'. The effect of the constructs on survival was greatest for AB1886. In AB1885, pSMKA and pSMK5' conferred approximately 7- and 16-fold greater survival, respectively, after 12 minutes of exposure, while in AB1884 and AB1886 the resistance conferred was even greater (up to 76-fold). The pSMK3' construct also conferred resistance to AB1885, and especially to AB1886 (9- and 31-fold increase in survival respectively).

4.4.6. Sensitivities of the transformants to hydrogen peroxide

The same transformants discussed above were tested for their ability to survive exposure to hydrogen peroxide (50 mM) for various time periods. A representative selection of survival curves is shown in figure 4.1.c, and the results are summarized in table 4.3. DK1 cells exhibit a high natural resistance to hydrogen peroxide, however, the lack of a shoulder region on the survival curve suggests a lack of an efficient repair system. DK1 cells harbouring pSMK3' exhibited a similar sensitivity to the oxidizing agent to cells harbouring the vector control. The presence of pSMK5' and pSMKA, however, resulted in an increase in the survival of the cells, particularly at longer exposure times. The relative increase in survival after 12 minutes of

exposure was 20- and 30-fold, respectively (Fig 4.1.c, panel 1). In the case of CC118, however, the presence of all three *katG* constructs conferred a disadvantage to the cells, and survival was reduced by more than 100-fold at longer exposure times. Strain CC118 has a different genetic background to the other 4 *E. coli* strains used, and it is possible that, in these cells, the expression of the *M. tuberculosis katG*, or portions thereof, results in negative regulation of the native *katG* gene. This would explain the reduced survival of the cells relative to cells harbouring the vector control, upon exposure to hydrogen peroxide. The hydrogen peroxide survival curves of the excision repair mutants, AB1884, AB1885, and AB1886, had a different shape to those of the *recA* mutants (Fig. 4.1.c, panel 3). Little killing was seen up to approximately 9 minutes of exposure, after which there was a dramatic reduction in survival as the repair mechanisms became saturated. It is evident that no resistance to hydrogen peroxide was conferred by the *katG* constructs in any of these mutants.

Table 4.3 Analysis of the hydrogen peroxide sensitivities of *E. coli* DNA repair mutants transformed with the *katG* constructs.

Strain	Construct	Surv. fraction (9 min)[1]	Rel. surv. x/c[2]
DK1	pMAL™-c2	$1.31 (\pm 0.05) \times 10^{-2}$	1.00
	pSMKA	$1.14 (\pm 0.20) \times 10^{-1}$	8.70
	pSMK5'	$4.24 (\pm 0.35) \times 10^{-2}$	3.20
	pSMK3'	$1.93 (\pm 0.56) \times 10^{-2}$	1.47
CC118	pMAL™-c2	$2.83 (\pm 2.78) \times 10^{-4}$	1.00
	pSMKA	$5.41 (\pm 1.70) \times 10^{-6}$	0.02
	pSMK5'	$2.67 (\pm 1.06) \times 10^{-6}$	0.009
	pSMK3'	$1.19 (\pm 0.35) \times 10^{-6}$	0.004
AB1884	pMAL™-c2	$3.49 (\pm 4.09) \times 10^{-2}$	1.00
	pSMKA	$4.12 (\pm 3.91) \times 10^{-2}$	1.18
	pSMK5'	$5.17 (\pm 2.69) \times 10^{-2}$	1.48
	pSMK3'	$3.42 (\pm 1.34) \times 10^{-2}$	0.98
AB1885	pMAL™-c2	$6.61 (\pm 9.19) \times 10^{-2}$	1.00
	pSMKA	$3.32 (\pm 4.63) \times 10^{-2}$	0.50
	pSMK5'	$1.82 (\pm 1.69) \times 10^{-4}$	0.003
	pSMK3'	$8.56 (\pm 9.81) \times 10^{-3}$	0.13
AB1886	pMAL™-c2	$3.27 (\pm 4.60) \times 10^{-2}$	1.00
	pSMKA	$1.11 (\pm 1.57) \times 10^{-1}$	3.39
	pSMK5'	$1.11 (\pm 0.60) \times 10^{-2}$	0.34
	pSMK3'	$3.73 (\pm 1.14) \times 10^{-2}$	1.14

[1] Surviving fraction after 9 minutes of exposure. The mean of three independent experiments ± the standard error of the mean.

[2] Relative survival i.e. the surviving fraction of the construct after 9 minutes of exposure divided by the surviving fraction of the control, pMAL™-c2, after 9 minutes of exposure.

4.4.7. UV sensitivity tests on different mycobacterial species

A selection of mycobacterial species were tested for their ability to survive exposure to short wavelength UV irradiation. The UV doses used in this experiment were high in comparison to those used in the experiments described above, due to the fact that the mycobacterial species were not mutant in any known DNA repair genes. The mycobacterial strains tested included *M. smegmatis*, *M. intracellulare*, *M. fortuitum*, *M. bovis* BCG, *M. tuberculosis* H37Ra, and *M. tuberculosis* H37Rv. Also tested were a *katG* mutant of *M. intracellulare* and its parent strain, 1403 (Marklund, personal communication), and 4 clinical isolates of *M. tuberculosis*, strain 255, 463, 519 (T. C. Victor, University of Stellenbosch S.A.), and 34 (Heym *et al.*, 1994). Strain 255 has no detectable mutations in the *katG* gene, strain 463 has mutations in codons 463 and 315, and strain 519 has a mutation in codon 315 of the *katG* gene (T.C. Victor, Personal communication). Strain 34 has a complete deletion of the *katG* gene (Heym *et al.*, 1994).

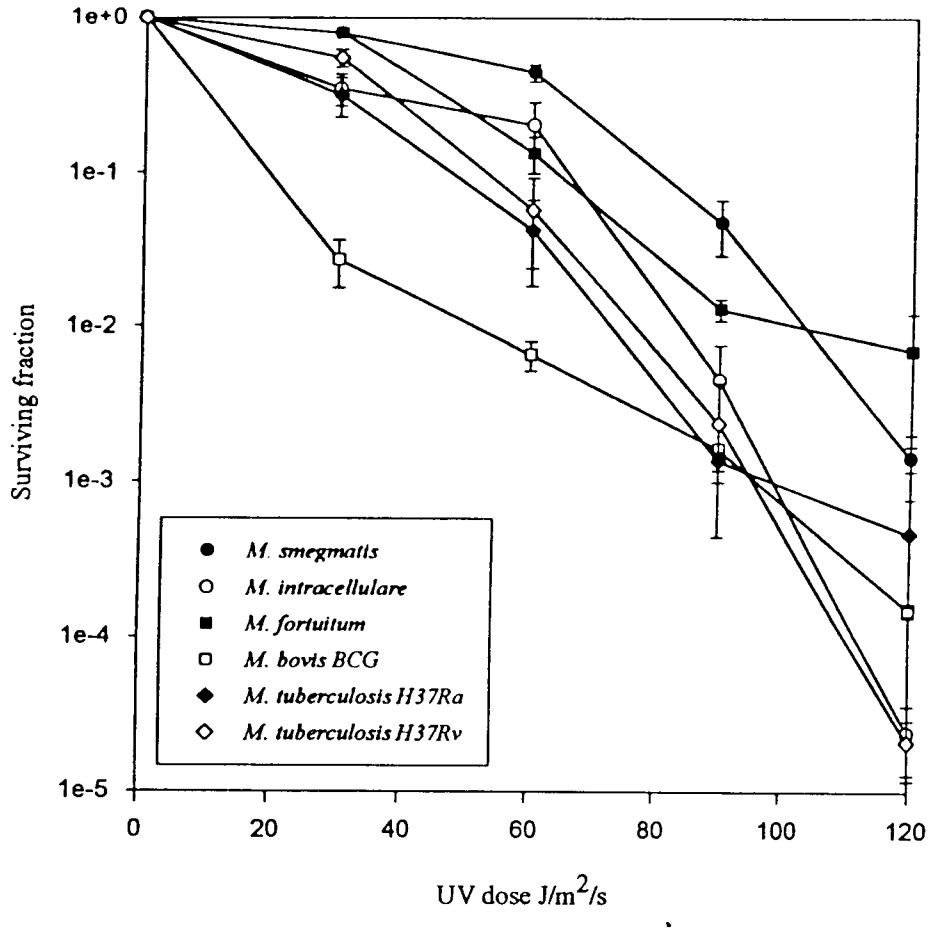


Figure 4.3 UV survival curves of different mycobacterial species. The curves represent the mean of 3 experiments. Error bars were calculated using the standard error of the mean.

In general, the mycobacterial strains exhibited a high natural resistance to UV irradiation. The order of sensitivity was found to be as follows: *M. bovis* BCG > *M. tuberculosis* H37Ra > *M. tuberculosis* H37Rv > *M. intracellulare* > *M. fortuitum* > *M. smegmatis*, although the differences in the sensitivities of these organisms were small. The *katG*$^+$ and *katG*$^-$ *M. intracellulare* strains exhibited no differences in their sensitivity to UV (Fig. 4.4.a). *M. tuberculosis* cells containing no detectable mutations in *katG* (strain 255) exhibited a similar sensitivity to cells containing a mutation in codons 463 and 315 of *katG* (Fig. 4.4.b). These strains both exhibited approximately a 36-fold reduction in survival at a UV dose of 60 J/m^2/s, relative to strain 519 and 34, but exhibited comparable survival to the latter strains at 120 J/m^2/s. Strains containing a mutation in codon 315 or complete deletion of the *katG* gene (strain 519 and 34, respectively) exhibited a higher resistance to UV than the laboratory strains, *M. tuberculosis* H37Ra and *M. tuberculosis* H37Rv.

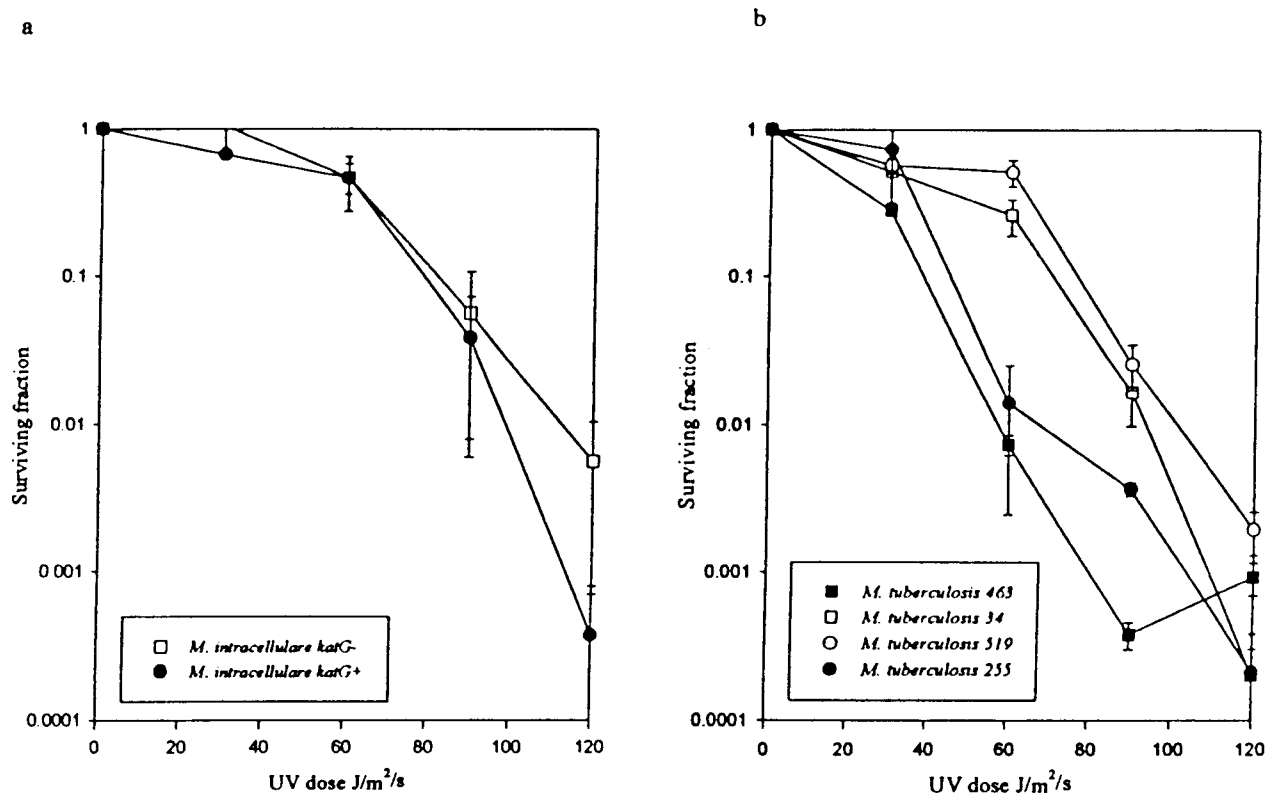


Figure 4.4 UV survival curves of mycobacterial *katG* mutants. The curves represent the mean of 3 experiments. Error bars were calculated using the standard error of the mean. Graph a) *M. intracellulare* 1403 *katG* mutant and parent strains, and graph b) *M. tuberculosis* clinical isolates.

4.5. DISCUSSION

It is evident from the results presented in this chapter that the *M. tuberculosis* KatG plays a role in the protection of DNA repair-deficient mutants of *E. coli* against DNA damaging agents. Possible explanations for these observations are found by examining the DNA repair mechanisms at work in each of the repair mutants in response to the DNA damaging agents.

In general, in repair-competent hosts, DNA damage is repaired as it occurs, due to the presence of efficient repair systems, and the cells survive. However, if the damaging agent is continually administered, the repair mechanisms become saturated, damage accumulation overcomes repair ability, and exponential killing occurs. The presence of functional repair mechanisms in cells, therefore determines the shapes of survival curves. The characteristic shoulder seen in survival curves is due to these efficient repair systems. When a threshold value is reached, however, the repair processes become saturated, and the DNA repair enzymes are progressively inactivated. This results in an exponential drop in cell survival, which is determined by the number of lethal lesions (Haynes, 1966; Word *et al.*, 1991). In repair-deficient hosts, if at least one repair mechanism is present, the survival curves will exhibit a shoulder as well, but the threshold value will be lower as the repair mechanisms are saturated sooner, and exponential killing occurs at lower doses of the damaging agent. The levelling out of the curve at higher doses is due to the induction of alternative repair systems or the saturation of damage sites on the DNA.

The most common structural defects caused by short wavelength UV irradiation are pyrimidine dimers, which block DNA synthesis (Setlow, 1966a; 1966b). Recovery from UV irradiation depends on direct activation of RecA in recombinational repair, and indirect action by derepressing other DNA repair genes in the SOS response. The lesions are repaired primarily by excision repair involving the UvrABC exinuclease, encoded by the *uvrA*, *uvrB*, and *uvrC* genes. In *E. coli* DK1 cells, the *recA* gene is completely deleted and, as a result, no recombinational repair is possible, and no induction of the SOS genes occurs. If these cells are treated with UV, therefore, the only DNA repair that can occur is a small amount of excision repair due to the basal levels of the UvrA, B, and C proteins. This is not sufficient to overcome the damage and exponential killing begins immediately after exposure. The presence of constructs expressing the *M. tuberculosis katG*, however, resulted in a reduced killing and a levelling off of the curve with a high surviving fraction. It is unlikely that KatG replaces RecA

in either the induction of the SOS response or in recombinational repair, which suggests that it may be enhancing the Uvr-dependent nucleotide excision repair process in these cells. In CC118, there is only a point mutation in the N-terminal recombinase domain of *recA*, therefore, although recombinational repair is impaired, induction of the SOS genes is probably unaffected. As a result the excision repair process is more efficient and CC118 cells are more resistant to UV damage. The survival patterns seen in the presence of the constructs are similar to those of DK1, but to a lesser extent, and this may be due to the masking of the positive action of KatG by having a RecA which retains some functions. The *M. tuberculosis* KatG may, therefore, also enhance excision repair in these mutants. In the excision repair mutants, AB1884, AB1885, and AB1886, no Uvr-dependent nucleotide excision repair is possible, therefore damage tolerance mechanisms, such as recombinational repair are responsible for continued cell survival. This repair mechanism is, however, quickly saturated and the exponential drop-off in cell survival occurs at low UV doses. No increase in survival was observed in the presence of the *katG* constructs in these mutants. If KatG does enhance the excision repair pathway, as suggested above, it would not be expected to have an effect on these mutants, since no Uvr-dependent nucleotide excision repair occurs in these cells.

Mitomycin C causes covalent cross-linking of complementary DNA strands which causes inhibition of DNA replication and transcription (Iyer and Szybalski, 1963). These DNA lesions are removed in a similar manner to those caused by UV, i.e. excision repair, except that both strands must be removed, leaving a gap in the DNA which requires replacing by recombination. The order of sensitivity of the mutants used in this study to mitomycin C is as follows: DK1 $\geq$ CC118 > AB1884 = AB1885 = AB1886, indicating that recombinational repair is the most effective means of removing this type of damage. DK1 and CC118 showed a high degree of sensitivity to mitomycin C, and the survival of the cells dropped exponentially immediately after exposure. The presence of the *katG* constructs resulted in enhanced survival. The cells are particularly sensitive to this agent, thus resulting in extremely low colony counts, particularly at higher exposure times. This increases the error margin, and may result in small effects of the *katG* constructs being exaggerated. In these mutants, KatG may enhance the excision repair pathway, however, the lesions resulting from removal of both DNA strands cannot be repaired due to the lack of a functional RecA. As a result, the apparent effect of KatG is small. The AB1884, AB1885, and AB1886 mutants exhibited a much higher natural resistance to

mitomycin C than the *recA* mutants, due to the presence of an efficient recombinational repair system. The standard errors on these curves were, therefore, smaller, and the results were easier to interpret. The presence of the *katG* constructs resulted in a significant increase in cell survival. The effect observed was greatest in the *uvrC* mutant, AB1886, which suggests complementation of defects in this function. It is possible that, in these cells, KatG either enhances the recombinational repair pathway, or that it enhances a Uvr-independent nucleotide excision repair pathway utilized by the cells.

As outlined in Chapter 2, reactive oxygen species attack and damage DNA in a number of ways. The exact mechanisms involved in repair of oxidative damage to DNA are not known but probably involve the oxygen radical scavengers as well as the DNA repair enzymes. It has been demonstrated, however, that, in *E. coli*, the radical-removing enzymes such as the catalases, peroxidases, and superoxide dismutases play a surprisingly small role in protection from killing by hydrogen peroxide (Carlsson and Carpenter, 1980). Thus survival after exposure relies on repair of DNA damage. The excision repair mutants exhibited a far higher resistance to hydrogen peroxide than the *recA* mutants, which suggests that recombinational repair is more important than the nucleotide excision system for the repair of oxidative damage. Survival of the cells in the presence of the KatG fusion proteins was enhanced only in the DK1 strain. This increase in survival was observed with an inactive KatG fusion (pSMK5'), suggesting that it was not due to the removal of toxic oxygen products by the catalase and peroxidase activities of the enzyme. The reasons for the enhancement of survival in DK1, but not CC118 are not clear. The observations suggest that KatG enhances an oxidative damage repair system that is not functional in the latter mutant. In the excision repair mutants, survival up to 9 minutes of exposure to hydrogen peroxide was comparable to that of the parent strain, therefore, any effects due to KatG would be masked.

The fact that resistance to the three DNA damaging agents was not enhanced in all five mutants, suggest that KatG does not increase the general repair capacity of the cell, but rather that it acts on specific pathways. Collectively, the results suggest that KatG may enhance the excision repair pathway in the *E. coli* cells. The possible mechanisms involved in this enhancement are speculative. The KatG may act by stabilizing intermediates in the excision repair pathway, for example the UvrB-[damaged-DNA] complex, which may be vulnerable after excision of the

short oligonucleotide. The effect of the protein is probably not due to direct complementation of the DNA repair genes. The resistance levels of the cells were not restored to wild type levels by the enzyme, therefore complementation was only partial. Another possible explanation for the observed results is that KatG protects DNA from damage rather than stimulating repair. This possibility could be tested in the future by exposure of the cells to the DNA damaging agent, and then induction of KatG. If KatG protects the cells by preventing DNA damage, the cells would be sensitive to the agent, since KatG would be expressed after the damage had occurred. The experiment would also provide evidence as to whether the repair enhancement is direct, i.e. is seen immediately after induction of KatG, or delayed due to the induction of other genes involved in DNA repair.

It is interesting that all three *katG* constructs showed similar effects in the various mutants in response to DNA damage. If the "repair enhancing" activity was encoded in the 5' portion of the gene, one would expect only pSMKA and pSMK5' to confer resistance, and that pSMK3' would have no effect. This is not the case, suggesting that both domains of the protein can carry out the function independently, although the 3' domain is less effective. A possible explanation for this is the duplication seen in the *katG* gene. The N- and C-terminal sequences are similar, and it has been suggested that the C-terminal arose from a gene duplication event (Welinder, 1991). It is possible, therefore, that both domains encode this function. The presence of both domains, however, did not have an additive effect, since mutants harbouring pSMKA generally exhibited similar resistances to those harbouring pSMK5'. A possible explanation for the activity being encoded by both domains of the protein is that the effect of the protein on the sensitivity of cells to DNA damaging agents is due to saturation of the intracellular iron pool, since both the 5' and 3' portions of KatG contain contact sites for heme. The fact that both the 5' and 3' regions of the KatG protein could produce the protective effects also suggests that the effects are not due to the removal of reactive oxygen species produced as a result of the DNA damaging agents, since neither of these domains of the protein are active on their own. This demonstrates that the activity observed in this study is independent of the catalase and peroxidase activities of the enzyme, and implies that the effects may be due to a novel, as yet undefined activity of the *M. tuberculosis* enzyme.

The question that is raised by this study is, why would *M. tuberculosis* have evolved an additional activity in its *katG* gene involved in DNA repair? As mentioned in Chapter 2, *M. tuberculosis* has some unusual features with respect to DNA repair and oxidative stress mechanisms, and it is possible that it has evolved this activity in order to compensate for these differences. A review of the literature on the *E. coli katG* gene, however, suggests that this feature may not be unique to the *M. tuberculosis katG* gene. A number of links have been observed between the presence of KatG in *E. coli* cells and their resistance to various DNA damaging agents. Hoerter *et al.* (1989) showed that, in an *E. coli* strain lacking both the Mn- and Fe-superoxide dismutase, near-UV (300-400 nm) radiation, which is normally only weakly mutagenic, induced a 9-fold increase in mutation relative to wild type cells, which was reduced to wild type levels when a plasmid carrying the *katG* gene was introduced into the double mutant. The same effect was seen in triple mutants which also lacked exonuclease III. The effect may have been due to the removal of toxic oxygen species by KatG, produced as a result of the irradiation, however, the indirect effects of ionizing radiation due to radical species contribute only partly to its lethal effects. Other mutagenic lesions in DNA result from the direct transfer of energy to components of the DNA, and the possibility that the *katG* gene enhances the repair of these cannot be excluded. KatF, which is required for the stationary phase induction of *katG*, has also been shown to play a regulatory role in near-UV resistance, and mutations in *katF* result in sensitivity to near-UV and hydrogen peroxide (Sammartano *et al.*, 1986; Mulvey *et al.*, 1990). Eisenstark (1989) has demonstrated that *katG* is induced in response to oxidative agents other than hydrogen peroxide, including near-UV radiation, and this may be due to the induction of *katF* under these conditions.

Greenberg and Demple (1988) observed that mutations which caused overproduction of KatG in Δ*oxyR* mutants resulted in an increase in resistance of the cells to redox cycling agents, such as paraquat and menadione, and to N-ethylmaleimide, and reduced the spontaneous mutation rate approximately 10-20-fold. The authors concluded that the scavenging enzymes play an important role in protecting aerobic cells from spontaneous mutations due to intrinsic oxidative damage, however, the effects seen may again be due to enhancement of DNA repair processes in addition to the removal of reactive oxygen species.

Farr *et al.* (1988) have reported that the induction of *katG* by OxyR is necessary, and sufficient, for the rapid recovery from inhibition of membrane function observed after exposure to hydrogen peroxide. The nature of the damage responsible for the loss of membrane function by hydrogen peroxide is not known. If it is directly due to damage to membrane components, then recovery of these functions would imply the existence of constitutive enzymes able to catalyze the reverse reaction. However only *katG* was induced here, and this suggests a possible role for the product in the reversal of damage. Cells mutant in *katG* also experienced a large decrease in uptake processes after near-UV exposure, however, a strain carrying a *katG*$^+$ plasmid showed considerably less reduction in uptake after near-UV exposure, and uptake quickly recovered after incubation. These observations with respect to the effect of *katG* on the survival or mutation rates of *E. coli* cells exposed to DNA damaging agents have all been suggested by the authors to be due to the production of reactive oxygen species by these agents, and their subsequent removal by the catalase or peroxidase activity of KatG. In the light of the results presented in this study with the *M. tuberculosis katG* gene, however, the above explanations may need to be reassessed. This study has demonstrated that the *katG* gene is actively involved in the survival of DNA repair mutants of *E. coli* exposed to various DNA damaging agents, and that this is independent of its catalase and peroxidase activities. All bacterial catalase-peroxidases are suggested to have arisen due to a gene duplication event (Welinder, 1991). As mentioned previously, the catalase and peroxidase activities are predicted to lie in the N-terminal domain of the protein, and, as yet, no definite function has been assigned to the C-terminal domain. This study suggests that the gene may encode an, as yet undefined, function involved in DNA repair.

The possibility of the *E. coli* KatG playing a role in DNA repair raises the question as to its role in the experiments described in this chapter. In all of the DNA repair mutants of *E. coli* used, the *katG* gene is unaltered. The gene product is, therefore, probably produced at normal levels, and may have its own effect on the survival of cells after exposure to the various DNA damaging agents. This effect will, however, be present in strains carrying the vector control as well as the fusion constructs. Thus what is being measured in these experiments is whether the *M. tuberculosis* KatG, which is present in excess to the *E. coli* KatG, has an additional effect on survival in comparison to cells not expressing this protein. It is possible that overexpression of the *M. tuberculosis katG* results in a decrease in the expression of the native *katG*, for example

in strain CC118, which is more sensitive to hydrogen peroxide in the presence of the *katG* constructs. If the native KatG also has an effect on the survival of the cells, then the decreased levels of the protein in these cells, in the presence of the *M. tuberculosis* KatG, further emphasizes the role played by the latter protein in protection against damage caused by UV irradiation and mitomycin C.

Another point which needs to be addressed is whether the overexpression of the foreign proteins is responsible for the effects observed in the experiments. Kanemori *et al.* (1994), for example, have demonstrated that the production of abnormal proteins in *E. coli* results in enhanced synthesis of heat shock proteins due to an increase in the stabilization of σ^{32}. In addition, Tsui *et al.* (1996) have demonstrated an induction of DNA repair pathways upon induction of the heat shock response. Thus it is possible that the modification of DNA repair seen in *E. coli* upon expression of the *M. tuberculosis katG* gene could be the result of non-specific effects due to the high level expression of a heterologous protein. However, overexpression of the MBP-LacZ fusion and the fusions produced by the pSMK3'a, pSMK3'b, pSMK3'$_{opp}$, and pSMKd constructs did not result in increased resistance to UV light. Only fusion proteins containing the *M. tuberculosis* KatG had this effect. Thus it is unlikely that the observed effects in these experiments are due to the non-specific effects described above.

The mycobacterial UV sensitivity tests in this study were all performed on species of mycobacteria that contain a KatG (see Chapter 1). The order of sensitivity of the different species was found to be as follows: *M. bovis* BCG > *M. tuberculosis* H37Ra > *M. tuberculosis* H37Rv > *M. intracellulare* > *M. fortuitum* > *M. smegmatis*, although the differences in the sensitivities of these organisms were small. All of the survival curves, except that of *M. bovis* BCG, exhibited a smoothly bending shoulder, characteristic of repair-competent hosts. This is due to the efficient repair of DNA damage by the cells. When a threshold UV dose is reached, the repair mechanisms become saturated and the survival of the cells is determined by the number of lethal lesions. The threshold UV doses were different for each of the mycobacterial species tested, but were followed by exponential killing of the cells in all. The reasons for the shoulderless curve of *M. bovis* BCG are unclear. This shape of curve is usually indicative of deficiencies in DNA repair. However, the survival of this species at higher UV doses (90 - 120 J/m^2/s) was comparable to that of the other mycobacterial species tested. This suggests that

these cells have a similar sensitivity to UV to other mycobacterial species, but that the induction or activation of repair mechanisms is delayed in this organism. This strain of *M. bovis* is attenuated and may be impaired in one or more of the enzymes involved in DNA repair. Deficiencies in DNA repair may, in fact, contribute to the attenuation of the organism.

No differences were observed in the UV sensitivities of *katG*$^+$ and *katG*$^-$ strains of *M. intracellulare*. The mutant strain was constructed by allelic exchange, and is, therefore, isogenic with the parent strain, except with respect to the *katG* gene. The results suggest that, in repair-competent hosts (i.e. cells in which no DNA repair genes have been mutated), the *katG* gene does not play a significant role in repair of damage caused by UV irradiation. The *M. tuberculosis* clinical isolates tested in this study are not isogenic, therefore, the effect of mutations in the *katG* gene on sensitivity to UV cannot be directly correlated. Strain 34, for example, has recently been shown to overexpress the AhpC product, due to mutations in the promoter region of the *ahpC* gene (Heym *et al.*, 1997). The increased expression of this product may compensate for the lack of KatG in this strain, thus masking any effects of deletion of the gene, and may be responsible for the higher levels of resistance of this strain to UV relative to the laboratory strains *M. tuberculosis* H37Ra and H37Rv. An *M. tuberculosis* isolate containing a mutation in codon 315 of *katG* also exhibited higher levels of resistance to UV than the laboratory strains. The mutation has been demonstrated to result in a 50% reduction in peroxidase activity (Rouse *et al.*, 1996). This strain may also contain upmutations in the *ahpC* or other genes, to compensate for the reduced KatG activity. The reasons for the lower survival of strains 255 and 463 in the middle range of UV doses used here are not clear. It may be due to more rapid saturation of repair processes. The levelling off of the curves at higher UV doses may be due to the induction of alternative repair processes or saturation of the damage sites on the DNA.

It is evident from the results presented in this chapter that the *katG* gene does not play a significant role in the sensitivity of mycobacterial cells to short wavelength UV irradiation. The experiments were all performed using repair-competent hosts which exhibit a high natural resistance to UV irradiation. Any effects of *katG* on the UV sensitivities of these strains may, therefore, be masked. Extrapolation of the role of *katG* in DNA repair to the mycobacteria will, therefore, require complementation studies with the gene in repair-deficient mycobacterial hosts.

The recent advances in the construction of defined mutants of both the rapid growing mycobacteria and the *M. tuberculosis* complex organisms should facilitate this kind of study in the future.

PART B

CHARACTERIZATION OF THE *M. TUBERCULOSIS KATG* PROMOTER AND REGULATION OF EXPRESSION

CHAPTER 5

INTRODUCTION

CONTENTS Pg.

CHAPTER 5

INTRODUCTION

Time, effort, and resources have been put into the sequencing of the *M. tuberculosis* and *M. leprae* genomes. Results from these projects will provide a great deal of information on both the genetics of the organisms, and their metabolic pathways. However, sequencing data alone will not identify the mechanisms involved in mycobacterial gene regulation. This requires molecular biological input. There are a number of different mechanisms of regulation of gene expression in prokaryotes. The majority of these occur at the level of transcription initiation. A closer look at the promoter structures and sequences, and cellular proteins which are involved in transcription initiation is, therefore, essential to the understanding of gene regulation.

5.1. TRANSCRIPTION INITIATION IN PROKARYOTES

5.1.1. The transcription machinery and promoter structures

Transcription initiation in prokaryotes has been extensively studied, and has been reviewed by McClure (1985) and Busby and Ebright (1994). Transcription initiation requires the RNA polymerase holoenzyme and specific promoter sequences on the DNA. The most widely studied RNA polymerase is that from *E. coli*. It consists of β, β', α and σ subunits. The core enzyme, $\beta\beta'\alpha_2$ (E), contains the catalytic machinery required for synthesis of the RNA transcript. When bound to the sigma factor, E(σ), the enzyme can bind specifically to promoter sites and initiate transcription. This specific binding occurs over a region of 65 bp (von Hippel *et al.*, 1984), and is followed by the formation of an open complex, which results in a topological unwinding of the DNA of 540° (Gamper and Hearst, 1982). The end of the initiation process occurs with the release of the σ factor, and only once the transcription complex has been translocated out of the promoter region, can another RNA polymerase molecule bind and repeat the process. The formation of the open complex is probably what determines the frequency of RNA chain initiation (McClure, 1985). *E. coli* has approximately

3000 molecules of the RNA polymerase core enzyme per cell (Burgess, 1976). Generally genes occur at one copy per cell, and only one enzyme is responsible for transcribing most of them. Some are transcribed once every few seconds, and others only once per generation or less (Hahn *et al.*, 1977). The relative rate of synthesis of a full length RNA product from a given promoter depends on the promoter strength. Many organisms have more than one σ factor, each specific for the initiation of a distinct class of genes. Alternative σ factors are required in many organisms for developmental growth or response to environmental conditions (McClure, 1985).

In *E. coli*, the vegetative σ factor, σ^{70}, is responsible for initiating transcription of the genes required for normal vegetative growth. A consensus sequence has been proposed for *E. coli* σ^{70} promoters, the -35 and -10 hexamers TTGACA and TATAAT, respectively (Rosenberg and Court, 1979), which are usually separated by 17-19 nucleotides. The TTG of the -35 hexamer, and TA---T of the -10 hexamer are highly conserved. In general, divergence from the consensus sequence results in decreased promoter activity. The strength of a promoter is, therefore, determined by the degree of similarity of its -10 and -35 regions to the consensus sequences, which themselves give maximal promoter strength. *E. coli* promoters that bind $E\sigma^{70}$ have at least 2 of the 3 most highly conserved bases at the -10 region, and at least 1 of the 3 at the -35 region. Changes in these highly conserved bases can result in a 100-fold or more decrease in promoter strength, while mutations in weakly conserved bases can result in a 10-fold decrease in strength (McClure, 1985). Although the -10 and -35 hexamers play a central role in expression, there are also some conserved and weakly conserved sequence homologies on either side of these hexamers, which may be important in promoter function. These regions of the DNA probably also interact with the RNA polymerase, and it is likely, in fact, that the entire 60 bp between -50 and +10 are important for transcription. For example, upstream of the -35 region in most *E. coli* promoters, is an 8-10 bp AT-rich region with a conserved A at position -45 (Hawley and McClure, 1983; Staden, 1984). The T at -18, T at -16, and the G at -15 are also weakly conserved. Promoters with spacings between the -10 and -35 regions of 17 are the strongest, however, those with spacings of 15 and 20 bp retain partial function. Of 263 *E. coli* promoters compiled and analyzed by Harley and Reynolds (1987), 92% had a spacing of 17 ± 1 bp between the -10 and -35 regions. Spacings of less than 16 or more than 18 result in the conserved contact points not lying on the same face of the DNA helix. Transcription usually

initiates with a purine (A or G), and a sequence of CAT is preferred around the transcription initiation codon, but no significant homologies are found downstream of +2. The spacing between the -10 hexamer and the transcription start site varies between 4 and 8 nucleotides, but is usually 6 or 7 (Hawley and McClure, 1983). There are, therefore, a number of factors which determine promoter strength, and the additivity rule states that the individual promoter strength contribution of base pairs, spacer length, DNA conformation, and electrostatic binding within the promoter can be added together to obtain total promoter strength (McClure, 1985).

The conformation of the DNA within the promoter region may play an important role in promoter function, and may be determined by specific sequences. A recently discovered third element in some *E. coli* promoters is an AT-rich sequence upstream of the -35 sequence (Ross *et al.*, 1993). Such a sequence, of approximately 20 bp, immediately upstream of the *E. coli rrn* -35 region, called the upstream or UP element, increases activity 30-fold, in the absence of other protein factors. The region adjacent to the UP element contains binding sites for the activator protein, Fis, which is not required for activation of transcription by the UP element, but increases the activity of the promoter by a factor of 10. The UP element must be in the correct orientation relative to the promoter but its position can be varied, provided that the face of the helix phasing is maintained. DNaseI footprinting revealed that the α subunit of RNA polymerase interacts directly with the UP element. Ross *et al.* (1993) found UP elements with the same properties in three other *E. coli* promoters, *rrnB* P2, *leuV* and RNA II. Although all UP elements are AT-rich, no consensus sequence is evident. Ramalingam *et al.* (1995) identified an AT-rich 81 bp sequence between -664 and -584 upstream of the translation initiation codon, which is required for temporal transcription of the *celA* gene of *Dictyostelium discoideum*. The sequence comprises two *cis* elements each active on its own. A factor binds in this region, and is involved in expression of *celA*. Another promoter which requires an AT-rich upstream region for maximal expression, is that of the *Bacillus subtilis spoVG* gene (Banner *et al.*, 1983). Fredrick and Helmann (1997) recently demonstrated that the UP element enhances the initial association of RNA polymerase with the DNA, and that this is independent of the presence of a σ factor. AT-rich sequences are a frequent feature of gram-positive promoters (Graves and Rabinowitz, 1986; Moran *et al.*, 1982). These sequences may be involved in modulating promoter strength.

Another region found to be important in some prokaryotic promoters is that immediately upstream of the -10 region. Keilty and Rosenberg (1987) showed that the T at position -15 and the G at -14 are essential for constitutive function of the lambda (λ) promoter *P*re. Mutations in these positions can be compensated for by the addition of the phage transcriptional activator protein, cII. The authors suggest that the extended -10 region may be an important contact site for RNA polymerase, which is critical for promoter recognition and function in the absence of a consensus -35 recognition sequence. The T and G at positions -15 and -14 are conserved in other promoters which do not require a specific -35 consensus sequence, for example, *gal* P1, *gal* P2, and the *omp*F, and *uvr*A promoters. The P22 *P*ant promoter contains highly conserved -10 and -35 regions. However, in mutants with severe -35 downmutations, promoter function could be recovered by replacement of an A with a G at position -14 (Grana *et al.*, 1985). Therefore, the extended -10 sequence is probably very important to promoters in the absence of conserved -35 sequences (Keilty and Rosenberg, 1987). Sabelnikov *et al.* (1995) showed that a promoter for the *Dpn*II restriction-modification system of *Streptococcus pneumoniae* consists of an extended -10 site (TaTGgTATAAT), with no required -35 site. In fact, *S. pneumoniae* promoters frequently require an extended -10 which can function without a -35 region (Sabelnikov *et al.*, 1995).

Lisser and Margalit (1994) also hypothesized that nucleotides surrounding the -10 and -35 regions and the presence of specific sequence-dependent, physical-chemical, and structural characteristics are required for promoter function. They demonstrated that both the consensus and flanking regions are less stable, more flexible, and show a greater tendency for B conformation, in comparison to random sequences. Curved DNA upstream of the -35 region has been implicated in the function of a large number of *E. coli* promoters, and a correlation has been found between promoter activity and sequence-dependent curvature. It has been demonstrated that sequence-dependent curvature used to replace the upstream binding sites for the catabolite activator protein (CAP) can stimulate transcription (Gartenberg and Crothers, 1991), suggesting a relationship between protein-mediated DNA bending and sequence-dependent DNA curvature in the initiation of transcription. In *B. subtilis*, the position of the curvature in relation to the promoter is essential for activation (McAllister and Achberger, 1989). Nickerson and Achberger (1995) showed that intrinsically curved DNA upstream of the -35 region affects the interaction of the RNA polymerase with supercoiled and relaxed DNA

templates. The authors proposed a model in which curved DNA upstream of the -35 region facilitates binding of the RNA polymerase. It was found that the rotational orientation, and not the linear displacement, was most important in stimulating promoter binding to the RNA polymerase in *B. subtilis* (Nickerson and Achberger, 1995).

DNA supercoiling has also been shown to have a direct effect on certain promoters, both positive and negative (Galán and Curtiss, 1990). It can also indirectly affect transcription initiation, by affecting the binding of ancillary proteins. For example, the CAP-cyclic AMP activated promoters are sensitive to drugs that inhibit DNA gyrase, an enzyme required for the introduction of negative supercoils in an energy-dependent manner (Sanzey, 1979), and supercoiling directly reduces the promoter strength of the *gyrA* and *gyrB* promoters encoding the DNA gyrase enzyme (Menzel and Gellert, 1983). In *E. coli*, transcription by RNA polymerase involves unwinding of the DNA at the site of initiation to form an open complex (McClure, 1985), and this may be facilitated by supercoiling of the DNA substrate (Pruss and Drlica, 1989).

The region upstream of genes also contains information required for translation of the mRNA transcripts into protein. Approximately 5-12 bp upstream of the translation initiation codon lies a purine (A+G) rich sequence of approximately 3-6 bp known as the Shine and Dalgarno sequence. The 3' end of the 16S rRNA is complementary to this sequence, and facilitates binding of the ribosome to the mRNA transcript (Shine and Dalgarno, 1974). Studies in *E. coli* have shown that gene expression also depends on the region downstream of the ATG (Jacques and Dreyfus, 1990). Jacques and Dreyfus (1990) review the properties of mRNAs which determine where, and how efficiently, translation is initiated. On average, translation initiation occurs every four seconds on *E. coli* mRNAs. Initiation rather than elongation is usually the rate-limiting step in translation, and depends on the mRNA (Ray and Pearson, 1975). However, ribosomes are packed close together on mRNAs (Miller *et al.*, 1970), therefore, if translation of certain codons is slow, elongation will become rate-limiting. Both the nucleotide composition, and the precise sequence around the translation start site contribute to the lack of local structure, which allows free accessibility to the ribosome. Certain other sequences may make favourable contacts with the translational machinery, thereby facilitating translation initiation. The tight coupling of transcription and translation is probably required for the prevention of

secondary structures forming in the mRNA, which could interfere with translation (Jacques and Dreyfus, 1990). Although the interaction of the ribosome binding site on the mRNA transcript with sequences of the 16S rRNA is regarded to be essential for translation of prokaryotic mRNA, a number of genes have been identified that encode mRNA lacking 5' leader sequences (Reviewed by Janssen, 1993). The mechanisms by which these mRNAs bind ribosomes and are translated have yet to be elucidated.

5.1.2. Regulation of transcription

Many bacterial promoters are regulated by activator proteins which bind within or upstream of the promoter and interact with the α subunit of RNA polymerase to increase expression (Ishihama, 1993). Raibaud and Schwartz (1984) review the positive control of transcription initiation in bacteria. Activators are usually negatively autoregulated, and are, therefore, only produced in small amounts. Promoters which require activation usually have poor homology at the -35 region, and deviate from the consensus sequence in the first T of the -10 region. All known activator binding sites are located near or upstream of the -35 region (Raibaud and Schwartz, 1984). An example is CAP, which binds to a site centred between positions -61 and -62 on the *lac* promoter and activates transcription. It has been suggested that CAP directs the α subunit of RNA polymerase to the DNA immediately upstream of the -35 region, compensating for the lack of a UP element (Zhou *et al.*, 1993a; 1993b). The location of the CAP binding site differs on different CAP-dependent promoters. It has been shown that the interaction between CAP and the α subunit is the same, and is provided by bending of the intervening DNA (Zhou *et al.*, 1994a; 1994b). Many activators which do not interact with the α subunit bind to sites which overlap the -35 region, for example λ cI, which binds at a site centred at position -42, and activates transcription of *P*rm (Hochschild *et al.*, 1983). Li *et al.* (1994) provide evidence that cI makes direct contact with a region of σ^{70}. The fact that there are at least two separate targets for activators, suggests that some promoters may be regulated by more than one activator. Activators may act by helping to recruit DNA-binding proteins to DNA, thereby supplementing sub-optimal protein-DNA interactions with protein-protein contacts. Therefore an activator must bind to a DNA site adjacent to or within a promoter, and must make energetically favourable protein-protein contacts with the RNA polymerase (Busby and Ebright, 1994).

Repressors bind to specific sites, operators, near the promoter and inhibit transcription initiation. Operators can be located as far as 60 bp upstream and 12 bp downstream of the promoter sequence, although many overlap the promoters. Inhibition is most likely due to a steric blocking mechanism. There are differences in the tightness of binding of repressors to operators. Very tightly binding repressors, such as *E. coli lac* and P22 *mnt*, result in absolute shut-off of chain initiation, while the weaker binders contribute to down modulation (McClure, 1985).

5.2. TRANSCRIPTION INITIATION IN MYCOBACTERIA

Our understanding of the regulation of gene expression in mycobacteria has been impeded by the lack of knowledge on mycobacterial transcription signals. Mycobacterial DNA, as mentioned previously, has a high G+C content. Since the G+C content of an organism's genome affects codon usage and promoter recognition sites (Nakayama *et al.*, 1989; Ohama *et al.*, 1987), it is expected that transcription and translation signals in mycobacteria will differ slightly from those in other bacteria, such as *E. coli*. GTG is often used as a start codon in mycobacteria, possibly due to the higher G+C content, and mycobacterial genes show a relatively high degree of codon bias, reflected by a predominance of G or C at position 3, especially in *M. tuberculosis*. This probably reflects the distribution of tRNA species in mycobacteria. The upstream regions of mycobacterial genes also have a higher G+C content than the corresponding genes from *E. coli* (Bashyam *et al.*, 1996). It has been suggested that mycobacteria have a low transcription rate and a low RNA content per unit DNA (Harshey and Ramakrishnan, 1977), probably due to a low rate of transcription initiation. Understanding the reasons for this low level of transcription, and the possible mechanisms of regulation of gene expression, requires a closer look at the mycobacterial transcription machinery, and the structure of mycobacterial promoters.

5.2.1. The mycobacterial transcription machinery

Several aspects of the mycobacterial transcription machinery have been studied. The *rpoB* and *rpoC* genes encoding the β and β' subunits of RNA polymerase, respectively, from *M. leprae*

and *M. tuberculosis* H37Rv have been sequenced (Honoré *et al.*, 1993; Miller *et al.*, 1994). The *M. tuberculosis* RpoB is 57% identical to that of *E. coli* at an amino acid level. A number of mycobacterial σ factors have also been identified. The *M. smegmatis mysA* and *mysB* genes encode proteins of approximately 52 and 36 kDa, respectively, which are involved in transcription initiation. The amino acid residues important for recognition of the -35 and -10 regions are identical in both, suggesting a similar promoter specificity. Both also show strong similarity to the principal σ factor, HrdB, of *Streptomyces aureofaciens* and *Streptomyces coelicolor*, although MysA is more similar, and is thought to be the principal σ factor (Predich *et al.*, 1995). Doukhan *et al.* (1995) sequenced the corresponding genes from *M. leprae* and *M. tuberculosis*. Between the *mysA* and *mysB* genes is a highly conserved ORF, approximately 1 kb upstream of *mysB*. Its function is unknown, but the very high hydrophobicity of the predicted protein suggests that it may be an integral membrane component with 7 putative transmembrane segments. A similar ORF is found in the same region in *M. smegmatis*. In both *M. tuberculosis* and *M. leprae*, immediately downstream of *mysB* is an ORF with high similarity to DtxR, the iron-binding repressor of iron uptake and production of toxin in *Corynebacterium diphtheriae*, which is discussed in section 5.4.2. Bashyam *et al.* (1996) compared the domains of the *M. smegmatis* MysA and MysB responsible for binding to the -10 (2.4) and -35 (4.2) regions, to those from the principal σ factors of *S. aureofaciens* (HrdB) and *E. coli* (RpoD). The amino acids of the 2.4 region were identical in MysA and HrdB but differed from RpoD in only 3 out of 23 positions, two of which are conserved changes. Therefore, the -10 binding regions are nearly identical, and would be expected to bind similar consensus sequences. There was, however, larger variation in the -35 binding domain. MysA and MysB differed from each other in 18 of the 39 amino acids of the 4.2 region, 12 of which are non-conserved. This suggests that they probably recognize and express genes with different -35 sequences. MysA and HrdB differed by 2 amino acids, but MysA and RpoD differed in 14 positions, 9 of which are non-conserved changes. There is, therefore, probably poor interaction between mycobacterial -35 sequences and the *E. coli* 4.2 region. These observations may explain the fact that mycobacterial promoters function better in *Streptomyces* than in *E. coli* (Kieser *et al.*, 1986). This is not unexpected, since mycobacteria and streptomycetes belong to the same bacterial order, i.e. Actinomycetales.

Mycobacteria may contain multiple σ factors, a situation found in *Streptomyces* (Westpheling *et al.*, 1985), which allow for greater transcriptional flexibility (Bashyam *et al.*, 1996). As

mentioned previously, alternative σ factors are a common means of regulating gene expression in bacteria. Demaio *et al.* (1996) identified and cloned the *sigF* gene from *M. tuberculosis*, which had significant similarity to the SigF sporulation σ factors from *S. coelicolor* and *B. subtilis*, and to SigB, a stress response σ factor from *B. subtilis*. The *sigF* gene was present in *M. avium* and *M. bovis* BCG, but not in *M. smegmatis* or *M. abscessus*. The authors showed that *sigF* mRNA is not present during exponential growth in *M. bovis* BCG, but is strongly induced during stationary phase, nitrogen depletion, and cold shock, and weakly induced during late exponential phase, oxidative stress, anaerobioses, and alcohol shock (Demaio *et al.*, 1996).

Levin and Hatfull (1993) purified the *M. smegmatis* RNA polymerase enzyme and were able to demonstrate efficient transcription initiation *in vitro*. SDS-PAGE analysis of the purified products revealed the characteristic β, β', and α subunits, as well as several minor proteins, which may represent σ factors. The authors were able to demonstrate *in vitro* transcription from the *M. bovis* BCG *hsp60* promoter (Stover *et al.*, 1991), using supercoiled plasmid DNA, and the RNA polymerase preparation. They also present evidence that the *M. smegmatis* RNA polymerase preparation directs transcription from the ColE1 replicon of *E. coli*, although this does not support plasmid replication in *M. smegmatis* (Snapper *et al.*, 1988). This suggests that the preparation contains a form of holoenzyme with a similar specificity to that of the major σ factor of *E. coli*. The *M. smegmatis* transcription machinery appears to use the *hsp60* promoter in a similar manner to BCG, but there may be some differences between the two systems, due to the fact that only one transcription start site is used in *M. smegmatis*. Bashyam *et al.* (1996), on the other hand, demonstrated that the efficiency and specificity of transcriptional recognition is conserved in *M. tuberculosis*, *M. smegmatis*, and *M. bovis* BCG. They suggest, therefore, that *M. smegmatis* may be used as a surrogate host, at least for studying the constitutively expressed *M. tuberculosis* genes. Predich *et al.* (1995) used a similar method to Levin and Hatfull to purify the RNA polymerase from *M. smegmatis*. They used an antibody specific for a conserved region of the *E. coli* σ^{70} to show that two σ factors (40 and 65 kDa) were present in the holoenzyme preparations, but absent after phosphocellulose chromatography. Peptide sequencing of fragments of these two proteins confirmed that the 65 kDa protein corresponded to MysA, and the 40 kDa protein to MysB. The authors were able to demonstrate *in vitro* transcription from linear and supercoiled forms of the heterologous *B. subtilis sin* P3 promoter, and the *M. smegmatis rrnB* promoter, using the *M. smegmatis*

RNA polymerase preparation. The *M. smegmatis* promoter was, however, not transcribed by a *B. subtilis* RNA polymerase preparation. The lack of requirement for a supercoiled template demonstrated here is in contrast to the observations of Levin and Hatfull (1993), that *in vitro* transcription of the *M. bovis* BCG *hsp60* promoter depends on the superhelical state of the template, although the method of template relaxation used in the two experiments differed. The results suggest that the *M. smegmatis* RNA polymerase has a broader specificity than that of *B. subtilis*, although this may be due to the fact that the preparation contains a mixture of two holoenzymes. A higher transcriptional flexibility may be useful in responding to variations in environmental conditions and for complex interactions with the host (Bashyam *et al.*, 1996).

5.2.2. Mycobacterial promoters

Little is known about the structure and activities of mycobacterial promoters. In general, they function poorly in *E. coli* (Das Gupta *et al.*, 1993; Sirokova *et al.*, 1989), although there are exceptions. Heat shock promoters, which are active in *E. coli*, are the most widely studied (Stover *et al.*, 1991). There are some sequence similarities between these and consensus promoters recognised by σ^{70} and σ^{32}. Comparative studies by Kieser *et al.* (1986) have shown that mycobacterial genes are better expressed in *Streptomyces* than in *E. coli*. Many mycobacterial promoters are similar to *Streptomyces* promoters, and are usually GC rich (69-78%) (Janssen and Bibb, 1990). Strohl (1992) compiled a list of 139 promoters from *Streptomyces*, and showed that 21% of them resembled the typical *E. coli* σ^{70} promoters, and most of these directed transcription of housekeeping genes. It is possible that a similar situation occurs in the mycobacteria. A number of mycobacterial promoters have either been used to drive expression of cloned genes in BCG, or have been examined in more detail. These are summarized in table 5.1. The transcriptional start sites of some of these have been mapped. Some resemble the *E. coli* σ^{70}-type promoter but most have a higher G+C content (65-70%) and different consensus sequences. Many are specifically regulated and may not be considered to be typical of mycobacterial promoters. Some of these promoters have been studied in more detail and have shed some light on mycobacterial transcription signals.

Table 5.1 Comparison of mycobacterial promoter sequences.

	-35		-10		R.B.S.		T.S.C.	Reference
M. tuberculosis								
# 65	TTGCCG	17	TTTCAT	9	GGAGGAA	10	ATG	(Shinnick, 1987)
# 10-12K	TTGAGT	17	TAGAGT	160	TGGAGGG	8	GTG	(Baird *et al.*, 1989)
# 38K	GAAATT	11	AAAGGT	5			GTG	(Andersen and Hansen, 1989)
* 35K	TCGACA		TAAATT					(O'Connor *et al.*, 1990)
# 32K	TTGAGA	16	AAGAAT	37	GGGAAGC	4	ATG	(Borremans *et al.*, 1989)
* v.a.p.	TTGACA		TATCAT					(Gupta and Tyagi, 1993)
* gyrB	TCGTGG		CATATT					(Madhusudan *et al.*, 1994)
85A	TTGACT	22	CGCCTG	71			ATG	(Kremer *et al.*, 1995)
or	TACACG	17	CGCCTG	71			ATG	
16S rRNA	TTGACT	18	TAGACT	7				(Ji *et al.*, 1994; Verma *et al.*, 1994)
cpn60	TGCTCA	17	GGCGGC	37			ATG	(Kong *et al.*, 1993)
* rpsL	TTGACC	18	TATTGT	8				(Kenney and Churchward, 1996)
***M. bovis* BCG**								
* 23 K	TCGTCG		TATGTT					(Yamaguchi *et al.*, 1989)
* 18 K	TTGAGG		TACAGG					(Terasaka *et al.*, 1989)
# alpha	TCGACA	18	TATGTT	27	AGGGGCA	5	ATG	(Matsuo *et al.*, 1988)
# MPB64	TCGTCG	17	TATGTT	69	GGAGGA	8	GTG	(Yamaguchi *et al.*, 1989)
* 64K	TTGCCG		TTTCAT					(Thole *et al.*, 1987)
rRNA	TTGACT	15	TATTAG	197				(Suzuki *et al.*, 1991)
hsp60A	TTGCAC	17	TAAGAA	172	CGGAGGA	11	ATG	(Stover *et al.*, 1991)
or		18	AAGAAT					(Dellagostin *et al.*, 1995)
hsp60B	TCGAAC	11	CCCGGT	214	CGGAGGA	11	ATG	(Dellagostin *et al.*, 1995)
mpb70	CCGATC	17	CATCAG	183			ATG	(Radford *et al.*, 1990)
* rpsL	TTGACC	18	TATTGT					(Kenney and Churchward, 1996)
M. leprae								
# 65K	TTGACA	18	TAAAAA	149	GAGGAA	10	ATG	(Mehra *et al.*, 1986)
18K	TTGCAT	18	CTATAT	58	GCGAGGT	8	ATG	
or		19	TATATC					(Dellagostin *et al.*, 1995)
# SOD	TGGGCG	18	GATTAT	59	AGGAAGG	6	GTG	(Thangaraj *et al.*, 1989)
# 36K	TTTCCT	17	TACGTT	140	GGAAGGT	5	ATG	(Thole *et al.*, 1990)
16S rRNA	TTGACT	16	ATTAAT					(Sela and Clarke-Curtiss,1991)
* 28 K	TGGTCA	14	TACCAT	75	GAGGAGA	4	GTG	(Dellagostin *et al.*, 1995)
* rpsL	TTGACC	18	TATTGT					(Honoré *et al.*, 1993)
M. smegmatis								
* asd	TTGCCC	12	CACGAT					(Cirillo *et al.*, 1994)
ask	CCCACG	17	ACGCTG	258			GTG	(Cirillo *et al.*, 1994)
* Acetam.	TTGACT		TATATT					(Mahenthiralingam *et al.*,1993)
* rrnB	TTGACT	17	TAACTT					(Predich *et al.*, 1995)
rpsL	TTGTTT	18	TATTGT					(Kenney and Churchward, 1996)
or	TTCGTC	18	TGTGGT					
M. paratuberculosis								
P_{AN}	TCGACA	17	TACACT	33	AGAAGGAG	6	ATG	(Murray *et al.*, 1992)
M. fortuitum								
bla	TTCAAA	19	TACGCT	8			ATG	(Timm *et al.*, 1994)
Mycobacteriophage M13								
* pKGR1	TCGACA	24	TCATCT					(Ramesh and Gopinathan, 1995)
* pKGR9	TCGTCA	15	TGTTGT					(Ramesh and Gopinathan, 1995)
* pKGR25	TGCACT	13	CACTAT					(Ramesh and Gopinathan, 1995)
* pKGR38	TCGACG	16	TCATCT					(Ramesh and Gopinathan, 1995)
* ORF1	TCGAGG	16	CGAACT					(Ramesh and Gopinathan, 1995)
* ORF2	TCGACA	22	CAATCT					(Ramesh and Gopinathan, 1995)
Mycobacteriophage L5								
71 P1	TCCGCA	19	TATCCT	54	GGAGGGA	3	ATG	(Nesbit *et al.*, 1995)
71 P2	TTGCTA	18	TACATT	161	GGAGGGA	3	ATG	(Nesbit *et al.*, 1995)
Pleft	TTGACA	18	CATTCT					(Nesbit *et al.*, 1995)

The positions of the -35 and -10 regions, the ribosome binding site (R.B.S.), and the translation start codon (T.S.C.) are indicated. The distances between the -35 and -10 regions, the -10 region and the R.B.S., and the R.B.S. and the T.S.C. are given.

\# Promoters identified by Dale and Patki (1990), using the method of Staden (1984) (see Dale and Patki for references). The transcription start sites for these have not been mapped.

* Transcription start sites have not been mapped. The list does not include the *M. smegmatis* and *M. tuberculosis* promoters sequenced by Bashyam *et al.* (1996).

Suzuki *et al.* (1991) identified two possible transcription intitiation sites upstream of the 16S rRNA gene of *M. bovis* BCG, one of which had promoter sequences similar to the *E. coli* consensus sequences, and the other of which contained a *B. subtilis*-like -10 sequence and no defined -35 region (Ogasawara *et al.*, 1983). *In vitro* transcription assays showed that a promoter recognized by the *E. coli* RNA polymerase is situated approximately 200 bp upstream of the gene. The strengths of the *E. coli* and *M. bovis* BCG *rrn* promoters were identical in *E. coli*. The *E. coli* RNA polymerase did not utilize the other putative promoter of the BCG *rrn*. This suggests that the second promoter may be recognized by a specific σ factor not present in *E. coli*. Cirillo *et al.* (1994) identified the transcription start site of the promoter region of the *M. smegmatis ask* gene. The -10 region for this did not correspond well to *E. coli* consensus sequences, although a region further upstream, corresponding to the -10 region of the weaker transcription start site, did show some correlation but was > 20 bp upstream of the transcription start site. The *ask* gene has a long leader sequence which may be important for regulation. There are several regions that may form large hairpin loops which need to be further studied with respect to their role in regulation. Dellagostin *et al.* (1995) identified the transcriptional start point of the *M. leprae* 18 kDa promoter region. Immediately upstream of this are putative -10 and -35 regions which are similar to *E. coli* σ^{70} consensus promoter sequences, except for a mismatch of a conserved T residue of the *E. coli* -10 hexamer. The authors suggest that the *M. leprae* -10 region may be a pentamer. Kremer *et al.* (1995) studied the promoter region of the *M. tuberculosis* 85A antigen gene. They made progressive deletions of the 5' end using nuclease Bal31. All of the deletions gave less expression than the full length fragment. Removal of the first 44 bps resulted in a 40% decrease in promoter activity. Further deletions revealed that the essential region lies between nucleotides -26 and -136, with respect to the initiation codon. Neither the full length nor the truncated forms could drive expression in *E. coli*. The transcriptional start site was located 63 bp upstream of the proposed ATG initiation codon. The -10 region showed 50% homology with the BCG *mpb70* promoter, and 50% homology with the putative -10 region of the *M. smegmatis ask* gene. The 85A antigen promoter -10 region also showed some similarities to some *Streptomyces* promoters which are not expressed in *E. coli*. Two putative -35 regions were identified. The one (17 bp from the -10 region) showed 50% sequence similarity with that of σ^{70} promoters, while the other (located 22 bp from the -10 region) showed 83% identity with the consensus sequence, and was identical to the -35 region of the *M. leprae* and *M. tuberculosis* 16S rDNA promoter regions. Increasing

the distance between the -35 and -10 regions by 64 bp did not affect the activity of the promoter, a characteristic which is similar to *Streptomyces* promoters, although the presence of the -35 region was essential.

Nesbit *et al.* (1995) studied the expression of gene 71 of mycobacteriophage L5 in detail. A number of gene 71-*lacZ* fusion reporter plasmids were constructed, containing differing amounts of the region upstream of the gene 71 translation initiation codon. The largest fusion (2607 bp) and one with 450 bp of upstream sequence were active in *M. smegmatis*, however, when only 244 bp were present, a reduction in activity was observed. Activities of the fusions were also measured in an L5 lysogen of *M. smegmatis*, and found to be only 30-40% of those in a non-lysogen, suggesting that expression may be regulated by prophage-encoded proteins. Three transcription start sites were found, 58, 166, and 267 bp upstream of the initiation codon of gene 71, corresponding to promoters P1, P2, and P3, respectively. P1 and P2 are recognized by *E . coli* RNA polymerase *in vitro*, while P3 is not, and the latter appears to be a weaker promoter. The authors showed that P3 is recognized by the *M. smegmatis* RNA polymerase *in vitro*, and that the sequences between -11 and -35 are required for promoter function. In a thermoinducible and normal L5 lysogen, all three promoters were active, however, induction of the thermoinducible lysogen resulted in a significant decrease in activity in all three. The authors identified another promoter, P_{left}, in the opposite orientation, which is repressed in lysogeny and induced in early lytic growth, and, therefore, may be regulated by gene 71. Significant amounts of P_{left} mRNA were also observed during late lytic growth, and this was shown to be due to the stability of the P_{left} mRNA transcript. The promoter is recognized by the *M. smegmatis* RNA polymerase *in vitro* and, therefore, does not require any phage-encoded activator proteins.

Ramesh and Gopinathan (1995) studied the promoter regions of the temperate, transducing mycobacteriophage 13. They constructed a phage 13 library in a vector containing the promoterless chloramphenicol acetyl transferase (CAT) gene, and isolated a total of 48 promoter-containing clones. Four of the strongest promoters were sequenced, and a -10 and -35 region was identified for each. They suggest that, in general, there is a good sequence similarity between *E. coli* and mycobacterial promoters at the -35 consensus, but significant variation at the -10 region. In *E. coli*, the distance between the -10 and -35 regions varies

between 17 and 19 bp, while in mycobacteria it varies between 13 and 24 bp. The authors used these and other mycobacterial promoter sequences to generate the following probable consensus; -35, T (100%), T (55%), G (100%), A (67%), C (75%), A (50%), and -10, T (70%), A (75%), T (60%), A (60%), A/T (40%), T (75%). However, confirmation of these sequences as -10 and -35 regions has not yet been done for all and, in many cases, the transcription start site has not been mapped. This study is limited to mycobacterial promoters that are active in *E. coli*, which may not be representative of most mycobacterial promoters.

Kenney and Churchward (1996) analyzed the promoter region of the *M. smegmatis rpsL* gene, encoding the ribosomal protein S12. They mapped the transcription start site to a G residue, 199 bp upstream of the translation initiation codon, and possibly a second 80 bp upstream of this. The 267 bp promoter region was tested using the *xylE* gene as a reporter. Deletion analysis and subcloning showed that sequences important for promoter activity lay between nucleotides 200 and 231, upstream of the start codon, and that a second weaker promoter may be present further upstream. Analysis of the region upstream of the transcription initiation site revealed two potential -10 hexamers, one 6 bp upstream of the transcription start site, which had 3/6 nucleotides matching the typical *E. coli* σ^{70} region, and one 9 bp upstream, which had 4/6 matching nucleotides. There were also two potential -35 hexamers, one 17 bp from the first -10 region, which had 2/6 matching nucleotides, and the other 18 bp from the second -10 region, which had 3/6 matching nucleotides. Kenney and Churchward used oligonucleotide-directed mutagenesis to change conserved positions in the -10 and -35 regions (all to a CG base pair). Only one change, T $\rightarrow$ C at position -9, abolished promoter activity, while a T $\rightarrow$ C change at position -10 reduced activity by 50%. No changes in the -35 region affected activity, in fact, a T $\rightarrow$ C change at -34 increased promoter activity by 10%. The results suggest that the second -10 region is recognized *in vivo*. The authors then used random PCR mutagenesis and found 5 different independent promoter mutations that abolished activity *in vivo*. All mutations were in the second -10 region; T $\rightarrow$ C at -14, substitution of either G or C at -3, deletion of T at -12 (or -11), and T $\rightarrow$ A at -9. These results suggest that the -35 region is not required for promoter function. Other promoters that do not require a -35 region have been reported (Chan and Busby, 1989; Keilty and Rosenberg, 1987; Kumar *et al.*, 1993; Ponnambalam *et al.*, 1988). These promoters usually have an extended -10 region, produced by a TGn motif immediately upstream of the -10 region. Kenney and Churchward (1996) note the presence of a TGG

immediately upstream of the *rpsL* -10 region. They found that mutation of these nucleotides resulted in a 15% reduction in promoter activity, and, if combined with mutations at -37 or -38, resulted in a 70% reduction in activity. This indicates that either an extended -10, or a functional -35 region is needed for normal promoter activity. It has been proposed that when a -35 region is absent, the extended -10 region is responsible for binding of RNA polymerase and other steps in transcription initiation (Chan and Busby, 1989; Keilty and Rosenberg, 1987). An alignment of the *rpsL* upstream sequences from *M. leprae*, *M. bovis*, *M. tuberculosis*, and *M. smegmatis*, revealed a 45% conservation in the untranslated mRNA leader sequence, but a 75% conservation of the -10 and -35 sequences (the -10 region was absolutely conserved). There was also sequence conservation around the translation start site, and a 20 bp sequence 140 bp upstream of the initiation codon. Only the *M. smegmatis* promoter contained the extended -10 TGn motif. The *M. bovis hsp60* promoter, P2 (Stover *et al.*, 1991), and the *M. leprae* 16S rDNA promoter (Sela and Clarke-Curtiss, 1991) also show an extended -10 motif. The slow growers, *M. leprae*, *M. tuberculosis*, and *M. bovis* do not have an extended -10 in the *rpsL* promoter, but have 5/6 matches with *E. coli* in the -35 region. This may reflect differences in the primary σ factors of fast and slow growing mycobacteria (Kenney and Churchward, 1996).

Das Gupta *et al.* (1993) constructed an *E. coli*-mycobacterial shuttle vector, pSD7, containing a CAT reporter cassette for studying mycobacterial promoters in a mycobacterial host. The authors made libraries of *M. tuberculosis* H37Rv and *M. smegmatis* in the vector and selected for clones expressing CAT. The frequency of promoter clones was 10-20% for *M. smegmatis* (350 altogether), and 1-2% for *M. tuberculosis* (125 altogether). The strengths of the promoters were determined by assaying for CAT activity. Most of the promoters from *M. tuberculosis* gave CAT activities between 5 and 100 nmol/min/mg protein, while most of the *M. smegmatis* promoters gave much higher activity (> 500 nmol/min/mg protein). None of the *M. tuberculosis*-derived promoters, and only 12% of the *M. smegmatis*-derived promoter plasmids, conferred chloramphenicol resistance in *E. coli*, and all exhibited very low levels of expression in this host (2-100-fold less than in *M. smegmatis*). Therefore, stronger promoters occur less frequently in *M. tuberculosis* than in *M. smegmatis*. This is also consistent with the lower frequency of isolation of promoters from *M. tuberculosis* (Das Gupta *et al.*, 1993), and may be due to a lower rate of transcription, lower RNA content/unit of DNA, and slower

growth rate (Harshey and Ramakrishnan, 1977). Bashyam *et al.* (1996) have analyzed the promoters isolated in this study further. Of these promoters, 75% were active in *M. smegmatis*, and approximately two thirds exhibited 10% or higher CAT activities in *M. smegmatis* than in *E. coli*. This indicates that the *M. smegmatis* transcription machinery can tolerate a larger variety of -35 sequences than the *E. coli* machinery. The authors sequenced 14 of the *M. smegmatis* and 10 of the *M. tuberculosis* promoters, and aligned them on the basis of their transcription start sites. All contained a conserved -10 region at nearly identical positions upstream of their transcription start sites. The conserved sequences were; T (100%), A (93%), T (50%), A (57%), a (43%), and T (71%) for *M. smegmatis*, and T (80%), A (90%), Y (60%), g (40%), A (60%), and T (100%) for *M. tuberculosis*, where Y indicates a pyrimidine base. The first, second, and sixth bases are most strongly conserved, as in *E. coli*. The less conserved bases tend more towards G and C substitutions. None of the -35 regions of the promoters studied here were homologous to the *E. coli* consensus sequence, and none were conserved in the mycobacteria. The authors suggest that the absence of a conserved -35 region is a distinctive feature of mycobacterial promoters. This is in contrast to the results of Ramesh and Gopinathan (1995), but is supported by the results of Sarkis *et al.* (1995), Kremer *et al.* (1995), and Kenney and Churchward (1996), and is also observed in *Streptomyces* species. Strohl (1992) reports that 80% of *Streptomyces* promoters contain diverse sequences in their -35 regions, and do not function in *E. coli*. Deletion analysis in one *M. smegmatis* and one *M. tuberculosis* promoter revealed that the -35 region alone is insufficient to support transcription, and that the -10 region is essential. In 9/14 *M. smegmatis* and 7/10 *M. tuberculosis* promoters, transcription initiated at a purine. Interestingly, the *M. tuberculosis* promoters had a higher G+C content (57%) from positions -1 to -50, with respect to the translation initiation codon, than the *M. smegmatis* promoters (43%), which may have a bearing on the lower strength of the *M. tuberculosis* promoters. The authors suggest that weaker promoters may have an advantage in pathogenicity (Bashyam *et al.*, 1996).

It is evident from the above-mentioned studies that mycobacterial transcriptional signals share some features in common with those from other, unrelated organisms. Most similarities, however, are found with the promoters of *Streptomyces*, a genus which is closely related to the mycobacteria. As yet, too few mycobacterial promoters have been characterized to make accurate predictions on promoter structures and consensus sequences. It is likely that

promoters from *M. tuberculosis* will contain a unique set of features, which account for the very low rate of transcription initiation, and possibly also the slow growth and pathogenicity of the organism.

5.2.3. The regulation of gene expression in mycobacteria

Some evidence exists for gene regulation mechanisms in mycobacteria which are similar to those described in other bacteria. The *M. tuberculosis* 65 kDa and 10-12 kDa heat shock proteins, for example, have potential heat shock promoter sequences which may be recognized by σ^{32} (Cowing *et al.*, 1985), and overlap the σ^{70} sequences. The BCG *hsp60* promoter is expressed constitutively when present on an extrachromosomal vector, but shows an increase in response to stress with acid, heat, and peroxide when present on an integrative vector, although expression is lower (Stover *et al.*, 1991). Levin and Hatfull (1993) found that transcription of the *hsp60* promoter by the *M. smegmatis* RNA polymerase relied on supercoiling of the DNA substrate. The authors suggest that the differences in expression noted by Stover *et al.* (1991) may be explained by the differences in the superhelical state of the DNA, which may play a direct role in the regulation of gene expression (Pruss and Drlica, 1985; 1989; Mark *et al.*, 1988; Galán and Curtiss, 1990). They also suggest that the high G+C content of mycobacterial DNA may be responsible for the need for DNA superhelicity, to unwind promoter regions. Bacterial virulence genes which respond to stress require DNA supercoiling for regulation (Higgins and Dorman, 1990; Higgens *et al.*, 1990; Mekalanos, 1992). Environmental stimuli such as osmolarity, temperature, oxygen levels, and starvation affect DNA supercoiling, and pathogenic organisms may use this system to respond to these stimuli (Galán and Curtiss, 1990).

Dale and Patki (1990) have identified a region that partially resembles a *pho* box, found upstream of genes regulated by phosphate, and a region with homology to the Fur binding site upstream of the *M. leprae* 28 kDa antigen. This is not a perfect match but is situated in the correct position, i.e. it overlaps the -35 region. Some evidence exists for the 28 kDa antigen being involved in iron uptake (Cherayil and Young, 1988; Hall *et al.*, 1987; Barclay and Wheeler, 1989), and a gene involved in iron regulation has been identified in *M. tuberculosis* (discussed in section 5.4.2). Other regulated genes in mycobacteria include glycerol 3-phosphate dehydrogenase for the utilization of glycerol, which is induced by glycerol (Segal

and Edwards, 1984); aspartokinase, which is repressed by threonine and methionine; homoserine dehydrogenase, which is repressed by threonine and valine (Karasevich and Butenko, 1966, cited by Segal and Edwards, 1984; Sritharan *et al.*, 1989; 1990); and amidase, which is induced by acetamide (Draper, 1967). As mentioned in Chapter 2, the promoter of the *M. tuberculosis recA* gene has recently been shown to be inducible in response to a variety of DNA damaging agents. This induction involves the LexA protein (Movahedzadeh *et al.*, 1997a; 1997b). There is also some evidence for the induction of some mycobacterial genes during *in vivo* growth in macrophages. Dellagostin *et al.* (1995) showed that the *M. leprae* 18 kDa promoter shows high levels of expression in BCG cells which have infected macrophages. This is consistent with the fact that the 18 kDa protein is a major antigen in natural infection, and suggests upregulation of the gene during infection. This was recently confirmed by McBride *et al.* (ASM Abstract U-103, 1997). These authors determined the minimum sequences required for *in vitro* expression and *in vivo* upregulation of the promoter. Post-translational modification has been studied to a small extent in mycobacteria. Several secreted proteins and possible signal peptide sequences have been identified, and there have been some studies on biotinylated proteins (Takayama and Qureshi, 1984; Collins *et al.*, 1990; Young *et al.*, 1985; Moss, 1987). These observations indicate that the mycobacteria possess some mechanisms of gene regulation in common with other bacteria. It is likely, though, that these organisms also contain a unique set of regulatory mechanisms which are responsible for some of their unique properties.

5.2.4. Tools for studying gene expression in mycobacteria

The significant increase in the information available on the regulation of gene expression in the mycobacteria over the past decade has been facilitated by the development of molecular tools and techniques for the genetic manipulation of these organisms. These methods are reviewed by Jacobs *et al.* (1991). *E. coli* and *Streptomyces lividans* have been used to study the expression of mycobacterial genes, with variable success (Jacobs *et al.*, 1986; Kieser *et al.*, 1986; Labidi *et al.*, 1985b; Thole *et al.*, 1985). It is evidently more ideal to study the expression and regulation of these genes in a homologous system, in which response to physiological constraints is expected to be similar. Molecular biological approaches to studying mycobacteria have been facilitated by the development of efficient electroporation methods for mycobacteria

(Snapper *et al.*, 1988; 1990a; 1990b), the isolation of high transformation frequency mutants of *M. smegmatis* (Snapper *et al.*, 1990a), and the characterization and use of the pAL5000 plasmid of *M. fortuitum*, which replicates autonomously in mycobacterial species (Labidi *et al.*, 1985a). The *oriM* of pAL5000 is located on a 1.8 kb fragment, and can direct plasmid replication in *M. smegmatis* and BCG (Stover *et al.*, 1991). Another naturally occurring mycobacterial plasmid is pMSC262, isolated from *M. scrofulaceum* (Goto *et al.*, 1991). The minimum replication region of this plasmid is contained in a 1.5 kb fragment, which has no homology with the replication region of pAL5000. Plasmid pMSC262 is compatible with pAL5000 and, like the latter, requires different regions for replication in *M. fortuitum* and *M. smegmatis* (Qin *et al.*, 1994). These plasmids have been used to construct a number of different vectors for use in the mycobacteria, many of which have possible applications in vaccine development. Some of these are reviewed briefly below.

Stover *et al.*, (1991) constructed a vector containing the following: the 1.8 kb *oriM* of pAL5000, in which some restriction endonuclease sites had been mutated using PCR mutagenesis; the *E. coli oriE* from pUC19; the kanamycin resistance gene from Tn*903*; and an expression cassette containing the BCG *hsp60* promoter, including the first 6 amino acids of this protein (Thole *et al.*, 1987), a multiple cloning site, and the *rrnAB*t1 transcriptional terminator (Sarmientos *et al.*, 1983). The resultant vector, pMV261, has a copy number of approximately 5 in BCG. To ensure that these plasmids are stably maintained in BCG, the authors replaced the *oriM* of pMV261 with a DNA fragment containing the attachment site (*attP*), and integrase gene (*int*) from the mycobacteriophage L5 (Lee *et al.*, 1991), to form the vector pMV361. The vector becomes stably integrated into the BCG host chromosome at the *attB* site, and is compatible with *oriM*-containing plasmids. The authors used the vectors to drive expression of antigen genes from a variety of pathogens, including HIV. Introduction of these into BCG resulted in significant levels of expression of the recombinant antigens, and a greater antibody response than to the antigens of BCG itself. An advantage of these recombinant BCG is that they can provide continuous boosts to the foreign antigen from a single immunization. Other similar vectors have been constructed using different transcription signals and antigens (Matsuo *et al.*, 1990; Aldovini and Young, 1991; Snapper *et al.*, 1988; Ranes *et al.*, 1990; Dellagostin *et al.*, 1993). A number of vectors have also been constructed for studying promoter sequences. These usually contain the following: a reporter gene with no promoter, ribosome binding site or

initiation codon; the *aph* gene from Tn*903* or Tn*5*, encoding resistance to kanamycin; the *E. coli oriE* from pUC18; and the *oriM* from pAL5000 (Barletta *et al.*, 1992; Das Gupta *et al.*, 1993; Curcic *et al.*, 1994).

Another useful tool for studying the genetics of mycobacteria and mycobacterial gene expression is the mycobacteriophage. Phage vectors, called shuttle phasmids, have been constructed from the mycobacteriophage TM4 and L1 (Jacobs *et al.*, 1987; Snapper *et al.*, 1988). These can replicate in mycobacteria as phage (undergo lysis and lysogeny), and in *E. coli* as cosmids, and can be packaged into λ phage heads to facilitate the cloning of additional genes (Jacobs *et al.*, 1989). They normally contain an *E. coli* plasmid, often a cosmid, inserted in a non-essential region of the phage genome. Shuttle cosmids, such as pYUB18, have also been constructed with the following features: *E. coli* and mycobacterial origins of replication; a kanamycin gene functional in *E. coli* and mycobacteria; a phage λ *cos* sequence, which allows packaging into λ phage heads; and unique restriction endonuclease sites for the construction of libraries. These cosmids allow the cloning of very large DNA fragments, and have been used extensively for constructing mycobacterial genomic libraries (Jacobs *et al.*, 1991).

Insertional mutagenesis is useful for the selection of mutants arising from a single mutagenesis event directly. Guilhot *et al.* (1994) were the first to report on the construction of mycobacterial insertional mutant libraries by transposition directly into *M. smegmatis*. They used transposon Tn*611* in a temperature-sensitive plasmid isolated from an *E. coli*-mycobacterial shuttle vector, containing the pAL5000 and pUC18 replicons. This system may be useful for the study of *M. smegmatis* genetics, and possibly those of other mycobacteria too. Random shuttle mutagenesis has been developed by Kalpana *et al.* (1991). It involves firstly cloning the desired gene into a vector that replicates in *E. coli*, but not in the organism of interest, secondly introduction of a selectable marker on a transposon into the gene of interest, and thirdly homologous recombination to return the mutagenized gene into the chromosome of the organism of interest. The authors used Tn*5* sequences containing the *neo* gene, which confers resistance to kanamycin in *E. coli* and mycobacteria, and hyper-resistance to neomycin. The transposon transposes preferentially into GC-rich sequences.

A number of different reporter genes have been used for studying expression from mycobacterial promoters. Curcic *et al.* (1994) and Kenney and Churchward (1996) used the *xylE* gene from *Pseudomonas putida* as a reporter for studying mycobacterial gene expression. The gene product, catechol 2,3-dioxygenase, converts catechol, a small aromatic compound, into a bright yellow product, which is easily measured. The advantages of *xylE* are: activity can be rapidly detected by spraying bacterial colonies with catechol; this assay does not alter the viability of the cells; the catechol is inexpensive, and penetrates the lipid-rich cell wall of mycobacteria, therefore analysis can be done on whole cells; the *xylE* gene product is stable; the gene derives from a GC rich organism and, therefore, may have a similar codon usage to mycobacteria; and the assay is highly reproducible (Curcic *et al.*, 1994). The *phoA* expression-secretion system is very useful for evaluating the efficiency of mycobacterial promoters and signal peptides, and has been used to study the promoter region of an *M. fortuitum* β-lactamase gene in *M. smegmatis* (Timm *et al.*, 1994), and the 85A antigen promoter from *M. tuberculosis* (Kremer *et al.*, 1995). The gene product is only functional if secreted, therefore, Lim *et al.* (1995) used the reporter in a vector, pJEM11, for isolating genes encoding exported proteins from mycobacteria. Another frequently used reporter is the *E. coli* β-galactosidase enzyme. It is easily detected, and is known to induce an antibody and T-cell proliferative response in the presence of IFA. It is, therefore, useful for studying the *in vivo* expression of mycobacterial promoters (Stover *et al.*, 1991; Murray *et al.*, 1992; Barletta *et al.*, 1992). The green flourescent protein (GFP) of the jellyfish *Aequorea victoria* has been used by Dhandayuthapani *et al.* (1995) for studying expression of the *ahpC* promoters from *M. tuberculosis* and *M. leprae*, *mtrA*, a response regulator from *M. tuberculosis*, *hsp60* from *M. bovis* BCG, and *tbprc3* from *M. tuberculosis*, in *M. smegmatis* and *M. bovis* BCG hosts. The reporter requires no exogenously added substrate or co-factors, only irradiation with blue light. The authors used epiflourescence microscopy, laser scanning confocal microscopy, flourescence spectroscopy, and flow cytometry to monitor expression. They were able to identify bacteria expressing GFP from a strong promoter, and to localize and isolate phagocytic vesicles containing mycobacteria. It may therefore be a useful tool for studying the interactions of mycobacteria with macrophages (Dhandayuthapani *et al.*, 1995). The chloramphenicol acetyl transferase gene (CAT) has also been used extensively as a reporter (Sela and Clarke-Curtiss, 1991; Das Gupta *et al.*, 1993; Verma *et al.*, 1994; Ramesh and Gopinathan, 1995). Selection for promoters driving expression of this gene is simple and rapid, and involves growing on chloramphenicol.

The luciferase enzyme is another useful reporter for studying gene expression. The luciferase genes from *Vibrio fisheri*, *Vibrio harveyi*, and the firefly, *Photinus pyrulis*, have been used for this purpose (Engelbrecht *et al.*, 1985; Legocki *et al.*, 1986; Schmetterer *et al.*, 1987; Carmi *et al.*, 1987). The bacterial luciferase system requires up to 7 genes in the *lux* operon, however, luminescence can be obtained from just the two subunits of the luciferase enzyme, encoded by the *luxA* and *luxB* genes (Meighen, 1988). In luminous bacteria, the luciferase enzyme catalyzes the following reaction: $RCHO + FMNH_2 + O_2 \rightarrow RCOOH + FMN + H_2O + h\nu$ (490 nm), in which R is an aliphatic moiety containing at least 7 carbon atoms, FMN is a flavin mononucleotide, and $FMNH_2$ is its reduced form. One luciferase molecule gives rise to approximately one photon of light, therefore very small amounts of luciferase can be detected using a luminometer. The reaction, however, requires the addition of exogenous substrates such as decanal. The substrate is generally added in excess to ensure that it is not limiting, however, it becomes strongly inhibiting at high concentrations, therefore complicating the interpretation of results. Other problems with decanal are that it is highly volatile, poorly soluble in aqueous systems, lipophilic, and is cytotoxic at higher concentrations (Blouin *et al.*, 1996; Olsson *et al.*, 1988). The bacterial luciferase has been used as a reporter in the mycobacteria (Gordon *et al.*, 1994).

The firefly luciferase enzyme, which catalyzes the oxidation of luciferin in the presence of ATP, has also been used as a reporter in the mycobacteria. Hickey *et al.* (1996) constructed vectors expressing high levels of luciferase in *M. bovis* BCG, and used them to monitor the *in vivo* growth of the recombinant BCG strain in mice. The authors were able to detect activities of standard antimycobacterial drugs directly in mice by reductions in spleen luminescence. Similarly, Arain *et al.* (1996) used strains of *M. tuberculosis* and *M. bovis* BCG expressing the firefly luciferase to assess the activities of INH and rifampin against mycobacteria in human macrophages. Sarkis *et al.* (1995) introduced the firefly luciferase gene into the tRNA region of the mycobacteriophage L5 genome. When used to infect *M. smegmatis*, these recombinant phages resulted in expression of the luciferase gene and light emission. The gene is fortuitously expressed continuously in lysogens surviving infection. These authors, again, suggest the use of these reporter phages for discrimination between drug-sensitive and drug-resistant strains of *M. smegmatis*, and possibly for rapid identification and classification of antimycobacterial agents.

5.3. THE REGULATION OF CATALASES IN PROKARYOTES

Outlined above is an overview of the information available on the regulation of gene expression in the mycobacteria, and the tools available for further studies. Many of these have been applied in this study to the examination of the *M. tuberculosis katG* promoter and its mechanisms of regulation. Chapter 1 outlines the importance of the *M. tuberculosis* KatG in INH sensitivity and in virulence. It would, therefore, be useful to improve our understanding of the way in which the gene is regulated in these organisms. This is the focus of the next few chapters. As mentioned in Chapter 1, a number of INH-resistant clinical isolates of *M. tuberculosis* have been described which have reduced catalase and peroxidase activity, but have no mutations in the *katG* gene. This suggests that they may have mutations in an unlinked regulatory gene. In order to provide clues as to the mechanisms of regulation of the *M. tuberculosis katG* gene, the regulation of catalases in other bacteria is reviewed.

5.3.1. The oxidative stress response

Central to the regulation of catalase expression in many bacteria is the oxidative stress response. Active oxygen molecules are produced as a byproduct of oxygen metabolism, and their production is enhanced by exposure to certain environments or by dietary or disease conditions. These radicals can oxidize membrane fatty acids and proteins, and damage DNA. When the degree of oxidative stress is greater than the ability of the cell to defend against it, oxygen toxicity results. Much progress has been made in understanding bacterial responses to oxidative stress, particularly in *E. coli* and *S. typhimurium*, and is reviewed by Farr and Kogoma (1991).

Molecular oxygen (O_2) has two unpaired electrons in its molecular orbitals with the same spin quantum number (parallel spins), i.e. is spin restricted, and cannot easily accept a pair of electrons. It is, therefore, unreactive with most compounds, except radicals. Molecular oxygen is a poor acceptor of one electron, but in the presence of metal catalysts, one-electron reduction can occur. Its reactivity increases with acceptance of 1, 2, or 3 electrons to form the superoxide radical ($O_2^{\cdot}$), hydrogen peroxide (H_2O_2), and hydroxyl radical ($OH^{\cdot}$), respectively. A 4-electron reduction results in the formation of water. Under acidic conditions, molecular oxygen can be protonated to form the hydroperoxyl radical, $HOO^{\cdot}$ (Bielski, 1985), which is much more

reactive than O_2^-. The superoxide radical oxidizes thiols and ascorbate (Hill and Allen, 1978), attacks proteins containing $(Fe\text{-}S)_4$ clusters (Gardner and Fridovich, 1991), and reduces transition metals and metal complexes. It spontaneously dismutates to H_2O_2 and O_2. Hydrogen peroxide acts as a weak oxidizing agent, and will attack thiol groups of proteins or reduced glutathione, and can react with some keto acids (Wefers and Sies, 1983). It reacts with reduced iron or copper ions in the Fenton reaction to generate hydroxyl radicals ($OH^{\cdot}$), which react with most molecules in the cell at diffusion-limited rates (Singh and Singh, 1982). Iron-oxygen complexes may also be important in oxidative damage. A superoxide anion may form a complex with Fe^{3+} to yield $[2Fe^{2+}O_2]$, which undergoes a series of reactions to form the ferryl radical $[2Fe^{2+}O]$ (Aust *et al.*, 1985), which is electron rich and has radical characteristics. Active oxygen species therefore arise from a variety of intracellular and extracellular sources, and can cause extensive damage to cells if not removed.

Organisms which grow under aerobic conditions have developed a number of defences against these active oxygen species. These include, enzymes which degrade these species, such as superoxide dismutases, which dismutate O_2^- to H_2O_2, catalases and peroxidases, which break down H_2O_2 to water and oxygen, and alkylhydroperoxide reductase, encoded by *ahpC* and *ahpF*, which reduces various organic hydroperoxides to their corresponding alcohols (Christman *et al.*, 1985; Tartaglia *et al.*, 1989). DNA repair enzymes, e.g. exonuclease III, DNA polymerase I, the RecBC nuclease, and RecA, repair damage to DNA caused by oxygen radicals that escape removal by the scavengers. As mentioned previously, when the concentration of active oxygen exceeds the basal level of the scavenging capacity of the cell, an oxidative stress results. Bacteria then invoke one or both of two known stimulons, the peroxide stress response, and the superoxide stress response. These responses are distinct and differ slightly in *E. coli* and *S. typhimurium*. The superoxide stress response involves the induction of approximately 30 proteins, at least 6 of which are regulated by SoxR and SoxS (Greenberg *et al.*, 1990; Wu and Weiss, 1991), in response to superoxide-generating agents, such as paraquat (Hassan and Fridovich, 1977). Other induced proteins include two associated with the HPI catalase, two proteins of Ahp, heat shock protein GroEL, and endonuclease IV, resulting in an increased DNA repair capacity (Farr *et al.*, 1985). The first enzyme identified as being induced is MnSOD (Gregory and Fridovich, 1973), which dismutates O_2^- to hydrogen peroxide, which is then thought to induce the peroxide stress response.

5.3.1.1. The peroxide stress response and OxyR

When bacteria are pretreated with low doses of hydrogen peroxide, they become resistant to subsequent doses that would otherwise be lethal. This is due to the induction of at least 30 proteins, including RecA (Greenberg and Demple, 1989). Other organic peroxides, such as *tert*-butyl hydroperoxide and cumene hydroperoxide, induce a similar response. Adaptation to hydrogen peroxide in *S. typhimurium* requires *de novo* protein synthesis, and is independent of RecA. There are two temporal classes of hydrogen peroxide inducible proteins, as revealed by pulse-labelling with ^{35}S-methionine. The early proteins, of which there are 12, are synthesized maximally during the first 10 minutes, while the late proteins, of which there are 18, are synthesized maximally 10-30 minutes after treatment with hydrogen peroxide. In both *E. coli* and *S. typhimurium*, at least 9 of the peroxide-induced proteins are positively regulated by OxyR. These include: two heat shock proteins; two electromorphs of HPI, encoded by the *katG* gene; the two subunits of Ahp, encoded by the *ahpCF* operon; glutathione reductase, encoded by *gorA*; *dps*, a 19 kDa DNA binding protein from starved cells, which has been implicated in the protection of DNA from oxidants; *oxyS*, which is a small, untranslated RNA, divergently transcribed from *oxyR* in *E. coli*; and superoxide dismutase (Christman *et al.*, 1985; Tartaglia *et al.*, 1989; Altuvia *et al.*, 1994; Toledano *et al.*, 1994; Storz *et al.*, 1990; Farr and Kogoma, 1991; VanBogelen *et al.*, 1987). Among the proteins induced by hydrogen peroxide independently of OxyR are some involved in DNA repair. The negatively autoregulated *oxyR* gene encodes a 34 kDa protein (Christman *et al.*, 1989), with significant homology to the LysR family of bacterial regulatory proteins (Henikoff *et al.*, 1988). Induction of the OxyR regulon by hydrogen peroxide does not involve an increase in OxyR expression. Instead, upon oxidative stress, OxyR is oxidized to a form which then activates expression of the OxyR regulon (Storz *et al.*, 1990). Both the reduced and oxidized forms of OxyR bind to the regulatory regions of *ahpCF*, *katG* and *oxyR*, but in a different manner. Storz *et al.* (1990) suggest that the conformational change resulting from oxidation alters the way in which it interacts with RNA polymerase. The precise mechanisms involved, however, have not yet been elucidated.

The OxyR-binding sites upstream of the *katG*, *ahpCF* operon, and *oxyR* genes of *E. coli* have been identified by DNase I footprinting. The *ahpC* site lies between -79 and -35, and that of *katG* lies between -76 and -35. In the *katG* promoter, 14/20 bases match the consensus

sequence recognized by the oxidized form of OxyR (Toledano *et al.*, 1994). The OxyR-binding sites extend into the -35 σ^{70} binding site, suggesting that OxyR interacts with the RNA polymerase to activate transcription. Members of the LysR family of transcriptional regulators usually bind to a site containing a core T-N_{11}-A motif, embedded within a larger dyadic sequence (Goethals *et al.*, 1992; Schell, 1993). The *E. coli* OxyR binds an ATAGnTnnnAnCTAT motif (Toledano *et al.*, 1994). OxyR also regulates expression of the phage Mu *mom* gene which encodes a DNA modification function (Bolker and Kahmann, 1989). The protein binds to a 5' region of the *mom* structural gene, containing 3 methylation sites, and prevents transcription, but only when unmethylated. It can, therefore, discriminate between methylated and unmethylated states of the *mom* promoter region, and there may be other OxyR-regulated genes whose repression or induction depends on correct methylation. The OxyR-binding sites of the *mom* and *oxyR* genes share significant homology, therefore, the modes of interaction of OxyR with DNA may differ depending on whether it functions as a repressor or activator (Bolker and Kahmann, 1989).

A greater understanding of the oxidative stress response has been facilitated by the isolation and characterization of *oxyR* mutants. Christman *et al.* (1985) isolated a hydrogen peroxide-resistant mutant of *S. typhimurium*, designated *oxyR1*. This strain was highly resistant to hydrogen peroxide and constitutively expressed 9 of the 12 early hydrogen peroxide-inducible proteins. It was also more resistant to organic hydroperoxides, such as cumene hydroperoxide and *t*-butyl hydroperoxide, and heat killing, but was more sensitive to killing by quinones, such as menadione and other oxidants including diamide, paraquat, and N-ethylmaleimide. The mutant had a 45-50x greater specific catalase activity than the wild type associated with HPI and HPII, elevated levels of glutathione reductase, MnSOD, and alkyl hydroperoxide reductase, and constitutively overexpressed three heat shock proteins. Morgan *et al.* (1986) also found that the steady-state level of *katG* mRNA increased more than 50-fold in the *Salmonella oxyR1* mutant, resulting in a 50-fold increase in catalase activity. Another Tn*1* deletion mutant, *oxyR2*, was hypersensitive to hydrogen peroxide, and did not induce the 9 proteins. It contained a missense mutation which inactivates the *oxyR* gene (Christman *et al.*, 1989). Kullik *et al.* (1995a) isolated 6 non-repressing *oxyR* mutants impaired in DNA binding. Five of these mapped near the N-terminal helix-turn-helix motif conserved among LysR family members. The same authors (1995b) used site-directed mutagenesis to show that a cysteine residue at position

199 is responsible for its redox sensitivity, and that mutations in the C-terminal region result in constitutive activity. They showed that activation under both oxidizing and reducing conditions is due to the ability of the mutants to induce co-operative binding of RNA polymerase. The active form of OxyR is readily reduced to an inactive form, and, therefore, only lasts until hydrogen peroxide has been decomposed by catalase (Storz *et al.*, 1990).

Oxidative stress also induces some proteins which are induced under other stresses, for example, heat shock and DNA damage (Christman *et al.*, 1985; Greenberg and Demple, 1989; Morgan *et al.*, 1986). GroEL and GroES are induced by peroxide and superoxide-mediated oxidative stress, heat shock, starvation, and the SOS response in *E. coli*. Cells which are adapted to peroxide stress, therefore, also exhibit increased resistance to heat shock. Glucose-starved, stationary phase cells of *E. coli* also have enhanced resistance to other stresses, including oxidative stress and heat (Jenkins *et al.*, 1988). Morgan *et al.* (1986) have, similarly, noted overlaps in the hydrogen peroxide, heat, ethanol, and nalidixic acid responses in *S. typhimurium*. Therefore several stress responses overlap, and the extent of the overlap varies. Each stress, however, induces its own characteristic group of proteins, therefore there is probably not a common intracellular signal, but it is possible that the genes have *cis*-acting regulatory sequences required for several forms of regulation.

5.3.2. The role of RpoS in the regulation of catalases

Apart from the oxidative stress response, other mechanisms of regulation of catalases exist in prokaryotes. The regulation of expression of catalase in *E. coli* has been extensively studied. Catalases are repressed under strictly anaerobic conditions (Morgan *et al.*, 1986), but are induced by a shift from anaerobic to aerobic growth (Gregory and Fridovich, 1974). *E. coli* produces two catalases, HPI and HPII, which are regulated differently. HPI, encoded by *katG*, is produced under both aerobic and anaerobic conditions, and is transcriptionally induced by OxyR (> 100-fold), in response to hydrogen peroxide, as a part of the oxidative stress response (Christman *et al.*, 1985; Loewen *et al.*, 1985; Storz *et al.*, 1990; Tao *et al.*, 1991). It is, therefore, also induced by reagents which produce hydrogen peroxide, and other oxidative agents, including near-UV irradiation (Eisenstark, 1989), the organic peroxides cumene hydroperoxide and *tert*-butyl-hydroperoxide (Belkin *et al.*, 1996), and superoxide-generating

drugs (Greenberg and Demple, 1989). Ascorbic acid, which reacts with oxygen in the presence of metal ions to produce dehydroascorbate and hydrogen peroxide, also induces *katG*. The presence of 5.7 mM ascorbic acid results in an 8-fold increase in catalase activity after 30 minutes in *E. coli* (Richter and Loewen, 1981). Similarly, catalase activity in exponential cultures of *S. typhimurium* is induced 4-5-fold after 1 hour exposure to 60 μM hydrogen peroxide in minimal glucose medium (Christman *et al.*, 1985; Kagaya *et al.*, 1992).

HPI activity also increases 6-fold in an *oxyR*^{c} strain, and 5-fold in an *oxyR*$^-$ deletion strain as cultures enter stationary phase. Growth phase-dependent induction begins at mid exponential phase, and reaches a maximum during late exponential phase, and is probably independent of OxyR (Mukhopadhyay and Schellhorn, 1994). Mukhopadhyay and Schellhorn (1994) treated exponential phase cells with spent media from cultures expressing high catalase and peroxidase activity. They observed a 3-4-fold increase in HPI peroxidase activity in both *oxyR*^{c} and *oxyR*$^-$ strains, relative to cells resuspended in fresh medium. These results suggest that extracellular metabolites which accumulate during growth are responsible for the growth phase-dependent induction of HPI. There is a marked increase in the extracellular concentration of acetate as the cells enter late exponential phase, and treatment of early exponential phase cells with medium containing 0 to 80 mM sodium acetate (pH 7.0) results in induction of HPI activity in an *oxyR*$^-$ (5-fold) and *oxyR*$^+$ (7-fold) strain. Other weak acids such as sodium benzoate, and sodium lactate also induce HPI activity, independently of OxyR (2-fold). The induction of HPI as the cells enter stationary phase depends on RpoS or KatF (Loewen *et al.*, 1985; Mulvey *et al.*, 1990; Mukhopadhyay and Schellhorn, 1994), a σ factor involved in the regulation of a group of stress proteins induced during the transition to stationary phase, and during starvation (Tanaka *et al.*, 1993). Cells mutant in *rpoS* exhibit a 9.75-fold lower expression of *katG* during stationary phase, and a significantly lower expression throughout the life cycle (Ivanova *et al.*, 1994). Tao *et al.* (1993) showed that wild type $E\sigma^{70}$ cannot bind to the *katG* promoter without OxyR. RpoS may act directly through binding of RNA polymerase, or indirectly by synthesis of a new transcriptional activator, which allows $E\sigma^{70}$ to bind to the *katG* promoter, in the absence of OxyR (Ivanova *et al.*, 1994). Ivanova *et al.* (1994) provide evidence that the basal levels of *katG* expression are regulated by RpoS. It is not known whether σ^S and σ^{70} recognize the same *katG* promoter. Greenberg and Demple (1988) and Loewen (1992) have isolated and characterized mutations in the *katG* promoter that cause overproduction of HPI in *oxyR*^{sup}

mutants. All contained a C → T change at position -58, which is approximately in the middle of the OxyR-binding site, separate from the RNA polymerase binding site. Loewen (1992) suggests that this mutation may enhance RNA polymerase binding in the absence of OxyR. Other possibilities are that it changes promoter specificity, with respect to the σ factor, or that there is another transcriptional activator protein that increases expression of *katG* without OxyR (Ivanova *et al.*, 1994). Kim and Kim (1995) demonstrated that non-*oxyR* regulators may be responsible for regulating transcription at the basal level, and that multiple binding proteins can specifically interact with the *katG* promoter in *E. coli*.

Synthesis of HPII, or KatE, in *E. coli* increases with increased respiratory activity, accumulation of tricarboxylic acid cycle intermediates (Loewen *et al.*, 1985), osmotic changes (Jenkins *et al.*, 1990), and entry of cells into stationary phase (Mulvey *et al.*, 1990), and relies largely on KatF. HPII expression is low during exponential phase, but is induced 30-fold upon entry into stationary phase (Mulvey *et al.*, 1990), independently of hydrogen peroxide or anaerobiosis. As noted for HPI, when exponentially growing cells are treated with spent medium from stationary phase cultures, HPII is induced (Schellhorn and Hassan, 1988; Mukhopadhyay and Schellhorn, 1994), and this is probably due to the accumulation of weak acids, such as acetate, in the supernatant (Schellhorn and Stones, 1992). The expression of KatE relies largely on RpoS. RpoS expression is induced by aromatic amino acids, and this in turn activates *katE* transcription, although the mechanism of induction by the amino acids is not known. Maximal induction of *rpoS* results in a 10-fold increase in the level of *katE*. Further increases in *katE* expression occur without any change in *rpoS* levels, suggesting that another factor may be involved (Mulvey *et al.*, 1990).

The two catalases of *E. coli* are, therefore, regulated in a different manner, but have at least one regulatory mechanism in common. Schellhorn and Stones (1992) suggest that signalling by the accumulation of weak acids may allow the cells to produce additional catalase during the switch from fermentative to oxidative metabolism, as cells begin to use weak acids as carbon sources. Oxidation of these acids in the presence of oxygen may result in partially reduced toxic oxygen species, such as hydrogen peroxide, which would require the catalases for detoxification. Acid-linked induction of HPI may also be important, for example, in macrophages which have as a part of their defence mechanisms a low pH and toxic levels of hydrogen peroxide.

Mukhopadhyay and Schellhorn (1994) suggest, therefore, that cytoplasmic acidification or alteration of the proton motive force may function as a metabolic signal for induction of the stationary-phase regulon. Other mechanisms of regulation of catalase may also occur. For example, Mukhopadhyay *et al.* (ASM Abstract H-150, 1997) recently demonstrated that the bacterial adaptive response gene, *barA*, encoding a membrane bound sensory kinase, may affect catalase expression in *E. coli*, possibly through its role in iron uptake. The regulation of catalase may be finely tuned by a complex network of regulatory proteins. It was recently shown that OxyS, a small, untranslated RNA, induced by OxyR, regulates the translation of *rpoS*. Thus, when OxyR is activated in response to hydrogen peroxide, both *katG* and *oxyS* are induced. The OxyS then inhibits *rpoS* expression, therefore preventing redundancy. During the stationary phase, OxyR is inactive, and RpoS is induced by starvation signals and, in turn, induces the catalases (Zhang *et al.*, ASM Abstract H-151, 1997).

5.3.3. Catalase synthesis in other bacteria

Catalases are regulated in a similar manner in other gram-negative and gram-positive bacteria, with some variations. The *Bacteroides fragilis* catalase enzyme, encoded by *katB*, is regulated in a similar manner to the *E. coli* HPII. It is upregulated by oxygenation and stationary phase but not by peroxides (Rocha and Smith, 1995). In contrast, expression of the catalase gene, *katX*, of *Xanthomonas oryzae pv. oryzae*, a plant pathogen, is highest during exponential phase. It is highly inducible by superoxide generators and methyl methane sulfonate (8-10-fold), and weakly inducible by hydrogen peroxide (2.5-fold) (Chamnongpol *et al.*, 1995). Interestingly, *katX* expression in *E. coli* was temperature- and growth phase-dependent (Mongkolsuk *et al.*, 1996). The HPII homologues of *Haemophilus influenzae*, *hktE*, and *Rhizobium meliloti*, *katA*, are induced by hydrogen peroxide (Bishai *et al.*, 1994; Hérouart *et al.*, 1996). The activity of the *Caulobacter crescentus* catalase-peroxidase, like that of *E. coli*, increases in the transition from exponential to stationary phase. Peroxidase activity increases approximately 100-fold after 50 to 70 hours, and this is partly responsible for the large increase in resistance to hydrogen peroxide killing in the stationary phase (Schnell and Steinman, 1995). In *Rhodobacter capsulatus*, catalase-peroxidase activity is 10-15-fold higher under aerobic-dark than anaerobic-light conditions, i.e. it is regulated by oxygen (Hochman and Shemesh, 1987). This is due to a 20-50-fold higher activity of the *cpeA* promoter region (Forkl *et al.*, 1996). No further

increases in activity occur when the cultures reach stationary phase. Treatment with 200 μM hydrogen peroxide also results in an 8-10-fold induction of expression from the promoter. Forkl *et al.* (1996) showed, using deletion analysis of the promoter region, that a palindromic sequence (-143 to -130, with respect to the transcription start site) was required for induction. The palindromic sequence (TCGGCGCGCGCCGA) resembles a recognition site for DNA binding regulatory proteins, and is probably required for induction of expression under the two conditions described. The region (-68 to -40) shows high homology to the complementary upstream region of the *oxyR* and *katG* genes of *E. coli* at which OxyR binds (Farr and Kogoma, 1991). This region may function as a second regulatory domain, since deletion results in a 2-3-fold reduction in LacZ activity. It does not, however, closely match the extended OxyR consensus sequence. The results support the idea that, during exponential growth, the enzyme acts as a catalase for the detoxification of hydrogen peroxide, while during stationary phase it may act mainly as a peroxidase for adaptation to starvation conditions (Forkl *et al.*, 1996).

In *B. subtilis*, the peroxide stress response is regulated by a repressor that senses both metal ion levels and hydrogen peroxide. The vegetative catalase, KatA, is induced by hydrogen peroxide, and upon entry of the cells into stationary phase. The post-exponential induction of KatA requires Arb, a transition state regulator (Strauch and Hoch, 1993), and is inhibited by Mn(II). The *B. subtilis mrgA* gene, which also forms a part of the peroxide response regulon, encodes a homologue of the stationary phase *E. coli* protein, Dps (Almirón *et al.*, 1992). Both *katA* and *mrgA* contain similar Per boxes, centred 15 bp upstream of the -35 region, which may be responsible for regulation (Chen *et al.*, 1995).

Pseudomonas aeruginosa possesses two catalases, KatA (170 kDa), which is produced throughout the growth phase, and KatB (228 kDa), which is produced in response to paraquat or hydrogen peroxide (Brown *et al.*, 1995). Catalase activity in this organism is also regulated by the *fur* gene product (discussed below). Mutations in the *P. aeruginosa fur* locus result in a decrease in total SOD and catalase activity, and increased sensitivity to hydrogen peroxide (Hassett *et al.*, 1996). The region upstream of the *katB* gene contains a sequence CGCATTGATATTCCTAAT at position -114 to -97, with respect to the ATG start codon, that has 13/19 matches with the consensus *E. coli* iron box, however, the authors were unable to demonstrate binding of the purified Fur to this region. Possibly katB is regulated by Fur, since

the mutants here resulted in increased activity, but indirectly, since it does not bind to the promoter region.

5.3.4. The role of Fur in the regulation of catalases

As mentioned in Chapter 1, catalase has been shown to be a virulence factor in several organisms (Mitchison *et al.*, 1960; Wilson *et al.*, 1995; Bishai *et al.*, 1994; Hochman *et al.*, 1992; Mandell, 1975; Filice, 1983; Beaman and Beaman, 1984; Zheng *et al.*, 1992). Another factor involved in virulence is the ability to acquire and store iron, which is an essential element for all bacteria, but is not readily available under aerobic conditions, at neutral pH (Litwin and Calderwood, 1993). Free iron is limited in mammalian hosts due to the presence of transferrin in the serum and lactoferrin in mucosal secretions (Bullen *et al.*, 1978). The available iron is also in the incorrect ionic form for bacterial use. The low iron concentrations, as well as other environmental conditions within host cells, such as temperature, calcium, osmolarity, and anaerobiosis, may be signals for the derepression of virulence factors (reviewed by Mekalanos, 1992), including those involved in iron metabolism (Litwin and Calderwood, 1993). The Fur protein controls iron acquistion in *E. coli* (Hantke, 1987), *S. typhimurium* (Foster and Hall, 1992), *Legionella pneumophila* (Hickey and Cianciotto, 1994), *Vibrio cholerae* (Litwin *et al.*, 1992), *Vibrio anguillarum* (Tolmasky *et al.*, 1994), *Campylobacter jejuni* (Woolridge *et al.*, 1994), *Yersinia pestis* (Staggs and Perry, 1991), *Neisseria meningitidis* (Thomas and Sparling, 1994), and *Neisseria gonorrhoeae* (Berish *et al.*, 1993). Fur also governs iron transport systems (Hantke, 1987), and regulates expression of *sodA* (Mn-SOD) in *E. coil* (Niederhoffer *et al.*, 1990), as well as a number of other virulence factors, such as the Shiga-like toxin of *E. coli* (Calderwood and Mekalanos, 1987; DiRita and Mekalanos, 1989; Hale, 1991), and the *P. aeruginosa* exotoxin A (Bjorn *et al.*, 1978). It is, therefore, essential for a number of cellular processes including iron metabolism, response to oxidative stress, the acid-tolerance response, sugar metabolism, and toxin synthesis (Litwin and Calderwood, 1993; Tsolis *et al.*, 1995; Hassett *et al.*, 1996). In the presence of sufficient iron, Fur complexes to Fe(II), and binds to a highly conserved Fur box upstream of iron-regulated genes, which overlaps the promoter, and blocks transcription. When iron is deficient, the genes are no longer repressed by Fur (de Lorenzo *et al.*, 1988; O'Sullivan *et al.*, 1994). Sequences regulated by Fur contain the "iron box" palindromic consensus sequence, GATAATGATAATCATTATC, (Bagg and Neilands,

1987; de Lorenzo *et al.*, 1987). Fur is also activated to bind to iron-regulated promoters by other divalent metal ions, including manganese, zinc, cobalt, copper, and cadmium (Bagg and Neilands, 1987).

Another regulator of the iron response is DtxR, which shares little significant amino acid sequence homology with Fur (Boyd *et al.*, 1990; Schmitt and Holmes, 1991). The *C. diphtheriae dtxR* gene encodes an iron-responsive repressor of the diphtheria toxin, *tox*, promoter, (Pappenheimer, 1977; Boyd *et al.*, 1990). It plays the same role as Fur, i.e. regulates siderophore expression (Schmitt and Holmes, 1991). The DtxR-binding site contains a region of dyad symmetry, and overlaps the -10 region of the *tox* promoter. The consensus binding site contains a perfect 9 bp inverted repeat, interrupted by a single base pair. It has no significant similarity to the *E. coli* Fur-binding site (Litwin and Calderwood, 1993), and the two proteins do not bind each others' operator sequences (Schmitt *et al.*, 1992; Schmitt and Holmes, 1991). A DtxR homologue has been found in *Brevibacterium lactofermentum* (Oguiza *et al.*, 1995), *S. lividans*, and *Streptomyces pilosus* (Günter-Seeboth and Schupp, 1995), and DtxR-like binding sites have been identified upstream of iron-regulated genes in *S. pilosus* (Günter *et al.*, 1993).

5.4. THE REGULATION OF CATALASE IN MYCOBACTERIA

5.4.1. The mycobacterial oxidative stress response

As is evident from Chapter 1, the study of the mechanisms of action of, and resistance to, isoniazid has led to an improved understanding of the oxidative stress response in the mycobacteria. In *M. leprae*, the *ahpC* gene is linked to, and divergently transcribed from the *oxyR* gene (Bergh and Cole, 1994). The same situation occurs in *M. tuberculosis*, except that the *oxyR* gene is altered. The *M. tuberculosis ahpC* gene has > 90% homology with that of *M. avium* and *M. leprae*, and is also very similar to other members of the *ahpC* family. The *oxyR-ahpC* intergenic region is also similar to that of *M. leprae*, however, the *oxyR* gene of *M. tuberculosis* is shorter and carries multiple lesions, including in-frame stop codons, deletions, and a frameshift, which prevent it from forming a single, contiguous translation product. The *oxyR* gene is, therefore, naturally inactivated in this organism. The *M. tuberculosis oxyR* gene is

also transcribed at a much lower level than that of *M. leprae*. A study of the *ahpC-oxyR* region in 20 clinical isolates of *M. tuberculosis*, as well as other members of the *M. tuberculosis* complex, revealed identical lesions in the *oxyR* gene in all, suggesting that they are specific to members of the *M. tuberculosis* complex (Sherman *et al.*, 1995; Deretic *et al.*, 1995). The latter authors suggest that the *oxyR* lesions must have occurred after the divergence from the common mycobacterial ancestor.

The important question is, therefore, in the absence of a functional OxyR, how does *M. tuberculosis* regulate the expression of *katG* and other OxyR-regulated genes? Deretic *et al.* (1995) showed that the *M. tuberculosis katG* gene was not induced by either high or low concentrations of hydrogen peroxide, suggesting that this gene, and probably *ahpC* too, is regulated differently from that in enteric bacteria. Growth of the organisms at 40°C resulted in a slight increase in KatG labelling in the presence of hydrogen peroxide, suggesting that a combination of factors is required for an appreciable induction of the gene. This is in contrast to the results of Sherman *et al.* (1995; 1996). These authors demonstrated a moderate increase in the mRNA and protein levels of *katG*, but no change in the levels of *ahpC*, in *M. tuberculosis* and BCG upon exposure to hydrogen peroxide. They also cloned the *M. tuberculosis katG* promoter upstream of the luciferase gene, and monitored expression in BCG. Expression increased 7-fold upon treatment with hydrogen peroxide, however, these exposures did not protect *M. tuberculosis* against subsequent lethal hydrogen peroxide doses, as they do in *E. coli*, *S typhimurium*, and *M. smegmatis* (Christman *et al.*, 1985; Rosner and Storz, 1994; Sherman *et al.*, 1995). Induction of *katG* did, however, allow a 35-fold increase in the survival of BCG exposed to cumene hydroperoxide. A *katG*-deleted strain showed a much higher sensitivity to oxidative stress, and no protection against organic peroxides could be induced by pre-exposure to subinhibitory hydrogen peroxide concentrations. This may explain the attenuated virulence resulting from loss of *katG*. The authors overexpressed the *ahpC* gene from *M. tuberculosis* Erdman in strain H37Rv, which was then tested for sensitivity to cumene hydroperoxide. The recombinant was significantly more resistant to the peroxide than the wild type. Similar results were obtained for BCG.

Sherman *et al.* (1995) and Dhandayuthapani *et al.* (1996) examined the oxidative stress response in *M. smegmatis*. The former authors demonstrated the induction of > 12 proteins in

M. smegmatis in response to hydrogen peroxide. The authors cloned and sequenced the *ahpC* gene from this organism, and showed that it encoded the 25 kDa peroxide-inducible protein. It has 63% identity with the *C. diphtheriae* DirA, which is functionally interchangeable with the AhpC from *E. coli* (Tai and Zhu, 1995). The AhpC enzyme is involved in the conversion of potentially harmful hydroperoxide derivatives of lipids and nucleic acids into corresponding alcohols (Christman *et al.*, 1985; Tartaglia *et al.*, 1989), and probably has the same function in the mycobacteria. The region upstream of the *ahpC* gene has no detectable similarities with the corresponding conserved region in other mycobacteria, and does not contain an *oxyR* gene. This suggests that the genomic organization of *M. smegmatis* is different to other mycobacteria. Pagan-Ramos *et al.* (ASM Abstract U-130, 1997) recently reported an inability to detect any sequences homologous to *oxyR* in *M. smegmatis*. Dhandayuthapani *et al.* (1996) also showed, using metabolic labelling of newly synthesized proteins, that a distinct pattern of proteins is induced in *M. smegmatis* upon exposure to hydrogen peroxide, most of which were also inducible by two alkyl peroxides, cumene hydroperoxide and *tert*-butyl hydroperoxide. Nine proteins were induced by peroxide altogether, of which 6 (17, 19, 25, 42, 64, and 68 kDa) responded to hydrogen peroxide treatment. A 50 kDa protein was induced only at very high concentrations of hydrogen peroxide (> 125 μM), but not by organic peroxides. Two of the proteins induced by both organic peroxides, 16 and 59 kDa, were not inducible by hydrogen peroxide. *M. smegmatis,* therefore, has a significant peroxide stress response.

Downstream of *ahpC* in *M. tuberculosis*, *M. bovis*, *M. smegmatis*, and *M. leprae* is an ORF encoding a putative protein of 18-19 kDa, *ahpX*, which shows no homology to the second component of alkyl hydroperoxide reductase, AhpF of *E. coli* and *S. typhimurium* (Sherman *et al.*, 1995). The *C. diphtheriae* DirA antibody recognizes the *M. tuberculosis* AhpC, but its basal levels are too low for detection by immunoblotting. Previous studies by Dhandayuthapani and co-workers (1995) have shown that the *M. tuberculosis ahpC* is transcribed at lower levels than that of *M. leprae,* and these studies showed that there are dramatic differences in the levels of AhpC in *M. smegmatis* and *M. tuberculosis*. The authors (1996) were also unable to detect AhpC in *M. aurum* ATCC 23366 by immunoblot analysis, even after exposure to hydrogen peroxide and organic peroxides. Interestingly, this species also contains no detectable *oxyR* sequences. Protein labelling studies showed that only one protein (18 kDa) was induced by peroxide stress in this organism. This was induced by organic peroxide but not H_2O_2. Both

M. marinum and *M. xenopi* produce detectable basal levels of AhpC, which are induced upon exposure to hydrogen peroxide (Pagan-Ramos *et al.*, ASM Abstract U-130, 1997).

Dhandayuthapani *et al.* (1997) recently examined the interaction of the *M. leprae* OxyR with the *oxyR-ahpC* intergenic region. They identified two transcription start sites for both the *ahpC* and *oxyR* genes of *M. leprae*. The second transcription start sites of each overlapped by 28 nucleotides, while the first transcription start site of *oxyR* coincided with the translation initiation codon of the *oxyR* gene. This phenomenon has been described in some *Streptomyces* promoters (Strohl, 1992). The authors overexpressed and purified the *M. leprae* OxyR in *E. coli*, and demonstrated binding of it to both the *M. tuberculosis* and *M. leprae oxyR-ahpC* intergenic region, although the former had a lower affinity for the protein. The OxyR-binding site was located in a region spanning the initiation codon of *oxyR*. The authors demonstrated a significant decrease in the transcription of the *M. leprae ahpC*, in the absence of the *M. leprae oxyR*. A putative OxyR recognition site, CTTATCACTCATACCGATAAG, was located 65 bp upstream of the first *ahpC* transcription start site. The authors examined the same region in other mycobacterial species and proposed the following mycobacterial OxyR consensus recognition sequence, CTTATCGGC-N_3-GCCGATAAG, which shows limited similarity to that of *E. coli*, but contains the LysR T-N_{11}-A motif. The sequence is well conserved in other mycobacteria, except *M. tuberculosis*. The *M. tuberculosis* OxyR-binding site deviates in the left side of the core motif, by substitution of the ATC with GGC, which destroys the T-N_{11}-A motif, and may account for its lower affinity for OxyR (Dhandayathupani *et al.*, 1997). Pagan-Ramos *et al.* (ASM Abstract U-130, 1997) recently characterized the *oxyR* and *ahpC* genes of *M. marinum* and *M. xenopi*. They compared the OxyR binding sites in these organisms with those of *M. leprae* and *M. tuberculosis*, and proposed a consensus mycobacterial OxyR recognition sequence, CTTATCGGt-N_3-ACCGATAAG, which differs slightly from that proposed by Dhandayathupani *et al.* (1997). Mutation of T and A nucleotides in the *M. leprae* sequence resulted in a decrease in the binding of the purified *M. leprae* OxyR protein, while mutation of the *M. tuberculosis* binding site to the consensus resulted in an increased binding (Pagan-Ramos *et al.*, ASM Abstract U-130, 1997).

It is evident from the above-mentioned studies that the oxidative stress response in most mycobacteria is similar to the well-characterized responses of *E. coli* and *S. typhimurium*. Due

to the lack of a functional OxyR, however, *M. tuberculosis* probably has no inducible oxidative stress response, and therefore relies on the constitutive defense offered by its inert and largely impermeable cell wall. Only KatG is peroxide-inducible but not to a sufficient extent to protect against hydrogen peroxide challenge. It appears quite paradoxical that organisms which require additional help in surviving oxidative stress imposed by host defence systems have an impaired system for induction of the oxidative stress response. It may be that these organisms are highly specialized, and express these genes constitutively at a high enough level to protect the cells during infection, or that they use other mechanisms of regulation of these genes. The lack of an oxidative stress response may explain the rapid loss of viability which results when *M. tuberculosis* is subjected to sudden changes in aeration conditions (Wayne, 1976).

5.4.2. The regulation of iron metabolism in the mycobacteria

The acquisition of iron by *M. tuberculosis* is probably essential for growth and virulence. Mycobacteria produce a number of proteins involved in the acquisition and storage of iron from the environment (Snow, 1970; Barclay and Ratledge, 1983; Wheeler and Ratledge, 1994; Fiss *et al.*, 1994). The presence of an iron acquisition system in mycobacteria suggests the presence of an iron regulation system. A DtxR homologue, IdeR, has been described in *M. tuberculosis* H37Rv, *M. bovis* BCG, *M. leprae*, *M. avium*, and *M. smegmatis* (Doukhan *et al.*, 1995). In *M. tuberculosis* and *M. leprae*, the *dtxR* gene is located immediately downstream of the *mysB* gene. The predicted product has an overall identity of 58% and conservation of 70% with the *C. diphtheriae* DtxR. The two proteins are 78% identical over 140 residues at the amino terminal. The N-terminal region of DtxR, which is conserved in mycobacteria, contains the DNA binding motif, the protein-protein interaction domain, and the metal binding site (Doukhan *et al.*, 1995; Qui *et al.*, 1995; Schiering *et al.*, 1995; Ding *et al.*, 1996). No function has been assigned to the 90 amino acid C-terminal region. The *fxbA* genes from *M. smegmatis* and *M. bovis* BCG have an iron-regulated promoter and a potential DtxR iron box (Fiss *et al.*, 1994), and Dale and Patki (1990) have identified a region with homology to the Fur binding site upstream of the *M. leprae* 28 kDa antigen. This is not a perfect match but is situated in the correct position, i.e. overlaps the -35 region.

Schmitt *et al.* (1995) showed that the *M. tuberculosis* IdeR regulates transcription from the *C. diphtheriae tox* promoter, in an iron-dependent manner, as efficiently as the *C. diphtheriae* DtxR. Dussurget *et al.* (1996) constructed isogenic IdeR-deficient mutants in *M. smegmatis*, which were defective in their ability to repress synthesis of exochelin and mycobactin. They showed a greater sensitivity to hydrogen peroxide, and had decreased catalase, peroxidase, and superoxide dismutase activities. The authors constructed two mutant strains, SM1 and SM3, which showed derepression of secreted siderophore production. Derepression was not complete, however, suggesting the existence of a second repressor. The *ideR* mutants did not show significant sensitivity to acid compared with the wild type strain, but did show an increased sensitivity to hydrogen peroxide on solid medium. Both mutants showed a dramatic decrease in Mn-SOD activity and catalase and peroxidase activity, associated with KatG, while a strain that had been complemented with the gene showed wild type activities. The KatE activity levels were unchanged. The authors therefore provide evidence that the mycobacterial IdeR is a repressor of exochelin and mycobactin synthesis in response to the environmental iron concentration, and that it is involved in the oxidative stress response. It is not known whether IdeR acts directly on the *sod* and *katG* promoters, or indirectly via other regulatory genes. The authors suggest that IdeR may be responsible for regulation of hydrogen peroxide-inducible expression of *katG* in *M. tuberculosis*. *E. coli fur* mutants are also more sensitive to hydrogen peroxide (Touati *et al.*, 1995), possibly due to an iron overload which generates ferryl and hydroxyl radicals through the Fenton reaction, which damage the DNA (Imlay *et al.*, 1988). Accumulation of intracellular free iron in the presence of oxygen generates oxidative stress and cellular damage, therefore, iron metabolism and oxidative stress are closely linked in enteric bacteria and mycobacteria (Dussurget *et al.*, 1996).

Some catalase regulatory mechanisms, therefore, exist in the mycobacteria, which are similar to those in other organisms. Mulder *et al.* (Manuscript in preparation) recently reported evidence that the *M. tuberculosis* MysB regulates synthesis of the *katG* gene in a similar manner to the *E. coli* RpoS, and suggests that MysB is the RpoS homologue of *M. tuberculosis*. This σ factor is probably responsible for the basal levels of *katG* expression in this organism, as well as the growth phase induction of the gene, noted in this study. Some evidence also exists for the post-transcriptional regulation of the gene. Rouse *et al.* (1996) have noted discrepancies between the amounts of *katG* mRNA transcript, and the levels of KatG protein, in some *katG* mutants of

M. tuberculosis. The authors suggest that the mutations may alter the stability of the *katG* mRNA transcript or KatG protein, and that the KatG protein synthesis may be regulated at a post-transcriptional level. Mutations in the *E. coli katG*, which alter the conformation, also reduce the stability of the active tetramer (Loewen *et al.*, 1990).

5.5. THE AIMS OF THIS STUDY

In an attempt to improve our understanding of both the regulation of gene expression in mycobacteria, and, particularly, the regulation of virulence factors, this study reports the examination of the *M. tuberculosis katG* promoter. The *katG* gene has been extensively studied with respect to its role in INH sensitivity, and the product has been purified and studied with respect to its enzyme activities, however, very few observations on the mechanisms of regulation of the gene have been reported. This study, therefore represents the first extensive report on this subject. The firefly luciferase gene was used as a reporter for characterizing the *kat*G promoter, and for testing under which conditions it is expressed in the mycobacteria. Both *M. smegmatis* and *M. bovis* BCG were used as mycobacterial hosts. The transcription start sites were mapped and putative -10 and -35 regions identified. Deletion analysis of the promoter-containing fragment revealed an upstream element required for optimal promoter expression. This element was also able to increase expression of an unrelated promoter, and may contain a binding site for regulatory proteins. The results of this study suggest a complex network of regulatory mechanisms required for expression from this promoter.

CHAPTER 6

CLONING AND IDENTIFICATION OF THE *M. TUBERCULOSIS KATG* PROMOTER AND CHARACTERIZATION OF CONDITIONS OF EXPRESSION

CONTENTS Pg.

CHAPTER 6

CLONING AND IDENTIFICATION OF THE *M. TUBERCULOSIS KATG* PROMOTER AND CHARACTERIZATION OF CONDITIONS OF EXPRESSION

6.1. SUMMARY

This chapter describes the construction of an *E. coli*-mycobacterial shuttle vector, pJCluc, containing the promoterless firefly luciferase gene as a reporter, for the examination of mycobacterial promoter sequences. The vector was used to characterize the *M. tuberculosis katG* promoter. A 1.9 kb region immediately upstream of the *katG* gene was amplified by PCR and cloned into pJCluc, and promoted expression of the luciferase gene in *E. coli* and *M. smegmatis*. A smaller promoter fragment (559 bp) promoted expression with equal efficiency, and was used in all further studies. Two transcription start sites were mapped by primer extension analysis, and used to identify putative -10 and -35 regions. These show strong similarity to previously identified mycobacterial promoters. *M. smegmatis* recombinants harbouring the promoter construct were used to determine under what conditions the *katG* promoter is expressed. Luciferase activities revealed that expression from the promoter peaks during late exponential phase, and declines during stationary phase. The promoter is induced by ascorbic acid, and is repressed by oxygen limitation and growth at elevated temperatures.

6.2. INTRODUCTION

The importance of the *M. tuberculosis katG* gene, and its role in virulence and drug susceptibility, have been emphasized in Chapter 1. Understanding the mechanisms of regulation of this gene may, therefore, provide insight into the processes that govern interaction with the host and mediate susceptibility to antimycobacterial agents, particularly isoniazid. At the same time, a study of this sort may provide more valuable information on the mycobacterial transcription signals and gene regulatory mechanisms.

As discussed in Chapter 5, studies on mycobacterial promoters in heterologous hosts are generally not successful. It is more favourable to examine mycobacterial transcription signals in a mycobacterial host, in which response to physiological constraints and regulatory mechanisms are conserved. However, due to various limitations in the genetic manipulation of mycobacteria, *E. coli* provides a more favourable host for the propagation and amplification of cloned and altered DNA fragments. For this reason, an *E. coli*-mycobacterial shuttle vector was employed. The requirements for the vector for this study were: a mycobacterial *ori* for replication and propagation in mycobacteria; an *E. coli ori* for replication and propagation in *E. coli*, an antibiotic resistance gene for selection of recombinants; a cloning site with selection for cloning of the reporter; and the reporter gene itself. The *E. coli*-mycobacterial shuttle vector, pJC86 (Jack Crawford, CDC Atlanta), provides all these requirements, except for the reporter. The firefly luciferase gene was chosen as a reporter, and was cloned into pJC86 to yield pJCluc. The source of the reporter gene was the pGEM®-*luc* plasmid (Promega). The firefly luciferase gene in pJCluc contains no promoter or ribosome binding site. These must be provided by the DNA fragment under investigation. Measurement of luciferase activity expressed off a cloned promoter, therefore, allows quantitation of transcription initiation, promoter strength, and efficiency of ribosome binding.

The primary advantage of the firefly luciferase as a reporter is the extreme sensitivity of the assay. Less than 10^{-20} moles (0.001 pg) of luciferase have been detected under ideal conditions (Wood, 1991a), which enables the study of low level expression, for example from weak promoters (Williams *et al.*, 1989). Sensitivity is approximately 100-fold greater than with the chloramphenicol acetyltransferase assay (Alam and Cook, 1990). Other advantages are that the assay is rapid, easy to perform and relatively inexpensive. Endogenous luminescence is absent from most biological samples, therefore there is no interference with the signal generated by luciferase, and no background. The luciferase gene encodes a 62 kDa protein which requires no post-translational modifications for activity, and can function immediately after translation (Wood *et al.*, 1984; de Wet *et al.*, 1985). It is also structurally very simplistic and contains no prosthetic groups, bound metals or disulfide bonds, and can therefore be expressed in virtually any host or cell type (Wood, 1990). The protein is stable at room temperature and loses only 5% of its activity after 10 hours. As mentioned in Chapter 5, the firefly luciferase catalyzes the oxidation of beetle luciferin with concomitant production of a photon of light. The oxidation

normally occurs from an enzyme intermediate, luciferyl-AMP, however, in the presence of co-enzyme A, it occurs through luciferyl-CoA, resulting in light production without the characteristic self-inhibition of luciferase (Wood, 1991b). In conventional luciferase assays (Wood, 1991a), a flash of light is generated, which rapidly decays after the enzyme and substrate are mixed. The Promega luciferase assay system, on the other hand, allows for greater enzymatic turnover of luciferase (Wood, 1991b), resulting in greater light intensity which remains nearly constant for several minutes. Thus this system does not require rapid mixing protocols or automated injection devices.

The vector constructed in this study provides a useful tool for the assessment of mycobacterial promoters and has been used to study the promoter regions of the *M. tuberculosis whiB* (Mulder *et al.*, Manuscript in preparation) and *cspA* (Ros Chapman, Unpublished) homologues, and the conditions of expression of the *M. tuberculosis mysB* gene (Mulder *et al.*, Manuscript in preparation). In this study, the vector facilitated the assessment of the regions essential for optimal promoter activity (Chapter 7), the mapping of the transcription start sites and identification of putative -10 and -35 sequences, and the quantitation of expression under different growth conditions and at different growth stages, of the *M. tuberculosis katG* promoter. Ideally, gene regulation should be studied in a single-copy situation, however, the pAL5000-based vector used in this study is present in the mycobacteria in a low copy number (3-5 copies per cell) and should, therefore, be adequate to use, unless a very tight regulatory network governs *katG* expression.

6.3. MATERIALS AND METHODS

The bacterial strains, plasmids, and culture conditions used are described in Appendix A. Details of the commonly used techniques and solutions are given in Appendices B and C, respectively.

6.3.1. Construction of pJCluc

The *E. coli*-mycobacterial shuttle vector, pJCluc, was constructed as shown in figure 6.1. The vector, pJC86 (Jack Crawford, CDC Atlanta), was prepared by digestion with *Kpn*I, removal of the resulting protruding ends with T4 DNA polymerase (DNA Blunting kit, Amersham Life Science) (see Appendix B), and subsequent digestion with *Hin*dIII. The 1748 bp *Hin*dIII-*Stu*I fragment from the luciferase cassette vector, pGEM®-*luc* (GenBank Accession number X65316) (Promega), containing the promoterless firefly luciferase gene, was ligated to pJC86 to form pJCluc. The construct $pJCluc_{opp}$, was made by insertion of a blunt-ended *luc* cassette into pJC86 digested with *Kpn*I and blunt-ended as described above. This construct contains the luciferase gene in the opposite orientation to that in pJCluc. *E. coli* cells harbouring recombinant plasmids were identified as white colonies on 2x YT agar plates containing X-GAL (100 μg/ml) and IPTG (0.3 mM) (Melford Laboratories, England), due to interruption of the pJC86 *lacZ* gene. The recombinant plasmids were verified by restriction analysis. The pUC19 forward and reverse 17-mer sequencing primers (Promega) were used to sequence over the junctions of the recombinant plasmids, to confirm the sequence and orientation of the inserts.

6.3.2. PCR amplification of the *katG* promoter region and cloning into pJCluc

Oligonucleotides were designed for the amplification of the region immediately upstream of the *M. tuberculosis katG* gene. The primers were based on the published sequence (GenBank Accession number X68081). Mismatches were introduced near the 5' end of each primer to generate convenient restriction endonuclease recognition sites. The forward primer was a 25-mer with the sequence: 5'-CTGGTAAGCttGGCCGCAAAACAGC-3', and the reverse primer was a 25-mer with the sequence: 5'-CACAGggaTCCTTCCAGGAGTTGGT-3'. The lower case nucleotides indicate the mismatches, and the underlined nucleotides represent the restriction endonuclease sites for *Hin*dIII and *Bam*HI, respectively. The complement of the GTG initiation codon of the *katG* gene is included at the 5' end of the reverse primer. The primers were predicted to amplify a 1943 bp fragment. PCR amplification was performed as described in Appendix B, using *M. tuberculosis* H37Rv genomic DNA (0.1 μg) as the template, and an annealing temperature of 50°C. Control reactions included those with no template DNA, or those with single primers only, to ensure that the PCR products were not the result of non-

specific binding of primers, or due to contaminating DNA. PCR products were separated by agarose gel electrophoresis, stained with ethidium bromide, and viewed under 310 nm UV transillumination. The 1943 bp PCR product was excised from an agarose gel, purified by the phenol squeeze method described in Appendix B, and digested to completion with the restriction endonucleases *Hin*dIII and *Bam*HI. The resulting fragment was cloned into both pJCluc and pUC18, which had been digested to completion with the same restriction endonucleases, to yield pJClucK and pKup, respectively (see Fig. A.1, Appendix A).

6.3.3. Subcloning of the *katG* promoter fragment

In order to localize the *katG* promoter region to a smaller fragment, the sequences upstream of the *Sma*I site of the pJClucK insert were deleted (see Fig. 6.2). To achieve this, the 559 bp *Sma*I-*Bam*HI fragment from pKup was cloned into the *Eco*RV and *Bam*HI restriction endonuclease sites of pBluescript-SK (Stratagene), to produce pKup0.6. Immediately upstream of the *Eco*RV restriction endonuclease site in this vector lies a *Hin*dIII restriction endonuclease site. This restriction endonuclease was, therefore, used, together with *Bam*HI, to remove the insert, which was then cloned into the *Hin*dIII and *Bam*HI restriction endonuclease sites of pJCluc to produce pJCLK0.6 (see Fig. A.2, Appendix A). This cloning strategy via pBluescript-SK resulted in the addition of 3 bp of vector sequence upstream of the *Sma*I restriction endonuclease site. The entire 559 bp insert of pKup0.6 was sequenced, using the pUC19 forward and reverse sequencing primers (Promega).

6.3.4. Electroporation of *M. smegmatis*

M. smegmatis cells were prepared for electroporation by the method of Jacobs *et al.* (1991). A stationary phase culture of *M. smegmatis* LR222 (Beggs *et al.*, 1995), in Middlebrook-ADC-Tween, was diluted 100-fold into fresh medium, and grown for 48 hours at 37°C with shaking. The cells were harvested by centrifugation, washed 3-4 times in ice-cold 10% (v/v) glycerol, and resuspended in a final volume of 2 ml/100 ml original culture volume in 10% (v/v) glycerol. For electroporation, 60 µl cells were combined with 140 µl of 10% (v/v) glycerol and 1 µg of DNA, in a 5 µl volume, and incubated on ice for 5 minutes. Electroporations were performed in 0.1

cm gap Gene Pulser® cuvettes (Bio-Rad Laboratories, CA U.S.A.) at 1.8 kV, 25 μF and 1000 Ω, using a Gene Pulser™ apparatus (Bio-Rad Laboratories, CA U.S.A.), which resulted in a field strength of 18.0 kV/cm. After electroporation, 900 μl of Middlebrook-ADC was added, and the cells were incubated at 37°C for 3-4 hours to facilitate expression of the kanamycin resistance gene. Transformants containing the recombinant plasmids were selected on Middlebrook-ADC agar plates containing kanamycin at a concentration of 50 μg/ml. A control electroporation with no added DNA was included to monitor the frequency of spontaneous mutation to kanamycin resistance After incubation at 37°C for three days, transformants were cultured in liquid medium for plasmid preparations. Plasmids were prepared as described in Appendix B and their structure verified by restriction analysis.

6.3.5. Preparation of cell extracts and luciferase activity assays

For the luciferase activity assays, cells were grown to late exponential phase in a 5 ml volume at 37°C with shaking. *E. coli* cell extracts were prepared by harvesting cells from 1 ml of culture and resuspension of the cells in 100 μl of ice-cold 1x Cell Culture Lysis Reagent (Luciferase Assay System, Promega). The cells were lyzed by sonication in ice water (3 bursts of 40 seconds each) in a Branson 1200 Ultrasonic Cleaner (Branson, Danbury CT). Cell debris was removed by centrifugation, and the supernatant equilibrated to room temperature before being assayed for luciferase activity. For the activity assays, 20 μl of cell extract was mixed with 100 μl of Luciferase Assay Reagent (Promega), which had also been equilibrated to room temperature. Light output was measured in a Bio-Orbit 1253 luminometer (Turku, Finland). *M. smegmatis* cell extracts were prepared by resuspension of cell pellets in Cell Culture Lysis Reagent, followed by processing in a FastPrep™ FP120 instrument (Savant Instruments Inc., Farmingdale, NY, U.S.A.) for two 40 second cycles at a setting of 6. Samples were incubated on ice before and after each cycle, to compensate for heat generated during processing. Unlyzed cells and cell debris were removed by centrifugation. The supernatants were assayed for luciferase activity as described above. The protein concentrations of the cell extracts were measured using the Bio-Rad Protein Assay kit (Bio-Rad Laboratories) (see Appendix B), and the luciferase activities are quoted as relative light units (RLU)/mg of protein used in each assay.

6.3.6. RNA extractions

RNA was extracted from *E. coli* and *M. smegmatis* using the FastPrep™ system (FastPrep™ FP120 instrument (Savant Instruments Inc., Farmingdale, NY, U.S.A.) and FastRNA® Kit-Blue (√Bio101 Inc., Vista, CA, U.S.A.)), according to the manufacturers' instructions. RNA pellets were resuspended in an appropriate volume of SAFE (FastRNA® Kit). RNA yields and quality were assessed by measurement of the optical density at 260 and 280 nm, and electrophoresis in a MOPS-formaldehyde gel (Appendix C). The concentration was calculated using the relationship OD_{260} = 1 for 40 μg/ml of RNA.

6.3.7. Primer extension analysis

Potential transcription start sites of the *M. tuberculosis katG* gene were identified using primer extension analysis. The method used was adapted from that of Sambrook *et al.* (1989). A complementary 18-mer primer was designed to anneal 58 bp downstream of the ATG initiation codon of the luciferase gene. The primer had the sequence: 5'-GCAGTTGCTCTCCAGCGG-3'. Approximately 100 ng of primer was end-labelled with [γ-^{32}P] dATP using polynucleotide kinase (Boehringer Mannheim), in a 30 μl reaction with the supplied buffer. After incubation at 30°C for 30 minutes, the unincorporated nucleotides were removed using a Promega G25 spin column, according to the manufacturers' instructions. RNA samples (100 μg) were precipitated with sodium acetate and ethanol, resuspended in 30 μl of RNA hybridization buffer (Ambion Inc., Texas U.S.A.), and added to 25 ng of labelled primer. The nucleic acids were denatured at 85°C for 10 minutes, and annealed overnight at 50°C. Approximate primer annealing temperatures were calculated according to the formula: $T_h = T_d - 5°C$, where T_h = annealing temperature and T_d = denaturing temperature = 4°C for every G or C and 2°C for every A or T. The annealed RNA was precipitated with sodium acetate and ethanol, and resuspended in diethyl pyrocarbonate (DEPC)-treated water. The primer was extended on the mRNA template at 37°C for 60 minutes, using 200 units of M-MLV Reverse Transcriptase (Promega), in a reaction containing 40 units of RNase inhibitor (Boehringer Mannheim), 1x Reverse Transcriptase buffer (Promega), and an equal mixture of dNTPs (1 mM each) (Boehringer Mannheim). The primer extension products were precipitated and resuspended in the Stop

buffer supplied in the Pharmacia T7 Sequencing kit. The products were denatured at 90°C for 2 minutes and separated on a 6% (w/v) polyacrylamide sequencing gel adjacent to dideoxy sequencing reactions primed with the same 18-mer primer. Gels were dried and the bands visualized by autoradiography.

6.3.8. ATP assays

ATP concentration was measured as an indication of cell mass, using the Promega Enliten™ Luciferase/Luciferin (L/L) bioluminescence detection reagent. The ATP levels were monitored by measuring the production of light when ATP, luciferin and oxygen were combined in the presence of the enzyme luciferase. The assay system was calibrated using an ATP standard of known concentration (Stanley *et al.*, 1989; Prioli *et al.*, 1985). For *M. smegmatis*, 55.8 µl of cells were added to 4.2 µl of 20% (w/v) trichloroacetic acid (TCA), and incubated on ice for 30 minutes. The precipitation was terminated by the addition of 50 µl of neutralization buffer (1 M tris-acetate pH 7.75) and 622 µl of distilled water. The light output generated from a mixture of 50 µl of L/L reagent in 347.5 µl of L/L buffer (see Appendix C), and 100 µl of TCA-treated sample, was measured in a Bio-Orbit 1253 luminometer (Turku, Finland) (reading B). A blank containing L/L reagent in L/L buffer was used to measure the background relative light units (reading A). The light output of a mixture of the L/L reagent, L/L buffer, TCA-treated sample, and 2.5 µl of 10^{-7} mM ATP represents reading C. The ATP concentration was calculated using the formula: [ATP] = (B-C)/A x 5 x 65.5914 x 10^{-10} mM. At higher cell densities, the TCA-treated samples were diluted for the assay and the ATP calculation adjusted accordingly.

6.4. RESULTS

6.4.1. Construction and testing of pJCluc

The *E. coli*-mycobacterial promoter probe vector, pJCluc, was constructed as shown in figure 6.1. The vector pJC86 contains the 2.8 kb *Eco*RV-*Hpa*I backbone of pAL5000, with the sequences necessary for replication in mycobacteria (Snapper *et al.*, 1990b), the *lacZ* gene and

oriE from pUC18, and the kanamycin resistance gene from Tn*5*, for selection of cells containing the plasmid (Jack Crawford, Personal Communication). The resultant construct, pJCluc, contains the *luc* gene in the opposite orientation to the *lacZ* gene, with no promoter or Shine-Dalgarno sequence. Upstream of the luciferase ATG are two unique restriction endonuclease sites for the cloning of promoter fragments, *Hin*dIII and *Bam*HI. The construct $pJCluc_{opp}$ contains the *luc* gene in the same orientation as *lacZ*. The constructs pJCluc, $pJCluc_{opp}$, and pJC86 were electroporated into *M. smegmatis* strain LR222. A control electroporation containing no DNA, yielded no kanamycin resistant colonies, suggesting that the frequency of spontaneous mutation to kanamycin resistance is below detectable levels under the conditions used in this study.

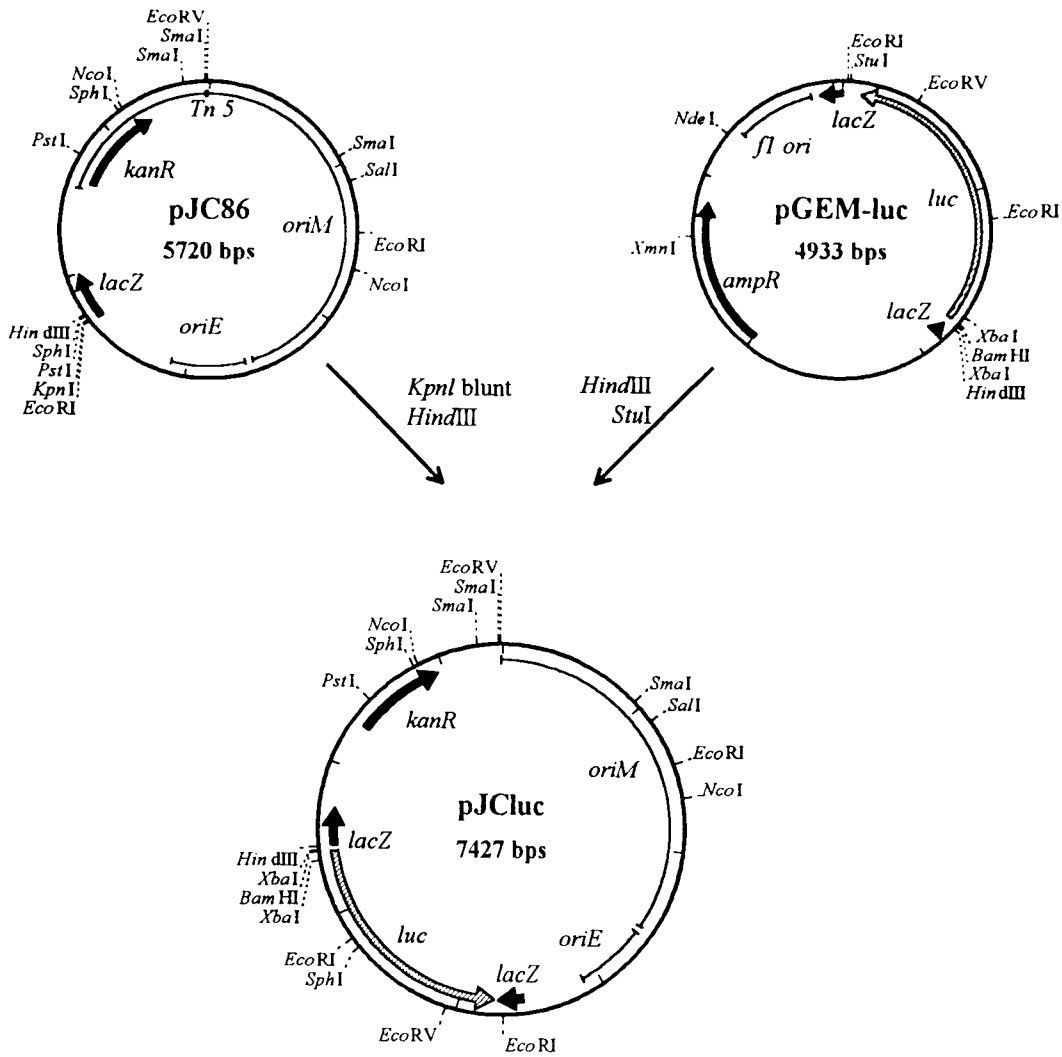


Figure 6.1 Schematic representation of the construction of pJCluc.

Late exponential-phase *E. coli* and *M. smegmatis* cells harbouring pJC86, pJCluc, and $pJCluc_{opp}$ were assayed for luciferase activity (Table 6.1). Cells containing pJC86 exhibited a luciferase activity of 0.001 RLU (results not shown), the same value obtained with the luciferase assay reagent alone, indicating no endogenous luminescence in either *E. coli* or *M. smegmatis*. Cells harbouring pJCluc exhibited a low level of activity in *E. coli* (44.55 RLU/mg protein), but slightly higher levels in *M. smegmatis* (142.74 RLU/mg protein). This small amount of activity may be due to readthrough from an upstream promoter, which appears to be more active in *M. smegmatis*, and may, therefore, be located on the pAL5000 region of the plasmid. The region upstream of the luciferase gene contains no known transcriptional terminators, thus allowing readthrough. However, it was demonstrated in later experiments that the luciferase activity resulting from pJCluc is extremely low in comparison to clones containing promoters. The $pJCluc_{opp}$ construct exhibited approximately 11-fold greater activity than pJCluc in *E. coli* (489.31 RLU/mg protein), but less than background activity in *M. smegmatis* (25.71 RLU/mg protein). In this construct, the luciferase gene is expressed from the *lacZ* promoter, which normally has low constitutive expression in *E. coli*, but is induced by IPTG. This promoter apparently functions inefficiently in mycobacterial cells.

6.4.2. Analysis of the sequence upstream of the *M. tuberculosis katG* gene

The region upstream of the *M. tuberculosis katG* gene (GenBank Accession number X68081) has been analyzed by Heym and co-workers (1993). These authors detected no open reading frames with high coding-probability values on either strand of the DNA fragment upstream of the *katG* gene. However, Song and Deretic have identified an open reading frame within this region which has 28, 27, and 25% identity, at an amino acid level, with the Fur proteins of *Pseudomonas aeruginosa*, *Legionella pneumophilus*, and *Bordetella pertussis*, respectively, and has been called FurA. The ORF encodes the signature of the putative metal-binding site, HHHxHxxCxxC (GenBank Accession number AF002194). Also within the *katG* upstream region are two copies of a 700 bp repeat, which are 68% identical, with the longest stretch of identity comprising 58 bp (Heym *et al.*, 1993) (Fig. 6.2). This sequence is not dispersed in the *M. tuberculosis* genome.

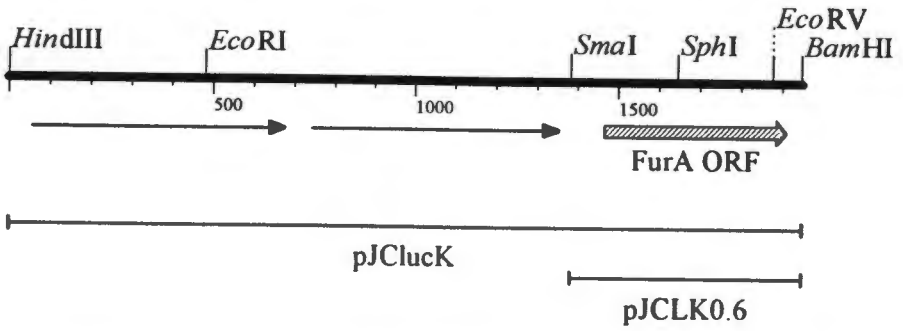


Figure 6.2 Linear restriction endonuclease map of the region upstream of the *M. tuberculosis katG* gene, corresponding to nucleotides 38 to 1981 of the published sequence (GenBank Accession number X68081). The positions of the repeat sequences are indicated with arrows. The FurA open reading frame is represented by a wide arrow. The *Hin*dIII and *Bam*HI restriction endonuclease sites were introduced by PCR. The regions present in pJClucK and pJCLK0.6 are shown.

6.4.3. Amplification and cloning of the region upstream of the *M. tuberculosis katG* gene

Oligonucleotides were designed for the amplification of a 1943 bp fragment immediately upstream of the *M. tuberculosis katG* gene. The reverse primer, which includes the *katG* start codon and a *Bam*HI restriction endonuclease site, was designed such that the distance between the *Bam*HI site and the GTG start codon was equal to that between the *Bam*HI site and the ATG of the luciferase gene in pJCluc, i.e. 2 bp. This was done to eliminate any differences in promoter function which may occur due to alterations in the distance between the translation initiation codon and the ribosome binding site. A large fragment was amplified to ensure the capture of all possible promoter and regulatory sequences. The primers were used in a PCR with *M. tuberculosis* H37Rv genomic DNA as the template. Reactions containing no template DNA or only single primers yielded no product, while those containing both primers at equal concentrations, together with the template DNA, resulted in a product of the predicted size (Fig. 6.3). The resulting 1943 bp fragment was cloned into the *Bam*HI and *Hin*dIII sites of both pJCluc and pUC18 to yield the constructs pJClucK and pKup, respectively (see Fig. A.1). Approximately 100 bp at the beginning and end of the insert of pKup were sequenced using the pUC19 forward and reverse sequencing primers (Promega), to ensure that the amplified product was as predicted. A smaller promoter fragment containing sequences downstream of the *Sma*I site was cloned separately into pJCluc, as shown in figure A.2 (Appendix A), to yield pJCLK0.6. Sequencing of the 559 bp insert in pKup0.6 revealed that the amplified fragment was identical to the published sequence, except for one A → T change at position -210 with

respect to the translation initiation codon. This base substitution may have been introduced as a result of an error in the PCR reaction, alternatively, it may be a polymorphism.

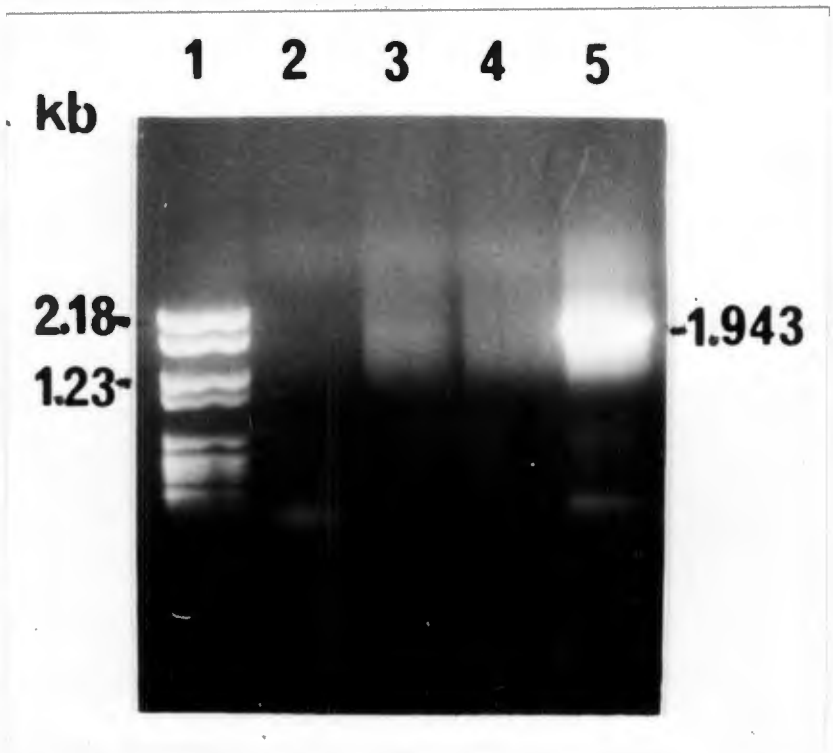


Figure 6.3 Agarose gel of the amplified *katG* promoter product and the control PCR reactions. The lanes were loaded as follows: 1) DNA molecular weight marker VI (Boehringer Mannheim); 2) No template DNA; 3) Forward primer only; 4) Reverse primer only; 5) Both primers and *M. tuberculosis* H37Rv template DNA.

6.4.4. Luciferase activity assays on the *katG* promoter constructs

Both pJClucK and pJCLK0.6 were electroporated into *M. smegmatis*, and tested for luciferase activity in *M. smegmatis* and *E. coli* (Table 6.1). *M. smegmatis* cells harbouring pJClucK and pJCLK0.6 exhibited high activity (41 787.32 and 38 545.22 RLU/mg protein, respectively), while *E. coli* transformants exhibited similar levels of activity to the uninduced *lacZ* promoter (380.71 and 474.45 RLU/mg protein, respectively). These results indicate that the *katG* promoter lies within the amplified fragment and is recognized by the *E. coli* and *M. smegmatis* transcription machinery. The luciferase activities of both *E. coli* and *M. smegmatis* cells harbouring pJClucK and pJCLK0.6 did not differ significantly, suggesting that the promoter region lies within the smaller *Sma*I-*Bam*HI fragment, and that no additional sequences required for optimal expression are located in the 1374 bp region upstream of the *Sma*I restriction endonuclease site. The smaller construct, pJCLK0.6, was therefore used for all further studies.

Table 6.1 Luciferase activities of the pJCluc constructs in *E. coli* and *M. smegmatis*.

	E.coli		*M. smegmatis*	
Construct	**RLU/mg protein[1]**	**Rel. act.[2]**	**RLU/mg protein[1]**	**Rel. act.[2]**
pJCluc	44.55 ± 17.83	1.00	142.74 ± 27.62	1.00
pJCluc$_{opp}$	489.31 ± 61.68	10.98	25.71 ± 12.72	0.18
pJClucK	380.71 ± 99.68	8.55	41787.32 ± 2372.55	292.76
pJCLK0.6	474.45 ± 13.06	10.65	38545.22 ± 10694.53	270.05

[1] The mean of three independent assays ± the standard error of the mean.
[2] Relative activity- normalized to the luciferase activity of pJCluc

6.4.5. Mapping of the transcription start site by primer extension analysis

Primer extension analysis was performed using an antisense strand primer complementary to sequences near the 5' end of the luciferase gene, and total RNA isolated from late exponential phase *M. smegmatis* and *E. coli* cells harbouring pJCLK0.6. In both *E. coli* and *M. smegmatis*, transcription was initiated at two sites located 47 (tssA) and 56 (tssB) bp upstream of the translation initiation codon (Fig. 6.4). The possibility that the bands may be due to a premature stop site of the reverse transcription reaction can probably be excluded, as sequencing of this region revealed no signs of band compressions, suggesting no tendency for secondary structure formation. The putative -10 and -35 sequences relative to these transcription start sites are represented in figure 6.5.a and b. The putative -10 hexamer for tssA, GACACT, lies 8 bp upstream of the transcription start site, and shows a high degree of similarity to other mycobacterial -10 sequences (see Table 5.1). It matches the -10 hexamer of the *M. paratuberculosis* P_{AN} promoter (Murray *et al.*, 1992) in 5/6 positions, and the -10 hexamers of the mycobacteriophage L5 gene 71 P2 (Nesbit *et al.*, 1995), the *M. tuberculosis* 16S rDNA (Ji *et al.*, 1994), the *M. fortuitum bla* (Timm *et al.*, 1994), and the *M. smegmatis* promoter S69, identified by Bashyam *et al.* (1996), in 4/6 positions. The first G of the hexamer is the least conserved. Associated with this -10 hexamer are three possible -35 sequences; 1) TCTATG 19 bp upstream, 2) TGTCCT 15 bp upstream, and 3) TCCTGA 13 bp upstream. Only the first is represented in figure 6.5.a. This hexamer shows 4/6 matches with the -35 sequences of the mycobacteriophage 13 pKGR38 and ORF1 (Ramesh and Gopinathan, 1995), while the second matches the putative *M. leprae* 36K promoter (Dale and Patki, 1990) in 5/6 positions, and the mycobacteriophage 13 pKGR25 (Ramesh and Gopinathan, 1995) in 4/6 positions. The third

-35 sequence shows 4/6 matches with the mycobacteriophage 13 pKGR9 (Ramesh and Gopinathan, 1995), mycobacteriophage L5 gene 71 P1 (Nesbit *et al.*, 1995), and *M. tuberculosis gyrB* (Madhusudan *et al.*, 1994) and *cpn60* (Kong *et al.*, 1993) -35 sequences.

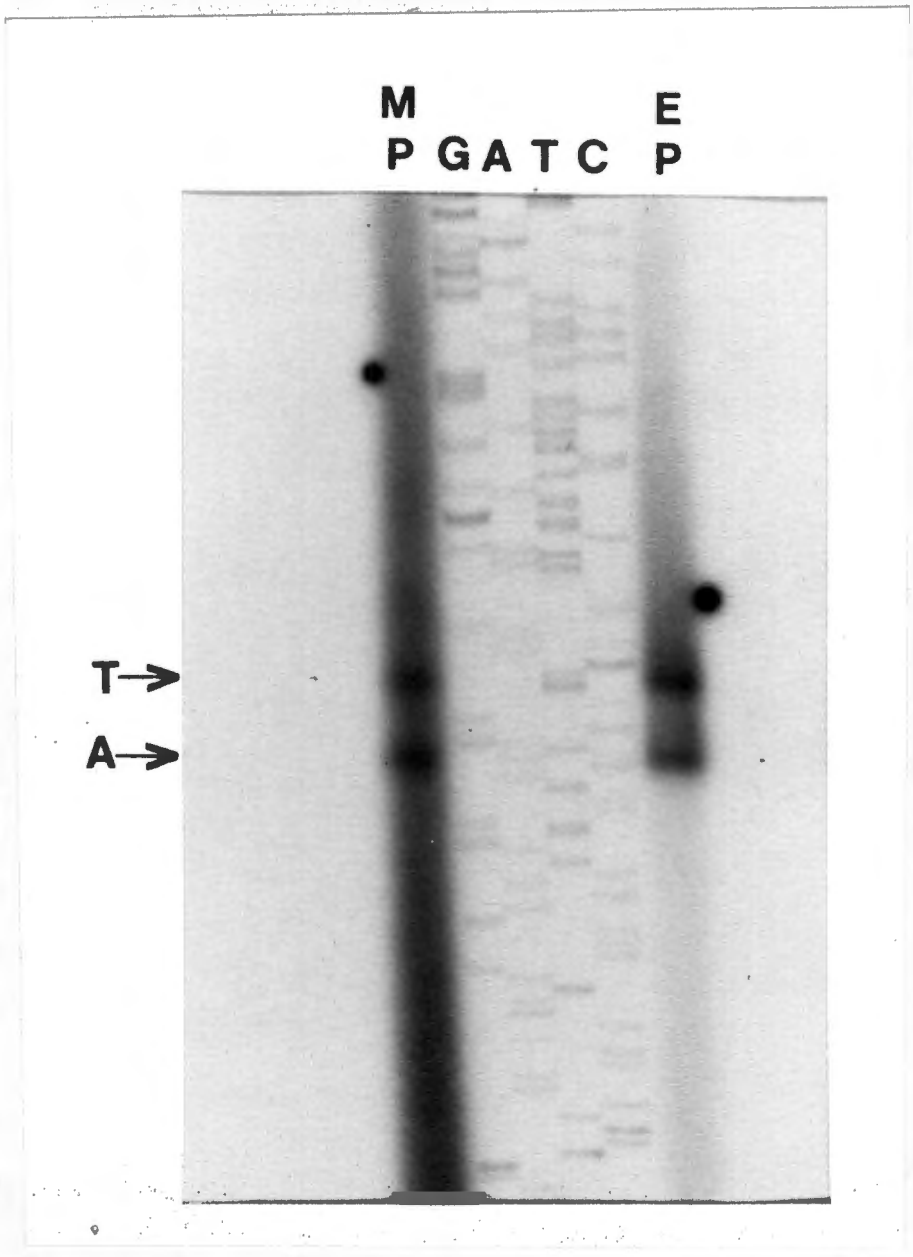


Figure 6.4 Primer extension analysis of the *M. tuberculosis katG* upstream region using RNA isolated during late exponential phase from *E. coli* and *M. smegmatis* cells harbouring pJCLK0.6. Two potential transcription start sites were identified, 47 and 56 bp upstream of the GTG translation initiation codon. M- *M. smegmatis* RNA; E- *E. coli* RNA.

The putative -10 hexamer for tssB, GATATC, lies 6 bp upstream of the transcription start site. This hexamer matches the *M. leprae* 18 kDa -10 hexamer (Dellagostin *et al.*, 1995) in 5/6 positions, and the *M. tuberculosis gyrB* (Madhusudan *et al.*, 1994), *M. smegmatis* acetamidase (Mahenthiralingam *et al.*, 1993), and *M. smegmatis* S30 (Bashyam *et al.*, 1996) -10 hexamers in 4/6 positions. Again, the first G of the hexamer is the least conserved. Associated with this -10 region are two possible -35 sequences, TCTACT, 22 bp upstream, and TACTGG, 20 bp upstream. The first, which is represented in figure 6.5.a, matches the -35 hexamers of a number of other mycobacterial promoters in 4/6 positions. These include the *M. bovis* rDNA (Suzuki *et al.*, 1991) and α antigen (Matsuo *et al.*, 1988), *M. leprae* 16S rDNA (Sela and Clarke-Curtiss, 1991), *M. smegmatis* acetamidase (Mahenthiralingam *et al.*, 1993) and *rrnB* (Predich *et al.*, 1995), the *M. paratuberculosis* P_{AN} (Murray *et al.*, 1992), *M. tuberculosis* 85A antigen (Kremer *et al.*, 1995), and some of the mycobacteriophage 13 promoters (Ramesh and Gopinathan, 1995). The second putative -35 hexamer matches that of the *M. tuberculosis gyrB* (Madhusudan *et al.*, 1994) in 4/6 positions. Some of the promoters to which the *katG* promoter has been compared have not had their transcription start sites mapped. The -10 and -35 regions for these promoters are, therefore, not confirmed, and are based purely on similarities to *E. coli* promoter sequences. However, the putative -10 and -35 sequences identified here for the *M. tuberculosis katG* promoter appear to be fairly well conserved.

a)

```
GCGAGGCACCCTGTCTGACGGCCTCGGACCATAACGGCTTCCTGTTGGACGAGGCGGAGG    1879
    -35B      -35A                -10B  -10A PB       PA
TCATCTACTGGGGTCTATGTCCTGATTGTTCGATATCCGACACTTCGCGATCACATCCGT    1939
                                          ↑       ↑
GATCACAGCCCGATAACACCAACTCCTGGAAGGAATGCT GTG CCC GAG              1987
                           R.B.S.       M   P   E
```

b)

	-35 region	spacing	-10 region	spacing	+1	spacing	start codon
P_A	TCTATG	19 bp	GACACT	8 bp	A	46 bp	GTG
	TGTCCT	15 bp	GACACT				
	TCCTGA	13 bp	GACACT				
P_B	TCTACT	22 bp	GATATC	6 bp	T	55 bp	GTG
	TACTGG	20 bp	GATATC				

Figure 6.5 a) The nucleotide sequence of the *M. tuberculosis katG* promoter region, including the first 3 codons, labelled with respect to the published sequence (GenBank Accession number X68081). The transcription start sites are indicated with arrows. The putative -10 and -35 hexamers for each transcription start site are underlined, and the putative ribosome binding site is in boldprint. b) Summary of the potential *katG* promoter sequences discussed in section 6.4.5.

6.4.6. Growth phase-dependent expression of the *katG* promoter in *M. smegmatis*

Expression of the *katG* promoter in *M. smegmatis* was monitored over a 68 hour period of growth by measurement of luciferase activity. The values used to plot the curves in figure 6.6 are from one experiment, but are representative of the trend observed in three independent experiments. Stationary phase cultures of *M. smegmatis*, harbouring pJCLK0.6, were diluted 100-fold into 500 ml of catalase-free Middlebrook-ADC-Tween, and incubated at 37°C with shaking. Growth was monitored using absorbance readings at 600 nm to measure cell density, and ATP concentration to measure cell mass. Samples were removed at various times and assayed for luciferase activity. The growth curves generated by absorbance readings and ATP concentration showed similar trends (Fig. 6.6). According to the curves, the cells reach late exponential phase at approximately 30 hours, after which both curves form a plateau, indicative of stationary phase. It is evident that expression of luciferase from the *katG* promoter is low in early exponential phase, but increases steadily and reaches a peak at late exponential phase, followed by a rapid decline as the cells enter stationary phase.

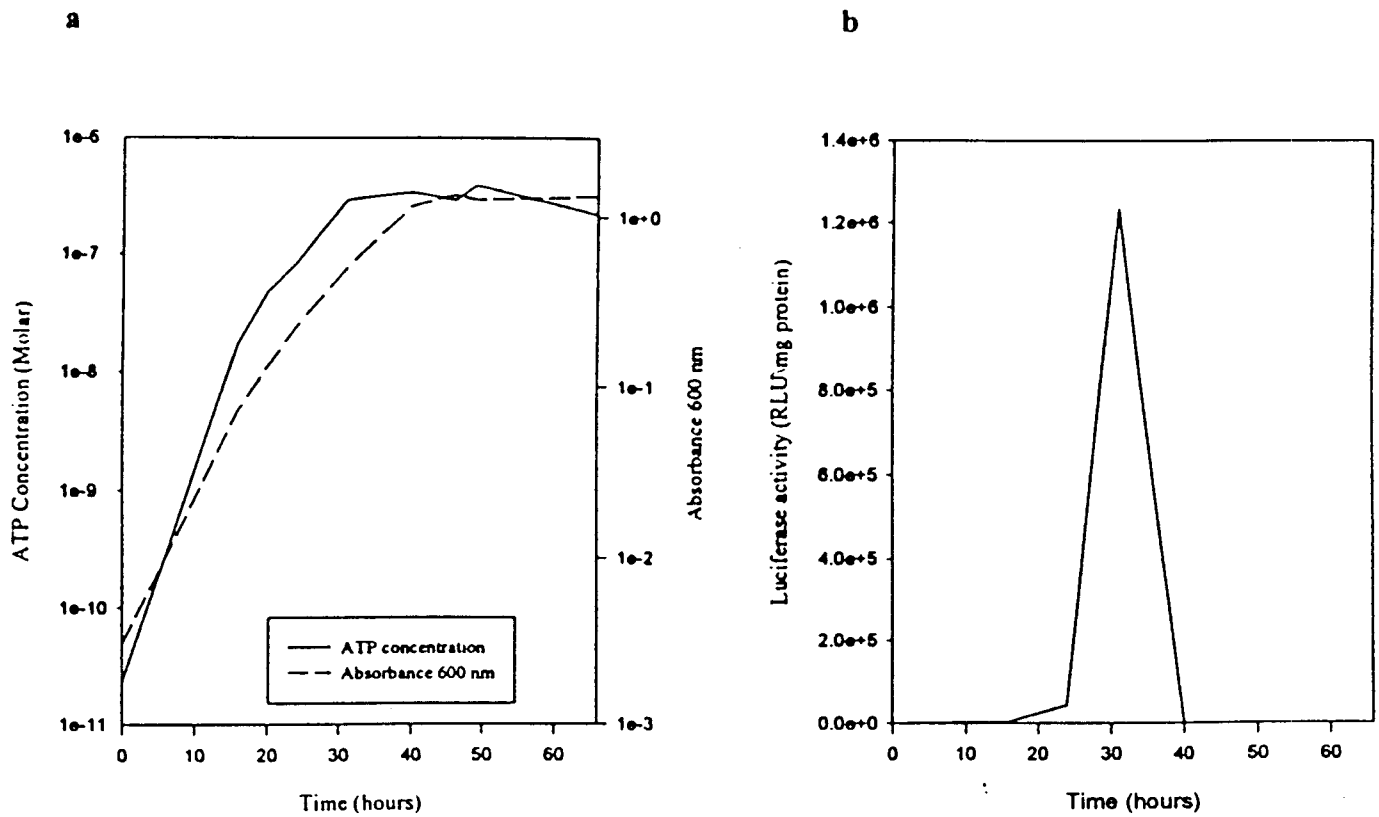


Figure 6.6 a) Growth curve of *M. smegmatis* (pJCLK0.6) plotted as a measure of ATP concentration and absorbance at 600 nm. b) Luciferase activity of the *M. tuberculosis katG* promoter monitored over a 68 hour growth period in *M. smegmatis*.

6.4.7. Expression of the *katG* promoter in response to stress in *M. smegmatis*

Previous reports on the regulation of expression of the *katG* gene in other bacteria have revealed a number of agents or conditions which regulate expression. Some of these were tested for their ability to influence the expression of the *M. tuberculosis katG*. For the experiment, 800 ml of an *M. smegmatis* (pJCLK0.6) culture were grown to mid exponential phase (24 hours), and split into 8 separate cultures of 100 ml each. The cells were harvested by centrifugation and resuspended in catalase-free Middlebrook-ADC-Tween with the following additions: 1) hydrogen peroxide to a final concentration of 10 mM; 2) ascorbic acid to a final concentration of 10 mM; 3) sodium acetate to a final concentration of 60 mM (pH 7.0); and 4) 2,2'-dipyridyl to a final concentration of 200 μM. The chemical 2,2'-dipyridyl is an iron chelator, and removes all available iron from the medium. In the 5th culture, the pH was adjusted to 4.0 with glacial acetic acid. The remaining three cultures were resuspended in medium with no additions. One of these was heat shocked by incubation at 45°C for 4 hours, one was incubated in a 100 ml bottle with no head volume and without agitation, to simulate oxygen limitation (Wayne and Sramek, 1994), and the last had no alterations and served as a control. For the standing culture, the cells that had settled at the bottom of the bottle were used as the anaerobic sample. Unless otherwise stated, the cultures were incubated at 37°C with shaking. After 4 hours, samples were removed and assayed for luciferase activity. The experiment was performed in triplicate, and the luciferase activities of one representative experiment are given in table 6.2. Treatment with 10 mM ascorbic acid resulted in an almost two-fold (86%) increase in luciferase activity. Growth under oxygen limiting conditions and at 45°C resulted in a 10.7- and 41.8-fold decrease in luciferase activity, respectively. The other changes in the culture conditions did not significantly affect luciferase activity.

Table 6.2 Luciferase activities of *M. smegmatis* (pJCLK0.6) exposed to various stresses.

Stress	RLU/μg protein	Rel. Act.[1]
Control	88.60	1.00
Hydrogen peroxide	87.82	0.99
Sodium acetate	80.02	0.90
Ascorbic acid	164.48	1.86
Oxygen limitation	8.25	0.09
↓ iron	78.26	0.89
↓ pH	79.52	0.90
45°C	2.12	0.02

[1] Relative activity- normalized to the luciferase activity of the control

6.5. DISCUSSION

In the luciferase assay system, light intensity is proportional to luciferase concentration in the range 10^{-16} to 10^{-8} M, and is a measure of the rate of catalysis by luciferase. The light output is, therefore, a measure of the levels of luciferase enzyme present, which is directly related to the rate of transcription of the luciferase gene by upstream promoters. In pJCluc, the level of transcription of the firefly luciferase gene was low, and the detectable activity was probably due to readthrough from an upstream promoter present in the vector. When a specific promoter fragment was cloned upstream of the gene, the levels of transcription, i.e. luciferase activity, increased at least 10-fold, and can, therefore, be attributed to the cloned promoter. Neither *E. coli* nor *M. smegmatis* cells exhibited endogenous luminescence.

The 1943 bp and 559 bp *M. tuberculosis katG* promoter fragments promoted expression of the luciferase gene with approximately equal efficiency under the conditions used in this study. This suggests that sequences upstream of the *Sma*I site are not essential for optimal promoter efficiency during normal growth. The possibility that this region may be involved in the regulation of the gene cannot, however, be excluded. The promoter functioned in *E. coli*, suggesting that the transcriptional signals are recognized by the *E. coli* transcription machinery. Expression in this host, however, was low, therefore the transcriptional signals probably do not closely match those of consensus *E. coli* promoters (discussed below). It is not known whether the RNA polymerase holoenzyme responsible for recognition of the *M. tuberculosis katG* promoter in this host is bound to the principal σ factor, σ^{70}, or an alternative σ factor, such as σ^{S}, which is responsible for transcription of the *E. coli katG* promoter during normal growth (Ivanova *et al.*, 1994). As mentioned in Chapter 5, there are a number of other mycobacterial promoters which have been shown to function in *E. coli*.

The activity of the *M. tuberculosis katG* promoter was high in *M. smegmatis* in comparison to its activity in *E. coli*. Its relative promoter strength in the mycobacterial host cannot be predicted without comparison with luciferase activities generated by other mycobacterial promoters in the same system. The activities of the promoter in *M. smegmatis* and *E. coli* can also not be accurately compared due to differences in the copy number of the plasmid in the different hosts. In *E. coli*, pUC18-based plasmids generally have a copy number of 500-700 (Minton *et al.*, 1988), while the pAL5000 plasmid only has a copy number of 3-5 in

M. smegmatis (Ranes *et al.*, 1990; Stover *et al.*, 1991). The luciferase activities measured in *M. smegmatis* should, therefore, be increased by a factor of approximately 100, to compensate for the differences in plasmid copy number. This study, however, is not concerned with absolute promoter strengths, but rather with relative promoter strengths.

Two transcription start sites were identified for the *M. tuberculosis katG* promoter in pJCLK0.6 in both *E. coli* and *M. smegmatis*. This suggests that the gene is transcribed off two promoters and that the same promoter sequences are recognized by both hosts. The presence of two promoters upstream of mycobacterial genes is not uncommon and has been reported for the *M. bovis* BCG *hsp60* (Stover *et al.*, 1991) and rDNA (Suzuki *et al.*, 1991), the *M. tuberculosis recA* (Movahedzadeh *et al.*, 1997b) and *whiB* homologues (Mulder *et al.*, Manuscript in preparation), and the *M. leprae oxyR* and *ahpC* genes (Dhandayathupani *et al.*, 1997). It is possible that the promoters are recognized by different RNA polymerase holoenzymes, and that they are utilized to different extents during growth. In primer extension experiments using different RNA preparations of the same recombinants, the proportions of the transcripts differed, possibly due to slight differences in the growth stage or conditions under which the RNA was isolated. This suggests that the two promoters are differentially expressed, but this observation will require further investigation. It was recently demonstrated that removal of the sequence between the ribosome binding site and transcription start site B of the *katG* promoter results in a 45-50% reduction in luciferase activity (Gavin Rush, Unpublished data). This suggests that the promoters are used to similar extents under the conditions of the above experiment. Primer extension analysis on total RNA isolated from *M. tuberculosis* H37Ra was not successful, possibly due to low levels of the *katG* mRNA transcript. However, the determination of the transcription start site from RNA isolated from a recombinant *M. smegmatis* and *M. bovis* BCG containing the promoter of interest in a promoter probe vector, has been shown to correlate with results obtained from native mRNA transcripts (Levin and Hatfull, 1993; Dhandayuthapani *et al.*, 1997; Movahedzadeh *et al.*, 1997b), and has been used as a substitute in many cases (Murray *et al.*, 1992; Kremer *et al.*, 1995; Nesbit *et al.*, 1995).

The putative -10 and -35 sequences identified for the *katG* promoter show significant homology to many mycobacterial promoters characterized thus far, and may, therefore, be "typical

mycobacterial promoters." The -10 hexamer associated with tssA matches the consensus *E. coli* sequence in 3/6 positions, two of which, A at position 2 and T at position 6, are highly conserved. All three putative -35 sequences match the *E. coli* consensus sequence in 2/6 positions, 1 of which, T at position 1, is highly conserved. The putative -10 hexamer for tssB also matches the *E. coli* consensus sequence in 3/6 positions, however, only one of these is highly conserved, the A at position 2. The first putative -35 hexamer for tssB matches the *E. coli* consensus in 3/6 positions, one of which, the T at position 1, is highly conserved, while the second putative -35 region has only the T at position 1 in common with the *E. coli* consensus sequence. The two putative *katG* promoters are, therefore, probably recognized and bound by $E\sigma^{70}$ in this host. The spacing between the -10 and -35 regions differs from that found in most *E. coli* promoters (17 ± 1) (Harley and Reynolds, 1987). This, together with the divergence from the consensus sequences, may be responsible for the poor efficiency of the promoters in *E. coli*. It has been noted that the distance between the -10 and -35 region is not critical in mycobacterial promoters (Kremer *et al.*, 1995), and varies between 13 and 24 bp (Ramesh and Gopinathan, 1995). Both -10 hexamers are least conserved in the first position, where the T is replaced by a G. This may be due to the high G+C content noted in many *M. tuberculosis* promoter sequences (Bashyam *et al.*, 1996). Bashyam *et al.* (1996) have noted that the less conserved bases in *M. tuberculosis* promoters tend towards G and C substitutions. *E. coli* promoters which deviate from the consensus in the first T of the -10 region often require activator proteins for maximum promoter activity (Raibaud and Schwartz, 1984). This possibility is investigated in Chapter 7. It is not known which -35 regions are utilized, or whether a -35 region is necessary for this promoter. Recent evidence suggests that this region is not essential in many mycobacterial promoters and can withstand large variation (Kenney and Churchward, 1996; Bashyam *et al.*, 1996). The confirmation of the identified sequences as promoter sequences, and their importance in promoter activity will require more in-depth analysis, such as site-directed mutagenesis.

The growth phase-dependent expression of the *M. tuberculosis katG* promoter in *M. smegmatis* is similar to that in other bacteria (see Chapter 5), and to the *M. smegmatis katG* (Mulder *et al.*, Manuscript in preparation). Many bacterial catalases are produced at low levels during exponential growth and are induced either during late exponential phase or during the transition to stationary phase (Mukhopadyay and Schellhorn, 1994; Rocha and Smith, 1995; Schnell and

Steinman, 1995). This probably reflects the function of the enzyme in the cells. The *M. tuberculosis* KatG may be produced in steadily increasing levels during exponential phase to detoxify the oxidative products of normal aerobic metabolism. As the cells enter stationary phase, the oxidative challenge increases resulting in an increased requirement for KatG for survival. As mentioned in Chapter 5, in *E. coli*, the induction of *katG* as the cells enter stationary phase is dependent on the stationary-phase or starvation response σ factor, KatF or RpoS (Loewen *et al.*, 1985; Mulvey *et al.*, 1990; Mukhopadyay and Schellhorn, 1994). The MysB has been identified as the putative KatF homologue in *M. tuberculosis* (Mulder *et al.*, Manuscript in preparation). It was shown that the levels of *mysB* mRNA in *M. tuberculosis* H37Ra increase during stationary phase. Analysis of the luciferase activities of *M. smegmatis* cells harbouring a pJCluc derivative containing the *M. tuberculosis mysB* promoter, revealed a similar growth phase-dependence. In fact, the trend seen in the luciferase activities relative to the growth phase was identical to that with the *katG* promoter. Expression was low during early exponential phase, peaked at late exponential phase or during the transition to stationary phase, and declined during stationary phase. The authors also demonstrated that overexpression of the *M. tuberculosis mysB* and *E. coli katF* genes in *M. smegmatis* resulted in 7.9- and 4.2-fold increases in the mRNA levels of the *M. smegmatis katG*, respectively. Similar increases in *katG* mRNA levels were demonstrated in *M. bovis* BCG overexpressing *mysB*, although to a smaller extent. These results suggest that the MysB is either directly or indirectly involved in the expression of the *katG* gene in the mycobacteria. The fact that the two genes show similar levels of induction during growth, suggests that MysB is responsible for the growth phase-dependence of the *katG* gene noted in this study, and that it may direct basal levels of transcription of the gene during normal growth.

The experiments designed to measure the effect of various stresses on expression of the *katG* promoter produced interesting results. In *E. coli*, most stress responses occur within the first hour, however, *M. smegmatis* has a much slower growth rate in comparison to *E. coli*, and may take longer to respond to stress. A heat shock response has been noted in *M. smegmatis* within the first hour (Bongi Ntolosi, Unpublished data), while it takes between 3 and 6 hours for the cells to respond to cold shock (Karen Shires, Unpublished data). A time period of 4 hours for exposure to the stresses in this experiment was chosen in order to accomodate these differences in response time and to allow for lag phases. It is possible that this time period is either too long or too short to accurately measure the effects, and may not represent maximal induction or

repression of the promoter. Resolution of these issues will require a more exhaustive series of experiments, involving exposure of the cells to each stress for a variety of time points. As revealed in the materials and methods, the cells were removed from the relevant media or stress conditions and the proteins extracted for the luciferase assay. Thus the luciferase activity measurements are not influenced by the nature of the stress. Instead, the luciferase activities are indicative of the amount of luciferase protein induced by exposure to the stress. The repression of the *katG* promoter under anaerobic conditions has also been noted for other catalases (Morgan *et al.*, 1986), and is probably due to a reduced requirement for the enzyme. The factor/s responsible for this repression, and the repression of the promoter at elevated temperatures, have not been identified. In *E. coli*, treatment with 5.7 mM ascorbic acid results in an 8-fold increase in catalase activity, and this is due to the reaction of ascorbic acid with oxygen, in the presence of metal ions, to produce dehydroascorbate and hydrogen peroxide (Richter and Loewen, 1981). The treatment of *M. smegmatis* (pJCLK0.6) with 10 mM ascorbic acid resulted in less than 2-fold induction of the *M. tuberculosis katG* promoter. This, together with the lack of induction of expression in the presence of hydrogen peroxide, suggests that the promoter is not significantly affected by hydrogen peroxide, and that the effect of ascorbic acid observed here is not due to the production of this agent. Belkin *et al.* (1996) have shown, using the *V. fischeri* luciferase as a reporter, that the *E. coli katG* promoter is induced approximately 20 minutes after exposure to hydrogen peroxide, and that induction peaks at 45 minutes and then declines. This observation raises the question whether the *M. tuberculosis katG* promoter is induced by hydrogen peroxide but that, either the lag phase had not yet ended, or the induction had already declined, at the stage at which the luciferase activity was measured in this study. The lack of induction of the promoter by hydrogen peroxide agrees with the results of Deretic *et al.* (1995), but not those of Sherman *et al.* (1996). The latter authors detected a 7-fold increase in expression from the *M. tuberculosis katG* promoter in an *M. bovis* BCG host in response to hydrogen peroxide, however, the exposures used did not protect *M. tuberculosis* against subsequent lethal doses. The differences in the response of the *katG* promoter to hydrogen peroxide may have been due to the use of different mycobacterial hosts. As mentioned in Chapter 5, *M. tuberculosis* does not have a functional OxyR, which is responsible for the hydrogen peroxide-dependent induction of the *katG* gene in most bacteria. In addition, no homologous *oxyR* sequences have been detected in *M. smegmatis*. These observations suggest that the *katG* gene is regulated differently in these organisms, and that induction of the

gene in response to hydrogen peroxide is unlikely, unless it occurs through an alternative mechanism. The treatment of *E. coli* with 0-80 mM sodium acetate (pH 7.0) induces catalase activity 7-fold (Mukhopadyay and Schellhorn, 1994). However, no induction of the *M. tuberculosis katG* promoter in response to sodium acetate was observed in this study, suggesting, again, that there are differences in the regulation of the gene in the two organisms.

Dussurget *et al.* (1996) have shown that IdeR-deficient mutants of *M. smegmatis* exhibit reduced catalase and peroxidase activities, associated with KatG. This suggests that the protein is responsible for the regulation of the *katG* gene in response to iron limitation in this organism, either directly by activating the *katG* promoter, or indirectly, via other regulatory genes. In this study, no changes in expression from the *katG* promoter were noted under iron limiting conditions. It is possible that the *M. tuberculosis katG* promoter is not recognized by the *M. smegmatis* IdeR or the factor/s induced by the protein, or that regulation in response to iron limitation occurs at a post-transcriptional level in this organism. Alternatively, the *M. smegmatis* IdeR may not be induced or active under the conditions used in this study.

The examination of the regulation of expression of the *M. tuberculosis katG* promoter in this study has been carried out in the homologous host, *M. smegmatis*. The *katG* gene is native to both *M. smegmatis* and *M. tuberculosis*, but numerous studies have shown important differences between the two mycobacterial species with respect to the oxidative stress response. Thus, although the experiments described above represent a good starting point for the study of the regulation of the gene, these studies need to be extended to the natural host, *M. tuberculosis*, in which there are likely to exist unique mechanisms of regulation during normal growth and in response to various stresses.

CHAPTER 7

IDENTIFICATION OF AN UPSTREAM ELEMENT REQUIRED FOR OPTIMAL EXPRESSION OF THE *KATG* PROMOTER

CONTENTS

CHAPTER 7

IDENTIFICATION OF AN UPSTREAM ELEMENT REQUIRED FOR OPTIMAL EXPRESSION OF THE *KATG* PROMOTER

7.1. SUMMARY

This chapter describes the identification of regions of the 559 bp *M. tuberculosis katG* promoter fragment that are essential for promoter activity. Deletions in the fragment were made using Bal31 nuclease and restriction endonucleases. Analysis of the luciferase reporter protein activities of the resulting constructs revealed that a region between 559 and 448 bp upstream of the translation initiation codon, designated the upstream activator region (UAR), is essential for promoter activity in *E. coli*, and is required for optimal activity in *M. smegmatis*. Primer extension analysis on selected deletions revealed that an alternative promoter is used in the absence of this region, and in the absence of the putative *katG* promoters identified by primer extension analysis in Chapter 6. Cloning of a fragment containing the *katG* UAR upstream of the *M. paratuberculosis* P_{AN} promoter, in pJCluc, resulted in a 14.7-fold increase in luciferase activity in *E. coli* and an 11.6-fold increase in *M. smegmatis*, relative to the P_{AN} promoter alone. Mobility shift assays revealed that one or more *M. smegmatis* proteins bind to the *katG* UAR. This region, therefore, acts as a transcriptional enhancer, possibly by binding to one or more regulatory proteins. The luciferase activities of a selection of promoter constructs were tested in *M. bovis* BCG. These constructs exhibited similar activities in this host as they did in *M. smegmatis*, suggesting that the recognition and utilization of promoters is similar in the two mycobacterial species.

7.2. INTRODUCTION

In the previous chapter, the *M. tuberculosis katG* promoter region was localized to a 559 bp fragment immediately upstream of the GTG initiation codon of the gene. Two putative promoters were identified, which showed some similarity to previously identified mycobacterial promoters. It was demonstrated that expression of the *katG* promoter is highest at the end of

exponential phase and declines during stationary phase, and that it is induced by ascorbic acid and repressed by oxygen limitation and growth at elevated temperatures.

Although the -10 and -35 promoter sequences are essential for expression, it is evident from the data presented in the introduction in Chapter 5 that a number of other sequences and regions adjacent to the promoter may also play a crucial role in promoter activity, and in determining promoter strength. For example, the AT-rich sequence identified upstream of the -35 sequence of the *E. coli rrnB* P2, *leuV* and RNA II promoters (Ross *et al.*, 1993). It was noted in Chapter 6 that the putative -10 hexamers of the *M. tuberculosis katG* promoters, P_A and P_B, deviate from the consensus sequence in the first position, which normally has a conserved T. *E. coli* promoters which deviate in this position usually require other upstream regions or activator proteins for optimal promoter activity (Raibaud and Schwartz, 1984). Therefore, other regions of the *M. tuberculosis katG* promoter fragment, which may play a role in promoter activity were examined. It was demonstrated that a region between 559 and 448 bp upstream of the translation initiation codon is essential for promoter activity in *E. coli*, and is required for optimal expression in *M. smegmatis*. This region lies within the 262 bp *Hin*dIII-*Sph*I fragment shown in figure 7.1, and will be referred to as the 'upstream activator region' or UAR. The 300 bp *Sph*I-*Bam*HI fragment, immediately downstream of this, contains the two putative promoters identified in Chapter 6, and is described as the promoter region (see Fig. 7.1). The possible mechanisms of transcription enhancement by the upstream activator region are discussed in section 7.5.

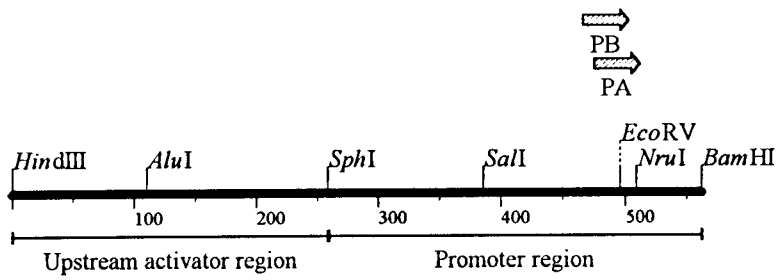


Figure 7.1 Linear restriction endonuclease map of the 562 bp *Hin*dIII-*Bam*HI insert of pJCLK0.6 and pKup0.6. The *Sma*I site on the 5' end of the insert was destroyed during cloning of the fragment into the *Eco*RV restriction endonuclease site of pBluescript-SK. The *Hin*dIII restriction endonuclease site originates from this cloning vector. The restriction endonuclease sites used in the construction of the deletions in this chapter are shown. The positions of the putative promoters, P_A and P_B, identified in Chapter 6, and the positions of the upstream activator and promoter regions are indicated.

The P_{AN} promoter was used in this study to determine whether the UAR of the *M. tuberculosis katG* could increase expression of an unrelated promoter. The promoter was isolated by Murray *et al.* (1992) from *M. paratuberculosis*. It was found upstream of an unknown open reading frame, ORF2, which may form part of the *M. paratuberculosis* IS900 insertion element. The promoter fragment was able to promote efficient expression of the firefly luciferase gene in pJCluc.

7.3. MATERIALS AND METHODS

The bacterial strains, plasmids, and culture conditions used are described in Appendix A. Details of the commonly used techniques and solutions are given in Appendices B and C, respectively.

7.3.1. Deletions in pJCLK0.6 using restriction endonucleases

Derivatives of pJCLK0.6 were constructed, making use of the restriction endonuclease sites present in the insert (see Fig. 7.1). Deletion p0.6DelEN$_{13}$ was constructed as follows: pKup0.6 was digested to completion with *Eco*RV and *Nru*I, both of which produce blunt ends, and religated to form pKup0.6D with the 13 intervening base pairs deleted. The 549 bp *Hin*dIII-*Bam*HI fragment from pKup0.6D was cloned into the *Hin*dIII and *Bam*HI restriction endonuclease sites of pJCluc to produce p0.6DelEN$_{13}$ (see Fig. A.3, Appendix A). Deletion p0.6DelAB$_{111}$ was constructed by cloning of the 448 bp *Alu*I-*Bam*HI fragment of the pKup0.6 insert into pJCluc digested to completion with *Hin*dIII, blunt-ended with the Amersham DNA Blunting kit, and digested to completion with *Bam*HI (see Fig. A.4). To construct deletion p0.6DelAS$_{155}$, the 300 bp *Sph*I-*Bam*HI fragment of pKup0.6 was cloned into the *Sph*I and *Bam*HI restriction endonuclease sites of pUCBM21, to produce pUC210.3. The 111 bp *Hin*dIII-*Alu*I fragment of pKup0.6 was then cloned into pUC210.3 which had been digested to completion with *Sph*I, blunt-ended, and digested to completion with *Hin*dIII, to produce pUC210.4. This construct essentially contains the insert from pKup0.6 with the 155 bp *Alu*I-*Sph*I fragment deleted. The insert of pUC210.4 was removed by digestion with *Hin*dIII and *Bam*HI, and cloned into the corresponding restriction endonuclease sites of pJCluc to form

p0.6DelAS$_{155}$ (see Fig. A.5). To delete the 106 bp *Sal*I-*Eco*RV fragment from the *katG* promoter region, pKup0.6 was digested with *Sal*I and *Eco*RV, blunt-ended, and religated in the absence of the 106 bp insert fragment to produce pKup0.5. The 456 bp *Hin*dIII-*Bam*HI insert from this construct was then cloned into the *Hin*dIII and *Bam*HI restriction endonuclease sites of pJCluc to produce p0.6DelSE$_{106}$ (see Fig. A.6). The inserts of all the constructs described above were sequenced over the relevant areas using the pUC19 forward sequencing primer (Promega), and a reverse primer complementary to sequences immediately downstream of the luciferase ATG translation start codon, to confirm that the appropriate sequences had been removed. The custom-synthesized luciferase reverse primer had the sequence: 5'-CTTTATGTTTTTGGCGTC-3'.

7.3.2. Electroporation of *M. smegmatis* and luciferase assays

The electroporation of *M. smegmatis* and the luciferase assays were performed as described in sections 6.3.4, and 6.3.5, respectively.

7.3.3. Bal31 nuclease deletions in pJCLK0.6

Bal31 nuclease was used to delete successive bases from the 5' end of the 559 bp *M. tuberculosis katG* promoter fragment. A pilot experiment using 3 μg of DNA and doubling dilutions of Bal31 nuclease was performed in order to determine the optimal enzyme concentration. In the scale-up, 30 μg of pJCLK0.6 was digested to completion with *Hin*dIII, extracted with phenol/chloroform/isoamyl alcohol and precipitated with ethanol. The linear fragment was incubated at 37°C with 3 units of Bal31 Nuclease (Boehringer Mannheim) in a reaction containing 0.02 M tris-Cl (pH 8.0), 0.012 M calcium chloride, 0.012 M magnesium chloride, 0.2 mM EDTA (pH 8.0), and 0.6 M sodium chloride. An aliquot was removed before addition of the enzyme, and every 90 seconds thereafter, for 10.5 minutes, and extracted with phenol/chloroform/isoamyl alcohol, and the DNA precipitated with ethanol. The DNA was religated and used to transform competent *E. coli* LKIII cells. Plasmids were prepared from transformants corresponding to the various time points and analyzed by restriction enzyme digestion. The inserts were sequenced using the reverse luciferase primer described in section

7.3.1. To test whether the reductions in luciferase activity observed were not due to the deletion of important vector sequences, the 262 bp *Hin*dIII-*Sph*I fragment of pJCLK0.6 (see Fig. 7.1) was removed, without alteration of any vector sequences. The cloning strategy was as follows: pKup0.6 was digested to completion with *Sph*I, the resulting protruding ends were filled in using the Amersham DNA Blunting kit, and the construct digested with *Bam*HI. The resulting 300 bp fragment was gel-purified and cloned into pJCluc digested to completion with *Hin*dIII, blunt-ended with T4 polymerase (Amersham DNA Blunting kit), and digested to completion with *Bam*HI, to form the construct pJCLK0.3 (see Fig. A.7, Appendix A). To test whether the region upstream of the *Sph*I site was able to promote expression of the luciferase gene on its own, this fragment was cloned separately into pJCluc. The cloning strategy was as follows: the 262 bp *Hin*dIII-*Sph*I fragment of pKup0.6 was cloned into the *Hin*dIII and *Sph*I restriction endonuclease sites of pUCBM21 to produce pUC21E. The insert was removed from this construct using the restriction endonucleases *Hin*dIII and *Bam*HI, and cloned into the corresponding restriction endonuclease sites of pJCluc to yield pJCLKE. Cloning of the fragment via pUC21E resulted in the addition of 56 bp of *lacZ* coding sequence between the *Sph*I restriction endonuclease site of the insert and the *Bam*HI restriction endonuclease site of pJCluc (see Fig. A.8).

7.3.4. Primer extension analysis

Primer extension analysis was performed as described in section 6.3.7, using total RNA isolated from *M. smegmatis* cells harbouring pJCLK0.3, p0.6DelSE$_{106}$, and p0.6DelEN$_{13}$. The synthetic oligonucleotide complementary to sequences near the 5' end of the luciferase gene, described in section 6.3.7, was used in the experiments.

7.3.5. Cloning of the *katG* UAR upstream of the P_{AN} promoter

The transcription enhancing properties of the *M. tuberculosis katG* UAR were tested on the P_{AN} promoter from *M. paratuberculosis*. The construct pANIIIW (Karen Kempsall, Glaxo-Wellcome), consisting of a 170 bp PCR fragment containing the P_{AN} promoter cloned into the T-tailed *Eco*RV restriction endonuclease site of the vector pT7Blue(R) (Novagen), was used as

a source of the promoter. The PCR fragment contains the P_{AN} ribosome binding site, but no ATG start codon. The 218 bp *Hin*dIII-*Bam*HI fragment from pANIIIW was cloned into the *Hin*dIII and *Bam*HI restriction endonuclease sites of pJCluc to form pJCLpan. The use of the vector restriction endonuclease sites resulted in the addition of 45 bp of vector sequence on the 5' end of the insert and 3 bp on the 3' end. In this construct, the P_{AN} promoter is in the correct orientation to promote expression of the luciferase gene. The following strategy was used to clone the *katG* UAR upstream of the P_{AN} promoter: The construct pKup0.6 was digested to completion with *Hin*dIII and *Sph*I to release the 262 bp activator region, which was purified by agarose gel electrophoresis and phenol extraction. This was then ligated to pANIIIW which had been digested to completion with *Sph*I. The linear ligated fragments were cleaned by phenol/chloroform extraction and ethanol precipitation, and digested to completion with *Bam*HI. The digested products were electrophoresed in a 2% (w/v) agarose gel and a fragment of 486 bp containing the *katG* UAR ligated to the P_{AN} promoter was excised. The fragment was purified using the phenol squeeze method described in Appendix B, and cloned into the *Hin*dIII and *Bam*HI sites of pJCluc to form pJCLpanE. The inserts of both pJCLpan and pJCLpanE were verified by sequencing with the pUC19 forward primer (Promega), and the reverse primer complementary to sequences near the 5' end of the luciferase gene described in section 7.3.1.

7.3.6. Gel mobility shift assays

Gel mobility shift assays were performed as described by Ausubel *et al.* (1987). The gel-purified 262 bp *Hin*dIII-*Sph*I and 300 bp *Sph*I-*Bam*HI fragments from pKup0.6 containing the *M. tuberculosis katG* activator and promoter regions, respectively, were used as templates. The fragments were radiolabelled by filling in of the 5' protruding ends generated by digestion with *Hin*dIII and *Bam*HI using Klenow polymerase (Boehringer Mannheim) and dTTP, dGTP, dATP, and [α-^{32}P] dCTP, in a buffer containing 10 mM tris-Cl (pH 7.6), 10 mM magnesium chloride, 50 mM sodium chloride, 1 mM dithiothreitol (DTT), and 50 µg/ml bovine serum albumin (BSA). An *M. smegmatis* cell extract was prepared, as described in section 6.3.5, from a late exponential phase culture. In this method, the cells are lyzed with Cell Culture Lysis Reagent (Promega) and by mechanical disruption in the FastPrep™ FP 120 instrument (Savant Instruments Inc., Farmingdale, NY, U.S.A.). The protein concentration of the extract was

measured with the Bio-Rad Protein Assay kit (Bio-Rad Laboratories, CA U.S.A.). Labelled template DNA was incubated with the cell extract for 15 minutes at 37°C in a binding reaction containing the following: 12% (v/v) glycerol, 300 μg/ml BSA, 4 mM tris-Cl (pH 7.9), 60 mM potassium chloride, 1 mM EDTA, 1 mM DTT, and 0.5 μg herring sperm DNA (Boehringer Mannheim). For competition studies, 1 μg of unlabelled competitor DNA was mixed with the labelled probe before the addition of the extract. The reactions were loaded on a 6% (w/v) non-denaturing polyacrylamide gel, and separated by electrophoresis in a high ionic strength tris-glycine buffer (see Appendix C) at a constant current of 30 mA The gel was dried and exposed to X-ray film.

7.3.7. Preparation and electroporation of *M. bovis* BCG

Electrocompetent *M. bovis* BCG cells were prepared and electroporated using the method of Jacobs *et al.* (1991), with minor variations. A stationary phase culture of *M. bovis* BCG grown in Middlebrook-ADC-Tween was diluted 100-fold in freshly prepared Middlebrook-ADC-Tween, and grown to mid-exponential phase (8 days) on a roller at 37°C. Cells were harvested by centrifugation at 4°C and washed 4-5 times in sterile, deionized water, to remove all traces of medium and salts. Washed cells were resuspended in a final volume of 1 ml of ice-cold 10% (v/v) glycerol/100 ml of cells (original volume). For electroporation, 60 μl of cells were incubated with 240 μl 10% (v/v) glycerol and 1 μg of DNA on ice for 10 minutes. Electroporations were performed in 0.1 cm gap Gene Pulser® cuvettes (Bio-Rad Laboratories) at 1.8 kV, 1000 Ω and 25 μF Ω, using a Gene Pulser™ apparatus (Bio-Rad Laboratories, CA U.S.A.), which resulted in a field strength of 18.0 kV/cm. The cells were incubated for a further 5 minutes on ice before the addition of 700 μl Middlebrook-ADC and incubation for 3-4 hours at 37°C without agitation. Expression mixes were plated on Middlebrook-ADC agar plates containing kanamycin at a concentration of 30 μg/ml and 0.01% (w/v) cycloheximide (Fluka Biochemicals), and incubated at 37°C for three weeks. Plasmids were prepared from the recombinants as described in Appendix B, and verified by restriction analysis.

7.4. RESULTS

7.4.1. Subcloning and testing of the *katG* upstream region

To determine whether other regions of the *M. tuberculosis katG* upstream region were required for promoter activity, a number of deletions were made using the restriction endonuclease sites present on the insert. The regions present in the various deletions and the resulting decreases in luciferase activity are summarized in figure 7.6 (see discussion). Deletion p0.6DelAB$_{111}$ has the first 111 bp up to the *Alu*I site deleted. p0.6DelEN$_{13}$ has the 13 bp *Eco*RV-*Nru*I fragment deleted. In p0.6DelAS$_{155}$ the 155 bp *Alu*I-*Sph*I fragment is deleted, and in p0.6DelSE$_{106}$, the 106 bp *Sal*I-*Eco*RV fragment is deleted. The constructs were electroporated into *M. smegmatis*, and tested for luciferase activity in *E. coli* and *M. smegmatis* (Table 7.1).

Table 7.1. Luciferase activities of the pJCLK0.6 deletion constructs.

	E.coli		*M. smegmatis*	
Construct	**RLU/mg protein[1]**	**Rel. Act.[2]**	**RLU/mg protein[1]**	**Rel. Act.[2]**
pJCLK0.6	8695.44 ± 1714.8	1.000	50291.26 ± 14354.2	1.000
p0.6DelEN$_{13}$	2787.34 ± 697.94	0.3206	38399.91 ± 9793.484	0.7636
p0.6DelAB$_{111}$	105.89 ± 23.00	0.0122	921.31 ± 93.10	0.0183
p0.6DelAS$_{155}$	6101.43 ± 837.54	0.7017	59506.86 ± 5043.60	1.1832
p0.6DelSE$_{106}$	2848.22 ± 726.80	0.3276	49606.81 ± 8819.99	0.9864

[1] The mean of three independent assays ± the standard error of the mean.
[2] Relative Activity- normalized to the luciferase activity of the parent construct, pJCLK0.6

Removal of 111 bp from the 5' end of the insert (p0.6DelAB$_{111}$) resulted in an 82.1- and 54.6-fold reduction in activity in *E. coli* and *M. smegmatis*, respectively, suggesting that this region is essential for optimal promoter activity. Removal of the 13 bp between the *Eco*RV and *Nru*I restriction endonuclease sites in the promoter region resulted in a 3.12-fold decrease in promoter activity in *E. coli*. This deletion removes the putative -10 hexamer associated with transcription start site A, as well as transcription start site B and a region of its putative -10 hexamer (see section 6.4.5). The deletion should, therefore, abolish promoter activity completely. In *M. smegmatis*, deletion of this region had no significant effect on expression. The transcription machinery evidently uses different promoter sequences for transcription initiation in this construct. Removal of sequences upstream of the *Eco*RV site gave a similar

result, i.e. a 3.05-fold decrease in activity in *E. coli*, and no significant change in *M. smegmatis*. Deletion of the *Sal*I-*Eco*RV fragment results in the removal of half of the -10 hexamer of transcription start site B and the -35 sequences of transcription start sites A and B. It is conceivable that, in this construct, the promoter associated with transcription start site A is used to promote expression, and that the presence of a -35 sequence is not essential. Alternatively other promoter sequences may be recognized in place of those previously identified. The absence of the 155 bp *Alu*I-*Sph*I fragment had no significant effect on promoter activity in either *E. coli* or *M. smegmatis*, which suggests that this region is not important for promoter activity. The sequences necessary for maximal promoter activity, therefore, lie upstream of the *Alu*I site (see Fig. 7.6). This region does not contain a promoter but may act as a regulatory region which enhances transcription initiation, i.e. an upstream activator region (UAR). The results also indicate that the distance between the UAR and the promoters is not critical.

7.4.2. Deletion analysis of the *katG* promoter fragment using Bal31 nuclease

In a further attempt to locate the sequences required for promoter activity, the pJCLK0.6 insert was progressively deleted from the *Hin*dIII restriction endonuclease site at the 5' end, using Bal31 nuclease. Bal31 nuclease possesses single strand-specific endonuclease activity and double strand-specific exonuclease activity, which removes successive DNA bases bidirectionally. As a result, the enzyme deleted some vector sequences upstream of the *Hin*dIII site. The deletion reaction was stopped at various time points to give a range of progressively shortened products. DNA from time points 4 and 6 yielded no transformants. One colony from time points 5, 7, and 8, and three from time point 3 were examined further. All six recombinants were sequenced to determine the size of the deletions (Table 7.2). In all the deletions, the successive removal of nucleotides in both directions was uneven. With the exception of deletions p0.6Del123 and p0.6Del283, the Bal31 nuclease enzyme removed more vector sequence than insert sequence. The reason for this is unclear. It may be due to the fact that the insert sequences have a higher G+C content and may form secondary structures, which are more resistant to digestion with Bal31 nuclease. To determine whether the deletion of vector sequences was responsible for the observed differences in luciferase activity, a construct was made containing the 300 bp *Sph*I-*Bam*HI *katG* promoter region in pJCluc. This construct, pJCLK0.3, contains approximately the same amount of insert sequence as p0.6Del283, but has

no vector sequences removed. The plasmid pJCLKE was constructed to determine whether the UAR on its own was able to promote expression of the luciferase gene. These constructs were electroporated into *M. smegmatis*. *E. coli* and *M. smegmatis* cells harbouring the deletion plasmids, as well as the undeleted and vector controls, pJCLK0.6 and pJCluc, respectively, were tested for luciferase activity (Table 7.2).

Table 7.2. Luciferase activities of derivatives of pJCLK0.6.

			Luciferase activity			
	No. bp deleted		*E. coli*		*M. smegmatis*	
Construct	**Insert**	**Vector**	**RLU/mg[1]**	**Rel. Act.[2]**	**RLU/mg[1]**	**Rel. Act.[2]**
pJCluc	-	-	132.30 ± 34.49	0.0838	106.93 ± 13.22	0.0105
pJCLK0.6	-	-	1578.00 ± 257.43	1.000	10164.19 ± 1270.98	1.000
p0.6Del0	0	300	1170.48 ± 52.73	0.7417	11830.69 ± 1010.69	1.1640
p0.6Del18	18	350	1671.54 ± 64.72	1.0593	6024.54 ± 304.03	0.5927
p0.6Del123	123	100	41.46 ± 4.06	0.0263	1585.16 ± 389.48	0.1560
p0.6Del167	167	268	45.84 ± 3.87	0.0290	905.43 ± 83.64	0.0891
p0.6Del173	173	489	18.39 ± 0.09	0.0117	14.28 ± 2.63	0.0014
p0.6Del283	283	271	45.56 ± 5.01	0.0289	972.10 ± 118.31	0.0956
pJCLK0.3	262	0	73.33 ± 9.58	0.0465	2640.14 ± 246.44	0.2597
pJCLKE	300	0	26.506 ± 12.77	0.0168	35.01 ± 8.21	0.0034

[1] The mean of three independent experiments ± the standard error of the mean.
[2] Relative Activity- normalized to the luciferase activity of the parent construct, pJCLK0.6

Removal of the first 18 bp (p0.6Del18) of the insert resulted in approximately a 40% reduction in promoter activity, while the removal of 123 bp from the 5' end (p0.6Del123) resulted in a reduction in luciferase activity to below background levels in *E. coli* (38-fold decrease) and a 6.4-fold reduction in *M. smegmatis*. Removal of a further 160 bp (p0.6Del283) had no additional effect on promoter activity in *E. coli*, but resulted in a 10.5-fold reduction in activity in *M. smegmatis*. These results suggest that the region up to 123 bp from the 5' end of the insert is essential for promoter activity in *E. coli*, and is required for maximal activity in *M. smegmatis*, and that the region between 123 and 283 bp from the 5' end is not additionally important. Deletion p0.6Del0 has no insert sequences deleted, but has 300 bp of vector sequence removed. The construct exhibited similar luciferase activities to the parent plasmid, pJCLK0.6, suggesting that deletion of at least the first 300 bp of vector sequence upstream of

the *katG* promoter fragment has no effect on promoter activity. In addition the construct pJCLK0.3, in which no vector sequences are deleted but 262 bp of insert are removed, exhibited a similar decrease in luciferase activity to p0.6Del123, p0.6Del167, and p0.6Del283. These results indicate that the decreases in luciferase activity are due to removal of essential insert fragments, rather than vector sequences. Deletion p0.6Del173, on the other hand, exhibited larger reductions in luciferase activities (83.3- and 714-fold in *E. coli* and *M. smegmatis*, respectively) than other deletions plasmids in which similar or smaller amounts of insert sequence had been removed. In this plasmid, 489 bp of vector has been removed, and this may have a significant effect on the stability or copy number of the plasmid. Therefore, in this case, the reduction in luciferase activity cannot be directly correlated with the absence of insert sequences. The below background luciferase activity of cells harbouring pJCLKE suggests that the 262 bp *Hin*dIII-*Sph*I UAR contains no promoter sequences capable of promoting expression of the luciferase gene. The MIC of kanamycin for *M. smegmatis* and *E. coli* cells harbouring a selection of the restriction endonuclease- and Bal31-generated deletion constructs were determined as a measure of the relative copy number of the relevant plasmids (results not shown). *E. coli* cells harbouring the plasmids were far more resistant to kanamycin than the *M. smegmatis* recombinants, as expected, however, there were no significant differences in the MICs of the cells harbouring different deletion constructs in either of the two species. Thus the effects of the deletions on the luciferase activities observed were not due to differences in the copy number of the plasmid.

7.4.3. Primer extension analysis of selected deletions

It was demonstrated in section 7.4.1, that the *M. tuberculosis katG* upstream region could promote expression of the luciferase gene in the absence of some of the putative promoter sequences identified in section 6.4.5. To test whether alternative promoter sequences were being employed in these cases, primer extension analysis was performed on total RNA extracted from *M. smegmatis* cells harbouring the deletions pJCLK0.3, p0.6DelEN_{13}, and p0.6DelSE_{106} (Fig. 7.2). In all three cases, transcription initiated at a C residue located 23 bp upstream of the translation initiation codon. Therefore, in the absence of the UAR or the usual promoter sequences, a third promoter on the *katG* fragment, designated P_C, becomes active.

Located 7 bp upstream of transcription start site C, is the putative -10 hexamer, CACAGC (Fig. 7.3). This matches the putative -10 hexamers of the *M. bovis* 18 kDa (Terasaka *et al.*, 1989) and MBP70 (Dale and Patki, 1990) genes in 4/6 positions. Associated with this are two putative -35 regions, 1) TTCGCG, 14 bp upstream of the -10 hexamer, and 2) TCCGAC, 22 bp upstream of the -10 hexamer. The first matches in 4/6 positions the -35 hexamers of the *M. leprae* SOD (Dale and Patki, 1990), the mycobacteriophage L5 gene 71 P1(Nesbit *et al.*, 1995), and the *M. smegmatis rpsL* (Kenney and Churchward, 1996) genes. The second matches in 4/6 positions the *M. smegmatis rpsL* (Kenney and Churchward, 1996), and the mycobacteriophage L5 gene 71 P1(Nesbit *et al.*, 1995) -35 hexamers.

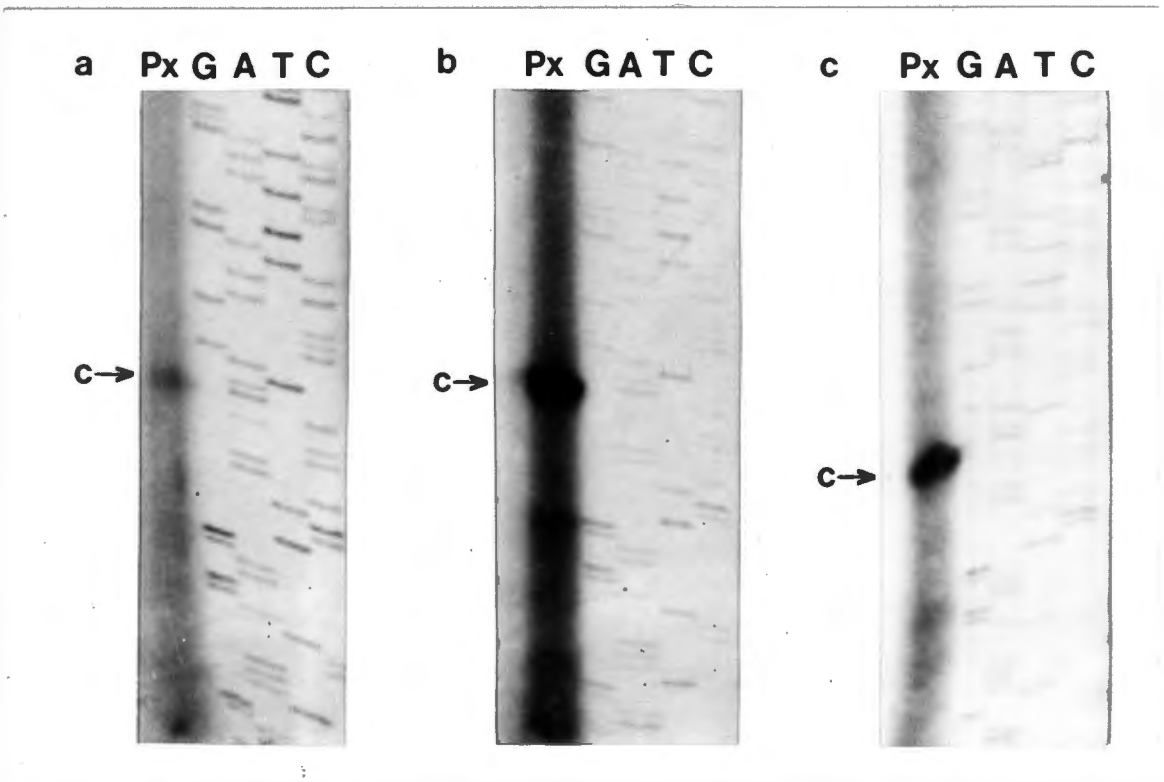


Figure 7.2 Primer extension analysis of the *M. tuberculosis katG* upstream region using RNA isolated during late exponential phase from *M. smegmatis* cells harbouring a) pJCLK0.3, b) p0.6DelEN$_{13}$, and c) p0.6DelSE$_{106}$. The potential transcription start site, tssC, is indicated with an arrow. This corresponds to a C residue, 23 bp upstream of the GTG translation initiation codon.

```
GGCTTCCTGTTGGACGAGGCGGAGGTCATCTACTGGGGTCTATGTCCTGATTGTTCGATA    1913
-35₂      -35₁                  -10          Pc          R.B.S.
TCCGACACTTCGCGATCACATCCGTGATCACAGCCCGATAACACCAACTCCTGGAAGGAA    1973
                                              ↑
TGCT GTG CCC GAG                                                1987
     M   P   E
```

Figure 7.3 The nucleotide sequence of the *M. tuberculosis katG* promoter region associated with tssC, including the first 3 codons, labelled with respect to the published sequence (GenBank Accession number X68081). The transcription start site is indicated with an arrow. The putative -10 and -35 hexamers are underlined, and the putative ribosome binding site is in boldprint

7.4.4. Testing of the *katG* UAR on the P_{AN} promoter.

The possibility that the 5' end of the insert of pJCLK0.6 contains a region which enhances transcription was tested by cloning the 262 bp *Hin*dIII-*Sph*I fragment from pKup0.6 (see Fig. 7.1) upstream of the *M. paratuberculosis* P_{AN} promoter. The M. *paratuberculosis* P_{AN} promoter was cloned into pJCluc on its own (pJCLpan) and downstream of the *Hin*dIII-*Sph*I *katG* UAR (pJCLpanE). The constructs were electroporated into *M. smegmatis* and cell extracts from the resulting transformants were tested for luciferase activity.

Table 7.3. Luciferase activities of the *katG* and P_{AN} promoters in the presence and absence of the *katG* UAR.

	E. coli		*M. smegmatis*	
Construct	**RLU/mg protein[1]**	**Rel. Act.[2]**	**RLU/mg protein[1]**	**Rel. Act.[2]**
pJCLK0.3	73.33 ± 9.58	1.00	2640.14 ± 246.44	1.00
pJCLK0.6	1897 ± 569.88	25.88	11029.62 ± 1297.62	4.18
pJCLpan	278.36 ± 71.79	1.00	9944.93 ± 1881.99	1.00
pJCLpanE	4085.67 ± 1407.4	14.68	115594.49 ± 24219.2	11.62

[1] The mean of three independent assays ± the standard error of the mean.
[2] Rel. Act. - Luciferase activity relative to the construct without the *katG* UAR

The presence of the *katG* UAR resulted in a 14.7-fold increase in expression of the P_{AN} promoter in *E. coli*, and an 11.6-fold increase in *M. smegmatis*. This result confirms the suggestion that the region acts as an enhancer of transcription initiation.

7.4.5. Gel mobility shift assays.

The 262 bp *Hin*dIII-*Sph*HI fragment from pKup0.6, containing the *katG* UAR, was used as a template in DNA binding studies with total cellular proteins extracted from *M. smegmatis* (Fig. 7.4). Incubation of the probe with increasing amounts of cell extract resulted in the formation of two DNA-protein complexes, C_1 and C_3, seen as shifts in the electrophoretic mobility of the probe. All of the probe was bound in the presence of 30 μg of extract. The addition of an excess of unlabelled non-specific competitor DNA (pUCBM21) had no effect on the formation of the DNA-protein complexes, suggesting that the binding of the probe to the protein/s in the cell extract was sequence-specific. The addition of unlabelled pKup0.6, however, also had no effect on complex formation. This construct contains the UAR present in the probe, as well as the downstream promoter region, and should result in full competition. The addition of a competitor fragment containing the same sequences as the probe (pUC21E), resulted in a further retardation in the mobility of the probe in the gel to form complex 2 (C_2), while the addition of competitor DNA containing the *Sph*I-*Bam*HI *katG* promoter region resulted in a small amount of competition for the bound protein of complex 1. These results suggest that one or more proteins in the *M. smegmatis* extract bind to the *katG* UAR, and that the promoter region binds to the same protein and partially competes for it if added on a plasmid. The addition of the unlabelled probe fragment should have also resulted in full competition for the protein and loss of the complex, but, instead, the size of the complex changed. This suggests that a new complex was formed, possibly containing both the labelled and unlabelled probe fragments bound together with one or more *M. smegmatis* proteins. This is indicative of co-operative binding, rather than competition. The formation of complex 3 was not inhibited by any competitor DNA fragments.

Similar results were observed using the *Sph*I-*Bam*HI promoter region as a template (Fig. 7.5). The addition of an *M. smegmatis* cell extract resulted in the formation of two complexes, C_1 and C_3. Unlabelled pUCBM21 DNA competed with the labelled template for the formation of both complexes. These may, therefore, be the result of binding of the RNA polymerase holoenzyme to the promoter sequences. The same holoenzyme could be expected to bind to the promoters present in pUCBM21. The addition of pKup0.6 had no effect on the formation of either complex. This is surprising, since the plasmid contains the same promoter region used as a template as well as the vector sequences of pUCBM21. The addition of unlabelled pUC21E

resulted in an increase in the size of complex 1, to form complex 2 (C_2), as noted for the *Hin*dIII-*Sph*I template, and an increase in the amount of complex 3, while the addition of unlabelled pUC210.3 resulted in partial competition for the bound proteins of complex 1 and 3. These results suggest that the upstream activator and the promoter regions both bind to the same protein/s in the cell extract, and that when both regions are present on the same fragment, binding of the protein/s to either region is reduced. The fragment containing the UAR and promoter region may contain an intrinsic bend which is not present when the two regions are cloned separately. This may explain the fact that the regions on their own have different effects on complex formation to when they are added on the same plasmid. The competitor DNAs were added in plasmid form, since this is the form of the promoter fragments in the expression studies. It is possible that the affinity and binding specificity of a DNA binding protein is affected by differences in DNA bending and superhelicity.

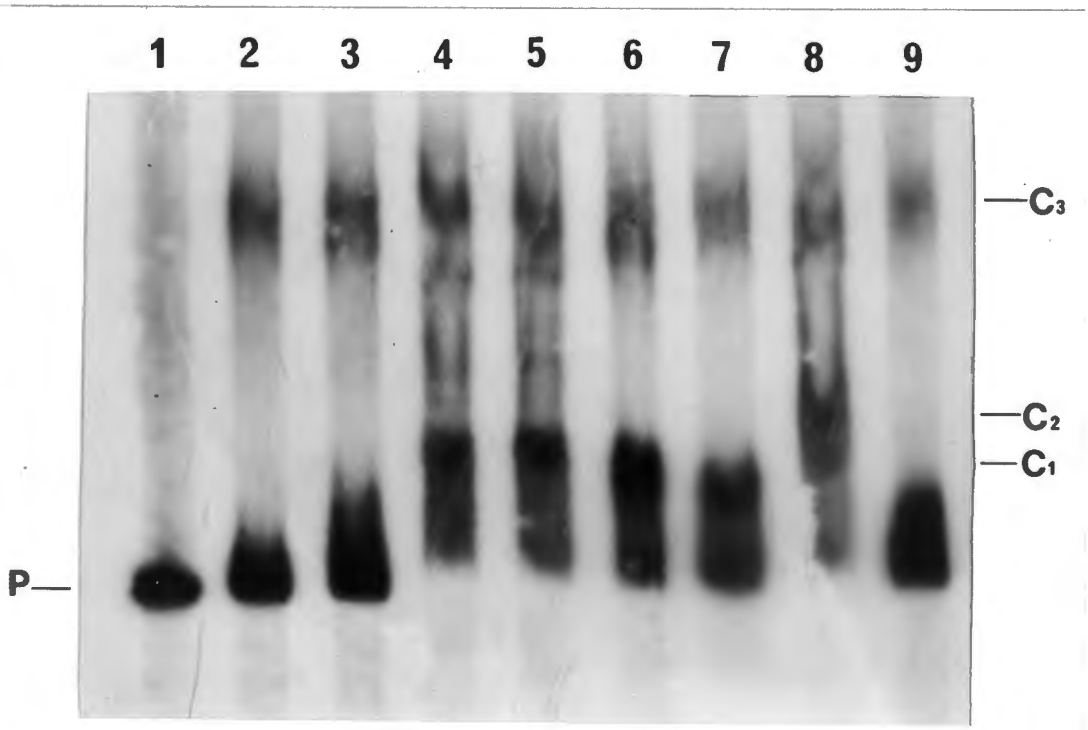


Figure 7.4 Mobility shift assays with the *Hin*dIII-*Sph*I *M. tuberculosis katG* UAR. The lanes were loaded as follows: 1) unbound template; 2) to 5) 10, 20, 30, and 40 μg of *M. smegmatis* cell extract, respectively; 6) to 9) 30 μg of *M. smegmatis* cell extract and 1μg unlabelled competitor DNA; 6) pUCBM21, 7) pKup0.6, 8) pUC21E and 9) pUC210.3. Labelled template DNA was included in all the reactions. P- unbound probe. C_1, C_2, and C_3- complexes 1, 2, and 3, respectively.

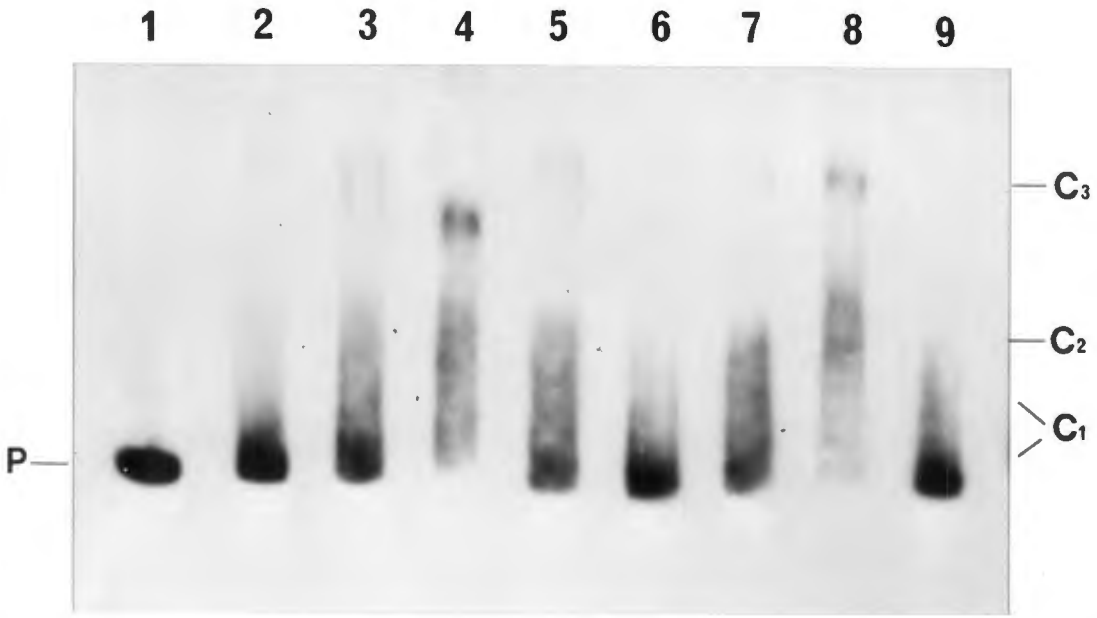


Figure 7.5 Mobility shift assays with the *Sph*I-*Bam*HI *M. tuberculosis katG* promoter region. The lanes were loaded as follows: 1) unbound template; 2) to 5) 10, 20, 30, and 40 μg of *M. smegmatis* cell extract, respectively; 6) to 9) 30 μg of *M. smegmatis* cell extract and 1μg unlabelled competitor DNA; 6) pUCBM21, 7) pKup0.6, 8) pUC21E and 9) pUC210.3. Labelled template DNA was included in all the reactions. P- unbound probe. C_1, C_2, and C_3- complexes 1, 2, and 3, respectively.

7.4.6. Electroporation and testing of the pJCluc constructs in *M. bovis* BCG

To determine whether the same sequences were required for the expression of the *M. tuberculosis katG* promoter in a slow-growing host, the following constructs were electroporated into *M. bovis* BCG and tested for luciferase activity (Table 7.4): pJCluc, pJCLK0.6, pJCLK0.3, pJCLKE, p0.6Del167, pJCLpan, and pJCLpanE.

Table 7.4 Luciferase activities of the pJCluc constructs in *M. bovis* BCG.

Construct	RLU/mg protein[1]
pJCluc	12.02 ± 2.563
pJCLK0.6	25506.05 ± 2409.78
pJCLK0.3	328.57 ± 53.42
pJCLKE	45.85 ± 7.73
p0.6Del167	3266.05 ± 1761.33
pJCLpan	4832.02 ± 1331.26
pJCLpanE	21423.90 ± 595.33

[1] The mean of three independent experiments ± the standard error of the mean.

The background luciferase activity resulting from transcriptional readthrough in pJCluc was low in *M. bovis* BCG. The activity of the *katG* promoter in this host was, however, comparable to that in *M. smegmatis*. Removal of the *katG* UAR (pJCLK0.3) resulted in a 77.63-fold decrease in luciferase activity, indicating that the region also plays an important role in promoter activity in this host. The Bal31 nuclease deletion plasmid, p0.6Del167, exhibited a 7.8-fold decrease in luciferase activity relative to the parent construct, pJCLK0.6. This plasmid contains an extra 93 bp of insert sequence relative to pJCLK0.3, which may explain the smaller reduction in luciferase activity. The low background levels of luciferase activity resulting from pJCLKE indicates, as shown before, that the *Hin*dIII-*Sph*I UAR of *katG* contains no active promoter sequences. The activity of the *M. paratuberculosis* P_{AN} promoter increased 4.4-fold in the presence of the *katG* UAR. This further confirms the importance of the region in enhancing transcription in this host.

7.5. DISCUSSION

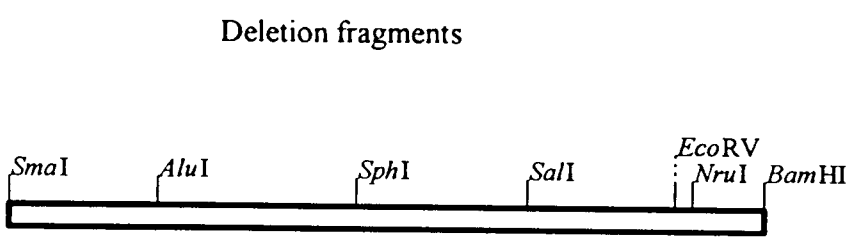
The construction of deletions in pJCLK0.6, making use of the restriction endonuclease sites present on the insert, facilitated the assessment of the relative importance of various regions of the insert in promoter activity in *E. coli* and *M. smegmatis*. The successive deletion of sequences from the 5' end of the 559 bp *M. tuberculosis katG* promoter fragment in pJCLK0.6, using Bal31 nuclease, confirmed the observed results. The various deletion plasmids constructed in this study, and the relative luciferase activities are summarized in figure 7.6.

The first 111 bp at the 5' end of the *Hin*dIII-*Bam*HI *M. tuberculosis katG* promoter fragment are essential for promoter activity in *E. coli*, and are required for optimal expression in *M. smegmatis*. Mobility shift assays revealed that one or more proteins from a total *M. smegmatis* extract bind to this region. The requirement for the region may, therefore, be due to necessary interactions between the protein bound to this region and the RNA polymerase holoenzyme bound to the promoters. As mentioned in Chapter 6, the first nucleotide of both the putative -10 hexamers of the *katG* promoter deviates from the consensus sequence, which normally has a conserved T in this position. *E. coli* promoters which deviate in this position generally require the help of activators for efficient transcription initiation. These activators normally bind to sites upstream of the -35 region (Raibaud and Schwartz, 1984). Activators

which bind further upstream of the -35 region, such as the CAP, may induce a bend in the intervening DNA, to facilitate interaction of the activator with the RNA polymerase bound at the -35 region, to activate transcription (Zhou *et al.*, 1994a; 1994b). It is possible, therefore, that a protein binds to the *katG* UAR, and facilitates the binding of the RNA polymerase holoenzyme further downstream, and the formation of the open complex for transcription initiation. The results of the mobility shift assays suggest the possibility of co-operative binding of the protein to the DNA. It is possible that the protein/s bind to the DNA in more than one region, and that the interaction of the subunits or different proteins bound to the different regions results in a bend in the intervening sequences. The protein/s evidently also bind to the downstream promoter region, which suggests an even more complicated scenario. The experiments reported here do not conclusively prove that binding of proteins to the UAR and promoter regions is sequence-specific. This will require further experimentation, which should preferably be carried out using proteins isolated from the native host, *M. tuberculosis*. Purification of the protein/s which bind to the *katG* fragments, DNase I footprinting analysis to locate the binding sites, and competition studies to determine the relative binding efficiencies of the protein/s to the different regions, will be required in order to elucidate the mechanism of transcription enhancement employed here.

In the absence of either the UAR, or the putative *katG* promoters, P_A and P_B, the *M. smegmatis* machinery recognizes and utilizes a third promoter, P_C. It is possible that a similar situation occurs in *E. coli*. This promoter apparently functions as efficiently as the other two in *M. smegmatis*, since the luciferase activities associated with the constructs p0.6DelEN$_{13}$ and p0.6DelSE$_{106}$ are comparable to that of the parent construct, pJCLK0.6, in which P_A and P_B are utilized. The presence of the UAR is, however, also essential for the optimal activity of promoter C, since *M. smegmatis* cells harbouring pJCLK0.3, which contains P_A and P_B but no UAR, exhibit a 4-fold reduction in luciferase activity relative to the undeleted parent plasmid, pJCLK0.6. If promoter P_C is used to promote expression of the luciferase gene in p0.6DelEN$_{13}$ and p0.6DelSE$_{106}$ in *E. coli*, then the promoter is evidently not as efficient as promoters P_A and P_B in this host, since *E. coli* cells harbouring these constructs exhibit a 3-fold lower luciferase activity than the parent plasmid, pJCLK0.6. The first putative -35 hexamer of P_C, TTCGCG, matches the consensus *E. coli* -35 region in 3/6 positions, the first two of which are highly conserved among *E. coli* σ^{70} promoters. The -10 hexamer, on the other hand, contains only one

of the highly conserved nucleotides of the *E. coli* consensus sequence, A at position 2. The spacings between the -10 and putative -35 regions also differ significantly from that found in most *E. coli* promoters (17 ± 1) (Harley and Reynolds, 1987). These factors may explain the reduced efficiency of transcription initiation in this host.



	Relative Decrease in Luciferase activity	
	M. smegmatis	*E. coli*
	1	1
p0.6Del123	6.4	38.1
p0.6Del167	11.2	34.4
p0.6Del173	712	85.8
p0.6Del283	10.5	34.6
p0.6DelAB111	54.6	82.1
p0.6DelEN13	1.3	3.1
p0.6DelAS155	0.85	1.4
p0.6DelSE106	1.01	3.1
pJCLKE	290	59.5
PJCLK0.3	3.85	21.5

Figure 7.6 Summary of results obtained from deletion analysis of the 559 bp *M. tuberculosis katG* promoter region.

The absence of the 155 bp *Alu*I-*Sph*I fragment had no significant effect on promoter activity in either *E. coli* or *M. smegmatis*. This indicates that this region is not important for promoter activity, and that the distance between the UAR and the promoter sequences is not critical. The sequences necessary for maximal promoter activity, therefore, lie upstream of the *Alu*I site (see Fig. 7.6). This region does not contain a promoter but may act as a regulatory region which enhances transcription initiation. Within this region, lies a 45 bp sequence (-523 to -478 with respect to the GTG translation initiation codon), which has an A+T content of 62.2%, and a 24 bp sequence (-523 to -499 with respect to the GTG translation initiation codon), with an A+T content of 66.67%. These A+T contents are high in comparison to normal *M. tuberculosis* promoter regions (Bashyam *et al.*, 1996), and may be analagous to the UP elements found upstream of some *E. coli* promoters. These elements are usually AT-rich and can increase promoter activity up to 30-fold (Ross *et al.*, 1993). These sequences are a frequent feature of gram-positive promoters, and may be involved in modulating promoter strength (Graves and Rabinowitz, 1986; Moran *et al.*, 1982). Movahedzadeh *et al.* (1997b) have recently identified a 41 bp region, located 269-310 bp upstream of the coding region of the *M. tuberculosis recA* gene, which is essential for expression from the promoter. The region on its own exhibited no promoter activity, and the authors suggest that it may contain a binding site for a transcription-activating protein. This region may, therefore, be analogous to, or perform the same function as that identified in this study. A database search of the *M. tuberculosis katG* UAR revealed that it is not dispersed in the *M. tuberculosis* genome, but that it is similar to the region upstream of the *M. fortuitum katG* gene. The 559 bp *katG* upstream sequence is 61% homologous to the corresponding region from *M. fortuitum*. In addition, the 24 bp AT-rich sequence of the *katG* UAR, described above, is 79.2% homologous to a region located 489 bp upstream of the *M. fortuitum katG* ATG translation initiation codon. This suggests possible similarities in the mode of regulation of the gene in the two mycobacterial hosts.

As mentioned in section 6.4.2, Song and Deretic have identified an open reading frame upstream of the *M. tuberculosis katG* gene with homology to the Fur protein. The translation start codon of this putative open reading frame lies just within the 45 bp AT-rich sequence described above, while the stop codon lies immediately downstream of the *katG* tssA. This overlap of the FurA ORF with the *katG* promoter region and UAR may have implications for the regulation of both genes. Transcription of either of the genes will interfere with transcription of the other. If

the *furA* promoter lies immediately upstream of the gene, within the *katG* UAR, then transcription of the former gene will regulate the availability of the UAR, and thus the expression of *katG* and the utilization of the different *katG* promoters. This suggests a possible coupling of the regulation of oxidative stress and iron metabolism genes in this organism, and warrants further investigation.

The fact that the pJCluc promoter constructs in this study exhibited similar luciferase activities in *M. bovis* BCG and *M. smegmatis* was not surprising. Dellagostin *et al.* (1995) also found no significant difference between promoter activity in *M. smegmatis* and *M. bovis* BCG. In addition, Bashyam *et al.* (1996) have demonstrated that the efficiency and specificity of transcriptional recognition is conserved in *M. tuberculosis*, *M. smegmatis*, and *M. bovis* BCG. *M. smegmatis* is, therefore, a suitable host for the preliminary characterization of *M. tuberculosis* promoters.

CHAPTER 8

GENERAL CONCLUSIONS

Several projects are in progress to sequence the *M. tuberculosis* genome. Results from these projects will provide a great deal of information on both the genetics of the organism, and its metabolic pathways. In addition, they will identify new avenues for biochemical and immunological investigation, and open up new leads for research in chemotherapy and drug design. However, sequencing data alone will not identify the physiological functions of the gene products or the mechanisms involved in their regulation. This requires further information based on molecular biological experimentation. In this study, molecular biological techniques have been used to identify a novel function of the *M. tuberculosis katG* gene and to examine the mechanisms of regulation of the gene.

The *M. tuberculosis katG* gene plays an important role in the susceptibility of the organism to the antituberculosis drug, isoniazid, and in the virulence of the organism. The gene has been extensively studied with respect to its catalase and peroxidase activities, its role in virulence, and its role in susceptibility to isoniazid, however, this study provides the first report on the involvement of the gene in DNA repair processes and on the in-depth analysis of the regulation of expression of the gene. The following conclusions were drawn from the results obtained in the respective chapters:

Chapters 3 and 4

- The *M. tuberculosis* KatG remains active when fused to the maltose binding protein. The dimerization and folding of the enzyme into the correct conformation for activity, therefore, can occur in the presence of large additions to the N-terminal end of the protein.
- Removal of either the proximal two thirds or the terminal third of the protein results in a complete loss of catalase and peroxidase activity. The active sites lie in the N-terminal domain of the protein, but the presence of the C-terminal domain is essential, possibly for subunit-subunit interaction, tertiary folding of the protein, or substrate interaction.

- The KatG enzyme probably contains a cleavage recognition site for an *E. coli* protease, and is cleaved immediately after translation. The KatG cleavage product does not exhibit catalase or peroxidase activity.

- The 2 putative ORFs downstream of the *M. tuberculosis katG* gene either do not encode proteins, or do not contain promoters which function efficiently in *E. coli*.

- Immediately downstream of the *M. tuberculosis katG* gene is a 7 bp inverted repeat with dyad symmetry which may act as a transcriptional terminator.

- The KatG protein enhances the survival of DNA repair mutants of *E. coli* after exposure to short wavelength UV light irradiation, mitomycin C, and hydrogen peroxide, possibly by enhancing excision repair pathways. This function is independent of its catalase and peroxidase activity, and, therefore, is not due to the removal of toxic oxygen species by the enzyme. KatG may stabilize intermediates in the excision repair pathway.

- In repair-competent mycobacterial hosts, KatG does not play an additive role in repair of DNA damage caused by exposure to short wavelength UV light irradiation. These cells possess efficient DNA repair systems, therefore, any involvement of KatG in the repair processes may be masked by the high survival rate of the cells.

Chapters 6 and 7:

- The sequences required for promoting expression of the *M. tuberculosis katG* gene are located within a 559 bp region immediately upstream of the gene. The sequences between 559 and 448 bp upstream of the translation initiation codon are essential for promoter activity in *E. coli* and are required for maximal expression from the promoter in *M. smegmatis*. This region is referred to in this study as the upstream activator region (UAR). The UAR contains a 45 bp AT-rich sequence which may be analogous to the AT-rich regions found upstream of some *E. coli* promoters, which enhance transcription. A similar region is present upstream of the *M. fortuitum katG* gene, and the gene may be regulated in a similar manner in the two mycobacterial species.

- One or more proteins from a total *M. smegmatis* extract bind to the *M. tuberculosis katG* UAR. These protein/s also bind to a region further downstream, possibly in a co-operative manner.

- The UAR is able to enhance transcription from the *M. paratuberculosis* P_{AN} promoter, in *E. coli*, *M. smegmatis*, and *M. bovis* BCG. The distance between the UAR and the promoter is not critical.

- The region between 448 and 293 bp upstream of the translation initiation codon is not required for promoter activity.

- The *M. tuberculosis katG* utilizes two promoters, which are equally efficient, for expression under normal growth conditions. A third promoter is active in the absence of the UAR or the other two promoter sequences. All three promoters share strong similarity with other mycobacterial promoters and contain at least one of the three conserved bases of *E. coli* -10 and -35 consensus sequences. These promoters, therefore, function in *E. coli*, and are probably transcribed by the $E\sigma^{70}$ RNA polymerase holoenzyme.

- Promoter activity peaks during late exponential phase and declines during stationary phase, and is induced by ascorbic acid and repressed by growth in oxygen limiting conditions and at elevated temperatures.

- The *M. tuberculosis katG* promoter shows similar activity in *M. smegmatis* and *M. bovis* BCG. In both hosts, the same regions upstream of the *katG* gene are required for optimal expression.

FUTURE PROSPECTS

A study of this kind is never complete. Each observation raises further questions, many of which warrant further investigation. For example, the extrapolation of the DNA repair activity of the KatG protein to mycobacteria. This will require the overexpression of the gene in DNA repair-deficient mutants of *M. tuberculosis*, *M. bovis* BCG, or *M. smegmatis*, once they become available, and testing the resulting transformants for sensitivity to a variety of DNA damaging agents. An investigation of the mechanism of enhancement of survival will require an in-depth

examination of the DNA lesions resulting from the DNA damaging agents, and their possible interaction with the KatG protein.

The regulation of the *M. tuberculosis katG* gene may be complex, and may involve a number of cellular factors. An investigation of the nature of the protein/s which bind to the upstream sequence, the DNA sequences involved in binding, and the role played by competition between two different regions for binding may shed more light on the mechanisms of regulation, and of transcription enhancement by the UAR. It is important that these experiments are performed in *M. tuberculosis*, since this organism probably has some unique regulatory mechanisms not present in *M. smegmatis*.

APPENDIX A

BACTERIAL STRAINS, PLASMIDS, AND CULTURE CONDITIONS

A.1. BACTERIAL STRAINS

The *E. coli* strains used in this study are summarized in table A.1, and the mycobacterial strains are summarized in table A.2.

Table A.1 *E. coli* strains used in this study.

Strain	Genotype	Reference
LKIII	*recA*, *lacI*$^-$, *lacZ*, *lacY*$^+$	Zabeau & Stanley, 1982
UM255	*katG*, *katE*	Mulvey *et al.*, 1988
MC1060	*recA*$^+$	Meissner *et al.*, 1987
DK1	*recA*(del), *recE*	ATCC 35691 Willis *et al.*, 1981
CC118	*recA1*, *recE*	Manoil and Beckwith, 1985
AB1157	*uvr*$^+$, *recE*(del)	Howard-Flanders *et al.*, 1966
AB1884	*uvrC34*, *recE*(del)	Howard-Flanders *et al.*, 1966
AB1885	*uvrB5*, *recE*(del)	Howard-Flanders *et al.*, 1966
AB1886	*uvrA6*, *recE*(del)	Howard-Flanders *et al.*, 1966
HB101	*recA13*, *recE*	Boyer and Roulland-Dussoix, 1969
JC5519	*recE*(del), *recB21*, *recC22*	A. J. Clarke
JC8679	*recB21*, *recC22*	ATCC 47001

Table A.2 Mycobacterial strains used in this study.

Strain	Description	Reference
M. smegmatis LR222	High frequency transformation mutant	Beggs *et al.*, 1995
M. fortuitum	ATCC 6841	ATCC[1]
M. intracellulare	ATCC 13950	ATCC[1]
M. intracellulare 1403 *katG*$^+$	*katG*$^+$ parent	Marklund, Stockholm Sweden
M. intracellulare 1403 *katG*$^-$	*katG*$^-$ mutant	Marklund, Stockholm Sweden
M. bovis BCG, Tokyo	Attenuated vaccine strain 172	State Vaccine Institute, Pinelands, S. A.
M. tuberculosis H37Rv	Virulent laboratory strain. ATCC 27294	ATCC[1]
M. tuberculosis H37Ra	Avirulent laboratory strain. ATCC 25177	ATCC[1]
M. tuberculosis 34	INH-resistant clinical isolate, *katG* (del)	Heym *et al.*, 1994
M. tuberculosis 255	INH-susceptible clinical isolate, no detectable mutations in *katG*	T.C. Victor, University of Stellenbosch, S.A.
M. tuberculosis 463	INH-resistant clinical isolate, mutations in codon 463 and 315 of *katG*	T.C. Victor, University of Stellenbosch, S.A.
M. tuberculosis 519	INH-resistant clinical isolate, mutation in codon 315 of *katG*	T.C. Victor, University of Stellenbosch, S.A.

[1] American Type Culture Collection (Rockville, Md.).

A. 2. PLASMIDS

The vectors used in this study are summarized in table A.3, and the recombinant plasmids constructed during the study are described in tables A.4, A.5, and A.6. The recombinant plasmids are listed according to the chapter in which they were described.

Table A.3 Vectors and cloning systems used in this study.

Plasmid	Size (bp)	Markers	Description	Reference
pUC18	2686	Amp^R, *lacZ*$^+$	M13-based cloning vector[1]	Messing, 1983; Yanisch-Perron *et al.*, 1985
pUC19	2686	Amp^R, *lacZ*$^+$	M13-based cloning vector[1]	Messing, 1983; Yanisch-Perron *et al.*, 1985
pUCBM21	2725	Amp^R, *lacZ*$^+$	M13-based cloning vector (Boehringer Mannheim)	Messing, 1983; Yanisch-Perron *et al.*, 1985
pBluescript-SK	2959	Amp^R, *lacZ*$^+$	Stratagene cloning vector	Short *et al.*, 1988
pMAL™-c2	6646	Amp^R, *lacZ*$^+$	Protein fusion expression and purification system	New England Biolabs
pGEM®-*luc*	4933	Amp^R, *lacZ*$^+$	Luciferase cassette vector	Promega
pEcoR252	3349	Amp^R, *Eco*RI	Suicide vector, derivative of pEcoR251	Zappe *et al.*, 1986
pJC86	5720	Kan^R, *lacZ*$^+$	*E. coli*-mycobacterial shuttle vector	Jack Crawford, CDC Atlanta U.S.A.

[1] pUC18 and pUC19 differ in the orientation of the multiple cloning site with respect to the *lacZ* gene.

Table A.4 Recombinant plasmids described in Chapter 3.

Construct	Size (bp)	Description
pD3	5096	1747 bp *Pst*I fragment from *M. tuberculosis* H37Rv in pEcoR252
pSNS100	4433	1747 bp *Pst*I fragment from *M. tuberculosis* H37Rv in pUC19
pSMK3'	8393	1747 bp *Pst*I fragment from pD3 in pMAL™-c2, in-frame with *malE*
pSMK3'$_{opp}$	8393	1747 bp *Pst*I fragment from pD3 in pMAL™-c2, in the opposite orientation to *malE*
pSMK3'a	8390	1747 bp *Pst*I fragment from pD3 in pMAL™-c2, in frame 2, with respect to *malE*
pSMK3'b	8396	1747 bp *Pst*I fragment from pD3 in pMAL™-c2, in frame 3, with respect to *malE*
pSMKd	7607	961 bp *Alu*I-*Pst*I fragment of the 1747 bp pD3 insert in pMAL™-c2
pSMKA	8888	2242 bp *Eco*RI- and *Xba*I-digested *katG* PCR product in pMAL™-c2
pSMK5'	8132	1486 bp *Eco*RI-*Pst*I fragment from pSMKA (5' end) in pMAL™-c2
pSMK3'*	7402	756 bp *Pst*I fragment from pSMKA (3' end) in pMAL™-c2

Table A.5 Recombinant plasmids described in Chapter 6.

Construct	Size (bp)	Description
pJCluc	7427	Promoterless luciferase gene in pJC86
pJCluc$_{opp}$	7427	Luciferase gene in pJC86, in opposite orientation to that in pJCluc
pJClucK	9370	1943 kb *katG* upstream region in pJCluc (*Hin*dIII-*Bam*HI)
pJCLK0.6	7956	559 bp *katG* upstream region in pJCluc (*Sma*I-*Bam*HI)
pKup	4629	1943 bp *katG* upstream region in pUC18 (*Hin*dIII-*Bam*HI)
pKup0.6	3518	559 bp *katG* upstream region in pBluescript-SK (*Sma*I-*Bam*HI)

Table A.6 Recombinant plasmids described in Chapter 7.

Construct	Size (bp)	Description
p0.6DelEN$_{13}$	7943	pJCLK0.6 with the 13 bp *Eco*RV-*Nru*I fragment deleted
p0.6DelAB$_{111}$	7845	pJCLK0.6 with the 111 bp *Hin*dIII-*Alu*I fragment deleted
p0.6DelAS$_{155}$	7801	pJCLK0.6 with the 155 bp *Alu*I-*Sph*I fragment deleted
p0.6DelSE$_{106}$	7850	pJCLK0.6 with the 106 bp *Sal*I-*Eco*RV fragment deleted
pUC210.3	3025	300 bp *Sph*I-*Bam*HI promoter fragment in pUCBM21
pUC210.4	3136	111 bp *Hin*dIII-*Alu*I fragment in pUC210.3
pKup0.6D	3506	pKup0.6 with the 13 bp *Eco*RV-*Nru*I fragment deleted
pKup0.5	3413	pKup0.6 with the 106 bp *Sal*I-*Eco*RV fragment deleted
p0.6Del0	7656	pJCLK0.6 with 0 bp insert and 300 bp vector deleted[1]
p0.6Del18	7588	pJCLK0.6 with 18 bp insert and 350 bp vector deleted[1]
p0.6Del123	7733	pJCLK0.6 with 123 bp insert and 100 bp vector deleted[1]
p0.6Del167	7521	pJCLK0.6 with 167 bp insert and 268 bp vector deleted[1]
p0.6Del173	7294	pJCLK0.6 with 173 bp insert and 489 bp vector deleted[1]
p0.6Del283	7402	pJCLK0.6 with 283 bp insert and 271 bp vector deleted[1]
pJCLKE	7689	262 bp *Hin*dIII-*Sph*I *katG* UAR[2] in pJCluc
pUC21E	2987	262 bp *Hin*dIII-*Sph*I *katG* UAR in pUCBM21
pJCLK0.3	7727	300 bp *Sph*I-*Bam*HI *katG* promoter fragment in pJCluc
pJCLpan	7646	218 bp P_{AN} promoter in pJCluc
pJCLpanE	7913	218 bp P_{AN} promoter downstream of the *katG* UAR in pJCluc

[1] Bal31 nuclease derivatives of pJCLK0.6.
[2] Upstream activator region.

A.3. CULTURE CONDITIONS

A.3.1 *E. coli*

E. coli cultures were grown aerobically in 2x YT (yeast tryptone) medium. Agar was added to the medium at a final concentration of 1.5% (w/v). The medium was supplemented with kanamycin (50 μg/ml), ampicillin (50 μg/ml), X-GAL (100 μg/ml), and IPTG (0.3 mM) where necessary (see Appendix C).

2 x Yeast Tryptone (YT)

16 g bacto-tryptone
10 g bacto-yeast extract
5 g sodium chloride
Make up to 1 litre with water and autoclave.

A.3.2 Mycobacteria

Mycobacterial cultures were grown aerobically in Middlebrook 7H9 medium supplemented with catalase-free ADC and Tween 80. Agar was added to the medium at a final concentration of 1.5% (w/v). The medium was supplemented with kanamycin (40 μg/ml), and cycloheximide (0.01% (w/v)) where necessary.

Middlebrook 7H9

4.7 g Middlebrook 7H9 (Difco)
2 g/l asparagine
0.5 g/l casitone
2.5 g/l sodium phosphate
1 g/l potassium phosphate
50 mg/l ferric ammonium citrate
10 mg/l magnesium sulphate
0.5 mg/l calcium chloride
0.1 mg/l zinc sulphate
0.1 mg/ml copper sulphate
Make up to 900 ml with water and autoclave.

ADC Supplement

5 g bovine albumin fraction V

2 g glucose

Make up to 100 ml with water and filter sterilize.

Middlebrook-ADC-Tween

900 ml Middlebrook 7H9

100 ml ADC supplement

0.05% Tween 80

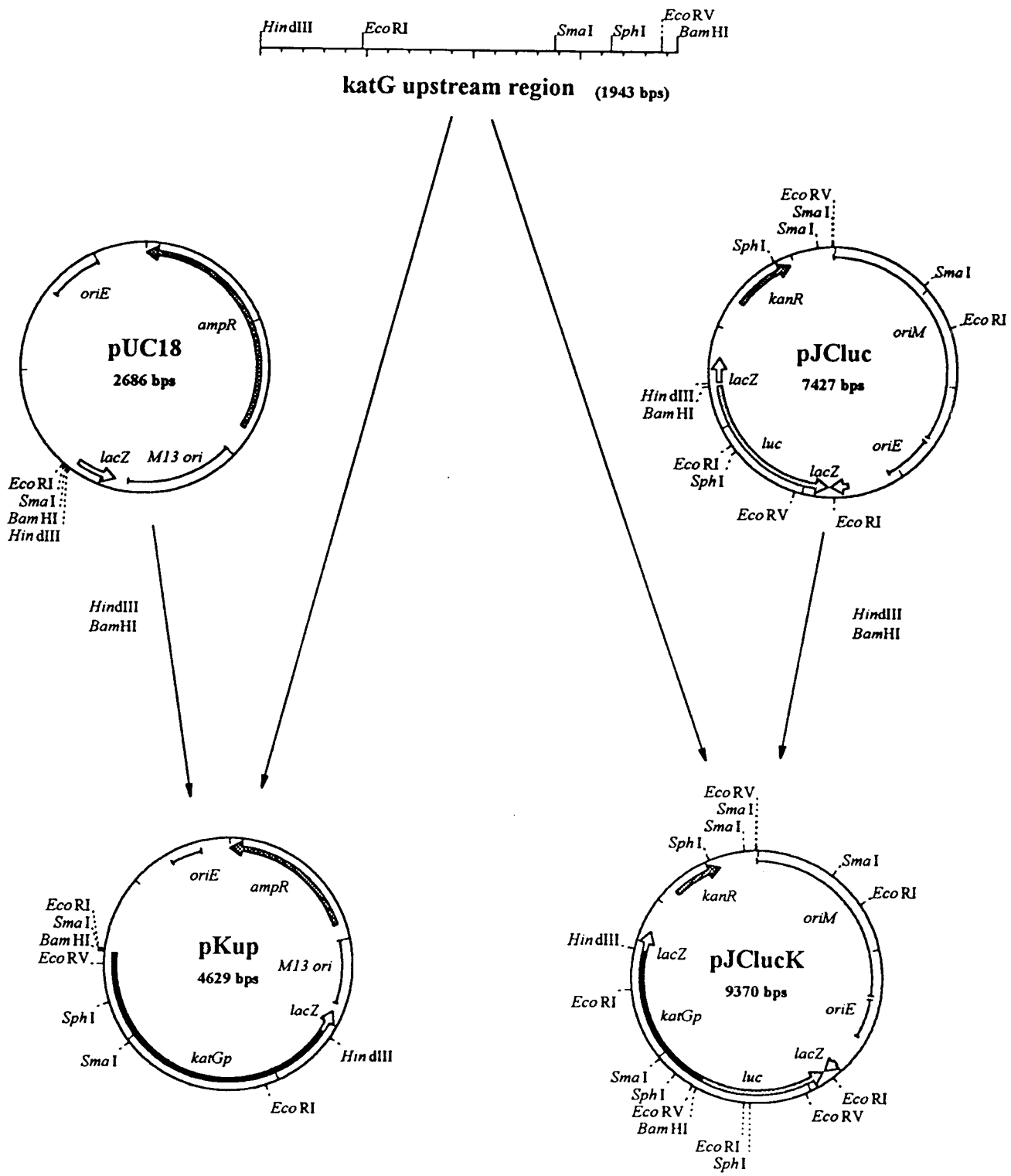


Figure A.1 Schematic representation of the construction of pKup and pJClucK (not to scale).

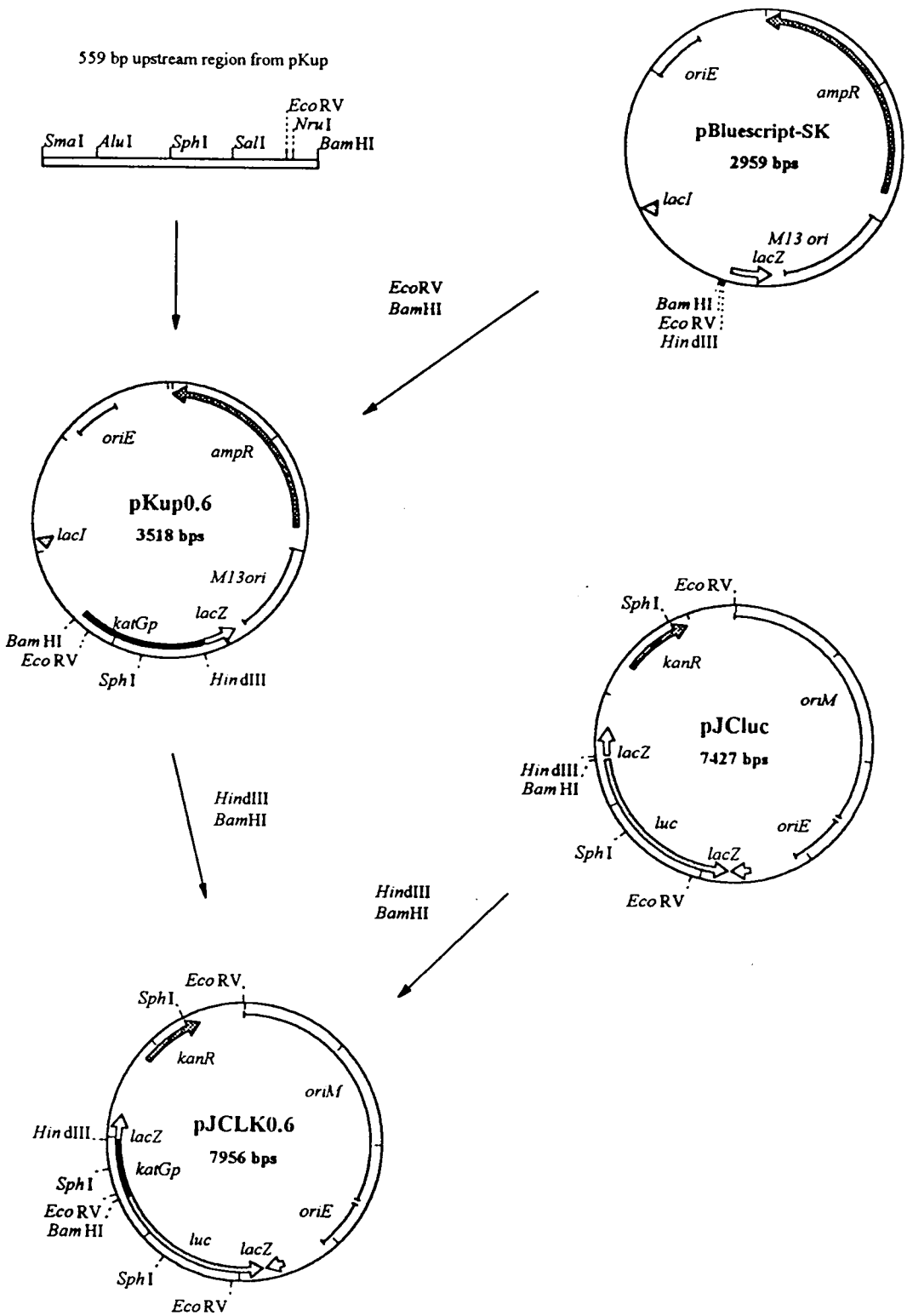
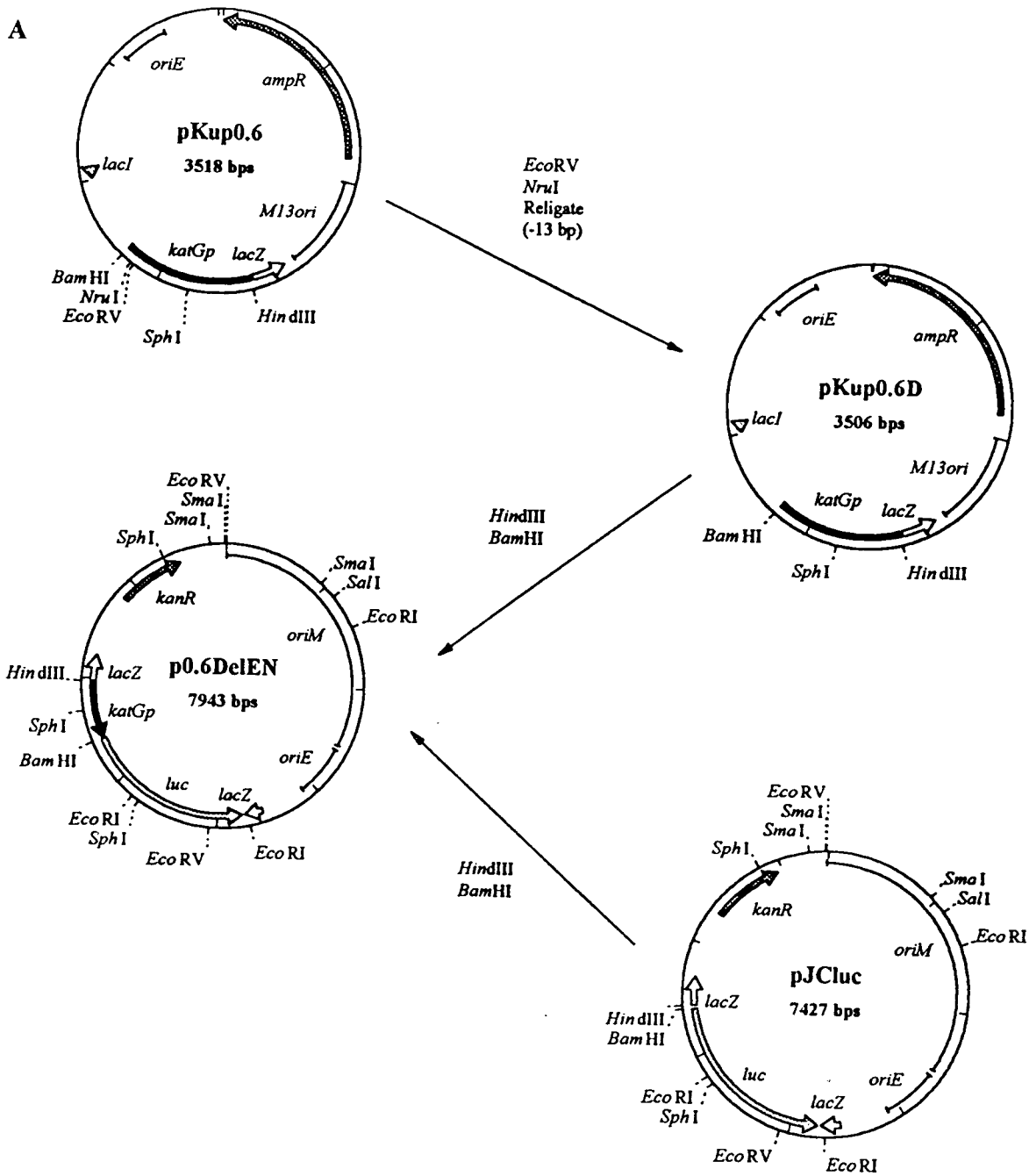


Figure A.2 Schematic representation of the construction of pJCLK0.6 (not to scale).

A

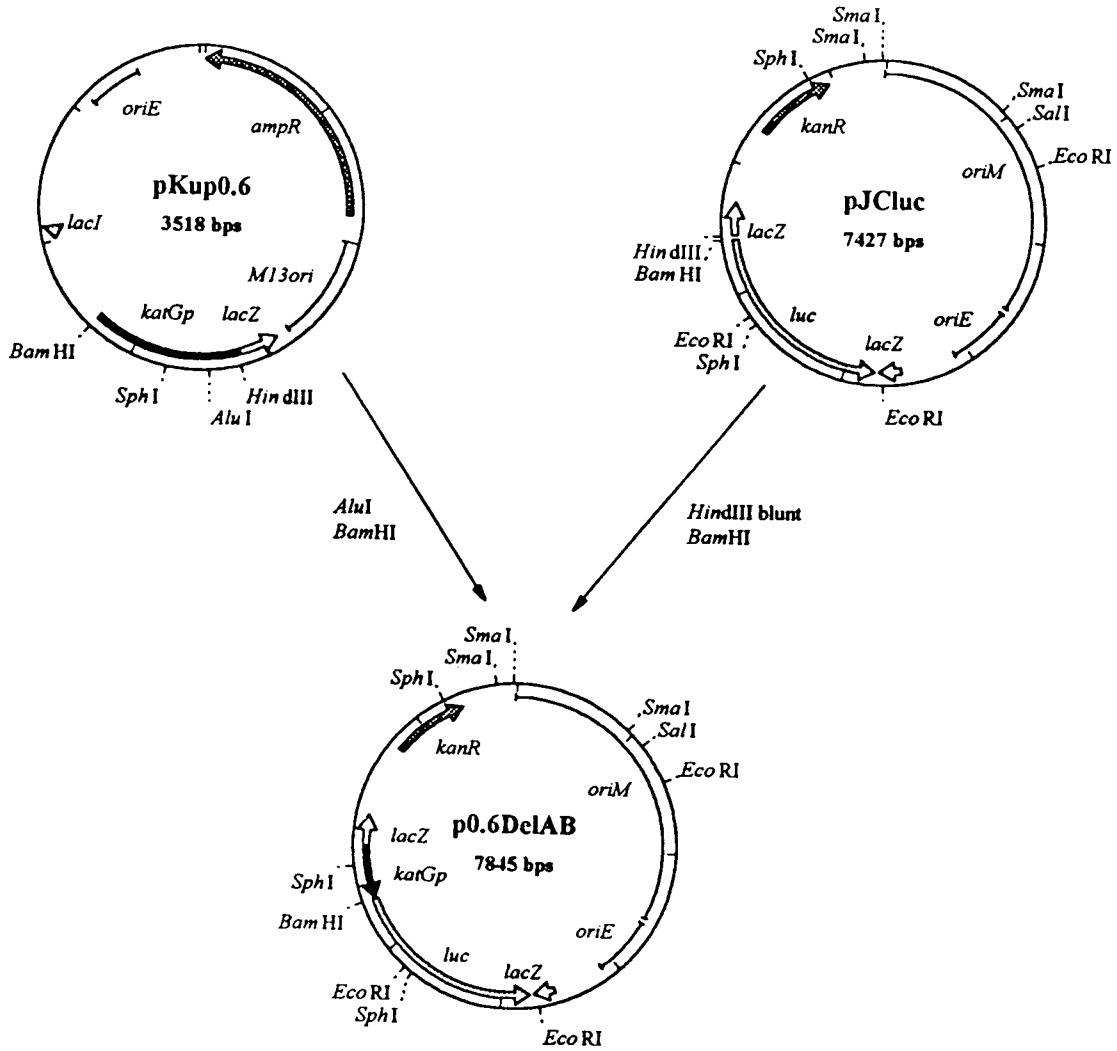


B

```
CCCGGGCGATCGGGTCCTAGCAGACGCCTGTCACGCTAGCCAAAGTCTTGACTGATTCCA  1475
GAAAAGGGAGTCATATTGTCTAGTGTGTCCTCTATACCGGACTACGCCGAACAGCTCCGG  1535
ACGGCCGACCTGCGCGTGACCCGACCGCGCGTCGCCGTCCTGGAAGCAGTGAATGCGCAT  1595
CCACACGCCGACACGGAAACGATTTTCGGTGCCGTGCGTTTTGCGCTGCCCGACGTATCC  1655
GGCAAGCCGTGTACGACGTGCTGCATGCCCTGACCGCCGCGGGCTTGGTGCGAAAGATCC  1715
AACCCTCGGGCTCCGTCGCGCGCTACGAGTCCAGGGTCGGCGACAACCACCATCACATCG  1775
TCTGCCGGTCTTGCGGGGTTATCGCCGATGTCGACTGTGCTGTTGGCGAGGCACCCTGTC  1835
TGACGGCCTCGGACCATAACGGCTTCCTGTTGGACGAGGCGGAGGTCATCTACTGGGGTC  1895
TATGTCCTGATTGTTCGAT[ATCCGACACTTCG]CGATCACATCCGTGATCACAGCCCGATA  1955
ACACCAACTCCTGGAAGGATCC                                         1977
```

Figure A.3 A.) Schematic representation of the construction of p0.6DelEN$_{13}$ (not to scale). B.) Nucleotide sequence of the *Sma*I-*Bam*HI insert of pJCLK0.6, labelled with respect to the published sequence (GenBank Accession Number X68081). The bases which have been deleted in p0.6DelEN$_{13}$ are in boldtype.

A

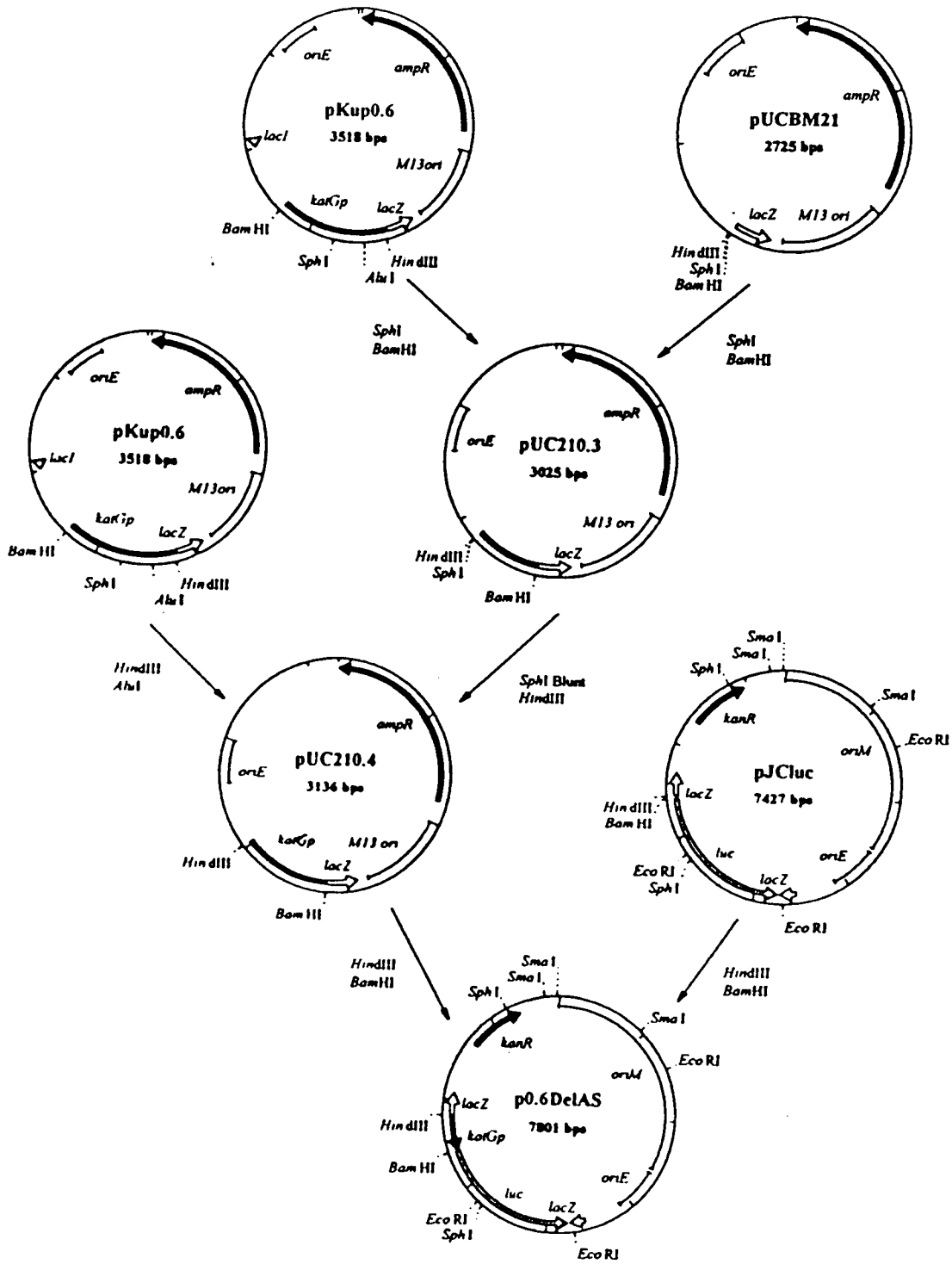


B

```
CCCGGGCGATCGGGTCCTAGCAGACGCCTGTCACGCTAGCCAAAGTCTTGACTGATTCCA  1475
GAAAAGGGAGTCATATTGTCTAGTGTGTCCTCTATACCGGACTACGCCGAACAGCTCCGG  1535
ACGGCCGACCTGCGCGTGACCCGACCGCGCGTCGCCGTCCTGGAAGCAGTGAATGCGCAT  1595
CCACACGCCGACACGGAAACGATTTTCGGTGCCGTGCGTTTTGCGCTGCCCGACGTATCC  1655
GGCAAGCCGTGTACGACGTGCTGCATGCCCTGACCGCCGCGGGCTTGGTGCGAAAGATCC  1715
AACCCTCGGGCTCCGTCGCGCGCTACGAGTCCAGGGTCGGCGACAACCACCATCACATCG  1775
TCTGCCGGTCTTGCGGGGTTATCGCCGATGTCGACTGTGCTGTTGGCGAGGCACCCTGTC  1835
TGACGGCCTCGGACCATAACGGCTTCCTGTTGGACGAGGCGGAGGTCATCTACTGGGGTC  1895
TATGTCCTGATTGTTCGATATCCGACACTTCGCGATCACATCCGTGATCACAGCCCGATA  1955
ACACCAACTCCTGGAAGGATCC                                         1977
```

Figure A.4 A.) Schematic representation of the construction of p0.6DelAB$_{111}$ (not to scale). B.) Nucleotide sequence of the *Sma*I-*Bam*HI insert of pJCLK0.6, labelled with respect to the published sequence (GenBank Accession Number X68081). The bases which have been deleted in p0.6DelAB$_{111}$ are in boldtype.

A

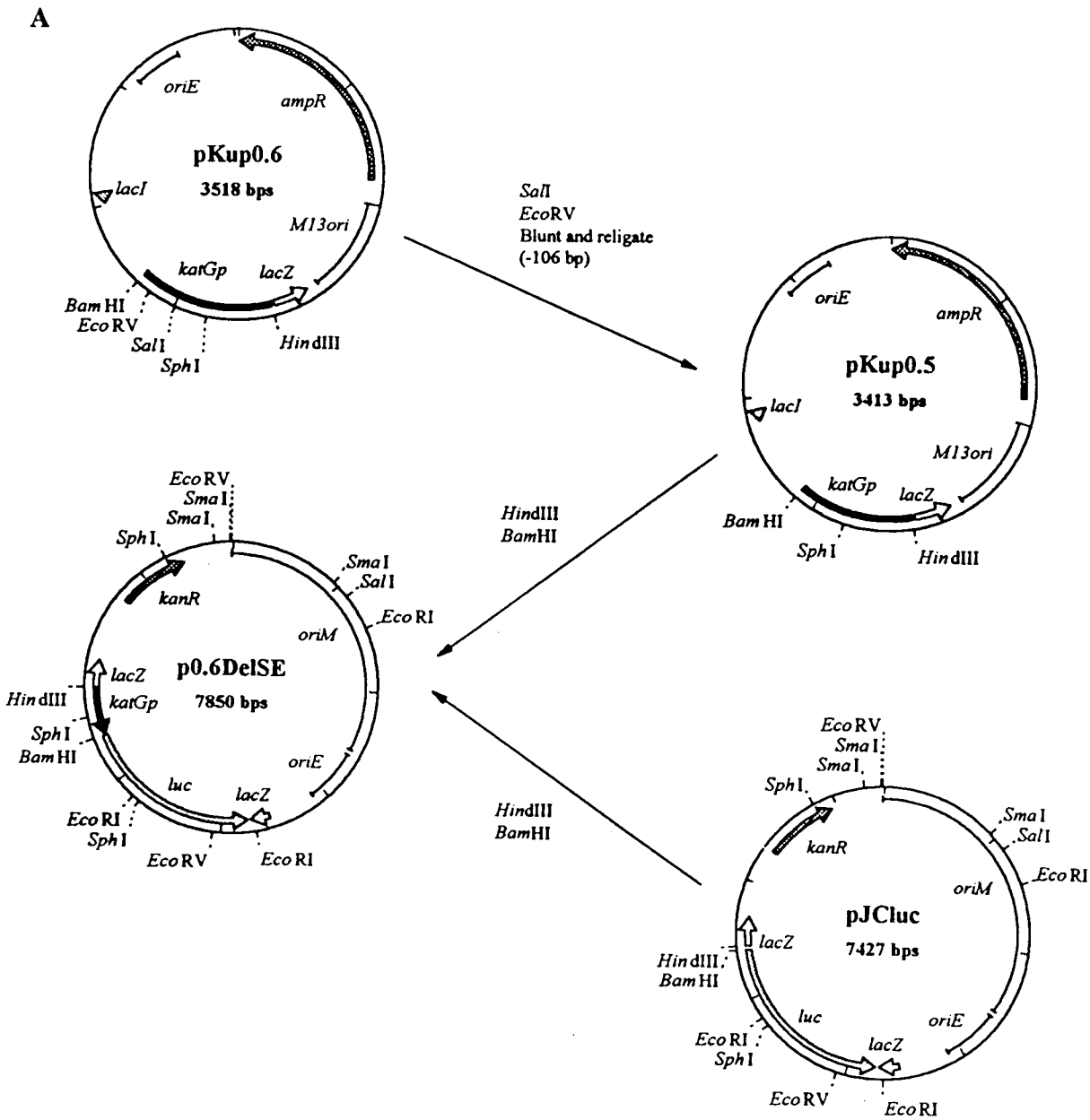


B

```
CCCGGGCGATCGGGTCCTAGCAGACGCCTGTCACGCTAGCCAAAGTCTTGACTGATTCCA  1475
GAAAAGGGAGTCATATTGTCTAGTGTGTCCTCTATACCGGACTACGCCGAACAGCTCCGG  1535
ACGGCCGACCTGCGCGTGACCCGACCGCGCGTCGCCGTCCTGGAAGCAGTGAATGCGCAT  1595
CCACACGCCGACACGGAAACGATTTTCGGTGCCGTGCGTTTTGCGCTGCCCGACGTATCC  1655
GGCAAGCCGTGTACGACGTGCTGCATGCCCTGACCGCCGCGGGCTTGGTGCGAAAGATCC  1715
AACCCTCGGGCTCCGTCGCGCGCTACGAGTCCAGGGTCGGCGACAACCACCATCACATCG  1775
TCTGCCGGTCTTGCGGGGTTATCGCCGATGTCGACTGTGCTGTTGGCGAGGCACCCTGTC  1835
TGACGGCCTCGGACCATAACGGCTTCCTGTTGGACGAGGCGGAGGTCATCTACTGGGGTC  1895
TATGTCCTGATTGTTCGATATCCGACACTTCGCGATCACATCCGTGATCACAGCCCGATA  1955
ACACCAACTCCTGGAAGGATCC                                        1977
```

Figure A.5 A.) Schematic representation of the construction of p0.6DelAS$_{155}$ (not to scale). B.) Nucleotide sequence of the *Sma*I-*Bam*HI insert of pJCLK0.6, labelled with respect to the published sequence (GenBank Accession Number X68081). The bases which have been deleted in p0.6DelAS$_{155}$ are in boldtype.

A

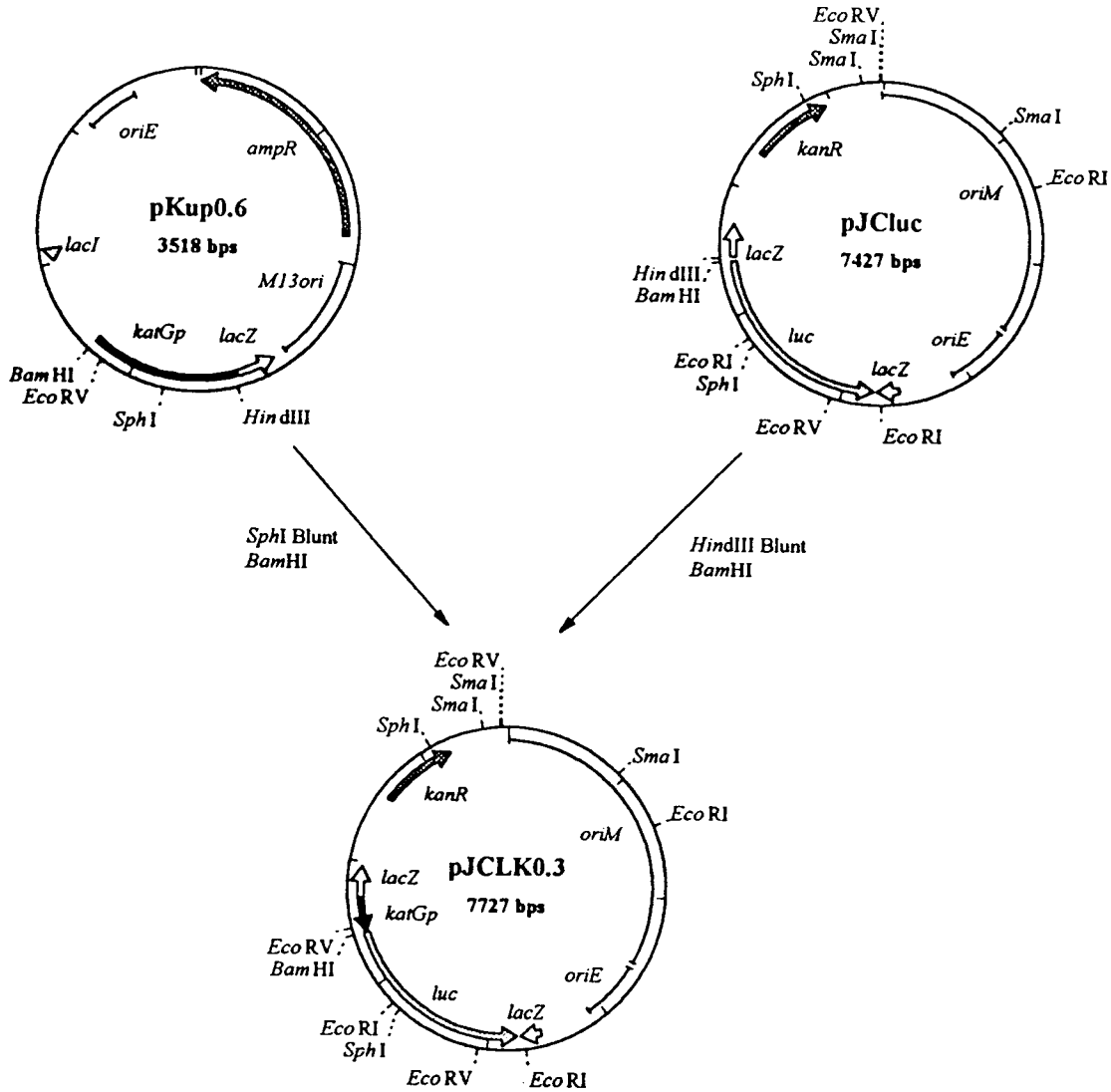


B

```
CCCGGGCGATCGGGTCCTAGCAGACGCCTGTCACGCTAGCCAAAGTCTTGACTGATTCCA  1475
GAAAAGGGAGTCATATTGTCTAGTGTGTCCTCTATACCGGACTACGCCGAACAGCTCCGG  1535
ACGGCCGACCTGCGCGTGACCCGACCGCGCGTCGCCGTCCTGGAAGCAGTGAATGCGCAT  1595
CCACACGCCGACACGGAAACGATTTTCGGTGCCGTGCGTTTTGCGCTGCCCGACGTATCC  1655
GGCAAGCCGTGTACGACGTGCTGCATGCCCTGACCGCCGCGGGCTTGGTGCGAAAGATCC  1715
AACCCTCGGGCTCCGTCGCGCGCTACGAGTCCAGGGTCGGCGACAACCACCATCACATCG  1775
TCTGCCGGTCTTGCGGGGTTATCGCCGATGTCGACTGTGCTGTTGGCGAGGCACCCTGTC  1835
TGACGGCCTCGGACCATAACGGCTTCCTGTTGGACGAGGCGGAGGTCATCTACTGGGGTC  1895
TATGTCCTGATTGTTCGATATCCGACACTTCGCGATCACATCCGTGATCACAGCCCGATA  1955
ACACCAACTCCTGGAAGGATCC                                         1977
```

Figure A.6 A.) Schematic representation of the construction of p0.6DelSE$_{106}$ (not to scale). B.) Nucleotide sequence of the *Sma*I-*Bam*HI insert of pJCLK0.6, labelled with respect to the published sequence (GenBank Accession Number X68081). The bases which have been deleted in p0.6DelSE$_{106}$ are in boldtype.

A

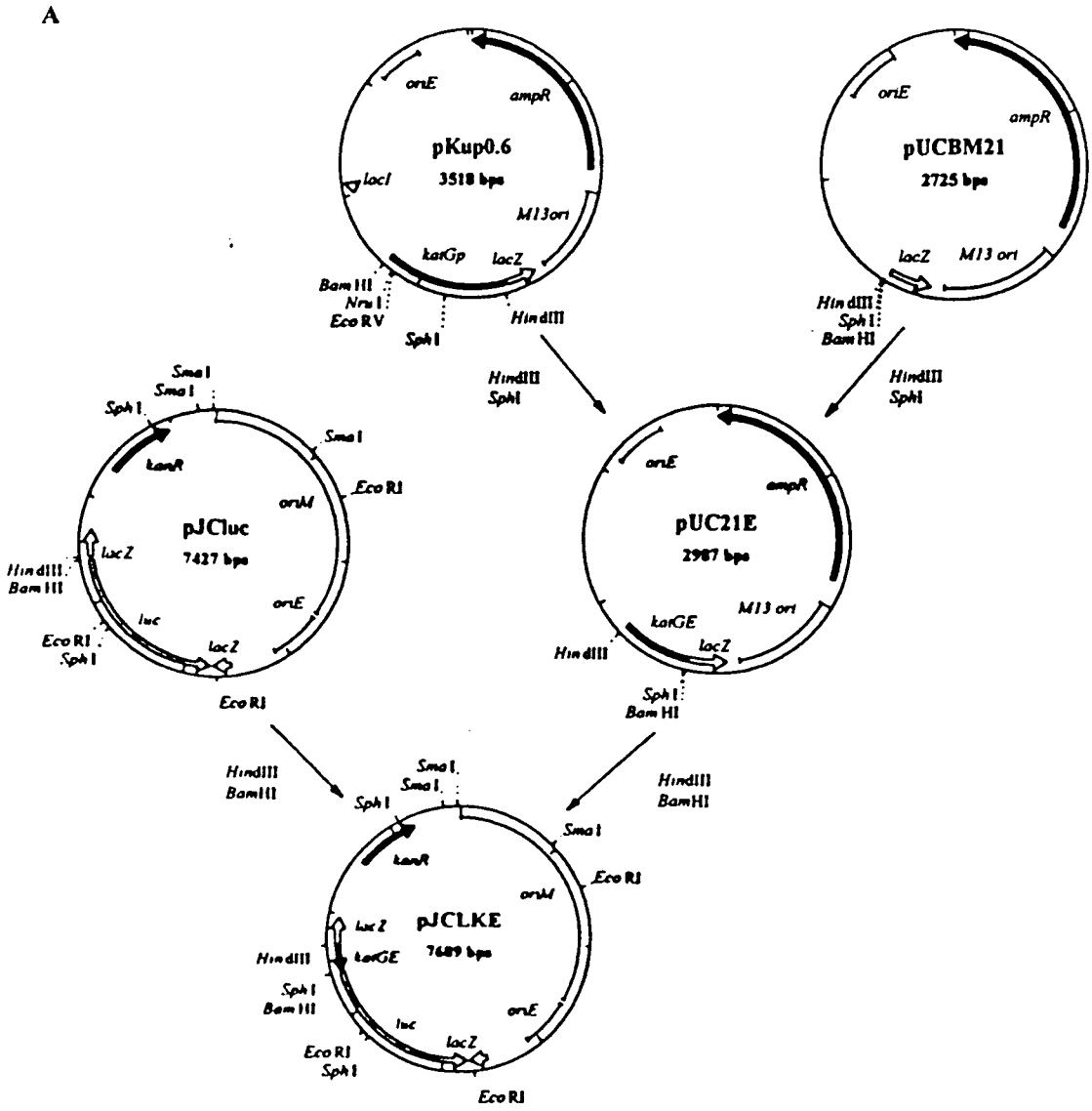


B

```
CCCGGGCGATCGGGTCCTAGCAGACGCCTGTCACGCTAGCCAAAGTCTTGACTGATTCCA  1475
GAAAAGGGAGTCATATTGTCTAGTGTGTCCTCTATACCGGACTACGCCGAACAGCTCCGG  1535
ACGGCCGACCTGCGCGTGACCCGACCGCGCGTCGCCGTCCTGGAAGCAGTGAATGCGCAT  1595
CCACACGCCGACACGGAAACGATTTTCGGTGCCGTGCGTTTTGCGCTGCCCGACGTATCC  1655
GGCAAGCCGTGTACGACGTGCTGCATGCCCTGACCGCCGCGGGCTTGGTGCGAAAGATCC  1715
AACCCTCGGGCTCCGTCGCGCGCTACGAGTCCAGGGTCGGCGACAACCACCATCACATCG  1775
TCTGCCGGTCTTGCGGGGTTATCGCCGATGTCGACTGTGCTGTTGGCGAGGCACCCTGTC  1835
TGACGGCCTCGGACCATAACGGCTTCCTGTTGGACGAGGCGGAGGTCATCTACTGGGGTC  1895
TATGTCCTGATTGTTCGATATCCGACACTTCGCGATCACATCCGTGATCACAGCCCGATA  1955
ACACCAACTCCTGGAAGGATCC                                         1977
```

Figure A.7 A.) Schematic representation of the construction of pJCLK0.3 (not to scale). B.) Nucleotide sequence of the *Sma*I-*Bam*HI insert of pJCLK0.6, labelled with respect to the published sequence (GenBank Accession Number X68081). The bases which have been deleted in pJCLK0.3 are in boldtype.

A



B

1

```
CCCGGGCGATCGGGTCCTAGCAGACGCCTGTCACGCTAGCCAAAGTCTTGACTGATTCCA  1475
GAAAAGGGAGTCATATTGTCTAGTGTGTCCTCTATACCGGACTACGCCGAACAGCTCCGG  1535
ACGGCCGACCTGCGCGTGACCCGACCGCGCGTCGCCGTCCTGGAAGCAGTGAATGCGCAT  1595
CCACACGCCGACACGGAAACGATTTTCGGTGCCGTGCGTTTTGCGCTGCCCGACGTATCC  1655
GGCAAGCCGTGTACGACGTGCTGCATGCCCTGACCGCCGCGGGCTTGGTGCGAAAGATCC  1715
AACCCTCGGGCTCCGTCGCGCGCTACGAGTCCAGGGTCGGCGACAACCACCATCACATCG  1775
TCTGCCGGTCTTGCGGGGTTATCGCCGATGTCGACTGTGCTGTTGGCGAGGCACCCTGTC  1835
TGACGGCCTCGGACCATAACGGCTTCCTGTTGGACGAGGCGGAGGTCATCTACTGGGGTC  1895
TATGTCCTGATTGTTCGATATCCGACACTTCGCGATCACATCCGTGATCACAGCCCGATA  1955
ACACCAACTCCTGGAAGGATCC                                         1977
```

2

```
CCCGGGCGATCGGGTCCTAGCAGACGCCTGTCACGCTAGCCAAAGTCTTGACTGATTCCA  1475
GAAAAGGGAGTCATATTGTCTAGTGTGTCCTCTATACCGGACTACGCCGAACAGCTCCGG  1535
ACGGCCGACCTGCGCGTGACCCGACCGCGCGTCGCCGTCCTGGAAGCAGTGAATGCGCAT  1595
CCACACGCCGACACGGAAACGATTTTCGGTGCCGTGCGTTTTGCGCTGCCCGACGTATCC  1655
GGCAAGCCGTGTACGACGTGCTGCATGCCCTGCAGAGCTCTAGAGTCGACGGGCCCGGTAC  1715
CGCGGCCGCACTTAAGTTACGCGTGGATCC                                 1745
```

Figure A.8 A.) Schematic representation of the construction of pJCLKE (not to scale). B.1.) Nucleotide sequence of the *Sma*I-*Bam*HI insert of pJCLK0.6, labelled with respect to the published sequence (GenBank Accession Number X68081). The bases which have been deleted in pJCLKE are in boldtype. 2.) Nucleotide sequence of the insert of pJCLKE. The bases from pUC21E which have been added to the insert are underlined.

APPENDIX B

GENERAL TECHNIQUES

The enzymes and kits used in this study were supplied by Boehringer Mannheim (Mannheim Germany), Promega (Madison, U.S.A.), Amersham (UK), and Pharmacia (Uppsala, Sweden). Chemicals were supplied by Melford Laboratories (England), Bio-Rad Laboratories (CA, U.S.A.), BDH Chemicals Ltd (Poole, England), Sigma Chemicals (St. Louis, U.S.A.), Fluka Biochemicals, and Merck (Darmstadt, Germany).

B.1. CLONING

B.1.1. Restriction endonuclease digestion

Restriction endonuclease digestions were carried out according to the method of Sambrook *et al.* (1989), in a 20 µl volume containing 0.5 - 10 µg of DNA in a 1x solution of the supplied restriction endonuclease incubation buffer, and 10 units of restriction endonuclease. The digests were incubated at 37°C for between 1 and 18 hours.

B.1.2. Blunt-ending with T4 DNA polymerase and Mung Bean nuclease

Cohesive ends resulting from restriction endonuclease digestion were blunt-ended using the Amersham DNA Blunting kit (Amersham Life Science, UK), according to the manufacturers' instructions. The T4 DNA polymerase from the kit fills in 5'-protruding ends with its 5'-3' polymerase activity and removes 3'-protruding ends with its 3'-5' exonuclease activity. Alternatively, cohesive ends were removed by digestion with Mung Bean nuclease (Boehringer Mannheim). The DNA was digested in a 50 µl reaction containing 5 µl of 10x Mung Bean digestion buffer (0.5 M sodium acetate pH 5, 0.3 M sodium chloride, 0.01 M zinc sulphate), and 10 units of Mung Bean nuclease. The reaction mix was incubated at 30°C for 30 minutes, and the DNA was recovered by phenol/chloroform/isoamyl alcohol (25:24:1) extraction and precipitation with ethanol (see B.1.4).

B.1.3. Agarose gel electrophoresis

For agarose gel electrophoresis, a horizontal, submerged gel system was used, as described by Sambrook *et al.* (1989), with a tris-acetate buffer (see Appendix C). Agarose concentrations varied from 0.8% (w/v) to 2% (w/v). Gel tracking dye (see Appendix C) was added to the DNA sample prior to electrophoresis. The DNA was stained with ethidium bromide (0.2 μg/ml), visualized using a 310 nm transilluminator, and photographed using a Mamiya RB67 Pro-S camera.

B.1.4. Phenol squeeze extraction of DNA from agarose gels

DNA was purified from agarose gels using a phenol extraction procedure (Seth, 1984). The gel band was finely chopped using a Feather surgical blade, an equal volume of phenol was added, and the sample was immediately frozen on dry ice for 10 minutes. The sample was centrifuged at 14 000 rpm for 10 minutes and the aqueous phase was recovered. The DNA was extracted twice more by adding an equal volume of phenol/chloroform/isoamyl alcohol (25:24:1), vortexing and centrifuging for 5 minutes. The aqueous phase was recovered and 1/10th of a volume of 3 M sodium acetate (pH 5.2) and 2.5 volumes of ice-cold absolute ethanol were added to precipitate the DNA. After incubation at -70°C for 15 minutes, the DNA was collected by centrifugation at 14 000 rpm for 15 minutes. The DNA was washed with 70% (v/v) ethanol, dried, and resuspended in a minimum volume of sterile distilled water.

B.1.5. Ligation reactions

Insert and vector DNA were recovered from restriction endonuclease digestion or blunt-end reactions by extraction with phenol/chloroform/isoamyl alcohol (25:24:1) and precipitation with ethanol. The DNA was ligated according to the method of Sambrook *et al.* (1989) in a 15 μl reaction containing 10x T4 DNA ligase buffer and 10 units of T4 DNA ligase (Boehringer Mannheim). Ligations were carried out at 16°C for 3-4 hours or overnight. In each ligation experiment, reactions were included which had no DNA or the vector self-ligated, as controls.

B.1.6. Preparation and transformation of competent *E. coli* cells

E. coli cells were made competent for the uptake of DNA according to the method of Chung and Miller (1988). Cells were grown in 5 ml of 2x YT broth to stationary phase. The culture was diluted 1/200 into 100 ml of 2x YT broth and grown to early exponential phase (OD_{600} = 0.5) at 37°C. The cells were collected by centrifugation at 4000 rpm for 5 minutes at 4°C, resuspended in 1/10th of the culture volume in ice cold TSB (2x YT broth containing 10% (w/v) PEG 4000, 5% (v/v) DMSO, 10 mM magnesium chloride, and 10 mM magnesium sulphate), and held on ice for 20 minutes before transformation. The competent *E. coli* cells were transformed by the addition of 2-5 µl of ligated DNA to 50 µl of cells. The cells were held on ice for 20 minutes and heat shocked at 42°C for 2 minutes. TSB (450 µl) containing 20 mM glucose was added to the transformation mix, and the cells were incubated at 37°C for 1 hour to facilitate expression of the antibiotic resistance gene. A 200 µl aliquot of the cells was plated onto 2x YT agar plates with the appropriate antibiotic selection. Competent cells were frozen in dry-ice and ethanol and stored at -70°C.

B.1.7. Plasmid preparation from *E. coli*

Plasmids were isolated from *E. coli* using the alkaline lysis method of Ish-Horowicz and Burke (1981). A 5 ml overnight culture was harvested by centrifugation and the pellet was resuspended in 200 µl of solution 1 (10x stock: 0.25 M tris-Cl pH 8, 0.5 M glucose, 0.1 M EDTA pH 8). The sample was kept at room temperature for 5 minutes, and 400 µl of freshly prepared solution 2 (0.2 M sodium hydroxide, 1% (w/v) SDS) was added. The sample was placed on ice for 5 minutes. This was followed by the addition of 300 µl of solution 3 (3 M potassium acetate, 2 M acetic acid, pH 4.8), and incubation on ice for a further 5 minutes. The precipitated proteins were removed by centrifugation at 14 000 rpm for 10 minutes and the DNA in the supernatant was precipitated by the addition of 0.8 volumes of isopropanol and recovered by centrifugation. The pellet was resuspended in water and reprecipitated with 1/10 of a volume of 3 M sodium acetate (pH 4.5), and 2.5 volumes of ice-cold 100% (v/v) ethanol. The pellet was washed with 70% (v/v) ethanol and resuspended in water. Large-scale DNA preparations were performed using the Nucleobond AX PC-kit 100 (Macherey-Nagel), according to the manufacturers' instructions.

B.2. DNA SEQUENCING

B.2.1. Denaturation of template DNA

Template DNA (5-10 µg in 32 µl of sterile distilled water) was denatured by the addition of 8 µl of 2 M sodium hydroxide and incubation for 10 minutes at room temperature. The denatured DNA was precipitated by the addition of 7 µl of 3 M sodium acetate (pH 4.8), 4 µl of distilled water and 120 µl of 100% (v/v) ethanol, and incubation on dry ice for 15 minutes. The DNA was recovered by centrifugation at 14 000 rpm for 15 minutes, washed with 70% (v/v) ethanol, and resuspended in 10 µl of sterile distilled water.

B.2.2. Annealing reaction

The denatured DNA was annealed to 2 µl of sequencing primer (10 pmoles) with 2µl of annealing buffer (T7 Sequencing kit, Pharmacia), by incubation at 65°C for 5 minutes, followed by incubation at 37°C for 10 minutes, and room temperature for 10 minutes. Sequencing reactions were performed using the dideoxynucleotide triphosphate chain termination method of Sanger *et al.* (1977), [α-^{35}S]dATP (1000-1500 Ci/mmol, Amersham Life Science, UK), and the T7 Sequencing kit (Pharmacia), according to the manufacturers' instructions.

B.2.3. Sequenase™ Version 2.1 protocol for GC-rich templates

The Sequenase™ Version 2.1 DNA Sequencing kit (United States Biochemical, Amersham Life Science, UK) was used for the resolution of band compressions. Templates for sequencing were denatured as described above, except that the DNA was resuspended in a final volume of 7 µl of water. The sequencing primer (5 pmoles) was added to the denatured template, together with 2 µl of 5x Sequenase Reaction Buffer, and the reaction mixture was heated to 65°C for 2 minutes, cooled slowly to 30°C, then placed on ice. Sequencing reactions were performed using the dITP and dGTP reagents of the Sequenase™ Version 2.1 kit, and the manufacturers' protocol modified for GC-rich templates. The dGTP and dITP reactions were loaded adjacent to each other on polyacrylamide gels.

B.2.4. Preparation of the acrylamide gel

The glass plates, slides and comb were washed thoroughly with 70% (v/v) ethanol and acetone, and clamped together using filter paper as the bottom spacer. The 6% (w/v) acrylamide solution was prepared by dissolving 48 g of urea in 10 ml of 10x NNB (see Appendix C), 15 ml of 40% (w/v) acrylamide/2% (w/v) bis-acrylamide, and 39 ml of distilled water. For polymerization, 350 µl of freshly prepared 13% (w/v) ammonium persulphate and 80 µl of TEMED were added and the gel was poured. The acrylamide was left to polymerize overnight.

B.2.5. Acrylamide gel electrophoresis

The acrylamide gel was "prerun" in a vertical apparatus at 20 watts (W) in 1 x NNB electrophoresis buffer for 20 minutes. The sequencing reactions were heated to 80°C for 2 minutes to denature the DNA, and 4 µl of each sample was loaded onto the gel in the order G, A, T, and C. The gel was run at 40 W. After electrophoresis, the gel was transferred onto blotting paper and dried under a vacuum on a Hoefer slab gel dryer Model SE1160 for one hour. The dried gel was then exposed to an Agfa Curix X-Ray film for autoradiography.

B.2.6. Computer analysis of DNA sequences

The following computer software was used for the analysis of nucleic acid and amino acid sequences:

GENEPRO™ Version 6.10: Riverside Scientific Enterprises; Bainbridge Enterprises, WA U.S.A. Program for the analysis of DNA sequences, open reading frames, restriction endonuclease sites, and for protein sequence analysis.

CLUSTAL W: Multiple sequence alignment program (Higgins *et al.*, 1996).

BLAST: Basic Local Alignment Research Tool at the National Centre for Biotechnology Information, National Institute of Health. Tool for searching the nucleotide and protein sequence database for homologous sequences (Altschul *et al.*, 1990).

B.3. POLYMERASE CHAIN REACTION

Taq polymerase, derived from *Thermus aquaticus* (Saiki *et al.*, 1988) (Boehringer Mannheim), was used in all the polymerase chain reactions (PCRs). The reactions were carried out in a 0.5 ml microfuge tube containing 5 µl of 10x Taq polymerase buffer, 0.4 µl of Taq DNA polymerase (2-3 units), 1 µl each of dATP, dGTP, dCTP and dTTP from 10 mM stock solutions, 10% (v/v) DMSO, 50-100 pmoles of primer, and 100-200 ng of template DNA. The volume was made up to 50 µl with sterile distilled water, and an overlay of 50 µl of paraffin was added to prevent evaporation. The PCR was carried out in a JDI High Performance Temperature Profiler, Model 8012. The template DNA was denatured by heating to 95°C for 5 minutes, and the temperature profile of the amplification was as follows: 30s at 94°C, 60s at the annealing temperature, and 90s at 72°C for 36 cycles, followed by further incubation at 72°C for 5 min.

B.4. ISOLATION OF CHROMOSOMAL DNA FROM MYCOBACTERIAL CELLS

A 100 ml culture of mycobacterial cells was harvested by centrifugation at 3000 rpm for 10 minutes at 4°C, and the pellet was resuspended in 500 µl of TES (10 mM tris, 1 mM EDTA, 0.15 M sodium chloride) (Jacobs *et al.*, 1991). The cells were disrupted by 2 cycles in a Biospec Mini-Beadbeater at 4200 rpm for 30 seconds each. The cell debris was removed by centrifugation. The supernatant was extracted twice with phenol/chloroform/isoamyl alcohol (25:24:1) and the top aqueous phase was recovered. The DNA was precipitated with 1/10 of a volume of 3 M sodium acetate and 2.5 volumes of ethanol, and recovered by centrifugation. The pellet was washed with 70% (v/v) ethanol and resuspended in sterile distilled water. The RNA was removed by digestion with RNase A, and the DNA was extracted twice more with phenol/chlorofom/isoamyl alcohol (25:24:1). The aqueous phase was collected, and the DNA was precipitated as described above. The DNA pellet was resuspended in 150-200 µl of sterile distilled water. The quality of the chromosomal DNA was checked by electrophoresis on a 0.8% (w/v) agarose gel, and the concentration of the DNA was determined by measuring the absorbance at 260 nm on a Beckman DU-40 Spectrophotometer. The concentration was calculated using the relationship $OD_{260} = 1$ for 50 µg of dsDNA.

B.5. PLASMID ISOLATION FROM MYCOBACTERIA

Plasmids were isolated from mycobacteria using the method of Jacobs *et al.* (1991). A 5 ml culture was grown to late exponential phase and the cells were harvested by centrifugation. The pellet was resuspended in 150 µl of GTE (25 mM tris-Cl pH 8, 10 mM EDTA, 50 mM glucose), containing 10 mg/ml of lysozyme, and incubated at 37°C overnight. Two volumes of a freshly prepared 0.2 M sodium hydroxide, 1% (w/v) SDS solution was added, and the suspension was incubated on ice for 10 minutes. After incubation, 225 µl of 3 M potassium acetate was added and the sample was incubated for a further 10 minutes on ice. The proteins and cell debris were removed by centrifugation at 14 000 rpm for 30 minutes, and the DNA in the supernatant was extracted with phenol/chloroform/isoamyl alcohol (25:24:1). The DNA was precipitated with 2 volumes of ethanol, and the pellet was washed with 70% (v/v) ethanol and resuspended in 25-50 µl of TE.

B.6. BIO-RAD ASSAY

Protein concentration was determined using a Bio-Rad *DC* Protein Assay kit. The protein standards, 0.2 mg/ml, 0.5 mg/ml, 0.75 mg/ml and 1.0 mg/ml, were prepared by dilution of the Bio-Rad bovine serum albumin protein standard (1.5 mg/ml) in the appropriate buffer. For each assay, 100 µl of solution A' (20 µl of solution S/ml of solution A) was mixed with 20 µl of protein sample, and 800 µl of solution B. The samples were incubated at room temperature for 15-20 minutes, and the OD at 750 nm determined. The OD_{750} values of the Bio-Rad protein standards were plotted against concentration to give a linear graph. The concentration of protein in the samples was determined from the graph.

B.7. SDS-POLYACRYLAMIDE GEL ELECTROPHORESIS

SDS-polyacrylamide gels were poured and electrophoresed in a Hoeffer gel pouring stand and a vertical electrophoresis apparatus. The gel preparation and electrophoresis were performed according to the method of Laemmli (1970). A resolving gel of acrylamide in 1.5 M tris (pH 8.8), 0.4% (w/v) SDS buffer was poured first and allowed to polymerize for at least 1 hour.

The stacking gel in a 0.5 M tris (pH 6.8), 0.4% (w/v) SDS buffer was then poured and allowed to polymerize overnight. Gels were electrophoresed at a constant current of 40 mA in a tris-glycine based buffer (10x solution: 15.15 g tris, 72.05 g glycine, 1% (w/v) SDS in 1 litre). Protein samples were mixed with a sample buffer (see Appendix C) and boiled for 5-10 minutes before loading. Proteins in the Pharmacia Low Molecular Weight Calibration kit were reconstituted in SDS-PAGE sample buffer, and used as molecular weight standards. Non-denaturing polyacrylamide gel electrophoresis was performed as described for SDS-PAGE, except that SDS was omitted from all buffers. SDS-polyacrylamide gels were stained with Coomassie® Brilliant Blue R-250 (Bio-Rad Laboratories) (1.25g Coomassie® Brilliant Blue R-250 in 50% (v/v) methanol, 10% (v/v) acetic acid) for 4-8 hours, and destained with 30% (v/v) methanol, 10% (v/v) acetic acid, and 3% glycerol.

B.8. IMMUNOBLOTTING

Immunoblotting was performed using a method adapted from Sambrook *et al.* (1989). The proteins of interest were electrophoresed in an SDS-polyacrylamide gel, as described above. The proteins were transferred to an Amersham Hybond-C Super nitrocellulose membrane in a transfer buffer (22.4 g glycine, 4.8 g tris, and 400 ml methanol in 2 litres of water), using an Omeg Scientific semi-dry blotter, at a constant current of 200 mA. After transfer, the membrane was washed twice in TBS (0.05 M tris-Cl pH 7.5, 0.15 M sodium chloride) and incubated in 1% (w/v) blocking solution (Chemiluminescence Blotting Substrate-POD kit, Boehringer Mannheim) for 1 hour. This was followed by incubation in 0.5% (w/v) blocking solution containing the primary antibody for 1 hour. The unbound antibody was removed by washing the membrane twice in TBST (TBS + 0.1% (v/v) Tween 20) for 10 minutes each. The membrane was soaked in 0.5% blocking solution for 2 x 10 minutes, then incubated with POD-labelled anti-Rabbit-IgG (Boehringer Mannheim) diluted 1 in 10 000 in 0.5% (w/v) blocking solution for 30 minutes. The secondary antibody was removed by washing the membrane four times in TBST for 15 minutes each. The kit detection solution was prepared by mixing detection reagents A and B in a 100:1 ratio and allowing them to stand at room temperature for 30 minutes before use. The bound antibodies on the membrane were detected by incubating in detection reagent for 1 minute and exposure to X-Ray film for 30-60 seconds.

APPENDIX C

BUFFERS AND SOLUTIONS

Chemicals used in this study were supplied by Melford Laboratories (England), Bio-Rad Laboratories (CA, U.S.A.), BDH Chemicals Ltd (Poole, England), Sigma Chemicals (St. Louis, U.S.A.), Fluka Biochemicals, and Merck (Darmstadt, Germany). The compositions of the commonly used solutions are given below.

Ampicillin

100 mg/ml stock solution:
2 g in 20 ml of water
Filter sterilize and store at -20°C.

Cell Culture Lysis Reagent (Promega)

25 mM tris-phosphate (pH 7.8)
2 mM DTT
2 mM 1,2-diaminocyclohexane-N,N,N',N'-tetraacetic acid
10% (v/v) glycerol
1% (v/v) Triton X-100

Column Buffer

20 mM tris-Cl
0.2 M sodium chloride
1 mM EDTA

Ethidium Bromide

10 mg/ml stock solution:
0.2 g ethidium bromide in 20 ml of water
Store in the dark.

Gel Tracking Dye

0.25 g bromophenol blue
40.0 g sucrose
4.0 ml 0.5 M EDTA (pH 8.0)
Make up to 100 ml with water.

High ionic strength running buffer

5x stock:
30.28 g tris
142.7 g glycine
3.92 g EDTA
Make up to 1 litre with water.

IPTG

0.056 g IPTG in 2 ml water
Filter sterilize and store at -20°C.

Luciferase/Luciferin L/L reagent (Promega)

purified luciferase
d-luciferin
BSA
DTT
tris-acetate (pH 7.75)
EDTA
magnesium acetate

L/L Buffer

0.1 M tris-acetate (pH 7.75)
2 mM EDTA

Luciferase Assay Reagent (Promega)

20 mM tricine
33.3 mM DTT
270 µM coenzyme A
470 µM luciferin
530 µM ATP
0.1 mM EDTA
1.07 mM magnesium carbonate
2.67 mM magnesium sulphate

MOPS running buffer

10x solution:
41.8 g MOPS
800 ml DEPC-water
Adjust pH to 7.0 with sodium hydroxide.
16.6 ml 3 M DEPC-sodium acetate
20 ml 0.5 M DEPC-EDTA (pH 8.0)
Make up to 1 l with DEPC-water.
For details of MOPS-formaldehyde electrophoresis see Ausubel *et al.* (1987).

NNB

10 x stock solution:
162 g tris base
27.5 g boric acid
9.3 g EDTA
Make up to 1 litre with water and autoclave.

Non-denaturing polyacrylamide gel mix

8 ml 5x tris-glycine stock (see high ionic strength running buffer)
6.67 ml 30% (w/v) acrylamide
1 ml 2% (w/v) bis-acrylamide
2 ml 50% glycerol
22.37 ml water

Phenol

Melt 500 g phenol
0.6 g 8-hydroxyquinoline
7.5 ml 2 M sodium hydroxide
Dissolve in 130 ml of water and 6 ml of 1 M tris-Cl (pH 7.6).

Potassium Phosphate Buffer

0.1 M solution (pH 7.0):
61.5 ml 1 M *di*-potassium hydrogen phosphate
38.5 ml 1M potassium dihydrogen phosphate

SDS-PAGE Sample Buffer

3% (w/v) SDS
5% (v/v) β-mercaptoethanol
10% (v/v) glycerol
0.625 M tris-Cl (pH 6.8)
0.01% (w/v) bromophenol blue

TAE

50x solution:
242 g tris base
57.1 ml glacial acetic acid
100 ml 0.5 M EDTA (pH 8.0)
Make up to 1 litre with water and autoclave.
Dilute 1/50 for use and adjust the pH to 7.4 with acetic acid.

X-GAL

0.04 g of X-gal in 2 ml of dimethyl formamide
Store at -20°C.

LITERATURE CITED

Abrams, J. J., and Webster, D. A. 1990. Purification, partial characterization, and possible role of catalase in the bacterium *Vitreoscilla. Arch. Biochem. Biophys.* **279**: 54-59.

Abratt, V. R., Jones, D. T., and Woods, D. R. 1985. Isolation and physiological characterization of mitomycin C-sensitive/UV-sensitive mutants in *Bacteroides fragilis. J. Gen. Microbiol.* **131**: 2479-2483.

Alam, J., and Cook, J. L. 1990. Reporter genes: application to the study of mammalian gene transcription. *Anal. Biochem.* **188**: 245-254.

Aldovini, A., Husson, R. N., and Young, R. A. 1993. The *uraA* locus and homologous recombination in *Mycobacterium bovis* BCG. *J. Bacteriol.* **175**: 7282-7289.

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

Allen, R. C., Yevich, S. J., Orth, R. W., and Steele, R. H. 1974. The superoxide anion and singlet molecular oxygen: their role in the microbicidal activity of the polymorphonuclear leukocyte. *Biochem. Biophys. Res. Commun.* **60**: 909-917.

Almirón, M., Link, A. J., Furlong, D., and Kolter, R. 1992. A novel DNA-binding protein with regulatory and protective roles in starved *Escherichia coli. Genes Dev.* **6**: 2646-2654 .

Alonso-Moraga, A., Bocanegra, A., Torres, J. M., Lopez-Barea, J., and Pueyo, C. 1987. Glutathione status and sensitivity to GSH-reacting compounds of an *Escherichia coli* strain deficient in glutathione metabolism and/or catalase activity. *Mol. Cell. Biochem.* **73**: 61-68.

Altamirano, M., Marostenmaki, J., Wong, A., Fitzgerald, M., Black, W. A., and Smith, J. A. 1994. Mutations in the catalase-peroxidase gene from isoniazid-resistant *Mycobacterium tuberculosis* isolates. *J. Infect. Dis.* **169**: 1162-1165.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, J. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**: 403-410.

Altuvia, S., Almiron, N., Huisman, G., Kolter, R., and Storz, G. 1994. The *dps* promoter is activated by OxyR during growth and by IHF and sigmaS in stationary phase. *Mol. Microbiol.* **13**: 265-272.

Amann, E., and Brosius, J. 1985. 'ATG vectors' for regulated high-level expression of cloned genes in *Escherichia coli. Gene* **40**: 183-190.

Ananthaswamy, H. N., and Eisenstark, A. 1977. Repair of hydrogen peroxide-induced single-strand breaks in *Escherichia coli* deoxyribonucleic acid. *J. Bacteriol.* **130**: 187-191.

Andersen, P. 1994. Effective vaccination of mice against *Mycobacterium tuberculosis* infection with a soluble mixture of secreted mycobacterial proteins. *Infect. Immun.* **62**: 2536-2544.

Andrew, M., O'Brodovich, H., Coates, G., Robertson, G. L., and Gray, G. W. 1985. Hypoxia-induced vasopressin release and coagulopathy in a normal subject. *Aviat. Space Environ. Med.* **56**: 1220-1223.

Arain, T. M., Resconi, A. E., Singh, D. C., and Stover, C. K. 1996. Reporter gene technology to assess activity of antimycobacterial agents in macrophages. *Antimicrob. Agents Chemother.* **40**: 1542-1544.

Aust, S. D., Morehouse, L. D., and Thomas, C. E. 1985. The role of metals in oxygen radical reactions. *Free Radic. Biol. Med.* **1**: 3-26.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (eds). 1987. Current Protocols in Molecular Biology. Wiley Interscience, Harvard Medical School. Vol 1 and 2.

Babior, B. M. 1978. Oxygen-dependent microbial killing by phagocytes (first of two parts). *N. Engl. J. Med.* **298**: 659-668.

Babior, B. M., Kepnes, R. S., and Curnutte, J. T. 1973. Biological defense mechanisms. The production by leukocytes of superoxide, a potential bactericidal agent. *J. Clin. Invest.* **52**: 741-744.

Bagg, A., and Neilands, J. B. 1987. Ferric uptake regulation protein acts as a repressor, employing iron (II) as a cofactor to bind the operator of an iron transport operon in *Escherichia coli. Biochemistry* **26**: 5471-5477.

Bainton, D. F., Ullyot, J. L., and Farquhar, M. G. 1971. The development of neutrophilic polymorphonuclear leukocytes in human bone marrow. *J. Exp. Med.* **134**: 907-934.

Bairoch, A. 1992. PROSITE: a dictionary of sites and patterns in proteins. *Nucleic Acids Res.* **20**: 2013-2018.

Banerjee, A., Dubnau, E., Quemard, A., Balasubramanian,V., Um, K. S., 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.

Banner, D. B. C., Moran, C. P., Jr., and Losick, R. 1983. Deletion analysis of a complex promoter for a developmentally regulated gene from *Bacillus subtilis. J. Mol. Biol.* **168**: 351.

Barclay, R., and Ratledge, C. 1983. Iron-binding compounds of *Mycobacterium avium, M. intracellulare, M. scrofulaceum*, and mycobactin-dependent *M. paratuberculosis* and *M. avium. J. Bacteriol.* **153**: 1138-1146.

Barclay, R. and Wheeler, P. R. 1989. In: C. Ratledge, J. Stanford and J. M. Grange (eds). *The Biology of the Mycobacteria.* pp. 37-106. Academic Press, London.

Barletta, R. G., Kim, D. D., Snapper, S. B., Bloom, B. R., and Jacobs, W. R., Jr. 1992. Identification of expression signals of the mycobacteriophages Bxb1, L1 and TM4 using the *Escherichia-Mycobacterium* shuttle plasmids pYUB75 and pYUB76 designed to create translational fusions to the *lacZ* gene. *J. Gen. Microbiol.* **138**: 23-30.

Barnes, P. F., Grisso, C. L., Abrams, J. S., Band, H., Rea, T. H., and Modlin, R. L. 1992. Gamma delta T lymphocytes in human tuberculosis. *J. Infect. Dis.* **165**: 506-512.

Bartholomew, W. R. 1968. Multiple catalase enzymes in two species of mycobacteria. *Am. Rev. Respir. Dis.* **97**: 710-712.

Bashyam, M. D., Kaushal, D., Das Gupta, S. K., and Tyagi, A. K. 1996. A study of the mycobacterial transcriptional apparatus: identification of novel features in promoter elements. *J. Bacteriol.* **178**: 4847-4853.

Batra, P. P. 1967. Mechanism of photoinduced carotenoid synthesis: Induction of carotenoid synthesis by antimycin A in the absence of light in *Mycobacterium marinum. J. Biol. Chem.* **242**: 5630-5635.

Beaman, L., and Beaman, B. L. 1984. The role of oxygen and its derivatives in microbial pathogenesis and host defense. *Annu. Rev. Microbiol.* **38**: 27-48.

Beaman, B. L., Scates, S. M., Moring, S. E., Deem, R., and Misra, H. P. 1983. Purification and properties of a unique superoxide dismutase from *Nocardia asteroides. J. Biol. Chem.* **258**: 91-96.

Beers, R. F., Jr., and Sizer, I. W. 1952. A spectrophotometric method for measuring the breakdown of hydrogen peroxide by catalase. *J. Biol. Chem.* **195**: 133-140.

Beggs, M. L., Crawford, J. T., and Eisenach, K. D. 1995. Isolation and sequencing of the replication region of *Mycobacterium avium* plasmid pLR7. *J. Bacteriol.* **177**: 4836-4840.

Bekierkunst, A. 1966. Nicotinamide adenine dinucleotide in tubercle bacilli exposed to isoniazid. *Science* **152**: 525.

Bekierkunst, A., and Bricker, A. 1967. Studies on the mode of action of isoniazid on mycobacteria. *Arch. Biochem. Biophys.* **122**: 385-392.

Belkin, S., Smulski, D. R., Vollmer, A. C., Van Dyk, T. K., and Larossa, R. A. 1996. Oxidative stress detection with *Escherichia coli* harbouring a *katG::lux* fusion. *Appl. Environ. Microbiol.* **62**: 2252-2256.

Bergh, S., and Cole, S. T. 1994. MycDB: an integrated mycobacterial database. *Mol. Microbiol.* **12**: 574-534.

Bergler, H., Hogenauer, G., and Turnowsky, F. 1992. Sequences of the *envM* gene and of two mutated alleles in *Escherichia coli*. *J. Gen. Microbiol.* **138**: 2093-2100.

Bergler, H., Wallner, P., Ebeling, A., Leitinger, B., Fuchsbichler, S., Aschauer, H., Kollenz, G., Hogenauer, G., and Turnowsky, F. 1994. Protein EnvM is the NADH-dependent enoyl-ACP reductase (Fabl) of *Escherichia coli*. *J. Biol. Chem.* **269**: 5493-5496.

Berish, S. A., Subbarao, S., Chen, C. Y., Trees, D. L., and Morse, S. A. 1993. Identification and cloning of a *fur* homolog from *Neisseria gonorrhoeae*. *Infect. Immun.* **61**: 4599-4606.

Bernstein, J., Lott, W. A., Steinberg, B. A., and Yale, H. L. 1952. Chemotherapy of experimental tuberculosis. V. Isonicotinic acid hydrazide (Nydrazide) and related compounds. *Am. Rev. Tuberc.* **65**: 357.

Bielski, B. H. J. 1985. Reactivity of HO_2/O_2^- radicals in aqueous solution. *Phys. Chem. Ref. Data* **14**: 1041-1045.

Bishai, W. R., Howard, N. S., Winkelstein, J. A., and Smith, H. O. 1994. Characterization and virulence analysis of catalase mutants of *Haemophilus influenzae*. *Infect. Immun.* **62**: 4855-4860.

Bjorn, M. J., Iglewski, B. H., Ives, S. K., Sadoff, J. C., and Vasil, M. L. 1978. Effect of iron on yields of exotoxin A in cultures of *Pseudomonas aeruginosa* PA-103. *Infect. Immun.* **19**: 785-791.

Bloch, A. B., Cauthen, G. M., Onorato, I. M., Dansbury, K. G., Kelly, G. D., Driver, C. R., and Snider, D. E., Jr. 1994. Nationwide survey of drug resistant tuberculosis in the United States. *JAMA* **271**: 665-671.

Bloch, C. A., Thorne, G. M., and Ausubel, F. M. 1989. General method for site-directed mutagenesis in *Escherichia coli* O18ac:K1:H7: deletion of the inducible superoxide dismutase gene, *sodA*, does not diminish bacteremia in neonatal rats. *Infect. Immun.* **57**: 2141-2148.

Bloom, B. R., and Fine, P. E. M. 1994. In: B. R. Bloom (ed). *Tuberculosis: Pathogenesis, Protection, and Control.* ASM Press, Washington D. C.

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

Blouin, K., Walker, S. G., Smit, J., and Turner, R. F. B. 1996. Characterization of *in vivo* reporter systems for gene expression and biosensor applications based on *luxAB* luciferase genes. *Appl. Environ. Microbiol.* **62**: 2013-2021.

Bodmer, T., Zurcher, G., Imboden, P., and Telenti, A. 1995. Mutation position and type of substitution in the β-subunit of the RNA polymerase influence *in-vitro* activity of rifamycins in rifampicin-resistant *Mycobacterium tuberculosis*. *J. Antimicrob. Chemother.* **35**: 345-348.

Bolker, M., and Kahmann, R. 1989. The *Escherichia coli* regulatory protein OxyR discriminates between methylated and unmethylated states of the phage Mu *mom* promoter. *EMBO J.* **8**: 2403-2410.

Böttger, E. C. 1994. Resistance to drugs targeting protein synthesis in mycobacteria. *Trends Microbiol.* **2**: 416-421.

Boyd, J., Oza, M. N., and Murphy, J. R. 1990. Molecular cloning and DNA sequence analysis of a diphtheria tox iron-dependent regulatory element (*dtxR*) from *Corynebacterium diphtheriae. Proc. Natl. Acad. Sci. USA* **87**: 5968-5972.

Boyer, H. W., and Roulland-Dussoix, D. 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli. J. Mol. Biol.* **41**: 459.

Breimer, L. H., and Lindahl, T. 1984. DNA glycosylase activities for thymine residues damaged by ring saturation, fragmentation, or ring contraction are functions of endonuclease III in *Escherichia coli. J. Biol. Chem.* **259**: 5543-5548.

Bridges, B. A. 1983. Psoralens and serendipity: aspects of the genetic toxicology of 8-methoxypsoralen. *Environ. Mutagen.* **5**: 329-339.

Bridges, B. A. 1984. Further characterization of repair of 8-methoxypsoralen cross-links in UV-excision-defective *E. coli. Mutat. Res.* **132**: 153-160.

Bridges, B. A., and Stannard, M. 1982. A new pathway for repair of cross-linkable 8-methoxypsoralen mono-adducts in *uvr* strains of *Escherichia coli. Mutat. Res.* **92**: 9-14.

Bridges, B. A., and Woodgate, R. 1985. Mutagenic repair in *Escherichia coli*: products of the *recA* gene and of the *umuD* and *umuC* genes act at different steps in UV-induced mutagenesis. *Proc. Natl. Acad. Sci. USA* **82**: 4193-4197.

Brown, S. M., Howell, M. L., Vasil, M. L., Anderson, A., and Hassett, D. J. 1995. Cloning and characterization of the *katB* gene of *Pseudomonas aeruginosa* encoding a hydrogen peroxide-inducible catalase: purification of KatB, cellular localization, and demonstration that it is essential for optimal resistance to hydrogen peroxide. *J. Bacteriol.* **177**: 6536-6544.

Brown-Peterson, N. J., and Salin, M. L. 1993. Purification of a catalase-peroxidase from *Halobacterium halobium*: Characterization of some unique properties of the halophilic enzyme. *J. Bacteriol.* **175**: 4197-4202.

Bullen, J. J., Rogers, H. J., and Griffiths, E. 1978. Role of iron in bacterial infection. *Curr. Top. Microbiol. Immunol.* **80**: 1-35.

Burgess, R. R. 1976. In: R. Losick and M. Chamberlin (eds). *RNA polymerase.* pp. 69-100. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York.

Busby, S., and Ebright, R. H. 1994. Promoter structure, promoter recognition, and transcription activation in prokaryotes. *Cell* **79**: 743-746.

Butler, W. R., and Kilburn, J. O. 1983. Susceptibility of *M. tuberculosis* to pyrazinamide and its relationship to pyrazinamidase activity. *Antimicrob. Agents Chemother.* **24**: 600-601.

Calderwood, S. B., and Mekalanos, J. J. 1987. Iron regulation of Shiga-like toxin expression in *Escherichia coli* is mediated by the *fur* locus. *J. Bacteriol.* **169**: 4759-4764.

Canetti, G. 1965. The J. Burns Amberson lecture. Present aspects of bacterial resistance in tuberculosis. *Am. Rev. Respir. Dis.* **92**: 687-703.

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

Carlsson, J., and Carpenter, V. S. 1980. The *recA*$^+$ gene product is more important than catalase and superoxide dismutase in protecting *Escherichia coli* against hydrogen peroxide toxicity. *J. Bacteriol.* **142**: 319-321.

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

Caron, P. R., Kushner, S. R., and Grossman, L. 1985. Involvement of helicase II (*uvrD* gene product) and DNA polymerase I in excision mediated by the *uvr*ABC protein complex. *Proc. Natl. Acad. Sci. USA* **82**: 4925-4929.

Cendrin, F., Jouve, H. M., Gaillard, J., Thibault, P., and Zaccai, G. 1994. Purification and properties of a halophilic catalase peroxidase from *Haloarcula marismortui. Biochim. Biophys. Acta* **1209**: 1-9.

Chalke, H. D. 1962. The impact of tuberculosis on history, literature and art. *Med. Hist.* **6**: 301-318.

Chamnongpol, S., Vattanaviboon, P., Loprasert, S., and Mongkolsuk, S. 1995. Atypical oxidative stress regulation of a *Xanthomonas oryzae* pv. *oryzae* monofunctional catalase. *Can. J. Microbiol.* **41**: 541-547.

Chan, B., and Busby, S. 1989. Recognition of nucleotide sequences at the *Escherichia coli* galactose operon P1 promoter by RNA polymerase. *Gene* **84**: 227-236.

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

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. *Proc. Natl. Acad. Sci. USA* **86**: 2453-2457.

Chen, L., Keramati, L., and Helmann, J. D. 1995. Coordinate regulation of *Bacillus subtilis* peroxide stress genes by hydrogen peroxide and metal ions. *Proc. Natl. Acad. Sci. USA* **92**: 8190-8194.

Cheo, D. L., Bayles, K. W., and Yasbin, R. E. 1991. Cloning and characterization of DNA damage-inducible promoter regions from *Bacillus subtilis. J. Bacteriol.* **173**: 1696-1703.

Cherayil, B. J. and Young, R. A. 1988. A 28-kDa protein from *Mycobacterium leprae* is a target of the human antibody response in lepromatous leprosy. *J. Immun.* **141**: 4370-4375.

Christman, M. F., Morgan, R. W., Jacobson, F. S., and Ames, B. N. 1985. Positive control of a regulon for defenses against oxidative stress and some heat-shock proteins in *Salmonella typhimurium. Cell* **41**: 753-762.

Christman, M. F., Storz, G., and Ames, B. N. 1989. OxyR, a positive regulator of hydrogen peroxide-inducible genes in *Escherichia coli* and *Salmonella typhimurium*, is homologous to a family of bacterial regulatory proteins. *Proc. Natl. Acad. Sci. USA* **86**: 3484-3488.

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

Cirillo, J. D., Weisbrod, T. R., Pascopella, L., Bloom, B. R., and Jacobs, W. R., Jr. 1994. Isolation and characterization of the aspartokinase and aspartate semialdehyde dehydrogenase operon from mycobacteria. *Mol. Microbiol.* **11**: 629-639.

Claiborne, A., and Fridovich, I. 1979. Purification of the *o*-dianisidine peroxidase from *Escherichia coli*. *Br. J. Biol. Chem.* **254**: 4245-4252.

Clare, D. A., Duong, M. N., Darr, D., Archibald, F., and Fridovich, I. 1984. Effects of molecular oxygen on detection of superoxide radical with nitroblue tetrazolium and on activity stains for catalase. *Anal. Biochem.* **140**: 532-537.

Clarke, G. A., Kelley, M. A., Grange, J. M., and Hill, M. C. 1987. The evolution of mycobacterial disease in human populations. A re-evaluation. *Current Anthropology.* **28**: 45-62.

Clarke-Curtiss, J. E., Jacobs, 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*. *J. Bacteriol.* **161**: 1093-1102.

Clemens, D. L., and Horwitz, M. A. 1996. The *Mycobacterium tuberculosis* phagosome interacts with early endosomes and is accessible to exogenously administered transferrin. *J. Exp. Med.* **184**: 1349-1355.

Cockerill, F. R., Uhl, J. R., Temesgen, Z., Zhang, Y., Stockman, L., Roberts, G. D., Williams, D. L., and Kline, B. C. 1995. Rapid identification of a point mutation of the *Mycobacterium tuberculosis* catalase-peroxidase (*katG*) gene associated with isoniazid resistance. *J. Infect. Dis.* **171**: 240-245.

Cohn, Z. A. 1963. The fate of bacteria within phagocytic cells. I. The degradation of isotopically labelled bacteria by polymorphonuclear leucocytes and macrophages. *J. Exp. Med.* **117**: 27-42.

Cohn, M. L., Kovitz, C., Oda, U., and Middlebrook, G. 1954. Studies on isoniazid and tubercle bacilli. II. The growth requirements, catalase activities, and pathogenic properties of isoniazid-resistant mutants. *Am. Rev. Tuberc.* **70**: 641-664.

Colditz, G. A., Brewer, T. F., Berkey, C. S., Wilson, M. E., Burdick, E., Fineberg, H. V., and Mosteller, F. 1994. Efficacy of BCG vaccine in the prevention of tuberculosis. Meta-analysis of the published literature. *J. Am. Med. Assoc.* **271**: 698-707.

Cole, R. S. 1973. Repair of DNA containing interstrand cross-links in *Escherichia coli*: sequential excision and recombination. *Proc. Natl. Acad. Sci. USA* **70**: 1064-1068.

Cole, S. T. 1994. Mechanisms of drug resistance in *Mycobacterium tuberculosis*. *Immunobiol.* **191**: 584-585.

Collins, F. M. 1971. Relative susceptibility of acid-fast and non-acid-fast bacteria to ultraviolet light. *Appl. Microbiol.* **21**: 411.

Collins, F. M. 1989. Mycobacterial disease, immunosuppression and acquired immunodeficiency syndrome. *Clin. Microbiol. Rev.* **2**: 360-377.

Collins, M. E., Patki, A., Wall, S., Nolan, A., Goodger, J., Woodward, M. J., and Dale, J. W. 1990. Cloning and characterization of the gene for the '19 kDa' antigen of *Mycobacterium bovis*. *J. Gen. Microbiol.* **136**: 1429-1436.

Colston, M. J., and Davis, E. O. 1994. Homologous recombination, DNA repair, and mycobacterial *recA* genes. In: B. R. Bloom (ed). *Tuberculosis: Pathogenesis, Protection and Control.* pp. 217-238. ASM Press, Washington D.C.

Cowing, D. W., Bardwell, J. C. A., Craig, E. A., Woolford, C., Hendrix, R. W., and Gross, C. A. 1985. Consensus sequence for *Escherichia coli* heat shock gene promoters. *Proc. Natl. Acad. Sci. USA* **82**: 2679-2683.

Cox, M. M., and Lehman, I. R. 1981. RecA protein of *Escherichia coli* promotes branch migration, a kinetically distinct phase of DNA strand exchange. *Proc. Natl. Acad. Sci. USA* **78**: 3433-3437.

Crane, D., Holmes, R., and Masters, C. 1982. Proteolytic modification of mouse liver catalase. *Biochem. Biophys. Res. Commun.* **104**: 1567-1572.

Cummins, S. L. 1920. Tuberculosis in primitive tribes and its bearing on the tuberculosis of civilized communities. *Int. J. Public Health* **1**: 137-171.

Cunningham, R. P., and Weiss, B. 1985. Endonuclease III (*nth*) mutants of *Escherichia coli. Proc. Natl. Acad. Sci. USA* **82**: 474-478.

Curcic, R., Dhandayuthapani, S., and Deretic, V. 1994. Gene expression in mycobacteria: transcriptional fusions based on *xylE* and analysis of the promoter region of the response regulator *mtrA* from *Mycobacterium tuberculosis. Mol. Microbiol.* **13**: 1057-1064.

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

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

David, H. L. 1973. Response of mycobacteria to ultraviolet light radiation. *Am. Rev. Respir. Dis.* **108**: 1175-1185.

David, H. L., Jones, W. D., Jr., and Newman, C. M. 1971. Ultraviolet light inactivation and photoreactivation in the mycobacteria. *Infect. Immunol.* 4: 318.

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. *J. Bacteriol.* **173**: 5653-5662.

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

Deisseroth, A., and Dounce, A. L. 1970. Catalase: physical and chemical properties, mechanism of catalysis and physiological role. *Phys. Rev.* **50**: 319-375.

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

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

de Lorenzo, V., Giovannini, F., Herrero, M., and Neilands, J. B. 1988. Metal ion regulation of gene expression. Fur repressor-operator interaction at the promoter region of the aerobactin system of pColV-K30. *J. Mol. Biol.* **203**: 875-884.

de Lorenzo, V., Wee, S., Herrero, M., and Neilands, J. B. 1987. Operator sequences of the aerobactin operon of plasmid ColV-K30 binding the ferric uptake regulation (Fur) repressor. *J. Bacteriol.* **169**: 2624-2630.

Demaio, J., Zhang, Y., Ko, C., Young, D. B., and Bishai, W. R. 1996. A stationary-phase stress-response sigma factor from *Mycobacterium tuberculosis. Proc. Natl. Acad. Sci. USA* **93**: 2790-2794.

Demple, B. 1991. Regulation of bacterial oxidative stress genes. *Annu. Rev. Genet.* **25**: 315-337.

Demple, B., and Halbrook, J. 1983. Inducible repair of oxidative DNA damage in *Escherichia coli. Nature (London)* **304**: 466-468.

Demple, B., Halbrook, J., and Linn, S. 1983. *Escherichia coli xthA* mutants are hypersensitive to hydrogen peroxide. *J. Bacteriol.* **153**: 1079-1082.

Demple, B., and Linn, S. 1980. DNA N-glycosylases and UV repair. *Nature (London)* **287**: 203-208.

Demple, B., and Linn, S. 1982. 5,6-saturated thymine lesions in DNA: production by ultraviolet light or hydrogen peroxide. *Nucleic Acids Res.* **10**: 3781-3789.

Department of National Health and Population Development. 1990. Tuberculosis Control Programme - 1988. *Epidemiol. Comm.* **17**: 3-13.

Deretic, V., Philipp, W., Dhandayuthapani, S., Mudd, M. H., Curcic, R., Garbe, T., Heym, B., Via, L. E., and Cole, S. T. 1995. *Mycobacterium tuberculosis* is a natural mutant with an inactivated oxidative-stress regulatory gene: implications for sensitivity to INH. *Mol. Microbiol.* **17**: 889-900.

Dessen, A., Quemard, A., Blanchard, J. S., Jacobs, W. R., Jr., and Sacchettini, J. C. 1995. Crystal structure and function of the isoniazid target of *Mycobacterium tuberculosis*. *Science* **267**: 1638-1641.

Devi, B. G., Shaila, M. S., Ramakrishnan, 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. *Biochem. J.* **149**: 187-197.

de Wet, J. R., Wood, K. V., Helinski, D. R., and DeLuca, M. 1985. Cloning of firefly luciferase cDNA and the expression of active luciferase in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **82**: 7870-7873.

Dhandayuthapani, S., Mudd, M. H., and Deretic, V. 1997. Interactions of OxyR with the promoter region of the *oxyR* and *ahpC* genes from *Mycobacterium leprae* and *Mycobacterium tuberculosis*. *J. Bacteriol.* **179**: 2401-2409.

Dhandayuthapani, S., Via, L. E., Thomas, C. A., Horowitz, P. M., Deretic, D., and Deretic, V. 1995. Green fluorescent protein as a marker for gene expression and cell biology of mycobacterial interactions with macrophages. *Mol. Microbiol.* **17**: 901-912.

Dhandayuthapani, S., Zhang, Y., Mudd, M. H., and Deretic, V. 1996. Oxidative stress response and its role in sensitivity to isoniazid in mycobacteria: characterization and inducibility of *ahpC* by peroxides in *Mycobacterium smegmatis* and lack of expression in *M. aurum* and *M. tuberculosis*. *J. Bacteriol.* **178**: 3641-3649.

Diaz, G. A., and Wayne, L. G. 1974. Isolation and characterization of catalase produced by *M. tuberculosis*. *Am. Rev. Respir. Dis.* **110**: 312-319.

Ding, X., Zeng, H., Schiering, N., Ringe, D., and Murphy, J. R. 1996. Identification of the primary metal ion-activation sites of the diphtheria *tox* repressor by X-ray crystallography and site-directed mutational analysis. *Nat. Struct. Biol.* **3**: 382-387.

DiRita, V. J., and Mekalanos, J. J. 1989. Genetic regulation of bacterial virulence. *Annu. Rev. Genet.* **23**: 455-482.

Dixon, D. A., and Kowalczykowski, S. C. 1991. Homologous pairing *in vitro* stimulated by the recombination hotspot Chi. *Cell* **66**: 361-371.

Douglass, J., and Steyn, L. M. 1993. A ribosomal gene mutation in streptomycin-resistant *Mycobacterium tuberculosis* isolates. *J. infect. Dis.* **167**: 1505-1506.

Doukhan, L., Predich, M., Nair, G., Dussurget, O., Mandic-Mulec, I., Cole, S. T., Smith, D. R., and Smith, I. 1995. Genomic organization of the mycobacterial sigma gene cluster. *Gene* **165**: 67-70.

Draper, P. 1967. The aliphatic acylamide amidohydrolase of *Mycobacterium smegmatis*: its inducible nature and relation to acyl-transfer to hydroxylamine. *J. Gen. Microbiol.* **46**: 111-123.

Dubnau, E., Soares, S., Huang, T. J., and Jacobs, W. R. 1996. Overproduction of mycobacterial ribosomal protein S13 induces catalase peroxidase activity and hypersensitivity to isoniazid in *Mycobacterium smegmatis*. *Gene* **170**: 17-22.

Dudley, M. A., and Willet, H. P. 1966. Comparison of nicotinamide adenine dinucleotide synthesis by *Mycobacterium tuberculosis* and *Mycobacterium bovis*. *Can. J. Microbiol.* **12**: 1225-1233.

Duncan, J., Hamilton, L., and Friedberg, E. C. 1976. Enzymatic degradation of uracil-containing DNA. *J. Virol.* **19**: 338-345.

Duncan, B. K., and Weiss, B. 1982. Specific mutator effects of *ung* (uracil-DNA glycosylase) mutations in *Escherichia coli*. *J. Bacteriol.* **151**: 750-755.

Dunford, H. B., and Stillman, J. S. 1976. On the function and mechanism of action of peroxidases. *Coord. Chem. Rev.* **19**: 187-251.

Dussurget, O., Rodriguez, M., and Smith, I. 1996. An *ideR* mutant of *Mycobacterium smegmatis* has derepressed siderophore production and an altered oxidative-stress response. *Mol. Microbiol.* **22**: 535-544.

Eisenstark, A. 1989. Bacterial genes involved in response to near-ultraviolet radiation. *Adv. Genet.* **26**: 99-147.

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

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

Farr, S. B., Natvig, D. O., and Kogoma, T. 1985. Toxicity and mutagenicity of plumbagin and the induction of a possible new DNA repair pathway in *Escherichia coli*. *J. Bacteriol.* **164**: 1309-1316.

Farr, S. B., Touati, D., and Kogoma, T. 1988 . Effects of oxygen stress on membrane functions in *Escherichia coli*: role of HPI catalase. *J. Bacteriol.* **170**: 1837-1842.

Ferrazoli, L., Palaci, M., Telles, M. A. D., Ueki, S. Y., Kritski, A., Marques, L. R. M., Ferreira, O. C., and Riley, L. W. 1995. Catalase expression, KatG, and MIC of isoniazid for *Mycobacterium tuberculosis* isolates from Sao Paulo, Brazil. *J. Infect. Dis.* **171**: 237-240.

Fields, B. S., Barbaree, J. M., Shotts, E. B., Jr., Feeley, J. C., Morrill, W. E., Sanden, G. N., and Dykstra, M. J. 1986. Comparison of guinea pig and protozoan models for determining virulence of *Legionella* species. *Infect. Immun.* **53**: 553-559.

Filice, G. A. 1983. Resistance of *Nocardia asteroides* to oxygen-dependent killing by neutrophils. *J. Infect. Dis.* **148**: 861-867.

Finken, M., Kirschner, P., Meier, A., Wrede, A., and Böttger, E. C. 1993. Molecular basis of streptomycin resistance in *Mycobacterium tuberculosis*: alterations of the ribosomal protein S12 gene and point mutations within a functional 16S ribosomal RNA psuedoknot. *Mol. Microbiol.* **9**: 1239-1246.

Finzel, B. C., Poulos, T. L., and Kraut, J. 1984. Crystal structure of yeast cytochrome *c* peroxidase refined at 1.7-A° resolution. *J. Biol. Chem.* **259**: 13027-13036.

Fiss, E. H., Yu, S., and Jacobs, W. R., Jr. 1994. Identification of genes involved in the sequestration of iron in mycobacterium: the ferric exochelin biosynthetic and uptake pathways. *Mol. Microbiol.* **14**: 557-569.

Fita, I., and Rossmann, M. G. 1985. The active centre of catalase. *J. Mol. Biol.* **185**: 21-37.

Fogliano, M., and Schendel, P. F. 1981. Evidence for the inducibility of the *uvrB* operon. *Nature (London)* **289**: 196-198.

Forkl. H., Drews, G., and Tadros, M. H. 1996. Promoter analysis of the catalase-peroxidase gene (*cpeA*) from *Rhodobacter capsulatus. FEMS Microbiol. Lett.* **137**: 169-174.

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

Fox, H. H. 1952. The chemical approach to the control of tuberculosis. *Science* **116**: 129.

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

Fredrick, K., and Helmann, J. D. 1997. RNA polymerase sigma factor determines start-site selection but is not required for upstream promoter element activation on heteroduplex (bubble) templates. *Proc. Natl. Acad. Sci. USA* (in Press).

Fridovich, I. 1986. Biological effects of superoxide radical. *Arch. Biochem. Biophys.* **247**: 1-11.

Friedberg, E. C., and King, J. J. 1969. Endonucleolytic cleavage of UV-irradiated DNA controlled by the V^+ gene in phage T4. *Biochem. Biophys. Res. Commun.* **37**: 646-651.

Friedberg, E. C., Walker, G. C., and Siede, W. (eds). 1995. DNA Repair and Mutagenesis. ASM Press, Washington D. C.

Frieden, T. R., Sterling, T., Pablos-Mendez, A., Kilburn, J. O., Cauthen, G. M., and Dooley, S. W. 1993. The emergence of drug-resistant tuberculosis in New York City. *N. Engl. J. Med.* **328**: 521-526.

Frothingham, R., Hills, H. G., and Wilson, K. H. 1994. Extensive DNA sequence conservation throughout the *Mycobacterium tuberculosis* complex. *J. Clin. Microbiol.* **32**: 1639-1643.

Fsihi, H., Vincent, V., and Cole, S. T. 1996. Homing events in the *gyrA* gene of some mycobacteria. *Proc. Natl. Acad. Sci. USA* **93**: 3410-3415.

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. Infect. Immun.* **52**: 50-55.

Galán, J. E., and Curtiss III, R. 1990. Expression of *Salmonella typhimurium* genes required for invasion is regulated by changes in DNA supercoiling. *Infect. Immun.* **58**: 1879-1885.

Galdston, I. 1942. The dynamics of epidemiology in relation to epidemic tuberculosis. *Am. Rev. Tuberc.* **45**: 609-615.

Gale, E. F., Cundliffe, E., Reynolds, P. E., Richmond, M. H., and Waring, M. J. 1981. The molecular basis of antibiotic action. John Wiley and Sons, Inc., New York.

Gamper, H. B., and Hearst, J. E. 1982. A topological model for transcription based on unwinding angle analysis of *E. coli* RNA polymerase binary, initiation and ternary complexes. *Cell* **29**: 81-90.

Gangadharam, P. R. J., Harold, F. M., and Schaefer, W. B. 1963. Selective inhibition of nucleic acid synthesis in *Mycobacterium tuberculosis* by isoniazid. *Nature (London)* **198**: 712-714.

Gangadharam, P. R. J., and Pratt, P. K. 1984. Susceptibility of *Mycobacterium intracellulare* to hydrogen peroxide. *Am. Rev. Respir. Dis.* **130**: 309-311.

Garbe, T. R., Hibler, N. S., and Deretic, V. 1996. Isoniazid induces expression of the antigen 85 complex in *Mycobacterium tuberculosis. Antimicrob. Agents Chemother.* **40**: 1754-1756.

Gardner, P., and Fridovich, I. 1991. Superoxide sensitivity of the *Escherichia coli* 6-phosphogluconate dehydratase. *J. Biol. Chem.* **266**: 1478-1483.

Gartenberg, M. R., and Crothers, D. M. 1991. Synthetic DNA bending sequences increase the rate of *in vitro* transcription initiation at the *Escherichia coli lac* promoter. *J. Mol. Biol.* **219**: 217-230.

Gheorghiu, M. 1990. The present and future role of BCG vaccine in tuberculosis control. *Biologicals* **18**: 135-141.

Glatthaar, E. 1982. Where have we gone wrong- a look at the future? *Conference on TB in South Africa: Consumption in the land of plenty, University of Cape Town, Cape Town, 2-4 August.*

Goerlich, O., Quillardet, P., and Hofnung, M. 1989. Induction of the SOS response by hydrogen peroxide in various *Escherichia coli* mutants with altered protection against oxidative DNA damage. *J. Bacteriol.* **171**: 6141-6147.

Goethals, K., VanMontagu, M., and Holsters, M. 1992. Conserved motifs in a divergent *nod* box of *Azorhizobium caulinodans* ORS571 reveal a common structure in promoters regulated by LysR-type proteins. *Proc. Natl. Acad. Sci. USA* **89**: 1646-1650.

Goldberg, I., and Hochman, A. 1989. Purification and characterization of a novel type of catalase from the bacterium *Klebsiella pneumoniae. Biochim. Biophys. Acta* **991**: 330-336.

Gordon, A. H., Hart, P. D., and Young, M. R. 1980. Ammonia inhibits phagosome-lysosome fusion in macrophages. *Nature (London)* **286**: 79-80.

Gordon, S., Parish, T., Roberts, I. S., and Andrew, P. W. 1994. The application of luciferase as a reporter of environmental regulation of gene expression in mycobacteria. *Lett. Appl. Microbiol.* **19**: 336-340.

Goto, M., Oka, S., Tachikawa, N., Kitada, K., Wada, M., Abe, C., Shimada, K., and Kimura, S. 1995. *katG* sequence deletion is not the major cause of isoniazid resistance in Japanese and Yemeni *Mycobacterium tuberculosis* isolates. *Mol. Cell. Probes* **9**: 433-439.

Goto, Y., Taniguchi, H., Udou, T., Mizuguchi, Y., and Tokunaga, T. 1991. Development of a new host vector system in mycobacteria. *FEMS Microbiol. Lett.* **83**: 277-282.

Goyal, M., Young, D., Zhang, Y., Jenkins, P. A., and Shaw, R. J. 1994. PCR amplification of variable sequence upstream of *katG* to subdivide strains of *Mycobacterium tuberculosis* complex. *J. Clin. Microbiol.* **32**: 3070-3071.

Grana, D., Youderian, P., and Susskind, M. 1985. Mutations that improve the *ant* promoter of *Salmonella* phage P22. *Genetics* **110**: 1.

Grange, J. M., and Bishop, P. J. 1982. "Uber tuberkulose": a tribute to Robert Koch's discovery of the tubercle bacillus, 1882. *Tubercle* **63**: 3-17.

Grange, J. M., and Collins, C. H. 1987. Bovine tubercle bacilli and disease in animals and man. *Epidemiol. Infect.* **92**: 221-234.

Grassberger, M. A., Turnowsky, F., and Hildebrand, J. 1984. Preparation and antibacterial activities of new 1,2,3-diazaborine derivatives and analogues. *J. Med. Chem.* **27**: 947.

Graves, M. C., and Rabinowitz, J. C. 1986. *In vivo* and *in vitro* transcription of the *Clostridium pasteurianum* ferredoxin gene. Evidence for "extended" promoter elements in gram-positive organisms. *J. Biol. Chem.* **261**: 11409-11415.

Greenberg, J. T., and Demple, B. 1988. Overproduction of peroxide-scavenging enzymes in *Escherichia coli* suppresses spontaneous mutagenesis and sensitivity to redox-cycling agents in *oxyR*⁻ mutants. *EMBO J.* 7: 2611-2617.

Greenberg, J. T., and Demple, B. 1989. A global response induced in *Escherichia coli* by redox-cycling agents overlaps with that induced by peroxide stress. *J. Bacteriol.* **171**: 3933-3939.

Greenberg, J. T., Monach, P., Chou, J. H., Josephy, P. D., and Demple, B. 1990. Positive control of a global antioxidant defense regulon activated by superoxide-generating agents in *Escherichia coli. Proc. Natl. Acad. Sci. USA* **87**: 6181-6185.

Gregory, E. M., and Fridovich, I. 1974. Visualization of catalase on acrylamide gels. *Analyt. Biochem.* **58**: 57-62.

Grosset, J. 1980. Bacteriologic basis of short-course chemotherapy for tuberculosis. *Clin. Chest Med.* **1**: 231-241.

Grossman, L., and Thiagalingam, S. 1993. Nucleotide excision repair, a tracking mechanism in search of damage. *J. Biol. Chem.* **268**: 16871-16874.

Grumbach, F. 1969. Experimental "*in vivo*" studies of new antituberculosis drugs: capreomycin, ethambutol, rifampicin. *Tubercle* **50**: 12-21.

Guan, C., Li. P., Riggs, P. D., and Inouye, H. 1987. Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. *Gene* **67**: 21-30.

Guilhot, C., Otal, I., Vanrompaey, I., Martin, C., and Gicquel, B. 1994. Efficient transposition in mycobacteria- construction of *Mycobacterium smegmatis* insertional mutant libraries. *J. Bacteriol.* **176**: 535-539.

Günter, K., Toupet, C., and Schupp, T. 1993. Characterization of an iron-regulated promoter involved in desferrioxamine B synthesis in *Streptomyces pilosus*: repressor binding site and homology to the diphtheria toxin gene promoter. *J. Bacteriol.* **175**: 3295-3302.

Günter-Seeboth, K., and Schupp, T. 1995. Cloning and sequence analysis of the *Corynebacterium diphtheria dtxR* homologue from *Streptomyces lividans* and *Streptomyces pilosus* encoding a putative iron repressor protein. *Gene* **166**: 117-119.

Gupta, S., and Tyagi, A. K. 1993. Sequence of a newly identified *Mycobacterium tuberculosis* gene encoding a protein with sequence homology to virulence-regulating proteins. *Gene* **126**: 157-158.

Haas, W. H., Schilke, K., Brand, J., Amtherle, B., Weyer, K., Fourie, P. B., Bretzel, G., Sticht-Groh, V., and Bremer, H. J. 1997. Molecular analysis of *katG* gene mutations in strains of *Mycobacterium tuberculosis* complex from Africa. *Antimicrob. Agents Chemother.* **41**: 1601-1603.

Hagensee, M. E., and Moses, R. E. 1989. Multiple pathways for repair of hydrogen peroxide-induced DNA damage in *Escherichia coli. J. Bacteriol.* **171**: 991-995.

Hahn, W. E., Pettijohn, D. E., and Van Ness, J. 1977. One strand equivalent of the *Escherichia coli* genome is transcribed: complexity and abundance classes of mRNA. *Science* **197**: 582-585.

Hale, T. L. 1991. Genetic basis for virulence in *Shigella* species. *Microbiol. Rev.* **55**: 206-224.

Hall, R. M., Sritharan, M., Messenger, A. J. M., and Ratledge, C. 1987. Iron transport in *Mycobacterium smegmatis*: occurrence of iron-regulated envelope proteins as potential receptors for iron uptake. *J. Gen. Microbiol.* **133**: 2107-2114.

Hantke, K. 1987. Ferrous iron transport mutants in *Escherichia coli* K12. *FEMS Microbiol. Lett.* **44**: 53-57.

Hare, R. 1967. In: D. Brothwell and A. T. Sandison (eds). *Diseases in Antiquity.* Chapter 8. Charles C. Thomas, Springfield.

Harley, C. B., and Reynolds, R. P. 1987. Analysis of *E. coli* promoter sequences. *Nucleic Acids Res.* **15**: 2343-2361.

Harshey, R. M., and Ramakrishnan, T. 1977. Rate of ribonucleic acid chain growth in *Mycobacterium tuberculosis* H37Rv. *J. Bacteriol.* **129**: 616-622.

Hassan, H. M. 1989. Microbial superoxide dismutases. *Adv. Genet.* **26**: 65-97.

Hassan, H. M., and Fridovich, I. 1977. Regulation of the synthesis of superoxide dismutase in *Escherichia coli. J. Biol. Chem.* **252**: 7667-7672.

Hassett, D. J., Charniga, L., and Cohen, M. S. 1990. *recA* and catalase in H_2O_2-mediated toxicity in *Neisseria gonorrhoeae. J. Bacteriol.* **172**: 7293-7295.

Hassett, D. J., Sokol, P. A., Howell, M. L., Ma, J., Schweizer, H. T., Oschner, U., and Vasil, M. L. 1996. Ferric uptake regulator (Fur) mutants of *Pseudomonas aeruginosa* demonstrate defective siderophore-mediated iron uptake, altered aerobic growth, and decreased superoxide dismutase and catalase activities. *J. Bacteriol.* **178**: 3996-4003.

Hawley, D. K., and McClure, W. R. 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res.* **11**: 2237-2255.

Haynes, R. H. 1966. The interpretation of microbial inactivation and recovery phenomena. *Radiat. Res.* **6**: S1.

Heimburger, A., and Eisenstark, A. 1988. Compartmentalizaton of catalases in *Escherichia coli. Biochem. Biophys. Res. Commun.* **154**: 392-395.

Henikoff, S., Haughn, G. W., Calvo, J. M., and Wallace, J. C. 1988. A large family of bacterial activator proteins. *Proc. Natl. Acad. Sci. USA* **85**: 6602-6606.

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

Hérouart, D., Sigaud, S., Moreau, S., Frendo, P., Touati, D., and Puppo, A. 1996. Cloning and characterization of the *katA* gene of *Rhizobium meliloti* encoding a hydrogen peroxide-inducible catalase. *J. Bacteriol.* **178**: 6802-6809.

Heym, B., Alzari, P. M., Honoré, N., and Cole, S. T. 1995. Missense mutations in the catalase-peroxidase gene, *katG*, are associated with isoniazid resistance in *Mycobacterium tuberculosis. Mol. Microbiol.* **15**: 235-245.

Heym, B., and Cole, S. T. 1992. Isolation and characterization of isoniazid-resistant mutants of *Mycobacterium smegmatis* and *M. aurum. Res. Microbiol.* **143**: 721-730.

Heym B., Honoré, N., Truffat-Perrot, C., Banerjee, A., Schura, C., Jacobs, W. R., Jr., van Embden, J. D. A., Grosset, J. H., and Cole, S. T. 1994. The implications of multidrug resistance for the future of short course chemotherapy of tuberculosis: a molecular study. *Lancet* **344**: 293-298.

Heym, B., Stavropoulos, E., Honoré, N., Domenech, P., Saint-Joanis, B., Wilson, T. M., Collins, D. M., Colston, J., and Cole, S. T. 1997. Effects of overexpression of the alkyl hydroperoxide reductase AhpC on the virulence and isoniazid resistance of *Mycobacterium tuberculosis. Infect. Immun.* **65**: 1395-1401.

Heym, B., Zhang Y., Poulet, S., Young, D., and Cole, S. T. 1993. Characterization of the *katG* gene encoding a catalase-peroxidase required for isoniazid susceptibility of *Mycobacterium tuberculosis. J. Bacteriol.* **175**: 4255-4259.

Hickey, M. J., Arain, T. M., Shawar, R. M., Humble, D. J., Langhorne, M. H., Morgenroth, J. N., and Stover, C. K. 1996. Luciferase *in vivo* expression technology- use of recombinant mycobacterial reporter strains to evaluate antimycobacterial activity in mice. *Antimicrob. Agents Chemother.* **40**: 400-407.

Hickey, E. K., and Cianciotto, N. P. 1994. Cloning and sequencing of the *Legionella pneumophila fur* gene. *Gene* **143**: 117-121.

Higgins, C. F., and Dorman, C. J. 1990. DNA supercoiling: a role in the regulation of gene expression and bacterial virulence. In: R. Rappuoli (ed). *Bacterial Protein Toxins.* pp. 293-301. Gustav Fischer, New York.

Higgins, C. F., Dorman, C. J., and Nibhriain, N. 1990. Environmental influences on DNA supercoiling: a novel mechanism for the regulation of gene expression. In: K. Drlica and M. Riley (eds). *Bacterial Chromosome.* pp. 421-434. ASM Press, Washington D. C.

Higgins, D. G., Thompson, J. D., and Gibson, T. J. 1996. Using CLUSTAL for multiple sequence alignments. *Meth. Enzymol.* **266**: 383-402.

Hill, O., and Allen, H. 1978. The chemistry of dioxygen and its reduction products. In: D. W. Fitzsimons (ed). *Oxygen free radicals and tissue damage.* p. 5-12. Elsevier Science Publishing, Inc., New York.

Hochman, A., Figueredo, A., and Wall, J. D. 1992. Physiological functions of hydroperoxidases in *Rhodobacter capsulatus. J. Bacteriol.* **174**: 3386-3391.

Hochman, A., and Goldberg, I. 1991. Purification and characterization of a catalase-peroxidase and a typical catalase from the bacterium *Klebsiella pneumoniae. Biochim. Biophys. Acta* **1077**: 299-307.

Hochman, A., and Shemesh, A. 1987. Purification and characterization of a catalase-peroxidase from the photosynthetic bacterium *Rhodopseudomonas capsulata. J. Biol. Chem.* **262**: 6871-6876.

Hochschild, A., Irwin, N., and Ptashne, M. 1983. Repressor structure and the mechanism of positive control. *Cell* **32**: 319-325.

Hoerter, J., Eisenstark, A., and Touati, D. 1989. Mutations by near-ultraviolet radiation in *Escherichia coli* strains lacking superoxide dismutase. *Mutat. Res.* **215**: 161-165.

Hok, T. T. 1964. A comparative study of the susceptibility to ethionamide, thiosemicarbazone, and isoniazid of tubercle bacilli from patients never treated with ethionamide or thiosemicarbazone. *Am. Rev. Respir. Dis.* **90**: 468-469.

Honoré, N., Bergh, S., Chanteau, S., Doucet-Populaire, F., Eiglmeier, K., Garnier, T., Georges, C., Launois, P., Limpaiboon, T., Newton, S., Niang, K., del Portillo, P., Ramesh, G. R., Reddi, P., Ridel, P. R., Sittisombut, N., Wu-Hunter, S., and Cole, S. T. 1993. Nucleotide sequence of the first cosmid from the *Mycobacterium leprae* genome project: structure and function of the Rif-Str regions. *Mol. Microbiol.* 7: 207-214.

Honoré, N., and Cole, S. T. 1993. Molecular basis of rifampin resistance in *Mycobacterium leprae. Antimicrob. Agents Chemother.* **37**: 414-418.

Honoré, N., and Cole, S. T. 1994. Streptomycin resistance in mycobacteria. *Antimicrob. Agents Chemother.* **38**: 238-242.

Howard-Flanders, P., Boyce, R. P., and Theriot, L. 1966. Three loci in *Escherichia coli* K-12 that control the excision of pyrimidine dimers and certain other mutagen products from DNA. *Genetics* **53**: 1119-1136.

Howard-Flanders, P., West, S. C., and Stasiak, A. J. 1984. Role of RecA spiral filaments in genetic recombination. *Nature (London)* **309**: 215-220.

Husain, I., Sancar, G. B., Holbrook, S. R., and Sancar, A. 1987. Mechanism of damage recognition by *Escherichia coli* DNA photolyase. *J. Biol. Chem.* **262**: 13188-13197.

Imaeda, 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. *J. Bacteriol.* **150**: 414-417.

Imlay, J. A., Chin, S. M., and Linn, S. 1988. Toxic DNA damage by hydrogen peroxide through the Fenton reaction *in vivo* and *in vitro*. *Science* **240**: 640-642.

Imlay, J. A., and Linn, S. 1987. Mutagenesis and stress responses induced in *Escherichia coli* by hydrogen peroxide. *J. Bacteriol.* **169**: 2967-2976.

Imlay, J. A., and Linn, S. 1988. DNA damage and oxygen radical toxicity. *Science* **240**: 1302-1309.

Inderlied, C. B. 1991. Antimycobacterial agents: *in-vitro* susceptibility testing, spectrums of activity, mechanisms of action and resistance, and assays for activity in biological fluids. In: V. Lorian (ed). *Antibiotics in Laboratory Medicine*. p. 134-197. Williams and Wilkins, Baltimore.

Inderlied, C. B., Kemper, C. A., and Bermudez, L. E. M. 1993. The *Mycobacterium avium* complex. *Clin. Microbiol. Rev.* **6**: 266-310.

Ishihama, A. 1993. Protein-protein communication within the transcription apparatus. *J. Bacteriol.* **175**: 2483-2489.

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

Ivanova, A., Miller, C., Glinsky, G., and Eisenstark, A. 1994. Role of *rpoS* (*katF*) in *oxyR*-independent regulation of hydroperoxidase I in *Escherichia coli*. *Mol. Microbiol.* **12**: 571-578.

Iyer, G. Y. N., Islam, M. F., and Quastel, J. H. 1961. Biochemical aspects of phagocytosis. *Nature (London)* **192**: 535-541.

Iyer, V. N., and Szybalski, W. 1963. A molecular mechanism of mitomycin action: linking of complementary DNA strands. *Proc. Natl. Acad. Sci. USA* **50**: 355-362.

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*. *J. Gen. Microbiol.* **104**: 37-45.

Jackett, P. S., Aber, V. R., Mitchison, R. D., and Lowrie, D. B. 1981a. The contribution of hydrogen peroxide resistance to the virulence of *Mycobacterium tuberculosis* during the first six days after intravenous infection of normal and BCG-vaccinated guinea pigs. *Brit. J. Exp. Pathol.* **62**: 34-40.

Jackett, P. S., Andrew, P. W., Aber, V. R., and Lowrie, D. B. 1981b. Hydrogen peroxide and superoxide release by alveolar macrophages from normal and BCG-vaccinated guinea pigs after intravenous challenge with *Mycobacterium tuberculosis*. *Brit. J. Exp. Pathol.* **62**: 419-428.

Jacobs, W. R., Docherty, M. A., Curtiss III, R., and Clarke-Curtiss, J. E. 1986. Expression of *Mycobacterium leprae* genes from a *Streptococcus mutans* promoter in *Escherichia coli* K-12. *Proc. Natl. Acad. Sci. USA* **83**: 1926-1930.

Jacobs, W. R., Jr., 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. *Meth. Enzymol.* **204**: 537-555.

Jacobs, W. R., Jr., Snapper, S. B., Tuckman, M., and Bloom, B. R. 1989. Mycobacteriophage vector systems. *Rev. Infect. Dis.* **11**: S404-S410.

Jacobs, W. R., Jr., Tuckman, M., and Bloom, B. R. 1987. Introduction of foreign DNA into mycobacteria using a shuttle phasmid. *Nature (London)* **327**: 532-536.

Jacques, N., and Dreyfus, M. 1990. Translation initiation in *Escherichia coli*: old and new questions. *Mol. Microbiol.* **4**: 1063-1067.

Janssen, G. R. 1993. Eubacterial, archaebacterial, and eucaryotic genes that encode leaderless mRNA. In: R. H. Baltz, G. D. Hegeman, and P. L. Skatrud (eds). *Industrial Microorganisms: Basic and Applied Molecular Genetics.* ASM Press, Washington D. C.

Janssen, G. R., and Bibb, M. J. 1990. Tandem promoters *tsrp*1 and *tsrp*2 direct transcription of the thiostrepton resistance gene (*tsr*) of *Streptomyces azureus*: transcriptional iniation from *tsrp*2 occurs after deletion of the -35 region. *Mol. Gen. Genet.* **221**: 339-346.

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

Jenkins, D. E., Chaisson, S. A., and Matin, A. 1990. Starvation-induced cross protection against osmotic challenge in *Escherichia coli*. *J. Bacteriol.* **172**: 2779-2781.

Jenkins, D. E., Schultz, J. E., and Matin, A. 1988. Starvation-induced cross protection against heat or hydrogen peroxide challenge in *Escherichia coli*. *J. Bacteriol.* **170**: 3910-3914.

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

Jin, D. J., and Gross, C. A. 1988. Mapping and sequencing of mutations in the *Escherichia coli rpoB* gene that lead to rifampicin resistance. *J. Mol. Biol.* **202**: 45-58.

Johnsson, K., Froland, W. A., and Schultz, P. G. 1997. Overexpression, purification, and characterization of the catalase-peroxidase KatG from *Mycobacterium tuberculosis*. *J. Biol. Chem.* **272**: 2834-2840.

Johnsson, K., King, D. S., and Schulz, P. G. 1995. Studies on the mechanism of action of isoniazid and ethionamide in the chemotherapy of tuberculosis. *J. Am. Chem. Soc.* **117**: 5009-5010.

Johnsson, K., and Schultz, P. G. 1994. Mechanistic studies of the oxidation of isoniazid by the catalase peroxidase from *Mycobacterium tuberculosis*. *J. Am. Chem. Soc.* **116**: 7425-7426.

Jones, D. T., and Woods, D. R. 1981. Effect of oxygen on liquid holding recovery of *Bacteroides fragilis*. *J. Bacteriol.* **145**: 1-7.

Jorns, M. S., Sancar, G. B., and Sancar, A. 1984. Identification of a neutral flavin radical and characterization of a second chromophore in *Escherichia coli* DNA photolyase. *Biochemistry* **23**: 2673-2679.

Kagaya, K., Miyakawa, Y., Watanabe, K., and Fukazawa, Y. 1992. Antigenic role of stress-induced catalase of *Salmonella typhimurium* in cell-mediated immunity. *Infect. Immun.* **60**: 1820-1825.

Kalpana, G. V., Bloom, B. R., and Jacobs, W. R., Jr. 1991. Insertional mutagenesis and illegitimate recombination in mycobacteria. *Proc. Natl. Acad. Sci. USA* **88**: 5433-5437.

Kanemori, M., Mori, H., and Yura, T. 1994. Induction of heat shock proteins by abnormal proteins results from stabilization and not increased synthesis of sigma 32 in *Escherichia coli*. *J. Bacteriol.* **176**: 5648-5653.

Kapur, V., Li, L. L., Hamrick, M. R., Plikaytis, B. B., Shinnick, T. M., Telenti, A., Jacobs, W. R., Banerjee, A., Cole, S. T., Yuen, K. Y., Clarridge, J. E., Kreiswirth, B. N., and Musser, J. M. 1995. Rapid *Mycobacterium* species assignment and unambiguous identification of mutations associated with antimicrobial resistance in *Mycobacterium tuberculosis* by automated DNA sequencing. *Arch. Pathol. Lab. Med.* **119**: 131-138.

Kapur, V., Li, L. L., Iordanescu, S., Hamrick, M. R., Wanger, A., Kreiswirth, B. N., and Musser, J. M. 1994. Characterization by automated DNA sequencing of mutations in the gene (*rpoB*) encoding the RNA polymerase β-subunit in rifampin-resistant *Mycobacterium tuberculosis* strains from New York City and Texas. *J. Clin. Microbiol.* **32**: 1095-1098.

Keilty, S., and Rosenberg, M. 1987. Constitutive function of a positively regulated promoter reveals new sequences essential for activity. *J. Biol. Chem.* **262**: 6389-6395.

Kellerman, O. K., and Ferenci, T. 1982. Maltose binding protein from *E. coli*. *Meth. Enzymol.* **90**: 459-463.

Kenney, T. J., and Churchward, G. 1994. Cloning and sequence analysis of the *rpsL* and *rpsG* genes of *Mycobacterium smegmatis* and characterization of mutations causing resistance to streptomycin. *J. Bacteriol.* **176**: 6153-6156.

Kenney, T. J., and Churchward, G. 1996. Genetic analysis of the *Mycobacterium smegmatis rpsL* promoter. *J. Bacteriol.* **178**: 3564-3571.

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

Kim, Y. K., and Kim, K. S. 1995. Multiple-binding proteins that can specifically interact with the OxyR-regulated promoter of the *katG* catalase gene in *Escherichia coli*. *Molecules and Cells* **5**: 119-125.

Klebanoff, S. J. 1980. In: R. Van Furth (ed). *Monuclear Phagocytes, Functional Aspects.* Nijhoff, Boston.

Koch, R. 1932. Die aetiologie der tuberculose. *Am. Rev. Tuberc.* **25**: 285-323. (Translated from the original 1882 article by Berna Pinner and Max Pinner.)

Kochi, A. 1991. The global tuberculosis situation and the new control strategy of the World Health Organization. *Tubercle* **72**: 1-6.

Kong, T. H., Coates, R. M., Butcher, P. D., Hickman, C. J., and Shinnick, T. M. 1993. *Mycobacterium tuberculosis* expresses two chaperonin-60 homologs. *Proc. Natl. Acad. Sci. USA* **90**: 2608-2612.

Konno, K., Feldman, F. M., and McDermott, W. 1967. Pyrazinamide susceptibilty of amidase activity of tubercle bacilli. *Am. Rev. Respir. Dis.* **95**: 461-469.

Konno, K., Nagayama, H., and Oka, S. 1959. Nicotinamidase in mycobacteria: A method for distinguishing bovine type tubercle bacilli from other mycobacteria. *Nature (London)* **184**: 1743-1744.

Konno, K., Oizumi, K., Shimizu, Y., Tamagawa, S., and Oka, S. 1966. Niacin metabolism in mycobacteria. Mechanism of excess niacin production by human tubercle bacilli. *Am. Rev. Respir. Dis.* **93**: 41-46.

Kornberg, A., and Baker, T. 1992. DNA Replication. W. H. Freeman and Co., New York.

Kremer, L., Baulard, A., Estaquier, J., Content, J., Capron, A., and Locht, C. 1995. . Analysis of the *Mycobacterium tuberculosis* 85A antigen promoter region. *J. Bacteriol.* **177**: 642-653.

Krishna-Murti, C. R. 1975. Isonicotinic acid hydrazide. In: J. W. Corcoran and F. A. Hahn (eds). *Antibiotics. Vol III. Mechansim of Action of Antimicrobial and Antitumor Agents.* Springer-Verlag, New York.

Krüger-Thiemer, E. 1958. Isonicotinic acid hypothesis of the antituberculous action of isoniazid. *Am. Rev. Tuberc.* 77: 364-367.

Kubica, G. P., Jones, W. D., Jr., Abbott, V. D., Beam, R. E., Kilburn, J. O., and Cater, J. C., Jr. 1966. Differential identification of mycobacteria.- I. Tests on catalase activity. *Am. Rev. Respir. Dis.* **94**: 400-405.

Kubica, G. P., and Pool, G. L. 1960. Studies on the catalase activity of acid-fast bacilli: I. An attempt to subgroup these organisms on the basis of their catalase activities at different temperatures and pH. *Am. Rev. Respir. Dis.* **81**: 387.

Kullberg, B. J., Vogels, M. T. E., and Vandermeer, J. W. M. 1994. Immunomodulators in bacterial and fungal infections- a review of their therapeutic value. *Clin. Immunother.* **1**: 43-55.

Kullik, I., Stevens, J., Toledano, M. B., and Storz, G. 1995b. Mutational analysis of redox-sensitive transcriptional regulator OxyR: regions important for DNA binding and multimerization. *J. Bacteriol.* **177**: 1285-1291.

Kullik, I., Toledano, M. B., Tartaglia, L. A., and Storz, G. 1995a. Mutational analysis of redox-sensitive transcriptional regulator OxyR: regions important for oxidation and transcriptional activation. *J. Bacteriol.* **177**: 1275-1284.

Kumar, S., Ganguly, N. K., and Kohli, K. K. 1994. Effect of isoniazid on glutathione biosynthesis and degradation in *Mycobacterium smegmatis. Folia Microbiol.* **39**: 571-575.

Kumar, A., Malloch, R. A., Fujita, N., Smillie, D. A., Ishihama, A., and Hayward, R. S. 1993. The minus 35-recognition sequence of *Escherichia coli* sigma 70 is inessential for initiation of transcription at an "extended minus 10" promoter. *J. Mol. Biol.* **232**: 406-418.

Kumar, R. A., Vaze, M. B., Chandra, N. R., Vijayan, M., and Muniyappa, K. 1996. Functional characterization of the precursor and spliced forms of RecA protein of *Mycobacterium tuberculosis. Biochemistry* **35**: 1793-1802.

Kushner, S. R., Maples, V. F., Easton, A., Farrance, I., and Peramachi, P. 1983. Physical, biochemical, and genetic characterization of the *uvrD* gene. In: E. C. Friedberg and B. A. Bridges (eds). *Cellular Responses to DNA Damage.* p. 153-159. Alan R. Liss, Inc., New York.

Kwon, H. H., Tomioka, H., and Saito, H. 1995. Distribution and characterization of β-lactamases of mycobacteria and related organisms. *Tubercle Lung Dis.* **76**: 141-148.

Labidi, A., David, H. L., and Roulland-Dussoix, D. 1985a. Restriction endonuclease mapping and cloning of *Mycobacterium fortuitum* var. *fortuitum* plasmid pAL5000. *Ann. Inst. Pasteur Microbiol.* **136**: 209-215.

Labidi, A., David, H. L., and Roulland-Dussoix, D. 1985b. Cloning and expression of mycobacterial plasmid DNA in *Escherichia coli. FEMS Microbiol. Lett.* **30**: 221-225.

Laemmli U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**: 680-685.

Lang, G., Spartalian, K., and Yonetani, T. 1976. Mossbauer spectroscopic study of compound ES of cytochrome c peroxidase. *Biochim. Biophys. Acta* **451**: 250.

Lee, M. H., Pascopella, L., Jacobs, W. R., Jr., and Hatfull, G. F. 1991. Site-specific integration of mycobacteriophage L5: integration-proficient vectors for *Mycobacterium smegmatis, Mycobacterium tuberculosis*, and bacille Calmette-Guerin. *Proc. Natl. Acad. Sci. USA* **88**: 3111-3115.

Legocki, R. P., Legocki, M., Baldwin, T. O., and Szalay, A. A. 1986. Bioluminescence in soybean root nodules: Demonstration of a general approach to assay gene expression *in vivo* by using bacterial luciferase. *Proc. Natl. Acad. Sci. USA* **83**: 9080-9084.

Leven, S., Heimberger, A., and Eisenstark, A. 1990. Catalase HPI influences membrane permeability in *Escherichia coli* following near-UV stress. *Biochem. Biophys. Res. Commun.* **171**: 1224-1228.

Levin, M. E., and Hatfull, G. F. 1993. *Mycobacterium smegmatis* RNA polymerase: DNA supercoiling, action of rifampicin and mechanism of rifampicin resistance. *Mol. Microbiol.* **8**: 277-285.

Levine, S. A. 1977. Isolation and characterization of catalase deficient mutants of *Salmonella typhimurium*. *Mol. Gen. Genet.* **150**: 205-209.

Li, M., Moyle, H. and Susskind, M. M. 1994. Target of the transcriptional activation function of phage lambda cI protein. *Science* **263**: 75-77.

Lichenstein, H. 1928. *Travels in Africa, 1803, 1804, 1805, 1806.* (Reprint of translation by A. Plumptre from the original German, The Van Rierberck Society, Cape Town, South Africa).

Lim, E. M., Rauzier, J., Timm, J., Torrea, G., Murray, A., Giquel, B., and Portnoi, P. 1995. Identification of *Mycobacterium tuberculosis* DNA sequences encoding exported proteins by using *phoA* gene fusions. *J. Bacteriol.* **177**: 59-65.

Lin, J. J., and Sancar, A. 1989. A new mechanism for repairing oxidative damage to DNA: (A)BC excinuclease removes AP sites and thymine glycols from DNA. *Biochemistry* **28**: 7979-7984.

Lin, J. J., and Sancar, A. 1992. Active site of (A)BC exinuclease. I. Evidence for 5' incision by UvrC through a catalytic site involving Asp399, Asp438, Asp466, and His538 residues. *J. Biol. Chem.* **267**: 17688-17692.

Lindahl, T. 1979. DNA glycosylases, endonucleases for apurinic/apyrimidinic sites, and base excision repair. *Prog. Nucleic Acids Res. Mol. Biol.* **22**: 135-192.

Lisser, S., and Margalit, H. 1994. Determination of common structural features in *Escherichia coli* promoters by computer analysis. *Eur. J. Biochem.* **223**: 823-830.

Little, J. W., and Mount, D. W. 1982. The SOS regulatory system of *Escherichia coli. Cell* **29**: 11-22.

Litwin, C. M., Boyko, S. A., and Calderwood, S. B. 1992. Cloning, sequencing, and transcriptional regulation of the *Vibrio cholerae fur* gene. *J. Bacteriol.* **174**: 1897-1903.

Litwin, C. M., and Calderwood, S. B. 1993. Role of iron in regulation of virulence genes. *Clin. Microbiol. Rev.* **6**: 137-149.

Ljungquist, S., Lindahl, T., and Howard-Flanders, P. 1976. Methyl methane sulfonate-sensitive mutant of *Escherichia coli* deficient in an endonuclease specific for apurinic sites in deoxyribonucleic acid. *J. Bacteriol.* **126**: 646-653.

Loewen, P. C. 1984. Isolation of catalase-deficient *Escherichia coli* mutants and genetic mapping of *katE*, a locus that affects catalase activity. *J. Bacteriol.* **157**: 622-626.

Loewen, P. C. 1992. Regulation of bacterial catalase synthesis. In: J. G. Scandalios (ed). *Molecular Biology of Free Radical Scavenging Systems.* pp. 97-115. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Loewen, P. C., and Stauffer, G. V. 1990. Nucleotide sequence of *katG* of *Salmonella typhimurium* LT2 and characterization of its product, hydroperoxidase I. *Mol. Gen. Genet.* **224**: 147-151.

Loewen, P. C., and Switala, J. 1986. Purification and characterization of catalase HPII from *Escherichia coli* K-12. *Biochem. Cell Biol.* **64**: 638-646.

Loewen, P. C., Switala, J., Smolenski, M., and Triggs-Raine, B. L. 1990. Molecular characterization of three mutations in *katG* affecting the activity of hydroperoxidase I of *Escherichia coli*. *Biochem. Cell Biol.* **68**:1037-1044.

Loewen, P. C., Switala, J., and Triggs-Raine, B. L. 1985. Catalases HPI and HPII in *Escherichia coli* are induced independently. *Arch. Biochem. Biophys.* **243**: 144-149.

Loprasert, S., Negoro, S., and Okada, H. 1988. Thermostable peroxidase from *Bacillus stearothermophilus*. *J. Gen. Microbiol.* **134**:1971-1976.

Loprasert, S., Negoro, S., and Okada, H. 1989. Cloning, nucleotide sequence, and expression in *Escherichia coli* of the *Bacillus stearothermophilus* peroxidase gene (*perA*). *J. Bacteriol.* **171**: 4871-4875.

Love, J. D., Nguyen, H. Y., Or, A., Attri, A. K., and Minton, K. W. 1986. UV-induced interstrand crosslinking of d(GT)n.d(CA)n is facilitated by a structural transition. *J. Biol. Chem.* **261**: 10051-10057.

Lowrie, D. B. 1983. How macrophages kill tubercle bacilli. *J. Med. Microbiol.* **16**: 1-12.

Lowrie, D. B., Tascon, R. E., Colston, M. J., and Silva, C. L. 1994. Towards a DNA vaccine against tuberculosis. *Vaccine* **12**: 1537-1540.

Ma, M., and Eaton, J. W. 1992. Multicellular oxidant defense in unicellular organisms. *Proc. Natl. Acad. Sci. USA* **89**: 7924-7928.

Mackaness, G. B. 1956. The intracellular activation of pyrazinamide and nicotinamide. *Am. Rev. Tuberc.* **74**: 718-728.

Madhusudan, K., Ramesh, V., and Nagaraja, V. 1994. Molecular cloning of *gyrA* and *gyrB* genes of *Mycobacterium tuberculosis*: analysis of nucleotide sequence. *Biochem. Mol. Biol. Intl.* **33**: 651-660.

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*. *J. Gen. Microbiol.* **139**: 575-583.

Maina, C. V., Riggs, P. D., Grandea A. G. III., Slatko, B. E., Moran, L. S., Tagliamonte, J. A., McReynolds, L. A., and Guan, C. 1988. A vector to express and purify foreign proteins in *Escherichia coli* by fusion to, and separation from, maltose binding protein. *Gene* **74**: 365-373.

Mandell, G. L. 1975. Catalase, superoxide dismutase, and virulence of *Staphylococcus aureus*: *in vitro* and *in vivo* studies with emphasis on staphylococcal-leukocyte interaction. *J. Clin. Invest.* **55**: 561-566.

Manoil, C., and Beckwith, J. 1985. TnphoA: a transposon probe for protein export signals. *Proc. Natl. Acad. Sci. USA* **82**: 8129-8133.

Marcinkeviciene, J. A., Magliozzo, R. S., and Blanchard, J. S. 1995. Purification and characterization of the *Mycobacterium smegmatis* catalase-peroxidase involved in isoniazid activation. *J. Biol. Chem.* **270**: 22290-22295.

Mark, S., Richardson, H., Higgins, C. F., and Lilley, D. M. 1988. DNA supercoiling and the *leu*-500 promoter mutation of *Salmonella typhimurium*. *EMBO J.* **7**: 1863-1869.

Masker, W., Hanawalt, P. C., and Shizuya, H. 1973. Role of DNA polymerase II in repair replication in *Escherichia coli*. *Nature New Biol.* **244**: 242-243.

Matsuo, K., Yamaguchi, R., Yamazaki, A., Tasaka, H., Terasaka, K., Totsuka, M., Kobayashi, K., Yukitake, H., and Yamada, T. 1990. Establishment of a foreign antigen secretion system in mycobacteria. *Infect. Immun.* **58**: 4049-4054.

McAllister, C. F., and Achberger, E. C. 1989. Rotational orientation of upstream curved DNA affects promoter function in *B. subtilis*. *J. Biol. Chem.* **264**: 10451-10456.

McClatchy, J. K. 1971. Mechanism of action of INH on *M. bovis*, strain BCG. *Infect. Immun.* **3**: 530.

McClure, W. R. 1985. Mechanism and control of transcription initiation in prokaryotes. *Annu. Rev. Biochem.* **54**: 171-204.

McClure, W. R., and Cech, C. L. 1978. On the mechanisms of rifampicin inhibition of RNA synthesis. *J. Biol. Chem.* **253**: 8949-8956.

McCord, J. M., and Fridovich, I. 1988. Superoxide dismutase: the first twenty years (1968-1988). *Free Radic. Biol. Med.* **5**: 363-369.

McDermott, W., and Tompsett, R. 1954. Activation of pyrazinamide and nicotinamide in acid environments *in vitro*. *Am. Rev. Tuberc.* **70**: 748.

McKeown, T. 1979. *The role of medicine: Dream, Mirage or Nemesis?* Oxford: Basil Blackwell.

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. *Res. Microbiol.* **142**: 451-463.

McVicar, N. 1932. The prevalence of certain diseases amongst the natives of the Ciskei. *S. A. Med. J.* **6**: 721.

Mdluli, K. E., Sherman, D. R., Hickey, M. J., Kreiswirth, B. N., Morris, S. L., Stover, C. K., and Barry, C. E. III. 1996. Biochemical and genetic data suggest that InhA is not the primary target for activated isoniazid in *Mycobacterium tuberculosis*. *J. Infect. Dis.* **174**: 1085-1090.

Medical Research Council. 1972. BCG vaccine and the vaccine prepared with BRC for the prevention of tuberculosis in adolescents and young adults. 4th report to the Medical Research Council by its Tuberculosis Vaccines Clinical Trials Committee. *Bull. WHO.* **46**: 371-385.

Meier, A., Kirschner, P., Bange, F. C., Vogel, U., and Böttger, E. C. 1994. Genetic alterations in streptomycin-resistant *Mycobacterium tuberculosis*- mapping of mutations conferring resistance. *Antimicrob. Agents Chemother.* **38**: 228-233.

Meighen, E. A. 1988. Enzymes and genes from the lux operons of bioluminescent bacteria. *Annu. Rev. Microbiol.* **42**: 151-176.

Meir, E., and Yagil, E. 1984. Catalase-negative mutants of *Escherichia coli*. *Curr. Microbiol.* **11**: 13-18.

Meissner, P. S., Sisk, W. P., and Bergman, M. L. 1987. Bacteriophage λ cloning system for the construction of directional cDNA libraries. *Proc. Natl. Acad. Sci. USA* **84**: 4171.

Mekalanos, J. J. 1992. Environmental signals controlling expression of virulence determinants in bacteria. *J. Bacteriol.* **174**: 1-7.

Menzel, R., and Gellert, M. 1983. Regulation of the genes for *E. coli* DNA gyrase: homeostatic control of DNA supercoiling. *Cell* **34**: 105-113.

Messing, J. 1983. New M13 vectors for cloning. *Meth. Enzymol.* **101**: 20-78.

Meyer, H., and Malley, J. 1912. Hydrazine derivatives of pyridine carboxylic acids. *Monatsh.* **33**: 393.

Middlebrook, G. 1952. Sterilization of tubercle bacilli by isonicotinic acid hydrazide and the incidence of variants resistant to the drug *in vitro*. *Am. Rev. Tuberc.* **65**: 765.

Middlebrook, G. 1954. Isoniazid resistance and catalase activity of tubercle bacilli: A preliminary report. *Am. Rev. Tuberc.* **69**: 471-472.

Middlebrook, G., and Cohn, M. L. 1953. Some observations on the pathogenicity of isoniazid-resistant variants of tubercle bacilli. *Science* **118**: 297-299.

Milano, A., De Rossi, E., Gusberti, L., Heym, B., Marone, P., and Riccardi, G. 1996. The *katE* gene, which encodes the catalase HPII of *Mycobacterium avium*. *Mol. Microbiol.* **19**: 113-123.

Milcarek, C., and Weiss, B. 1972. Mutants of *E. coli* with altered deoxyribonucleases. 1. Isolation and characterization of mutants for exonuclease III. *J. Mol. Biol.* **68**: 303-318.

Miller, L. P., Crawford, J. T., and Shinnick, T. M. 1994. The *rpoB* gene of *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **38**: 805-811.

Miller, O. L., Jr., Hamkalo, B. A., and Thomas, C. A., Jr. 1970. Visualization of bacterial genes in action. *Science* **169**: 392-395.

Minnikin, D. E. 1982. In: C. Ratledge and J. Stanford (eds). *The Biology of the Mycobacteria.* pp. 95-184. Academic Press, San Diego.

Minton, N. P., Oultram, J. D., Brehm, J. K., and Atkinson, T. 1988. The replication proteins of plasmids pE194 and pLS1 have N-terminal homology. *Nucleic Acids Res.* **16**: 3101.

Mitchison, D. A., Selkon, J. B., and Lloyd, J. 1963. Virulence in guinea pig, susceptibility to hydrogen peroxide, and catalase activity of isoniazid-sensitive tubercle bacilli from South Indian and British patients. *J. Pathol. Bacteriol.* **86**: 377-386.

Mitchison, D. A., Wallace, J. G., Bhatia, A. L., Selkon, J. B., Subbaiah, J., and Lancaster, M. C. 1960. A comparison of the virulence in guinea pigs of South India and British tubercle bacilli. *Tubercle* **41**: 1-22.

Moazed, D., and Noller, H. F. 1987. Interaction of antibiotics with functional sites in 16S ribosomal RNA. *Nature (London)* **327**: 389-394.

Mongkolsuk, S., Loprasert, S., Vattanaviboon, P., Chanvanichayachai, C., Chamnongpol, S., and Supsamran, N. 1996. Heterologous growth phase- and temperature-dependent expression and H_2O_2 toxicity protection of a superoxide-inducible monofunctional catalase gene from *Xanthomonas oryzae* pv. *oryzae*. *J. Bacteriol.* **178**: 3578-3584.

Moran, C. P., Lang, N., LeGrice, S. F. J., Lee, G., Stephens, M., Sonenshein, A. L., Pero, J., and Losick, R. 1982. Nucleotide sequences that signal the initiation of transcription and translation in *Bacillus subtilis*. *Mol. Gen. Genet.* **186**: 339-346.

Morgan, R. W., Christman, M. F., Jacobson, F. S., Storz, G., and Ames, B. N. 1986. Hydrogen-peroxide-inducible proteins in *Salmonella typhimurium* overlap with heat-shock and other stress proteins. *Proc. Natl. Acad. Sci. USA* **83**: 8059-8063.

Morris, S. L., Nair, J., and Rouse, D. A. 1992. The catalase-peroxidase of *Mycobacterium intracellulare*: nucleotide sequence analysis and expression in *Escherichia coli*. *J. Gen. Microbiol.* **138**: 2363-2370.

Morse, D. 1961. Prehistoric tuberculosis in America. *Am. Rev. Respir. Dis.* **83**: 489-504.

Moss, M. T. 1987. Cloning and expression of mycobacterial genes in *Escherichia coli.* PhD thesis, University of Surrey.

Movahedzadeh, F., Colston, M. J., and Davis, E. O. 1997a. Characterization of *Mycobacterium tuberculosis* LexA: Recognition of a Cheo (*Bacillus*-type SOS) box. *Microbiology* **143**: 929-936.

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

Mukhopadhyay, S., and Schellhorn, H. E. 1994. Induction of *Escherichia coli* hydroperoxidase I by acetate and other weak acids *J Bacteriol* **176**: 2300-2307.

Mulvey, M. R., Sorby, P. A., Triggs-Raine, B. L., and Loewen, P. C. 1988. Cloning and physical characterization of *katE* and *katF* required for catalase HPII expression in *Escherichia coli. Gene* **73**: 337-345.

Mulvey, M. R., Switala, J., Borys, A., and Loewen, P. C. 1990. Regulation of transcription of *katE* and *katF* in *Escherichia coli. J. Bacteriol.* **172**: 6713-6720.

Murohashi, T., and Yoshida, K. 1965. Effect of ultraviolet irradiation on the acid-fastness and gram-positivity of mycobacteria. *Am. Rev. Respir. Dis.* **92**: 817.

Murohashi, T., and Yoshida, K. 1968. Effect of ultraviolet irradiation on the acid-fastness of drug-resistant mutants of tubercle bacilli, with special reference to the virulence of isoniazid-resistant strains. *Am. Rev. Respir. Dis.* **97**: 283-291.

Murray, C. J. L., Styblo, K., and Rouillon, A. 1990. Tuberculosis in developing countries: burden, intervention and cost. *Bull. Int. Union Tuberc.* **65**: 6-24.

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* β-galactosidase in *Mycobacterium bovis* BCG using an expression system isolated from *Mycobacterium paratuberculosis* which induced humoral and cellular immune responses. *Mol. Microbiol.* **6**: 3331-3342.

Musser, J. M. 1995. Antimicrobial agent resistance in mycobacteria: molecular genetic insights. *Clin. Microbiol. Rev.* **8**: 496-514.

Mutsuda, M., Ishikawa, T., Takeda, T., and Shigeoka, S. 1996. The catalase-peroxidase of *Synechococcus* PCC 7942: Purification, nucleotide sequence analysis and expression in *Escherichia coli. Biochem. J.* **316**: 251-257.

Nadler, V., Goldberg, I., and Hochman, A. 1986. Comparative study of bacterial catalases. *Biochim. Biophys. Acta* **882**: 234-241.

Nagai, K., and Thogersen, H. C. 1984. Generation of β-globin by sequence-specific proteolysis of a hybrid protein produced in *Escherichia coli. Nature (London)* **309**: 810-812.

Nagai, K., and Thogersen, H. C. 1987. Synthesis and sequence-specific proteolysis of hybrid proteins produced in *Escherichia coli. Meth. Enzymol.* **153**: 461-481.

Nagy, J. M., Cass, A. E. G., and Brown, K. A. 1995. Progress in the characterization of catalase-peroxidase from *Mycobacterium tuberculosis. Biochem. Soc. Transactions* **23**: S152- S152.

Nair, J., Rouse, D. A., and Morris, S. L. 1992. Nucleotide sequence analysis and serological characterization of the *Mycobacterium intracellulare* homologue of the *Mycobacterium tuberculosis* 19 kDa antigen. *Mol. Microbiol.* **6**: 1431-1439.

Nair, J., Rouse, D. A., and Morris, S. L. 1993. Nucleotide sequence analysis and serologic characterization of a 27-kilodalton *Mycobacterium intracellulare* lipoprotein. *Infect. Immun.* **61**: 1074-1081.

Nair, S. 1993. Identification and molecular characterization of a novel RecA from *Mycobacterium tuberculosis*. PhD thesis, University of Cape Town, South Africa.

Nair, S., and Steyn, L. M. 1991. Cloning and expression in *Escherichia coli* of a *recA* homologue from *Mycobacterium tuberculosis*. *J. Gen. Microbiol.* **137**: 2409-2414.

Nakata, N., Matsuoka, M., Kashiwabara, Y., Okada, N., and Sasakawa, C. 1997. Nucleotide sequence of the *Mycobacterium leprae katG* region. *J. Bacteriol.* **179**: 3053-3057.

Nakayama, N., Fujita, N., Ohama, T., Osawa, S., and Ishihama, A. 1989. *Micrococcus luteus*, a bacterium with a high genomic G+C content, contains *Escherichia coli*-type promoters. *Mol. Gen. Genet.* **218**: 384-389.

Narayanan Nair, C., MacKay-Scollay, E. M., Ramachandran, K., Selkon, J. B., Tripathy, S. P., Mitchison, D. A., and Dickinson, J. M. 1964. Virulence in the guinea pig and susceptibility to hydrogen peroxide of isoniazid-sensitive tubercle bacilli from South Indian patients. *Tubercle* **45**: 345-353.

Nesbit, C. E., Levin, M. E., Donnelly-Wu, M. K., and Hatfull, G. F. 1995. Transcriptional regulation of repressor synthesis in mycobacteriophage L5. *Mol. Microbiol.* **17**: 1045-1056.

Nickerson, C. A., and Achberger, E. C. 1995. Role of curved DNA in binding of *Escherichia coli* RNA polymerase to promoters. *J. Bacteriol.* **177**: 5756-5761.

Niederhoffer, E. C., Naranjo, C. M., Bradley, K. L., and, Fee, J. L. 1990. Control of *Escherichia coli* superoxide dismutase (*sodA* and *sodB*) genes by the ferric uptake regulation (*fur*) locus. *J. Bacteriol.* **172**: 1930-1938.

O'Brien, S., Jackett, P. S., Lowrie, D. B., and Andrew, P. W. 1991. Guinea pig alveolar macrophages kill *Mycobacterium tuberculosis in vitro*, but killing is independent of susceptibility to hydrogen peroxide or triggering of the respiratory burst. *Microb. Pathog.* **10**: 199-207.

O'Connor, S. P., Rumschlag, H. S. and Mayer, L. W. 1990. Nucleotide sequence of the gene encoding the 35-kDa protein of *Mycobacterium tuberculosis*. *Res. Microbiol.* **141**: 407-423.

O'Donnell, M. A., Aldovini, A., Duda, R. B., Yang, B., Szilvasi, A., Young, R. A., and Dewolf, W. C. 1994. Recombinant *Mycobacterium bovis* BCG secreting functional interleukin-2 enhances gamma interferon production by splenocytes. *Infect. Immun.* **62**: 2508-2514.

Offe, H. A., Siefkin, W., and Domagk, G. 1952. The tuberculostatic activity of hydrazine derivatives from pyridine carboxylic acids and carbonyl compounds. *Z. Naturforsch.* **76**: 462.

Ogasawara, N., Moriya, S., and Yoshikawa, H. 1983. Structure and organisation of rRNA operons in the region of the replication origin of the *Bacillus subtilis* chromosome. *Nucleic Acids Res.* **11**: 6301-6318.

Oguiza, J. A., Tao, X., Marcos, A. T., Martin, J. F., and Murphy, J. R. 1995. Molecular cloning, DNA sequence analysis, and characterization of the *Corynebacterium diphtheriae dtxR* homolog from *Brevibacterium lactofermentum*. *J. Bacteriol.* **177**: 465-467.

Oh, E. Y., and Grossman, L. 1986. The effect of *Escherichia coli* Uvr protein binding on the topology of supercoiled DNA. *Nucleic Acids Res.* **14**: 8557-8571.

Ohama, T. F., Ysamao, F., Muto, A., and Osawa, S. 1987. Organisation and codon usage of the streptomycin operon in *Micrococcus luteus*, a bacterium with a high G+C content. *J. Bacteriol.* **169**: 4770-4777.

Olivera, B. M. 1971. The DNA joining enzyme from *Escherichia coli*. *Meth. Enzymol.* **21**: 311.

Olsson, O., Konez, C., and Szalay, A. A. 1988. The use of the *luxA* gene of the bacterial luciferase operon as a reporter gene. *Mol. Gen. Genet.* **215**: 1-9.

Ordway, D. J., Sonnenberg, M. G., Donahue, S. A., Belisle, J. T., and Orme, I. M. 1995. Drug-resistant strains of *Mycobacterium tuberculosis* exhibit a range of virulence for mice. *Infect. Immun.* **63**: 741-743.

Orren, D. K., and Sancar, A. 1989. The (A)BC exinuclease of *Escherichia coli* has only the UvrB and UvrC subunits in the incision complex. *Proc. Natl. Acad. Sci. USA* **86**: 5237-5241.

Orren, D. K., Selby, C. P., Hearst, J. E., and Sancar, A. 1992. Post-incision steps of nucleotide excision repair in *Escherichia coli*. Disassembly of the UvrBC-DNA complex by helicase II and DNA polymerase I. *J. Biol. Chem.* **267**: 780-788.

O'Sullivan, D. J., Dowling, D. N., de Lorenzo, V., and O'Gara, F. 1994. *Escherichia coli* ferric uptake regulator (Fur) can mediate regulation of a pseudomonad iron-regulated promoter. *FEMS Microbiol. Lett.* **117**: 327-332.

Palmer, C. E., and Long, M. W. 1966. Effects of infection with atypical mycobacteria on BCG vaccination and tuberculosis. *Am. Rev. Respir. Dis.* **94**: 553-568.

Pappenheimer, A. M., Jr. 1977. Diphtheria toxin. *Annu. Rev. Biochem.* **46**: 69-94.

Pfuetze, K. H., Pyle, M. M., Hinshaw, H. C., and Feldman, W. H. 1955. The first clinical trial of streptomycin in human tuberculosis. *Am. Rev. Tuberc.* **71**: 752-754.

Ponnambalam, S., Chan, B., and Busby, S. 1988. Functional analysis of different sequence elements in the *Escherichia coli* galactose operon P2 promoter. *Mol. Microbiol.* **2**: 165-172.

Poulos, T. L., and Kraut, J. 1980. A hypothetical model of the cytochrome *c* peroxidase . cytochrome *c* electron transfer complex. *J. Biol. Chem.* **255**: 10322-10330.

Predich, M., Doukhan, L., Nair, G., and Smith, I. 1995. Characterization of RNA polymerase and two sigma-factor genes from *Mycobacterium smegmatis*. *Mol. Microbiol.* **15**: 355-366.

Prescott, D. J., and Vagelos, P. R. 1972. Acyl carrier protein. *Adv. Enzymol.* **36**: 269-311.

Pretorius, G. S., van Helden, P. D., Sirgel, F., Eisenach, K. D., and Victor, T. C. 1995. Mutations in *katG* gene sequences in isoniazid-resistant clinical isolates of *Mycobacterium tuberculosis* are rare. *Antimicrob. Agents Chemother.* **39**: 2276-2281.

Prioli, R. P., Tanna, A., and Brown, I. N. 1985. Rapid methods for counting mycobacteria: comparison of methods for extraction of mycobacterial adenosine triphosphate (ATP) determined by firefly luciferase assay. *Tubercle* **66**: 99-108.

Pruss, G. J., and Drlica, K. 1985. DNA supercoiling and suppression of the *leu*-500 promoter mutation. *J. Bacteriol.* **164**: 947-949.

Pruss, G. J., and Drlica, K. 1989. DNA supercoiling and prokaryotic transcription. *Cell* **56**: 521-523.

Qin, M., Taniguchi, H., and Mizuguchi, Y. 1994. Analysis of the replication region of a mycobacterial plasmid, pMSC262. *J. Bacteriol.* **176**: 419-425.

Quemard, A., Lacave, C., and Laneelle, G. 1991. Isoniazid inhibition of mycolic acid synthesis by cell extracts of sensitive and resistant strains of *Mycobacterium aurum*. *Antimicrob. Agents Chemother.* **35**: 1035-1039.

Quemard, A., Sacchettini, J. C., Dessen, A., Vilcheze, C., Bittman, R., Jacobs, W. R., and Blanchard, J. S. 1995. Enzymatic characterization of the target for isoniazid in *Mycobacterium tuberculosis*. *Biochemistry* **34**: 8235-8241.

Qui, X., Verlinde, C. L. M. J., Zhang, S., Schmitt, M. P., Holmes, R. K., and Hol, W. G. J. 1995. Three-dimensional structure of the diphtheria toxin repressor in complex with divalent cation co-repressors. *Structure* **3**: 87-100.

Radford, A. J., Wood, P. R., Billman-Jacobe, H., Geysen, H. M., Mason, T. J., and Tribbick, G. 1990. Epitope mapping of the *Mycobacterium bovis* secretory protein MPB70 using overlapping peptide analysis. *J. Gen. Microbiol.* **136**: 265-272.

Radman, M. 1974. Phenomenology of an inducible mutagenic DNA repair pathway in *Escherichia coli*: SOS repair hypothesis. In: L. Prakash, F. Sherman, M. Miller, C. Lawrence, and H. W. Tabor (eds). *Molecular and Environmental Aspects of Mutagenesis*. pp. 128-142. Charles C. Thomas, Springfield, Ill.

Raibaud, O., and Schwartz, M. 1984. Positive control of transcription initiation in bacteria. *Annu. Rev. Genet.* **18**: 173-206.

Ramalingam, R., Blume, J. E., Ganguly, K., and Ennis, H. L. 1995. AT-rich upstream sequence elements regulate spore germination-specific expression of the *Dictyostelium discoideum celA* gene. *Nucleic Acids Res.* **23**: 3018-3025.

Ramesh, G. R., and Gopinathan, K. P. 1995. Cloning and characterization of mycobacteriophage 13 promoters. *Ind. J. Biochem. Biophys.* **32**: 361-367.

Ranes, M. G., Rauzier, J., Lagranderie, M., Gheorghiu, M., and Gicquel, B. 1990. Functional analysis of pAL5000, a plasmid from *Mycobacterium fortuitum*: construction of a "mini" *Mycobacterium-Escherichia coli* shuttle vector. *J. Bacteriol.* **172**: 2793-2797.

Ray, P. N., and Pearson, M. L. 1975. Functional inactivation of bacteriophage lambda morphogenetic gene in RNA. *Nature (London)* **253**: 647-650.

Revel, V., Cambau, E., Jarlier, V., and Sougakoff, W. 1994. Characterization of mutations in *Mycobacterium smegmatis* involved in resistance to fluoroquinolones. *Antimicrob. Agents Chemother.* **38:** 1991-1996.

Reynolds, P. E. 1989. Structure, biochemistry and mechanism of action of glycopeptide antibiotics. *Eur. J. Clin. Microbiol. Infect. Dis.* **8**: 943-950.

Richter, H. E., and Loewen, P. C. 1981. Induction of catalase in *Escherichia coli* by ascorbic acid involves hydrogen peroxide. *Biochem. Biophys. Res. Commun.* **100**: 1039-1046.

Riddell, R. W., Stewart, S. M., and Somner, A. R. 1960. Ethionamide. *Br. Med. J.* **2**: 1207-1208.

Riley, L. W. 1995. Determinants of cell entry and intracellular survival of *Mycobacterium tuberculosis*. *Trends Microbiol.* **3**: 27-30.

Riley, R. L., Mills, C. L., Nyka, W., Weinstock, N., Storey, P. B., Sultan, L. K., Riley, M. C., and Wells, W. F. 1959. Aerial dissemination of pulmonary tuberculosis: a two year study of contagion in a tuberculosis ward. *Am. J. Hyg.* **70**: 185.

Robbi, M., and Lazarow, P. B. 1978. Synthesis of catalase in two cell-free protein-synthesizing systems and in rat liver. *Proc. Natl. Acad. Sci. USA* **75**: 4344-4348.

Robitzek, E. H., and Selikoff, I. J. 1952. Hydrazine derivatives of isonicotinic acid (Rifimon, Marsilid) in the treatment of active progressive caseous-pneumonic tuberculosis. *Am. Rev. Tuberc.* **65**: 402-428.

Rocha, E. R., and Smith, C. J. 1995. Biochemical and genetic analyses of a catalase from the anaerobic bacterium *Bacteroides fragilis*. *J. Bacteriol.* **177**: 3111-3119.

Roche, P. W., Triccas, J. A., and Winter, N. 1995. BCG vaccination against tuberculosis: past disappointments and future hopes. *Trends Microbiol.* **3**: 397-401.

Rosenberg, M., and Court, D. 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. *Annu. Rev. Genet.* **13**: 319-353.

Rosner, J. L. 1993. Susceptibilities of *oxyR* regulon mutants of *Escherichia coli* and *Salmonella typhimurium* to isoniazid. *Antimicrob. Agents Chemother.* **37**: 2251-2253.

Rosner, J. L., and Storz, G. 1994. Effects of peroxides on susceptibilities of *Escherichia coli* and *Mycobacterium smegmatis* to isoniazid. *Antimicrob. Agents Chemother.* **38**: 1829-1833.

Ross, W., Gosink, K. K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K., and Gourse, R. L. 1993. A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science* **262**: 1407-1413.

Rouse, D. A., DeVito, J. A., Li, Z. M., Byer, H., and Morris, S. L. 1996. Site-directed mutagenesis of the *katG* gene of *Mycobacterium tuberculosis*: effects on catalase-peroxidase activities and isoniazid resistance. *Mol. Microbiol.* **22**: 583-592.

Rouse, D. A., Li, Z. M., Bai, G. H., and Morris, S. L. 1995. Characterization of the *katG* and *inhA* genes of isoniazid-resistant clinical isolates of *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **39**: 2472-2477.

Rouse, D. A., and Morris, S. L. 1995. Molecular mechanisms of isoniazid resistance in *Mycobacterium tuberculosis* and *Mycobacterium bovis*. *Infect. Immun.* **63**: 1427-1433.

Runyon, E. H., Wayne, L. G., and Kubica, G. P. 1974. In: R. E. Buchanan and N. E. Gibbons (eds). *Bergey's Manual of Determinative Bacteriology*. 8th Ed. The Williams and Wilkins Co., Baltimore.

Sabelnikov, A. G., Greenberg, B., and Lacks, S. A. 1995. An extended -10 promoter alone directs transcription of the DpnII operon of *Streptococcus pneumoniae*. *J. Mol. Biol.* **250**: 144-155.

Sager, P., Schalimtzek, M., and Moller-Christensen, V. 1972. A case of tuberculosis spondylosa in the Danish Neolithic Age. *Danish Med. Bull.* **19**: 176-180.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuch, 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 catalase hydroperoxidase II in *Escherichia coli* are both regulated by the *katF* gene product. *Proc. Natl. Acad. Sci. USA* **86**: 3271-3275.

Sakula, A. 1982. Robert Koch: centenary of the discovery of the tubercle bacillus, 1882. *Bull. Int. Union Tuberc.* **57**: 111-116.

Salo, W. L., Aufderheide, A. C., Biukstra, J., and Holcomb, T. A. 1994. Identification of *Mycobacterium tuberculosis* DNA in a pre-Columbian Peruvian mummy. *Proc. Natl. Acad. Sci. USA* **91**: 2091-2094.

Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989. Molecular cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York.

Sammartano, L. J., Tuveson, R. W., and Davenport, R. 1986. Control of sensitivity to inactivation by H_2O_2 and broad-spectrum near-UV radiation by the *Escherichia coli katF* locus. *J. Bacteriol.* **168**: 13-21.

Sancar, A. 1994. Structure and function of DNA photolyase. *Biochemistry* **33**: 2-9.

Sancar, A., Franklin, K. A., and Sancar, G. B. 1984. *Escherichia coli* DNA photolyase stimulates UvrABC excision nuclease *in vitro*. *Proc. Natl. Acad. Sci. USA* **81**: 7397-7401.

Sancar, A., and Hearst, J. E. 1993. Molecular matchmakers. *Science* **259**: 1415-1420.

Sancar, A., and Sancar, G. B. 1984 *Escherichia coli* DNA photolyase is a flavoprotein. *J. Mol. Biol.* **172**: 223-227.

Sancar, G. B., and Sancar, A. 1987. Structure and function of DNA photolyases. *Trends Biochem. Sci.* **12**: 259-261.

Sancar, A., and Sancar, G. B. 1988. DNA repair enzymes. *Annu. Rev. Biochem.* **57**: 29-67.

Sanger, F., Nicklen, S., and Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**: 5463-5467.

Sanzey, B. 1979. Modulation of gene expression by drugs affecting deoxyribonucleic acid gyrase. *J. Bacteriol.* **138**: 40-47.

Sarkis, G. J., Jacobs, W. R., and Hatful, G. F. 1995. L5 luciferase reporter mycobacteriophages: a sensitive tool for the detection and assay of live mycobacteria. *Mol. Microbiol.* **15**: 1055-1067.

Sarmientos, P., Sylvester, J. E., Contente, S., and Cashel, M. 1983. Differential stringent control of the tandem *E. coli* ribosomal RNA promoters from the *rrnA* operon expressed *in vivo* in multicopy plasmids. *Cell* **32**: 1337-1346.

Saunders, B. C., Holmes, A. G., and Stark, B. P. 1964. Peroxidase. The properties and uses of a versatile enzyme and of some related catalysis. Butterworth and Co., Ltd., London.

Schatz, A., Bugie, E., and Waksman, A. 1944. Streptomycin, a substance exhibiting antibiotic activity against gram-positive and gram-negative bacteria. *Proc. Soc. Exp. Biol. Med.* **55**: 66.

Schell, M. A. 1993. Molecular biology of the LysR family of transcriptional regulators. *Annu. Rev. Microbiol.* **47**: 597-626.

Schellhorn, H. E. 1994. Regulation of hydroperoxidase (catalase) expression in *Escherichia coli*. *FEMS Microbiol. Lett.* **131**: 113-119.

Schellhorn, H. E., and Hassan, H. M. 1988. Transcriptional regulation of *katE* in *Escherichia coli* K-12. *J. Bacteriol.* **170**: 4286-4292.

Schellhorn, H. E., and Stones, V. L. 1992. Regulation of *katF* and *katE* in *Escherichia coli* K-12 by weak acids. *J. Bacteriol.* **174**: 4769-4776.

Schiering, N., Tao, X., Zeng, H., Murphy, J. R., Petsko, G. A., and Ringe, D. 1995. Structures of the apo- and metal ion-activated forms of the diphtheria tox repressor from *Corynebacterium diphtheriae*. *Proc. Natl. Acad. Sci. USA* **92**: 9843-9850.

Schlegel, H. G. 1986. General Microbiology. 6th Ed. Cambridge University Press, Cambridge.

Schmetterer, G., Wolk, C. P., and Shay, J. 1986. Expression of luciferases from *Vibrio harveyi* and *Vibrio fisheri* in filamentous cyanobacteria. *J. Bacteriol.* **167**: 411-414.

Schmitt, M. P., and Holmes, R. K. 1991. Iron-dependent regulation of diphtheria toxin and siderophore expression by the cloned *Corynebacterium diphtheriae* repressor gene *dtxR* in *C. diphtheriae* C7 strains. *Infect. Immun.* **59**: 1899-1904.

Schmitt, M. P., Predich, M., Doukhan, L., Smith, I., and Holmes, R. K. 1995. Characterization of an iron-dependent regulatory protein (IdeR) of *Mycobacterium tuberculosis* as a functional homolog of the diphtheria toxin repressor (DtxR) from *Corynebacterium diphtheriae*. *Infect. Immun.* **63**: 4284-4289.

Schmitt, M. P., Twiddy, E. M., and Holmes, R. K. 1992. Purification and characterization of the diphtheria toxin repressor. *Proc. Natl. Acad. Sci. USA* **89**: 7576-7580.

Schnell, S., and Steinman, H. M. 1995. Function and stationary-phase induction of periplasmic copper-zinc superoxide dismutase and catalase/peroxidase in *Caulobacter crescentus*. *J. Bacteriol.* **177**: 5924-5929.

Schonbaum, G. R., and Chance, B. 1976. Catalase. In: P. D. Boyer (ed). *The Enzymes*. Vol.13. Academic Press, New York.

Scorpio, A., Lindholm-Levy, P., Heifets, L., Gilman, R., Siddiqi, S., Cynamon, M., and Zhang, Y. 1997. Characterization of *pncA* mutations in pyrazinamide-resistant *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **41**: 540-543.

Scorpio, A., and Zhang, Y. 1996. Mutations in *pncA*, a gene encoding pyrazinamidase/nicotinamidase, cause resistance to the antituberculous drug pyrazinamide in tubercle bacillus. *Nature Med.* **2**: 662-667.

Segal, W., and Edwards, B. S. 1984. In: G. P. Kubica and L. G. Wayne (eds). *The Mycobacteria: a Sourcebook*. Marcel Dekker, New York and Basel.

Sela, S., and Clark-Curtiss, J. E. 1991. Cloning and characterization of the *Mycobacterium leprae* ribosomal RNA promoter in *Escherichia coli*. *Gene* **98**: 123-127.

Selby, C. P., and Sancar, A. 1990. Structure and function of the (A)BC exinuclease of *Escherichia coli*. *Mutat. Res.* **236**: 203-211.

Sellers, M. I., Nakamura, R., and Tokunaga, T. 1970. The effects of ultraviolet radiation on mycobacteriophages and their infectious DNA's. *J. Gen. Virol.* **7**: 23.

Seth, A. 1984. A new method for linker ligation. *Gene Analysis Techniques.* **1**: 99-103.

Setlow, R. B. 1966a. The molecular basis of biological effects of ultraviolet radiation and photoreactivation. *Curr. Top. Radiat. Res.* **2**: 195-248.

Setlow, R. B. 1966b. Cyclobutane-type pyrimidine dimers in polynucleotides. *Science* **153**: 379-386.

Setlow, R. B. 1968. The photochemistry, photobiology, and repair of polynucleotides. *Prog. Nucleic Acid Res. Mol. Biol.* **8**: 257-295.

Setlow, R. B., and Carrier, W. L. 1966. Pyrimidine dimers in ultraviolet-irradiated DNA's. *J. Mol. Biol.* **17**: 237.

Sherman, D. R., Mdluli, K., Hickey, M. J., Arain, T. M., Morris, S. L., Barry, C. E. III., and Stover, C. K. 1996. Compensatory *ahpC* gene expression in isoniazid-resistant *Mycobacterium tuberculosis*. *Science* **272**: 1641-1643.

Sherman, D. R., Sabo, P. J., Hickey, M. J., Arain, T. M., Mahairas, G. G., Yuan, Y., Barry, C. E. III., and Stover, C. K. 1995. Disparate responses to oxidative stress in saprophytic and pathogenic mycobacteria. *Proc. Natl. Acad. Sci. USA* **92**: 6625-6629.

Shim, T. S. 1997. Isoniazid resistance and the point mutation of codon 463 of the *katG* gene of *Mycobacterium tuberculosis. J. Korean Med. Sci.* **12**: 92-98.

Shine, J., and Dalgarno, L. 1974. The 3'-terminal sequence of *E. coli* 16S rRNA: complementarity to nonsense triplets and ribosome binding sites. *Proc. Natl. Acad. Sci. USA* **71**: 1342-1346.

Shinnick, T. M. 1987. The 65-kilodalton antigen of *Mycobacterium tuberculosis. J. Bacteriol.* **169**: 1080-1088.

Shinnick, T. M., Plikaytis, B. B., Hyche, A. D., Van Landingham, R. M., and Walker, L. L. 1989. The *Mycobacterium tuberculosis* BCG-α protein has homology with the *Escherichia coli* GroES protein. *Nucleic Acids Res.* **17**: 1254.

Shoeb, H. A., Bowman, B. U., Jr., Ottolenghi, A. C., and Merola, A. J. 1985a. Peroxidase-mediated oxidation of isoniazid. *Antimicrob. Agents Chemother.* **27**: 399-403.

Shoeb, H. A., Bowman, B. U., Jr., Ottolenghi, A. C., and Merola, A. J. 1985b. Evidence for the generation of active oxygen by isoniazid treatment of extracts of *Mycobacterium tuberculosis* H37Ra. *Antimicrob. Agents Chemother.* **27**: 404-407.

Short, J. M., Fernandez, J. M., Sorge, J. A., and Huse, W. D. 1988. Lambda ZAP: a bacteriophage lambda expression vector with *in vivo* excision properties. *Nucleic Acids Res.* **16**: 7583-7600.

Singh, A., and Singh, H. 1982. Time scale and nature of radiation biological damage. *Prog. Biophys. Mol. Biol.* **39**: 69-107.

Sirokova, T. D., Bardarov, S. S., Kriakov, J. I., and Markov, K. I. 1989. Molecular cloning of mycobacterial promoters in *E. coli. FEMS Microbiol. Lett.* **59**:153-156.

Sivaraja, M., Goodin, D. B., Smith, M., and Hoffman, B. M. 1989. Identification by ENDOR of Trp^{191} as the free-radical site in cytochrome c peroxidase compound ES. *Science* **245**: 738-740.

Sladek, K. M., Dynlacht, B., and Howard-Flanders, P. 1988. Characterization and partial purification of a novel activity from *Escherichia coli* that acts on psoralen mono-adducts. *J. Cell Biochem.* **12**: 276.

Snapper, S. B., Bloom, B. R., and Jacobs, W. R. 1990a. Molecular genetic approaches to mycobacterial investigation. In: J. McFadden (ed). *Molecular Biology of the Mycobacteria.* pp. 199-218. Surrey University Press, London.

Snapper, S. B., Lugosi, L., Jekkel, A., Melton, R. E., Kieser, T., Bloom, B. R., and Jacobs, W. R. 1988. Lysogeny and transformation in mycobacteria: stable expression of foreign genes. *Proc. Natl. Acad. Sci. USA* **85**: 6987-6991.

Snapper, S. B., Melton, R. E, Mustafa, S., Kieser, T., and Jacobs, W. R., Jr. 1990b. Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis. Mol. Microbiol.* **4**: 1911-1919.

Snow, G. A., 1970. Mycobactins: iron-chelating growth factors from mycobacteria. *Bacteriol. Rev.* **34**: 99-125.

Sreevatsan, S., Pan, X., Zhang, Y., Deretic, V., and Musser, J. M. 1997. Analysis of the *oxyR-ahpC* region in isoniazid-resistant and -susceptible *Mycobacterium tuberculosis* organisms recovered from diseased humans and animals in diverse localities. *Antimicrob. Agents Chemother.* **41**: 600-606.

Sreevatsan, S., Pan, X., Zhang, Y., Kreiswirth, B. N., and Musser, J. M. 1997. Mutations associated with pyrazinamide resistance in *pncA* of *Mycobacterium tuberculosis* complex organisms. *Antimicrob. Agents Chemother.* **41**: 636-640.

Sritharan, M., Wheeler, P. R., and Ratledge, C. 1989. Metabolism of aspartate in *Mycobacterium smegmatis*. *Eur. J. Biochem.* **180**: 587-593.

Sritharan, M., Wheeler, P. R., and Ratledge, C. 1990. Aspartate metabolism in *Mycobacterium avium* grown in host tissue and axenically and in *Mycobacterium leprae*. *J. Gen. Microbiol.* **136**: 203-209.

Staden, R. 1984. Computer methods to locate signals in nucleic acid sequences. *Nucleic Acids Res.* **12**: 505-519.

Staggs, T. M., and Perry, R. D. 1991. Identification and cloning of a *fur* regulatory gene in *Yersinia pestis*. *J. Bacteriol.* **173**: 417-425.

Stanford, J. L., Stanford, C. A., Rook, G. A. W., and Grange, J. M. 1994. Immunotherapy for tuberculosis-investigative and practical aspects. *Clin. Immunother.* **1**: 430-440.

Stanley, P. E., McCarthy, B. J., and Smither, R. (eds). 1989. Rapid Methods in Microbiology. Society for Applied Bacteriology, Blackwell Scientific Publications.

Stoeckle, M. Y., Guan, L., Riegler, N., Weitzman, I., Kreiswirth, B., Kornblum, J., Laraque, F., and Riley, L. W. 1993. Catalase-peroxidase gene sequences in isoniazid-sensitive and -resistant strains of *Mycobacterium tuberculosis* from New York City. *J. Inf. Dis.* **168**: 1063-1065.

Storz, G., Tartaglia, L. A., and Ames, B. N. 1990. Transcriptional regulation of oxidative stress-inducible genes: direct activation by oxidation. *Science* **248**: 189-194.

Stover, C. K., de la Cruz, V. F., Fuerst, T. R., Burlein, J. E., Benson, L. A., Bennett, L. T., Bansal, G. P., Young, J. F., Lee, M. H., Hatfull, G. F., Snapper, S. B., Barletta, R. G., Jacobs, W. R., Jr., and Bloom. B. R. 1991. New use of BCG for recombinant vaccines. *Nature (London)* **351**: 456-460.

Strauch, M. A., and Hoch, J. A. 1993. A positive feedback loop controls transcription of the *spoOF* gene, a component of the sporulation phosphorelay in *Bacillus subtilis*. *Mol. Microbiol.* 7: 967-974 .

Strohl, W. R. 1992. Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. *Nucleic Acids Res.* **20**: 961-974.

Sturgill-Koszycki, S., Schaible, U. E., and Russell, D. G. 1996. *Mycobacterium*-containing phagosomes are accessible to early endosomes and reflect a transitional state in normal phagosome biogenesis. *EMBO J.* **15**: 6960-6968.

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.

Suzuki, Y., Nagata, A., and Yamada, T. 1991. Analysis of the promoter region in the rRNA operon from *Mycobacterium bovis BCG*. *Antonie van Leeuwenhoek* **60**: 7-11.

Switala, J., Triggs-Raine, B. L., and Loewen, P. C. 1990. Homology among bacterial catalase genes. *Can. J. Microbiol.* **36**: 728-731.

Tai, S. S., and Zhu, Y. Y. 1995. Cloning of a *Corynebacterium diphtheriae* iron-repressible gene that shares homology with the AhpC subunit of alkyl hydroperoxide of *Salmonella typhimurium*. *J. Bacteriol.* **177**: 3512-3517.

Takayama, K., and Qureshi, N. 1984. In: G. B. Kubica and L. G. Wayne (eds). *The Mycobacteria: a Sourcebook.* pp. 315-344. Marcel Dekker, New York and Basel.

Takayama, K., Schnoes, H. K., Armstrong, E. L., and Boyle, R. W. 1975. Site of inhibitory action of isoniazid in the synthesis of mycolic acids in *Mycobacterium tuberculosis. J. Lipid Res.* **16**: 308-317.

Takayama, K., Wang, L., and David, H. L. 1972. Effect of isoniazid on the *in vivo* mycolic acid synthesis, cell growth, and viability of *Mycobacterium tuberculosis. Antimicrob. Agents Chemother.* **2**: 29-35.

Takiff, H. E., Salazar, L., Guerrero, C., Philipp, W., Huang, W. M., Kreiswirth, B., Cole, S. T., Jacobs, W. R., and Telenti, A. 1994. Cloning and nucleotide sequence of *Mycobacterium tuberculosis gyrA* and *gyrB* genes and detection of quinolone resistance mutations. *Antimicrob. Agents Chemother.* **38**: 773-780.

Tanaka, K., Takayanagi, Y., Fujita, N., Ishihama, A., and Takahashi, H. 1993. Heterogeneity of the principal sigma factor in *Escherichia coli*: the *rpoS* gene product sigma38 is a secondary principal sigma factor of RNA polymerase in stationary phase *Escherichia coli. Proc. Natl. Acad. Sci. USA* **90**: 3511-3515.

Tao, K., Fujita, N., and Ishihama, I. 1993. Involvement of the RNA polymerase α subunit C-terminal region in cooperative interaction and transcriptional activation with OxyR protein. *Mol. Microbiol.* **7**: 859-864.

Tao, K., Makino, K., Yonei, S., Nakata, A., and Shinagawa, H. 1991. Purification and characterization of the *Escherichia coli* OxyR protein, the positive regulator for a hydrogen peroxide-inducible regulon. *J. Biochem.* **109**: 262-266.

Tartaglia, L. A., Storz, G., and Ames, B. N. 1989. Identification and molecular analysis of *oxyR*-regulated promoters important for the bacterial adaptation to oxidative stress. *J. Mol. Biol.* **210**: 709-719.

Tauber, A. I., and Babior, B. M. 1978. O_2^- and host defense: the production and fate of O_2^- in neutrophils. *Photochem. Photobiol.* **28**: 701-709.

Telenti, A., Philipp, W. J., Sreevatsan, S., Bernasconi, C., Stockbauer, K. E., Wieles, B., Musser, J. M., and Jacobs, W. R. 1997. The *emb* operon, a gene cluster of *Mycobacterium tuberculosis* involved in resistance to ethambutol. *Nature Med.* **3**: 567-570.

Terasaka, K., Yamaguchi, R., Matsuo, K., Yamazaki, A., Nagai, S., and Yamada, T. 1989. Complete nucleotide sequence of immunogenic protein MPB70 from *Mycobacterium bovis* BCG. *FEMS Microbiol. Lett.* **49**: 273-276.

Thole, J. E. R., Dauwerse, H. G., Das, P. K., Groothuis, 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. Infect. Immun.* **50**: 800-806.

Thole, J. E. R., Keulen, W. J., Kolk, A. H. J., Groothuis. D. G., Berwald, L. J., Tiesjema, R. H., and van Embden, J. D. A. 1987. Characterization, sequence determination, and immunogenicity of a 64-kilodalton protein of *Mycobacterium bovis* BCG expressed in *Escherichia coli* K-12. *Infect. Immun.* **55**: 1466-1475.

Thomas, C. E., and Sparling, P. F. 1994. Identification and cloning of a *fur* homolog from *Neisseria meningitidis. Mol. Microbiol.* **11**: 725-737.

Timm, J. L., Lim, M., and Gicquel, B. 1994. *Escherichia coli* mycobacteria shuttle vectors for operon and gene fusions to *lacZ*: the pJEM series. *J. Bacteriol.* 176: 6749-6753.

Toledano, M. B., Kullik, I., Trinh, F., Baird, P. T., Schneider, T. D., and Storz, G. 1994. Redox-dependent shift of OxyR-DNA contacts along an extended DNA-binding site: a mechanism for differential promoter selection. *Cell* **78**: 897-909.

Tolmasky, M. E., Wertheimer, A. M., Actis, L. A., and Crosa, J. H. 1994. Characterization of the *Vibrio anguillarum fur* gene: role in regulation of expression of the FatA outer membrane protein and catechols. *J. Bacteriol.* **176**: 213-220.

Touati, D., Jacques, M., Tardat, B., Bouchard, L., and Despied, S. 1995. Lethal oxidative damage and mutagenesis are generated by iron in Δ*fur* mutants of *Escherichia coli*: protective role of superoxide dismutase. *J. Bacteriol.* **177**: 2305-2314.

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*. *J. Bacteriol.* **170**: 4415-4419.

Tsolis, R. M., Baümler, A. J., Stojiljkovic, I., and Heffron, F. 1995. Fur regulon of *Salmonella typhimurium*: identification of new iron-regulated genes. *J. Bacteriol.* **177**: 4628-4637.

Tsui, H. C., Feng, G., and Winkler, M. E. 1996. Transcription of the *mutL* repair, *miaA* tRNA modification, *hfq* pleiotropic regulator, and *hflA* region protease genes of *Escherichia coli* K-12 from clustered sigma 32-specific promoters during heat shock. *J. Bacteriol.* **178**: 5719-5731.

Tsukamura, M., and Tsukamura, S. 1963. Isotopic studies on the effect of isoniazid on protein synthesis of mycobacteria. *Jap. J. Tuberc.* **11**: 14.

Turnowsky, F., Fuchs, K., Jeschek, C., and Hogenauer, G. 1989. *envM* genes of *Salmonella typhimurium* and *Escherichia coli*. *J. Bacteriol.* **171**: 6555-6565.

VanBogelen, 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*. *J. Bacteriol.* **169**: 26-32.

Van Houten, B. 1990. Nucleotide excision repair in *Escherichia coli*. *Microbiol. Rev.* **54**: 18-51.

Van Houten, B., and Snowden, A. 1993. Mechanism of action of the *Escherichia coli* UvrABC nuclease: clues to the damage recognition problem. *Bioessays* **15**: 51-59.

Van Zyl, J. M., and Van der Walt, B. J. 1994. Apparent hydroxyl radical generation without transition metal catalysis and tyrosine nitration during oxidation of the anti-tubercular drug, isonicotinic acid hydrazide. *Biochem. Pharmacol.* **48**: 2033-2042.

Verbist, L., and Gyselen, A. 1968. Antituberculous activity of rifampin *in vitro* and *in vivo* and the concentrations attained in human blood. *Am. Rev. Respir. Dis.* **98**: 923-932.

Verma, A., Kinger, A. K., and Tyagi, J. S. 1994. Functional analysis of transcription of the *Mycobacterium tuberculosis* 16S rDNA-encoding gene. *Gene* **148**: 113-118.

Victor, T. C., Pretorius, G. S., Felix, J. V., Jordaan, A. M., van Helden, P. D., and Eisenach, K. D. 1996. *katG* mutations in isoniazid-resistant strains of *Mycobacterium tuberculosis* are not infrequent. *Antimicrob. Agents Chemother.* **40**: 1572.

von Hippel, P. H., Bear, D. G., Morgan, W. D., and McSwiggen, J. A. 1984. Protein-nucleic acid interactions in transcription: a molecular analysis. *Annu. Rev. Biochem.* **53**: 389-446.

Walker, G. C. 1984. Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli*. *Microbiol. Rev.* **48**: 60-93.

Walker, G. C. 1985. Inducible DNA repair systems. *Annu. Rev. Biochem.* **54**: 425-457.

Walker, G. C. 1987. The SOS response of *Escherichia coli* and *Salmonella typhimurium* In: F. C. Neidhart (ed). *Cellular and Molecular Biology*. ASM Press, Washington D. C.

Walker, L., and Lowrie, D. B. 1981. Killing of *Mycobacterium microti* by immunologically activated macrophages. *Nature (London)* **293**: 69-70.

Wang, T. C. V., and Smith, K. C. 1983. Mechanisms for *recF*-dependent and *recB*-dependent pathways of postreplication repair in UV-irradiated *Escherichia coli uvrB*. *J. Bacteriol.* **156**: 1093-1098.

Wayne, L. G. 1962. Two varieties of *M. kansasii* with different clinical significance. *Am. Rev. Respir. Dis.* **86**: 651-656.

Wayne, L. G. 1976. Dynamics of submerged growth of *Mycobacterium tuberculosis* under aerobic and microaerophilic conditions. *Am. Rev. Respir. Dis.* **114**: 807-811.

Wayne, L. G. 1977. Synchronized replication of *Mycobacterium tuberculosis*. *Infect. Immun.* **17**: 528-530.

Wayne, L. G., and Diaz, G. A. 1982. Serological, taxinomic and kinetic studies of the T and M classes of mycobacterial catalase. *Int. J. System. Bacteriol.* **32**: 296-304.

Wayne, L. G., and Diaz, G. A. 1986. A double staining method for differentiating between two classes of mycobacterial catalase in polyacrylamide electrophoresis gels. *Anal. Biochem.* **157**: 89-92.

Wayne, L. G., and Diaz, G. A. 1988. Detection of a novel catalase in extracts of *Mycobacterium avium* and *Mycobacterium intracellulare*. *Infect. Immun.* **56**: 936-941.

Wayne, L. G., Diaz, G. A., and Doubek, J. R. 1968. Acquired isoniazid resistance and catalase activity of mycobacteria. *Am. Rev. Respir. Dis.* **97**: 909-913.

Wayne, L. G., and Sramek, H. A. 1994. Metranidazole is bactericidal to dormant cells of *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **38**: 2054-2058.

Wefers, H., and Sies, H. 1983. Oxidation of glutathione by the superoxide radical to the disulfide and sulfonate yielding singlet oxygen. *Eur. J. Biochem.* **127**: 29-36.

Welch, D. F., Sword, C. P., Brehm, S., and Dusanic, D. 1979. Relationship between superoxide dismutase and pathogenic mechanisms of *Listeria monocytogenes*. *Infect. Immun.* **23**: 863-872.

Welinder, K. G. 1991. Bacterial catalase-peroxidases are gene duplicated members of the plant peroxidase superfamily. *Biochim. Biophys. Acta* **1080**: 215-220.

Welinder, K. G. 1992. Superfamily of plant, fungal and bacterial peroxidases. *Curr. Opin. Struct. Biol.* **2**: 388-393.

Wells, W. F. 1955. Airborne Contagion and Air Hygiene. Harvard University Press, Cambridge, Mass.

West, S. C., Cassuto, E., and Howard-Flanders, P. 1981. Mechanism of *E. coli* RecA protein-directed strand exchanges in post-replication repair of DNA. *Nature (London)* **294**: 659-662.

Westpheling, J., Ranes, M., and Losick, R. 1985. RNA polymerase heterogeneity in *Streptomyces coelicolor*. *Nature (London)* **313**: 22-27.

Weyer, K., and Kleeberg, H. H. 1992. Primary and acquired drug resistance in adult black patients with tuberculosis in South Africa: results of a continuous national drug surveillance programme involvement. *Tuberc. Lung Dis.* **73**: 106-112.

Wheeler, P. R., and Ratledge, C. 1994. Metabolism of *Mycobacterium tuberculosis*. In: B. Bloom (ed). *Tuberculosis: Pathogenesis, Protection, and Control*. p.353-385. ASM Press, Washington D. C.

Williams, T. M., Burlein, J. E., Ogden, S., Kricka, L. J., and Kant, J. A. 1989. Advantages of firefly luciferase as a reporter gene: application to the interleukin-2 gene promoter. *Anal. Biochem.* **176**: 28-32.

Willis, D. K., Uhlin, B. E., Amini, K. S., and Clark, A. J. 1981. Physical mapping of the *srl recA* region of *Escherichia coli*: analysis of Tn10 generated insertions and deletions. *Mol. Gen. Genet.* **183**: 497-504.

Wilson, T. M., and Collins, D. M. 1996. *ahpC*, a gene involved in isoniazid resistance of the *Mycobacterium tuberculosis* complex. *Mol. Microbiol.* **19**: 1025-1034.

Wilson, T. M., de Lisle, G. W., and Collins, D. M. 1995. Effect of *inhA* and *katG* on isoniazid resistance and virulence of *Mycobacterium bovis*. *Mol. Microbiol.* **15**: 1009-1015.

Winder, F. G. 1960. Catalase and peroxidase in mycobacteria: Possible relationship to the mode of action of isoniazid. *Am. Rev. Respir. Dis.* **81**: 68-78.

Winder, F. G. 1982. Mode of action of the antimycobacterial agents and associated aspects of the molecular biology of mycobacteria. In: C. Ratledge and J. Stanford (eds). *The Biology of Mycobacteria.* Academic Press, New York.

Winder, F. G., and Brennan, P. 1965. Effect of isoniazid on lipid metabolism of *Mycobacterium tuberculosis*. *Biochem. J.* **96**: 77.

Winder, F. G., and Collins, P. 1968. The effect of isoniazid on nicotinamide nucleotide levels in *Mycobacterium bovis*, strain BCG. *Am. Rev. Respir. Dis.* **97**: 719-720.

Winder, F. G., and Collins, P. 1969. The effect of isoniazid on nicotinamide nucleotide concentrations in tubercle bacilli: A correction and further observations. *Am. Rev. Respir. Dis.* **100**: 101-103.

Winder, F. G., and Collins, P. 1970. Inhibition by isoniazid of synthesis of mycolic acids in *Mycobacterium tuberculosis*. *J. Gen. Microbiol.* **63**: 41-48.

Winder, F. G., and Rooney, S. A. 1968. Effects of isoniazid on the triglycerides of BCG. *Am. Rev. Respir. Dis.* **97**: 938-940.

Witkin, E. M. 1976. Ultraviolet mutagenesis and inducible DNA repair in *Escherichia coli*. *Bacteriol. Rev.* **40**: 869-907.

Wolinsky, E., Smith, M. M., and Steenken, W., Jr. 1956. Isoniazid susceptibility, catalase activity, and guinea pig virulence of recently isolated cultures of tubercle bacilli. *Am. Rev. Tuberc.* **73**: 768-772.

Wood, K. V. 1990. Luc genes: introduction of colour into bioluminescence assays. *J. Biolumin. Chemilumin.* **5**: 107-114.

Wood, K. V. 1991a. Recent advances and prospects for use of Beetle luciferases as genetic reporters. In: P. Stanley and L. Kricka (eds). *Bioluminescence and chemiluminescence: current status.* John Wiley and sons, Chichester.

Wood, K. V. 1991b. The origin of Beetle luciferases. In: P. Stanley and L. Kricka (eds). *Bioluminescence and chemiluminescence: current status.* John Wiley and sons, Chichester.

Wood, K. V., de Wet, J. R., Dewji, N., and DeLuca, M. 1984. Synthesis of active firefly luciferase by *in vitro* translation of RNA obtained from adult lanterns. *Biochem. Biophys. Res. Commun.* **124**: 592-596.

Woolridge, K. G., Williams, P. H., and Ketley, J. M. 1994. Iron-responsive genetic regulation in *Campylobacter jejuni*: cloning and characterization of a *fur* homolog. *J. Bacteriol.* **176**: 5852-5856.

Word, J. F., Limali, C. L., Calabrojones, P. M., and Aguitera, J. 1991. An examination of the repair saturation hypothesis for describing shouldered survival curves. *Rad. Res.* **127**: 90-96.

World Health Organization. 1989. Global Programme on AIDS and Tuberculosis Programme, World Health Organization, in collaboration with the International Union Against Tuberculosis and Lung Disease. *Bull. Int. Union Tuberc. Lung Dis.* **64**: 8-11.

World Health Organization. 1994. TB: a global emergency: WHO report on the TB epidemic. WHO Tuberculosis Program Information. World Health Organization, Washington D. C.

Wu. J., and Weiss, B. 1991. Two divergently transcribed genes, *soxR* and *soxS*, control a superoxide response regulon in *Escherichia coli. J. Bacteriol.* **173**: 2864-2871.

Yajko, D. M., and Weiss, B. 1975. Mutations simultaneously affecting endonuclease II and exonuclease III in *E. coli. Proc. Natl. Acad. Sci. USA* **72**: 688-692.

Yamaguchi, R., Matsuo, K., Yamazaki, A., Abe, C., Nagai, S., Terasaka, K., and Yamada, T. 1989. Cloning and characterization of the gene for immunogenic protein MPB64 of *Mycobacterium bovis* BCG. *Infect. Immun.* 57: 283-288.

Yanisch-Perron, 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.

Yoakum, G., and Eisenstark, A. 1972. Toxicity of L-tryptophan photoproduct on recombinationless (*rec*) mutants of *Salmonella typhimurium. J. Bacteriol.* **112**: 653-655.

Yonetani, T. 1976. In: P. D. Boyer (ed). *The Enzymes.* Vol. 13. Academic Press, New York.

Youatt, J. 1960. The uptake of isoniazid and related compounds by mycobacteria. *Aust. J. Exp. Biol. Med. Sci.* **38**: 331.

Youatt, J. 1961a. The formation of 4-pyridinemethanol from isoniazid by mycobacteria. *Aust. J. Chem.* **14**: 308.

Youatt, J. 1961b. Pigments produced by mycobacteria on exposure to isoniazid. *Aust. J. Exp. Biol. Med. Sci.* **39**: 93-100.

Youatt, J. 1969. A review of the action of isoniazid. *Am. Rev. Respir. Dis.* **99**: 729.

Youatt, J., and Tham, S. H. 1969a. Radioactive content of *Mycobacterium tuberculosis* after exposure to ^{14}C-isoniazid. *Am. Rev. Respir. Dis.* **100**: 77-78.

Youatt, J., and Tham, S. H. 1969b. An enzyme system of *Mycobacterium tuberculosis* which reacts specifically with isoniazid: II. Correlation of this reaction with the binding and metabolism of isoniazid. *Am. Rev. Respir. Dis.* **100**: 31-37.

Young, R. A., Bloom, B. R., Grosskinsky, C. M., Ivanyi, J., Thomas, D., and Davis, R. W. 1985. Dissection of *Mycobacterium tuberculosis* antigens using recombinant DNA. *Proc. Natl. Acad. Sci. USA* **82**: 2583-2587.

Young, D. B., and Cole, S. T. 1993. Minireview: Leprosy, tuberculosis, and the new genetics. *J. Bacteriol.* **175**: 1-6.

Young, L. S., and Wormser, G. P. 1994. The resurgence of tuberculosis. *Scand. J. Infect. Dis.* **S93**: 9-19.

Yuan, Y., Lee, R. E., Besra, G. S., Belisle, J. T., and Barry, C. E. III. 1995. Identification of a gene involved in the biosynthesis of cyclopropanated mycolic acids in *Mycobacterium tuberculosis. Proc. Natl. Acad. Sci. USA* **92**: 6630-6634.

Yumoto, I., Fukumori, Y., and Yamanaka, T. 1990. Purification and characterization of catalase from a facultative alkalophilic *Bacillus*. *J. Biochem.* **108**: 583-587.

Zabeau, M., and Stanley, K. K. 1982. Enhanced expression of Cro-β-galactosidase fusion proteins under the control of the Pr promoter of bacteriophage lambda. *EMBO J.* **1**: 1217-1224.

Zabinski, R. F., and Blanchard, J. S. 1997. The requirement for manganese and oxygen in the isoniazid-dependent inactivation of *Mycobacterium tuberculosis* enoyl reductase. *J. Am. Chem. Soc.* **119**: 2331-2332.

Zappe, H., Jones, D. T., and Woods, D. R. 1986. Cloning and expression of *Clostridium acetobutylicum* endoglucanase, cellobiase and amino acid biosynthesis genes in *Escherichia coli*. *J. Gen. Microbiol.* **132**: 1367-1372.

Zhang, Y., Dhandayathupani, S., and Deretic, V. 1996. Molecular basis for the exquisite sensitivity of *Mycobacterium tuberculosis* to isoniazid. *Proc. Natl. Acad. Sci. USA* **93**: 13212-13216.

Zhang, Y., Heym, B., Allen, B., Young, D., and Cole, S. T. 1992. The catalase-peroxidase gene and isoniazid resistance of *Mycobacterium tuberculosis*. *Nature (London)* **358**: 591-593.

Zhang, Y., Lathigra, R., Garbe, T., Catty, D., and Young, D. 1991. Genetic analysis of superoxide dismutase, the 23 kilodalton antigen of *Mycobacterium tuberculosis*. *Mol. Microbiol.* **5**: 381-391.

Zhang, Y., and Young, D. B. 1993. Molecular mechanisms of isoniazid: a drug at the front line of tuberculosis control. *Trends Microbiol.* **110**: 109-113.

Zhang, Y., and Young, D. 1994. Strain variation in the *katG* region of *Mycobacterium tuberculosis*. *Mol. Microbiol.* **14**: 301-308.

Zheng, H., Hassett, D. J., Bean, K., and Cohen, M. S. 1992. Regulation of catalase in *Neisseria gonorrhoeae*. *J. Clin. Invest.* **90**: 1000-1006.

Zhou, Y., Busby, S., and Ebright, R. 1993b. Identification of the functional subunit of a dimeric transcription activator protein by use of oriented heterodimers. *Cell* **73**: 375-379.

Zhou, Y., Merkel, T., and Ebright, R. 1994a. Characterization of the activating region of *Escherichia coli* catabolite gene activator protein (CAP). II. Role at Class I and class II CAP-dependent promoters. *J. Mol. Biol.* **243**: 603-610.

Zhou, Y., Pendergrast, S., Bell, A., Williams, R., Busby, S., and Ebright, R. 1994b. The functional subunit of a dimeric transcription activator protein depends on promoter architecture. *EMBO J.* **13**: 4549-4557.

Zhou, Y., Zhang, X., and Ebright, R. 1993a. Identification of the activating region of catabolite gene activator protein (CAP): isolation and characterization of mutants of CAP specifically defective in transcription activation. *Proc. Natl. Acad. Sci. USA* **90**: 6081-6085.

Zimmerman, M. R. 1979. Pulmonary and osseous tuberculosis in an Egyptian mummy. *Bull. New York Acad. Med.* **55**: 604-608.