

The Gene Cloning and Biochemical and Genetic
Analysis of the Ribonuclease HI from
Mycobacterium smegmatis

Stephanie Susan Dawes

A Dissertation Submitted to the Faculty of Science,
University of the Witwatersrand, Johannesburg,
for the Degree of Doctor of Philosophy,
June, 1998.

ABSTRACT

Two members of the genus *Mycobacteria* are the causative agents of tuberculosis and leprosy, and many others are opportunistic pathogens in immune-compromised individuals. An alarming resurgence of tuberculosis in First World countries, coupled with the high toll of mortality that this disease exacts in the Developing World, has led the World Health Organisation to declare tuberculosis a global emergency. Relatively little is known, however, about the molecular biology of the mycobacteria. Elucidation of mechanisms involved in the regulation of fundamental processes in this genus, such as DNA replication, are expected to provide a basis for understanding mechanisms of pathogenesis in the mycobacteria. RNase HI, encoded by *rnhA*, plays a central role in DNA replication in *Escherichia coli* by fixing initiation of DNA replication to one origin. The presence of a similar activity in the mycobacteria was therefore investigated by probing a genetically manipulable member of the mycobacteria, *M. smegmatis*, for a homologue of the *rnhA* gene. Degenerate primers based on blocks of amino acids conserved in other bacterial RNase HI homologues had been shown to amplify an internal portion of *rnhA* from *M. smegmatis* (Mizrahi *et al.*, 1993). This PCR generated probe was used in this study to isolate the entire gene from *M. smegmatis*, which was cloned as two overlapping fragments and sequenced. The deduced amino acid sequence of *M. smegmatis rnhA* coded for an RNase HI of 159 amino acids which showed 50% identity to *E. coli* RNase HI. A recombinant form of *M. smegmatis* RNase HI was expressed as part of a fusion protein to the maltose binding protein to facilitate *in vitro* biochemical characterisation of the protein. The recombinant RNase HI exhibited hydrolytic activity specific for the RNA strand of an RNA:DNA hybrid and generated a similar distribution of hydrolysis products to that of *E. coli* RNase HI.

The regulation of RNase HI was probed at the post-translational level, and at the level of transcription. Observations by Foster and Marinus (1992) had suggested a possible regulation of cellular levels of *E. coli* RNase HI by the molecular chaperonins, DnaJ and DnaK. Preliminary experiments, therefore, using the *E. coli* homologues of RNase HI, DnaJ and DnaK were performed, and a direct interaction between RNase HI

and DnaK, and between RNase HI and DnaJ was confirmed. As a high degree of homology exists between the *E. coli* and *M. smegmatis* homologues, a similar interaction may exist in the mycobacteria

Transcription of *M. smegmatis rnhA* proceeds from its own promoter and the transcriptional start point was mapped. This promoter region is not recognised by *E. coli*, and no *E. coli* -10 or -35 consensus motifs in this promoter region are apparent. Unusual direct repeat structures observed in the promoter region of *M. smegmatis rnhA* appear not to affect regulation of transcription.

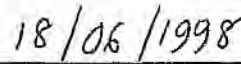
The chromosomal context of *M. smegmatis rnhA* differs from the overlapping divergent promoter organisation with *dnaQ* observed for the Gram negative bacteria *E. coli*, *H. pylori* and *S. typhi*, and from the varied contexts of other actinomycete *rnhA* genes. *M. smegmatis rnhA* does not appear to be part of an operon, but an unexplained polar effect on unknown neighbouring gene(s) was observed on insertional inactivation of *rnhA* with a resistance marker. Attempts to create a knockout mutant of *rnhA* by homologous recombination were unsuccessful, suggesting that *rnhA* may encode an essential function in *M. smegmatis*.

DECLARATION

I declare that this thesis is my own unaided work. It is being submitted for the degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg. It has not been submitted for any degree or examination at any other university.



Stephanie Susan Dawes



Date

ACKNOWLEDGEMENTS

I would like to express my gratitude and appreciation to my supervisor, Professor Valerie Mizrahi, for her enthusiasm, dedication and inspiration and most particularly for her availability for discussions and help with problem solving.

I would like to thank Professor Robert Crouch in whose facilities (Laboratory of Molecular Genetics, National Institute of Child Health and Human Development, NIH, Bethesda, MD) and under whose guidance the protein interaction experiments were conducted. Purified *E. coli* RNase HI for these experiments was kindly provided by Dr. Wu-Po Ma.

I would also like to thank Professor Graham Hatfull (Department of Biological Sciences, University of Pittsburgh, PA) for provision of *M. smegmatis* RNA polymerase and the facilities to perform the run-off transcription experiments, and Dr. Chadd Nesbit, under whose guidance these experiments were performed.

I would also like to thank my family for their support in what must have seemed a never-ending project, my friends and colleagues in the Molecular Biology Unit for their enthusiasm, support and willingness to help in many ways -- especially Helena, may I be there for you, too! -- and Dr. Pelle Stolt, my important other, for his enthusiasm, thought-provoking discussions and excellent technical advice.

I would also like to thank my friends Christopher Wilson and Jean Snijders for opening their home to me so that I could write up without distractions, and Christopher in particular for his stalwart presence, encouragement and reasoned advice from the beginning to the end.

I acknowledge and appreciate the funding that was given to me by the University of the Witwatersrand, the South African Institute for Medical Research, the Foundation for Research Development and the Glaxo-Wellcome Action TB Initiative.

PUBLICATIONS FROM THIS THESIS

Dawes SS, Crouch RJ, Morris SL, Mizrahi V (1995) Cloning, sequence analysis, overproduction in *Escherichia coli* and enzymatic characterization of the RNase HI from *Mycobacterium smegmatis*. *Gene* 165:71-75.

Mizrahi V, Huberts P, Dawes SS, Dudding LR (1993) A PCR method for the amplification of the *gyrA*, *polA*, and *mha* gene segments from mycobacteria. *Gene* 136:287-290.

TABLE OF CONTENTS

ABSTRACT	i
DECLARATION.....	iii
ACKNOWLEDGEMENTS	iv
PUBLICATIONS FROM THIS THESIS	v
TABLE OF CONTENTS	vi
LIST OF TABLES	xi
LIST OF FIGURES	xii
1. INTRODUCTION	1
1.1 MYCOBACTERIA AND MAN	1
The disease burden	1
Prophylaxis and prevention of disease	2
TB in the Developing World	3
TB in the First World	3
Needs identified	4
1.2 MYCOBACTERIA AND MOLECULAR BIOLOGY	5
A surrogate host	5
DNA delivery systems	6
Generation of mutants for analysis of gene function	9
Advances in screening for drug susceptibility	10
New drugs and new vaccines	11
Progress in understanding mechanisms of pathogenesis	12
Chromosomal replication in the mycobacteria	14
1.3 RNase HI: GENETIC AND BIOCHEMICAL STUDIES	16
General RNases HI	16
Cellular roles of RNase HI	17
3-D structure of RNase HI	18
Substrate binding	20
Catalytic site	20

Regulation of <i>rnhA</i>	21
Phenotypes of mutants with altered levels of RNase HI	22
1.4 AIMS OF THIS STUDY	24
2. MATERIALS AND METHODS	25
2.1 REAGENTS AND MANUFACTURED ITEMS	25
2.2 BACTERIAL STRAINS AND GROWTH CONDITIONS.....	25
<i>Escherichia coli</i>	26
<i>Mycobacterium smegmatis</i>	26
2.3 PLASMIDS AND BACTERIOPHAGES	26
2.4 TRANSFORMATION OF BACTERIA BY ELECTROPORATION	28
Preparation of electrocompetent <i>M. smegmatis</i>	28
Preparation of electrocompetent <i>E. coli</i>	28
Electroporation of DNA	28
2.5 CHEMICAL TRANSFORMATION OF <i>E. coli</i>	29
Strains DH5 α , TB1, MIC3037 and GM161	29
JM101	29
2.6 DNA EXTRACTIONS	30
Rapid isolation of <i>E. coli</i> plasmids	30
Chromosomal DNA from <i>M. smegmatis</i>	30
ssDNA from M13 phage	31
Uracil-enriched ssDNA from M13 phage	31
2.7 DNA MANIPULATIONS	32
Digestions with restriction enzymes	32
Blunt-ending of 5'-overhangs	32
Polymerase chain reaction (PCR)	32
Polishing of PCR products	32
Oligodeoxyribonucleotide-directed mutagenesis	32
Recovery of DNA from agarose gels	34
Recovery of labeled RNA from polyacrylamide gels	34

2.8 <i>IN VITRO</i> TRANSCRIPTION	34
SP6/T7 run-off transcription	34
Transcription with <i>M. smegmatis</i> RNA polymerase	34
2.9 PREPARATION OF LABELED NUCLEIC ACID	35
(a) Probes	35
5'-phosphorylation of oligodeoxyribonucleotides	35
Random-primed labeling	35
Primer extension for S1 protection assay	35
(b) Substrates for RNase and DNase assays	36
Labeled RNA	36
RNA-DNA and RNA-RNA hybrid formation	36
2.10 DNA SEQUENCING	36
2.11 ELECTROPHORESIS TECHNIQUES	37
Polyacrylamide gel electrophoresis (PAGE)	37
SDS-PAGE	38
Agarose gel electrophoresis of DNA	38
Agarose gel electrophoresis of RNA	38
2.12 HYBRIDISATION TECHNIQUES	38
Southern blotting	38
(a) Transfer	38
Gel fractionated DNA	38
Colony lifts	39
(b) Hybridisation	39
S1 nuclease protection assay	39
2.13 ISOLATION OF RNA FROM MYCOBACTERIA	40
2.14 PROTEIN EXTRACTION AND PURIFICATION	40
Total cell protein from mycobacteria	40
Purification of recombinant MBP::RNase HI	41
2.15 PROTEIN BIOCHEMICAL ASSAYS	41
Gel renaturation assay of RNase H activity	41
Electrophoretic separation of RNase HI hydrolysis products	42

Surface plasmon resonance	42
Catechol 2,3 dioxygenase (CDO) activity assay	43
3. RESULTS	44
3.1 CLONING AND SEQUENCE ANALYSIS OF THE	
<i>M. smegmatis rnhA</i> GENE	44
Gene cloning of the <i>M. smegmatis rnhA</i> gene	44
Identification of the <i>rnhA</i> ORF	44
The chromosomal context of <i>rnhA</i>	48
3.2 PROMOTER ANALYSIS OF <i>rnhA</i>	50
<i>rnhA</i> transcription is driven off its own promoter	50
Identification of the transcriptional start site	51
Functional analysis of the <i>rnhA</i> promoter region	51
3.3 ATTEMPTS TO CREATE A KNOCKOUT MUTANT OF <i>rnhA</i> BY	
HOMOLOGOUS RECOMBINATION	54
Construction of a plasmid for <i>rnhA</i> gene disruption	54
Allelic replacement of <i>rnhA</i> by <i>rnhA::aph</i> confers a lethal phenotype ...	55
<i>M. smegmatis rnhA</i> in trans cannot complement <i>rnhA::aph</i>	58
Recombination at the <i>rnhA</i> allele with <i>rnhA::aph</i> is polarised	60
Allelic replacement at the <i>rnhA</i> locus	63
Allelic replacement of <i>rnhA</i> by <i>rnhA::fs</i> confers a lethal phenotype	64
3.4 BIOCHEMICAL CHARACTERISATION OF RECOMBINANT	
<i>M. smegmatis</i> RNase HI	66
Construction of the <i>rnhA</i> expression cassette	66
Purification of recombinant MBP::RNase HI	68
Site selectivity of substrate cleavage of MBP::RNase HI	69
3.5 INTERACTION OF RNase HI WITH HEAT SHOCK PROTEINS	72
DnaJ binds to σ^{32}	72
DnaJ interacts with RNase HI	74
Interaction of DnaK with RNase HI	76
Interaction of DnaK with DnaJ and RNase HI	76

4. DISCUSSION	80
<i>M. smegmatis rnhA</i> does not cross hybridise to <i>M. tuberculosis rnhA</i>	80
<i>M. smegmatis rnhA</i> encodes an RNase HI of 159 amino acids	80
<i>M. smegmatis</i> RNase HI belongs to the Class I family of RNases HI	82
Key structural and catalytic residues are conserved between	
<i>M. smegmatis</i> RNase HI and <i>E. coli</i> and <i>T. thermophilus</i> RNases HI	83
<i>M. smegmatis</i> RNase HI specifically hydrolyses the RNA strand of an	
RNA-DNA hybrid	86
Transcription of <i>M. smegmatis rnhA</i> is driven by its own promoter	87
<i>M. smegmatis rnhA</i> -10 and -35 regions do not show conservation with	
<i>E. coli</i> consensus motifs	87
The <i>M. smegmatis rnhA</i> promoter is not recognised in <i>E. coli</i>	89
Analysis of the relative strength of the <i>rnhA</i> promoter	90
The promoter region of <i>M. smegmatis rnhA</i> extends beyond the tandem	
direct repeats	91
Transcription of <i>M. smegmatis rnhA</i> is not affected by induction of the	
SOS response	92
<i>E. coli</i> RNase HI interacts with DnaJ and DnaK	93
Insertion of <i>aph</i> in <i>rnhA</i> exerts a lethal polar effect	97
RNase HI is essential in <i>M. smegmatis</i>	99
Conclusions	103
 5. APPENDICES	 104
APPENDIX A: Reagents and suppliers	104
APPENDIX B: Media and general solutions	106
APPENDIX C: Abbreviations	108
APPENDIX D: Properties of <i>M. smegmatis</i> RNase HI	110
 6. REFERENCES	 111

LIST OF TABLES

Table I : Bacterial strains	25
Table II : <i>E. coli</i> and <i>E. coli-M. smegmatis</i> shuttle plasmids	27
Table III : Electroporation parameters	29
Table IV : Synthetic oligodeoxyribonucleotides	33
Table V : PAGE parameters	37

LIST OF FIGURES

Fig. 1. Ribbon structure of <i>E. coli</i> RNase HI	19
Fig. 2. Probe of <i>M. smegmatis</i> and <i>M. tuberculosis</i> genomic DNA for <i>rnhA</i>	45
Fig. 3. Restriction maps of pSD1 and pSD2	46
Fig. 4. Amino acid alignment of prokaryotic RNases HI	47
Fig. 5. The chromosomal context of <i>M. smegmatis</i> <i>rnhA</i>	49
Fig. 6. <i>In vitro</i> transcription with <i>M. smegmatis</i> RNA polymerase	50
Fig. 7. Determination of the transcriptional start site	52
Fig. 8. Promoter probe constructs	53
Fig. 9. Homologous recombination with pPBA	55
Fig. 10. Homologous recombination with pPBAS	57
Fig. 11. Complementation of <i>mc</i> ² 155::pPBAS with <i>M. smegmatis</i> RNase HI	59
Fig. 12. Types of SCO events with pPBAS	60
Fig. 13. Location of <i>hORF</i>	62
Fig. 14. Homologous recombination with pPABS	63
Fig. 15. Homologous recombination with pUSH- δ B	65
Fig. 16. Construction of the <i>rnhA</i> expression cassette	67
Fig. 17. Overexpression, purification and RNase HI activity gel analysis of MBP:: β -GAL and MBP::RNase HI	69
Fig. 18. RNase HI liquid hydrolysis assay	70
Fig. 19. PAGE analysis of <i>E. coli</i> RNase HI and MBP::RNase HI hydrolysis products	71
Fig. 20. DnaJ binds to σ^{32}	73
Fig. 21. Interaction of RNase HI with DnaJ at 37°C	75
Fig. 22. Interaction of RNase HI with DnaK at 37°C	77
Fig. 23. Interaction of DnaJ and DnaK with RNase HI at 37°C	78

1. INTRODUCTION

1.1 MYCOBACTERIA AND MAN

The disease burden

The Mycobacteria fall under the order Actinomycetales which are characteristically Gram positive, filamentous, spore forming bacteria. This order includes the Streptomyces, among which are commercially important antibiotic producers, the Rhodococci, opportunistic pathogens of man and animal (Hart *et al.*, 1988; Schaal and Lee, 1992), the Corynebacteria, the Nocardia, the Gordona and the Tsukamurella (Goodfellow and Cross, 1984). The Mycobacteria are distinguished in this group by their unusual cell wall structure (reviewed by Lee *et al.*, 1996). *Mycobacterium tuberculosis* and *M. leprae* are perhaps the most well known of this group, as they are the causative agents of tuberculosis (TB) and leprosy, respectively. There are 1 billion individuals infected with *M. tuberculosis*, with 8 million new cases, and an estimated 3 million deaths, per year (Fenton and Vermeulen, 1996; Sudre *et al.*, 1992). In addition, it is estimated that there are presently 1.3 million cases of infection with *M. leprae*, with 0.5 million new cases reported each year (Bland, 1996). Well-organised leprosy control programs are in place, however, and the elimination of this disease (defined as less than 1 case in 10 000) is predicted for the year 2000 (Noordeen, 1996).

The disease burden as a consequence of infection with *M. tuberculosis* alone is ranked seventh among the top ten causes of Disability-Adjusted Life Years (DALYs), and this ranking is not expected to improve before 2020 (Murray and Lopez, 1996). Infections by *M. szulgai* and *M. africanum*, which also cause tubercle-like disease in man, add to the human disease burden (Maloney *et al.*, 1987) and large economic losses due to weight loss, lowered fertility, lowered milk production and death result from infection of ruminants by *M. bovis* and *M. paratuberculosis* (Cocito *et al.*, 1994). In addition, many of the typically saprophytic mycobacterial species, such as *M. smegmatis*, *M. goodii*, *M. genavense*, *M. chelonae*, *M. celatum*, *M. fortuitum*, *M. kansasii*, *M. xenopi* and *M. intracellulare*, can also cause disseminated infection in immune-compromised individuals (Fournier *et al.*, 1998; Bonomo *et al.*, 1998; Rusconi *et al.*, 1997; But: 1998; el-Helou *et*

al., 1997; Albrecht *et al.*, 1997; Gadre 1997), as can the animal pathogens such as *M. bovis* and *M. avium* (Ellner *et al.*, 1991; Brettle, 1997; Guerrero *et al.*, 1997; Hocqueloux *et al.*, 1998).

Prophylaxis and prevention of disease

The burden of disease imposed by the Mycobacteria has created the need to understand the biology of this genus, and to generate means of control. TB and leprosy have been documented diseases of man since 2500 BC, but preventative prophylaxis against TB, in the form of vaccination with an attenuated strain of *M. bovis*, namely bacille Calmette-Guérin (BCG), only became available in the 1920s. Since 1948, this vaccine has been administered worldwide, and it is currently administered to over 100 million children each year (Guleria *et al.*, 1996). BCG over the years has built an impressive record of safety and has also been found to afford some protection against leprosy (Karonga Prevention Trial Group, 1996).

Subsequently, treatment of active infection with *M. tuberculosis* became possible with the discovery of streptomycin in 1945 and isoniazid in 1952. Chemical prophylaxis against leprosy also became a reality with the advent of dapsone in 1950. However, single drug therapy soon resulted in the emergence of drug resistant *M. leprae* and *M. tuberculosis* bacilli. Fortunately, multi-drug treatment initiated by the WHO in 1981, based on a combination of dapsone, clofazamine and rifampin, remains effective against leprosy and has resulted in the curing of 8 million patients to date (Nakajima, 1996).

Leprosy, however, is not as infectious as TB, since prolonged physical contact with patients is usually necessary for the spread of infection. In contrast, TB is spread very easily in aerosols and it has been shown in animal models that inhalation of one bacillus can be sufficient to cause infection (Bloom and Murray, 1992). As a testimony to its pervasiveness, it is estimated that one third of the world's population is latently infected with *M. tuberculosis*. The risk of developing clinical TB in such infected immunocompetent individuals is a 10% lifetime risk (Styblo, 1991), whereas immune compromised individuals run an 8% per year risk of the same (Selwyn *et al.*, 1989).

TB in the Developing World

Poor standards of living, combined with delayed diagnosis and lack of facilities for effective treatment, has meant that TB continues to exact a high toll of mortality in the Developing World; 95% of the new cases of TB per year occur in developing countries. BCG vaccination has been the strategy of choice to prevent TB in many countries in Europe and the Developing World, but the efficiency of vaccination has proved to be variable. Efficiencies range from 77% protection in the UK (Hart and Sutherland, 1977), to 0% in India (Tuberculosis Prevention Trial, Madras, 1980) and Malawi (Karonga Prevention Trial Group, 1996). This variance has been ascribed to the exposure of populations to resident environmental mycobacteria resulting in inherent levels of immunity amongst the population which vary from country to country depending on levels, distribution, and immunogenicity of the resident mycobacterial species (Day, 1996; Fine *et al.*, 1995). In some populations, vaccination with BCG provides 50%-75% protection against leprosy, but none against TB (Karonga Prevention Trial Group, 1996). Compounding the effect of the variable efficacy of BCG is the rising number of cases of TB with co-infection by human immunodeficiency virus (HIV). It has been noted that infection of macrophages by *M. tuberculosis* can activate latent HIV infection (Toossi *et al.*, 1997; Garrait *et al.*, 1997), and the suppression of the immune system by HIV may, in turn, reactivate latent infection by *M. tuberculosis* which leads to rapid progression of disease (reviewed in Murray, 1997). In the ten years from 1990 to 2000 it is expected that the number of cases of TB attributable to HIV infection will rise from 4% to 14% (Lienhardt and Rodrigues, 1997).

TB in the First World

In contrast to the situation in the Developing World, the introduction of chemotherapy, together with an improvement in living standards, helped to reduce the incidence of TB in the USA to the extent that many considered the disease all but eradicated in this First World country. This led to the dismantling of many clinics and the collapse of follow-up programmes. In 1985, however, a reversal in the decline of TB cases was noted with an increasing number of cases associated with HIV infection, alcohol and

drug abuse, and the decrease in living standards associated with increasing levels of homelessness (Brudney and Dobkin, 1991). Many patients undergoing the six-month antibiotic treatment regimen were lost to follow-up, and were re-admitted with reactivation of latent infections arising from non-compliance with the treatment. This alarming resurgence of TB in First World countries resulted in the classification of the disease as a global emergency by the World Health Organisation (WHO) in 1993. Non-compliance of patients with antibiotic treatment regimes has also resulted in the emergence of strains of *M. tuberculosis* that are resistant to one or more of the front line drugs used to treat the disease. This has contributed to the upsurge in cases of TB as the time-consuming diagnosis of multi-drug-resistant (MDR) strains of *M. tuberculosis* has resulted in outbreaks of MDR TB in hospitals and hospices. Co-infection of individuals with MDR TB and HIV results in an accelerated progression of the disease with fatality following within 2 months in 80% of cases.

Needs identified

Despite the availability of BCG and effective chemotherapy, TB still claims more lives worldwide than any other infectious agent (Kochi, 1991) and as such, demands: a) the development of systems to identify drug resistant strains more rapidly; b) the development of new drugs and vaccines to treat and prevent TB; and c) an improved understanding of the mechanisms of pathogenicity and host-bacteria interactions.

Progress towards meeting these needs has been slow. A knowledge of mycobacterial biochemistry and genetics is central to achieving these ends, but the introduction of successful antibiotic treatment of TB occurred concurrently with the discovery of the superhelical structure of DNA, and hence before the explosion of molecular biological information and availability of associated tools for manipulating bacterial systems. Initial investigations therefore focussed mainly on the cloning of antigenic determinants of *M. leprae* with the aim of making BCG a more effective vaccine against leprosy. The resurgence of TB in First World countries, followed by the categorisation of TB as a global emergency by the WHO, resulted in increased availability of funding in recent years for mycobacterial research with a concomitant explosion of

knowledge in this field. The genome of *M. tuberculosis* has been fully sequenced, and the genome of *M. leprae* is partially complete (Smith *et al.*, 1997; Cole *et al.*, 1998). The combination of these data, together with knowledge of systems gained from the study of other bacteria, has been a powerful aid for goal oriented research.

1.2 MYCOBACTERIA AND MOLECULAR BIOLOGY

A surrogate host

The typically slow growth rate of these pathogenic species (*M. tuberculosis* doubles every 18-24 hours and *M. leprae* every 13-14 days (Smith *et al.*, 1997)), combined with the refractoriness of *M. leprae* to *in vitro* culture, has hampered molecular biological characterisation of these bacteria. Conventional systems which utilise *E. coli* as a heterologous host for gene expression and regulation have had limited success. Although many of the highly conserved members of the heat shock family of proteins and cell wall-associated lipoproteins have been cloned and expressed in *E. coli* (reviewed in Young *et al.*, 1992), recent studies have shown that of *M. tuberculosis* promoters active in the heterologous mycobacterial host *M. smegmatis*, only 12% were active in *E. coli* (Das Gupta *et al.*, 1993). This may be due to the difference in G+C content of the genomes: *E. coli* has a G+C content of 44%, while mycobacterial G+C contents range from 58% for *M. leprae* (Smith *et al.*, 1997) to 65% for *M. smegmatis* (Clark-Curtiss, 1990).

Alternative hosts with high G+C content such as *Streptomyces* sp. and *Corynebacterium glutamicum* have been utilised to express mycobacterial DNA (Kieser *et al.*, 1986; Salim *et al.*, 1997), but these organisms are as difficult to work with as are the more amenable species from the mycobacterial group. Mycobacteria have also been found to glycosylate proteins, a relatively rare occurrence amongst prokaryotes. Dobos *et al.* (1995) provided the first study to show unequivocally the O-glycosylation of a threonine residue in a mycobacterial protein, the *M. tuberculosis* secreted protein MPT32. The entire protein has subsequently been mapped and threonine residues in both the N-terminal and C-terminal portions of the polypeptide were found to be glycosylated (Dobos *et al.*, 1996). Studies on the 19 kDa antigen from *M. tuberculosis* suggested that the role of glycosylation in the N-terminal part of this protein is to protect a proteolytically sensitive

region (Herrmann *et al.*, 1996). Post-translational modification of a recombinant form of the *M. tuberculosis* antigen MPT83 (which carries a putative signal motif for O-glycosylation) was observed in *M. smegmatis*, but not in *E. coli* (Hewinson *et al.*, 1996). Attempts to express and secrete four major *M. tuberculosis* extracellular proteins were unsuccessful in *E. coli* or *B. subtilis* based systems, but accurate processing of leader peptides and secretion of these proteins was achieved in *M. smegmatis* (Harth *et al.*, 1997). Glycosylation of proteins is expected to play an important role in the determination of the antigenicity of cell wall components and pathogenicity of the mycobacteria, and the usefulness of non-mycobacterial hosts for further study in this area is therefore limited.

Development and characterisation of a non-pathogenic and relatively fast growing member of the mycobacteria such as *M. smegmatis* into a heterologous system was therefore considered the method of choice.

DNA delivery systems

Genetic manipulation of the mycobacteria by classical transformation methods was initially hampered by the formidable barrier presented by their cell wall, and the early identification of a number of mycobacteriophages, initially used to type mycobacteria, provided the opportunity to investigate manipulation of the mycobacteria through transfection. Efficient uptake of DNA was demonstrated by transfection with the lytic mycobacteriophage D29 (Gicquel-Sanzey *et al.*, 1989), so attempts were made to manipulate the temperate mycobacteriophages TM4, L1 and L5 for use as vectors in order to by-pass the barrier of the mycobacterial cell wall (Jacobs *et al.*, 1987; Snapper *et al.*, 1988; Lee *et al.*, 1991). Unfortunately, although the lytic mycobacterial phage D29 plaques on a broad host range of mycobacteria (Russell *et al.*, 1963), the temperate phage L5 appeared to be confined to the fast growing mycobacterial species which has so far restricted the utility of such vectors (Sarkis *et al.*, 1995). Recently however, the discovery of the absolute requirement for calcium ions for L5 infection of *M. bovis* BCG has extended the field of phage manipulation of the mycobacteria to the slow growers (Fullner and Hatfull, 1997).

The discovery of a naturally occurring 5 kb plasmid in *M. fortuitum*, pAL5000, also provided the opportunity for plasmid-based systems to be developed (Labidi *et al.*, 1985a,b). The mode of replication exhibited by pAL5000 shows similarity to the ColE2 type plasmids found in the Enterobacteria and replication appears dependent on host DNA polymerase I (Gordhan and Mizrahi, 1997). pAL5000 has been found to replicate both in *M. bovis* BCG and *M. smegmatis*, unlike the *M. avium* plasmid pLR7 and the *M. scrofulaceum* plasmid pMSC262 which are unable to replicate in *M. smegmatis*. *Mycobacterium-E. coli* shuttle plasmids have been constructed using pAL5000 most frequently as the mycobacterial origin of replication (Jacobs *et al.*, 1987; Gicquel-Sanzey *et al.*, 1989; Ranes *et al.*, 1990; Radford and Hodgson, 1991; Fuerst *et al.*, 1991; Hinshelwood and Stoker 1992; Guilhot *et al.*, 1994; Villar and Benitez, 1994; Gavigan *et al.*, 1995; Matsumoto *et al.*, 1996), although shuttle plasmids based on phage D29, pLR7 and pMSC262 origins of replication, have also been constructed (David *et al.*, 1992; Beggs *et al.*, 1995; Goto *et al.*, 1991; Qin *et al.*, 1994). Recently, a cryptic plasmid, pJAZ38, showing similarities to pLR7 and pMSC262, has been isolated from *M. fortuitum* (Gavigan *et al.*, 1997). This plasmid is stably inherited in the absence of antibiotic selection and can replicate in *M. smegmatis* and thus shows potential for a new set of vectors for plasmid DNA delivery.

Initial attempts to transform *M. smegmatis* to streptomycin resistance by chemical means (in the presence of calcium and magnesium ions), and after mild irradiation with UV light, were not very successful (Norgard and Imaeda, 1978) and although spheroplast fusion was demonstrated in *M. aurum* (Rastogi *et al.*, 1983), no transformants of *M. smegmatis* could be obtained using protoplasts (Gicquel-Sanzey *et al.*, 1989). The phenomenon of natural competence displayed by *Bacillus subtilis* appears absent in the mycobacteria, as determined by the apparent lack of a protein competence factor, and by the observation that similar transformation efficiencies were achieved through all stages of growth (Norgard and Imaeda, 1978).

Electroporation of plasmid DNA into the mycobacteria provided a means of obtaining transformants, albeit initially at low frequencies - ca. 10 transformants per microgram of plasmid DNA (Snapper *et al.*, 1988). However, the isolation of an efficient

plasmid transformation mutant of *M. smegmatis* by Snapper *et al.* (1990), that allowed recovery of more than 10^5 transformants per microgram of plasmid DNA, provided a boost for plasmid-based delivery of DNA. Delivery of DNA into the slow growing mycobacteria is relatively inefficient although electroporation efficiencies of *M. bovis* BCG were increased from 10-100 colonies/microgram of DNA to 10^3 - 10^4 if the plasmid used for transformation was passaged through *M. bovis* beforehand (Lugosi *et al.*, 1989). This is probably a consequence of the methylation systems present in the mycobacteria (Hemavathy and Nagaraja, 1995). Efficient delivery of suicide plasmids into *M. bovis*, *M. tuberculosis* and *M. intracellulare* remains problematical although recent optimisation of protocols has raised this efficiency by several orders of magnitude (Aldovini *et al.*, 1993; Wards and Collins, 1996).

Useful markers for the selection of transformants include *aph* from *E. coli* Tn903 or Tn5 (Snapper *et al.*, 1988; Lugosi *et al.*, 1989), the hygromycin resistance gene, *hyg*, from *S. hygroscopicus* (Lydiate *et al.*, 1989; Garbe *et al.*, 1994), the mercury resistance (*mer*) genes from *Pseudomonas aeruginosa* and *Serratia marcescens* (Baulard *et al.*, 1995) and the apramycin resistance gene (Paget and Davies, 1996). Other markers include *sacB* from *Bacillus subtilis* which allows for counter selection of DNA bearing this gene (Pelicic *et al.*, 1996) as does a temperature-sensitive (*ts*) origin of replication in the delivery DNA (Pelicic *et al.*, 1997; Guillhot *et al.*, 1994).

Reporter genes that have found utility in mycobacterial systems include the transcriptional fusion reporter gene *xyIE* from *Pseudomonas putida* which encodes catechol 2,3 dioxygenase (Curcio *et al.*, 1994), chloramphenicol acetyl transferase (CAT) (Das Gupta *et al.*, 1993; Verma *et al.*, 1994), green fluorescent protein (GFP) from *Aequorea victoria* (Dhandayuthapani *et al.*, 1995; Kremer *et al.*, 1995a), alkaline phosphatase (*phoA*) from *E. coli* (Kremer *et al.*, 1995b), β -galactosidase (*lacZ*) reporter gene from *E. coli* (Barletta *et al.*, 1990; Barletta *et al.*, 1992; Timm *et al.*, 1994a; Timm *et al.*, 1994b; Dellagostin *et al.*, 1995) and the luciferase (*lux*) genes from firefly or *Vibrio cholerae* (Andrew and Roberts, 1993; Cooksey *et al.*, 1993; Jacobs *et al.*, 1993).

Generation of mutants for analysis of gene function

Mutagenesis of the mycobacteria has been achieved using N-methyl-N'-nitro-N-nitrosoguanidine or ultra violet irradiation (Rastogi *et al.*, 1983; Fiss *et al.*, 1994; Klamm *et al.*, 1998), but generally has not proved efficient; the cell wall provides a barrier to DNA-damaging agents, efficiencies are low, and isolation of single colonies is difficult due to the tendency of the cells to clump. It is also difficult to analyse chemically induced mutant phenotypes in that more than one mutation may have been generated in the chromosome. Transposon mutagenesis therefore was considered a more suitable method to generate libraries of mutants. Initial attempts to generate a transposon library of *M. smegmatis* used mutagenesis of a library of *M. smegmatis* in *E. coli*; this library was then transferred to the chromosome of *M. smegmatis* by homologous recombination (Kalpana *et al.*, 1991). Libraries of *M. smegmatis* using mycobacterial transposons have also been made (Gicquel, 1994) but the combination of the low efficiency of transformation coupled with the low frequency of transposition have made representative libraries difficult to obtain, especially for the slow growers (McAdam *et al.*, 1995).

The counter-selectable ts plasmid pB4D* has facilitated construction of such libraries in *M. smegmatis* (Guilhot *et al.*, 1994), but transfer of this technology to the slow-growing mycobacteria was complicated by the fact that *sacB* and the ts origin of replication of pB4D*, although effective in *M. smegmatis*, were only weakly counter-selectable in the slow growing mycobacteria (Pelicic *et al.*, 1997). However, the combination of these two markers in one system has been shown to be able to select efficiently for recombination events in *M. tuberculosis*, and a representative transposon library has finally been made for this organism (Pelicic *et al.*, 1997). In parallel, Bardarov *et al.* (1997) have generated transposon libraries of *M. phlei*, *M. bovis* BCG and *M. tuberculosis* utilising ts shuttle plasmids, based on D29 and TM4 phages, that replicate in *E. coli* as plasmids and infect mycobacteria at a permissive temperature. Growth of infected cells at the non-permissive temperature then allows for the selection of transposition events.

Such libraries, however, are only useful if mutant phenotypes can be easily identified. It is also important to be able to locate and analyse specific genes, and to be

able generate specific mutants. Target genes in the mycobacteria have been identified and cloned using heterologous probes (Davis *et al.*, 1991) or degenerate primers based on bacterial homologues (Ma *et al.*, 1997; Gordhan *et al.*, 1996; Rajagopalan *et al.*, 1995; Dawes *et al.*, 1995; Mizrahi *et al.*, 1993) or monoclonal antibodies (Lefevre *et al.*, 1997; Collins *et al.*, 1990), and by complementation of the appropriate *M. smegmatis* mutants (Telenti *et al.*, 1997a; Jackson *et al.*, 1996; Fiss *et al.*, 1994) or *E. coli* mutants (Cirillo *et al.*, 1994; Timm *et al.*, 1992). Although targeted gene replacement by homologous recombination with an insertionally inactivated allele has become routine in *M. smegmatis*, again, transfer of this technology to *M. tuberculosis* and *M. bovis* has not been without its difficulties. A high degree of illegitimate recombination was observed in these species (Kalpana *et al.*, 1991; Aldovini *et al.*, 1993) and, although it has been suggested that this might be useful to generate libraries of mutants (Kalpana *et al.*, 1991), this phenomenon renders isolation of true homologous recombinants virtually impossible in the absence of a counter-selectable system or a mutant phenotype that is easy to discern. However, targeted allele replacements in *M. bovis* BCG have been reported (Reyrat *et al.*, 1995; Norman *et al.*, 1995), and allele replacement of *leuD* using a long linear substrate was achieved in *M. tuberculosis* (Balasubramanian, 1996). The application of *sacB* as a counter-selectable marker has furthered achievements in this arena (Azad *et al.*, 1996; Pelicic *et al.*, 1996b,c; Azad *et al.*, 1997) and, used in conjunction with a *ts* origin of replication as a further counter-selectable agent, has led to another report of targeted allele replacement in *M. tuberculosis* (Pelicic *et al.*, 1997).

These advances in the genetic manipulation of the mycobacteria have facilitated the development of systems to address the needs previously defined.

Advances in screening for drug susceptibility

A number of drug targets have been identified and methods of resistance to these drugs have been characterised. Mycobacteria appear to develop resistance to drugs by accumulating point mutations in the chromosome (Lety *et al.*, 1997; Scorpio *et al.*, 1997; Meier *et al.*, 1994, 1996; Morris *et al.*, 1995; Nash and Inderlied, 1995; Alangaden *et al.*, 1995). Although this has led to the development of PCR-based systems for the

identification of some drug resistant clinical isolates (Piatek *et al.*, 1998; Telenti *et al.*, 1997b; De Beenhouwer *et al.*, 1995), the development of such diagnostic strategies in many cases is complicated since the location of many of the point mutations varies widely (O'Brien *et al.*, 1996).

A more general method to assay for drug susceptibility is based on the development of phage vectors that express the firefly luciferase gene (Carriere *et al.*, 1997; Sarkis *et al.*, 1995; Jacobs *et al.*, 1993). Lysogens of *M. tuberculosis* surviving phage infection continuously express this gene and subsequent survival of these bacteria in the presence of antibiotic is detected by the emission of light.

New drugs and new vaccines

The development of new antimycobacterial drugs has three main foci (reviewed in Barry, 1997): a) the optimisation of the action of existing drugs to shorten duration of treatment, b) the development/discovery of novel drugs to counter resistance to current drugs, and c) the development of drugs that will target dormant bacteria.

Progress, however, has been slow: Klopman *et al.* (1993) have developed a number of new quinolones that show promise of superior activity, and a novel class of anti-mycobacterial drugs, oxazolidinones, has been detected from a directed chemical modification program (Zurenko *et al.*, 1996), but the efficacy of these drugs in clinical trials has yet to be reported.

Attempts have also been made to improve the efficacy of BCG as a vaccine by engineering the bacterium to express cytokines (Murray *et al.*, 1996). However, the concept of BCG as a safe vaccine has been questioned due to the increasing prevalence of HIV-AIDS: although BCG has had an impressive record of safety so far, cases of disseminated infection due to *M. bovis* BCG have been reported in immune-compromised individuals (reviewed by Talbot *et al.*, 1997) although a comprehensive study by Toida (1993) indicated that such cases may be over-reported due to insufficient rigour in the identification of the *M. bovis* strain causing the infection.

Auxotrophic mutants of BCG have been investigated as alternative vaccine strains to address this issue (Guleria *et al.*, 1996; Bange *et al.*, 1996) but subunit vaccines, such

as one based on the secreted 30 kDa protein of *M. tuberculosis* (Sinha and Khuller, 1997), together with DNA vaccines based on secreted antigens (such as Ag85, ESAT-6, hsp65 and hsp70), show promise to become the new generation vaccines against TB (reviewed by Lowrie *et al.*, 1997). Alternate vectors, such as recombinant vaccinia viruses that afford protection in mice against challenge by *M. tuberculosis* by expression of the 39 kDa or 19 kDa secreted glyco-lipoproteins of *M. tuberculosis* (Zhu *et al.*, 1997) and alternative mycobacterial strains, such as *M. vaccae* and *M. habera* which have been proposed as candidate vaccine strains for TB and leprosy respectively (Marsh *et al.*, 1997; Skinner *et al.*, 1997; Singh *et al.*, 1997), are also under investigation.

Unfortunately, although these developments hold some preliminary promise, few advances have been made in the development of new drugs, and clinical implementation of a new vaccine is considered decades away (Efferen, 1997).

Progress in understanding mechanisms of pathogenesis

Very little is known about the molecular mechanisms of pathogenesis and host-bacteria interactions although host responses to infection have been well documented. Areas of host-bacteria interactions under investigation include those of uptake of the bacteria by the macrophage, the ability of the bacteria to multiply within the macrophage, and, under conditions of adequate host immune response, the ability of the bacteria to remain quiescent for extended periods of time until reactivation of disease occurs.

Pathogenic mycobacteria have been shown to subvert host immune systems to enhance uptake by macrophages (Schorey *et al.*, 1997) in addition to their ability to convert phagocytic ability on normally non-phagocytic cells (Arruda *et al.*, 1993). The search for virulence factors has targeted components of the mycobacterial cell wall since the cell wall provides the initial contact between the pathogen and the macrophage. Glycolipid components such as lipoarabinomannans have been shown to reduce phagosome-lysosome fusion after engulfment by phagocytes and to stimulate production of TNF α to varying degrees depending on the virulence of the strain tested (Kang and Schlesinger, 1998; Venisse *et al.*, 1995; Brown and Taffet, 1995). Mycoloyl glycolipids (such as cord factor, sulfolipids and penta acyl trehaloses) have also been shown to enhance markedly

phagocytosis and subsequent survival in macrophages by inhibition of phagosome-lysosome fusion (Fujiwara, 1997). Glycolipid profiles and levels of TUBag production show marked differences from strain to strain and it has been suggested that analysis of such profiles may provide a basis for virulent strain differentiation (Chaicumpar *et al.*, 1997; Constant *et al.*, 1995).

Despite the wealth of information available on virulence factors of other intracellular pathogens, few of the homologues of these genes found in the mycobacteria appear to be virulence genes in isolation. Although enteric pathogens have been found to possess plasmid-encoded virulence factors (reviewed by Bebora, 1997), virulence plasmids in the mycobacteria have only been reported in *M. avium* (Gangadharam *et al.*, 1988). Some chromosomally-encoded virulence factors have, however, been identified such as catalase-peroxidase (*katG*) in *M. bovis*, and the principal sigma factor (*rpoV*) and *mce* in members of the *M. tuberculosis* complex (Wilson *et al.*, 1995; Collins *et al.*, 1995; Parker *et al.*, 1995), and a number of potential virulence factors have been or are under investigation such as phospholipase C (Leao *et al.*, 1995), superoxide dismutase (Escuyer *et al.*, 1996), mycosiderin acid synthase (Azad *et al.*, 1996), urease (Reyrat *et al.*, 1996; Clemens *et al.*, 1995), and haemolysin (Deshpande *et al.*, 1997; Udou, 1994).

The environment encountered by the bacterium within the macrophage is also an area of intense interest as the mycobacteria manage to replicate under conditions that must be limiting for a number of factors. In this regard, research has focussed on pathways employed by bacteria under micro-aerophilic conditions (Cunningham and Spreadbury, 1998), reduced availability of nutrients and trace elements (Beall *et al.*, 1998; Rodriguez *et al.*, 1998), as well as investigations into the profiles of mycobacterial secreted proteins (Raynaud *et al.*, 1998) and genes that are upregulated upon infection of macrophages. Such studies have yielded a gene from *M. avium*, *mig*, that is upregulated under such conditions (Plum *et al.*, 1997; Plum and Clark-Curtiss, 1994) although the *M. tuberculosis* homologue of this gene appears to be constitutively expressed (Spreadbury and Alli, 1998). A number of promoter fragments from *M. marinum* have been found to be selectively active in macrophage cells and homologous DNA to one of these fragments has

been identified in *M. tuberculosis* (Ramakrishnan and Falkow, 1998). The identity of the gene involved, however, is not yet known.

A more holistic approach has been favoured by some researchers who have compared genetic and protein profiles of virulent and avirulent strains; Subtractive hybridisation studies comparing avirulent *M. bovis* BCG and virulent *M. bovis* have suggested that a regulatory mutation involving the expression of at least 10 proteins is responsible for the attenuation of BCG (Mahairas *et al.*, 1996). A similar study has been carried out on avirulent and virulent *M. intracellulare* (Brooks *et al.*, 1998). Libraries of the virulent *M. tuberculosis* strain H37Rv have been constructed and used to transform the avirulent *M. tuberculosis* strain H37Ra to virulence with limited success (Pascopella *et al.*, 1994).

An intriguing aspect of mycobacterial pathogenesis is the slow growth rate of these pathogens. All mycobacteria grow slowly; even under favourable laboratory conditions, the fastest growing mycobacteria take 3 hours to divide. However, the pathogens *M. leprae* and *M. tuberculosis* show extremely slow growth rates with doubling times of 13-14 days and 18-24 hours respectively. A wide ranging study of metabolism in *M. tuberculosis* has indicated that nucleic acid biosynthesis is rate limiting for growth, and both replication of the genome and transcription of RNA genes have been found to be related to generation time (reviewed in Wheeler and Ratledge, 1994). Whether the ability of these bacteria to cause disease or enter a persistent state is related in some manner to this slow growth is uncertain, but elucidation of the molecular mechanisms that regulate DNA replication and transcription of RNA genes during the different growth phases exhibited by the mycobacteria is the focus of several lines of investigation.

Chromosomal replication in the mycobacteria

In the *E. coli* model of DNA replication, initiation of replication occurs from *oriC* where the initiator protein, DnaA, binds to regions within *oriC* and facilitates melting of the duplex. DnaB helicase then enters the duplex, primers are synthesised by DnaG primase and bidirectional chromosomal replication directed by DNA polymerase III (DNA PolIII) proceeds. This mode of replication is dependent on transcription and protein

synthesis. Maintenance of replication at *oriC* is ensured by the action of RNase HI which degrades RNA-DNA hybrids at other locations in the chromosome that result from transcription and could act as primers for replication from alternative sites (reviewed in Messer and Weigel, 1996).

The corresponding *ori* regions of *M. leprae*, *M. tuberculosis* and *M. smegmatis* have been cloned and sequenced (Klann *et al.*, 1998; Qin *et al.*, 1997; Fsihi *et al.*, 1996b; Salazar *et al.*, 1996) and show a typical Gram positive pattern of organisation of the genes, viz. *rnpA-rpmH-dnaA-dnaN-recF-gyrB-gyrA* (Salazar *et al.*, 1996). However the *dnaA-dnaN* regions of *M. leprae* and *M. tuberculosis* were unable to promote autonomous replication as has been shown for the corresponding region in *M. smegmatis* (Salazar *et al.*, 1996). As is the case with other bacteria, cell division and replication of the chromosome have been shown to be linked in *M. smegmatis* (Klann *et al.*, 1998). Another factor related to the growth rate of *M. tuberculosis* has been shown to be the number of ribosomes per cell (Winder and Rooney, 1970). This number appears to be regulated both by the number of *rrn* operons, and by the number of promoters associated with each operon (Gonzalez-y-Merchand *et al.*, 1997).

Some of the enzymes involved in DNA replication in the mycobacteria have been cloned and characterised, such as topoisomerase I (Yang *et al.*, 1996), DNA PolI (Huberts and Mizrahi, 1995), DnaA (Rajagopalan *et al.*, 1995) and *gyrA* and *gyrB* from *M. tuberculosis* (Takiff *et al.*, 1994, Madhusudan *et al.*, 1994). Other genes that perform essential functions in DNA replication have been identified such as *ssb*, *dnaB*, *dnaG* and the subunit genes that constitute DNA PolIII (Fsihi *et al.*, 1996b; Klann *et al.*, 1998; Mizrahi, unpublished). Although a high degree of conservation between pathways and enzymes involved in DNA replication has been noted amongst bacterial species, interesting differences in some of the replication enzymes from the mycobacteria have been noted. The DNA polymerase I (DNA PolI) from *M. tuberculosis* has been shown to lack 3'-5' exonuclease activity and to be retroviral-like in non-discrimination between dNTPs and ddNTPs (Mizrahi and Huberts, 1996) and *polA* from *M. leprae* appears to lack a functional promoter (Gordhan and Mizrahi, 1997). A number of GyrA proteins from mycobacteria, including *M. leprae*, are distinguished by the presence of inteins (Fsihi *et*

al., 1996a; Kumar *et al.*, 1996) as is the DnaB helicase. *dnaG* has been identified in a locus different to the macromolecular synthesis operon in which it is found in *E. coli* (Klann *et al.*, 1998), and no rRNA operon has been identified in the *oriC* region of *M. leprae* (Fsihi *et al.*, 1996). In addition, it appears that the biologically active Class I ribonucleotide reductase (RR), an enzyme whose function is rate limiting for DNA replication, is encoded in *M. tuberculosis* by a homologue of the *E. coli mrdEF* genes which encode the RR subunits usually associated with micro-aerophilic growth (Yang *et al.*, 1997).

Alternative modes of replication become important under conditions of stress. As has been shown in *E. coli*, induction of the SOS response and entry into stationary phase result in transient modes of stable DNA replication (SDR), termed induced SDR (iSDR) and nSDR respectively (reviewed in Kogoma, 1997). nSDR is thought to have an important role in the survival of stationary phase cells, and appears mechanistically similar to the constitutive SDR (cSDR) found in cells lacking RNase HI (Hong *et al.*, 1996).

A number of enzymes involved in replication have not yet been characterised in the mycobacteria, and some *E. coli* homologues are notably absent as deduced by analysis of the genome sequence of *M. tuberculosis* (Mizrahi, pers. comm.). A homologue of *rnhA*, the gene encoding RNase HI, has not yet been annotated in the *M. tuberculosis* genome database, and there is a paucity of information about this mycobacterial homologue. The differences already noted in a number of enzymes involved in replication in the mycobacteria, together with the key role that RNase HI plays in the maintenance of replication at *oriC* under favourable conditions, and the modulation of this action that must occur to allow alternate modes of replication to become active under conditions of stress, led to us therefore to examine the presence and roles of RNase HI in the mycobacteria.

1.3 RNase HI: GENETIC AND BIOCHEMICAL STUDIES

General RNases H

RNases H are endonucleases that specifically degrade the RNA strand of an RNA:DNA hybrid to produce short oligonucleotides with 5'-phosphate and 3'-hydroxyl

groups. RNase H activity appears to be ubiquitous and is found in eukaryotes, prokaryotes and viruses (Crouch and Dirksen, 1982). This activity in prokaryotes and eukaryotes is coded for by at least two RNases H, RNase HI and RNase HII, which bear little homology to each other at the amino acid level. The eukaryotic family of RNases HI consists of large polypeptides of 68 - 90 kDa, with an RNase H domain, similar in size and amino acid sequence to *E. coli* RNase HI, fused to domains of unknown function. The RNase H activity found in retroviruses is associated with a domain of reverse transcriptase (RT) (Wintersberger, 1990). Although the primary amino acid sequence of HIV RNase H has only 20% identity to *E. coli* RNase HI, residues identified as important for catalysis are conserved (Doolittle *et al.*, 1989). RNase HI appears to constitute the primary RNase H function in prokaryotes since in the *E. coli* model, RNase HII activity is almost undetectable in comparison (Itaya, 1990). In addition to the RNases H, other enzymes have been shown to possess RNase H activity, such as *E. coli* exonuclease III (Keller and Crouch, 1972) and the 5'-3' exonuclease activity of DNA polymerase I (Kornberg and Baker, 1992), showing the high redundancy, and underscoring the importance of RNase H function to the cell.

Cellular roles of RNase HI

Although *E. coli* RNase HI has been extensively biochemically characterised, the extent of the role that this enzyme plays in the cell is not completely clear. However, the generation of *rnhA* mutants has facilitated elucidation of some of these roles.

1. Initiation at *oriC*: Initiation of replication of the chromosome in *E. coli* usually occurs at a specific site termed *oriC*. Mutants lacking RNase HI, however, have been shown to be able to initiate replication from a number of other sites in the chromosome, termed *oriK*, independently of protein synthesis or DnaA protein, but dependent on transcription and the recombinase activity of RecA (reviewed by Asai and Kogoma, 1994). This mode of replication has been termed constitutive stable DNA replication (cSDR) and allows cells to grow on minimal media at 30-40% of the wild-type rate with a two-fold decreased DNA content (von Meyenburg *et al.*, 1987). These alternative sites of initiation are thought to arise from RecA-catalysed annealing of RNA transcripts to the template

DNA strand, leading to the formation of an R-loop (Cao and Kogoma, 1993, Kogoma *et al.*, 1993). RNA polymerase has also been implicated in the production of such hybrids, as mutants of the β -subunit of RNA polymerase have been shown to alter the number of R-loops in the cell (Kogoma, 1994). RNase HI has been shown not be directly involved in replication at *oriC* (Hong and Kogoma, 1993) and thus appears only to prevent opportunistic initiation of replication at *oriK*s by degrading such non-*oriC* RNA:DNA hybrids.

2. Removal of RNA primers in lagging strand replication: Although this function appears primarily to be assigned to DNA PolI, mutant analysis has indicated that RNase HI has a minor role in this regard (Kitani *et al.*, 1985; Ogawa and Okazaki, 1984).

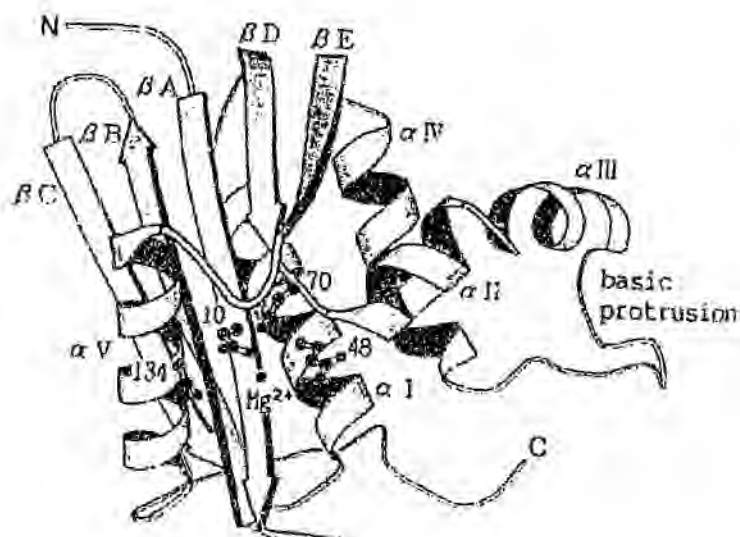
3. ColE1 plasmid replication: The formation of a persistent R-loop at the origin of replication (*ori*) of ColE1 type plasmids is essential for plasmid replication. Transcription of RNA II allows for hybridisation of this RNA to a region near the *ori*. RNase HI cleaves this hybrid and creates a primer to allow replication of the plasmid by DNA PolI (Itoh and Tomizawa, 1980).

4. msDNA synthesis: Retrons present in some strains of *E. coli* and other bacteria have been found to encode a reverse transcriptase (RT) which produces multicopy single stranded DNA (msDNA) covalently linked to RNA (reviewed in Inouye and Inouye, 1993). This RT consists only of a polymerase domain and lacks an RNase H domain. Host RNase HI is therefore necessary for correct processing of msDNA (Lima and Lim, 1995; Shimamoto *et al.*, 1995)

3-D structure of RNase HI

E. coli RNase HI is a 17.5 kDa protein consisting of 155 amino acids (Kanaya and Crouch, 1983). The three dimensional structure has been elucidated at 2Å by Yang *et al.* (1990) and at 1.48Å by Katayanagi *et al.* (1992) and is shown diagrammatically in Fig. 1. The tertiary structures of *Thermus thermophilus* RNase HI (Ishikawa *et al.*, 1993a) and the RNase H domain of HIV RT (Davies *et al.*, 1991) have also been elucidated and found to be almost completely superimposable on that of *E. coli* RNase HI. Holm and Sander (1994) proposed that the unique topology observed in RNases HI (that of a single-domain

(a)



(b)

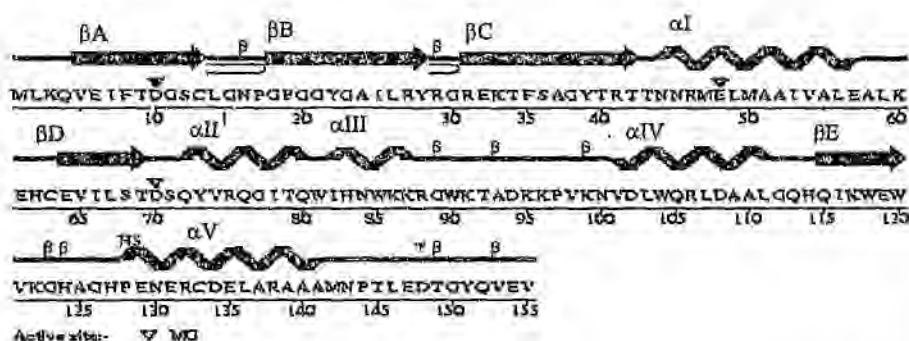


Fig. 1. a) Diagrammatic representation of *E. coli* RNase HI crystallised with Mg^{2+} (Ribbon model taken from Katayanagi *et al.*, 1993). Active site residues are numbered and the side chains involved in binding of the Mg^{2+} ion are shown. β -sheets are numbered A-E and α -helices are numbered I-V. (b) A corresponding extended ribbon diagram with *E. coli* RNase HI amino acid sequence (<http://www.biochem.ucl.ac.za/bsm/pdbsum/index.html>)

α - β protein) be classified as an RNase HI fold. It has been recently shown that the crystal structures of RuvC resolvase and HIV-1 integrase share similarities in topology to that of RNase HI (Ariyoshi *et al.*, 1994; Dyda *et al.*, 1994; Yang and Steitz, 1995) and it has been proposed that this topology might be common amongst a super-family of nucleotidyl transferases which are involved in DNA repair or replication (Akasako *et al.*, 1995). Important characteristics of the three dimensional structure include a catalytic site formed by the conjunction of Asp¹⁰, Glu⁴⁸ and Asp⁷⁰ (Kanaya *et al.*, 1990a), and a basic protrusion or handle region, containing a cluster of 7 basic amino acids, that is involved in the binding of substrate (Kanaya *et al.*, 1991c; Keck and Marqusee, 1995; Keck and Marqusee, 1996).

Substrate binding

The minimum size of RNA:DNA hybrid that has been recognised is a tetramer (Kanaya and Ikehara, 1993), but studies on 12 bp substrates suggest that the basic protrusion binds DNA residues (whether this DNA is hybridised to RNA or not) 6-7 b upstream of the site of cleavage of the RNA:DNA hybrid (Kanaya and Kanaya, 1995). These observations have been extended by Lima and Crooke (1997) who demonstrated the ability of *E. coli* RNase HI to cleave ssRNA extending 3' from an RNA:DNA hybrid. It has been shown that binding of substrate and specificity of cleavage are separable (Stahl *et al.*, 1994; Keck and Marqusee, 1996). *E. coli* RNase HI binds RNA:RNA and DNA:DNA duplexes in addition to RNA:DNA hybrids (Oda *et al.*, 1993a), but catalytic activity is specific for RNA:DNA hybrids (Berkower *et al.*, 1973; Crouch and Dirksen, 1982).

Catalytic site

A single divalent cation was found to be required for catalysis (Uchiyama *et al.*, 1994a; Katayanagi *et al.*, 1993b) in contrast to the two-metal ion binding observed with HIV-RT (Davies *et al.*, 1991). Mg²⁺, Mn²⁺, Co²⁺, Cd²⁺ and Zn²⁺ are all able to catalyse cleavage of the phosphodiester bond although Mn²⁺ and Zn²⁺ appeared to facilitate cleavage to a much lesser extent (Uchiyama *et al.*, 1994a). A structural role for the metal ion, for example the re-orientation of active site residues as suggested by Katayanagi *et al.* (1993b), has been excluded for this metal cofactor and its role in catalysis is proposed to

be the stabilisation of the transition state of the enzyme by hydrogen bonding with waters of solvation (Black *et al.*, 1996). The roles of the residues forming the catalytic site have been probed: Asp¹⁰ is suggested to be either the proton donor to the RNA moiety or the binding partner of the Mg²⁺ cofactor and Asp⁷⁰ is thought to be the proton acceptor from the water molecule (Oda *et al.*, 1994). Crystallisation of the Mg²⁺ bound enzyme indicated that the Mg²⁺ ion is in a position to be co-ordinated by the carboxyl group of Glu⁴⁸ (Katayanagi *et al.*, 1992; Katayanagi *et al.*, 1993b). Although His¹²⁴ and Asp¹³⁴ initially appeared not to be directly involved in catalysis (Kanaya *et al.*, 1990a), a role for these residues in the catalytic activity of the enzyme has subsequently been proposed; His¹²⁴ is proposed to change its conformation in the catalytic reaction and remove a proton from a catalytically essential carboxylate – the carboxyl oxygen at the delta 1 position of Asp¹³⁴ (Oda *et al.*, 1993b; Haruki *et al.*, 1994). In this scenario, the roles of Asp¹⁰ and Glu⁴⁸ would therefore be to bind Mg²⁺ (Oda *et al.*, 1994). However, Kanaya *et al.* (1996) have subsequently proposed that the negative repulsion between Asp¹⁰ and Asp⁷⁰ is responsible for the Mg²⁺ binding and that Glu⁴⁸ therefore functions to anchor a water molecule. The enzyme, and not the Mg²⁺ ion, is proposed to interact with the phosphate residue at the cleavage site (Uchiyama *et al.*, 1994b). The nucleoside 3' to the cleaved phosphodiester bond was found to act both as a proton donor and a proton acceptor, and the second nucleoside 5' to the cleaved bond was found to act as a proton donor in the enzyme-substrate complex (Iwai *et al.*, 1995).

Regulation of rnhA

In *E. coli* and *S. typhimurium*, *rnhA* and *dnaQ* (the gene encoding the ϵ (proofreading) subunit of DNA PolIII) share an overlapping divergent promoter region (Horiuchi *et al.*, 1981). Separation of the *rnhA* promoter region from the *dnaQ* region resulted in higher levels of transcription of *rnhA*, suggesting promoter interference between these genes (Nomura *et al.*, 1985). Bialy and Kogoma (1986) found no change in RNase HII levels of cells induced for the SOS response with nalidixic acid or UV-irradiation. However, Casaregola *et al.* (1987) observed a *recA*-dependent inhibition of *rnhA* transcription during UV-induced SOS response. Quinones *et al.* (1987) also noted

that transcription of *rnhA* is unaffected in *recA441* strains induced for the SOS response, but found levels of transcription of *rnhA* are reduced in *lexA(Def)* mutants which are constitutively derepressed for the SOS response, and are almost undetectable in *lexA recA441* strains. Foster and Marinus (1992) did not observe such a dramatic effect on transcription levels of *rnhA* with strains transiently induced for the SOS response however, but noted that mutations in the chaperone proteins DnaJ and DnaK, which are also induced as part of the SOS response, altered levels of RNase HI activity in the cells. These authors postulated a post-translational regulation, either directly or indirectly, of levels of RNase HI by these proteins.

Phenotypes of mutants with altered levels of RNase HI

No effect on plasmid maintenance or on growth rate was observed with a 20-fold over-expression of RNase HI (Kanaya and Crouch, 1984). Overproduction of RNase HI, however, was shown to result in an increased sensitivity to UV irradiation (Bockrath *et al.*, 1987) and inhibition of induced mutagenesis (Foster *et al.*, 1989).

As noted previously, null mutants of *rnhA* in *E. coli* have been shown to exhibit cSDR. *rnhA* mutants also constitutively express the SOS response due to the persistence of the single stranded (ss) DNA arising as DNA is displaced by the formation of R-loops (Kogoma *et al.*, 1993). Homologous recombination at certain sites in the *E. coli* chromosome is also enhanced in *rnhA* mutants (Nishitani *et al.*, 1993). In addition, although ColE1 type plasmid replication is still supported in such mutants, plasmid multimers are readily formed (Fukuoh *et al.*, 1997).

A number of mutants in conjunction with *rnhA* in *E. coli* have been shown to exhibit lethal phenotypes:

1. A combination of *recG* and *rnhA* is lethal (Hong *et al.*, 1995). RecG is a helicase that specifically disociates synthetic X and Y junctions and is thought to catalyse branch migration of Holliday junctions. *recG* mutants also exhibit cSDR and induced SDR (iSDR), and chronically express the SOS response due to the non-resolution of Holliday junctions which block the replication fork. It has been subsequently been shown that RecG is able to resolve R-loops in an ATP-dependent manner (Fukuoh *et al.*, 1997).

Interestingly, a *recG-dna4* double mutant is inviable (Hong *et al.*, 1995) whereas a *dna4-rnhA* double mutant is not. These authors suggest that the cSDR activity seen in *recG* mutants (half that seen in *rnhA* mutants) may not be sufficient for survival. This implies that removal of R-loops is an essential function in the cell.

2. A combination of *recB* or *recC* and *rnhA* is lethal (Itaya and Crouch, 1991a; Kogoma *et al.*, 1993). *recB270*(Ts), or *recC271*(Ts), in conjunction with *rnhA-339::cat* renders *E. coli* temperature-sensitive for growth. Lethality due to a defect in the repair function of these enzymes was implicated by complementation studies (Itaya and Crouch, 1991a). In contrast to these observations, the recombination function (helicase activity) of RecBCD was shown to be essential for survival as *rnhA recD* mutants were viable whereas a *recB* mutation that inactivates both the exonuclease V and recombination function of RecBCD was lethal in conjunction with *rnhA::cat* (Kogoma *et al.*, 1993). It has been shown that complementation of a *recG-rnhA* double mutant requires more RNase HI than does complementation of a *recB-rnhA* double mutant, implying that *recG* mutants require more RNase HI activity for survival than do *recB* mutants (Hong *et al.*, 1995). This is perhaps not surprising as it has been shown that 0.1% of wild-type RNase HI activity is sufficient to alleviate the temperature sensitive phenotype of *recB(ts)* mutants, while 10% of the same is not sufficient to alleviate cSDR (Itaya and Crouch, 1991b).

3. A combination of *rnhA* and *polA* is lethal (Kogoma *et al.*, 1993; Hong *et al.*, 1995). A *polA12* mutant is deficient in nick translation activity at 42°C (Monk and Kinross, 1972) and *rnhA-polA12* mutants exhibit temperature sensitive growth. The lethal combination of *rnhA-polA* can be suppressed by a *lexA*(Def) mutation that inactivates the LexA repressor (Kogoma and Maldonado, 1997).

4. A combination of *topA* and *rnhA* is lethal. (Drolet *et al.*, 1995). Topoisomerase I (encoded by *topA*) relaxes negatively supercoiled DNA and has been shown to suppress R-loop formation during transcription. Overexpression of RNase HI can partially compensate for a *topA* mutant, as can a defect in DNA gyrase. However, a combination of *rnhA* and *topA* is lethal.

5. A combination of *sbcB15* and *rnhA* retards growth (Itaya and Crouch, 1991a). *sbcB* encodes exonuclease I, an enzyme that hydrolyses ssDNA in a 3'-5' direction and has been shown to comprise the major DNA deoxyribophosphodiesterase activity in *E. coli* (Sandigursky and Franklin, 1992). This enzyme has been implicated in the DNA repair and recombination in pathways mediated by RecBC (ExoV). An *sbcB15* mutation that reduces (but does not completely eliminate) the exonuclease activity of this enzyme was found to dramatically affect plating efficiency of *rnhA* mutants.

From this mutant data, it appears that RNase HI plays an important, albeit biochemically redundant role in *E. coli*, specifically in the resolution of R-loops. Whether an analogous situation exists in the mycobacteria remains to be investigated.

1.4 AIMS OF THIS STUDY

In light of the roles of RNase HI in *E. coli*, and the importance of further analysis of enzymes involved in replication and repair in the mycobacteria, it was therefore of interest to this author to investigate the role(s) of RNase HI in *Mycobacterium smegmatis*.

The aims of this work therefore were to :

1. isolate and analyse the *rnhA* gene from *M. smegmatis*,
2. characterise the gene product biochemically and
3. investigate the role of RNase HI in the mycobacteria by analysing a knockout mutant of *rnhA* in *M. smegmatis*.

2. MATERIALS AND METHODS

2.1 Reagents and manufactured items

Reagents and suppliers are detailed in Appendix A. Solutions specific to certain procedures are detailed within the protocol, otherwise media composition and general solutions are detailed in Appendix B. A list of abbreviations can be found in Appendix C.

2.2 Bacterial strains and growth conditions

All strains used in this study were stored at -70°C in 50% glycerol. A list of the strains used is detailed in Table I.

Table I: Bacterial strains

STRAIN	GENOTYPE/MARKERS	REFERENCE/SOURCE
<i>E. coli</i> strains		
DH5 α	<i>supE44 ΔlacU169 (Φ80 <i>lacZ</i>ΔM15) <i>hsdR17 recA1 endA1 gyrA96 thi-1 relA1</i></i>	Hanahan (1983)
JM101	<i>supE thi-1 Δ(<i>lac-proAB</i>) F' [<i>traD36 proAB⁺ lacI^f lacZ</i>ΔM15]</i>	Yanisch-Perron <i>et al.</i> (1985)
CJ236	<i>dut1 ung1 thi-1 relA1 (pCJ105[<i>cam^r</i>F'])</i>	Kunkel <i>et al.</i> (1987)
GM161	<i>dam-4 supE44 hsdS1</i>	E. R. Dabbs [†]
TB1	<i>ara Δ(<i>lac-proAB</i>) rpsL (Φ80 <i>lacZ</i>ΔM15) <i>hsdR</i></i>	NEB, Inc. Beverly, MA Johnston <i>et al.</i> (1986)
MIC3037	<i>rnh-339::cat recC271 (Ts) F⁺ supE44 supF58 lacY1 trpK55 galK2 galT22 meiB1 hsdR14(<i>r_k⁻ m_k⁺</i>)</i>	Itaya and Crouch (1991a,b)
<i>M. smegmatis</i> strains		
mc ² 155	A highly transformable mutant of mc ² 6	Snapper <i>et al.</i> (1990)
LR222		TB Unit [§]

[†] E. R. Dabbs, Genetics Department, University of the Witwatersrand, PO WITS Johannesburg, 2000.

[§] Department of Microbiology, S.A.I.M.R. Central, PO Box 1038, Johannesburg, 2000.

Escherichia coli : All *E. coli* strains except JM101 were recovered by streaking onto Luria-Bertani agar (LA) plates and were incubated overnight at 37°C. All strains were propagated at 37°C except MIC3037 which was propagated at 30°C. Liquid cultures of *E. coli* were grown in Luria-Bertani broth (LB) with vigorous shaking (350-400 rpm) in a New Brunswick Series 25 Incubator Shaker. JM101 was propagated on minimal media (MM), and media for CJ236 included chloramphenicol at 25 µg/ml, both to select for the F⁺ episome.

Mycobacterium smegmatis : *M. smegmatis* strains were recovered by streaking onto Middlebrook agar (MA) plates (Jacobs *et al.*, 1991) or LA plates and were incubated for 3-4 days. *M. smegmatis* liquid cultures were grown with gentle shaking (100 rpm) in a New Brunswick Innova 4000 Incubator Shaker in Middlebrook Tween medium (M-TW) (Jacobs *et al.*, 1991) or in LB.

2.3 Plasmids and Bacteriophages

Plasmids were maintained frozen in the appropriate host at -70°C. A list of plasmids used is detailed in Table II. Strains were recovered as described in Section 2.2. Selective pressure for the maintenance of plasmid was applied by the inclusion of antibiotic at the following concentrations: *E. coli*: ampicillin (Amp) 100 µg/ml, kanamycin (Km) 50 µg/ml, hygromycin B (Hyg) 200 µg/ml; and *M. smegmatis* : Km 10 µg/ml and Hyg 50 µg/ml. 5% sucrose was included in media for negative selection of clones containing the *sacB* gene, and 50 µl IPTG (100 mM in water) and 50 µl X-gal (20 mg/ml in dimethyl formamide) were spread on plates or included in top agar (0.9 % NaCl/0.75 % agar) for blue/white colour selection when necessary. Plasmid DNA was isolated as described in Section 2.6. After purification, plasmid DNA was stored in water at -20°C. M13 phage was grown and prepared as described in Section 2.6. Phage constructs were stored either as infectious particles in 0.22 µm filter-sterilised supernatant at 4°C, or as replicative form (RF) DNA in water at -20°C.

Table II: *E. coli* and *E. coli*-*M. smegmatis* shuttle plasmids

PLASMID	REMARKS	SOURCE/REF.
<u><i>E. coli</i> plasmids</u>		
pGEM-3zf(+)	<i>bla lacZ</i>	Promega, Madison, WI
pMAL-C2	<i>bla malE lacZ</i>	NEB, Beverly, MA
M13mp18/19 (RF)	<i>lacZ'</i>	Amersham, UK
pUS	<i>sacB</i> cloned as a <i>Pst</i> I cassette into pBluescript (Stratagene, La Jolla, CA)	P. Stolt, unpublished
pSD1	3.6 kb <i>Pst</i> I fragment containing 350 bp of <i>rnhA</i> and upstream sequence cloned into pGEM3zf(+).	This work
pSD2	2.3 kb <i>Bam</i> HI fragment containing 471 bp of <i>rnhA</i> and downstream sequence cloned into pGEM3zf(+).	This work
pBA	pSD2 with the <i>aph</i> cassette cloned into the <i>Xho</i> I site in <i>rnhA</i>	This work
pPA	pSD1 with the <i>aph</i> cassette cloned into the <i>Bgl</i> II site upstream of <i>rnhA</i>	This work
pPBA	<i>Bam</i> HI fragment from pBA cloned into pSD1.	This work
pPAB	<i>Bam</i> HI fragment from pSD2 cloned into pA.	This work
pPBAS	The <i>Eco</i> RI fragment from pPBA cloned into pUS.	This work
pPABS	The <i>Eco</i> RI fragment from pPAB cloned into pUS.	This work
pUSH	<i>Hyg</i> cloned as a <i>Bam</i> HI- <i>Bgl</i> II cassette into the <i>Bam</i> HI site of pUS.	This work
pPS	720 bp <i>Pst</i> I- <i>Sac</i> I subclone in pGEM3Zf(+) of pSD1	This work
pPS- δ B	pPS with the <i>Bam</i> HI site blunt-ended and religated	This work
pUSH- δ B	The 720 bp <i>Eco</i> RI- <i>Hind</i> III fragment from pPS- δ B cloned into the <i>Sma</i> I site of pUSH	This work
pMAL::RNH	<i>rnhA</i> gene cloned into pMAL-C2 as a translational fusion with the <i>malE</i> gene.	This work
<u><i>E. coli</i>-<i>M. smegmatis</i> shuttle plasmids</u>		
pRCX3	<i>aph xylE</i>	Curcio <i>et al.</i> (1994)
pOLYG	pBluescript polylinker cloned into the <i>Kpn</i> I site of p16RI (Garbo <i>et al.</i> (1994))	P. O'Gaora, unpublished
pRNH350	<i>rnhA</i> , truncated at the <i>Pst</i> I site, cloned with 350 bp of promoter region as a transcriptional fusion with the <i>xylE</i> gene in pRCX3.	S. Durbach
pRNH112	<i>rnhA</i> with 112 bp of promoter region cloned as per pRNH350	This work
pRNH82	<i>rnhA</i> with 82 bp of promoter region cloned as per pRNH350	This work
pRNHA	The <i>Bam</i> HI- <i>Bgl</i> II promoter region plus the <i>Bam</i> HI- <i>Hind</i> III <i>rnhA</i> cassette from pMAL::RNH cloned into <i>Bam</i> HI- <i>Hind</i> III pOLYG.	This work
pRNH- <i>lacZ</i>	The <i>Bam</i> HI- <i>Bgl</i> II promoter region containing the first three codons of <i>rnhA</i> , cloned in frame with <i>lacZ</i> into the <i>Bam</i> HI site of pOLYG.	This work

2.4 Transformation of bacteria by electroporation

Preparation of electrocompetent *M. smegmatis* (Jacobs *et al.*, 1991): M-TW medium (100 ml) was inoculated with 1 ml of a stationary phase culture of mc²155 and incubated with gentle shaking at 37°C until an OD₆₀₀=0.9-1.5 was reached. The cells were chilled on ice for 15 min. The cells were harvested at 1500 xg for 5 min at 4°C in a pre-chilled JA20 rotor (Beckman). The pellet was washed in 100 ml ice cold 1% glycerol and re-harvested. Subsequent wash volumes were 50 ml, 25 ml, 10 ml and 2 ml with a final resuspension in 400 µl 1% glycerol. Cells were dispensed into 80 µl aliquots and either flash frozen at -70°C, or used immediately for electroporation. Frozen cells were thawed on ice before use.

Preparation of electrocompetent *E. coli* (as detailed in the Bio-Rad Gene Pulser manual): LB medium (50 ml) was inoculated with 0.5 ml of an overnight culture of *E. coli* DH5α. The culture was grown with vigorous shaking until an OD₆₀₀ = 0.6 was reached. The cells were chilled on ice for 15 min. The cells were harvested at 1500 xg for 5 min. The cells were re-suspended gently in 50 ml of ice-cold 10% glycerol. The cells were harvested as before and washed again in 25 ml and then 1 ml of 10% glycerol. The cells were re-suspended in a final volume of 150 µl 10% glycerol. Cells were dispensed in 40 µl aliquots and either flash frozen on dry ice, or used immediately. Frozen cells were thawed on ice before use.

Electroporation of DNA : Electroporation of competent cells was carried out using the Bio-Rad Gene Pulser™ (Bio-Rad Laboratories, Hercules, CA) and 0.2 cm electrode gap cuvettes from BTX (San Diego, CA) for *E. coli* DH5α cells, and 0.2 cm electrode gap cuvettes from Bio-Rad Laboratories for *M. smegmatis* mc²155. Settings and recovery procedures were as shown in Table III. Transformations of mc²155 with non-replicating vectors used 100 ng -1 µg of vector DNA, while transformations of mc²155 with replicating vectors typically used 10-20 ng. *E. coli* DH5α cells were transformed with 10-20 ng of a ligation reaction, or a vector-containing colony of mc²155 was mixed with the competent *E. coli* cells; the plasmid was then electroduced from the mc²155 cells into the *E. coli* cells under *E. coli* electroporation conditions (Baulard *et al.*, 1992).

Table III: Electroporation parameters

VECTOR	HOST	SETTINGS	RECOVERY MEDIUM	RECOVERY	VOLUME PLATED
non-replicating	mc ² 155	1000Ω, 35μF, 2.5 kV	1 ml M-TW	2.5 h, 37°C	1 ml
replicating	mc ² 155	1000Ω, 35μF, 2.5 kV	1 ml M-TW	2.5 h, 37°C	100 μl
replicating	DH5α	200Ω, 35μF, 2.5 kV	1 ml 2xTY	1 h, 37°C	5 μl

2.5 Chemical transformation of *E. coli* (Sambrook *et al.*, 1989)

Strains DH5α, TB1, MIC3037 and GM161: LB (50 ml) was inoculated with 500 μl of an overnight culture. The cells were grown with vigorous shaking to OD₆₀₀ = 0.4. The cells were chilled on ice for 15 min and then harvested at 1500 xg for 5 min at 4°C. The cells were gently re-suspended in 10 ml ice cold 0.1 M CaCl₂ and incubated on ice for 1 h. The cells were re-harvested, and re-suspended in 1 ml 0.1 M CaCl₂. Three microlitres (20 ng) of a ligation reaction were added to a 50 μl aliquot of the cells and the mixture was incubated for a further 20 min on ice. The cells were then heat-shocked at 42°C for 90 s. LB (200 μl) was added and the cells were allowed to recover for 1 h at 37°C without shaking, before plating on selective media.

JM101 : LB (50 ml) was inoculated with 500 μl of an overnight culture. The cells were grown with vigorous shaking to an OD₆₀₀ = 0.4. The cells were chilled on ice for 15 min and then harvested at 1500 xg for 5 min at 4°C. The cells were gently re-suspended in 10 ml of ice cold 0.1 M MgCl₂. The cells were re-harvested and gently re-suspended in 10 ml of ice cold 0.1 M CaCl₂ and incubated on ice for 1 h. The cells were re-harvested, and re-suspended in 1 ml 0.1 M CaCl₂. Typically 20 ng of a ligation reaction was added to a 50 μl aliquot of the cells, and the mixture was incubated for a further 20 min on ice. The cells were then heat-shocked at 42°C for 90 s. Four volumes of a late log-phase lawn culture (OD₆₀₀ = 1.0) was added to the cells and the mixture was transferred to 3 ml of λ phage agar at 56°C and IPTG and X-gal were added as described in Section 2.3 if blue/white colour selection was required. The cells were distributed evenly

in the agar by vortexing briefly and the agar was poured onto LA plates. After the top agar had solidified, the plates were incubated overnight at 37°C.

2.6 DNA extractions

Rapid isolation of *E. coli* plasmids (Sambrook *et al.*, 1989): Colonies were picked and inoculated into 1.5 ml of LB containing the appropriate selective agent. Cultures were grown to stationary phase and harvested at 4°C in microfuge tubes by centrifugation at 10 000 xg for 5 min. The pellets were re-suspended in 100 µl of solution I (50 mM Tris-HCl, pH 8.0/10mM EDTA) and the cells were lysed on ice for 5 min after the addition of 200 µl of solution II (0.2 M NaOH/1% SDS). The mixture was neutralised by the addition of 150 µl solution III (3 M KOAc, pH 5.5) and the mixture was allowed to incubate at room temperature for 5 min. The mixture was centrifuged for 10 min at 4°C and the supernatant retained. RNase A was added to a final concentration of 100 µg/ml and the mixture was incubated at 42°C for 5 min. The supernatant was extracted once with 400 µl 1:1 phenol:chloroform and once with 400 µl 24:1 chloroform:isoamylalcohol. DNA was precipitated by the addition of 1ml 98% ethanol and pelleted by centrifugation at 10 000 xg for 20 min at 4°C. The pellet was washed with 1 ml 70% ethanol, vacuum-dried, and re-suspended in 20 µl water.

Chromosomal DNA from *M. smegmatis* (De Wit *et al.*, 1990): Colonies were inoculated into 15 ml culture tubes containing 2 ml LB with or without selective agent and grown to early stationary phase. Cells were harvested by centrifugation in microfuge tubes at 10 000 xg for 5 min at 4°C. The cell pellets were thoroughly re-suspended in 500 µl TES (10 mM Tris-HCl pH8.0/0.1 mM EDTA/150 mM NaCl) by pipetting. The cells were incubated at 70°C for 30 min, phenol (600 µl, pH 8) and 90 µl 10% SDS were added to each tube and they were incubated horizontally with shaking at 37°C for 3 h. The lysis mixture was then centrifuged at 10 000 xg for 5 min at room temperature and the supernatant retained. The supernatant was then extracted twice with 1 ml 24:1 chloroform:isoamylalcohol. Chromosomal DNA was precipitated by the addition of 1 ml 98% ethanol and pelleted by centrifugation at 10 000 xg at 4°C for 10 min. The DNA was vacuum dried and re-suspended in 50 µl water. An equal volume of 2.5 M NaCl/20%

PEG 6000 was added and the DNA was re-precipitated on ice for 10 min. The DNA was recovered by centrifugation at 10 000 xg at 4°C for 10 min and the pellet was washed with 1 ml 70% ethanol. The pellet was dried as before and re-suspended in 30 µl water. DNA concentrations ranged from 0.1-1 µg/µl.

Single stranded (ss) DNA from M13 phage (Yanisch-Perron *et al.*, 1985):

Bulk preparation of ssDNA involved the preparation of multiple 1.5 ml aliquots of supernatant in the same fashion that single plaques were screened; 1.5 ml of LB was seeded with 0.1 volumes of a stationary phase culture of JM101 and inoculated either with a plaque picked from a fresh plate, or with 1 µl infectious supernatant (10^{12} pfu/ml). Cultures were grown at 37°C with vigorous aeration for 5 h. The cells were pelleted by centrifugation at 10 000 xg at 4°C for 5 min. Supernatant containing the ssDNA was drawn off immediately to minimise contamination of the preparation by cellular debris. The phage was precipitated from the supernatant by the addition of 200 µl 20% PEG 6000/2.5 M NaCl and incubation on ice for 10 min. Phage was pelleted by centrifugation at 10 000 xg for 10 min at 4°C. The supernatant was drawn off and the pellet re-suspended in 100 µl TE. After incubation on ice for 5 min, the mixture was centrifuged for 5 min at 4°C to pellet cell debris. The supernatant was recovered and NaOAc (pH 6.5) was added to give a final concentration of 0.3M. The mixture was extracted once with 250 µl 1:1 phenol:chloroform and once with 50 µl chloroform. The DNA was ethanol precipitated and re-suspended in 15 µl of water. The replicative form of the phage was prepared from the cells following the protocol for plasmid DNA, although the preparation contained a mixture of ss and dsDNA.

Uracil-enriched ssDNA from M13 phage (Kunkel *et al.*, 1987): *E. coli* strain CJ236 was streaked to single colonies and grown overnight on LA containing Cm at 25 µg/ml. A single colony was picked and inoculated into 10 ml of 2xTY and grown for 7 h. 1.5 ml of this culture was used to seed 20 ml of 2xTY containing 35 µg/ml Cm, 1 µg/ml uridine and an inoculum of infectious supernatant of the appropriate M13mp19 construct. The cells were grown overnight and the uracil-rich ssDNA was recovered by the usual ssDNA preparation protocol.

2.7 DNA manipulations

Digestions with restriction enzymes : Plasmid DNA was digested in a total volume of 20 μ l of the appropriate buffer supplied with the enzyme. The digestion was incubated at 37°C for 1 h, unless otherwise directed by the manufacturer. If necessary, 1 U of calf intestinal phosphatase was added to the reaction and the reaction was incubated for a further 45 min. The reaction volumes were made up to 100 μ l and NaOAc (pH 6.5) was added to give a final concentration of 0.3M. The enzymes were denatured by extraction with 1:1 phenol:chloroform and the DNA was ethanol-precipitated and re-suspended in water. Mycobacterial chromosomal DNA was digested in a total volume of 30 μ l at 37°C for 3 h or overnight.

Blunt ending of 5'-overhangs : Linearised DNA was re-suspended in 20 μ l of 1 x Klenow reaction buffer and 200 μ M dNTPs. One unit Klenow enzyme was added and the reaction was incubated at 37°C for 45 min.

Polymerase chain reaction (PCR): Typically, amplification of chromosomal DNA using degenerate primers was as follows: 1 μ M each of forward and reverse primers, 200 μ M dNTPs, 100-200 ng chromosomal DNA and 0.4 U Taq DNA polymerase in 50 μ l of supplied buffer was subjected to cycle amplification in a Hybaid Omnigene Temperature Cycler under the following conditions: 96°C for 5 min, 30 cycles of (96°C for 1 min, 50-60°C for 1 min, 72°C for 1 min) and then 72°C for 5 min. The ramping rate was set at 0.01 °C/s. Primer concentrations and annealing temperatures were varied with different degenerate primer sets and with specific primer pairs to reduce non-specific amplification. Primers used for PCR are detailed in Table IV.

Polishing of PCR products : The PCR reaction mixture was heated to 96°C and slowly cooled to 37°C over 30 min. $MgCl_2$ to a final concentration of 5 mM was added together with 1 U of Klenow enzyme. The reaction mixture was incubated at 37°C for 45 min. The polished product was recovered either by ethanol precipitation or from a low melting agarose (LMA) gel.

Oligodeoxyribonucleotide-directed mutagenesis (Kunkel *et al.*, 1987): The primer PROM-F1 (Table IV) containing the mismatch was 5'-phosphorylated as detailed and then heat inactivated at 65°C for 10 min. The following hybridisation and

and then heat inactivated at 65°C for 10 min. The following hybridisation and polymerisation reactions utilised buffers and protocols from the BioRad Mutagen Kit: 5'-phosphorylated oligonucleotide (18 ng) was hybridised to 0.5 µg of appropriate uracil-rich ssDNA in 5 µl of 1 x Annealing Buffer by heating the reaction in a 65°C water bath and allowing it to slowly cool to 37°C. 0.6 µl of 10 x Synthesis Buffer, 0.2 U T7 DNA polymerase and 0.2 U of T4 DNA ligase were added to make the reaction mixture up to 6 µl. The reaction was incubated for 3 min at room temperature and at 37°C for 90 min. The entire reaction mix was used to transform competent JM101. Plaques were screened by restriction analysis of (RF)DNA preparations.

Table IV: Synthetic oligodeoxyribonucleotides

NAME	SEQUENCE (5' - 3')	FEATURES *
PCR/Sequencing		
SGREX-F1	gtgagccaggatcccgatc	404-424, <i>Bam</i> HI site (underlined), sense (s)
SGREX-R1	gggggaagcctacggagcgaagtaccgc	862-880, <i>Hind</i> III site (underlined), G-clamp, antisense (as),
SGREX-R2	cccgggaagcctcaacggagcgaagtaccgc	864-883, <i>Hind</i> III site (underlined), G/C-clamp, stop codon (bold), as.
SGREX-R3	aagggaagcctcaacggagcgaagtaccgc	862-883, <i>Hind</i> III site (underlined), G-rich clamp, stop codon (bold), as.
SGRNH-R1	tcaaggcctcgagcgcca	570-587, as
SGRSEQ-R3	tggcactcgaatgatgaa	1024-1042, as
Male	ggcgcgcagactgctgatgaagcc	As NEB catalogue #1237, s.
Site-directed mutagenesis		
STEF-1	ggcccagatctcagcgcccgcgcgtcggg	318-341, <i>Bgl</i> II site (underlined), G/C clamp, s.
PROM1-F1	agcgcggagatctcagcgc	284-302, <i>Bgl</i> II site (underlined), mismatch (bold), s
S1-protection		
S1-BAM	gcacctgctcagctc	400-417, <i>Bam</i> HI site (underlined), start codon (bold), as

* Numbering indicates position of primer with respect to the sense strand and corresponds to that of L11918.

Recovery of DNA from LMA gels (as directed by Boehringer-Mannheim): DNA was digested and fractionated on a LMA agarose gel. The band of interest was excised and weighed. An appropriate amount of 25 x agarase buffer was added to the gel slice and the agarose liquified by incubation at 65°C for 15 min. The liquified agarose was transferred to 45°C, 1 U agarase was added and the reaction incubated for 45 min. Oligosaccharides were precipitated by the addition of 0.1 volumes of 3M NaOAc pH 6.5, incubation on ice for 5 min and centrifugation at 10 000 xg at 4°C for 5 min. DNA was ethanol-precipitated and re-suspended in water.

Recovery of labeled RNA from polyacrylamide gels (PAG) (D'Alessio, 1992): RNA was electrophoresed as described in Section 2.11. The band of interest was localised by autoradiography and excised from the gel. The gel slice was re-suspended in DEPC-treated water and incubated at 37°C for 15 min to leach salts from the gel. The gel slice was transferred to 1 ml PAG elution buffer (0.5 M NH₄OAc/10 mM MgOAc/1 mM EDTA/0.1% SDS) and incubated overnight at 37°C. The RNA was ethanol-precipitated, washed with 70% ethanol and stored lyophilised at -20°C until needed.

2.8 *In vitro* transcription

SP6/T7 run-off transcription (Mizrahi, 1989): The DNA substrate was linearised with an enzyme that cut downstream of the requisite transcription start, extracted with 1:1 phenol:chloroform and ethanol-precipitated. The run-off transcription reaction was done in a total volume of 30 µl reaction buffer containing substrate DNA at 200 ng/µl, 500 µM each of ATP, CTP and GTP, 500 µM UTP, 10mM DTT, 1 U rRNasin® and 60 U SP6 RNA polymerase or 60 U T7 RNA polymerase and was incubated at 40°C for 2 h. DNase I (1 U) was added and the reaction was incubated a further 15 min at 37°C.

Transcription with *M. smegmatis* RNA polymerase (Levin and Hatfull, 1993): Template DNA (3-5 µg) was incubated with 1 µl *M. smegmatis* RNA polymerase (kindly provided by G. Hatfull) and 1 U rRNasin in the presence of 40 mM Tris-HCl pH8/10 mM MgCl₂/0.1 mM EDTA/0.1 M KCl/400 µM dNTPs/2 mM DTT/100 µg/ml BSA at 37°C for 30 min.

2.9 Preparation of labeled nucleic acid

a) Probes

5'-phosphorylation of oligodeoxyribonucleotides: One picomole of oligonucleotide was 5'-phosphorylated in a 10 μ l reaction volume containing 30 μ Ci [γ^{32} P] ATP and 1 U polynucleotide kinase (PNK) in addition to the supplied buffer. The reaction was incubated at 37°C for 45 min. Oligonucleotides 5'-phosphorylated with unlabelled ATP were incubated as above in a 20 μ l volume containing 1mM ATP, 30 μ M oligonucleotide and 1 U PNK in addition to the supplied buffer.

Random-primed labelling: DNA to be labeled was re-suspended in 9 μ l of water and heat-denatured at 96°C for 10 min. The DNA was snap-cooled on ice, and then labeled with [α^{32} P] dCTP using the Random Primed Labeling Kit and the protocol suggested by the manufacturers. After incubation of the reaction mix at 37°C for 30 min, the reaction was stopped by the addition of 50 μ l TE. The labeled DNA was separated from unincorporated nucleotides by a Sephadex spin column: A one ml G-25 or G-50 Sephadex spin column was poured in 1 ml syringe plugged with glass wool. The column was equilibrated by the addition of 70 μ l of TE to the column and centrifugation of the column at 4000 rpm for 3 min in an IEC clinical centrifuge until the volume of TE recovered matched the input volume. The labeling reaction was then applied to the spin column and centrifuged as above. The flow-through was collected, and the labeled DNA was stored at -20°C.

Primer extension for S1 protection assays: Five picomoles of 5'-phosphorylated oligonucleotide in 5 μ l phosphorylation reaction mix was incubated at 65°C for 10 min to heat-inactivate the PNK, and added to a reaction mix containing 5-10 μ g appropriate ssDNA, 100 μ M dNTPs in Klenow reaction buffer. The mixture was again heated to 65°C and slowly cooled to room temperature to allow hybridisation of the oligonucleotide to the ssDNA. Klenow enzyme (5 U) was added and the reaction incubated at 37°C for 30 min. A Sephadex G-50 column was poured in a 1ml syringe plugged with glass wool. The reaction mix was applied to the column and high molecular weight product was isolated by the collection of 100 μ l fractions and analysis of these

fractions on a 6% polyacrylamide gel. The details of the primer used for the generation of probe for the S1 protection assay are shown in Table IV.

b) Substrates for RNase and DNase assays

Labeled RNA: Labeled RNA was generated by run-off transcription using either SP6 or T7 RNAP as described in Section 2.8, except that 50 μ Ci [α - 32 P] UTP was substituted for unlabeled UTP in the reaction mix. The labeled RNA transcript was separated from the nucleotides by electrophoresis of the products on a 6% polyacrylamide gel, visualisation of the labeled products by autoradiography and elution of the desired band as described in Section 2.7. Unlabeled transcript was ethanol-precipitated, re-suspended in DEPC-treated water and stored at -20°C.

RNA-DNA and RNA-RNA hybrid formation (Mizrahi, 1989): Labeled RNA was hybridised with the appropriate M13 ssDNA or RNA as follows: 10 μ l of labeled RNA was added to an excess of ss substrate (typically 5-10 μ g) in hybridisation buffer (37 mM Tris-HCl pH 8.0/37 mM KCl/3.7 mM MgCl₂). The reaction mix was heated to 65°C and slowly cooled to room temperature. Formation of hybrid was monitored by electrophoresis of the products on a 5% non-denaturing polyacrylamide gel.

2.10 DNA Sequencing (Sanger *et al.*, 1977):

ssDNA prepared from appropriate M13 phage sub-clones was used for most sequencing reactions. ssDNA (2-5 μ g) was used for each sequencing reaction, and dGTP and dITP reactions were run in parallel to resolve pauses and compressions respectively. The Sequenase DNA Sequencing Kit was used throughout and all reactions were performed according to the manufacturer's directions. The -40 primer supplied with the kit was used for all initial sequencing of sub-clones, but occasionally lack of suitable restriction sites necessitated design of primers for some sequencing reactions (Table IV). One sequencing reaction utilised denatured plasmid DNA as a template (Yie *et al.*, 1993): 5 μ g plasmid DNA was denatured in 25 μ l TE by the addition of 15 μ l 2 N NaOH and 110 μ l water and incubation of the mixture at 96°C for 3 min. The mixture was transferred to an ice

bath and NaOAc (pH5.2) was added to give a final concentration of 0.3 M. Ethanol was added to give a final concentration of 70% and the DNA was precipitated at -70°C for 15 min. The DNA was pelleted by centrifugation at 10 000 xg at 4°C for 20 min and re-suspended in Sequenase reaction buffer with 180 pmol primer. The sequencing reaction was otherwise that followed for ssDNA template. Sequence analysis and manipulations utilised the Lasergene Sequence Analysis Software (DNASTAR Inc., Madison, WI) and database searches utilised the BLAST family of algorithms (Altschul *et al.*, 1997; Altschul *et al.*, 1990).

2.11 Electrophoresis techniques

Polyacrylamide gel electrophoresis (PAGE) : Polyacrylamide gels for the electrophoresis of nucleic acids were run on a vertical electrophoresis apparatus from BRL (Gaithersburg) in TBE buffer (0.178 M Tris base/17.8 mM boric acid/ 0.4 mM EDTA). A Model S2 apparatus was used for sequencing and analysis of RNase HI hydrolysis products, and a Model V16 apparatus was used for monitoring hybrid formation and analysis and elution of labeled fragments. Samples were mixed with denaturing PAGE loading buffer (95 % formamide/20 mM EDTA/0.025 % bromophenol blue/0.025 % xylene cyanol) and heat denatured for 3 min at 75°C prior to loading. Samples run on non-denaturing gels were mixed with agarose gel loading buffer (30 % glycerol in water/0.025% bromophenol blue) and loaded directly. The power supply used for electrophoresis was a Pharmacia ECPS 2000/300. Electrophoresis parameters for the various applications are detailed in Table V.

Table V: PAGE parameters

technique	% acrylamide	[urea], M	voltage
sequencing	6	7	2 kV
hybrid formation	5	none	500V
RNase HI assay	15	8	2 kV
RNA elution	6	7	500V

SDS-PAGE: All SDS-PAGE gels were run on a BRL Model V16 vertical gel electrophoresis apparatus powered by a Consol E445 power pack. Protein samples were mixed with 3 x SDS-PAGE loading buffer (0.1875 M Tris-HCl pH6.8, 6% SDS, 15% β -mercaptoethanol, 30% glycerol, 0.006% bromophenol blue) and heat denatured at 80°C for 3 min prior to loading. Gels were either run for 14-16 h at a constant voltage of 44 V or for 30 min at 80 V, 30 min at 90 V and then for 4 h at 125 V. The gels were stained for 30 min at room temperature in 200 ml Coomassie Brilliant Blue stain (0.5% Coomassie Brilliant Blue in a 5:5:2 water:methanol:glacial acetic acid solution) and destained in a 30% methanol/10% acetic acid solution with gentle agitation.

Agarose gel electrophoresis of DNA : Samples were loaded onto an agarose gel (0.8-2% agarose in TAE) in a buffer containing 5% glycerol and 0.004% bromophenol blue. The gel was run in a Model E321 horizontal gel apparatus (Consol, Belgium) at 80-100V for 1 h to 2 h in TAE buffer. Ethidium bromide (0.3 μ g/ml) was included in both the running buffer and the agarose gel. Chromosomal digests and plasmid screens were run on 0.8% gels, while resolution of PCR products necessitated 1.5-2% agarose gels. LMA gels consisted of 0.8 - 1% low-melting agarose in TAE.

Agarose gel electrophoresis of RNA : RNA was run on 1.5% agarose gels containing 0.1% SDS. The gels were run as above in TAE without ethidium bromide. After electrophoresis, the gel was soaked in DEPC-water for 30 min to leach the SDS and then stained with ethidium bromide (0.3 μ g/ml in TAE) for 30 min to visualise the RNA.

2.12 Hybridisation techniques

Southern Blotting (Southern (1975), with modifications):

(a) Transfer

Gel-fractionated DNA : DNA was restricted and the products were fractionated on a 0.8% agarose gel. The agarose gel was soaked in 0.25 M HCl for 15 min, washed with distilled water, then soaked in 0.5 M NaOH/1.5 M NaCl for 15 min. The gel was washed in distilled water again and the fractionated DNA transferred to a nylon membrane (HybondTM-N) using the TE 22 Mini Transfor (Hoeffer Scientific) at 0.5 A for 2 h in TBE buffer. The DNA was cross-linked to the membrane using a UV Stratalinker 1800.

Colony lifts : Membranes were placed onto chilled transformation plates and the orientation of the filter with respect to the plate was marked. The filters were then transferred, colony side up, to filter paper soaked in the following solutions as follows: 10% SDS for 3 min, 1.5 M NaCl/1.5 M NaOH for 5 min, 0.5 M Tris-Cl/1.5 M NaCl pH 7.4 for 5 min and 2 x SSC for 1-2 min. The DNA was cross-linked to the membrane using a UV Stratalinker 1800. The membranes were then incubated in 10 ml of pre-wash solution (5 x SSC/0.5% SDS/1 mM EDTA) in roller bottles in a Techne Hybridiser HB-1 at 50°C for 30 min. Cell debris was wiped gently from the filter with a tissue soaked in pre-wash solution.

(b) Hybridisation

The membranes were pre-hybridised in 5 ml hybridisation buffer (0.5 % SDS/6 x SSC/5 x Denhardt's/50 % deionised formamide) containing 10 µg/ml heat-denatured salmon-sperm DNA. Labeled probe was denatured at 96°C for 10 min and added to the pre-hybridisation solution. The membranes were incubated with the probe at 42°C for 17-18 h in roller bottles in a Techne Hybridiser HB-1. The hybridisation solution was then removed and the membranes were washed as follows:

Solution	Duration	Repeats	Temperature
2 x SSC, 0.1 % SDS	15 min	2	42°C
0.5 x SSC, 0.1 % SDS	15 min	1	42°C
0.1 x SSC, 0.1 % SDS	15 min	1	42°C
0.1 x SSC, 1 % SDS	30 min	1	65°C

The filter was then autoradiographed at -70°C for a period of up to one week.

S1 nuclease protection assay (Reiter and Palm (1990) with modifications by Stolt and Zillig (1993)): Total cell RNA (20 - 30 µg) was ethanol-precipitated with 15 000 - 50 000 cpm of 5' end-labeled ssDNA together with 10 µg of carrier tRNA. The pellet was washed with 70% ethanol and re-suspended in 15 µl of 50 mM PIPES pH7.0/5mM EDTA /3 M Na-TCA. The mixture was overlaid with 1 drop of mineral oil and incubated at 75°C for 1 minute and then at 45°C overnight. A 300 µl aliquot of 1 x S1 nuclease

buffer (280 mM NaCl/4.5 mM ZnSO₄/30 mM NaOAc pH6.5) containing 400 U S1 nuclease was added to the hybridisation mixture and the reaction was incubated at 37°C for 1 h. The reaction was quenched by the addition of 75 µl of S1 stop solution (2.5 M NH₄OAc/50 mM EDTA/26 ng/µl carrier DNA) and the nucleic acid was ethanol-precipitated. The pellet was dried and re-suspended in 5 µl of formamide loading dye, denatured at 80°C for 3 min and loaded on a 6% PAG together with the corresponding sequencing reaction.

2.13 Isolation of RNA from mycobacteria (Ausubel *et al.* (1989) with modifications by Levin and Hatfull (1993))

A 250 µl aliquot of a stationary phase culture of *M. smegmatis* was used to inoculate 25 ml of M-TW. The culture was grown overnight to an OD₆₀₀ = 0.6. The cells were harvested in two 15 ml Elkay culture tubes at 4000 rpm in an IEC Clinical Centrifuge for 10 min. A 3 ml aliquot of chilled lysis buffer (4 M guanidinium thiocyanate/13 mM sodium citrate/2.6% sarcosyl/1 mM β-mercaptoethanol) was added to each tube and the pellet sonicated at 40 W for 30 s. A 400 µl aliquot of 2 M NaOAc (pH 4.4) was added to each tube, and 2 ml hydroxyquinolone phenol at 65°C (water-saturated, pH 4.3) was added immediately afterwards. The tubes were shaken vigorously and 1 ml chloroform was added and the contents mixed. The tubes were incubated on ice for 5 - 10 min and then centrifuged at 10 000 xg for 10 min to separate phases. The supernatant was retained and the RNA precipitated by the addition of 0.6 volumes of isopropanol and centrifugation at 10 000 xg at 4°C for 20 min. The pellet was washed with 70% ethanol and air-dried. The RNA was re-suspended in a total of 200 µl of DEPC-treated water.

2.14 Protein extraction and purification

Total cell protein from mycobacteria : *M. smegmatis* cells were grown to an OD₆₀₀ = 1.5. Aliquots (3 ml) were pelleted and re-suspended in 1 ml wash buffer (50 mM potassium phosphate, pH7.5). The cells were re-pelleted and washed. The cells were then re-suspended in 750 µl wash buffer containing 10% acetone in Bio-101 lysis tubes (Savant) containing lysing matrices. The cells were subjected to mechanical lysis in the

Bio-101 FastPrep 120 at a speed of 5.5 for 2 x 15 s. The tubes were placed on ice until chilled and the supernatant transferred to microfuge tubes. The supernatant was cleared by centrifugation at 10 000 xg for 10 min. A 100 µl aliquot of this supernatant was used for quantification of total protein. Protein concentrations were determined using the Bradford assay as supplied in the BIO-RAD Protein Assay Kit using BSA as the protein standard.

Purification of recombinant MBP::RNase H1 (as directed by NEB); *E. coli* TB1 [pMAL::RNH] was grown overnight in LB (100 µg/ml Amp), 800 µl of the overnight culture was used to inoculate 80 ml of rich growth medium (1% tryptone/0.5% yeast extract/0.5% NaCl/0.2% glucose/100 µg/ml Amp). The cells were grown until $OD_{600} = 0.5$. IPTG to a final concentration of 300 mM was added to induce expression of recombinant protein and the cells were further grown for 1 h at 37°C. The cells were harvested by centrifugation at 1500 x g for 20 min at 4°C. The pellet was re-suspended in 5 ml column buffer (20 mM Tris-HCl pH7.4/1.17% NaCl/1mM EDTA) and frozen overnight at -20°C. The cells were then thawed on ice and sonicated in an ice water bath for 15 second pulses at 40 W for a total of 3 min. The mixture was centrifuged at 9000 x g for 20 min and the supernatant retained. The recombinant protein was extracted from the supernatant by passing the supernatant over a 15 ml amylose resin column poured in a 50 ml syringe. The column was washed with 60 ml column buffer and recombinant protein eluted by passing column buffer containing 10 mM maltose through the column; 3 ml fractions were collected and the peak fractions (OD_{280}) were pooled. Recombinant protein from these fractions was concentrated 10 fold by centrifugation over Ultrafree MC filters (Millipore) and stored at -20°C in 25 mM HEPES-KOH/50 mM KCl/1 mM DTT/50% glycerol at a final concentration of 14 nM. Protein concentrations were determined using the Bradford assay as above.

2.15 Protein biochemical assays

Gel renaturation assay of RNase H activity (Tanese and Goff, 1988; Mizrahi, 1989): An SDS-PAGE gel polymerised with labeled RNA-DNA hybrid was used to size-fractionate either total cell protein or purified recombinant protein. Protein samples were

usually loaded in parallel so that one half of the gel could be visualised by Coomassie Blue staining while the other half was assayed for RNase H activity: Protein was renatured by washing the gel at room temperature in 200 ml aliquots of 50 mM Tris-HCl pH8.0/50 mM NaCl/8 mM MgCl₂. DTT to a final concentration of 2 mM was added to this buffer immediately prior to washing. The buffer was changed after 2 h and then twice a day for 4 - 5 days. The gel was wrapped in Gladwrap and autoradiographed without drying at -70°C. for 1-3 d.

Electrophoretic separation of RNase HI hydrolysis products (Mizrahi, 1989): 15 µl of RNA-DNA hybrid was added to reaction buffer to make a final concentration of 60 mM Tris-HCl, 60 mM KCl, 10 mM MgCl₂ and 5 mM DTT in a total volume of 40 µl. A 6 µl aliquot was mixed with 4 µl 1× PAG loading buffer as a zero time point, and subsequent 6 µl aliquots were similarly quenched at 30 s, 1, 2, 4, 8, and 16 min after the addition of either 1 µl *E. coli* RNase HI or 1 µl recombinant MBP::RNase HI and incubation at 37°C. Aliquots were then loaded on a 15 % denaturing polyacrylamide gel and electrophoresed at 2 kV for 90 min.

Surface plasmon resonance : Procedures for the immobilisation of ligand on the surface of a CM5 chip and detection of ligate/ligand interaction by the BIAcore™ SPR Biosensor apparatus (Pharmacia Biosensor AB, Uppsala, Sweden) were as recommended by the manufacturer. Experiments were either run at 25°C or 37°C. Briefly, after equilibration of the machine with HBS buffer (10 mM HEPES pH7.4/150 mM NaCl/0.05% surfactant P20), the dextran-coated gold surface of the chip was activated and blocked using the amine coupling kit; 30 µl freshly mixed 1:1 NHS:EDC (both 0.1 M in water prior to mixing) was flowed over the surface of the chip at a flow rate of 2-5 µl/minute. After activation, *E. coli* σ³² (30 µg/ml in formate pH3.0) or DnaJ (32 µg/ml in formate pH3.0) or *E. coli* DnaK (32 µg/ml in formate pH3.0) or RNase HI (50 µg/ml in 10 mM NaOAc pH4.0) was covalently linked to the activated surface by flowing 30 µl of the protein through the machine at a flow rate of 5 µl/min. The surface was then blocked by 30 µl of 1 M ethanolamine, pH 8.5, at the same flow rate. Ligand at varying concentrations in either HBS buffer, or A buffer (HBS/1 mM ATP/1 mM DTT) or A-ATP

buffer (HBS/1 mM DTT) was flowed at a flow rate of 5-15 μ l/minute over the chip. Sensorgrams depicting changes in surface plasmon resonance from protein-protein interactions were generated and analysed using the BIAcore™ software. Protein interactions were disrupted by passage of either 2 M NaCl, or 5 mM HCl over the chip surface.

Catechol 2,3-dioxygenase (CDO) activity assay : A 10 μ l aliquot of total cell protein extract was added to 750 μ l of 50 mM potassium phosphate buffer, pH 7.4 and the absorbance zeroed on a Shimadzu spectrophotometer at 375 nm. A 5.3 μ l aliquot of 10 mM catechol was added to the mixture and mixed thoroughly. The change in OD₃₇₅ was monitored at 10 second intervals. Activity (in mU/mg total cell protein) was calculated using the following formula:

$$\frac{\Delta A^{375} \cdot v}{\epsilon \cdot l \cdot t \cdot \text{mg}}$$

where ΔA^{375} = change in absorbance measured at 375 nm
 v = volume of reaction in litres (typically 0.8 ml)
 l = path length in cm (1 cm)
 t = time in min over which ΔA^{375} was measured
 ϵ = extinction coefficient (4.4×10^4)
 mg = mg of cell protein present in the reaction

3. RESULTS

3.1 Cloning and sequence analysis of the *M. smegmatis* *rnhA* gene

Gene cloning of the *M. smegmatis* *rnhA* gene. A 200 bp probe for the *M. smegmatis* *rnhA* gene had been generated by PCR using degenerate primers based on conserved blocks of amino acids between other bacterial RNase HI homologues (Mizrahi *et al.*, 1993). This probe was used to isolate fragments of chromosomal DNA from *M. smegmatis* LR222 containing the *rnhA* gene. *M. smegmatis* LR222 and *M. tuberculosis* H37Rv chromosomal DNA was restricted, size fractionated and probed with the PCR product (Fig. 2). The probe hybridised to only one fragment in each of the LR222 digests indicating the presence of a single copy of the *rnhA* gene. No cross hybridisation of the probe to *M. tuberculosis* DNA was observed. Of the *M. smegmatis* digests, the *Bam*HI and *Pst*I digests yielded products of approximately 2.3 kb and 3.6 kb, respectively, that were considered to be of a suitable size for cloning.

A band of fragments of the size corresponding to the region of hybridisation of the probe was purified from both the fractionated *Bam*HI and *Pst*I digests of LR222, and shotgun cloned into appropriately digested pGEM@3zf(+). A colony lift of the ligations transformed into DH5 α was performed and positive clones were isolated by Southern analysis. The 3.6 kb clone isolated from the *Pst*I-digested DNA and the 2.3 kb clone from the *Bam*HI-digested DNA were termed pSD1 and pSD2 respectively. The clones were analysed by further restriction digestions coupled with Southern blotting and, as shown in Fig. 3, were found to overlap by 320 bp. Appropriate fragments were subcloned into M13mp18 and M13mp19 for DNA sequence analysis, the strategy of which is also shown in Fig. 3.

Identification of the *rnhA* ORF. The DNA sequence of the 1571 bp region from 250 bp upstream of the *Kpn*I site of pSD1 to 263 bp downstream of the second *Pst*I site in pSD2 was determined and has been deposited in Genbank under the accession number L11918. Sequence data from the *Kpn*I site to the *Sac*I site on pSD1 was obtained on one strand only. The remaining region, containing the *rnhA* gene, 398 bp of upstream sequence and 188 bp of downstream sequence, was sequenced on both strands with a

redundancy of at least 3/bp. The base content of the DNA was 17.92 %A, 13.12 %T, 35 %G and 33.96 %C. The 69 % G+C content of the DNA was consistent with the G+C ratio determined for *M. smegmatis* DNA (Clark-Curtiss, 1990). An open reading frame of 480 bp encoding a possible RNase HII of 159 amino acids was found. No ATG (methionine) start codon was found, but two possible GTG (valine) start codons, separated by 36 bp, were noted. A GTG start is often preferred over ATG in GC-rich organisms (Dale and Patki, 1990) and a bias in favour of cytosine as the base immediately preceding these GTG start codons has been previously noted (Honoré *et al.*, 1993).

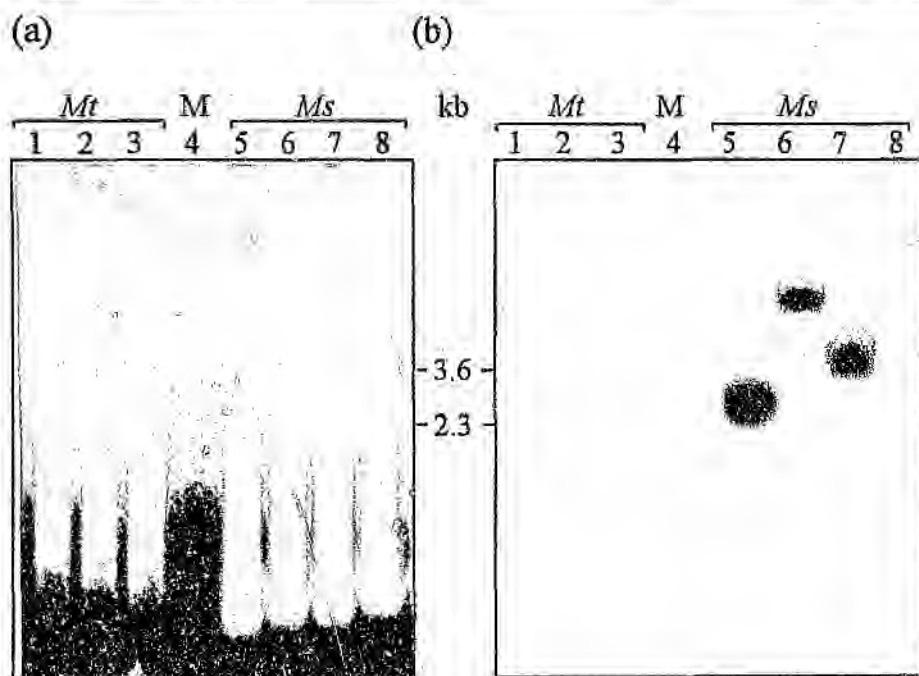


Fig. 2 (a) Agarose gel electrophoresis of restricted genomic DNA (5µg per lane) from *Mycobacterium tuberculosis* H37Rv (*Mt*; lanes 1,2 and 3 digested with *Bam*HI, *Eco*RI and *Pst*I respectively) and *M. smegmatis* LR222 (*Ms*; lanes 5-8, digested with *Bam*HI, *Eco*RI, *Pst*I and *Hind*III respectively) run with DNA size markers (*M*; lane 4 - BM molecular weight marker III), Southern blotted and (b) probed with the 200 bp PCR probe homologous to part of *M. smegmatis* *rnhA*. Sizes indicated are in kb.

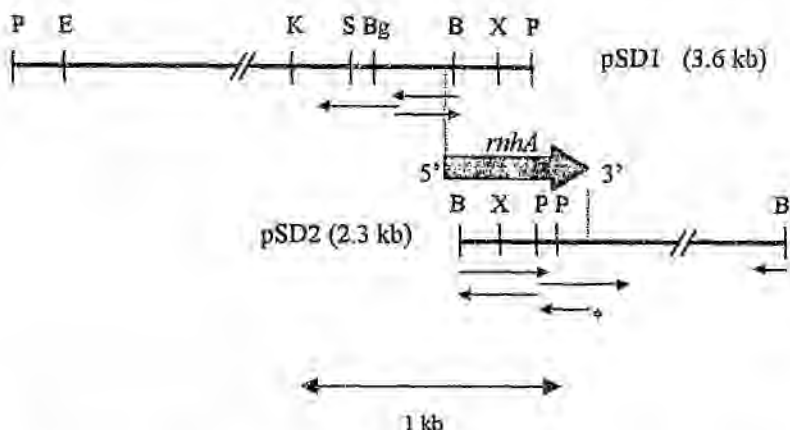


Fig. 3. Restriction maps of pSD1 and pSD2. Vector sequence is not shown. The position of the *rnhA* ORF with respect to pSD1 and pSD2 is indicated by the box labelled *rnhA*. Restriction sites used for subcloning fragments into M13mp18/19 for sequencing are shown, together with the sequencing strategy as indicated by arrows. The restriction sites are as follows: B, *Bam*HI; P, *Pst*I; E, *Eco*RI; S, *Sac*I; Bg, *Bgl*II; K, *Kpn*I; X, *Xho*I. The universal primer was used for all sequencing except where indicated by an asterisk (*) where a primer (SGRSEQ-R3, Table IV) was designed based on mycobacterial sequence. Sequence data has been deposited in GenBank under the accession number L11918.

Assignment of the second GTG as the start codon of *rnhA* was based on these considerations, in conjunction with alignment of the deduced amino acid sequence (159 amino acids) with other bacterial RNase HI homologues (Fig. 4). The RNases HI of *M. tuberculosis* (Genbank Acc. No. Z70692) and *M. leprae* (Genbank Acc. No. L78818) were identified by homology searches of the Sanger database (http://www.sanger.ac.uk/Projects/M_tuberculosis/) and Genome Therapeutics database (<http://www.cric.com/htdocs/sequences/>) respectively. These mycobacterial RNase HI homologues differ in length from the other bacterial RNases HI as a result of the carboxy-terminal fusion of a protein domain with similarity to the phosphoglycerate mutase family of protein 3. The RNase HI domains, however, show high similarity to the other Gram-positive RNases HI.

The deduced amino acid sequence of the *M. smegmatis* *rnhA* ORF has 50 %

		<u>BA</u>	<u>BB</u>	<u>BC</u>	<u>αI</u>
MsmhA	1	VSQDP---	V I I H T D G G C R P N P G P G G W G A V L R H R E H V R E N ---	FGGEAAVTSNNRMELTAPIM	
TtmhA	1	VNPSPRKRVALFTDGCACLNPGPGGWAALLRFHAHEK-L---	L5GGEAWTTNNRMELKAAIE		
EcmhA	1	ML----	KQVEIFTDGSCLGNPGPGGYGAILRYRGREK-T---	FSAGYTRTTNNRMELMAATV	
StmhA	1	ML----	KQVEIFTDGSCLGNPGPGGYGAILRYRGHEK-T---	FSEGYTLTTNNRMELMAATV	
BamhA	1	ML----	KLYKMFSDGSCLGNPGSGGYGTILRYKLHEK-I---	LTSGFFLTNNRMELMGVIC	
EfmhA	1	M-----	LRIYVDAATKGNPGESGGG-IYYLTDDGR---	QQLHIVPLGIVSNHEAEFKVLE	
BsmhA	1	M-----	PTEIYVDGASAGNPGPSGIG-I-FIKHEGKA-ESFSIFIGVHTNQAEFLALIE		
MlmhA	1	V-----	KVITLADGSGRGNPGPAGYGA VVWIAORSAYLTETKQATGRATNNVAEYRALIA		
MtmhA	1	V-----	KVYTEADGSGRGNPGPAGYGA VVWTAHSTYLAESKQATGRATNNVAEYRGLIA		
		<u>BD</u>	<u>αII</u>	<u>αIII</u>	<u>αIV</u>
MsmhA	57	ALEAL----	TRPVTVHLYTDSITYVRNGITK---	VLGWERNGWMTAAKQPVKNVDLWORLDAA	
TtmhA	59	GLKAL----	KEPCEVDLYTDSHYLKKAFTEGWLEGWRKRGWRTAEGKPVKNRDLWEALLLA		
EcmhA	55	ALEAL----	KEHCEVILSTDSQYVRGIGITQ-WIHNWYKRGWKTAOKKPVKNVDLWORLDAA		
StmhA	55	ALEAL----	KEHCEVILSTDSQYVRGIGITQ-WIHNWYKRGWKTAOKKPVKNVDLWORLDAA		
BamhA	55	GLESL----	KESCIVEITIDSQYVKGQITN-WIATWEKKKWKTTKKKLIKNDLWLRLINAV		
EfmhA	52	ALKQAIANEDNQTVLLHSDSKIVVQTIEKNY-----	AKNEKYQPYLAEY		
BsmhA	53	GMK--LCATRGYQSVSFRTSDIVERATELEM-----	VKNITFOFFVEEJ		
MlmhA	56	GLDDA--VKMGATEAEVLMDSKLVVEQ----	MSGRWK-----	VKHPDLIELYVHA	
MtmhA	56	GLDDA--VKLGATEAAVLMDSKI VVFG----	MSGRWK-----	VKHPDLKLYVOA	
		<u>BE</u>	<u>αV</u>		
MsmhA	113	CARHOVEWFYKGHSGTIGONELADELATRELOEAVGLTTSSAGTSLR.		159	
TtmhA	116	MAPHRVRHFYKGHGTGHPENERVDREARRQAQSQAKTPCPPRAPTLFEEA.		165	
EcmhA	111	LGOHQIKWEWYKGHAGHPENERCDELARAAAMNPTLEDITGYOVEV.		155	
StmhA	111	LGOHQIKWYVWKGHAGHPENERCDELARAAAMNPTQEDSGYQAEA.		155	
BamhA	111	IKNHHITWFWVKAHMGHLENERCDKJARQSAQSPSVKOFFYENNFYQKNL.		160	
EfmhA	103	QOLEKNFPLLLIKWLPESQNKAAADMLARDALOKFYPNKK.		135	
BsmhA	102	IRLKAAFFLFFIKWIPKONQKADLLAKEAIRLNEKN.		132	
MlmhA	106	OTLASRLASVSYTWIPTRNSRADRLANEAMDAAASCAEGRYDQAATOLAQSTKAVAKESQ			
MtmhA	106	DALASQFRRINYEWPVRARNTYADRLANDAMDAASAAADADP-----	AKIVATESP		
MlmhA	160	TAPGWTGARGTPTRLLI LRHGQTELSVORRYSGRGNPALNNVGWROARAAARYFAQRGGI			
MtmhA	153	TSPGWTGARGTPTRLLLL RHGQTELSQRRYSGRGNPGLNEVGWROYGAAAGYLARRGGI			
MlmhA	220	AGVYSSPLORAYDTAATVADLLGANLTYDDDLIETDFGAWEGLTFAEAAAARDPOLHHCWL			
MtmhA	213	AAVYSSPLORAYDTAVTAARALALDVVYDDDLVETDFGAWEGLTFAEAAERDPELHRRWL			
MlmhA	280	HDATATLPPGGESFDDVLGRYCSARARI IAEYOGETYLVVSHVTPIKMLRLALDAGAGIL			
MtmhA	273	QDTSITPPGGESFDDVLRRVRRGRDRI IGYEGATVLVSHVTPIKMLRLALDAGSGVL			
MlmhA	340	YRLHLDLGLSLSTAEFYPDGASSVRLVNOTGYL.		371	
MtmhA	333	YRLHLDLGLSLSTAEFYADGASSVRLVNOTGYL.		364	

Fig. 4. Comparison of the RNase HI of *M. smegmatis* (MsmhA) with those of the Gram-negative bacteria *Thermus thermophilus* (TtmhA; Ilaya and Kondo, 1991), *E. coli* (EcmhA; Kanaya and Crouch, 1983), *Salmonella typhimurium* (StmhA; Ilaya et al., 1991), and *Buchnera aphidicola* (BamhA; Munson et al., 1993), and with those of the Gram-positive bacteria *Enterococcus faecalis* (EfmhA; Bensing and Dunny, 1993), *Bacillus subtilis* (BsmhA; Iwakura et al., 1988), *M. leprae* (MlmhA; Genbank Acc. No. L78818) and *M. tuberculosis* (MtmhA; Genbank Acc. No. Z70692). The RNase HI domains of the *M. leprae* and *M. tuberculosis* proteins are aligned with those of the other RNases HI, and the extended carboxy-termini of these proteins are aligned with each other. Residues essential for catalysis are red, and identical or highly conserved (8/9) residues are blue. Amino acid sequences forming beta sheets (β A-E) or alpha helices (α I-V) are overlined.

identity to *E. coli* RNase HI (Kanaya and Crouch, 1983). In addition, the three invariant acidic residues that comprise the catalytic nucleus of the RNase HI enzyme (Kanaya *et al.*, 1990a), corresponding to Asp¹⁰, Glu⁴⁶ and Asp⁷⁰ in *E. coli* RNase HI, are present in the *M. smegmatis* RNase HI, suggesting that *rnhA* encodes a functional RNase H. The *T. thermophilus* RNase HI and *M. smegmatis* RNase HI share similar identity (52% and 50% respectively) to *E. coli* RNase HI, and have 44% identity to each other. Since the crystal structure of *T. thermophilus* RNase HI has been found to be very similar to that of the *E. coli* enzyme (Ishikawa *et al.*, 1993a) the three dimensional structure of the *M. smegmatis* RNase HI might therefore be predicted to be very similar to these enzymes.

The chromosomal context of *rnhA*. Analysis of sequence upstream of *M. smegmatis* *rnhA* indicated an ORF with gene-coding potential in a co-directional orientation to *rnhA* (designated ORF1). The 5' region of pSD1 was not sequenced sufficiently far to identify the start codon of ORF1, but the deduced 300 carboxy-terminal amino acid sequence showed no homology to known proteins in the non-redundant Genbank-EMBL-DDBJ database (May 1988). The sequence of the coding strand, including 389 bp of upstream sequence, together with the protein translation of the carboxy terminus of the possible upstream ORF and the *rnhA* gene, is shown in Fig. 5. The end of the putative upstream ORF and the beginning of the *rnhA* ORF is separated by 209 bp. Within this region, three closely related 25 bp direct repeats, each separated by 1 bp, were found 38 bp upstream of the start of the *rnhA* ORF. Two mutually-exclusive stem-loop structures were predicted to form within these direct repeats such that the free energy of base pairing (ΔG) for IR1 = -17.6 kcal.mol⁻¹ and IR2 = -21.0 kcal.mol⁻¹ (calculated according to the method of Tinoco *et al.*, (1973)). An additional pair of 9 bp direct repeats was located flanking the tandem direct repeats. A putative stem-loop structure, consisting of a stem with 8 out of 9 matches and a loop of 9 bp (ΔG = -15.8 kcal.mol⁻¹), was located 98 bp downstream of the translation stop of the *rnhA* ORF and was postulated to serve as a putative factor-independent transcriptional terminator. Stop codons were also noted in all three reading frames immediately following the putative transcriptional terminator, arguing against a possible operon structure with any downstream genes.

						<i>Bgl</i> II	
GCCTGCGCGC	GAACGCGACG	GCGAGCTGGG	CAGGCAGCCG	ATCCGCCAGA	TCTTCTTCCG		74
A S P R T R R R A G	D A A D P P D	L L P					
GTTCGCGCAG	GTGATCATCG	AGGTGGTGGG	TGCACCLGAC	GCGGTGGGCG	ACGGACCGTC		134
V R R G D H R G G G	C T R R R G R R T V						
GTCGCTGTGG	GGCATCACGT	ACGTGGTCCA	CGACATCGAC	GCCAACGCGAG	CATTTTTCGG		194
V A V G H H V R G R	R H R R H R S I F R						
TGACAACGCG	CTGCCGGTCA	AAGATGCCGT	ACAGCCCGGC	CGCCGGATAC	GCCACGCTAC		254
						<i>IR1</i>	
		DR			DR1		
GTGGCGACCG	GCTCGGGCTG	AGCGTGCACA	GCGCGGTGAT	CTCAGCGCCC	GCGCGTGGT		314
						<i>IR1</i>	
		DR2			DR3		
GGTGATCTCG	ACGCCCGCGC	GTGGGGATG	ATCTTGACGC	CCGCGCGTGG	GTGAACCACC		374
						<i>IR2</i>	
		DR			<i>Bqm</i> HI		
GGCACACGCG	CGTTGCCTAT	CCTCTGAGCG	TGAGCCAGGA	TCCCGTCATC	ATCCACACCG		434
ACGGCGGGCTG	CCGCCCAAC	CCGGGCCCGG	GCGGCTGGGG	TGCGGTGCTA	CGGCACCGCG		494
D G G C R P N P G P	G G W G A V L R H R						
AGCACGTCCG	CGAGATGTTT	GCGCGCGAAG	CGGCCGTGAC	CAGCAACAAC	CGCATGGAGC		554
E H V R E M F G G E	A A V T S N N R M E						
TGACGGCCCC	GATCATGGCG	CTCGAGGCGT	TGACCCGGCC	CGTCAGGGTG	CACCTGTACA		614
L T A P I M A L E A	L T R P V T V H L Y						
CCGACAGCAC	CTACGTGCGC	AACGGCATCA	CCAAGTGGGT	GCTGGGCTGG	GAGCGCAACG		674
T D S T Y V R N G I	T K W Y L G W E R N						
						<i>Pst</i> I	
GCTGGATGAC	CGCAGCCAAG	CAGCCCGTGA	ACAATGTGCA	CCTGTGGCAG	GCGCTGCAGG		734
G W M T A A K D P V	K N V D L W D R L D						
CCGCGTGGCG	GCGCCACCAG	GTGAGTGGT	TCTGGGTGAA	GCGGCACTCG	GGCATCGGCG		794
A A C A R H Q V E W	F W V K G H S G I G						
ACAACGAGCT	GCGCGACGAA	CTCGCGACCC	GCGGACTGCA	GGAGGCAGTG	GGCCTCACCA		854
D N E L A D E L A T	R G L D E A V G L T						
CCAGTAGTGC	GGTACTTTCG	CTCGGTGAG	CTTGTGCCC	GCCAGGCCGA	CCACCACGAG		914
T S S A G T S L R							
						<i>TT</i>	
GATCACCGCG	GCGGTGCGGA	CCAGGACGAC	GAATCCACAG	GCGGGATCT	CGGGCACCCAC		974
GCGGATCATC	GCGGTGCGCG	CGGCCCGTGA	GTGAGACATC	CTGAGCAGCT	TCATCATTCG		1034
GAGTGCCAGT	GCGCCGGCCA	CCGCACCGGC	CCACAACGCG	TCGAGAGCCG	TCATACCGAC		1094

Fig. 5. Nucleotide sequence (corresponding to the numbering of L11918) and amino acid translation of the carboxy terminus of the upstream ORF1 and of the *rmlA* gene. Direct repeat (DR) regions are indicated by arrows, where the tandemly arranged 25 bp repeats are labeled DR1, DR2 and DR3, and the complementary regions of the inverted repeats (IR) are indicated by boxed arrows. The putative factor independent transcription terminator (TT) is shown similarly. Stop codons are indicated by asterisks (*).

3.2 Promoter analysis of *rnhA*

rnhA transcription is driven off its own promoter. No apparent “-10” region upstream of *rnhA* could be identified by sequence homology to the general consensus TANNNT (Verma *et al.*, 1994; Kenney and Churchward, 1996; Gonzalez-y-Merchand *et al.*, 1997). In addition, no obvious -35 region, or purine-rich tract that could function as a ribosome binding site, was apparent in the sequence upstream of *rnhA*. The lack of identifiable -10 or -35 regions and the presence of a putative ORF upstream of *rnhA* was suggestive of a possible operon arrangement. However, the spacing of 207 bp between the two ORFs was also considered to be sufficient to contain a promoter region for the *rnhA* gene (C. Nesbitt, pers. comm., Das Gupta *et al.*, 1993). In addition, attenuation of

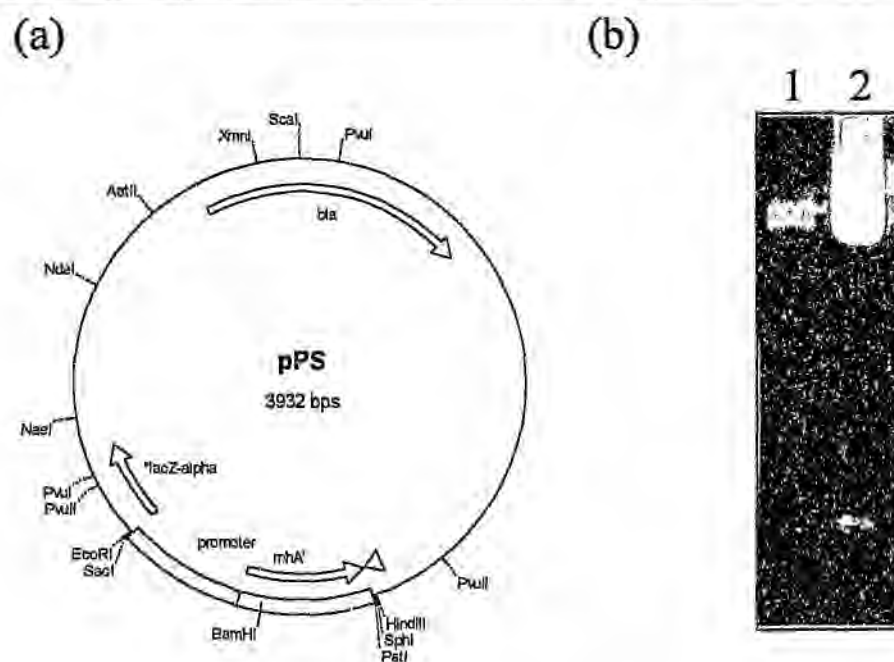


Fig. 6 (a) Putative promoter-containing *SacI*-*PstI* fragment cloned into pGEM3zf(+) consisting of the last 189 bp of ORF 1, 207 bp of intervening sequence and the first 331 bp of *rnhA* and (b) agarose gel electrophoresis of transcription products produced by *M. smegmatis* RNA polymerase off this template, incubated with and without RNase A (lanes 1 and 2 respectively).

transcriptional read-through from ORF1, if not termination, would probably be achieved by one of the stem-loop structures formed within the direct repeat region. The presence of a promoter region between ORF1 and *rnhA* was probed by analysing the *in vitro* transcription products of *M. smegmatis* RNA polymerase using a supercoiled 727 bp *Pst*I/*Sac*I subclone of pSD1 (pPS) as the substrate (Fig. 6a). This plasmid construct contained carboxy-terminal sequence of ORF1, 207 bp of sequence between ORF1 and *rnhA*, and the first 331 bp of *rnhA* and, as shown in Fig. 6(b), was shown to direct transcription.

Identification of the transcriptional start site. S1 nuclease protection assays are considered to be more robust than extension of a primer using reverse transcriptase, especially if secondary structures are likely to form in the transcription region (Strohl, 1992). The primer S1-BAM (Table IV), designed around the start codon of *rnhA*, was therefore used to generate high molecular weight ssDNA, labelled at the 5'-end, to probe total cell RNA for the transcriptional start site of the *rnhA* gene. Two different preparations of total cellular RNA (guanidinium thiocyanate and French press/CsCl preparations, the latter a kind gift from P. Stolt) were probed, and the transcriptional start site was mapped to a T residue located 56 bp upstream from the start of the *rnhA* gene (Fig. 7a). This transcriptional start lies within the direct repeat motifs in the loop region of the 3'-most stem-loop structure. Possible -10 and -35 regions, consisting of TCGGGG and TGGTGA respectively, were identified.

Functional analysis of the *rnhA* promoter region. Since no obvious regulatory elements could be identified from sequence data, and the transcriptional start point mapped within the direct repeat region, it was hypothesised that a) the entire promoter region for the *rnhA* gene was contained within the direct repeat region and b) that preferential formation of one or the other of the stem-loop structures could form part of the regulatory region of the *rnhA* gene; the formation of both stem loop structures *in vitro* as evidenced by the polymerase pausing observed during sequencing reactions (Fig. 7a) was suggestive of the potential formation of such structures *in vivo*. Promoter probe constructs designed to : a) terminate immediately upstream of the 5'-most direct repeat (DR1) and b) promote formation of one stem-loop structure exclusively by deletion of one

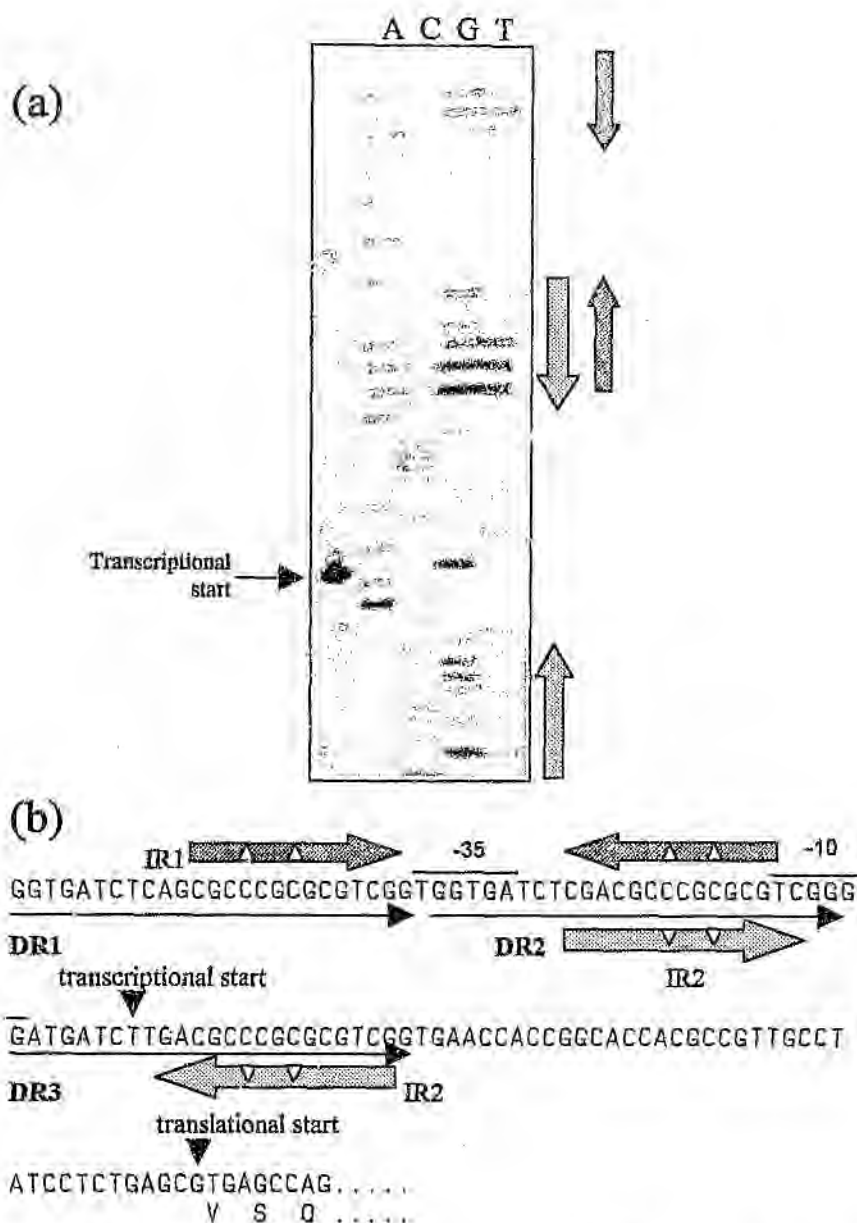


Fig. 7. Location of the transcriptional start. (a) product of S1 nuclease protection assay using the end-labeled primer S1-BAM, run with the corresponding sequence generated by S1-BAM. (b) Position of the transcriptional start, relative to the direct repeat structures, 56 bp upstream of the translational start. Inverted repeat structures (IR) are shown by similarly shaded box arrows and residues involved in pairing are boxed. The direct repeats (DR) are indicated by line arrows. Possible -10 and -35 regions are overlined.

of the regions of complementarity, were used to test this hypothesis (Fig. 8). The assay for promoter activity utilised the mycobacterial promoter-probe vector, pRCX3 (Curcio *et al.*, 1994), which contains the promoter-less *xylE* gene cloned downstream of a multiple cloning site. *xylE* encodes 2,3-catechol dioxygenase (CDO) which converts catechol to 2-

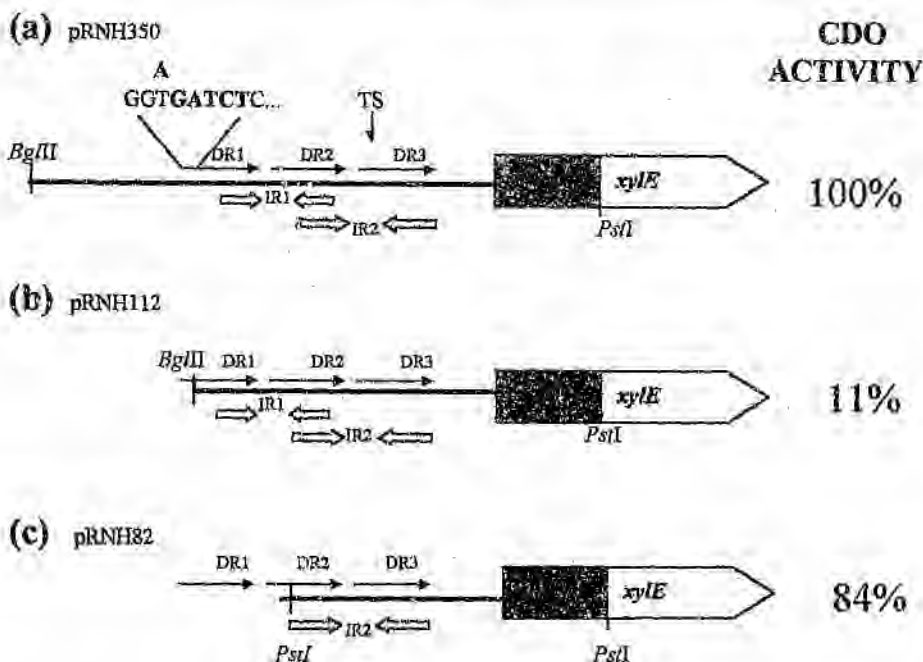


Fig. 8. Detail of the promoter regions contained in the three promoter constructs pRNH350, pRNH112 and pRNH82, cloned into the multiple cloning site of pRCX3, in a co-directional orientation with the promoter-less *xylE* gene to form a transcriptional fusion with the *xylE* gene. (a) The first 9 bp of DR1 and DR2 are shown above direct repeat 1 (DR1) in the pRNH350 construct. (b) pRNH112 was constructed by introducing a *Bgl*III site in the beginning of DR1 by site-directed mutagenesis using the primer PROM1-F1 (Table IV). (c) A *Bgl*III site was introduced by PCR mutagenesis into the beginning of DR2 using the primers STEF-1 and SGREX-R3 (Table IV) to create pRNH82, but the cloned amplicon was truncated before the *Bgl*III site. It was cloned as a *Pst*I fragment instead. The catechol dioxygenase (CDO) activity directed by each promoter construct is expressed as a percentage of the WT promoter, pRNH350.

hydroxymuconic semialdehyde, a product with a bright yellow colour that can be quantitatively assayed by monitoring product development at 375nm. Deletion mutants of the *rnhA* promoter region were constructed by introducing a *Bgl*II site in a) the first 8 bp of DR1 by site-directed mutagenesis using the primer PROM1-F1 (Table IV) and cloning the product as a *Bgl*II-*Pst*I fragment into pRCX3 to form pRNH12, and in b) the first 8 bp of DR2 by PCR amplification of pPS using the primers STEF-1 and SGREX-R3 (Table IV). The introduction of the *Bgl*II site in DR2 was predicted to generate a T→A substitution at the -31 position. Sequence confirmation of the PCR product, however, showed truncation of the amplicon so that the promoter construct contained only 25 bp of original sequence upstream of the transcriptional start. Although an extended -10 motif of TGN could allow for transcription in the absence of a -35 region (Bashyam *et al.*, 1998; Kenney and Churchward, 1996; Kumar *et al.*, 1993) this motif could not be identified in the *rnhA* promoter region. However, the region of complementation for the formation of the stem-loop structure was intact, so this construct was cloned as a *Pst*I fragment into pRCX3 to form pRNH82 and assayed for promoter activity. The 680 bp *Bgl*II-*Pst*I fragment of pSD1 was similarly cloned (S.I. Durbach) to form pRNH350 and was used as the wild-type (WT) promoter control.

These promoter constructs were electroporated into *M. smegmatis*. Early log phase cells were assayed for CDO activity. Activity of the WT promoter represented by pRNH350 was calculated at 180 mU/mg. pRNH12 exhibited 11% of the activity of pRNH350, whereas pRNH82 showed 84 % of WT activity.

3.3 Attempts to create a knock-out mutant of *rnhA* by homologous recombination

Construction of a plasmid for *rnhA* gene disruption. The elucidation of the cellular roles of *E. coli* RNase HI has been facilitated by the generation of mutants deficient in RNase HI (Horiuchi *et al.*, 1984; Kanaya and Crouch 1984; Naito *et al.*, 1984; Ogawa and Okazaki, 1984; Torrey *et al.*, 1984). A similar strategy to probe the role of RNase HI in *M. smegmatis* was devised. A construct analogous to the null allele *rnh-339::cat* of *E. coli* (Kanaya and Crouch, 1984) was created by insertion of the *aph* gene, encoding an aminoglycoside phosphotransferase which confers resistance to kanamycin

(Km), into *rnhA* 165 bp downstream of the translational start. The cloning strategy for the homologous recombination substrate pPBA is shown in Fig. 9. Briefly, the *aph* gene from Tn905 was obtained as a 1.3 kb *Bam*HI cassette from pY6002 (Husson *et al.*, 1990), end-filled and cloned into the unique *Xho*I site of pSD2 to create pBA. The 3.5 kb *Bam*HI fragment from pBA was excised and cloned into pSD1, forming pPBA. This plasmid contains approximately 2 kb and 3.4 kb of homologous mycobacterial DNA either side of the *aph* cassette in the *E. coli* plasmid pGEM3zf(+). This plasmid does not replicate in mycobacteria and therefore must recombine with the chromosome to confer Km resistance on transformants. A single recombination event, or single cross over (SCO), results in the tandem arrangement of the two alleles in the chromosome, while two recombination events, or double cross over (DCO), one on either side of the disruption, results in the allelic replacement of the WT allele by the disrupted allele.

Allelic replacement of *rnhA* by *rnhA::aph* confers a lethal phenotype. *M. smegmatis* mc²155 was electroporated with pPBA and recombinants were selected on Km. Recombinants were screened for the replacement of the *rnhA* allele by Southern analysis of chromosomal DNA restricted with *Bam*HI (Fig. 9b), or by PCR analysis by amplification of chromosomal DNA with primers flanking the site of insertion of the *aph* cassette (Fig. 9c). Southern analysis of colonies showed the presence of both the WT allele and the disrupted allele in all colonies screened, while PCR analysis generally preferentially amplified only the smaller 488 bp WT allele. Occasionally the 1776 bp disrupted allele was preferentially amplified by PCR, but subsequent Southern analysis of the chromosomal DNA confirmed the presence of the WT allele in all cases. In total, twenty six colonies were screened, but no DCOs were recovered.

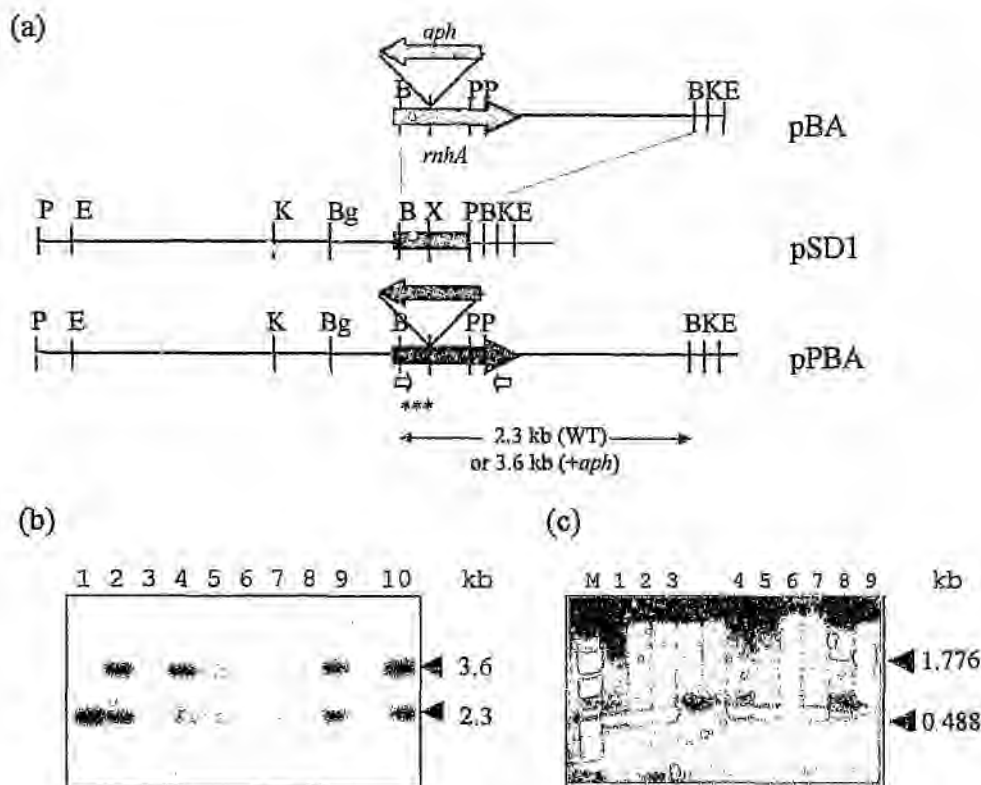


Fig. 9. Homologous recombination with pPBA. (a) pPBA was constructed by cloning the *Bam*HI-digested *aph* cassette into the blunt-ended *Xho*I site of pSD2 to form pBA. The 350 bp *Bam*HI fragment of pSD1 was replaced by the *Bam*HI fragment from pBA to form pPBA. (b) Southern blot of *Bam*HI chromosomal digests of *mc*²¹⁵⁵ (lane 1), and transformants obtained with pPBA (lanes 2-10), probed with the 200 bp PCR product. The sizes expected for the WT and disrupted allele are indicated in (a) and the localisation of the probe is shown by asterisks (*). (c) PCR amplification of *mc*²¹⁵⁵ (lane 1) and transformants obtained with pPBA (lanes 2-9) using primers SGREX-F1 and SGREX-R1 (Table IV, ⇔ ⇐) flanking the *rnhA* gene as shown in (a). The expected sizes of the WT and disrupted allele are 488 bp and 1776 bp respectively, and the relevant amplification bands are shown by arrows. M, DNA Molecular Weight Marker VI. Abbreviations for restriction sites are as follows: B, *Bam*HI; P, *Pst*I; E, *Eco*RI; Bg, *Bgl*II; K, *Kpn*I; X, *Xho*I.

The recovery of allelic replacement events in the mycobacteria has been facilitated by the inclusion of a counter-selectable marker, *sacB*, from *Bacillus subtilis* Pelicic *et al.*, 1996a,b,c; Ramakrishnan *et al.*, 1997). Expression of *sacB* confers lethality on cells grown in the presence of sucrose. The 6.6 kb *EcoRI* fragment from pPBA was therefore subcloned into pUS, a pBluescript-based vector containing the *sacB* marker cloned as a *PstI* cassette (P. Stolt, unpublished) to form pPBAS (Fig. 10).

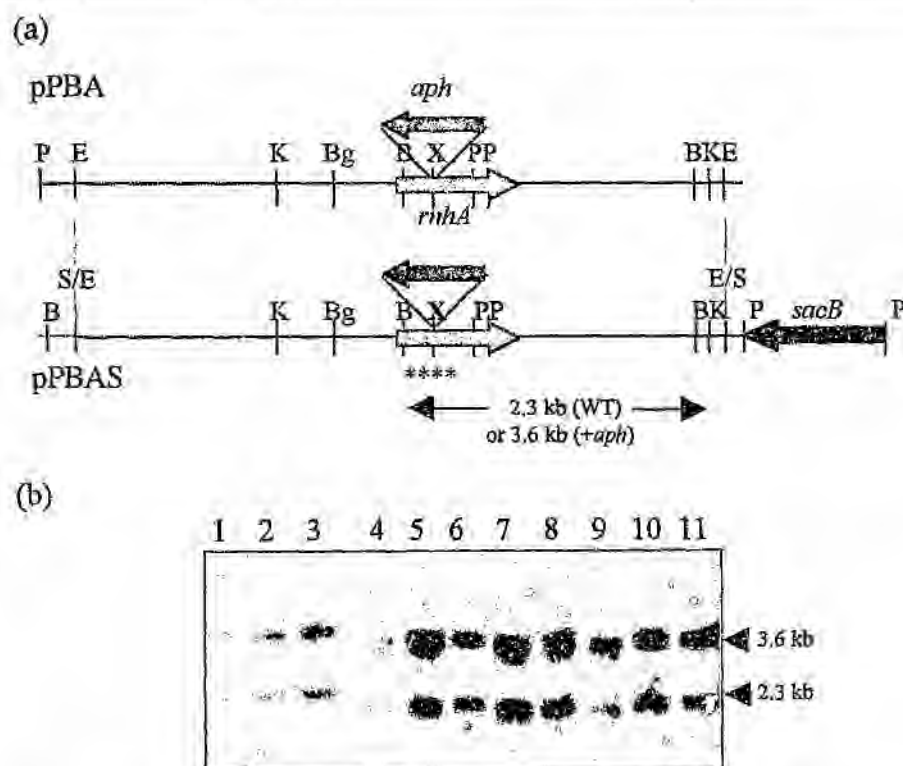


Fig. 10. Homologous recombination with pPBAS which contains the counter-selectable marker, *sacB*. (a) Cloning strategy: The *EcoRI* fragment from pPBA was excised from pPBA, blunt-ended and cloned into the *SmaI* site of pUS to form pPBAS. (b) Southern of *BamHI* chromosomal digests of pPBAS transformants selected on sucrose (lanes 1-11) probed with the 200 bp PCR probe. The localisation of the probe is shown in (a) by asterisks (*), and the fragment sizes expected from a *BamHI* digest are indicated. Abbreviations for restriction sites are as follows: B, *BamHI*; Bg, *BglII*; E, *EcoRI*; K, *KpnI*; P, *PstI*; S, *SmaI*; X, *XbaI*.

M. smegmatis mc²155 was electroporated with this construct and transformants were plated on sucrose. No survivors on sucrose were obtained. A SCO was then passaged on Km to enrich for a second recombination event, and plated to single colonies on sucrose. Eighteen survivors were screened by Southern analysis. All were found to have retained both *rnhA* alleles and were therefore considered to be spontaneous sucrose resistant (Suc^R) *sacB* mutants (Fig. 10b). This inability to recover DCOs after application of positive selection strongly suggested that insertional inactivation of *rnhA* conferred a lethal phenotype.

M. smegmatis rnhA in trans cannot complement rnhA::aph. Although *Mycoplasma genitalium* appears not to possess an RNase H (Fraser *et al.*, 1995), *rnhA* mutants of *E. coli* are unrecoverable, or growth is severely compromised, in certain genetic backgrounds (Hong *et al.*, 1995; Drolet *et al.*, 1995; Itaya and Crouch, 1991a). The possibility that the inactivation of *rnhA* in *M. smegmatis* conferred a lethal phenotype was therefore addressed. RNase HI supplied *in trans* by a replicating plasmid should allow recovery of a DCO within the chromosome. Due to the lack of suitable restriction sites downstream of *rnhA*, an *rnhA* cassette for use in a complementation experiment was amplified using the primers SGREX-F1 and SGREX-R3 where the reverse primer contained a *Hind*III site downstream of the stop codon of the gene, and the forward primer hybridised to the beginning of the gene, just upstream of a *Bam*HI site. (Fig. 16a). pRNHA was therefore constructed by combining the *Bam*HI-*H*⁺ *d*III fragment from the PCR amplification product with the *Bgl*II-*Bam*HI fragment containing the promoter region from pSD1 and cloning the resulting *Bgl*II-*Hind*III fragment into *Bam*HI-*Hind*III pOLYG (Fig. 11a). A SCO, obtained from homologous recombination of mc²155 with pPBA, was electroporated with pRNHA, and a transformant selected on Km was plated to single colonies on sucrose. Of 20 colonies screened, all retained the WT *rnhA* allele and were considered to be spontaneous Suc^R *sacB* mutants (Fig. 11b), although some appeared to have recombined with pRNHA (Fig. 11c). Although experimental evidence confirming production of active RNase HI from pRNHA was not obtained, the biased recovery of *sacB* mutants with *rnhA* supplied *in trans* suggested that the lethal effect of insertional inactivation of *rnhA* did not arise directly from inactivation of the *rnhA*

allele, but was due to a polar effect on a neighbouring gene. This led us to examine the SCOs obtained with pPBAS in more detail.

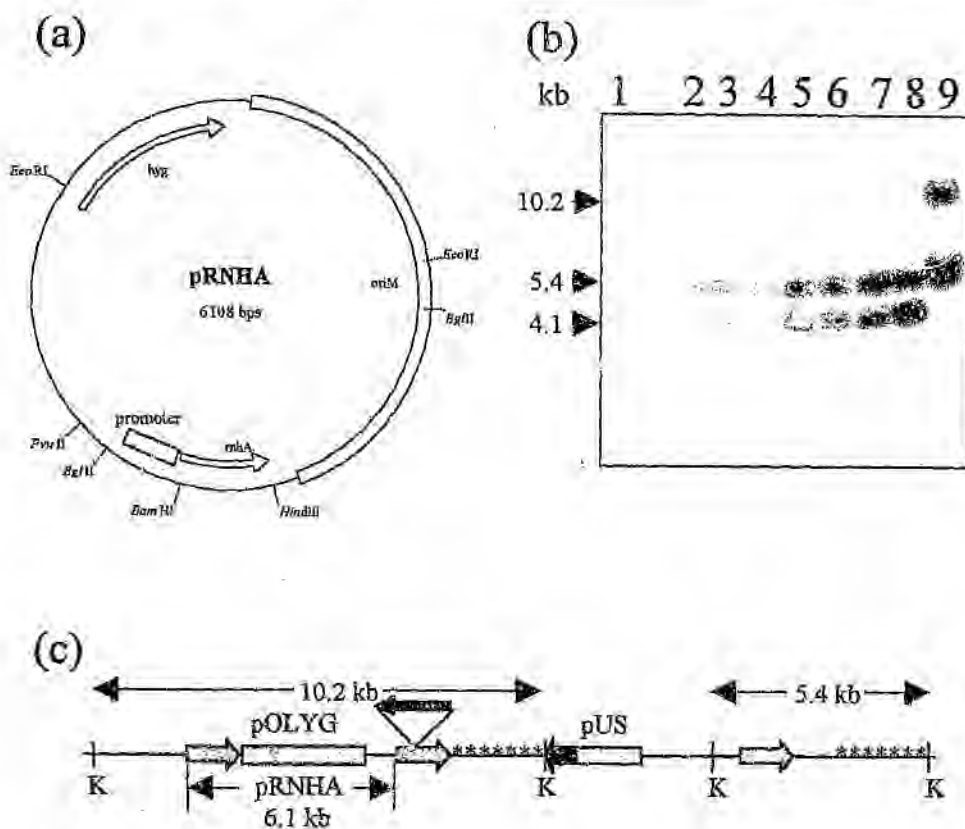


Fig. 11. Complementation of *mc*²155::pPBAS with *M. smegmatis* RNase HI. (a) The *rnhA* gene was cloned as a *Bam*HI-*Hind*III cassette from pMAL::RNH with the 350 bp *Bgl*II-*Bam*HI promoter region from pSD1 cloned ahead of the gene into pOLYG. (b) Southern analysis of a *Kpn*I digest of chromosomal DNA from *mc*²155 (lane 1) and *mc*²155::pPBAS [pRNHA] sucrose survivors (lanes 2-9) with integration of pRNHA at the *rnhA* locus (lanes 4 and 9). The blot was probed with a 2 kb *Stu*I-*Bam*HI fragment from pSD2. (c) Representation of the possible homologous recombination of pRNHA with the chromosome of *mc*²155::pPBAS. Sizes of the alleles expected with a *Kpn*I (K) digest and the *Stu*I-*Bam*HI probe are shown in (c). The new expected allele size after homologous recombination with pRNHA is indicated and the hybridisation region of the probe is shown by asterisks (*).

Recombination at the *rnhA* allele with *rnhA::aph* is polarised. Southern analysis of SCO events using the enzyme *Bam*HI cannot distinguish between SCO events occurring upstream of the *aph* cassette (Type I SCO), or downstream of the *aph* insertion (Type II SCO). The restriction enzyme site *Kpn*I occurs once in the cloned DNA and once in the vector and hence restriction analysis with this enzyme allows differentiation between Type I and Type II SCO events (Fig. 12a).

(a)

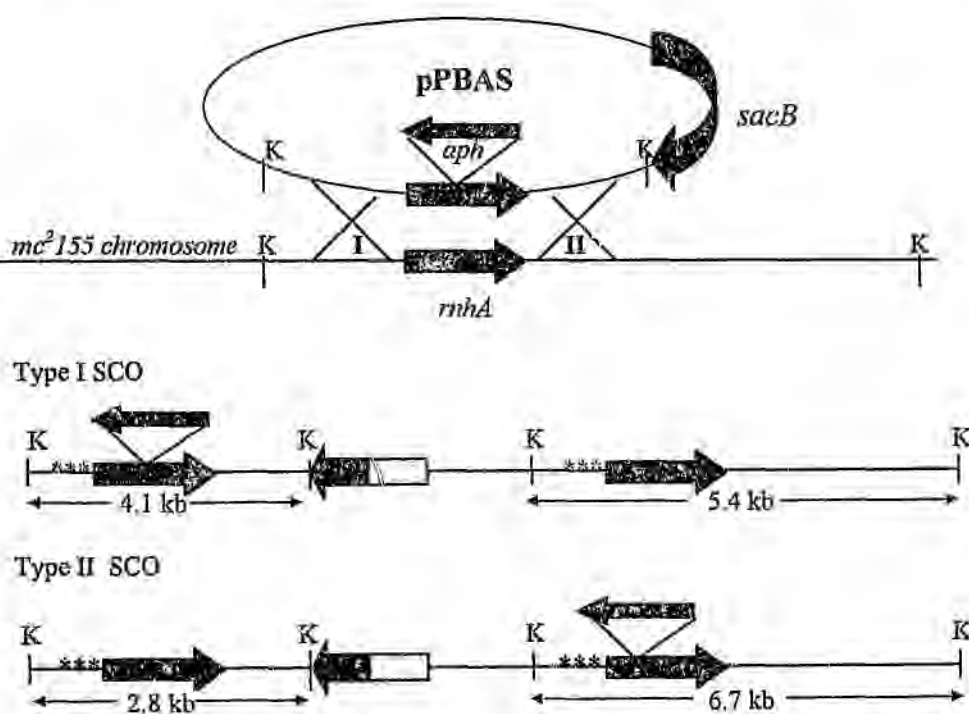


Fig 12(a). Types of SCO events with pPBAS: Diagrammatic representation of SCO Type I or Type II between the chromosomal DNA and the homologous region of the plasmid pPBAS, with the resultant chromosomal arrangement and sizes of the different alleles expected from a *Kpn*I (*K*) digest. Localisation of the probe is shown by asterisks (*).

(b)

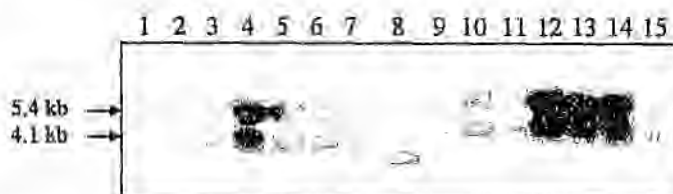
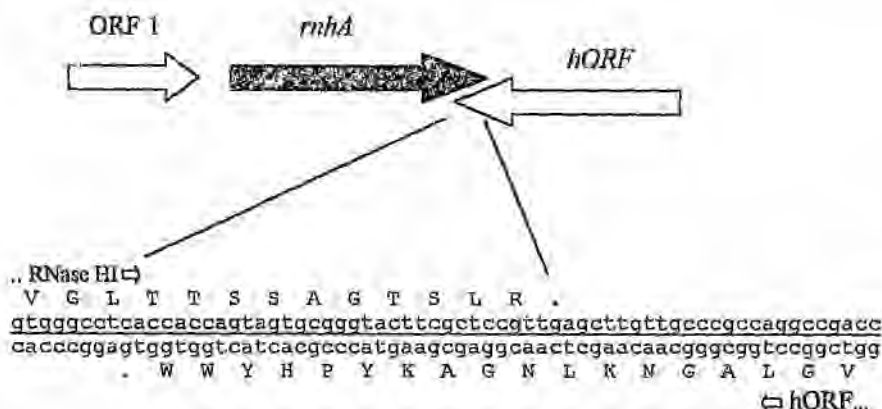


Fig. 12 (b) Types of SCO events with pPBAS: Southern analysis of *KpnI*-digested chromosomal DNA of *mc²155* (lane 1) and 14 recombinants selected on Km and sucrose (lanes 2-15) probed with the 400 bp *Bam*HI-*Sac*I probe. A deletion mutant was recovered (lane 8) but was not characterised further. The localisation of the probe is shown by asterisks (*) in 12(a), and the fragment sizes obtained with a *KpnI* (K) digest are indicated by arrows.

A Type I SCO yields restriction products of 4.1 and 5.4 kb, whereas a Type II SCO yields products of 6.7 or 2.8 kb. Chromosomal DNA from 25 colonies selected on Km was restricted with *KpnI* and subjected to Southern analysis. Only Type I SCOs were recovered (Fig. 12b). This bias in the recovery of SCOs, together with the inability to recover DCOs after application of positive selection, strongly suggested that the lethal polar effect was exerted by a combination of the insertion of the *aph* cassette and the truncation of the homologous DNA by the vector DNA inserted at the *Bam*HI site 2 kb downstream of the *rnhA* gene. This polar effect could be exerted by abolition of transcriptional read-through from *rnhA* to an essential gene in one allele in conjunction with truncation of this downstream gene in the other allele. Conversely, if this essential gene was encoded on the opposite strand and shared part of the DNA segment encoding *rnhA*, its promoter region could be truncated by vector sequence in one allele, and disrupted by the *aph* cassette in the other allele. Although no open reading frame could be discerned downstream of *rnhA*, sequence information from the complementary strand of *rnhA* indicated the presence of an ORF, designated *hORF*, that overlaps the last 30 bp of *rnhA* (Fig. 13a). *hORF* displays substantial homology to a hypothetical ORF from *Synechocystis* sp. PCC6803 (Kaneko *et al.*, 1995) and to the N-terminal part of a hypothetical protein, ORF378, from *Pseudomonas stutzeri* (Glockner and Zumft, 1996).

(Fig. 13b). Sequence from the 3' junction of the cloned DNA did not show any homology to any DNA or protein sequence in the Genbank and EMBL databases (May, 1998).

(a)



(b)

```

S.sp: 40 .....SFLAISICAYLTIVKTNSPLLMAPEFGATSVLIFGVPSPLAQPRNVIGGNLLAALV 94
cons:   S::A::I A LT T: P L::P::A:: L: G :: PL:QPRNVIGG:::A:V
hORF: 1  VVETIVSLVALAIASLTAITHDPWLIPPLAASAAIVGGCAHFFLSQPRNVIGGHVMSAIV 11
cons:   L :P :ASA L ::: P :QP :VIGG::: :
P.st: 54 .....LVAPAGASAVLLFAASSSPFAQPWSVIGGNLIATVI 89

S.sp: 95  SLTVLHFLGSSPMAMGAAVSLAIGIMQFTGTLHFPSPGAVALVMM--TR-PDWNF-IITPT 151
cons:   : V : A A :IA:G M : H:P: A A:: : P W:F :L T
hORF: 62  GVMVGMTALDALWAGAVAGALALGMMKLLRMSHSPAAATAMIAVV--PEIPRWEEFVLVGT 120
cons:   GV :G:T:L A:A: L L : L H:P:AA A::AVV PEI :L
P.st: 90  GVSIGLTSLPMSMVAALCVCLVLVCLFSLRCLHFPGAALALVAVVGSPEIHQLGYALLYPV 150

S.sp: 152  LEGSILLVLCVAVFNNLAERTYPKHWL. 179
cons:   ::V A N A YP W
hORF: 121  AAVILVVVGLAGNKLNGAK---YPHYWW. 145
cons:   A L:::A N YP :
P.st: 151  AFNSLLLVALLVNNLTR-HPYKPRIL..... 177 (total length 378)

```

Fig. 13.(a) Arrangement of *hORF* on the complementary strand to *rnhA* and detail of the overlapping region. (b) Alignment of the complete *hORF* with the relevant complementary regions of the hypothetical protein from *Synechocystis* sp. (S.sp; Genbank Acc. No. D64005) and ORF378 from *Pseudomonas stutzeri* (P.st; Genbank Acc. No. Z73914). The consensus (*cons.*) is indicated between *hORF* and S.sp. and between *hORF* and P.st. with similar amino acids denoted by a colon (:).

Allelic replacement at the *rnhA* locus. On the basis of these results, a control substrate to test homologous recombination at the *rnhA* locus was created (Fig. 14):

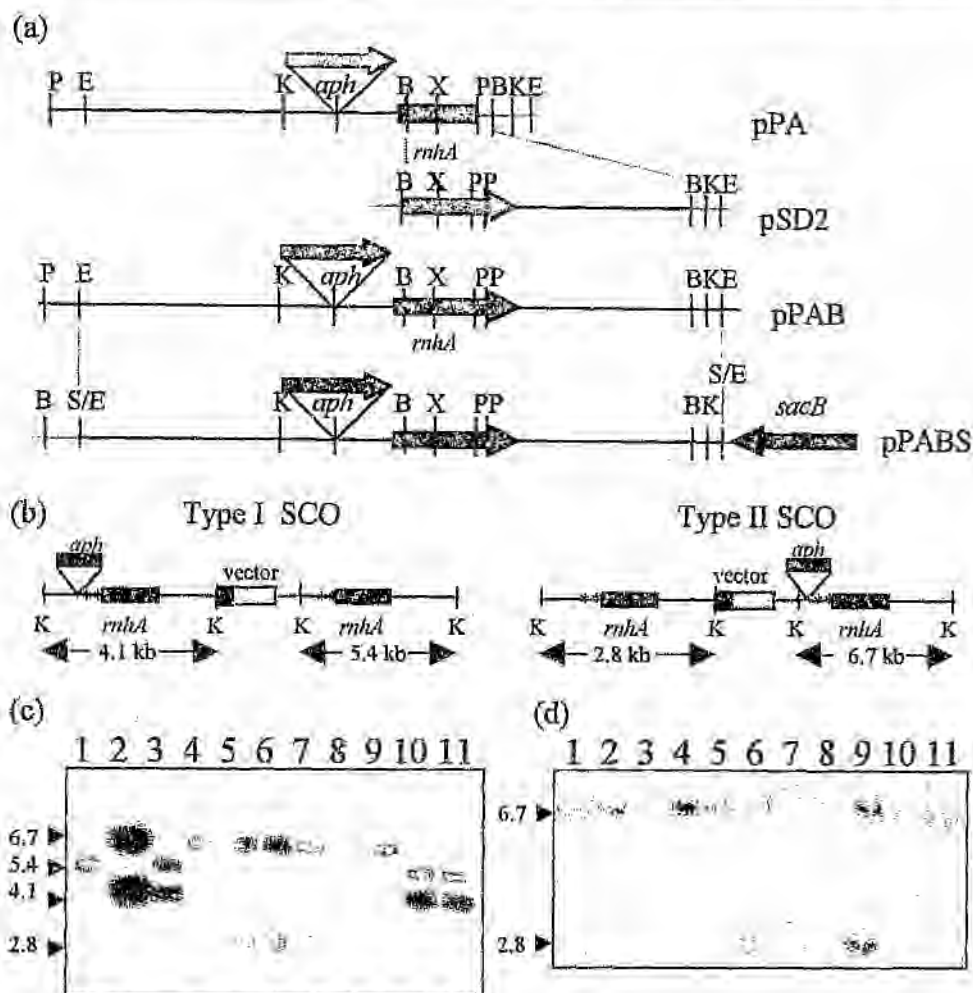


Fig. 14. Homologous recombination with pABS. (a) The *aph* cassette was cloned into the *Bgl*II (Bg) site of pSD1 to form pPA. The 350 bp *Bam*HI (B) fragment from pPA was replaced with the 2.3 kb *Bam*HI fragment from pSD2 to form pPAB. The 6.6 kb *Eco*RI (E) fragment from pPAB was blunt-ended and cloned into the unique *Sma*I (S) site of pUS to form pABS. (b) Diagram of the chromosomal arrangement of the different alleles and sizes expected from a *Kpn*I (K) digestion of Type I or Type II SCO with pABS. Southern of chromosomal DNA digested with *Kpn*I of (c) *mc*²¹⁵⁵ (lane 1) and recombinants selected on Km (lanes 2-11) and (d) colonies from a SCO selected on sucrose (lanes 1-11), probed with the 400 bp *Bam*HI-*Sac*I probe (*). The sizes (in kb) of the bands obtained are indicated by arrows. Abbreviations of restriction sites not already mentioned are as follows: P, *Pst*I; X, *Xho*I.

The plasmid pPABS was constructed such that the *aph* cassette was inserted upstream of the promoter region of *rnhA* so as to a) negate a possible transcriptional polar effect arising from interruption of the *rnhA* gene and b) relieve the cell of a potential negative effect arising directly from inactivation of the *rnhA* gene. The *aph* cassette was therefore cloned into the *Bgl*III site of pSD1 (a 5' shift of 513 bp from its previous location within the *rnhA* gene) to form pPA and the *Bam*HI cassette from pSD2 was subcloned into pPA to form pPAB (Fig. 14a). The 6.6 kb *Eco*RI fragment from pPAB was then subcloned into pUS to form pPABS and this construct was electroporated into *M. smegmatis*. Ten colonies obtained from the transformation were analysed for Type I or Type II SCOs; both types of SCO were recovered, indicating that alleviation of the polar effect observed with pPBAS had occurred (Fig. 14c). One of the SCOs was plated to single colonies on sucrose, and the majority of colonies recovered (10/12) were found to be DCOs although some (2/12) spontaneous *Suc*^R *sacB* mutants were also recovered (Fig. 14d).

Allelic replacement of *rnhA* with *rnhA::fs* confers a lethal phenotype. Since no adverse effects on recombination were observed with pPABS, a further construct was devised that introduced a frameshift mutation to inactivate the *rnhA* gene (Fig 15a). The *Bam*HI site 9 bp downstream of the start of the *rnhA* gene in pPS was blunt-ended and re-ligated to form pPS- δ B. This manipulation created a *Clal* site in place of the *Bam*HI site and the frameshift mutation was confirmed by digestion with *Clal* (data not shown). Since the site of the *aph* insertion in pPABS was in the *Bgl*III site, it was considered that sufficient sequence upstream of *rnhA* was present in pPS- δ B to give an analogous recombination phenotype to that of pPABS. A similar amount of sequence downstream (320 bp) of the frameshift as upstream (400 bp) was present in this clone, therefore the *Eco*RI-*Hind*III fragment from pPS- δ B was blunt-ended and cloned into the *Sma*I site of pUSH to form pUSH- δ B. Homologous recombination was achieved with this substrate albeit at a 2-fold lower transformation efficiency than with the constructs with longer flanking sequence. Only SCOs were initially recovered due to the presence of the resistance cassette in the vector. All SCOs recovered were Type I SCOs (Fig. 15b). A SCO was then passaged on LA without antibiotic to allow for a DCO event to occur, and plated on sucrose. Twenty survivors on sucrose were patched onto Hyg to distinguish

DCOs (8/20) from spontaneous *Suc^R sacB* mutants (12/20). Southern analysis of 7 DCOs showed that all DCOs had reverted to a WT *rnhA* allele (Fig. 15c).

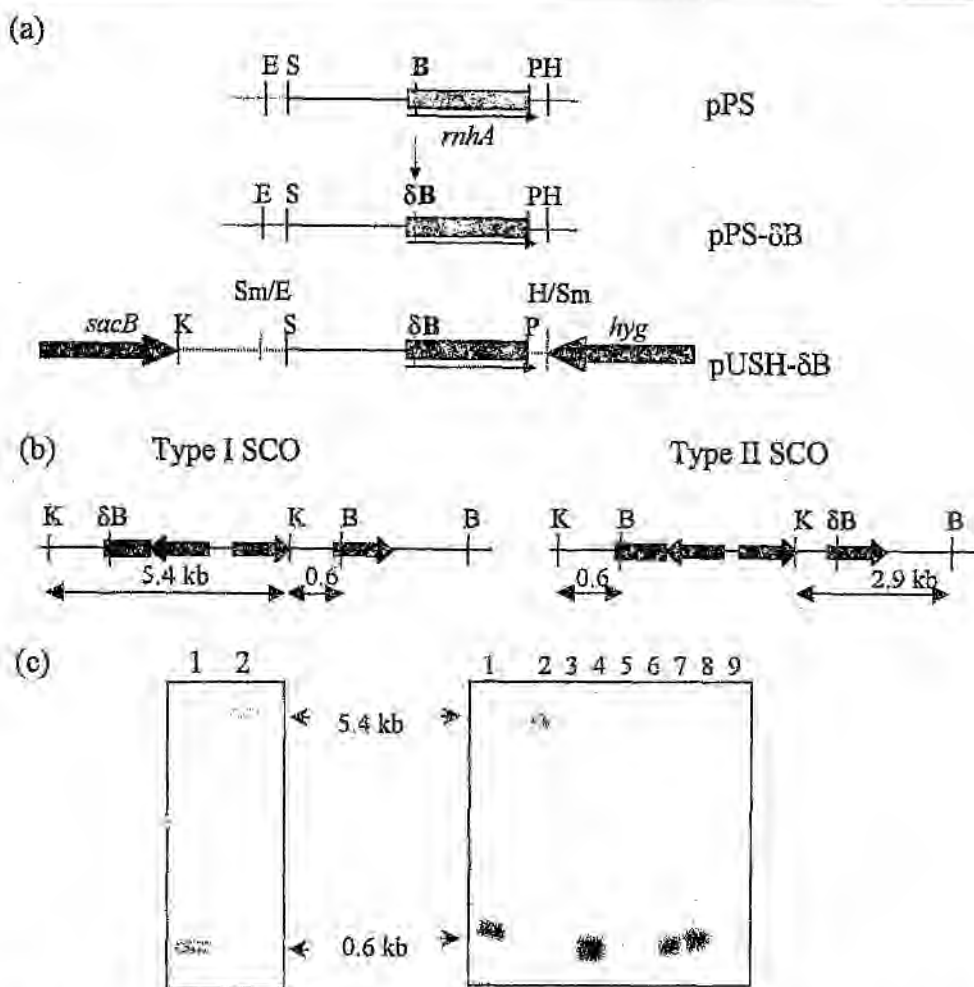


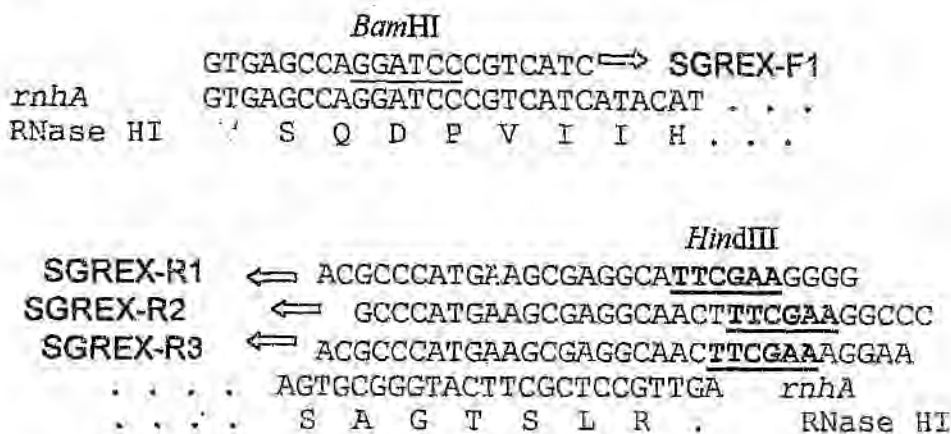
Fig. 15. Homologous recombination with pUSH- δ B. (a) pUSH- δ B was constructed by firstly blunting the unique *Bam*HI site in pPS to form pPS- δ B. The *Eco*RI (E)-*Hind*III (H) fragment from pPS- δ B was then subcloned into the *Sma*I (Sm) site of pUSH to form pUSH- δ B. (b) Diagram of the chromosomal arrangement of the different alleles and sizes expected from a Type I or Type II SCO with pUSH- δ B after *Bam*HI (B)-*Kpn*I (K) digestion. (c) Southern of a *Bam*HI-*Kpn*I digest of chromosomal DNA from *mc*³155 (lane 1), Type I SCO (lane 2) and DCOs obtained after sucrose selection (lanes 3-9). Sizes indicated are in kb. P; *Pst*I, S; *Sac*I.

It was considered possible that an *rnhA* mutant might be broth or temperature sensitive, so the experiment was repeated by passaging a SCO on MA at 30°C before selecting on MA+sucrose at 30°C. Again, only WT alleles were recovered from 20 DCOs screened (data not shown). This suggests that *rnhA* is essential in *M. smegmatis* under the selection conditions employed, although final confirmation of this awaits a complementation experiment where *rnhA* is supplied *in trans* on a plasmid.

3.4 Biochemical characterisation of recombinant *M. smegmatis* RNase HI

Construction of the *rnhA* expression cassette. Recombinant *M. smegmatis* RNase HI was overproduced as a maltose binding protein-RNase HI fusion protein using the pMAL-C2 system. The pMAL-C2 vector contains the *malE* gene, encoding a maltose binding protein (MBP), ahead of a multiple cloning site. An in-frame fusion of *malE* and *rnhA* allows purification of fusion protein by affinity chromatography after induction of protein production by IPTG. In addition, the pMAL-C2 vector encodes a Factor Xa cleavage site immediately before the multiple cloning site to allow cleavage of the MBP domain from the recombinant fusion protein. Primers SGREX-F1 and SGREX-R1 were designed (Fig. 16a) to amplify the coding region of the *rnhA* gene. The reverse primer, SGREX-R1, was designed with a *HindIII* site to allow for directional cloning in *XmnI*-*HindIII*-digested pMAL-C2. The primers specifically amplified a band from chromosomal DNA at an annealing temperature of 60°C. However, a new reverse primer (SGREX-R2) was subsequently designed to include a stop codon before the *HindIII* site. The primer pair SGREX-F1 and SGREX-R2 did not amplify the expected 488 bp fragment from chromosomal DNA although a specific 900 bp fragment was amplified at the same annealing temperature (data not shown). This fragment was not characterised further. An additional reverse primer, SGREX-R3, was therefore designed which amplified DNA both from chromosomal DNA and from the PCR product of SGREX-F1 and SGREX-R1. Non-specific bands were observed with the chromosomal amplification products, whereas only specific amplification was observed with the original PCR amplicon as a substrate reaction (data not shown). The amplification product from the PCR substrate was restricted with *HindIII* and cloned into pMAL-C2 to generate pMAL::RNH (Fig. 16b).

(a)



(b)

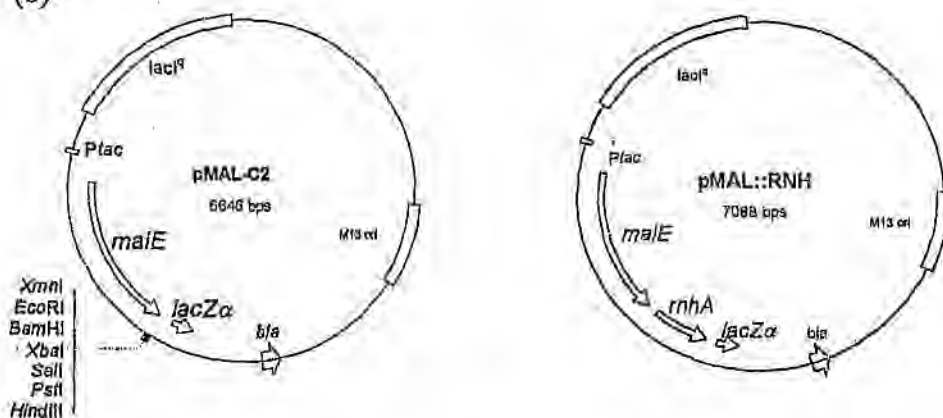


Fig. 16. Construction of the *rnhA* expression cassette. (a) Design of the PCR primers used to amplify the *rnhA* gene. The restriction sites present in the primers are underlined. The 5' end and the 3' end of the *rnhA* gene, and the beginning and end of the amino acid translation is shown underneath the primers. (b) The expression vectors pMAL-C2, containing the in-frame *malE::lacZα* gene fusion, and pMAL::RNH, containing the *malE::rnhA* gene fusion.

Clones containing the amplified *rnhA* gene were isolated by Southern analysis of restriction digests of putative clones. One positive clone was picked. The sequence of the blunt-end junction at the beginning of the gene was confirmed by ds sequencing of pMAL::RNH with the MalE primer (Table IV). The *Bam*HI-*Hind*III fragment containing the rest of the amplified region was subcloned into M13 for confirmatory sequencing. Four sequence errors resulting in three amino acid substitutions (S73G, H116L, F121L) were found, but the wild type sequence was restored by replacing the 298 bp *Bam*HI-*Sal*I fragment, followed by the 102 bp *Pst*I fragment from pMAL::RNH, with the corresponding restriction fragments from pSD2. The final sequence was confirmed by subcloning the *Bam*HI-*Hind*III fragment into M13 and sequencing the entire ORF.

Purification of recombinant MBP::RNase HI. Expression of recombinant MBP::RNase HI was induced in *E. coli* TB1 cells carrying pMAL::RNH by the addition of IPTG. In parallel, TB1 cells containing pMAL-C2 were induced to overexpress recombinant MBP:: β -GAL. Whole cell extracts of uninduced cells, and cells induced for MBP:: β -GAL and MBP::RNase HI, were analysed on an SDS-PAGE gel (Fig. 17a). Induced protein separated at the sizes expected for the two recombinant proteins. MBP:: β -GAL and MBP::RNase HI were purified by affinity chromatography. An activity gel was prepared by electrophoresis of the whole cell extracts and the purified protein fractions from both procedures on a denaturing SDS-PAGE gel co-polymerised with a radio-labeled RNA-DNA substrate. Renaturation of the protein *in situ* allowed for localised hydrolysis of the substrate. Products of both of the purification procedures showed no discernable contamination with endogenous *E. coli* RNase HI, while RNase H activity was found to be associated with the 60 kDa MBP::RNase HI fusion protein (Fig. 17b).

The substrate specificity of the hydrolysis reaction was analysed by activity gel analysis using radio-labelled RNA-RNA substrate or ssRNA substrate instead. No hydrolysis of either substrate was noted (data not shown). The recombinant protein was cleaved with Factor Xa, but although the correct size cleavage products were visible on a gel, indicating that there are no factor Xa cleavage sites within RNase HI (data not shown), no RNase H activity was associated with the cleaved product on an activity gel. However, the positive

control consisting of commercial *E. coli* RNase HI showed no activity either, suggesting that no conclusion as to the RNase HI activity of the cleaved recombinant RNase HI could be drawn. All biochemical characterisation was therefore carried out using the fusion protein.

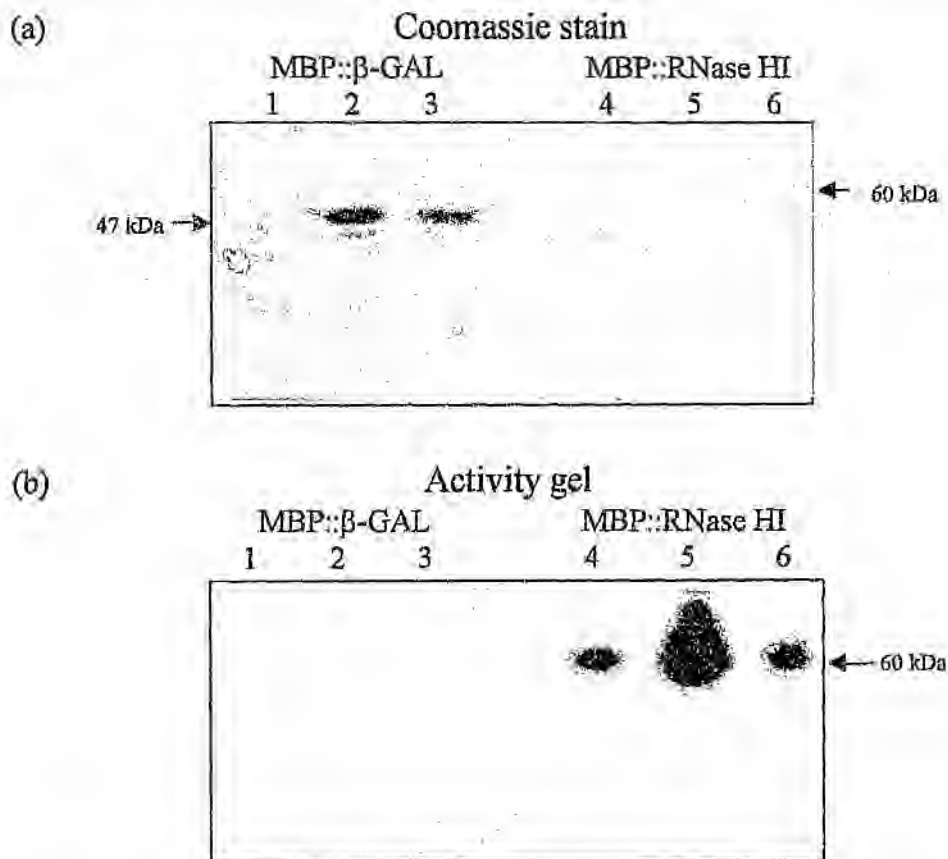


Fig. 17. (a) Coomassie stain of SDS-PAGE analysis of the overexpression and purification of MBP:: β -GAL and MBP::RNase HI fusion proteins. Aliquots of un-induced cell extract of TB1 cells carrying pMAL-C2 and pMAL::RNH (lanes 1 and 4, respectively), induced cell extract showing overproduction of the 47 kDa MBP:: β -GAL and the 60 kDa MBP::RNase HI fusion proteins (lanes 2 and 5, respectively), and purified MBP:: β -GAL and MBP::RNase HI (lanes 3 and 6, respectively). (b) Negative reproduction of an activity gel run in parallel with the Coomassie-stained gel. Lanes 1-6 are as above. RNase HI activity (arrowed) is associated with the 60 kDa MBP::RNase HI fusion product.

Site selectivity of substrate cleavage of MBP::RNase HI. To compare the cleavage site selectivity of MBP::RNase HI and *E. coli* RNase HI, a substrate suitable for use in a liquid hydrolysis assay was synthesised. A 71 bp RNA-DNA hybrid was generated by annealing labelled RNA from SP6 RNA polymerase run-off transcription of pGEM-GAG A (Mizrahi, 1989) to complementary ssDNA from the appropriate M13 subclone (Fig. 18a). Recombinant MBP:: β -GAL showed no hydrolysis of this substrate in comparison to the activity shown by MBP::RNase HI (Fig. 18b).

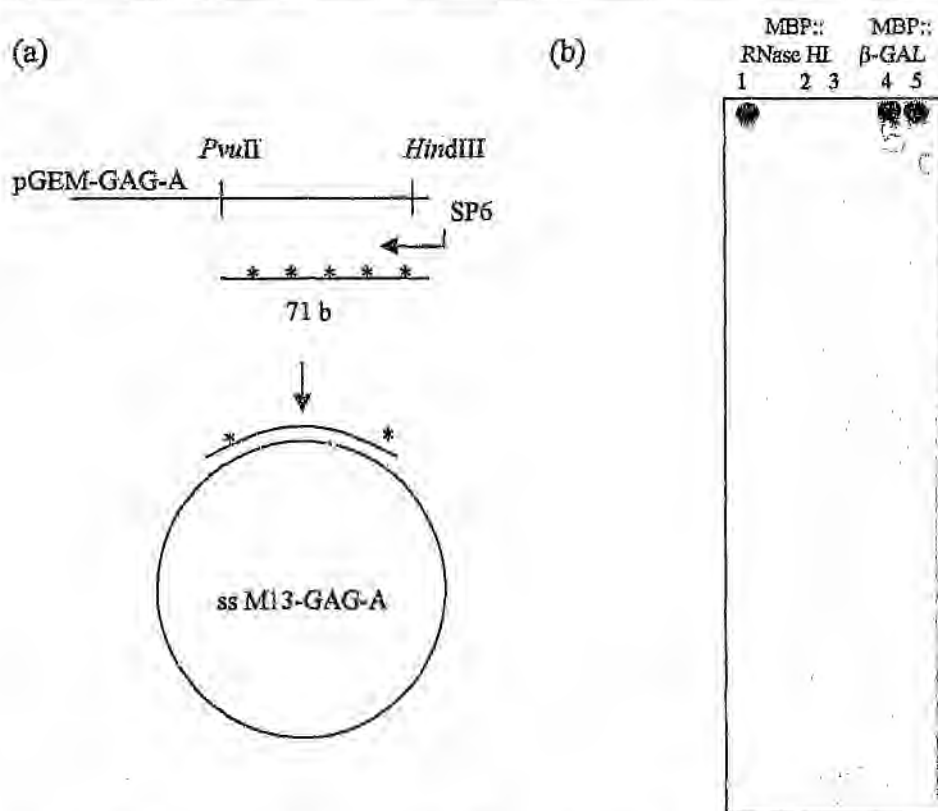


Fig. 18. RNase HI liquid hydrolysis assay (a) A 71 base radio-labeled RNA substrate was generated by SP6 run-off transcription of *Pvu*II-restricted pGEM-GAG-A. The labelled RNA was hybridised to complementary ssDNA (in the form of the appropriate M13 clone) to generate the substrate for the liquid RNase HI assay. (b) Hydrolysis of the RNA-DNA hybrid by MBP::RNase HI after a 0 min (lane 1), 1 min (lane 2) and 10 min (lane 3) incubation. No hydrolysis of the substrate was seen with MBP:: β -GAL after either 1 min or 10 min (lanes 4 and 5 respectively).

A 'footprint', representing the product distribution generated by the RNase HI activities of *E. coli* RNase HI and MBP::RNase HI, was therefore obtained by quenching aliquots of the reactions at different time points and fractionating the products on a PAG (Fig. 19).



Fig. 19. PAG electrophoresis of a time course of hydrolysis products generated by *E. coli* RNase HI (lanes 1-7) and MBP::RNase HI (lanes 8-14). Aliquots of the reaction mix were quenched at 0 min (lanes 1 and 8), 0.5 min (lanes 2 and 9), 1 min (lanes 3 and 10), 2 min (lanes 4 and 11), 4 min (lanes 5 and 12), 8 min (lanes 6 and 13) and 16 min (lanes 7 and 14). M; poly-dT₁₂₋₁₈.

E. coli RNase HI and MBP::RNase HI showed similar product distributions. Similar rates of cleavage were obtained by *E. coli* RNase and MBP::RNase HI.

3.5 Interaction of RNase HI with heat shock proteins

The molecular chaperones, DnaK and DnaJ, have been implicated in the control, either directly or indirectly, of the levels of RNase HI at the post-translational level (Foster and Marinus, 1992). Since the proteins of the DnaK chaperonin machine have been found to interact with a small number of proteins in their native state (Wewrzynów and Zylicz, 1995), it was of interest to investigate the possibility that the chaperonin proteins might also interact directly with native RNase HI. The interaction of RNase H with DnaK and DnaJ was investigated using *E. coli* proteins. Protein-protein interaction was monitored in real time by surface plasmon resonance (SPR) technology utilising the BIAcore™ SPR apparatus (Malmqvist, 1993; Johnsson *et al.*, 1991). Briefly, protein is immobilised by covalent linkage to a dextran matrix on the surface of a chip within a miniaturised integrated fluidic cartridge. The change in the resonance angle of the light reflected off this surface, caused by binding of protein, is detected by a diode array, translated into response units (RU), and is directly proportional to the amount of protein bound to the surface. The maximum amount of analyte expected to bind to protein immobilised to give a specific RU can be calculated by comparing the molecular weights of the proteins involved, and the number of binding sites available for interaction. Analyte is flowed over immobilised protein and a sensorgram depicting the interaction of the analyte with the bound protein is generated.

DnaJ binds to σ^{32} . DnaJ has been shown to form a stable complex with the transcription factor σ^{32} (Liberek *et al.* 1995), and this interaction was used as a positive control to set initial parameters to investigate the interaction of RNase HI with the heat shock proteins. DnaJ was immobilised to give a relative response of 6000 RU. A 1:1 ratio of binding of σ^{32} to DnaJ would give a maximum response (R_{max}) of 4682 RU, although σ^{32} has been shown to bind in a 1:2 ratio with DnaJ under certain conditions (Liberek and Georgopoulos, 1993) in which case an R_{max} of 2341 RU could be expected. Binding of σ^{32} to DnaJ was monitored at 25°C and 37°C in HBS buffer and A buffer. No interaction

was found at 25°C in HBS buffer pH7 or pH8, or in A buffer. Relatively weak interaction was observed at 37°C in HBS buffer (pH8) while no interaction was observed with HBS (pH7) (data not shown). The maximum in response units (5007 RU), representing complete saturation of bound DnaJ by σ^{32} , was recorded at 37°C in A buffer (Fig. 19). Hence buffer A was used in subsequent experiments and all interactions were monitored at 37°C.

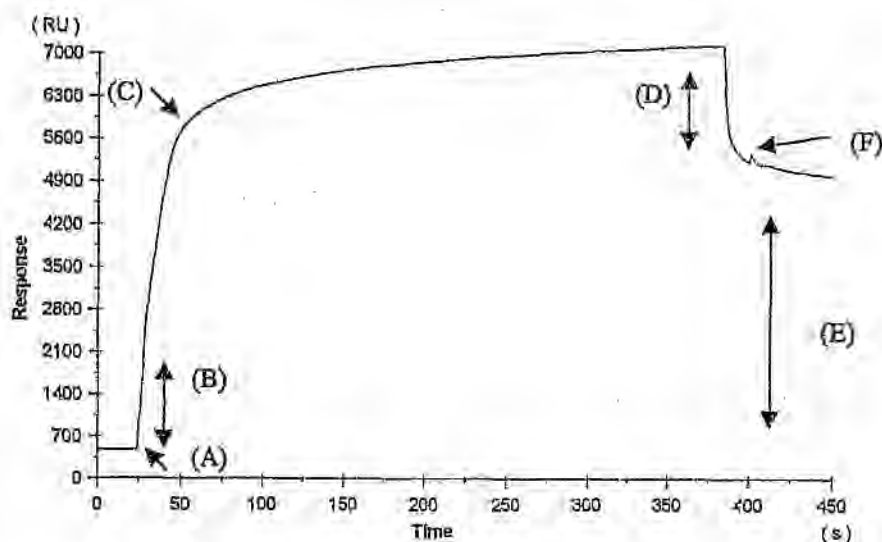


Fig. 20. DnaJ binds to σ^{32} : σ^{32} (25 $\mu\text{g/ml}$ in A buffer) was flowed over DnaJ (immobilised at 6000 RU) at 5 $\mu\text{l/min}$ at 37°C. The addition of analyte is marked by arrow (A). The analyte buffer contributes initially to the rise in refractive index noted (B), followed by a curve representing association of analyte protein with the immobilised protein (C). As the analyte leaves the cell, the buffer contribution towards the refractive index is removed (D), and the ΔRU observed (E) represents the remaining associated protein. The curve following the removal of the analyte mix (F) represents the dissociation of the bound protein and its removal from the cell by the wash buffer. After exit of the analyte mix, $\Delta\text{RU}=5007$ which represents 106% of the maximum expected binding if DnaJ and σ^{32} associate with a 1:1 ratio.

DnaJ interacts with RNase HI. RNase HI was immobilised to give 1000 RU and DnaJ was flowed over the bound RNase HI. Binding of DnaJ to RNase HI was initially observed in HBS buffer (pH8) (Fig. 21a). No significant difference was observed between HBS buffer and buffer A for this reaction (data not shown), therefore A buffer was used for further experiments for ease of comparison with other data. Observation of low affinity interactions can be facilitated by increasing the concentrations of analyte and ligand, and, since the off-rate is relatively fast, increasing the flow rate of buffer through the microfluidic cartridge allows analysis of more of the dissociation phase (Boniface and Davis, 1994). Most subsequent experiments therefore utilised a flow rate of 15 $\mu\text{l}/\text{min}$.

The specificity of the DnaJ and RNase HI interaction was then tested by monitoring the amount of RNase HI bound to different concentrations of immobilised DnaJ. DnaJ was immobilised to give relative responses of 1500 RU or 4500 RU, RNase HI as the analyte was flowed over the surface of each chip at a concentration of 100 $\mu\text{g}/\text{ml}$ at a rate of 15 $\mu\text{l}/\text{min}$ (Fig. 21b). An increase in the concentration of immobilised DnaJ resulted in a proportional increase in the amount of bound RNase HI, indicating that the binding observed was a function of bound DnaJ.

RNase HI was then flowed in different concentrations over immobilised DnaJ (Fig. 21c). An increase in binding proportional to the increase in concentration of RNase HI was observed. This provided confirmation of the specificity of the binding observed, since different concentrations of RNase HI flowed over an empty surface showed negligible increase in binding (data not shown).

A competitor experiment was designed where RNase HI was included in the analyte buffer together with DnaJ, to provide competition for DnaJ in the liquid phase to the bound RNase HI (Fig. 21d). Interestingly, an increase in the RUs was noted instead of a decrease, suggesting that one molecule of DnaJ may bind more than one molecule of RNase HI.

Data produced by the sensorgrams depicted in Fig. 21(c) were analysed by Sergei Khilko using the BIAlogue Kinetics Evaluation Software. An association constant of 630 (± 272) $\text{M}^{-1}\cdot\text{sec}^{-1}$ for the interaction of RNase HI and DnaJ was calculated. The analytical software was unable to compute a dissociation constant for the reaction due to

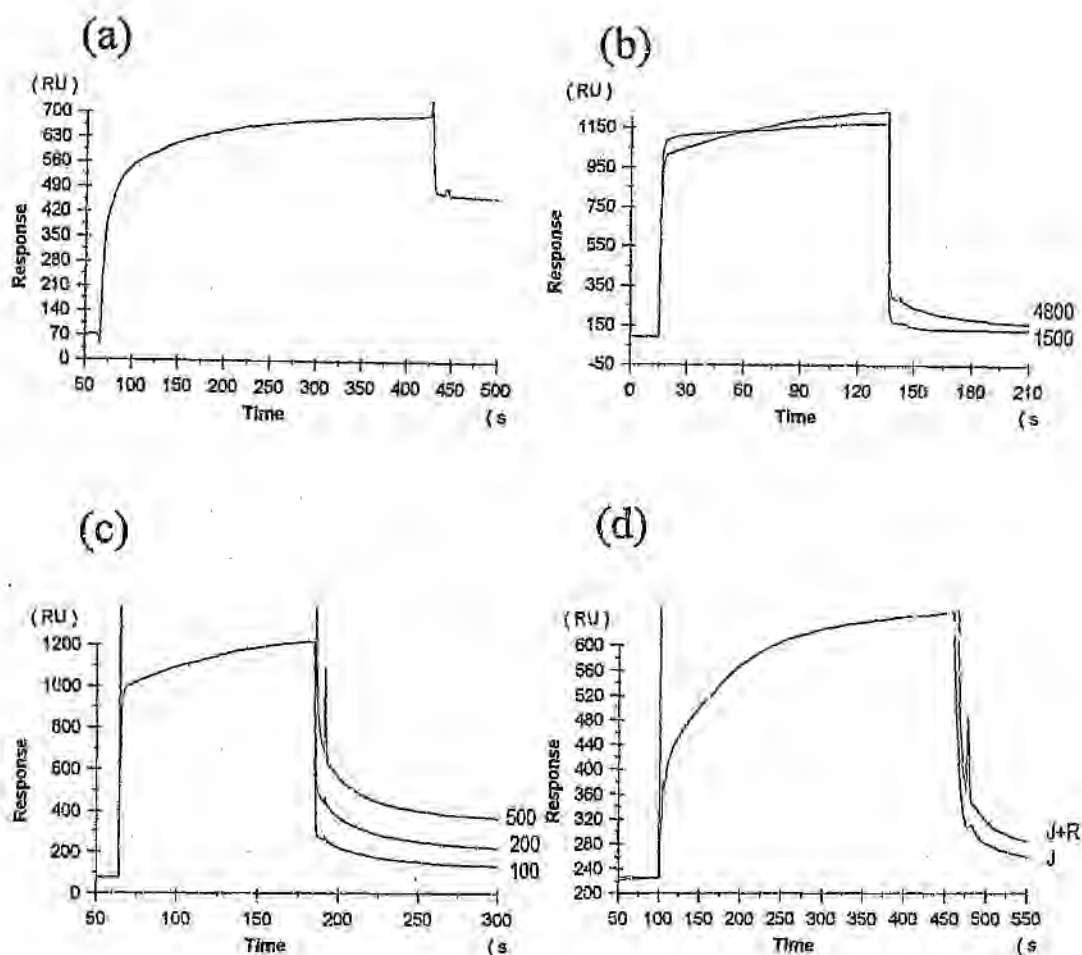


Fig. 21. Interaction of RNase HI with DnaJ at 37°C. (a) DnaJ (30 µg/ml in HBS pH8) was flowed over RNase HI (immobilised at 1000 RU) at a flow rate of 5 µl/min. The change in RU (410) observed represented 18% of the maximum binding possible if RNase HI and DnaJ bind in a 1:1 ratio. (b) RNase HI (100 µg/ml in A buffer) was flowed over DnaJ (immobilised at 1500 RU or 4800 RU) at 15 µl/min. The change in RUs of 97.8 and 238.7 respectively, represented 15% and 11% of the maximum binding expected. (c) RNase HI (100 µg/ml, 200 µg/ml, and 500 µg/ml in A buffer) was flowed over DnaJ (immobilised at 4800 RU) at 15 µl/min. At RNase HI concentrations of 100 µg/ml, 200 µg/ml and 500 µg/ml, 11%, 22% and 50% of R_{max} was noted, respectively. (d) RNase HI was immobilised at 1200 RU and either DnaJ alone (15 µg/ml in A buffer), or DnaJ (15 µg/ml) together with RNase HI (96 µg/ml) in A buffer, was flowed over immobilised RNase HI at 5 µl/min. DnaJ alone bound at 8 % of R_{max} ($\Delta RU=215$), while DnaJ and RNase HI together showed increased binding ($\Delta RU=378$).

inherent analytical limitations of the software for sensorgrams produced by the manual input of a queue of instructions. These limitations exist due to small delays in the reading, translation and execution of these instructions compared to encoded instructions in an automated program. This delay has a particular impact on initial measurements in the dissociation phase. However, a dissociation constant in the range of $10^{-2} - 10^{-3} \text{ sec}^{-1}$ was estimated (S. Khilko, pers. comm.) from the same sensorgram set.

Interaction of DnaK with RNase HI. Although DnaJ is considered generally to facilitate the binding of DnaK to proteins, in some instances DnaK has been shown to recognise native proteins, such as lambda P, more efficiently than DnaJ does (Wawrzynów and Zylicz, 1995; Wawrzynów *et al.*, 1995). This interaction however, was shown to be inhibited by ATP. A possible interaction of RNase HI and DnaK was therefore investigated in A buffer without ATP (A-ATP).

DnaK was immobilised at 4500 and 6000 RU. RNase HI at 325 µg/ml was flowed over these surfaces in A-ATP buffer (Fig. 22a). RNase HI was shown to bind very weakly to DnaK. Different concentrations of RNase HI were then flowed over a fixed amount of immobilised DnaK (Fig. 22b). Again, specificity of binding was deduced from the increase in RU observed.

A competitor experiment was performed where DnaK, together with RNase HI, was flowed over bound DnaK. In contrast to the competitor experiment performed with DnaJ and RNase HI, inclusion of DnaK in the analyte mix decreased the binding observed with RNase HI alone, further supporting the specificity of the binding seen.

Interaction of DnaK with DnaJ and RNase HI. It has been shown that DnaJ has a role in 'tagging' substrate proteins to facilitate binding of these proteins by DnaK. Unlike the binding of DnaJ, or DnaK, to substrate proteins, the DnaJ-dependent activation of DnaK is absolutely dependent on the hydrolysis of ATP (Wawrzynów *et al.*, 1995; Lieberk *et al.*, 1995; reviewed by Georgopoulos *et al.*, 1994). Interaction of RNase H with DnaK in the presence and absence of DnaJ, with and without ATP, was therefore investigated (Fig. 23). DnaK was immobilised to give a relative response ~4800 RU.

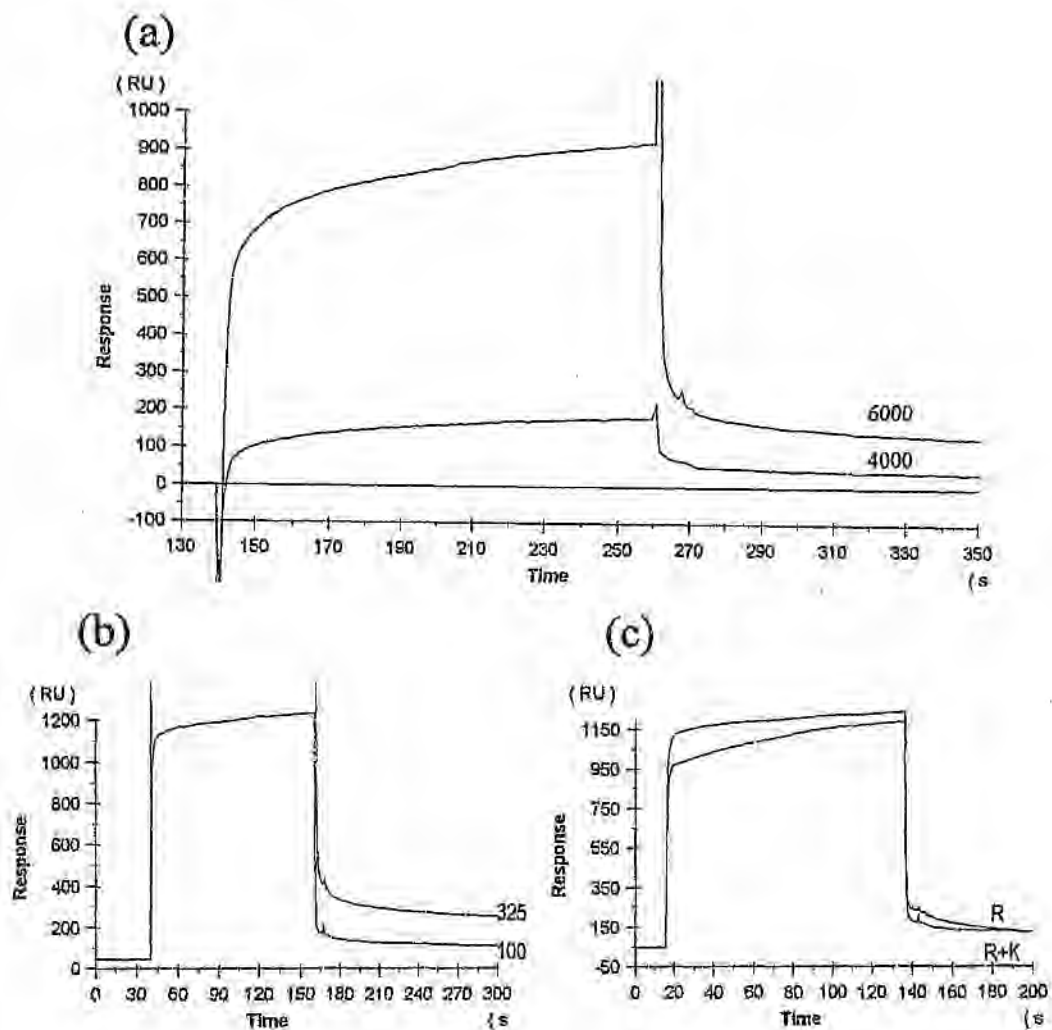


Fig. 22. Interaction of RNase HI with DnaK at 37°C in A buffer without ATP. (a) RNase HI (100 µg/ml) was flowed over DnaK immobilised at either 4000 RU or 6000 RU. Binding was observed at 4% and 12% of the maximum expected. (b) RNase HI (either 100 µg/ml or 325 µg/ml) was flowed over DnaK immobilised at 6000 RU, and binding was observed at 12% and 35% of the maximum, respectively. (c) DnaK was immobilised at 4000 RU. RNase HI alone (100 µg/ml), or RNase HI (100 µg/ml) together with DnaK (80 µg/ml), was flowed over the immobilised DnaK. RNase HI alone exhibited 4% of RU_{max} whereas a decrease in binding to 1.7% was observed when DnaK was included in the analyte mix.

DnaJ (25 $\mu\text{g/ml}$ in A buffer) was flowed at 5 $\mu\text{l/min}$ over the bound DnaK. Very little binding was observed (2% of R_{max}). In contrast, RNase HI (48 $\mu\text{g/ml}$ in A buffer) exhibited 20% of the maximum binding possible. The two proteins were then combined in the analyte mix. The binding observed with the combined protein as analyte was 50 % greater than that expected from simple addition of the binding data of the single analyte curves, suggesting that a synergistic effect on binding was obtained by including DnaJ together with RNase HI.

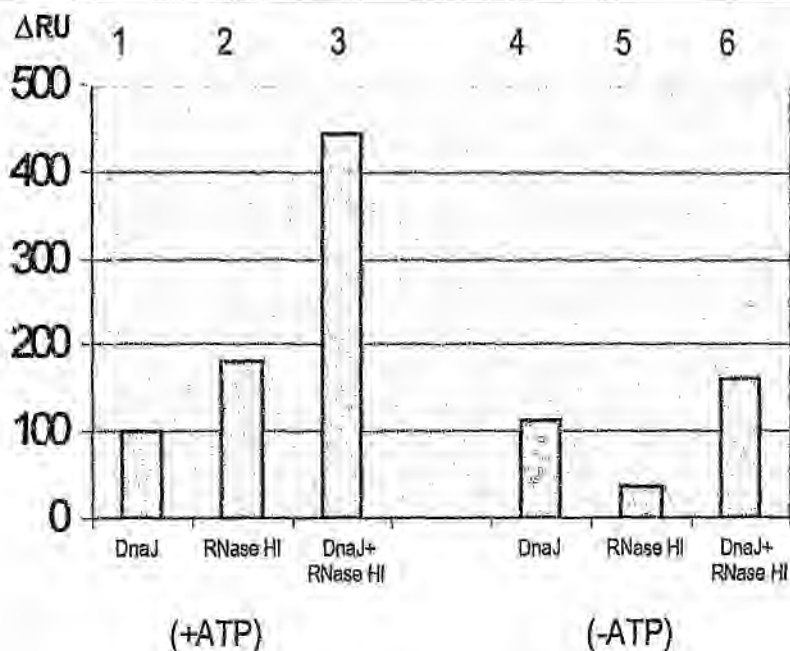


Fig. 23. Interaction of DnaJ and DnaK with RNase HI at 37°C. DnaJ (25 $\mu\text{g/ml}$) (column 1), or RNase HI (48 $\mu\text{g/ml}$) (column 2), or DnaJ (25 $\mu\text{g/ml}$) together with RNase HI (48 $\mu\text{g/ml}$) (column 3), in A buffer + ATP was flowed over DnaK (immobilised at 4800 RU) at a flow rate of 5 $\mu\text{l/min}$. DnaJ showed binding at 2% of R_{max} whereas RNase HI exhibited binding at 20% of R_{max} . The combined analyte mix of RNase HI and DnaJ showed binding at 50 % greater than that expected from combining the separate binding data. DnaJ (100 $\mu\text{g/ml}$) (column 4), or RNase HI (100 $\mu\text{g/ml}$) (column 5), or DnaJ (100 $\mu\text{g/ml}$) together with RNase HI (100 $\mu\text{g/ml}$) (column 6), in A buffer without ATP was flowed over the same surface at 15 $\mu\text{l/min}$. DnaJ exhibited binding at 6 % of R_{max} , and RNase HI exhibited binding at 4% of R_{max} . The combined analyte showed no increase over that expected from combining the separate binding data.

Similar experiments were performed in buffer without ATP. The concentration of analyte was increased so that weak binding could be observed, and the flow rate of the analyte was also increased to 15 $\mu\text{l}/\text{min}$. DnaJ (100 $\mu\text{g}/\text{ml}$) was flowed over bound DnaK. Binding at 6% of R_{max} was observed. RNase HI (100 $\mu\text{g}/\text{ml}$) was analysed similarly and exhibited 4% of the maximum binding. Combination of the two analytes resulted in no more binding than that expected from the individual binding data, suggesting that the synergism in binding observed in the presence of ATP, does not occur in the absence of ATP. DnaJ, together with RNase HI, bound DnaK in the presence of ATP 2.7-fold more strongly than in the absence of ATP, even though the analyte concentrations were more than doubled for the latter experiment.

4. DISCUSSION

M. smegmatis rnhA does not cross hybridise to *M. tuberculosis rnhA*

Although amplification of part of the *rnhA* of *M. smegmatis* was achieved using degenerate primers based on conserved blocks of amino acids from other bacterial RNases HI (Mizrahi *et al.*, 1993), this methodology was not successful for the amplification of an *rnhA* homologue from *M. tuberculosis*. However, a number of genes have been cloned using DNA probes from homologues identified in other bacteria (Alcendor *et al.*, 1995; Davis *et al.*, 1991). These genes were more than 69 % identical at the DNA level over the probe area. Interestingly, a probe consisting of an internal portion of the *rnhA* from *M. smegmatis* failed to hybridise to *M. tuberculosis* chromosomal DNA. Sequence information subsequently obtained from the *M. tuberculosis* genome sequence project allowed the identification of a putative *rnhA* from *M. tuberculosis*, which has no discernable homology at the DNA level to *M. smegmatis rnhA*, but has 22% identity at the amino acid level. In contrast, *M. tuberculosis rnhA* showed 56% identity to the *Streptomyces galbus* and *S. lincolnensis rnhA* genes (partially sequenced; Genbank Acc. Nos. X90705 and X95703 respectively) at the DNA level over the first third of the genes and a correspondingly higher identity at the amino acid level.

M. smegmatis rnhA encodes an RNase HI of 159 amino acids

The PCR probe was successfully used to isolate two overlapping clones from size-fractionated libraries of *M. smegmatis*. Hybridisation of these fragments to *M. smegmatis* DNA was observed, confirming their origin. Sequence information from these clones indicated a G+C content of 69%, consistent with the G+C content of *M. smegmatis*. Translation of this sequence allowed identification of an open reading frame with substantial homology to the amino acid sequences of other RNases HI.

A 90% occurrence of G or C was noted for the third position of each codon in the *rnhA* ORF (See Appendix D). Codon usage within the *rnhA* gene conformed to general mycobacterial codon usage: consistent with previous observations for some *M. tuberculosis* and *M. leprae* genes (Dale and Patki, 1990), the preferred codon for arginine is CGC, and AGG for arginine is notably absent. Further, and counter-intuitive with

respect to the G/C bias noted in the third position, marked preferences for GGC or GGU codons for glycine as opposed to GGG, and ACC rather than ACG for threonine, were also noted.

Two possible start codons encoding valine (GTG) were identified. GTG start codons appear to be almost as common as ATG starts in the mycobacteria, probably as a result of the high G+C content of their genomes (Dale and Patki, 1990; Honoré *et al.*, 1993). However, a study of 12 *M. leprae* start codons and 44 Streptomycete start codons indicated a preference for cytosine as the residue immediately preceding a start codon (Honoré *et al.*, 1993; Strohl, 1992) which favoured the assignment of the second GTG as the translational start. In addition, alignment of *M. smegmatis* RNase HI with the RNase HI homologues from *E. coli* and *T. thermophilus* suggested that a translational start at the second GTG would conserve the N-terminal with respect to these homologues. No ribosome binding site (RBS) with good homology to the 3' end of the 16S ribosomal RNA of *M. smegmatis*, 3'-OH UCUUJCCUC (Ji *et al.*, 1994), could be found upstream of either possible start codon. However, a lack of secondary structure in the 5' terminus of the mRNA has been found to promote translation initiation in *E. coli* to the extent that the length of Shine-Delgarno complementarity becomes irrelevant with sufficiently unstructured mRNAs (Draper, 1996). No obvious secondary structure in the 5' end of the mRNA of *rhid* was apparent, which might compensate therefore for the lack of a good RBS. In this case, GG ($\Delta G = -5.0 \text{ kcal.mol}^{-1}$), located 27 bp from the initiation codon, or GAG ($\Delta G = -4.4 \text{ kcal.mol}^{-1}$), located 1 bp from the initiation codon, might suffice (Fig. 5).

Interestingly, a region with requisite homology, AG---AGGA, was found immediately downstream of the assigned start ($\Delta G = -9.6 \text{ kcal.mol}^{-1}$). This correlates well with the free energy of base pairing for the *E. coli* prototypical RBS ($\Delta G = -9.4 \text{ kcal.mol}^{-1}$). Although *E. coli* initiation of translation is most efficient if the Shine-Delgarno sequence is located between 4 and 14 bp upstream of the start codon, a number of genes in the Streptomycetes and Mycobacteria have transcription start points that map to the initiation codon of the gene (Strohl, 1992; Timm *et al.*, 1994b). This suggests that ribosome assembly can therefore also occur downstream of the start codon. No experimental work regarding RBS location has been reported in the mycobacteria,

however, and since the initiation codon of *rnhA* has not been confirmed by amino terminal sequencing of *M. smegmatis* RNase HI, the location of a possible RBS for *M. smegmatis rnhA* therefore remains entirely speculative.

M. smegmatis RNase HI belongs to the Class I family of RNases HI

A search of the DNA sequence and protein databases (Genbank+EMBL+DDBJ+PDB and TIGR) for RNase HI homologues yielded a number of putative RNases HI, such as the YpdQ of *B. subtilis* and EbsB of *E. faecalis* (annotated, however, as potential cell wall biosynthesis enzymes) and those of *M. tuberculosis* and *M. leprae*. RNase HI homologues are present in all of the complete genome sequences available, with the exception of *Mycoplasma genitalium* which apparently does not possess either RNase HI or HII (Fraser *et al.*, 1995). In addition, genetic and biochemical evidence suggests that *Streptococcus pneumoniae* only possesses RNase HII (Zhang *et al.*, 1997). Interestingly, *M. tuberculosis rnhA* appears to be fused to a carboxy-terminal domain that bears high homology to the phosphoglycerate mutase (*pgm*) family of proteins. This organisation is conserved in *M. leprae* and is reminiscent of eukaryotic RNases HI which are commonly fused to other domains.

The location of *M. smegmatis rnhA* with respect to neighbouring genes appears unique. Although the identity of upstream and downstream genes is obscure, it is clear that the divergent overlapping promoter organisation observed in *E. coli*, *S. typhimurium* and *H. pylori* is not conserved in *M. smegmatis rnhA*. Similarly, the location of *Streptomyces galbus* and *S. lincolnensis rnhA* genes adjacent to, but divergent from, the *melC* operon, differs from that of *M. smegmatis rnhA*, as does the genetic organisation conserved between *M. tuberculosis* and *M. leprae*. Similarly, no homology with surrounding genes was observed with *B. subtilis rnhA* or *E. faecalis rnhA* or *M. smegmatis rnhA*.

Alignment of these bacterial RNases HI together with the RNase HI domains of the *M. tuberculosis* and *M. leprae* proteins highlighted interesting differences between the enzymes. *E. coli*, *S. typhimurium*, *T. thermophilus*, *M. smegmatis* and *B. aphidicola* appear to fall into one group of RNases HI, here denoted as Class I RNases HI, whereas *E. faecalis*, *B. subtilis*, *M. tuberculosis* and *M. leprae* RNases HI make up a second

group, here denoted as Class II. The Class II homologues appear to possess an extended loop between αI and βD , but have a reduced loop region between βE and αV (Fig. 1 and Fig. 4). More importantly, however, Class II homologues appear to lack αIII completely, part of the region defined as the basic protrusion (αIII + a further 10 amino acids), and which has been shown to be important for binding of the nucleic acid substrate (Keck and Marqusee, 1996, 1997; Iwai *et al.*, 1996). The RNase H domain of HIV-RT completely lacks this basic protrusion, and expression of the RNase H domain separately from the polymerase domain abolishes RNase H activity (Hostomsky *et al.*, 1991). In contrast, the murine leukemia virus (MuLV)-RT RNase H domain possesses part of this region and can be expressed separately from the polymerase domain and retain RNase H activity (Zhan and Crouch, 1997; Tanese and Goff, 1988). It may therefore be significant that only two RNases H have been isolated from a *B. subtilis* library by complementation of an *rnhA*- mutant of *E. coli* where neither of these RNases H correspond to YpdQ (Mizrahi and Itaya, in press). It is possible, then, that neither the *E. faecalis* or *B. subtilis* *rnhA* homologues encode a functional RNase HI, and thus represent pseudogenes. In this respect, the fusion of the *M. tuberculosis* and *M. leprae* RNase HI domains to another domain may be critical for retention of activity.

Comparison of the *M. smegmatis* RNase HI with its bacterial homologues placed this enzyme in the Class I RNase HI family. Identity to *E. coli* RNase HI at the amino acid level was calculated at 50% (78/155), whereas identity to the *M. tuberculosis* RNase HI domain was calculated at 22% (35/159). It is certainly interesting that *M. smegmatis* RNase HI exhibits higher homology to the Gram negative RNases HI than to its mycobacterial homologues. Partial sequences available for the RNases HI of *Streptomyces lincolnensis* and *S. galbus* show higher homology to the *M. leprae* and *M. tuberculosis* RNases HI than does *M. smegmatis* RNase HI, but nonetheless also fall into Class I.

Key structural and catalytic residues are conserved between M. smegmatis RNase HI and E. coli and T. thermophilus RNases HI

Due to the extensive amino acid homology between *M. smegmatis* RNase HI and *E. coli* and *T. thermophilus* RNases HI, a finer comparison of the conserved residues

between these enzymes was undertaken. The resolution of the crystal structure of *E. coli* RNase HI has allowed the identification of a number of intramolecular interactions between amino acids within this enzyme (Katayanagi *et al.*, 1992). Many of these amino acids, and presumably therefore the relevant intramolecular interactions of these amino acids, are conserved in *M. smegmatis* RNase HI. In particular, a number of structural interactions between the β sheets and the α helices have been conserved, such as the contacts between Val⁵(β A)-Leu⁵⁹(α I), Leu²⁶(β B)-Leu⁵⁹(α I), Val⁶⁵(β D)-Leu⁵⁹(α I), Ile⁷(β A)-Leu⁵⁶(α I), Ile⁷(β A)-Ala¹¹⁰(α IV), Leu⁶⁷(β I)-Ala⁵⁵(α I) and Trp¹¹⁸(β E)-Ala⁵⁵(α I) (Katayanagi *et al.*, 1992; positions correspond to the numbering of *E. coli* RNase HI). In addition, in *E. coli* RNase HI, Arg⁴⁶ is thought to coordinate with Asn¹⁰⁰, Asp¹⁰² and Asp¹⁴⁸ to form a centrally located ionic network which may contribute to general stabilisation of the molecule. All these residues are conserved in *T. thermophilus* and *M. smegmatis* RNase HI except for Asp¹⁴⁸ which has been replaced non-conservatively by Pro and Thr respectively. Other structural features, which have been assigned to Arg²⁷ and Arg¹³⁸ with respect to interaction of α V with β A, β C, and β D, appear mostly not to be conserved in either *T. thermophilus* or *M. smegmatis*. However, all six tryptophan residues found in *E. coli* RNase HI are conserved in *M. smegmatis* and *T. thermophilus*, with Trp⁹⁰ and Trp¹¹⁸ highly conserved among most bacterial homologues. All the tryptophan residues are clustered in the hydrophobic cleft of the enzyme and are oriented favourably for hydrophobic interactions. *M. smegmatis* and *T. thermophilus* RNase HI have an additional tryptophan residue, a conservative substitution of Tyr²² in *E. coli* RNase HI, which is also situated in the hydrophobic cleft. His¹¹⁴ has been shown to be important for stability (Kanaya *et al.*, 1991a) and is conserved in all Class I homologues.

Some residues that are highly conserved among RNase HI homologues, but have no apparent function in maintaining the structure of the enzyme, are involved in the binding and hydrolysis of the nucleic acid substrate. The motif NPG is invariant among all of the bacterial RNase HI homologues aligned, although it is completely absent from the RNase H domains of 11 reverse transcriptases (Kanaya *et al.*, 1990a). The Asn¹⁶ residue of this motif particularly is expected to play an important role in a steric interaction with the substrate, together with Cys¹³, Asn⁴⁴, Asn⁴⁵ and Gln⁷² (Katayanagi *et al.*, 1991). Of these, Asn¹⁶ and Asn⁴⁴ are invariant, and Cys¹³ is conserved amongst the

Class I homologues, Gln⁷², however, shows considerable variation in comparative residues amongst the bacterial RNase HI homologues. Cys¹³ and Asn⁴⁴ are expected to be involved in hydrogen bonds with the 2'-hydroxyls of the ribonucleotide chain and Asn⁴⁵ is thought to interact directly with the bases in the minor groove of the substrate (Kanaya *et al.*, 1991b). The role of Cys¹³, however, can be replaced by other residues as was shown by Kanaya *et al.* (1990b) who engineered a completely cysteine-free enzyme, substituted where appropriate by alanine, that exhibited wild-type activity. Cys⁶⁴ is not conserved in *M. smegmatis*, although it appears mostly conserved in other Class I RNase HI homologues and Cys¹³³ is only conserved between *E. coli* and *S. typhimurium*. *M. smegmatis* and *T. thermophilus* have one additional Cys residue corresponding to Lys¹¹¹ and Thr¹⁴⁹ in *E. coli* RNase HI, respectively. *E. faecalis* RNase HI does not possess any cysteine residues. Thr⁴³, instead, was proposed to have a more important role in binding of substrate (Kanaya *et al.*, 1991b), and certainly appears to be more highly conserved: Ser is the only conservative substitution observed for Thr⁴³ in *M. smegmatis*, *E. faecalis* and *B. subtilis* RNase HI. In *E. coli*, the basic protrusion contains two turns, TADK and VKNV, both of which expose lysine residues to solvent and hence are expected to be involved in binding of substrate. These are conserved in *M. smegmatis* as TAAK and VKNV.

Residues that have been identified as critical for catalytic activity in *E. coli* RNase HI, that is Asp¹⁰, Glu⁴⁸, Asp⁷⁰ and to a lesser extent His¹²⁴ and Asp¹³⁴, are conserved in the *M. smegmatis* homologue. The carbonyl oxygen of Gly¹¹ has been implicated in the co-ordination of metal ion binding and this amino acid is conserved amongst all homologues except for *E. faecalis* RNase HI in which alanine is substituted for this residue.

A number of residues (11) that are not conserved between *M. smegmatis* and *E. coli* or *S. typhimurium* RNase HI, are conserved between *M. smegmatis* and *T. thermophilus*. Of these, the substitution of Pro in the place of His⁶² has been shown to increase the stability of the *E. coli* enzyme, possibly by decreasing the entropy of the unfolded state of the enzyme (Akasaka *et al.*, 1995). Comparisons of the crystal structures of *E. coli* RNase HI and *T. thermophilus* RNase HI have shown that the enzymes adopt highly similar tertiary structures (Kanaya and Ikehara, 1993). On the

basis of the conservation of key structural amino acids, therefore, as well as those involved in binding and catalysis, it was concluded that *M. smegmatis* RNase HI was extremely likely to adopt an analogous tertiary structure to those of *E. coli* and *T. thermophilus* RNases HI, and to exhibit similar catalytic activity.

M. smegmatis RNase HI specifically hydrolyses the RNA strand of a RNA-DNA hybrid

Investigation of the biochemical attributes of the *M. smegmatis* enzyme was facilitated by the construction of an expression cassette which allowed overproduction and purification of a recombinant fusion protein. The recombinant protein (expressed as a soluble carboxyl terminal fusion to the maltose binding protein) was shown to possess RNase H activity in an activity gel analysis. As no activity was observed on a RNA-RNA substrate or ssRNA, it was concluded that the *M. smegmatis* enzyme was specific for the RNA strand of a RNA-DNA hybrid, as has been shown for the RNase HI of *E. coli*. RNase H activity has been classed as predominantly endonucleolytic in *E. coli* RNase HI and predominantly exonucleolytic in HIV-RT (Mizrahi, 1989), as demonstrated by PAGE analysis of the distribution of hydrolysis products generated by the activities of these enzymes on a heteroduplex substrate.

As *M. smegmatis* RNase HI was expected to exhibit a predominantly endonucleolytic activity, hydrolysis of such a substrate was used as a basis for the comparison of the activities of these two enzymes. As expected, PAGE analysis of the hydrolysis products generated by both enzymes showed very little difference in the distribution of the products. However, slight differences were noticeable in the predominance of certain size bands. It was difficult to analyse this result in terms of the substrate used since it was uniformly labeled and is of sufficient sequence complexity to make deductions of differences in site selectivities between the two enzymes spurious. However, Mizrahi (1989) was able to show the existence of strong site selectivity in the cleavage patterns of HIV-RT on an end-labeled substrate. Slight differences in site selectivity between *M. smegmatis* and *E. coli* RNases HI might therefore be discernable using a similar methodology. However, it was felt that such analysis was outside the scope of this project and of limited utility in investigating the role of RNase HI in *M.*

smegmatis. Further studies were directed therefore at investigating the regulation of *rnhA*.

Transcription of M. smegmatis rnhA is driven by its own promoter

A number of determinations of transcriptional starts for mycobacterial genes have been made in the past few years (Kong *et al.*, 1993; Cirillo *et al.*, 1994; Dellagostin *et al.*, 1995; Bashyam *et al.*, 1996; Gonzalez-y-Merchand *et al.*, 1996; Kenney and Churchward, 1996). The methods used have mostly relied on primer extension analysis. In this study it was felt that the presence of strong stem loop structures in the putative promoter region could terminate such primer extensions prematurely (Strohl, 1992; P. Stolt, pers. comm.; Cirillo *et al.*, 1994), and hence a more robust method which employed hybridisation of the mRNA to a large ssDNA probe was chosen (Stolt and Zili g, 1993).

In this study, determination of the transcriptional start point for *rnhA* rRNA pinpointed a T residue 56 bp upstream of the translational start. Transcription starts have been found from 0 bp (Timm *et al.*, 1994b), 29 bp (Kong *et al.*, 1993), 199 bp (Kenney and Churchward, 1996) to 273 bp (Cirillo *et al.*, 1994) but appear to average around 93 bp. In agreement with this is a comparison of *Streptomyces* promoters which indicated that in most genes, transcription is initiated within 100 bp of the translational start, although variation has been found within 0 and 345 bp (Strohl, 1992). A strong preference has been noted for a purine, most often G, for the initiation of transcription (Bashyam *et al.*, 1996; Bashyam and Tyagi, 1998; Jackson *et al.*, 1996; Kenney and Churchward, 1996), but a majority of transcripts in *M. paratuberculosis* promoters analysed (5/8) (Bannantine *et al.*, 1997) and a small number of transcripts in *M. smegmatis* (3/14) have also been found to initiate at a T residue (Bashyam *et al.*, 1996).

M. smegmatis rnhA -10 and -35 regions do not show conservation with E. coli consensus motifs

Bashyam *et al.* (1996) analysed 24 promoter regions from *M. smegmatis* and *M. tuberculosis* genes. A highly conserved -10 region corresponding to TANNNT was identified in 13/14 of the *M. smegmatis* promoter regions and in 7/10 of the *M. tuberculosis* promoter regions. This consensus sequence has been identified in a further 8

promoter regions of genes from *M. leprae* (Kenney and Churchward, 1996), *M. tuberculosis* (Jackson *et al.*, 1996; Kenney and Churchward, 1996; Curcio *et al.*, 1994), *M. bovis* (Kenney and Churchward, 1996) and *M. smegmatis* (Kenney and Churchward, 1996, Verma *et al.*, 1994) and is similar to that of *E. coli* (TATAAT) where the first, second and sixth bases of the hexameric sequence appear most important for function (Rosenberg and Court, 1979). An analogous situation appears to exist in the mycobacteria and, indeed, regions responsible for the binding of sigma factors to the -10 region are identical in MysA and HrdB, the principal sigma factors of *M. smegmatis* and *S. aureofasciens* respectively, and nearly identical (20/23 amino acids conserved) to RpoD, the principal sigma factor of *E. coli* (Bashyam *et al.*, 1996).

Despite this conservation, no obvious -10 site in the promoter region of *rnhA* could be identified. However, a hexamer consisting of TCGGGG was considered a candidate -10 region since the combination of TC for the first two bases has been found in a -10 region, albeit generally uncommon, and deviation from a T residue in the sixth position is most often a G (3/4 deviations) (Bashyam *et al.*, 1996). An alternate -10 consensus sequence consisting of CGGCC(C/G) has been derived from a comparison of *M. paratuberculosis* promoters (Bannantine *et al.*, 1997) suggesting that a single consensus for this region may be difficult to define.

Similarly, a range of -35 regions appears to exist, from those that correspond well to the *E. coli* consensus, to those that bear no resemblance to it. The general lack of consensus with the *E. coli* -35 motif has been ascribed to the fact that the recognition region of the sigma factor for the -35 hexamer differs substantially in MysA as opposed to RpoD, although MysA and HrdB retain high similarity. It might be expected then, that the -35 regions of mycobacterial promoters would conform more closely to those of the streptomycete -35 regions. However, Streptomycete -35 regions have been noted to vary widely, although a consensus of TTGAC in one class of promoters has been identified which conforms very closely to that of *E. coli* TTGACA (Bourn and Babb, 1995).

In this study, a hexamer consisting of TGGTGA is positioned with a 16 bp spacer region from the putative -10 hexamer. This hexamer has 50% identity with the *E. coli* consensus and the spacing between the -10 and -35 hexamers is consistent with the 17±1

bp variation noted in *E. coli* promoters, but is slightly shorter than the preferred 18-19 bp spacer observed with Streptomyces promoters (Bourn and Babb, 1995).

Although strong conservation of a -10 region is apparent in many mycobacterial promoters, a small number of promoter regions have been identified that show no homology to known sequences (Hsieh *et al.*, 1996, Kong *et al.*, 1993; Honoré *et al.*, 1993). A similar finding that 66/159 promoters in Streptomyces show no discernable homology to known consensus sequences, or indeed to each other, has been ascribed to coordination of gene function by a large number of σ factors (Bourn and Babb, 1995; Strohl, 1992) which may recognise differing subsets of promoters. An analogous situation may exist in the mycobacteria (Bannantine *et al.*, 1997). In support of this, the presence of thirteen sigma factors has been deduced from analysis of the genome sequence of *M. tuberculosis* (Cole, 1998).

In this study, *in vitro* transcription from a plasmid containing the promoter region of *rnhA* was obtained with purified *M. smegmatis* RNA polymerase, but the presence of only one type of sigma factor in this preparation could not be verified (Levin and Hatfull, 1993). Subsequently Predich *et al.* (1995) showed the presence of two sigma factors, MysA and MysB, in *M. smegmatis* RNA polymerase holoenzyme purified in a similar manner. Since both of these sigma factors show highest homology to the σ^{70} family of transcription factors, it appears likely that the *M. smegmatis rnhA* promoter may be recognised by a vegetative sigma factor.

The M. smegmatis rnhA promoter is not recognised in E. coli

The lack of *E. coli*-like transcriptional signals suggested that the *rnhA* promoter would not be recognised by the *E. coli* transcriptional machinery and, as expected, an in-frame fusion of the *rnhA* promoter plus the first two codons of *M. smegmatis rnhA* to *E. coli lacZ* on a multicopy plasmid (pRNH-*lacZ*) did not produce discernable galactosidase activity in *E. coli* (data not shown). In addition, complementation of the temperature sensitive phenotype of MIC3037 could not be obtained with *rnhA* driven off its own promoter (data not shown). A number of mycobacterial promoters are not recognised in *E. coli* such as the Ag85, *recA* and *ahpC* promoters of *M. tuberculosis* (Kremer *et al.*,

1995b; Davis *et al.*, 1991; Wilson and Collins, 1996), and the *sod* promoter of *M. leprae* (Thangaraj *et al.*, 1990).

Analysis of the relative strength of the M. smegmatis rnhA promoter

RNA synthesis in the mycobacteria has been found to be greatly reduced in comparison to *E. coli* and it has been hypothesised that the rate of initiation of transcription in the mycobacteria is probably very low (Harshey and Ramakrishnan 1977). A comparative study of the -50 regions of 14 *M. smegmatis* promoters and 10 *M. tuberculosis* promoters indicated a marked preference for a lower G+C content in *M. smegmatis* (43%) as compared to 57% in *M. tuberculosis* promoter regions (Bashyam *et al.*, 1996). This is especially significant since the average G+C content of *M. smegmatis* is higher than that of *M. tuberculosis*.

These authors and others (Das Gupta *et al.*, 1993) speculated that a high G+C promoter content could result in lower levels of gene expression which might explain the apparent absence of strong promoters in *M. tuberculosis* as compared to those found in *M. smegmatis*. Accordingly, retention of a similar preference was investigated in *M. smegmatis rnhA*. In contrast to these preceding studies, however, analysis of the G+C content of the -50 region of *M. smegmatis rnhA* indicated a G+C content of 75%. Inference of the strength of this promoter, therefore, based on the speculation that a high G+C content is concomitant with low expression levels, suggested that *M. smegmatis rnhA* is expressed at a low level indeed.

This was investigated therefore by fusing a reporter gene, *xylE*, to the promoter region of *rnhA* and assaying CDO activity. The results of such an assay show that CDO activity is barely detectable in cells in the lag phase, but expression becomes detectable with greater numbers of cells and remains proportional to total cell protein in exponentially growing cells. Although levels of *rnhA* expression were not investigated in stationary phase cells, this appears consistent with the findings of others that levels of RNase HII appear constant in the cell (Bialy and Kogoma, 1986). Comparison of another mycobacterial promoter, *P_{AN}* (Murray *et al.*, 1992), with the *rnhA* promoter, showed that relative levels of CDO activity in *M. smegmatis* cells were 10-fold higher with the *rnhA* promoter fusion than with the *P_{AN}* construct (data not shown). Interestingly, the G+C

content of the latter promoter region was calculated to be 54 %. The difference in levels of expression between the P_{AN} and *rnhA* promoters, therefore, does not appear to be due to differences in the G+C content of the promoter regions. This discrepancy in levels of expression could be due to less efficient recognition of this heterologous promoter region by *M. smegmatis* RNA polymerase. However, the P_{AN} promoter region contains well defined *E. coli*-like -10 and -35 consensus regions and therefore should be efficiently transcribed in *M. smegmatis*, and indeed, has been reported to drive high levels of expression of a *lacZ* reporter construct in *M. smegmatis* (Murray *et al.*, 1992).

Although not determined in this study, the *rpsL* promoter of *M. smegmatis* (which drives expression of the ribosomal protein S12) has been characterised using a *xylE* fusion as a reporter construct (Kenney and Churchward, 1996) and was found to drive > 1500 mU/mg of CDO activity. This, in turn, is 10-fold higher than the activity supported by the *rnhA* promoter. Other promoter strengths determined by *xylE* reporter construct activities include the *hsp60* promoter from *M. bovis* BCG and the *mtaA* promoter from *M. tuberculosis* H37Rv (Curcio *et al.*, 1994). These promoters were found to drive levels of expression in *M. smegmatis* of approximately 433 and 1430 mU/mg of CDO activity, respectively. Interestingly, the *hsp60* promoter was observed to be approximately twice as active in the heterologous host *M. smegmatis* than in *M. bovis* BCG (Curcio *et al.*, 1994). Based on this evidence therefore, the *rnhA* promoter appears to be of medium strength.

The promoter region of M. smegmatis rnhA extends beyond the tandem direct repeats

The transcription start point of the *rnhA* gene is located within a region comprising a number of direct repeat motifs. There is a cluster of three 25bp direct repeats, each separated by one base pair, flanked in turn by two 9 bp direct repeats, one of which overlaps the translational start. The three 25 bp direct repeats have, in turn, inverted repeats within these motifs, which have the potential to form two mutually exclusive stem loop structures.

A further hypothesis, therefore, provoked by the presence of these potentially strong stem-loop structures in the promoter region and by the relatively high level of expression of *rnhA*, was formulated. It was theorised that formation of a stem loop

structure in the promoter region could perhaps facilitate access of RNA polymerase to the promoter region and allow transcription of *rnhA* at a higher level than would otherwise be anticipated by consideration solely of the G+C content of this region. A similar suggestion has been made for the presence of two inverted repeat structures found in the -10 and -35 region of the *purC* promoter in *M. tuberculosis* (Jackson *et al.*, 1996).

Construction of deletion mutants of the promoter region was therefore undertaken to investigate this hypothesis. The first deletion mutant was designed to test whether the entire promoter region was contained in the direct repeat region, and thus was truncated immediately prior to the beginning of the three direct repeats. Analysis of the activity of this promoter construct showed a marked decrease in transcription levels, indicating that sequence upstream of these repeats was required for wild-type transcription levels. Interestingly, the deletion construct designed to test whether forced formation of the stem loop structure containing the transcriptional start influenced promoter activity showed enhanced levels of activity compared to the construct containing all three repeats.

Although the transcriptional start sites for these promoter constructs were not mapped, the increased level of transcription from the construct containing only two repeats was considered to be probably an artefact of cloning, since construction of the deletion mutant inserted vector DNA in the expected -35 region of the promoter. It has been shown that *MysA* has a relaxed requirement for -35 regions, and such a region fortuitously provided by vector sequence could have promoted the higher level of transcription observed. It seems unlikely that a new -10 region was created by the fusion where a functional -35 region in addition was also provided by the vector. Analysis of the vector sequence at the expected -35 position revealed a hexamer consisting of GGGGGA. Interestingly, a -35 region consisting of GGGGGG has been shown by mutational analysis to be an important determinant of transcriptional activity in a *Streptomyces* promoter (Strohl, 1992).

Transcription of M. smegmatis rnhA is not affected by induction of the SOS response

Although a number of genes have been noted to be regulated a) by direct repeats; in particular, the promoter regions of two chitinase genes from *Streptomyces plicatus* each contain a pair of 12 bp direct repeats that overlap the RNA polymerase binding site

and have been shown to interact with a sequence-specific DNA binding factor (Delic *et al.*, 1992) or b) by indirect repeats; three inverted repeats in the promoter region of *S. coelicolor dnaK* are responsible for the repression of the operon (Bucca *et al.*, 1997), the *rnhA* promoter was not expected to be highly inducible or repressed. Although Casaregola *et al.* (1987) observed a RecA-dependent decrease in expression of *rnhA* upon induction of the SOS response by UV irradiation, Quinones *et al.* (1989) noted little variation in transcriptional levels of *rnhA* upon treatment of *E. coli* with alkylating agents. Although cellular levels of *E. coli* RNase HI appear to be little affected by induction of the SOS system (Bialy and Kogoma, 1986), the promoter region of *E. coli rnhA* lacks similar repeat regions. Therefore, the effect of induction of the SOS response by mitomycin C on transcription levels of *M. smegmatis rnhA* was analysed (S. Durbach, unpublished). However, no change in transcription levels of *M. smegmatis rnhA* was observed.

Interestingly, marked changes in the transcription levels of *E. coli rnhA* were observed in a mutant fully induced for the SOS response (Quinones *et al.*, 1987) and it has been postulated that transcription of *rnhA* may be regulated by genes under the control of RecA, but independent of LexA (Altshuler, 1993; Casaregola *et al.*, 1987). An analogous experiment with *M. smegmatis rnhA* awaits the production of a constitutive *recA* mutant of *M. smegmatis* (S. Durbach, work in progress). The possible induction of *rnhA* by other agents was not investigated.

A number of small repetitive DNA sequences have been reported in the mycobacteria (discussed in Supply *et al.*, 1997). The tandem direct repeats observed upstream of *M. smegmatis rnhA* do not appear to fit into any of the previously described classes and no evidence for further copies of such a sequence could be obtained from a database search. These direct repeats may perhaps have resulted from polymerase slippage. In support of this is the observation that each direct repeat differs conservatively by 1 bp from the previous repeat.

E. coli RNase HI interacts with DnaJ and DnaK

The observation by Foster and Marinus (1992) that levels of RNase HI in *E. coli* appear to be regulated in some manner by the molecular chaperones DnaJ and DnaK

provided a further avenue for investigation of the regulation of RNase HI at the post-translational level. As this observation was made in *E. coli*, it appeared a prerequisite to verify a direct interaction between the *E. coli* proteins before testing the homologues in *M. smegmatis* for analogous activity.

E. coli DnaK and DnaJ have been found to bind native proteins in their separate capacities with differing affinities, as well as in concert as part of their chaperonin function. DnaJ associates stably with *E. coli* σ^{32} and plasmid-encoded RepA, and, with lower affinity, with DnaB helicase, phage λ P and λ O proteins (reviewed by Wawrzynów and Zylicz, 1995; Wickner, 1990; Georgopoulos *et al.*, 1990; Liberek *et al.*, 1990; Gamer *et al.*, 1992), whereas DnaK binds *E. coli* σ^{32} , GrpE, phage λ P protein and human p53 (Wawrzynów and Zylicz, 1995; Buchberger *et al.*, 1994; Liberek *et al.*, 1992; Gamer *et al.*, 1992).

In this study, comparison of a number of binding curves obtained from interactions of *E. coli* RNase HI with DnaJ and DnaK indicated specific interaction both between RNase HI and DnaJ, and RNase HI and DnaK. The latter interaction was substantially weaker than the former. Association of DnaK with RNase HI appeared to be enhanced by the addition of ATP (data not shown). This is in agreement with the observations by Schmid *et al.* (1994) who noted an increase in the rate of DnaK-substrate formation in the presence of ATP. However, these authors also noted a concomitant decrease in complex stability under the same conditions. Hydrolysis of ATP does not appear to be the cause of this rapid turnover of complex, but rather a conformational change engendered in DnaK by the binding of ATP. Surprisingly, in this study the RNase HI-DnaK complex appeared more stable in the presence of ATP than in its absence. Studies by Packschies *et al.* (1997) showed that a DnaK-ADP-protein complex is generally stable in the absence of GrpE which has been shown to facilitate release of the bound peptide by interchanging bound ADP for ATP. DnaK has a weak residual ATPase activity (1 molecule of ATP hydrolysed per molecule of DnaK in 10 minutes) that might have been stimulated in some fashion by the manner of immobilisation of DnaK to the dextran surface. Alternatively, it is possible that binding of RNase HI itself stimulated this ATPase activity. The weak ATPase activity of DnaK has been shown to

be stimulated 2-20 fold in the presence of a peptide substrate (Jordan and McMacken, 1995).

Comparison of the association curves of RNase HI with DnaJ as compared to RNase HI with DnaK in the presence of ATP indicated that RNase HI associates with DnaJ with higher affinity than with DnaK (data not shown). This is in agreement with the suggestion that DnaK has a broad degenerate substrate specificity, and that DnaJ provides specificity to the chaperonin machine (Gamer *et al.*, 1992). The ability of DnaJ to 'tag' RNase HI for enhanced recognition by DnaK was therefore investigated. A general model for the actions of the molecules of the DnaK chaperonin machine has been described and was used as the basis for the experiments in this study: DnaK interacts with substrate protein in the presence of ATP - this interaction is low affinity, fast binding and releasing. DnaJ binds to DnaK, triggering a conformational change in DnaK, which stimulates the ATPase activity of DnaK. Maximal activation of ATP hydrolysis by DnaK is dependent on the simultaneous binding of DnaK to protein substrate and to DnaJ (Karzai and McMacken, 1996). Hydrolysis of ATP allows the formation of a stable, high affinity DnaK-protein-ADP complex. DnaJ is not necessarily part of this complex, but if present in stoichiometric amounts will remain bound to the DnaK-protein-ADP complex (Liberek *et al.*, 1995). GrpE then exchanges ATP for the bound ADP (Fackschies *et al.*, 1997) and the DnaK-protein complex reverts to a low affinity interaction and the protein is released. Peirpaoli *et al.* (1997) have determined that the step in this cycle underlying the chaperone effect is the DnaJ triggered conformational change in DnaK which allows hydrolysis of ATP and the formation of a stable complex.

In agreement with this model, a synergistic effect on DnaK-RNase HI complex formation was observed in the presence of ATP and DnaJ. It is not known whether DnaJ was part of this complex, but since equimolar concentrations of RNase HI and DnaJ were used, it is likely that DnaJ remained bound to the DnaK-RNase HI complex (Liberek *et al.*, 1995). Although formation of a σ^{32} -DnaK complex occurred with only catalytic amounts of DnaJ, there are many indications that chaperonin interactions with σ^{32} do not conform to general rules (Georgopoulos *et al.*, 1994). Therefore, it would be interesting to see whether sub-stoichiometric amounts of DnaJ would facilitate binding of RNase HI to DnaK.

DnaJ did not appear to enhance binding between DnaK and RNase HI in the absence of ATP. Instead the association observed appeared to be simple addition of the separate binding data. This might be expected since a) the site of DnaJ and DnaK interaction is separate from that of DnaK-protein binding (Georgopoulos *et al.*, 1994) and b) from observations that hydrolysis of ATP is essential for the formation of a stable complex.

DnaK and DnaJ interact with a number of proteins in their denatured forms, but with a relatively low number of proteins in their native state (Wawrzynów and Zylicz, 1995). This study increases that number. The significance of this interaction, however, is unclear. Comparison of the binding observed between σ^{32} and DnaJ, and RNase HI and the chaperonins clearly indicates that, under the conditions used for binding, DnaJ and DnaK have a far lower affinity for RNase HI than for σ^{32} . It is therefore unlikely that DnaJ and DnaK sequester RNase HI in a similar manner to σ^{32} . However, DnaK and DnaJ have been shown to actively strip RNAP holoenzyme of σ^{32} , and it is unlikely that a similar affinity is necessary for other protein interactions.

Foster and Marinus (1992) postulate that perhaps the observed regulation of RNase HI by DnaK and DnaJ in their experiments occurs due to the association of the chaperonins with the nascent peptides produced from the ribosomes as has been shown for the N-terminal 17 amino acids of bovine rhodanese (Kudlicki *et al.*, 1996). This would assist production of a mature protein and hence enhance levels of RNase HI. DnaK is known to associate preferentially with positively charged hydrophobic peptides in an extended conformation (de Crouy-Chanel *et al.*, 1996). The first 20 amino acids of RNase HI might therefore sequester the chaperones in a noticeable fashion.

It is also possible that a weak interaction of chaperones with native proteins is ubiquitous, and merely a function of their role in the refolding of denatured proteins. In support of this, it has been shown that DnaK binds to alpha helical peptides - the most common secondary structure in folded proteins (de Crouy-Chanel *et al.*, 1997). RNase HI has a number of exposed alpha helical regions that could facilitate this binding.

The interaction between RNase HI and DnaK and DnaJ, however, does appear to be biologically significant: comparison of these low affinity interactions with others reported from BLAcore studies showed that the calculated association constant and

estimated dissociation constant for RNase HI and DnaJ are very similar in degree to those calculated for the interaction of major histocompatibility class II molecules with T-cell receptor molecules (Boniface and Davis, 1994).

Therefore weak interaction of RNase HI with DnaK and DnaJ may possibly be explained by the role that RNase HI plays in the replication of the chromosome. Chaperonins may be associated with the R-loop structures that can allow replication of the chromosome to occur at *oriK*'s and may therefore target RNase HI to some degree to these alternate origins. Oppositely, overproduction of DnaK and DnaJ upon induction of a stress response might be sufficient to sequester cellular RNase HI at levels such that RecA (which is also present in increased levels due to the SOS response) and other enzymes are allowed to promote replication from alternate origins.

Whether an analogous situation exists in the mycobacteria remains to be determined. However, the high conservation of heat shock proteins and the observed conservation of *M. smegmatis* RNase HI suggests that this interaction, and hence its physiological relevance, would also be conserved in the mycobacteria. The role of RNase HI *in vivo* in the mycobacteria was therefore investigated by experiments designed to create a knockout mutant of *rnhA*.

Insertion of aph within the coding region of rnhA exerts a lethal polar effect

This work attempted to construct a null mutant of *M. smegmatis* for *rnhA* in an analogous fashion to *E. coli rnhA-339::cat* (Kanaya and Crouch, 1984). Homologous recombination was readily achieved with this construct. However, no allele replacements could be obtained, even under the positive selection pressure exerted by the expression of *sacB*. A detailed analysis of the homologous recombinants showed an unequivocal bias towards one type of recombination event: that which conserved the linkage of the wild-type *rnhA* allele with uninterrupted downstream sequence. The disrupted *rnhA* allele therefore segregated with downstream sequence truncated by vector sequence. This bias, coupled with the failure to achieve allele replacement, suggested that the disruption of the *rnhA* gene exerted a strong polar effect on a gene or genes downstream of *rnhA* which resulted in a lethal phenotype upon allele replacement or recombination to form a Type II SCO. This hypothesis was further supported by the failure of *rnhA* supplied *in trans* to

complement this phenotype. The orientation of the *aph* cassette upon insertion would not be expected to contribute differentially to the phenotype of the knock out mutant as it contains its own transcriptional terminator. In addition, any read-through from this cassette in this instance would affect genes upstream of *rnhA* as it is cloned in an opposite orientation to the *rnhA* gene. Restoration of the wild-type *rnhA* gene by removal of the disruption to a site upstream of the promoter region of *rnhA* (a shift of 0.5 kb) allowed recovery of a number of double cross over events, in addition to both types of SCOs, which indicated that alleviation of the polar effect had occurred. It appeared that either a) transcriptional read-through from *rnhA* was critical for an essential gene downstream, where such a gene spanned the downstream cloning site and hence was truncated by vector sequence in the cloning process, or b) an essential gene was present on the complementary strand to *rnhA*, where the regulatory region or the N-terminus of this gene was truncated by vector sequence, and the C-terminus of this gene was somehow disrupted by insertion of *aph* into *rnhA*.

Sequence analysis of the region downstream of the *rnhA* gene failed to reveal either case. No co-directional open reading frame could be found downstream of *rnhA*. The presence of a strong stem loop structure approximately 100 bp downstream of *rnhA* argued in favour of transcriptional termination after transcription of the *rnhA* gene. This argument was supported by the presence of stop codons in all three reading frames immediately following this putative transcriptional terminator. However, sequence analysis did reveal the presence of a small open reading frame on the complementary strand, overlapping the reading frame of *rnhA* by 10 amino acids. The function of the 15 kDa protein (*hORF*) encoded by this reading frame is unknown. It shows a substantial degree of homology at the DNA level, and a high degree of homology at the amino acid level, to a hypothetical open reading frame in *Synechocystis* sp., and to the N-terminal of a hypothetical protein from *Pseudomonas stutzeri*. No other homologues of these proteins have been found in the database, and no known protein motifs could be identified. Therefore, no speculation as to function could be made.

Despite the discovery of this hypothetical ORF (*hORF*), the nature of the polar effect exerted by the *aph* cassette upon disruption of *rnhA* remains unresolved. The site of *aph* insertion is within the first third of *rnhA*, and *hORF* overlaps the last 30 bps of

rnhA only. It is difficult to see how this could have such a drastic effect on expression of *hORF*. In addition, the site of truncation of downstream sequence appeared to play a role in exertion of the polar effect, but sequence analysis of this region was not revealing. Neither DNA sequence nor amino acid sequence showed significant homology to any sequence deposited in the databases.

RNase HI is essential in M. smegmatis

Insertional disruptions of genes in *E. coli* have been shown to exert diverse effects (Link *et al.*, 1997). In particular, insertional inactivation of *hdeA* by *aph* in one site allocated an essential phenotype to the gene, whereas insertional inactivation in an alternate site of the same gene contradicted this finding. In addition, insertional inactivation of *yjbJ* in a similar fashion showed that the inserted allele actually bestowed improved growth on the host strain over both the wild type and deleted alleles of the gene. These authors suggest that in-frame deletion of a gene is perhaps the most reliable way to construct a null allele.

Since the exact nature of the polar effect was unknown, it was considered that a frameshift mutation in *rnhA* would be the least intrusive way to introduce a null allele, as a deletion of the *rnhA* gene might generate unforeseen complications (Link *et al.*, 1997). *M. tuberculosis* appears to lack a mismatch repair system (Mizrahi and Andersen, 1998), and experimental evidence has indicated that a 1 base pair alteration (S. Durbach, unpublished results), or a frameshift mutation (Pelicic *et al.*, 1996c) could be introduced stably into the genome of *M. smegmatis* by homologous recombination. Accordingly, a frameshift mutation in the beginning of *rnhA* was introduced into a homologous recombination substrate. Analysis of a small number of recombinants with this construct again indicated a bias towards one type of single cross over. In this instance, however, a bias towards one type of cross over could be explained by the fact that the other type of single cross over effectively generated an *rnhA*- mutant. This appeared to indicate that a *rnhA*- mutant would be selected against. Again, only allele replacements restoring wild type function could be recovered, even at lowered temperatures and on minimal medium. The results of these experiments strongly suggest that *rnhA* in *M. smegmatis* encodes an essential function under a variety of conditions.

An analogous situation has not been found for *rnhA* mutants of *E. coli* or for the *RNH1* null mutant of the diploid trypanosomatid *Crithidia fasciculata* (Ray and Hines, 1995). As yet, no other null mutants of *rnhA* have been reported, RNase HI appears to play an important role in the removal of R-loops (Drolet *et al.*, 1995; Hong *et al.*, 1995), and in repair mediated by either the RecBCD or RecF pathways (Itaya and Crouch, 1991a). In the absence of RNase HI, key enzymes involved in the initiation of cSDR become essential, such as DNA PolI, RecG and TopA (Kogoma and Maldonado, 1997; Drolet *et al.*, 1995; Hong *et al.*, 1995), the SOS response is chronically induced by the persistence of ssDNA, and the DNA repair functions of RecBCD and possibly also those of ExoI (*sbcB*), also become crucial (Itaya and Crouch, 1991a). The presence of some of these enzymes has already been investigated in the mycobacteria.

Topoisomerase I and DNA PolI from *M. tuberculosis* have been cloned and characterised (Yang *et al.*, 1996; Huberts and Mizrahi, 1995; Mizrahi and Huberts, 1996). Biochemical analysis of recombinant *M. tuberculosis* Topoisomerase I suggested that its activity is similar to that of *E. coli* Topoisomerase I although it appears to lack a zinc finger motif. *M. tuberculosis* DNA polymerase I appears to lack a 3'-5' proofreading activity, but possesses the 5'-3' nuclease activity which has been implicated as essential in *rnhA* mutants of *E. coli*. In addition, a null mutant of *polA* in *M. smegmatis* has shown DNA PolI to be functionally similar to *E. coli* DNA PolI (Gordhan *et al.*, 1996). *recB*, *recC*, and *recD* homologues are present in *M. smegmatis* (K. Derbyshire, pers. comm.). The *M. tuberculosis* homologues of these genes, and of *recG*, exhibit high homology to their *E. coli* counterparts, suggesting that the lethal phenotype of an *M. smegmatis rnhA* mutant probably does not result from either the absence or alteration of function of these enzymes in *M. smegmatis*.

In contrast, a search for a homologue of *sbcB* in *M. tuberculosis* was unsuccessful. This gene does not appear to be ubiquitous as it is also absent from *Helicobacter pylori*, *Aquifex aeolicus* and *Mycoplasma genitalium*. Exonuclease I, more recently annotated as *SbcB*, is encoded by *sbcB* in *E. coli*. This enzyme processively degrades ssDNA in a 3'-5' direction and has also been shown to have an associated deoxyribophosphodiesterase activity (Sandigursky and Franklin, 1992). It has been implicated in the DNA repair and recombination pathways mediated by RecBCD (Meisel

and Roth, 1996). Although *rnhA sbcB* mutants are viable, albeit severely compromised for growth, the *sbcB15* allele used to test viability of the double mutant has been shown to retain at least 26% activity (Phillips *et al.*, 1988). Hence it could be argued that complete absence of *sbcB* activity might be lethal in conjunction with *rnhA*. Certainly, prolonged incubation of selection plates for homologous recombinants of *M. smegmatis* did not reveal late colonies.

In the light of the absence of *sbcB* in *M. tuberculosis* it is possible therefore that this enzyme is also absent in *M. smegmatis* and that this is the reason for the inviability of an *rnhA* mutant of *M. smegmatis*. However, a very low level of RNase HI activity (<0.1%) is sufficient to complement the *recBC* temperature sensitive phenotype (Itaya and Crouch, 1991b), suggesting that *rnhA* activity in repair, although essential, is perhaps a minor role for the enzyme. In addition, *sbcB* appears not to be ubiquitous, suggesting that it does not have a major role in metabolism, but that possibly in *E. coli*, *S. typhimurium* and *B. subtilis* it may play an essential role in the RecBCD pathway of repair. It should be noted however, that SbcB exhibits a very strong association with RecA and SSB (Kowalczykowski *et al.*, 1994; Molineux and Gefner, 1975). *xonA* alleles of the *sbcB* gene also show a phenotype of illegitimate recombination (Allgood and Silhavy, 1991) that does not require functional RecA. The *xonA* allele of *sbcB* is defective for exonuclease I activity, but either retains or has acquired a second undefined activity (Phillips *et al.*, 1988). These authors suggested that the *xonA* gene product interferes directly with proteins involved in the RecF pathway. Therefore, an as yet undefined role may exist for SbcB. Although *Mycoplasma genitalium* has been noted not to have RNase HI or RNase HII homologues, or a *sbcB* homologue (Fraser *et al.*, 1995), the presence of an enzyme that may have RNase H activity has been identified (Koonin *et al.*, 1996). In addition, the lack of many systems in this organism argues that the arguably lethal phenotype generated by this combination in *E. coli* would not necessarily be observed in the mycoplasmas.

In contrast to the low levels of RNase HI required to complement *E. coli rnhA recBC*, 10% activity was insufficient to restore replication from *oriC* in *E. coli rnhA*, suggesting that the major role for RNase HI activity lies in this area. cSDR has only been demonstrated in *E. coli* as an artificial system induced by genetic defects such as *rnhA*,

recG and *sdrT* (reviewed by Kogoma, 1997). This alternative mode of DNA replication has not yet been demonstrated in any other bacterium. However, recent evidence suggests that *E. coli* cells entering stationary phase after a nutritional shift-up display a cSDP-type mode of DNA replication (Hong *et al.*, 1996). It has been suggested that two different modes of replication exist as a means to combat starved physiological states (eg in the soil) and states of maximum growth (ie in the host) (Kogoma, 1997). This author, together with Wintersberger (1990) postulate that originally, initiation from *oriK* and *oriC* were two modes of replication available, and that the acquisition of RNase H activity optimised the balance between the two systems. It is therefore possible that a similar situation exists in other bacteria. Enzymes known to stabilise or promote R-loops (RecA, RNA polymerase, gyrase), those known to resolve or counteract this formation (RNase HI, TopA and RecG) together with those involved in the bi-directional replication of the chromosome from these R-loops, (PriA, DNA PolI, DNA PolIII, DnaB and DnaG) appear conserved in the mycobacteria. If an analogous situation exists in the mycobacteria, it is conceivable that a different balance might exist between these enzymes for the resolution of R-loops. Loss of RNase HI activity might therefore be sufficient to cause loss of viability. The situation where an enzyme is inessential in *E. coli* but essential in another bacterium is not unknown: In *B. subtilis* for example, the Hbsu protein which is involved in DNA repair and recombination, is essential for viability, unlike its counterpart Hu in *E. coli* (Micka and Marahiel, 1992; Fernández *et al.*, 1997).

CONCLUSIONS

1. *M. smegmatis* possesses an *rnhA* homologue that encodes a protein highly homologous to its Gram negative counterparts.
2. This RNase HI appears similar at the biochemical level to *E. coli* RNase HI in that hydrolysis is predominantly endonucleolytic, and is specific for the RNA strand of an RNA-DNA hybrid.
3. The chromosomal context of this enzyme appears not to be conserved with other *rnhA* genes; in particular, it does not share an overlapping divergent promoter region with *dnaQ* as is found in *H. pylori*, *S. typhimurium* and *E. coli*.
4. Transcription of *rnhA* appears constitutive and is not affected by induction of the SOS response.
5. *E. coli* RNase HI interacts specifically, albeit weakly, with the molecular chaperonins, DnaJ and DnaK.
6. RNase HI activity is probably essential in *M. smegmatis*.

Appendix A

Reagents and Suppliers

Restriction endonucleases and restriction enzyme buffers, agarase, Klenow enzyme, RNase A, calf intestinal phosphatase, Taq DNA polymerase, S1 nuclease, Random Primed DNA Labeling Kit, bovine serum albumin (BSA), fish sperm DNA, chloramphenicol, hygromycin B, ampicillin, dithiothreitol (DTT), cesium chloride, isopropyl- α -D-galactopyranoside (IPTG), 5-bromo-4-chloro-3-indolyl- α -D-thiogalactopyranoside (X-gal), DNA molecular weight markers III and VI, and deoxyribonucleotide triphosphates (dNTP's) were from Boehringer Mannheim, Mannheim, Germany. T4 DNA ligase, T4 polynucleotide kinase, rRNasin and *E. coli* RNase HI were from Promega. Ammonium acetate, boric acid, n-butanol, calcium chloride, Coomassie brilliant blue, dimethylsulphoxide (DMSO), isoamyl alcohol, dichlorodimethylsilane, glucose, glycerol, guanidium chloride, guanidinium thiocyanate, 2-[4-(2-hydroxyethyl)-1-piperazinyl]ethanesulphonic acid (HEPES), magnesium acetate, magnesium chloride, phenol, piperazine-1,4-bis(2-ethanesulphonic acid) (PIPES), polyethyleneglycol (PEG) 6000, potassium acetate, potassium chloride, potassium dihydrogen phosphate, saccharose, sodium acetate, sodium dihydrogen phosphate, and 1,1,1-trichloroethane were from Merck, Darmstadt, Germany. Amberlite resin beads, catechol, ethylenediaminetetraacetic acid (EDTA), glass beads, mineral oil, polyoxyethylenesorbitan monooleate (Tween 80), Q-Sepharose, Sephadex G-25, thiamine, Tris-borate-EDTA (TBE) buffer, Trizma base and urea were obtained from Sigma, St Louis, MO, USA. Dimethyl formamide, ethanol, formamide, β -mercaptoethanol, sarcosyl, sodium citrate, and tryptophan were from BDH, Poole, UK. Bromophenol blue was obtained from Gurr Ltd., London, UK. Chloroform was obtained from Holpro Lavasz, Midrand, South Africa. (dT)₁₂₋₁₈ was from Pharmacia, Uppsala, Sweden. 3M XDA Trimax X-ray film was obtained from X-ray Imaging Services, Rustenburg, South Africa. Filter paper and 30 000 NMWL Ultrafree MC filter units were obtained from Whatman International, Maidstone, UK. Hybond-N nylon filters, kanamycin sulphate and Sequenase Version 2.0 DNA Sequencing Kit

were obtained from Amersham. Glasswool was from Fluka, Buchs, Switzerland. Gel drying paper was from Hoefer Scientific Instruments, San Francisco, CA, USA. Methanol was from SMM Chemicals, Alberton, South Africa. Scintillation cocktail was obtained from Packard, Downer's Grove, IL, USA. Sodium dodecyl sulphate (SDS) was from Behring Diagnostics, La Jolla, CA, USA. Sodium chloride was obtained from Associated Chemical Enterprises, Johannesburg, South Africa. Agar, bacto-agar, Middlebrook 7H9 broth, Middlebrook 7H10 agar and maltose were from Difco, Detroit, MI, USA. Yeast extract and tryptone were obtained from Oxoid Ltd, Basingstoke, UK. N,N,N',N',-tetramethylethylenediamine (TEMED), ammonium persulphate, acrylamide, N,N'-methylene-bis-acrylamide (Bis), SDS-PAGE molecular weight standards, xylene cyanol, Protein Assay Kit and Muta-Gene Plagemid *in vitro* Mutagenesis Kit were from Bio-Rad, Hercules, CA, USA. Sodium hydroxide, isopropanol and hydrochloric acid were from Labchem, Boksburg, South Africa. SeaKem®LE agarose and SeaPlaque®GTG® low melting agarose were obtained from FMC Bioproducts, Rockland, ME, USA. [α - 35 S] deoxyadenosine triphosphate (1 000 Ci/mmol), [α - 32 P] deoxycytosine triphosphate (3 000 Ci/mmol) and [γ - 32 P] adenosine triphosphate (6 000 Ci/mmol) were from New England Nuclear, Boston, MA, USA, and from Amersham International. Microcentrifuge tubes were from Eppendorf, Hamburg, Germany. The Fast-Link DNA ligation Kit was from Epicentre Technologies, Madison, WI. Amylose resin was from New England Biolabs Inc., Beverly, MA. Sterivex-GS 0.22 μ m filters were from Millipore S.A., Molsheim, France, and UltraFree MC filters were from Millipore Corporation, Bedford, MA.

APPENDIX B

Media and General Solutions

Final volume (made up with deionised water) is one litre, unless otherwise stated. Media were sterilised by autoclaving at 121°C for 20 minutes.

Luria Bertani Agar plates (LA)

10 g tryptone powder; 5 g yeast extract; 10 g sodium chloride; 15 g agar powder.

Luria Bertani broth (LB)

10 g tryptone powder; 5 g yeast extract; 10 g sodium chloride.

Middlebrook-Tween broth (M-TW)

4.7 g Middlebrook 7H9 broth base; (2 ml glycerol; 0.85 g sodium chloride)-autoclaved separately; 2.5 ml 20% Tween 80 – filter sterilised and added after autoclaving.

Middlebrook Agar plates (MA)

19 g Middlebrook 7H10 agar powder.

Minimal Medium (MM)

1 x M9 salts; 1 mM magnesium sulphate; 0.1 mM calcium chloride; 1 mM thiamine.HCL; 0.2 % glucose.

2TY broth

16 g tryptone powder; 10 g yeast extract; 5 g sodium chloride.

Antibiotics (filter sterilised)

ampicillin: 35 mg/ml in 50% ethanol. Stored at -20°C.

kanamycin: 50 mg/ml in water. Stored at 4°C.

chloramphenicol: 35 mg/ml in 50% ethanol. Stored at -20°C.

hygromycin B: supplied as a 50 mg/ml solution in phosphate buffered saline. Stored at 4°C.

Denhardt's Reagent (50x stock)

10 g Ficoll (Type 400); 10 g polyvinylpyrrolidone; 10 g bovine serum albumin (Fraction V). Filter sterilised. Stored at -20°C.

20 x SSC

175.3 g sodium chloride; 88.2 g sodium citrate; pH adjusted to 7.0 with NaOH.

TE

10 mM Tris pH8.0; 10 mM EDTA

10 x M9 salts

128 g Na₂HPO₄·7H₂O; 30 g KH₂PO₄; 5 g NaCl; 10 g NH₄Cl.

40% acrylamide stock for PAGE (100 ml)

38 g acrylamide; 2 g bis-acrylamide; (deionised with 5 g Amberlite resin and filtered through Whatmann 3MM filter paper). Stored at 4°C.

30% acrylamide stock for SDS-PAGE (100 ml)

30 g acrylamide; 0.8 g bis-acrylamide; (deionised with 5 g Amberlite resin and filtered through Whatmann 3MM filter paper). Stored at 4°C.

SDS-PAGE stacking gel

5% acrylamide; 0.176 M Tris pH6.8; 0.1% SDS; 0.035 % TEMED; 25 mM ammonium persulphate.

SDS-PAGE resolving gel

10% acrylamide; 0.38 M Tris pH6.8; 0.1 % SDS; 0.035 % TEMED; 25 mM ammonium persulphate.

APPENDIX C

Abbreviations

AIDS	acquired immune deficiency syndrome
Amp	ampicillin
<i>aph</i>	gene encoding Km ^R
ATP	adenosine triphosphate
BCG	bacille Calmette-Guérin
<i>bla</i>	gene encoding Amp ^R
bp	base pair
BSA	bovine serum albumin
CDO	catechol 2,3-dioxygenase
CTP	cytosine triphosphate
Cm	chloramphenicol
d	deoxy
DCO	double cross over
DEPC	diethyl pyrocarbonate
DNA	deoxyribonucleic acid
dNTPs	deoxyribonucleotide triphosphates
ds	double stranded
DTT	dithiothreitol
EDC	<i>N</i> -ethyl- <i>N'</i> -((3-di-methylamino)propyl)-carbodiimide hydrochloride
EDTA	ethylenediaminetetraacetic acid
GTP	guanine triphosphate
HEPES	2-[4-(2-hydroxyethyl)-1-piperazinyl]ethanesulphonic acid
HIV	human immunodeficiency virus
<i>hyg</i>	gene encoding Hyg ^R
Hyg	hygromycin B
IPTG	isopropyl- α -D-galactopyranoside
kb	kilobase(s)
Km	kanamycin
LA	Luria-Bertani agar
<i>lacZ</i>	gene encoding β -galactosidase
LB	Luria-Bertani broth
LMA	low melting agarose
MA	Middlebrook agar
<i>malE</i>	gene encoding MBP
MBP	maltose binding protein
MDR	multidrug resistant
MM	minimal medium
M-TW	Middlebrook-Tween medium
Na-TCA	sodium trichloroacetic acid
NEB	New England Biolabs
NHS	<i>N</i> -hydroxysuccinimide
OD	optical density

ORF	open reading frame
PAG	polyacrylamide gel
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEG	polyethylene glycol
PIPES	piperazine-1,4-bis(2-ethanesulphonic acid)
PNK	polynucleotide kinase
R	resistant/resistance
RBS	ribosome binding site
RF	replicative form
RNA	ribonucleic acid
RNase	ribonuclease
<i>rnhA</i>	gene encoding RNase HI
RT	reverse transcriptase
<i>sacB</i>	gene encoding levansucrase
SCO	single cross over
SDS	sodium dodecyl laurel sulphate
ss	single stranded
Suc	sucrose
TB	tuberculosis
TEMED	N,N,N',N'-tetramethylethylenediamine
TTP	thymine triphosphate
Tween	polyoxyethylene sorbitan monooleate
USB	United States Biochemical
UTP	uridine triphosphate
X-gal	5-bromo-4-chloro-3-indolyl- α -D-thiogalactopyranoside
<i>xylE</i>	gene encoding catechol-2,3-dioxygenase
WHO	World Health Organisation
WT	wild type

APPENDIX D

Predicted properties of *M. smegmatis* RNase HI as determined by the EDITSEQ module of LASERGENE.

Translate DNA Sequence L11918.SEQ(404,883)
With Standard Genetic Code

Molecular Weight 17495.81 Daltons
159 Amino Acids
16 Strongly Basic(+) Amino Acids (K,R)
16 Strongly Acidic(-) Amino Acids (D,E)
56 Hydrophobic Amino Acids (A,I,L,F,W,Y)
37 Polar Amino Acids (N,C,Q,S,T,V)

7.456 Isoelectric Point
0.869 Charge at PH 7.0

Total number of bases translated is 480

% A = 17.92 [86]
% G = 35.00 [168]
% T = 13.12 [63]
% C = 33.96 [163]
% Ambiguous = 0.00 [0]

% A+T = 31.04 [149]
% C+G = 68.96 [331]

Davis, Botstein, Roth Melting Temp C. 92.13
Wallace Temp C 1948.00

Codon usage table of *M. smegmatis* *rnhA*:

gca	Ala (A)	2	#	cag	Gln (Q)	6	#	uug	Leu (L)	1	#	uaa	Ter (.)	0
gcc	Ala (A)	6	#	---	Gln (Q)	6	#	---	Leu (L)	13	#	uag	Ter (.)	0
gcg	Ala (A)	7	#	gaa	Glu (E)	2	#	aaa	Lys (K)	0	#	uga	Ter (.)	1
gcu	Ala (A)	0	#	gag	Glu (E)	8	#	aag	Lys (K)	4	#	---	Ter (.)	1
---	Ala (A)	15	#	---	Glu (E)	10	#	---	Lys (K)	4	#	aca	Thr (T)	0
aga	Arg (R)	0	#	gga	Gly (G)	1	#	aug	Met (M)	4	#	acc	Thr (T)	10
agg	Arg (R)	0	#	ggc	Gly (G)	13	#	---	Met (M)	4	#	acg	Thr (T)	2
cga	Arg (R)	0	#	ggg	Gly (G)	1	#	uuc	Phe (F)	2	#	acu	Thr (T)	1
cgc	Arg (R)	9	#	ggu	Gly (G)	2	#	uuu	Phe (F)	0	#	---	Thr (T)	13
cgg	Arg (R)	2	#	---	Gly (G)	17	#	---	Phe (F)	2	#	ugg	Trp (W)	7
cgu	Arg (R)	1	#	cac	His (H)	6	#	cca	Pro (P)	0	#	---	Trp (W)	7
---	Arg (R)	12	#	cau	His (H)	0	#	ccc	Pro (P)	5	#	uac	Tyr (Y)	2
aac	Asn (N)	6	#	---	His (H)	6	#	ccg	Pro (P)	2	#	uau	Tyr (Y)	0
aau	Asn (N)	1	#	aua	Ile (I)	0	#	ccu	Pro (P)	0	#	---	Tyr (Y)	2
---	Asn (N)	7	#	auc	Ile (I)	5	#	---	Pro (P)	7	#	gua	Val (V)	0
gac	Asp (D)	5	#	auu	Ile (I)	0	#	agc	Ser (S)	3	#	guc	Val (V)	5
gau	Asp (D)	1	#	---	Ile (I)	5	#	agu	Ser (S)	2	#	gug	Val (V)	9
---	Asp (D)	6	#	cua	Leu (L)	1	#	uca	Ser (S)	0	#	guu	Val (V)	0
ugc	Cys (C)	2	#	cuc	Leu (L)	4	#	ucc	Ser (S)	0	#	---	Val (V)	14
ugu	Cys (C)	0	#	cug	Leu (L)	7	#	ucg	Ser (S)	2	#	nnn	???(X)	0
---	Cys (C)	2	#	cuu	Leu (L)	0	#	ucu	Ser (S)	0	#	TOTAL		160
caa	Gln (Q)	0	#	uaa	Leu (L)	0	#	---	Ser (S)	7	#			

5. REFERENCES

- Akasako A, Haruki M, Oobatake M, Kanaya S (1995) High resistance of *Escherichia coli* ribonuclease HI variant with quintuple thermostabilizing mutations to thermal denaturation, acid denaturation, and proteolytic degradation. *Biochemistry* 34:8115-8122.
- Alangaden GJ, Manavathu EK, Vakulenko SB, Zvonok NM, Lerner SA (1995) Characterisation of fluoroquinolone-resistant mutant strains of *Mycobacterium tuberculosis* selected in the laboratory and isolated from patients. *Antimicrob. Agents Chemother.* 39:1700-1703.
- Albrecht H, Rusch-Gardes S, Stellbrink HJ, Greten H, Jackle S (1997) Disseminated *Mycobacterium genavense* infection as a cause of pseudo-Whipple's disease and sclerosing cholangitis. *Clin. Infect. Dis.* 25: 742-743.
- Alcendor DJ, Chapman GC, Beaman BL (1995) Isolation, sequencing and expression of the superoxide dismutase-encoding gene (*sod*) of *Nocardia asteroides* strain GUH-2. *Gene* 164:143-147.
- Aldovini A, Husson RN, Young RA (1993) The *uraA* locus and homologous recombination in *Mycobacterium bovis* BCG. *J. Bacteriol.* 175:7282-7289.
- Allgood ND, Silhavy TJ (1991) *Escherichia coli xonA* (*sbcB*) mutants enhance illegitimate recombination. *Genetics* 127:671-680.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389-3402.
- Altshuler M (1993) Recovery of DNA replication in UV-damaged *Escherichia coli*. *Mutat. Res.* 294:91-100.
- Andrew PW, Roberts IS (1993) Construction of a bioluminescent mycobacterium and its use for assay of antimycobacterial agents. *J. Clin. Microbiol.* 31:2251-2254.
- Ariyoshi M, Vassilyev DG, Iwasaki H, Nakamura H, Shinagawa H, Morikawa K (1994) Atomic structure of the RuvC resolvase: a holliday junction-specific endonuclease from *E. coli*. *Cell* 78:1063-1072.

- Arruda S, Bomfim G, Knights R, Huima-Byron T, Riley LW (1993) Cloning of an *M. tuberculosis* DNA fragment associated with entry and survival inside cells. *Science* 261:1454-1457.
- Asai T, Kogoma T (1994) D-loops and R-loops: alternative mechanisms for the initiation of chromosome replication in *Escherichia coli*. *J. Bacteriol.* 176:1807-1812.
- Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (1989) *Current Protocols in Molecular Biology*. New York: John Wiley and Sons.
- Azad AK, Sirakova TD, Rogers LM, Kolattukudy PE (1996) Targeted replacement of the mycocerosic acid synthase gene in *Mycobacterium bovis* BCG produces a mutant that lacks mycosides. *Proc. Natl. Acad. Sci. USA* 93:4787-4792.
- Azad AK, Sirakova TD, Fernandes ND, Kolattukudy PE (1997) Gene knockout reveals a novel gene cluster for the synthesis of a class of cell wall lipids unique to pathogenic mycobacteria. *J. Biol. Chem.* 272:16741-16745.
- Balasubramanian V, Pavelka MS, Bardarov SS, Martin J, Weisbrod TR, McAdam RA, Bloom BR, Jacobs WR Jr (1996) Allelic exchange in *Mycobacterium tuberculosis* with long linear recombination substrates. *J. Bacteriol.* 178:273-279.
- Bange FC, Brown AM, Jacobs WR Jr (1996) Leucine auxotrophy restricts growth of *Mycobacterium bovis* BCG in macrophages. *Infect. Immun.* 64: 1794-1799.
- Bannantine JP, Barletta RG, Thoen CO, Andrews RE Jr (1997) Identification of *Mycobacterium paratuberculosis* gene expression signals. *Microbiology* 143:921-928.
- Bardarov S, Kraikov J, Carriere C, Yu S, Vaamonde C, McAdam RA, Bloom BR, Hatfull GF, Jacobs WR Jr (1997) Conditionally replicating mycobacteriophages: A system for transposon delivery to *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. USA* 94:10961-10966.
- Barletta RG, Kim DD, Snapper SB, Bloom BR, Jacobs WR Jr (1992) Identification of expression signals of the mycobacteriophages Bxb1, L1, and TM4 using *Escherichia-Mycobacterium* shuttle plasmids pYUB75 and pYUB76 designed to create translational fusions to the *lacZ* gene. *J. Gen. Microbiol.* 138:23-30.

- Barletta RG, Snapper B, Cirillo JD, Connell ND, Kim DD, Jacobs WR, Bloom BR** (1990) Recombinant BCG as a candidate oral vaccine vector. *Res. Microbiol.* 141:931-939.
- Barry CE 3rd** (1997) New horizons in the treatment of tuberculosis. *Biochem. Pharmacol.* 54:1165-1172.
- Bashyam MD, Kaushal D, Dasgupta SK, Tyagi AK** (1996) A study of mycobacterial transcriptional apparatus: identification of novel features in promoter elements. *J. Bacteriol.* 178:4847-4853.
- Bashyam MD, Tyagi AK** (1998) Identification and analysis of "extended -10" promoters from mycobacteria. *J. Bacteriol.* 180:2568-2573.
- Baulard A, Jourdan C, Mercenier A, Locht C** (1992) Rapid mycobacterial plasmid analysis by electrofusion between *Mycobacterium* spp. and *Escherichia coli*. *Nucleic Acids Res.* 20:4105.
- Baulard A, Escuyer V, Haddad N, Kremer L, Locht C, Berche P** (1995) Mercury resistance as a selective marker for recombinant mycobacteria. *Microbiology* 141:1045-1050.
- Beall DS, King CH, Quinn FD** (1998) Identification of *Mycobacterium tuberculosis* genes expressed under anoxic conditions. TB:Molecular Mechanisms and Immunologic Aspects; Keystone, CO; April 2-8.
- Bebora LC** (1997) Role of plasmids in the virulence of enteric bacteria. *East Afr. Med. J.* 74:444-446.
- Beggs ML, Crawford JT, Eisenach KD** (1995) Isolation and sequencing of the replication region of the *Mycobacterium avium* plasmid pLR7. *J. Bacteriol.* 177:4836-4840.
- Bensing BA, Dunnay GM** (1993) Cloning and molecular analysis of genes affecting expression of binding substance, the recipient-encoded receptor(s) mediating mating aggregate formation in *Enterococcus faecalis*. *J. Bacteriol.* 175:7421-7429.
- Berkower I, Leis J, Hurwitz J** (1973) Isolation and characterization of an endonuclease from *Escherichia coli* specific for ribonucleic acid in ribonucleic acid-deoxyribonucleic acid hybrid structures. *J. Biol. Chem.* 248:5914-5921.
- Bialy H, Kogoma T** (1986) RNase H is not involved in induction of stable DNA replication in *Escherichia coli*. *J. Bacteriol.* 165:321-323.

- Black CB, Foster M, Cowan JA** (1996) Mechanism of metal-promoted catalysis of nucleic acid hydrolysis by *Escherichia coli* ribonuclease H. *JBC* 1:500-506.
- Bland J** (1996) Facts about leprosy. *World Health* 49(3):6-7.
- Bloom BR, Murray C.JL** (1992) Tuberculosis: Commentary on a reemergent killer. *Science* 257:1055-1064.
- Bockrath R, Wolff L, Farr A, Crouch RJ** (1987) Amplified RNase H activity in *Escherichia coli* B/r increases sensitivity to ultraviolet radiation. *Genetics* 115:33-40.
- Boniface JJ, Davis MM** (1994) The kinetics of binding of peptide/MHC complexes to T-cell receptors: application of surface plasmon resonance to a low affinity measurement. *Methods: A companion to Methods Enzymol.* 6:168-176.
- Bonomo RA, Briggs JM, Gross W, Hassan M, Graham RC, Butler WR, Salata RA** (1998) *Mycobacterium celatum* infection in a patient with AIDS. *Clin Infect Dis.* 26: 243-245.
- Boura WR, Babb B** (1995) Computer assisted identification and classification of streptomycete promoters. *Nucleic Acids Res.* 23:3696-3703.
- Brettle RF** (1997) *Mycobacterium avium intracellulare* infection in patients with HIV or AIDS. *J Antimicrob Chemother.* 40:156-160.
- Brown MC, Taffet SM** (1995) Lipoarabinomannans derived from different strains of *Mycobacterium tuberculosis* differentially stimulate the activation of NF-kappa B and KBF1 in murine macrophages. *Infect. Immun.* 63:1960-1968.
- Brudney K, Dobkin J** (1991) Resurgent Tuberculosis in New York City: Human Immunodeficiency Virus, Homelessness and the Decline of Tuberculosis Control Programs. *Am. Rev. Respir. Dis.* 144:745-749.
- Bucca G, Hindle Z, Smith CP** (1997) Regulation of the *dnaK* operon of *Streptomyces coelicolor* A3(2) is governed by HspR, an autoregulatory repressor protein. *J. Bacteriol.* 179:5999-6004.
- Buchberger A, Schröder H, Büttner M, Valencia A, Bukau B** (1994) A conserved loop in the ATPase domain of the DnaK chaperone is essential for stable binding of GrpE. *Nat. Struct. Biol.* 1:95-101.
- Butt AA** (1998) Cervical adenitis due to *Mycobacterium fortuitum* in patients with acquired immunodeficiency syndrome. *Am. J. Med. Sci.* 315:50-55.
- Cao Y, Kogoma T** (1993) Requirement for the polymerization and 5'-3' exonuclease activities of DNA polymerase I in initiation of DNA replication at *oriK* sites in

- the absence of RecA in *Escherichia coli rnhA* mutants. *J. Bacteriol.* 175:7254-7259.
- Carriere C, Riska PF, Zimhony O, Kriakov J, Bardarov S, Burns J, Chzn J, Jacobs WR Jr (1997) Conditionally replicating luciferase reporter phages: improved sensitivity for rapid detection and assessment of drug susceptibility of *Mycobacterium tuberculosis*. *J. Clin. Microbiol.* 35:3232-3239.
- Casaregola S, Khidhir M, Holland IB (1987) Effects of modulation of RNase H production on the recovery of DNA synthesis following UV-irradiation in *Escherichia coli*. *Mol. Gen. Genet.* 209:494-498.
- Chaicumpar K, Fujiwara N, Nishimura O, Hotta H, Pan JW, Takahashi M, Abe C, Yano I (1997) Studies of polymorphic DNA fingerprinting and lipid pattern of *Mycobacterium tuberculosis* patient isolates in Japan. *Microbiol. Immunol.* 41:107-119.
- Cirillo JD, Weisbrod TR, Pascopella L, Bloom BR, Jacobs WR Jr (1994) Isolation and characterisation of the aspartokinase and aspartate semialdehyde dehydrogenase operon from mycobacteria. *Mol. Microbiol.* 11:629-639.
- Clark-Curtiss J (1990) Genome structure of the mycobacteria. In: *Molecular Biology of the Mycobacteria*. (ed. McFadden J) London: Surrey University Press. pp. 77-96.
- Clemens DL, Lee BY, Horwitz MA (1995) Purification, characterization, and genetic analysis of *Mycobacterium tuberculosis* urease, a potentially critical determinant of host-pathogen interaction. *J. Bacteriol.* 177:5644-5652.
- Cocito C, Gilot P, Coene M, de Kesel M, Poupert P, Vannuffel P (1994) Paratuberculosis. *Clin. Microbiol. Rev.* 7:328-345.
- Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE III, Tekaiia F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Krogh A, McLean J, Moule S, Murphy L, Oliver K, Osborne J, Quail MA, Rajandream M-A, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulston JE, Taylor K, Whitehead S, BarrellBG (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393:537-544.

- Collins DM, Kawakami RP, de Lisle GW, Pascopella L, Bloom BR, Jacobs WR Jr** (1995) Mutation of the principal sigma factor causes loss of virulence in a strain of the *Mycobacterium tuberculosis* complex. *Proc. Natl. Acad. Sci. USA* 92:8036-8040.
- Collins ME, Patil A, Wall S, Nolan A, Goodger J, Woodward MJ, Dale JW** (1990) Cloning and characterization of the gene for the '19 kDa' antigen of *Mycobacterium bovis*. *J. Gen. Microbiol.* 136:1429-1436.
- Constant P, Poquet Y, Peyrat MA, Davodeau F, Bonneville M, Fournie JJ** (1995) The antituberculous *Mycobacterium bovis* BCG vaccine is an attenuated mycobacterial producer of phosphorylated nonpeptidic antigens for human gamma delta T cells. *Infect. Immun.* 63:4628-4633.
- Cooksey RC, Crawford JT, Jacobs WR Jr, Shinnick TM** (1993) A rapid method for screening antimicrobial agents for activities against a strain of *Mycobacterium tuberculosis* expressing firefly luciferase. *Antimicrob. Agents Chemother.* 37:1348-1352.
- Crouch RJ, Dirksen ML** (1982) Ribonucleases H. In: *Nucleases* (ed. Lin SM, Roberts RJ) Cold Spring Harbor, New York: Cold Spring harbor Laboratory Press. pp. 211-241.
- Cunningham AF, Spreadbury CL** (1998) Mycobacterial stationary phase induced by low oxygen tension: Cell wall thickening and localisation of the 16-kilodalton α -crystallin homolog. *J. Bacteriol.* 180:801-808.
- Curcio R, Dhandayuthapani S, Deretic V** (1994) Gene expression in the mycobacteria: transcriptional fusions based on *xyle* and analysis of the response regulator *mtrA* from *Mycobacterium tuberculosis*. *Mol. Microbiol.* 13:1057-1064.
- D'Alessio JM** (1982) RNA sequencing. In: *Gel electrophoresis of nucleic acids*. (ed. Rickwood D, Hames BD). Oxford: IRL Press. pp.182-183.
- Dale JW, Patil A** (1990) Mycobacterial gene expression and regulation. In: *Molecular Biology of the Mycobacteria*. (ed. McFadden J) London: Surrey University Press. pp. 173-198.
- Das Gupta SK, Bashyam MD, Tyagi AK** (1993) Cloning and assessment of mycobacterial promoters by using a plasmid shuttle vector. *J. Bacteriol.* 175:5186-5192.

- David M, Lubinsky-Mink S, Ben-Zvi A, Ulitzur S, Kuhn J, Suissa M (1992) A stable *Escherichia coli*-*Mycobacterium smegmatis* plasmid shuttle vector containing the mycobacteriophage D29 origin. *Plasmid* 28:267-271.
- Davies JF, Hostomska Z, Hostomsky Z, Jordan SR, Matthews D (1991) Crystal structure of the RNase H domain of HIV-1 reverse transcriptase. *Science* 252:88-95.
- Davis EO, Sedgwick SG, Colston MJ (1991) Novel structure of the *recA* locus of *Mycobacterium tuberculosis* implies processing of the gene product. *J. Bacteriol.* 173:5633-5662.
- Dawes SS, Crouch RJ, Morris SL, Mizrahi V (1995) Cloning, sequence analysis, overproduction in *Escherichia coli* and enzymatic characterization of the RNase HI from *Mycobacterium smegmatis*. *Gene* 165:71-75.
- Day, M (1996) 'Poor' vaccine ruled out in TB puzzle. *New Scientist*, 13 July, pp 14.
- De Beenhouwer H, Lhiang Z, Jannes G, Mijs W, Machtelinckx L, Rossau R, Traore H (1995) Rapid detection of rifampicin resistance in sputum and biopsy specimens from tuberculosis patients by PCR and line probe assay. *Tuber. Lung Dis.* 76:425-430.
- de Crouy-Chanel A, Hodges RS, Kohiyama M, Richarme G (1997) DnaJ potentiates the interaction between DnaK and alpha-helical peptides. *Biochem. Biophys. Res. Commun.* 233:627-630.
- de Crouy-Chanel A, Kohiyama M, Richarme G (1996) Specificity of DnaK for arginine/lysine and effect of DnaJ on the amino acid specificity of DnaK. *J. Biol. Chem.* 271:15486-15490.
- De Wit D, Steyn L, Shoemaker S, Sogin M (1990) Direct detection of *Mycobacterium tuberculosis* in clinical specimens by DNA amplification. *J.Clin. Microbiol.* 28:2437-2441.
- Delic I, Robbins P, Westpheling J (1992) Direct repeat sequences are implicated in the regulation of two *Streptomyces* chitinase promoters that are subject to carbon catabolite control. *Proc. Natl. Acad. Sci. USA* 89:1885-1889.
- Dellagostin OA, Esposito G, Eales LJ, Dale JW, McFadden J (1995) Activity of mycobacterial promoters during intracellular and extracellular growth. *Microbiology* 141:1785-1792.

- Deshpande RG, Khan MB, Bhat DA, Navalkar RG (1997)** Isolation of a contact-dependent haemolysin from *Mycobacterium tuberculosis*. *J. Med. Microbiol.* 46:233-238.
- Dhandayuthapani S, Via LE, Thomas CA, Horowitz PM, Deretic D, 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.
- Dobos KM, Khoo KH, Swiderek KM, Brennan PJ, Belisle JT (1996)** Definition of the full extent of glycosylation of the 45-kilodalton glycoprotein of *Mycobacterium tuberculosis*. *J. Bacteriol.* 178:2498-2506.
- Dobos KM, Swiderek K, Khoo KH, Brennan PJ, Belisle JT (1995)** Evidence for glycosylation sites on the 45-kilodalton glycoprotein of *Mycobacterium tuberculosis*. *Infect. Immun.* 63:2846-2853.
- Doolittle R, Feng DF, Johnson MS, McClure MA (1989)** Origins and evolutionary relationships of retroviruses. *Q. Rev. Biol.* 64:1-30.
- Draper, DE (1996)** Translation initiation. In: *Escherichia coli and Salmonella typhimurium. Cellular and Molecular Biology*. 2nd Ed. (ed. Neidhardt FC, Curtiss R, Ingraham JL, Lin ECC, Low KB, Magasanik B, Reznikoff WS, Riley M, Schaechter M, Umberger HE). Washington DC: ASM Press. pp. 902-908.
- Drolet M, Phoenix P, Menzel R, Masse E, Liu LF, Crouch RJ (1995)** Overexpression of RNase III partially complements the growth defect of an *Escherichia coli* delta *topA* mutant: R-loop formation is a major problem in the absence of DNA topoisomerase I. *Proc Natl Acad Sci USA* 92:3526-3530.
- Dyda F, Hickman AB, Jenkins TM, Engelman A, Craigie R, Davies DR (1994)** Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science* 266:1981-1986.
- Efferen LS (1997)** Tuberculosis update: will good news become bad news? *Curr. Opin. Pulm. Med.* 3:131-138.
- el-Helou P, Rachlis A, Fong I, Walmsley S, Phillips A, Salit I, Simor AE (1997)** *Mycobacterium xenopi* infection in patients with human immunodeficiency virus infection. *Clin Infect Dis.* 25:206-210.
- Ellner JJ, Goldberger MJ, Parenti DM (1991)** *Mycobacterium avium* infection and AIDS: a therapeutic dilemma in rapid evolution. *J. Infect. Dis.* 163:1326-1335.

- Escuyer V, Haddad N, Frehel C, Berche P (1996)** Molecular characterization of a surface-exposed superoxide dismutase from *Mycobacterium avium*. *Microb. Pathog.* 20:41-55.
- Fenton MJ, Vermeulen MW (1996)** Immunopathology of tuberculosis: roles of macrophages and monocytes. *Infect. Immun.* 64:683-690.
- Fernandez S, Rojo F, Alonso JC (1997)** The *Bacillus subtilis* chromatin-associated protein Hbsu is involved in DNA repair and recombination. *Mol. Microbiol.* 23:1169-1179.
- Fine PE (1995)** Variation in protection by BCG: implications of and for heterologous immunity. *Lancet* 346:1339-1345.
- Fiss EH, Yu S, Jacobs WR Jr (1994)** Identification of genes involved in the sequestration of iron in mycobacteria: the ferric exochelin biosynthetic and uptake pathways. *Mol. Microbiol.* 14:337-569.
- Foster PL, Marinus MG (1992)** Levels of epsilon, an essential replication subunit of *Escherichia coli* DNA polymerase III, are controlled by heat shock proteins. *J. Bacteriol.* 174:7509-7516.
- Foster PL, Sullivan AD, Franklin SB (1989)** Presence of the *dnaQ-rnh* divergent transcriptional unit on a multicopy plasmid inhibits induced mutagenesis in *Escherichia coli*. *J. Bacteriol.* 171:3144-3151.
- Fournier S, Plaioux G, Vincent V (1998)** *Mycobacterium genavense* and cutaneous disease in AIDS. *Ann. Intern. Med.* 128:409.
- Fraser CM, Gocayne JD, White O, Adams MD, Clayton RA, Fleischmann RD, Bult CJ, Kerlavage AR, Sutton G, Kelley JM, Fritchman JL, Weidman JE, Small KV, Sandusky M, Fuhrman JL, Nguyen LT, Utterback TR, Saudek DM, Phillips CA, Merrick JM, Tomb J-F, Dougherty BA, Bott KF, Hu P-C, Lucier TS, Peterson SN, Smith HO, Hutchison CAI, Venter JC (1995)** The minimal gene complement of *Mycoplasma genitalium*. *Science* 270:397-403.
- Fsihi H, Cole ST (1995)** The *Mycobacterium leprae* genome: systematic sequence analysis identifies key catabolic enzymes, ATP-dependent transport systems and a novel *polA* locus associated with genome variability. *Mol. Microbiol.* 16:909-919.
- Fsihi H, Vincent V, Cole ST (1996a)** Homing events in the *gyrA* of some mycobacteria. *Proc. Natl. Acad. Sci. USA* 93:3410-3415.

- Fsihi H, De Rossi E, Salazar L, Cantoni R, Labo M, Riccardi G, Takiff HE, Eiglmeier K, Bergh S, Cole ST (1996b)** Gene arrangement and organization in a approximately 76 kb fragment encompassing the *oriC* region of the chromosome of *Mycobacterium leprae*. *Microbiology* **142**:3147-3161.
- Fuerst TR, Stover CK, de la Cruz VF (1991)** Development of BCG as a live recombinant vector system: potential use as an HIV vaccine. *Biotechnol. Ther.* **2**:159-178.
- Fujiwara N (1997)** Distribution of antigenic glycolipids among *Mycobacterium tuberculosis* strains and their contribution to virulence. *Kekkaku* **72**:193-205.
- Fulcuoh A, Iwasaki H, Ishioka K, Ghinagawa H (1997)** ATP-dependent resolution of R-loops at the ColE1 replication origin by *Escherichia coli* RecG protein, a Holliday junction-specific helicase. *EMBO J.* **16**:203-209.
- Fullner KJ, Hatfull GF (1997)** Mycobacteriophage L5 infection of *Mycobacterium bovis* BCG: implications for phage genetics in the slow growing mycobacteria. *Mol. Microbiol.* **26**:755-766.
- Gadre DV (1997)** HIV associated chronic atypical osteomyelitis by *Mycobacterium fortuitum-chelonae* group—a case report. *Indian J. Med. Sci.* **51**: 161-163.
- Gamer J, Hermann B, Bukau B (1992)** Physical interaction between heat shock proteins DnaK, DnaJ, and GrpE and the bacterial heat shock transcription factor σ^{32} . *Cell* **69**:833-842.
- Gangadharam PR, Perumal VK, Crawford JT, Bates JH (1988)** Association of plasmids and virulence of *Mycobacterium avium* complex. *Am. Rev. Respir. Dis.* **137**:212-214.
- Garbe TR, Barathi J, Barnini S, Zhang Y, Abou-Zeid C, Tang D, Mukherjee R, Young DB (1994)** Transformation of mycobacterial species using hygromycin resistance as selectable marker. *Microbiology* **140**:133-138.
- Garrat V, Cadranet J, Esvant N, Herry I, Morinet P, Mayaud C, Israel-Biet D (1997)** Tuberculosis generates a microenvironment enhancing the productive infection of local lymphocytes by HIV. *J. Immunol.* **159**:2824-2830.
- Gavigan JA, Guilhot C, Gicquel B, Martin C (1995)** Use of conjugative and thermosensitive cloning vectors for transposon delivery to *Mycobacterium smegmatis*. *FEMS Microbiol. Lett.* **127**:35-39.

- Gavigan JA, Ainsa JA, Perez E, Otal I, Martin C** (1997) Isolation by genetic labeling of a new mycobacterial plasmid, pJAZ38, from *Mycobacterium fortuitum*. *J. Bacteriol.* **179**:4115-4122.
- Georgopolous C, Ang D, Liberek K, Zylicz M** (1990) Properties of the *Escherichia coli* heat shock proteins and their role in bacteriophage λ growth. In: *Stress Proteins in Biology and Medicine*. (ed. Morimoto RI, Tissieres A, Georgopolous C). Cold Spring Harbor, NY: Cold Spring Harbor Press. pp. 191-221.
- Georgopolous C, Liberek K, Zylicz M, Ang D** (1994) Properties of the heat shock proteins of *Escherichia coli* and the autoregulation of the heat shock response. In: *The Biology of Heat Shock Proteins and Molecular Chaperones*. (ed. Morimoto RI, Tissieres A, Georgopolous C). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press. pp. 209-249.
- Gicquel B** (1994) Towards new mycobacterial vaccines. *Dev. Biol. Stand.* **82**:171-178.
- Gicquel-Sanzey B, Moniz-Pereira J, Gheorghiu M, Ranzier J** (1989) Structure of pAL5000, a plasmid from *M. fortuitum* and its utilization in transformation of mycobacteria. *Acta Leprol* **7**:208-211.
- Glockner AB, Zumft WG** (1996) Sequence analysis of an internal 9.72-kb segment from the 30-kb denitrification gene cluster of *Pseudomonas stutzeri*. *Biochim. Biophys. Acta* **1277**:6-12.
- Gonzalez-y-Merchand JA, Garcia MJ, Gonzalez-Rico S, Colston MJ, Cox RA** (1997) Strategies used by pathogenic and nonpathogenic mycobacteria to synthesize rRNA. *J. Bacteriol.* **179**:6949-6958.
- Goodfellow M, Cross T** (1984) Classification. In: *The Biology of the Actinomycetes*. (ed. M Goodfellow, M Mordarski, ST Williams) London: Academic Press. pp. 7-164.
- Gordhan BG, Andersen SJ, De Meyer AR, Mizrahi V** (1996) Construction by homologous recombination and phenotypic characterization of a DNA polymerase domain *polA* mutant of *Mycobacterium smegmatis*. *Gene* **178**:125-130.
- Gordhan BG, Mizrahi V** (1997) The RLEP-flanked *polA* gene from *Mycobacterium leprae* is not transcribed in *Mycobacterium smegmatis*. *Gene* **187**:63-66.

- Goto Y, Taniguchi H, Udon T, Mizuguchi Y, Tokunaga T (1991) Development of a new host vector system in mycobacteria. *FEMS Microbiol. Lett.* 67:277-282.
- Guerrero A, Coño J, Fortun J, Navas E, Quereda C, Asensio A, Canon J, Blazquez J, Gomez-Mampaso E (1997) Nosocomial transmission of *Mycobacterium bovis* resistant to 11 drugs in people with advanced HIV-1 infection. *Lancet.* 350:1738-1742.
- Guilhot C, Gicquel B, Martin C (1992) Temperature-sensitive mutants of the *Mycobacterium* plasmid pAL5000. *FEMS Microbiol. Lett.* 77:181-186.
- Guilhot C, Ota I, Van Rompaey I, Martin C, Gicquel B (1994) Efficient transposition in mycobacteria: construction of *Mycobacterium smegmatis* insertional mutant libraries. *J. Bacteriol.* 176:535-539.
- Guleria I, Teitelbaum R, McAdam RA, Kalpana G, Jacobs WR Jr, Bloom BR (1996) Auxotrophic vaccines for tuberculosis. *Nat. Med.* 2:334-337.
- Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166:557-580.
- Harshey RM, Ramakrishnan T (1977) Rate of ribonucleic acid chain growth in *Mycobacterium tuberculosis* H37Rv. *J. Bacteriol.* 129:616-622.
- Hart DHL, Peel MM, Andrew JH, Burdon JGW (1988) Lung infection caused by *Rhodococcus*. *Aust. NZ. J. Med.* 18:790-801.
- Hart PD and Sutherland I (1977) BCG and vole bacillus vaccines in the prevention of tuberculosis in adolescence and early adult life. *Br. Med J.* 22:293-295.
- Harth G, Lee BY, Horwitz MA (1997) High-level heterologous expression and secretion in rapidly growing non-pathogenic mycobacteria of four major *Mycobacterium tuberculosis* extracellular proteins considered to be leading vaccine candidates and drug targets. *Infect. Immun.* 65:2321-2328.
- Haruki M, Noguchi E, Nakai C, Liu YY, Oobatake M, Itaya M, Kanaya S (1994) Investigating the role of conserved residue Asp134 in *Escherichia coli* ribonuclease HI by site-directed random mutagenesis. *Eur. J. Biochem.* 220:623-631.
- Hemavathy KC, Nagaraja V (1995) DNA methylation in mycobacteria: absence of methylation at GATC (Dam) and CCA/TGG (Dcm) sequences. *FEMS Immunol. Med. Microbiol.* 11:291-296.

- Herrmann JL, O'Gaora P, Gallagher A, Thole JE, Young DB (1996) Bacterial glycoproteins: a link between glycosylation and proteolytic cleavage of a 19 kDa antigen from *Mycobacterium tuberculosis*. *EMBO J.* **15**:3547-3554.
- Hewinson RG, Michell SL, Russell WP, McAdam RA, Jacobs WR Jr (1996) Molecular characterization of MPT83: a seroreactive antigen of *Mycobacterium tuberculosis* with homology to MPT70. *Scand. J. Immunol.* **43**:490-499.
- Hiushelwood S, Stoker NG (1992) An *Escherichia coli*-*Mycobacterium* shuttle cosmid vector, pMSC1. *Gene* **110**:115-118.
- Hocqueloux L, Lesprit P, Herrmann JL, de La Blanchardiere A, Zagdanski AM, Decazes JM, Modai J (1998) Pulmonary *Mycobacterium avium* complex disease without dissemination in HIV-infected patients. *Chest.* **113**: 542-548.
- Holm L, Sander C (1994) Searching protein structure databases has come of age. *Proteins* **19**:165-173.
- Hong X, Cadwell GW, Kogoma T (1995) *Escherichia coli* RecG and RecA proteins in R-loop formation. *EMBO J.* **14**:2385-2392.
- Hong X, Cadwell GW, Kogoma T (1996) Activation of stable DNA replication in rapidly growing *Escherichia coli* cells at the time of entry to stationary phase. *Mol. Microbiol.* **21**:953-961.
- Honoré N, Bergh S, Chanteau S, Doucet-Populaire F, Eiglmeier K, Garnier T, Georges C, Launois P, Limpaihoon T, Newton S, *et al.*, (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.
- Horiuchi T, Maki H, Maruyama M, Sekiguchi M (1981) Identification of the *dnaQ* gene product and location of the structural gene for RNase H of *Escherichia coli* by cloning of the genes. *Proc. Natl. Acad. Sci. USA* **78**:3770-3774.
- Horiuchi T, Maki H, Sekiguchi M (1984) RNase H-defective mutants of *Escherichia coli*: a possible discriminatory role of RNase H in initiation of DNA replication. *Mol. Gen. Genet.* **195**:17-22.
- Hostomsky Z, Hostomska Z, Hudson GO, Moomaw EW, Nodes BR (1991) Reconstitution *in vitro* of RNase H activity by using purified N-terminal and C-terminal domains of human immunodeficiency virus type 1 reverse transcriptase. *Proc. Natl. Acad. Sci. USA* **88**:1148-1152.

- Hsieh PC, Shenoy BC, Samols D, Phillips NF (1996) Cloning, expression, and characterization of polyphosphate glucokinase from *Mycobacterium tuberculosis*. *J. Biol. Chem.* 271:4909-4915.
- Huberts P, Mizrahi V (1995) Cloning and sequence analysis of the gene encoding the DNA polymerase I from *Mycobacterium tuberculosis*. *Gene* 164:133-136.
- Husson RN, James BE, Young RA (1990) Gene replacement and expression of foreign DNA in mycobacteria. *J. Bacteriol.* 172:519-524.
- Inouye S, Inouye M (1993) The retron: a bacterial retroelement required for the synthesis of msDNA. *Curr. Opin. Genet. Dev.* 3:713-718.
- Ishikawa K, Okumura M, Katayanagi K, Kimura S, Kanaya S, Nakamura H, Morikawa K (1993) Crystal structure of ribonuclease H from *Thermus thermophilus* HB8 refined at 2.8 Å resolution. *J. Mol. Biol.* 230:529-542.
- Itaya M (1990) Isolation and characterization of a second RNase H (RNase HII) of *Escherichia coli* K12 encoded by the *rnhB* gene. *Proc. Natl. Acad. Sci. USA* 87:8587-8591.
- Itaya M, Crouch RJ (1991a) A combination of RNase II (*rnh*) and *recBCD* or *sbcB* mutations in *Escherichia coli* K12 adversely affects growth. *Mol. Gen. Genet.* 227:424-432.
- Itaya M, Crouch RJ (1991b) Correlation of activity with phenotypes of *Escherichia coli* partial function mutants of *rnh*, the gene encoding RNase H. *Mol. Gen. Genet.* 227:433-437.
- Itaya M, Kondo K (1991) Molecular cloning of a ribonuclease H (RNase HI) gene from an extreme thermophile *Thermus thermophilus* HB8: a thermostable RNase H can functionally replace the *Escherichia coli* enzyme *in vivo*. *Nucleic Acids Res.* 19:4443-4449.
- Itaya M, McKelvin D, Chatterjee SK, Crouch RJ (1991) Selective cloning of genes encoding RNase H from *Salmonella typhimurium*, *Saccharomyces cerevisiae* and *Escherichia coli* *rnh* mutant. *Mol. Gen. Genet.* 227:438-445.
- Itoh T, Tomizawa J (1980) Formation of an RNA primer for initiation of replication of ColE1 DNA by ribonuclease H. *Proc. Natl. Acad. Sci. USA* 77:2450-2454.
- Iwai S, Wakasa M, Ohtsuka E, Kanaya S, Kidera A, Nakamura H (1996) Interaction of the basic protrusion of *Escherichia coli* RNase HI with its substrate. *J. Mol. Biol.* 263:699-706.

- Iwai S, Kataoka S, Wakasa M, Ohtsuka E, Nakamura H (1995) Recognition of 2'-hydroxyl groups by *Escherichia coli* ribonuclease HI. *FEBS Lett* 368:315-320.
- Iwakura M, Kawata M, Tsuda K, Tanaka T (1988) Nucleotide sequence of the thymidylate synthase B and dihydrofolate reductase genes contained in one *Bacillus subtilis* operon. *Gene* 64:9-20.
- Jackson M, Berthet F-X, Otal I, Rauzier J, Martin C, Gicquel B, Guilhot C (1996) The *Mycobacterium tuberculosis* purine biosynthetic pathway: isolation and characterization of the *purC* and *purL* genes. *Microbiology* 142:2439-2447.
- Jacobs WR, Tuckman M, Bloom BR (1987) Introduction of foreign DNA into mycobacteria using a shuttle plasmid. *Nature* 327:532-535.
- Jacobs WR Jr, Ganjam KV, Cirillo JD, Pascopella L, Snapper SB, Udani RA, Jones W, Barletta RG, Bloom BR (1991) Genetic systems for the mycobacteria. *Methods Enzymol.* 204:537-555.
- Jacobs WR Jr, Barletta RG, Udani R, Chan J, Kalkut G, Sosne G, Keiser T, Sarkis GJ, Hatfull GF, Bloom BR (1993) Rapid assessment of susceptibilities of *Mycobacterium tuberculosis* by means of luciferase reporter phages. *Science* 260:819-822.
- Ji YE, Colston MJ, Cox RA (1994) The ribosomal RNA (*rri*) operons of fast-growing mycobacteria: primary and secondary structures and their relation to *rrn* operons of pathogenic slow-growers. *Microbiology* 140:2829-2840.
- Johnsson B, Löfås S, Lindqvist G (1991) Immobilization of proteins to a carboxymethyl-dextran modified gold surface for biospecific interaction analysis in surface plasmon resonance. *Anal. Biochem.* 198:268-277.
- Johnston TC, Thompson RB, Baldwin TO (1986) Nucleotide sequence of the *luxB* gene of *Vibrio harveyi* and the complete amino acid sequence of the beta subunit of bacterial luciferase. *J. Biol. Chem.* 261:4805-4811.
- Kalpana GV, Bloom BR, Jacobs WR Jr (1991) Insertional mutagenesis and illegitimate recombination in mycobacteria. *Proc. Natl. Acad. Sci. USA* 88:5433-5437.
- Kanaya E, Kanaya S (1995) Kinetic analysis of *Escherichia coli* RNase HI using oligomeric DNA/RNA substrates suggests an alternative mechanism between the enzyme and the substrate. *Eur. J. Biochem.* 231:557-562.

- Kanaya S, Crouch RJ (1983) DNA sequence of the gene coding for *Escherichia coli* ribonuclease H. *J. Biol. Chem.* 258:1276-1281.
- Kanaya S, Crouch RJ (1984) The *rnh* gene is essential for growth of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 81:3447-3451.
- Kanaya S, Ikehara M (1993) Structure and function of RNase HI from *Escherichia coli*. In: *Nucleic Acids and Molecular Biology* (ed, Eckstein F, Lilley DMJ) Vol. 7. Berlin, Heidelberg: Springer-Verlag. pp. 285-302.
- Kanaya S, Kohara A, Miura Y, Sekiguchi A, Iwai S, Inoue H, Ohtsuka E, Ikehara M (1990a) Identification of the amino acid residues involved in an active site of *Escherichia coli* ribonuclease H by site-directed mutagenesis. *J. Biol. Chem.* 265:4615-4621.
- Kanaya S, Kimura S, Katsuda C, Ikehara M (1990b) Role of cysteine residues in ribonuclease H from *Escherichia coli*. Site-directed mutagenesis and chemical modification. *Biochem. J.* 271:59-66.
- Kanaya S, Katayanagi K, Morikawa K, Inoue H, Ohtsuka E, Ikehara M (1991a) Effect of mutagenesis at each of five histidine residues on enzymatic activity and stability of ribonuclease H from *Escherichia coli*. *Eur. J. Biochem.* 198:437-440.
- Kanaya S, Katsuda C, Kimura S, Nakai T, Kitakuni E, Nakamura H, Katayanagi K, Morikawa K, Ikehara M (1991b) Stabilization of *Escherichia coli* ribonuclease H by introduction of an artificial disulfide bond. *J. Biol. Chem.* 266:6038-6044.
- Kanaya S, Katsuda-Nakai C, Ikehara M (1991c) Importance of the positive charge cluster in *Escherichia coli* ribonuclease HI for the effective binding of the substrate. *J. Biol. Chem.* 266:11621-11627.
- Kanaya S, Oobatake M, Liu Y (1996) Thermal stability of *Escherichia coli* ribonuclease HI and its active site mutants in the presence and absence of the Mg^{2+} ion. Proposal of a novel catalytic role for Glu48. *J. Biol. Chem.* 271:32729-32736.
- Kaneko T, Tanaka A, Sato S, Kotani H, Sazuka T, Miyajima N, Sugiura M, Tabata S (1995) Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. I. Sequence features in the 1 Mb region from map positions 64% to 92% of the genome. *DNA Res.* 2:153-166.

- Kang BK, Schlesinger LS (1998) *M. tuberculosis* lipoarabinomannan is a microbial determinant that reduces phagosome-lysosome fusion following phagocytosis events in human macrophages. TB:Molecular Mechanisms and Immunologic Aspects; Keystone, CO; April 2-8.
- Karonga Prevention Trial Group (1996) Randomised controlled trial of single BCG, repeated BCG, or combined BCG and killed *Mycobacterium leprae* vaccine for prevention of leprosy and tuberculosis in Malawi. *Lancet* 348:17-24.
- Karzai AW, McMacken R (1996) A bipartite signaling mechanism involved in DnaJ- mediated activation of the *Escherichia coli* DnaK protein. *J. Biol. Chem.* 271:11236-11246.
- Katayanagi K, Miyagawa M, Matsushima M, Ishikawa M, Kanaya S, Nakamura H, Ikehara M, Matsuzaki T, Morikawa K (1992) Structural details of ribonuclease H from *Escherichia coli* as refined to an atomic resolution. *J. Mol. Biol.* 223:1029-1052.
- Katayanagi K, Okumura M, Morikawa K (1993) Crystal structure of *Escherichia coli* RNase H in complex with Mg^{2+} at 2.8 Å resolution: proof for a single Mg^{2+} -binding site. *Proteins* 17:337-346.
- Keck JL, Marqusee S (1995) Substitution of a highly basic helix/loop sequence into the RNase H domain of human immunodeficiency virus reverse transcriptase restores its Mn^{2+} -dependent RNase H activity. *Proc. Natl. Acad. Sci. USA* 92:2740-2744.
- Keck JL, Marqusee S (1996) The putative substrate recognition loop of *Escherichia coli* ribonuclease H is not essential for activity. *J. Biol. Chem.* 271:19883-19887.
- Kenney TJ, Churchward G (1996) Genetic analysis of the *Mycobacterium smegmatis* *rpsL* promoter. *J. Bacteriol.* 178:3564-3571.
- Kieser T, Moss MT, Dale JW, Hopwood DA (1986) Cloning and expression of *Mycobacterium bovis* BCG DNA in "*Streptomyces lividans*". *J. Bacteriol.* 168:72-80.
- Kitani T, Yoda K, Ogawa T, Okazaki T (1985) Evidence that discontinuous DNA replication in *Escherichia coli* is primed by approximately 10 to 12 residues of RNA starting with a purine. *J. Mol. Biol.* 184:45-52.

- Klajn AG, Belanger AE, Albanes-De Mello A, Lee JY, Hatfull GF (1998) Characterization of the *dnaG* locus in *Mycobacterium smegmatis* reveals linkage of DNA replication and cell division. *J. Bacteriol.* 180:65-72.
- Klopman G, Wang S, Jacobs MR, Ellner JJ (1993) Anti-*Mycobacterium avium* activity of quinolones: structure-activity relationship studies. *Antimicrob. Agents Chemother.* 37:1807-1815.
- Kochi A (1991) The global tuberculosis situation and the new control strategy of the World Health Organization. *Tubercle* 72:1-6.
- Kogoma T (1994) *Escherichia coli* RNA polymerase mutants that enhance or diminish the SOS response constitutively expressed in the absence of RNase H1 activity. *J. Bacteriol.* 176:1521-1523
- Kogoma T (1997) Stable DNA replication: interplay between DNA replication, homologous recombination, and transcription. *Microbiol. Mol. Biol. Rev.* 61:212-238.
- Kogoma T, Hong X, Cadwell GW, Barnard KG, Asai T (1993) Requirement of homologous recombination functions for viability of the *Escherichia coli* cell that lacks RNase H1 and exonuclease V activities. *Biochimie* 75:89-99.
- Kogoma T, Maldonado RR (1997) DNA polymerase I in constitutive stable DNA replication in *Escherichia coli*. *J. Bacteriol.* 179:2109-2115.
- Kong TH, Coates ARM, Butcher PD, Hickman CJ, Shinnick TM (1993) *Mycobacterium tuberculosis* expresses two chaperonin-60 homologs. *Microbiology* 90:2608-2612.
- Koonin EV, Mushegian AR, Rudd KE (1996) Sequencing and analysis of bacterial genomes. *Curr. Biol.* 6:404-416.
- Kornberg A, Baker TA (1992) DNA replication. W. H. Freeman & Co., New York.
- Kowalczykowski SC, Dixon DA, Eggleston AK, Lauder SD, Rehrauer WM (1994) Biochemistry of homologous recombination in *Escherichia coli*. *Microbiol Rev.* 58:401-465.
- Kremer L, Baulard A, Estaquier J, Poulain-Godefroy O, Locht C (1995a) Green fluorescent protein as a new expression marker in mycobacteria. *Mol. Microbiol* 17:913-922.
- Kremer L, Baulard A, Estaquier J, Content J, Capron A, Locht C (1995b) Analysis of the *Mycobacterium tuberculosis* 85A antigen promoter region. *J. Bacteriol.* 177:642-653.

- Kudlicki W, Odom OW, Kramer G, Hardesty B (1996) Binding of an N-terminal rhodanese peptide to DnaJ and to ribosomes. *J. Biol. Chem.* 271:31150-31155.
- Kumar A, Malloch RA, Fujita N, Smillie DA, Ishihama A, Hayward RS (1993) The minus 35-recognition region of *Escherichia coli* sigma 70 is inessential for initiation of transcription at an "extended minus 10" promoter. *J. Mol. Biol.* 232:406-418.
- Kumar RA, Vaze MB, Chandra NR, Vijayan M, Muniyappa K (1996) Functional characterization of the precursor and spliced forms of RecA protein from *Mycobacterium tuberculosis*. *Biochemistry* 35:1793-1802.
- Kunkel TA, Roberts JD, Zakour RA (1987) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Meth. Enzymol.* 154:367-383.
- Labidi A, David HL, Roulland-Dussoix D (1985a) Restriction endonuclease mapping and cloning of the *Mycobacterium fortuitum* var. *fortuitum* plasmid pAL5000. *Ann. Inst. Pasteur-Microbiol.* 136:209-215.
- Labidi A, David HL, Roulland-Dussoix D (1985b) Cloning and expression of mycobacterial plasmid DNA in *Escherichia coli*. *FEMS Microbiol. Lett.* 30: 221-225.
- Leao SC, Rocha CL, Murillo LA, Parra CA, Patarroyo ME (1995) A species-specific nucleotide sequence of *Mycobacterium tuberculosis* encodes a protein that exhibits hemolytic activity when expressed in *Escherichia coli*. *Infect. Immun.* 63:4301-4306.
- Lee MH, Pascopella L, Jacobs WRJ, Hatfull GF (1991) Site-specific integration of mycobacteriophage L5: Integration-proficient vectors for *Mycobacterium smegmatis*, *Mycobacterium tuberculosis*, and bacille Calmette-Guérin. *Proc. Natl. Acad. Sci. USA* 88:3111-3115.
- Lee RE, Brennan PJ, Besra GS (1996) *Mycobacterium tuberculosis* cell envelope. *Curr. Top. Microbiol. Immunol.* 215:131-156.
- Lefevre P, Braibant M, de Wit L, Kalai M, Roeper D, Grotzinger J, Delville JP, Piers P, Ooms J, Huygen K, Content J (1997) Three different putative phosphate transporters are encoded by the *Mycobacterium tuberculosis* genome and are present at the surface of *Mycobacterium bovis* BCG. *J. Bacteriol.* 179:2900-2906.

- Lety MA, Nair S, Berche P, Escuyer V** (1997) A single point mutation in the *embB* gene is responsible for resistance to ethambutol in *Mycobacterium smegmatis*. *Antimicrob. Agents Chemother.* 41:2629-2633.
- Levin ME, Hatfull GF** (1993) *Mycobacterium smegmatis* RNA polymerase:DNA supercoiling, action of rifampicin and mechanism of rifampicin resistance. *Mol. Microbiol.* 8:277-285.
- Liberek K, Osipiuk J, Zylicz M, Ang D, Skorko J, Georgopoulos C** (1990) Physical interactions between bacteriophage and *Escherichia coli* proteins required for initiation of λ DNA replication. *J. Biol. Chem.* 265:3022-3029.
- Liberek K, Galitski TP, Zylicz M, Georgopoulos C** (1992) The DnaK chaperone modulates the heat shock response of *Escherichia coli* by binding to the σ^{32} transcription factor. *Proc. Natl. Acad. Sci. USA* 89:3516-3520.
- Liberek K, Georgopoulos C** (1993) Autoregulation of the *Escherichia coli* heat shock response by the DnaK and DnaJ heat shock proteins. *Proc. Natl. Acad. Sci. USA* 90:11019-11023.
- Liberek K, Wall D, Georgopoulos C** (1995) The DnaJ chaperone catalytically activates the DnaK chaperone to preferentially bind the σ^{32} heat shock transcriptional regulator. *Biochemistry* 92:6224-6228.
- Lienhardt C, Rodrigues LC** (1997) Estimation of the impact of the human immunodeficiency virus infection on tuberculosis: tuberculosis risks revisited ? *Int. J. Tuberc. Lung Dis.* 1:196-204.
- Lima TM, Lim D** (1995) Isolation and characterization of host mutants defective in msDNA synthesis: role of ribonuclease H in msDNA synthesis. *Plasmid* 33:235-238.
- Lima WF, Crooke ST** (1997) Cleavage of single strand RNA adjacent to RNA-DNA duplex regions by *Escherichia coli* RNase HI. *J. Biol. Chem.* 272:27513-27516.
- Link AJ, Phillips D, Church GM** (1997) Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization. *J. Bacteriol.* 179:6228-6237.
- Lowrie DB, Silva CL, Tascon RE** (1997) Genetic vaccination against tuberculosis. *Springer Semin. Immunopathol.* 19:161-173.

- Lugosi L, Jacobs WR Jr, Bloom BR (1989) Genetic transformation of BCG. *Tubercle* 70:159-170.
- Ma Y, Mills JA, Belisle JT, Vissa V, Howell M, Bowlin K, Scherman MS, McNeil M (1997) Determination of the pathway for rhamnose biosynthesis in mycobacteria: cloning, sequencing and expression of the *Mycobacterium tuberculosis* gene encoding alpha-D-glucose-1-phosphate thymidyltransferase. *Microbiology* 143:937-945.
- Madhusudan K, Ramesh V, Nagaraja V (1994) Molecular cloning of *gyrA* and *gyrB* genes of *Mycobacterium tuberculosis*: analysis of nucleotide sequence. *Biochem. Mol. Biol. Int.* 33:651-660.
- Mahairas GG, Sabo PJ, Hickey MJ, Singh DC, Stover CK (1996) Molecular analysis of genetic differences between *Mycobacterium bovis* BCG and virulent *M. bovis*. *J. Bacteriol.* 178:1274-1282.
- Malmqvist M (1993) Biospecific interaction analysis using biosensor technology. *Nature* 361:186-187.
- Maloney JM, Gregg CR, Stephens DS, Manian FA, Rimland D (1987) Infections caused by *Mycobacterium szulgai* in humans. *Rev. Infect. Dis.* 9:1120-1126.
- Marsh BJ, Fordham von Reyn C, Arbeit RD, Morin P (1997) Immunisation of HIV-infected adults with a three dose series of inactivated *Mycobacterium vaccae*. *Am. J. Med. Sci.* 313:377-38.
- Matsumoto S, Tamaki M, Yukitake H, Matsuo T, Naito M, Teraoka H, Yamada T (1996) A stable *Escherichia coli*-mycobacteria shuttle vector 'pSO246' in *Mycobacterium bovis* BCG. *FEMS Microbiol. Lett.* 135:237-243.
- McAdam RA, Weisbrod TR, Martin J, Scuderi JD, Brown AM, Cirillo JD, Bloom BR, Jacobs WR Jr (1995) *In vivo* growth characteristics of leucine and methionine auxotrophic mutants of *Mycobacterium bovis* BCG generated by transposon mutagenesis. *Infect. Immun.* 63:1004-1012.
- Meier A, Kirschner P, Bange FC, Vogel V, Bottger EC (1994) Genetic alterations in streptomycin resistant *Mycobacterium tuberculosis*: mapping of mutations conferring resistance. *Antimicrob Agents Chemother.* 38:228-233.
- Meier A, Heifets L, Wallace RJ Jr, Zhang Y, Brown BA, Sander P, Bottger EC (1996) Molecular mechanisms of clarithromycin resistance in *Mycobacterium avium*: observation of multiple 23S rDNA mutations in a clonal population. *J. Infect. Dis.* 174:354-360.

- Messer W, Weigel C (1996) In: *Escherichia coli* and *Salmonella typhimurium: Cellular and Molecular Biology*. (eds. Neidhardt, FC, Curtiss R, Ingraham JL, Lin ECC, Low KB, Magasanik B, Reznikoff WS, Riley M, Schaechter M, Umberger HE) Washington DC: ASM Press. pp. 1579-1601.
- Micka B, Marahiel MA (1992) The DNA-binding protein Hbsu is essential for normal growth and development in *Bacillus subtilis*. *Biochimie* 74:641-650.
- Miesel L, Roth JR (1996) Evidence that SbcB and RecF pathway functions contribute to RecBCD-dependent transductional recombination. *J. Bacteriol.* 178:3146-3155.
- Mizrahi V (1989) Analysis of the ribonuclease H activity of HIV-1 reverse transcriptase using RNA-DNA hybrid substrates from the gag region of HIV-1. *Biochemistry* 28:9088-9094.
- Mizrahi V, Andersen S (1998) DNA repair in *Mycobacterium tuberculosis*. What have we learnt from the genome sequence? *Mol. Microbiol.* (in press).
- Mizrahi V, Huberts P (1996) Deoxy- and dideoxynucleotide discrimination and identification of critical 5' nuclease domain residues of the DNA polymerase I from *Mycobacterium tuberculosis*. *Nucleic Acids Res.* 24:4845-4852.
- Mizrahi V, Huberts P, Dawes SS, Dudding LR (1993) A PCR method for the amplification of the *gyrA*, *polA* and *rnhA* gene segments from mycobacteria. *Gene* 136:287-290.
- Mizrahi V, Itaya M (1998) Other enzymes that degrade the RNA strand of RNA-DNA hybrids. In *Ribonucleases H*. (ed. Crouch RJ, Toulmé JJ) INSERM Editions (Paris) and John Libbey Ltd. in press.
- Molineux JJ, Gefters ML (1975) Properties of *Escherichia coli* DNA-binding (unwinding) protein. Interaction with nucleolytic enzymes and DNA. *J. Mol. Biol.* 98:811-825.
- Monk M, Kinross J (1972) Conditional lethality of *recA* and *recB* derivatives of a strain of *Escherichia coli* K-12 with a temperature-sensitive deoxyribonucleic acid polymerase I. *J. Bacteriol.* 109:971-978.
- Morris S, Han Bai G, Suffys P, Portillo-Gomez L, Fairchok M, Rouse D (1995) Molecular mechanisms of multiple drug resistance in clinical isolates of *Mycobacterium tuberculosis*. *J. Infect. Dis.* 171:954-960.
- Munson MA, Baumann L, Baumann P (1993) *Buchnera aphidicola* (a prokaryotic endosymbiont of aphids) contains a putative 16S rRNA operon unlinked to the

- 23S rRNA-encoding gene: sequence determination, and promoter and terminator analysis. *Gene* 137:171-178.
- Murray A, Winter N, Lagranderie M, Hill DF, Rauzier J, Timm J, Leclerc C, Moriarty KM, Gheorghiu M, 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.
- Murray CJL, Lopez AD (1996) Evidence-based health policy - lessons from the global burden of disease study. *Science* 274:740-743.
- Murray JF (1997) Tuberculosis and HIV infection: global perspectives. *Respirology* 2:209-213.
- Murray PJ, Aldovini A, Young RA (1996) Manipulation and potentiation of antimycobacterial immunity using recombinant bacille Calmette-Guerin strains that secrete cytokines. *Proc. Natl. Acad. Sci. USA* 93:934-939.
- Naito S, Kitani T, Ogawa T, Okazaki T, Uchida H (1984) *Escherichia coli* mutants suppressing replication-defective mutations of the ColE1 plasmid. *Proc. Natl. Acad. Sci. USA*. 81:550-554.
- Nakajima H (1996) Leprosy and guinea worm disease. *World Health* 49(3):3.
- Nash, KA, Inderlied CB (1995) Genetic basis of macrolide resistance in *Mycobacterium avium* isolated from patients with disseminated disease. *Antimicrob. Agents Chemother.* 39:2625-2630.
- Nishitani H, Hadaki M, Horiuchi T (1993) Specific chromosomal sites enhancing homologous recombination in *Escherichia coli* mutants defective in RNase H. *Mol. Gen. Genet.* 240:307-314.
- Nomura T, Fujita N, Ishihama A (1985) Promoter selectivity of *E. coli* RNA polymerase: analysis of the convergently-transcribed *dnaQ-rnh* genes. *Nucleic Acids Res.* 13:7647-7661.
- Noordeen SK (1996) Eliminating leprosy - a winnable war. *World Health* 49(3):4-5.
- Norgard MV, Imaeda T (1978) Physiological factors involved in the transformation of *Mycobacterium smegmatis*. *J. Bacteriol.* 133:1254-1262.
- Norman E, Dellagostin OA, McFadden J, Dale JW (1995) Gene replacement by homologous recombination in *Mycobacterium bovis* BCG. *Mol. Microbiol.* 16:755-760.

- O'Brien KL, Dietz HC, Romagnoli M, Eiden J (1996) Evaluation of *inhA* gene and catalase-peroxidase gene among Isoniazid-sensitive and resistant *Mycobacterium tuberculosis* isolates. *Mol. Cell. Probes* 10:1-6.
- Oda Y, Iwai S, Ohtsuka E, Ishikawa M, Ikehara M, Nakamura H (1993a) Binding of nucleic acids to *E. coli* RNase HI observed by NMR and CD spectroscopy. *Nucleic Acids Res.* 21:4690-4695.
- Oda Y, Yoshida M, Kanaya S (1993b) Role of histidine 124 in the catalytic function of ribonuclease HI from *Escherichia coli*. *J. Biol. Chem* 268:88-92.
- Oda Y, Yamazaki T, Nagayama K, Kanaya S, Kuroda Y, Nakamura H (1994) Individual ionisation constants of all the carboxyl groups in ribonuclease HI from *Escherichia coli* determined by NMR. *Biochemistry* 33:5275-5284.
- Ogawa T, Okazaki T (1984) Function of RNase H in DNA replication revealed by RNase H defective mutants of *Escherichia coli*. *Mol. Gen. Genet.* 193:231-237.
- Packschies L, Theyssen R, Buchberger A, Bukau B, Goody RS, Reinstein J (1997) GrpE accelerates nucleotide exchange of the molecular chaperone DnaK with an associative displacement mechanism. *Biochemistry* 36:3417-3422.
- Paget E, Davies J (1996) Apramycin resistance as a selective marker for gene transfer in mycobacteria. *J. Bacteriol.* 178:6357-6360
- Parker SL, Tsai YL, Palmer CJ (1995) Comparison of PCR-generated fragments of the *mce* gene from *Mycobacterium tuberculosis*, *M. avium*, *M. intracellulare* and *M. scrofulaceum*. *Clin. Diagn. Lab. Immunol.* 2:770-775.
- Pascope L, Collins FM, Martin JM, Lee MH, Hatfull GF, Stover CK, Bloom BR, Jacobs WR Jr (1994) Use of *in vivo* complementation in *Mycobacterium tuberculosis* to identify a genomic fragment associated with virulence. *Infect. Immun.* 62:1313-1319.
- Pellic V, Reyra J-M, Gicquel B (1996a) Expression of the *Bacillus subtilis* *sacB* gene confers sucrose sensitivity on mycobacteria. *J. Bacteriol.* 178:1197-1199.
- Pellic V, Reyra J-M, Gicquel B (1996b) Positive selection of allelic exchange mutants in *Mycobacterium bovis* BCG. *FEMS Microbiol. Lett.* 144:161-166.
- Pellic V, Reyra J-M, Gicquel B (1996c) Generation of unmarked directed mutations in mycobacteria, using sucrose counter-selectable suicide vectors. *Mol. Microbiol.* 20:919-925.

- Pelcic V, Jackson M, Reyat J-M, Jacobs WR Jr, Gicquel B, Guilhot C (1997)** Efficient allelic exchange and transposon mutagenesis in *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. USA* 94: 10955-10960.
- Phillips GJ, Prasher DC, Kushner SR (1988)** Physical and biochemical characterization of cloned *shcB* and *xonA* mutations from *Escherichia coli* K-12. *J. Bacteriol.* 170:2089-2094.
- Piatek AS, Tyagi S, Pol AC, Telenti A, Miller LP, Kramer FR, Alland D (1998)** Molecular beacon sequence analysis for detecting drug resistance in *Mycobacterium tuberculosis*. *Nat. Biotechnol.* 16:359-363.
- Pierpaoli EV, Sandmeier E, Baici A, Schonfeld HJ, Gisler S, Christen P (1997)** The power stroke of the DnaK/DnaJ/GrpE molecular chaperone system. *J. Mol. Biol.* 269:757-768.
- Plum G, Brenden M, Clark-Curtiss JE, Pulverer G (1997)** Cloning, sequencing, and expression of the *mig* gene of *Mycobacterium avium*, which codes for a secreted macrophage-induced protein. *Infect. Immun.* 65:4548-4557.
- Plum G, Clark-Curtiss JE (1994)** Induction of *Mycobacterium avium* gene expression following phagocytosis by human macrophages. *Infect. Immun.* 62:476-483.
- Predich M, Doukhan L, Nair G, Smith I (1995)** Characterization of RNA polymerase and two sigma factor genes from *Mycobacterium smegmatis*. *Mol. Microbiol.* 15:355-366.
- Qin M, Taniguchi H, Mizuguchi Y (1994)** Analysis of the replication region of a mycobacterial plasmid, pMSC262. *J. Bacteriol.* 176:419-425.
- Qin MH, Madiraju MV, Zachariah S, Rajagopalan M (1997)** Characterization of the *oriC* region of *Mycobacterium smegmatis*. *J. Bacteriol.* 179:6311-6317.
- Quinones A, Kucherer C, Piechocki R, Messer W (1987)** Reduced transcription of the *rnh* gene in *Escherichia coli* mutants expressing the SOS regulon constitutively. *Mol. Gen. Genet.* 206:95-100.
- Quinones A, Kaasch J, Kaasch M, Messer W (1989)** Induction of *dnaN* and *dnaQ* gene expression in *Escherichia coli* by alkylation damage to DNA. *EMBO J.* 8:587-593.
- Radford AJ, Hodgson ALM (1991)** Construction and characterization of a *Mycobacterium-Escherichia coli* shuttle vector. *Plasmid* 25:149-153.

- Rajagopalan M, Qin MH, Steingrube VA, Nash DR, Wallace RJ Jr, Madiraju MV (1995) Amplification and cloning of the *Mycobacterium tuberculosis dnaA* gene. *Gene* 163:75-79.
- Ramakrishnan L, Falkow S (1998) Macrophage specific gene expression in *Mycobacterium marinum*. TB:Molecular Mechanisms and Immunologic Aspects; Keystone, CO; April 2-8.
- Ramakrishnan L, Tran HT, Federspiel NA, Falkow S (1997) A *criB* homolog essential for photochromogenicity in *Mycobacterium marinum*: isolation, characterization, and gene disruption via homologous recombination. *J. Bacteriol.* 179:5862-5868.
- Ranes MG, Rauzier J, Lagranderie M, Gheorghiu M, 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.
- Rastogi N, David HL, Rafidinarivo E (1983) Spheroplast fusion as a mode of genetic recombination in mycobacteria. *J. Gen. Microbiol.* 129:1227-1237.
- Ray DS, Hines JC (1995) Disruption of the *Crithidia fasciculata RNH1* gene results in the loss of two active forms of ribonuclease H. *Nucleic Acids Res.* 23:2526-2530.
- Raynaud C, Etienne G, Peyron P, Lancelle MA, Daffe M (1998) Extracellular enzyme activities potentially involved in the pathogenicity of *Mycobacterium tuberculosis*. *Microbiology* 144:577-587.
- Reiter W-D, Palm F (1990) Identification and characterisation of a defective SSV1 genome integrated into a tRNA gene in the archaebacterium *Sulfolobus* sp. B12. *Mol. Gen. Genet.* 221:65-71.
- Reyrat J-M, Berthet F-X, Gicquel B (1995) The urease locus of *Mycobacterium tuberculosis* and its utilization for the demonstration of allelic exchange in *Mycobacterium bovis* bacillus Calmette-Guérin. *Proc. Natl. Acad. Sci. USA* 92:8768-8772.
- Reyrat J-M, Lopez-Ramirez G, Ofredo C, Gicquel B, Winter N (1996) Urease activity does not contribute dramatically to persistence of *Mycobacterium bovis* Bacillus Calmette-Guérin. *Infect. Immun.* 64:3934-3936.

- Rodriguez GM, Gold B, Dussurget O, Smith I (1998) Identification of iron regulated genes in *M. tuberculosis*. TB:Molecular Mechanisms and Immunologic Aspects; Keystone, CO; April 2-8.
- Rosenberg M, Court D (1979) Regulatory sequences involved in the promotion and termination of RNA transcription. *Annu. Rev. Genet.* 13:319-353.
- Rusconi S, Gori A, Vago L, Marchetti G, Franzetti F (1997) Cutaneous infection caused by *Mycobacterium gordonae* in a human immunodeficiency virus-infected patient receiving antimycobacterial treatment. *Clin. Infect. Dis.* 25:1490-1491.
- Russell RL, Jann GJ, Froman S (1963) Lysogeny in the mycobacteria. II. Some phage-host relationships of the lysogenic mycobacteria. *Am. Rev. Resp. Dis.* 88:528-538.
- Salazar L, Fsihi H, de Rossi E, Riccardi G, Rios C, Cole ST, Takiff HE (1996) Organization of the origins of replication of the chromosomes of *Mycobacterium smegmatis*, *Mycobacterium leprae* and *Mycobacterium tuberculosis* and isolation of a functional origin from *M. smegmatis*. *Mol. Microbiol.* 20:283-293.
- Salim K, Haedens V, Content J, Leblon G, Huygen K (1997) Heterologous expression of the *Mycobacterium tuberculosis* gene encoding antigen 85A in *Corynebacterium glutamicum* Appl. Environ. Microbiol. 63:4392-4400.
- Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual. 2nd Ed. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
- Sandigursky M, Franklin WA (1992) DNA deoxyribophosphodiesterase of *Escherichia coli* is associated with exonuclease I. *Nucleic Acids Res.* 20:4699-4703.
- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.
- Sarkis GJ, Jacobs WR Jr, Hatfull GF (1995) L5 luciferase reporter mycobacteriophages: a sensitive tool for the detection and assay of live mycobacteria. *Mol. Microbiol.* 15:1055-1067.
- Schaal KP, Lee H-J (1992) Actinomycete infection in humans – a review. *Gene* 115:201-211.

- Schmid D, Baici A, Gehring H, Christen P (1994) Kinetics of molecular chaperone action. *Science* 263:971-973.
- Schorey JS, Carroll MC, Brown EJ (1997) A macrophage invasion mechanism of pathogenic mycobacteria. *Science* 277:1091-1093.
- Selwyn PA, Hartel D, Lewis VA, Schoenbaum EE, Vermund SH, Klein RS, Walker AT, Friedland GH (1989) A prospective study of the risk of tuberculosis among intravenous drug abusers with human immunodeficiency virus infection. *N. Engl. J. Med.* 320:545-550.
- Shimamoto T, Shimada M, Inouye M, Inouye S (1995) The role of ribonuclease H in multicopy single-stranded DNA synthesis in retron-Ec73 and retron-Ec107 of *Escherichia coli* J. *Bacteriol.* 177:264-267.
- Singh NB, Gupta HP, Srivastava A, Kandpal H, Srivastava UM (1997) Lymphostimulatory and delayed-type hypersensitivity responses to a candidate leprosy vaccine strain: *Mycobacterium habana*. *Lepr. Rev.* 68:125-130.
- Sinha RK, Khuller GK (1997) The protective efficacy of a liposomal encapsulated 30 kDa secretory protein of *Mycobacterium tuberculosis* H37Ra against tuberculosis in mice. *Immunol. Cell Biol.* 75:461-466.
- Skinner MA, Yuan S, Prestidge R, Chuk D, Watson JD, Tan PL (1997) Immunisation with heat-killed *Mycobacterium vaccae* stimulates CD8+ cytotoxic T cells specific for macrophages infected with *Mycobacterium tuberculosis*. *Infect. Immun.* 65:4525-4530.
- Smith DR, Richterich P, Rubenfield M, Rice PW, Butler C, Lee H-M, Kirst S, Gundersen K, Abendschan K, Xu Q, Chung M, Deloughery C, Aldredge T, Maher J, Lundstrom R, Tulig C, Falls K, Imrich J, Torrey D, Engelstein M, Breton G, Madan D, Nietupski R, Seitz B, Connelly S, McDougall S, Safer H, Gibson R, Doucette L, Eiglmeier K, Bergh S, Cole ST, Robison K, Richterich L, Johnson J, Church GM, Mao J (1997) Multiplex sequencing of 1.5 Mb of the *Mycobacterium leprae* genome. *Genome Research* 7:802-819.
- Snapper SB, Lugosi L, Jekkel A, Melton RE, Kieser T, Bloom BR, Jacobs WR Jr (1988) Lysogeny and transformation in the mycobacteria: stable expression of foreign genes. *Proc. Natl. Acad. Sci. USA* 85:6987-6991.

- Snapper SB, Melton RE, Keiser T, Mustafa S, Jacobs WR Jr** (1990) Isolation and characterisation of high efficiency plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol. Microbiol.* 4:1911-1919.
- Southern EM** (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503-517.
- Spreadbury C, Alli OAT** (1998) The *Mycobacterium tuberculosis* homolog of the *M. avium* *mig* gene is not specifically expressed in the macrophage. TB:Molecular Mechanisms and Immunologic Aspects; Keystone, CO; April 2-8.
- Stahl SJ, Kaufman JD, Vikić-Topić S, Crouch RJ, Wingfield PT** (1994) Construction of an enzymatically active ribonuclease H domain of human immunodeficiency virus type 1 reverse transcriptase. *Protein Eng* 7:103-1108.
- Stolt P, W Zillig** (1993) Antisense RNA mediates transcriptional processing in an archaebacterium, indicating a novel kind of RNase activity. *Mol. Microbiol.* 7: 875-882.
- Ströhl WR** (1992) Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. *Nucleic Acids Res.* 20:961-974.
- Styblo K** (1991). Epidemiology of tuberculosis. *Selected Papers R. Netherlands Tuberc. Assoc.* 24:1-136.
- Sudre P, ten Dam G, Kocki A** (1992) Tuberculosis: a global overview of the situation today. *Bull WHO* 70:149-159.
- Supply P, Magdalena J, Himpens S, Locht C** (1997) Identification of novel intergenic repetitive units in a mycobacterial two-component system operon. *Mol. Microbiol.* 26:991-1003.
- Takiff HE, Salazar L, Guerrero C, Phillip W, Huang WM, Kreiswirth B, Cole ST, Jacobs WR Jr, Telenti A** (1994) Cloning and nucleotide sequence of the *Mycobacterium tuberculosis gyrA* and *gyrB* genes and the detection of quinolone resistance mechanisms. *Antimicrob. Agents Chemother.* 38:773-780.
- Talbot EA, Perkins MD, Fagundes S, Silva M, Frothingham R** (1997) Disseminated bacille Calmette-Guérin disease after vaccination: case report and review. *Clin. Infect. Dis.* 24:1139-1146.
- Tanese N, Goff SP** (1988) Domain structure of the Moloney murine leukemia virus reverse transcriptase: mutational analysis and separate expression of the DNA polymerase and RNase H activities. *Proc. Natl. Acad. Sci. USA* 85:1777-1781.

- Telenti A, Phillip WJ, Sreevatsan S, Bernasconi C, Stockbauer KE, Wiele B, Musser JM, Jacobs WR Jr (1997a) The *emb* operon, a gene cluster of *Mycobacterium tuberculosis* involved in resistance to ethambutol. *Nat. Med.* 3:567-570.
- Telenti A, Hunore N, Bernasconi C, March J, Ortega A, Heym B, Takiff HE, Cole ST (1997b) Genotypic assessment of isoniazid and rifampin resistance in *Mycobacterium tuberculosis*: a blind study at reference laboratory level. *J. Clin. Microbiol.* 35:719-723.
- Thangaraj HS, Lamb FI, Davis EO, Jenner PJ, Jeyakumar LH, Colston MJ (1990) Identification, sequencing, and expression of *Mycobacterium leprae* superoxide dismutase, a major antigen. *Infect. Immun.* 58:1937-1942.
- Timm J, Van Rompaey I, Tricot C, Massae M, Haeseleer F, Fauconnier A, Stalon V, Bollen A, Jacobs P (1992) Molecular cloning, characterization and purification of ornithine carbamoyltransferase from *Mycobacterium bovis* BCG. *Mol. Gen. Genet.* 234:475-480.
- Timm J, Lim EM, Gicquel B (1994a) *Escherichia coli* - mycobacteria shuttle vectors for operon and gene fusions to *lacZ*: the pJEM series. *J. Bacteriol.* 176:6749-6753.
- Timm J, Perilli MG, Duez C, Trias J, Orefici G, Fattorini L, Amicosante G, Oratore A, Joris B, Frere JM, Pugsley AP, Gicquel B (1994b) Transcription and expression analysis, using *lacZ* and *phoA* gene fusions, of *Mycobacterium fortuitum* β -lactamase genes cloned from a natural isolate and a high level β -lactamase producer. *Mol. Microbiol.* 12:491-504.
- Tinoco I Jr, Borer PN, Dengler B, Levin MD, Uhlenbeck OC, Crothers DM, Bralla J (1973) Improved estimation of secondary structure in ribonucleic acids. *Nat. New Biol.* 246:40-41.
- Toida I (1993) HIV-infection, AIDS and BCG vaccination. *Kekkaku* 68:435-444.
- Toossi Z, Nicolacakis K, Xia L, Ferrari NA, Rich EA (1997) Activation of latent HIV-1 by *Mycobacterium tuberculosis* and its purified protein derivative in alveolar macrophages from HIV-infected individuals *in vitro*. *J. Acquir. Immune Defic. Syndr. Hum. Retrovirol.* 15:325-331.

- Torrey TA, Athlung T, Kogoma T (1984)** *dnaA* suppressor (*dasF*) mutants of *Escherichia coli* are stable DNA replication (*sdrA/rnh*) mutants. *Mol. Gen. Genet.* **196**:350-355.
- Tuberculosis Prevention Trial, Madras (1980)** Trial of BCG vaccines in South India for tuberculosis prevention. *Indian J. Med. Res.* **72**:1-74.
- Uchiyama Y, Iwai S, Ueno Y, Ikehara M, Ohtsuka E (1994a)** Role of the Mg^{2+} ion in the *Escherichia coli* ribonuclease HI reaction. *J. Biochem (Tokyo)* **116**:1322-1329.
- Uchiyama Y, Miura Y, Inoue H, Ohtsuka E, Ueno Y, Ikehara M, Iwai S (1994b)** Studies of the interactions between *Escherichia coli* RNase HI and its substrate. *J. Mol. Biol.* **243**:782-791.
- Udou T (1994)** Extracellular hemolytic activity in rapidly growing mycobacteria. *Can. J. Microbiol.* **40**:318-321.
- Venisse A, Fournie JJ, Puzo G (1995)** Mannosylated lipoarabinomannan interacts with phagocytes. *Eur. J. Biochem.* **231**:440-447.
- Verma A, Kinger AK, Tyagi JS (1994)** Functional analysis of transcription of the *Mycobacterium tuberculosis* 16S rDNA-encoding gene. *Gene* **148**:113-118.
- Villar C, Benítez J (1994)** Perspectives and limitations of the mycobacterial expression system. *Arch. Med. Res.* **25**:451-453.
- von Meyenburg K, Boye E, Skarstad K, Koppes L, Kogoma T (1987)** Mode of initiation of constitutive stable DNA replication in RNase H-defective mutants of *Escherichia coli* K-12. *J. Bacteriol.* **169**:2650-2658.
- Wards BJ, Collins DM (1996)** Electroporation at elevated temperatures substantially improves transformation efficiency of slow-growing mycobacteria. *FEMS Microbiol. Lett.* **145**:101-105.
- Wawrzynów A, Zylicz M (1995)** Divergent effects of ATP on the binding of the DnaK and DnaJ chaperones to each other, or to their various native and denatured protein substrates. *J. Biol. Chem.* **270**:19300-19306.
- Wawrzynów A, Banecki B, Wall D, Liberek K, Georgopoulos C, Zylicz M (1995)** ATP hydrolysis is required for the DnaJ-dependent activation of DnaK chaperone for binding to both native and denatured protein substrates. *J. Biol. Chem.* **270**:19307-19311.

- Wheeler FR, Radledge C** (1994) Metabolism of *Mycobacterium tuberculosis*. In: *Tuberculosis: pathogenesis, protection, and control*. (ed. Bloom BR) Washington, DC: ASM Press. pp. 353-358.
- Wickner SJ** (1990) Three *Escherichia coli* heat shock proteins are required for P1 plasmid DNA replication: formation of an active complex between *E. coli* DnaJ protein and the P1 initiator protein. *Proc. Natl. Acad. Sci. USA* 87:2690-2694.
- Wilson TM, Collins DM** (1996) *ahpC*, a gene involved in isoniazid resistance of the *Mycobacterium tuberculosis* complex. *Mol. Microbiol.* 19:1025-1034.
- Wilson TM, de Lisle GW, Collins DM** (1995) Effect of *inhA* and *katG* on isoniazid resistance and virulence of *Mycobacterium bovis*. *Mol. Microbiol.* 15:1009-1015.
- Winder EG, Rooney SA** (1970) Effects of nitrogenous components of the medium on the carbohydrate and nucleic acid content of *Mycobacterium tuberculosis* BCG. *J. Gen. Microbiol.* 63:29-39.
- Wintersberger U** (1990) Ribonucleases H of retroviral and cellular origin. *Pharmacol. Ther.* 48:259-280.
- Yang F, Curran SC, Li L, Avarbock D, Graf JD, Chua M-M, Lu G, Salem J, Rubin H** (1997) Characterization of two genes encoding the *Mycobacterium tuberculosis* ribonucleotide reductase small subunit. *J. Bacteriol.* 179:6408-6415.
- Yang F, Lu G, Rubin H** (1996) Cloning, expression, purification and characterization of DNA topoisomerase I of *Mycobacterium tuberculosis*. *Gene* 178:63-69.
- Yang W, Steitz TA** (1995) Recombining the structure of HIV integrase, RuvC and RNase H. *Structure* 3:131-134.
- Yang W, Hendrickson WA, Crouch RJ, Satow Y** (1990) Structure of ribonuclease H phased at 2 Å resolution by MAD analysis of the selenomethionyl protein. *Science* 249:1398-1405.
- Yanisch-Perron C, Vieira J, Messing J** (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and pUC19 vectors. *Gene* 33:103-119.
- Yie Y, Wei Z, Tien P** (1993) A simplified and reliable protocol for plasmid DNA sequencing: fast miniprep and denaturation. *Nucleic Acids Res.* 21:361.

- Young DB, Kaufmann SHE, Hermans PWM, Thole JER (1992)** Mycobacterial protein antigens: a compilation. *Mol. Microbiol.* 6:133-145.
- Zhang Y-B, Ayalew S, Lacks SA (1997)** The *rnhB* gene encoding RNase HII of *Streptococcus pneumoniae* and evidence of conserved motifs in eucaryotic genes. *J. Bacteriol.* 179:3828-3836.
- Zhu X, Venkataprasad N, Ivanyi J, Vordermeier HM (1997)** Vaccination with recombinant vaccinia viruses protects mice against *Mycobacterium tuberculosis* infection. *Immunology* 92:429-431.
- Ziemienowicz A, Skowrya D, Zeilstra-Ryalls J, Fayet O, Georgopoulos C, Zylicz M (1993)** Both the *Escherichia coli* chaperone systems, GroEL/GroES and DnaK/DnaJ/GrpE, can reactivate heat-treated RNA polymerase. Different mechanisms for the same activity. *J. Biol. Chem.* 268:25425-25431.
- Zurenko GE, Yagi BH, Schaadt RD, Allison JW, Kilburn JO, Glickman SE, Hutchinson DK, Barbachyn MR, Brickner SJ (1996)** *In vitro* activities of U-100592 and U-100766, novel oxalidinone antibacterial agents. *Antimicrob. Agents Chemother.* 40:839-845.

Author: Dawes, Stephanie Susan.

Name of thesis: The gene cloning and biochemical and genetic analysis of the ribonuclease HI from Mycobacterium smegmatis.

PUBLISHER:

University of the Witwatersrand, Johannesburg

©2015

LEGALNOTICES:

Copyright Notice: All materials on the University of the Witwatersrand, Johannesburg Library website are protected by South African copyright law and may not be distributed, transmitted, displayed or otherwise published in any format, without the prior written permission of the copyright owner.

Disclaimer and Terms of Use: Provided that you maintain all copyright and other notices contained therein, you may download material (one machine readable copy and one print copy per page) for your personal and/or educational non-commercial use only.

The University of the Witwatersrand, Johannesburg, is not responsible for any errors or omissions and excludes any and all liability for any errors in or omissions from the information on the Library website.