Molecular and biochemical characterisation of key enzymes involved in mycolic acid biosynthesis from *Mycobacterium tuberculosis*

A thesis submitted to the
University of Newcastle upon Tyne
for the degree of Doctor of Philosophy

August 2004

Alistair K. Brown, B.Sc. (Hons)

Abstract

Mycolic acids are the dominant feature of the *Mycobacterium tuberculosis* cell wall, providing the basis for its lipid-rich permeability barrier. These α-alkyl, β-hydroxy fatty acids are thought to be formed by the Claisen-type condensation of a long C_{56} meromycolic acid and a shorter C_{24}-C_{26} fatty acid. These component fatty acids are produced *via* a combination of type I and II fatty acid synthase (FAS) systems. The C_{16}-C_{26} fatty acyl products of FAS-I are elongated by FAS-II with simultaneous modification to form meromycolic acids, which are then condensed with the C_{24}-C_{26} fatty acyl chain. These studies aimed to characterise key enzymes of FAS-II (mtFabH, KasA) and enzymes possibly involved in the Claisen-type condensation reaction to form mycolatic acids (AccD enzymes, Pks13, FadD32).

The β-ketoacyl ACP synthase (KAS) III (mtFabH) is proposed to link FAS-I and FAS-II, catalyzing the condensation of FAS-I-derived acyl-CoA with malonyl-Acyl Carrier Protein (ACP). The acyl-CoA chain length specificity of mtFabH was assessed *in vitro* . When using *E. coli,* the preferred substrates were C_{12}- and C_{14}-CoA. However, with the mycobacterial ACP (AcpM), the enzyme was able to utilise longer (up tp C_{20}) acyl-CoA chains. The substitution of residues implicated in acyl-CoA chain length specificity totally abrogated overall KAS activity and reduced the transacylation activity of the enzyme. Mutation of the proposed catalytic triad residues confirmed that Cys122 is essential for transacylation and His258 is essential for malonyl-AcpM decarboxylation.

KasA, which belongs to the FAS-II system, utilises palmitoyl-ACP rather than short-chain acyl-ACP primers. Purified recombinant KasA had *in vitro* KAS activity that was highly sensitive to cerulenin, a well-known KAS inhibitor. Mutation of proposed catalytic residues Cys171, His311, Lys340 and His345 inactivated the enzyme completely.

Four putative *accD* genes were found in *Corynebacterium glutamicum*. Overexpression of each gene resulted in increased acyl-CoA dependent $^{14}CO_2$ fixation *in vitro*, providing evidence that the *accD* genes encode a family of carboxyltransferases. Disruption of either *accD2* or *accD3* led to complete and specific loss of mycolic acids. These two carboxyltransferases are also retained in all Corynebacterianeae, including *M. leprae*, and probably provide a carboxylated intermediate for condensation of the mero-chain and α-branch directed by the *pks13*-encoded polyketide synthase.

Declaration

The work recorded in this thesis was carried out in the School of Cell and Molecular Biosciences at the University of Newcastle upon Tyne, Newcastle upon Tyne, U.K. during the period October 2000 to May 2002 and in the School of Biosciences at the University of Birmingham, Birmingahm, U. K. during the period April 2002 to September 2003. The work in this thesis is original except where acknowledged by reference.

No portion of the work is being, or has been submitted for a degree, diploma or any other qualification at any other university.

Acknowledgements

Through my four years as a PhD student a number of people played influential roles not only with my work but life in general. I must say I have thoroughly enjoyed my time as a PhD student in Newcastle and in Birmingham, shame we had to move from the Toon. Firstly I would like to thank my supervisor Professor Gurdyal S. Besra for offering me my PhD and for his support and guidance throughout my four years.

Secondly I would like to thank Drs. Laurent Kremer, William N. Maughan, Madhavan Nampoothiri and Sudagar Gurcha for their guidance, patience and advice, and also all my other lab colleagues past and present at both Newcastle and Birmingham for the joy and individuality that they brought to such enjoyable places to work.

I would also like to thank my family for their support throughout the many years that I've been a student and also to my long suffering girlfriend Jaqui who has kept me sane and put up with me for so long.

Lastly but most importantly to Dr Lynn G. Dover, even though you're a makem, I could not have done this without your support and guidance. Over the past four years we've become good friends and some of the best memories that I will take away with me involve you in some way or another. To mention a few, your wedding, Washington D.C., not C.D., and our many excursions to local boozers wherever we are! Your humour, good nature and generosity have inspired me even when everything was going wrong. You helped me through difficult times in my life and words can't express my gratitude, so thanks again Lynn you're a star.

Contents

List of Figures

List of Tables

Abbreviations

αα	amino Acids
A	adenine
ACP	acyl carrier protein
AG	arabinogalactan
AIDS	Acquired Immuno Deficiency Syndrome
APS	Ammonium persulphate
Ara*f*	arabinose in the furanose ring form
bp	base pairs
BSA	bovine serum albumin
C	cytosine
C_n	Carbon, where n = number of
CoA	Coenzyme A (pantothenic acid)
CDC	Centre for Disease Control
CER	cerulenin
CPM	counts *per* minute
°C	degrees
DAT	diacyl trehalose
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DMSO	dimethyl sulfoxide
DOTS	directly odserved therapy
DPA	Decaprenyl-monophosphoprenyl-Ara*f*
DPM	decays *per* minute
DTT	dithiothreitol
EMB	Ethambutol
ETH	Ethionamide
FAD	flavin adenine dinucleotide
FAME	fatty acid methyl esters
FAS	fatty acid synthase
G	guanine
g	grams
Gal	galactose
GlcNAc	N-acetylglucosamine
hr	hour(s)
His-tag	6 Histidine residue tag
HIV	human immunodeficiency syndrome
IPTG	isopropyl-β-thiogalactopyranoside
IC_{50}	50 % inhibition concentration
INH	Isoniazid
Kas	β-Ketoacyl ACP synthase
KAN	kanamycin
kb	kilobase pairs
kDa	kilodalton
KO	knock out
LAM	lipoarabinomannan
Lau	Lauroyl
L	Litre

LB	Luria-Bertani
LU	Linker unit
M	Molar
M	*Mycobacterium*
mAGP	mycolyl arabinogalcatan peptidoglycan
MAME	mycolic acid methyl esters
Mb	mega bases
MCAT	malonyl-CoA : ACP transacylase
ml	millilitre
mM	millimolar
MDR	multi-drug resistant
mt	*Mycobacterium tuberculosis*
µg	microgram
µl	micolitre
µM	micromolar
µm	micrometre
mAG	mycoyl arabinogalactan
MIC	Minimum inhibition concentrations
min	minute
MPI	mannosyl-phosphatidyl-*myo*-inositol
Mur	muramic acid
MurNAc	N-acetylmuramic acid
Myc-PL	6-*O*-mycolyl-β-D-mannopyranosyl-1-monophosphoheptaprenol
NAD^+	nicotinamide adenine dinucleotide
NADH	reduced nicotinamide adenine dinucleotide
$NADP^+$	nicotinamide adenine dinucleotide phosphate
NADPH	reduced nicotinamide adenine dinucleotide phosphate
nm	nanometres
ORF	Open reading frame
OD	Optical density
p	pyranose form
Pal	Palmitoyl
PAGE	polyacrilamide gel electropheresis
PAT	pentaacyl trehalose
PCR	polymerase chain reaction
PDIM	phthiocerol dimycocerosate
PG	peptidoglycan
PIM_6	phosphoinositol hexamannoside
PKS	Polyketide synthase
PPM	polyprenolmonophosporylmannose
PZA	Pyrazinamide
Rha	Rhamnose
RIF	rifampin
RNA	Ribonucleic acid
rRNA	ribosomal Ribonucleic acid
rpm	revolutions *per* minute
s	seconds
SDS	sodium dodecyl sulphate
SL	Sulphated tetra-acyl trehalose
spp	species

T	thymine
TAT	triacyl trehalsoe
TDM	trehalose 6,6'-dimycolate
TEMED	N,N,N',N'-tetramethylethylenendiamine
TLC	thin layer chromotography
TLM	thiolactomycin
TMM	trehalose monomycolate
U	Uracil
UDP	uridine diphosphate
v/v	volume/volume
WHO	World Health Organisation
w/v	weight *per* volume
x g	times gravity

Published work associated with this thesis

Kremer, L., Dover, L.G., Carrere, S., Nampoothiri, K. M., Lesjean, S., Brown, A. K., Brennan, P. J., Minnikin, D. E., Locht, C., & Besra, G. S. (2002) Mycolic acid biosynthesis and enzymic characterization of the beta-ketoacyl-ACP synthase A-condensing enzyme from *Mycobacterium tuberculosis*. *Biochem J* **364,** 423-30.

Gande, R., Gibson, K. J., Brown, A. K., Krumbach, K., Dover, L. G., Sahm, H., Shioyama, S., Oikawa, T., Besra, G. S., & Eggeling, L. (2004) Acyl-CoA carboxylases (*accD2* and *accD3*) together with a unique polyketide synthase (Cg-pks) are key to mycolic acid biosynthesis in Corynebacterianeae like *Corynebacterium glutamicum* and *Mycobacterium tuberculosis*. *J Biol Chem*. **279 (43),** 44847-57.

Brown, A. K., Kremer, L., Dover, L. G., & Besra, G. S. (2004) Probing the mechanism of action of the *Mycobacterium tuberculosis* β-ketoacyl-ACP synthase III *mt*FabH: identification of key residues involved in catalysis, substrate specificity and interaction with the acyl carrier protein. *J Biol Chem*. (*Submitted*)

CHAPTER 1

INTRODUCTION

1.1 General introduction

In terms of infectious disease *Mycobacterium tuberculosis*, as the cause of tuberculosis (TB), is the leading cause of morbidity and mortality worldwide accounting for 26 % of all preventable adult deaths globally (Ramaswamy & Musser, 1998). Approximately 8 million infections and 2 million deaths are recorded each year (WHO, 2003). Chemotherapy, mainly through the use of isoniazid (INH) and rifampin (RIF), coupled with the use of the BCG vaccine has been the method of choice to combat the increasing TB problem. Clinically isolated strains of *M. tuberculosis* resistant to specific antimycobacterial agents are recovered readily from both immuno-competent and immuno-compromised patients. More recently, multi-drug resistant strains of *M. tuberculosis* (MDR-TB) have emerged, severely hindering efforts towards prevention and control. Increases in the median rates of primary resistance to front-line drugs have been observed, with rates for INH , streptomycin, RIF and ethambutol (EMB) being recorded as 4.1 %, 3.5 %, 0.2 % and 0.1 % respectively (Cohn *et al.*, 1997). The median rates of acquired secondary resistance were observed to be much higher than primary resistance.

The current treatment procedure for *M. tuberculosis* uses an initial 2-month period of daily therapy with INH, RIF, and pyrazinamide (PZA), plus either streptomycin or EMB. During this period sputa tend to be culture-negative. This regimen is then followed by daily therapy with INH and one other primary drug for the next 4 months, resulting with successful therapy in clearing of shadows on chest X-rays. Ethionamide (ETH), cycloserine, *p*-amino-salicylic acid, thioacetazone, kanamycin, capreomycin, viomycin, amikacin and fluoroquinolones are used as secondary or alternative agents to treat TB infections (Stratton & Reed, 1986).

New strategies and approaches are needed to fight against TB. The elucidation of the *M. tuberculosis* H37Rv and CDC1551 (Cole *et al.*, 1998; Fleischmann *et al.*, 2002) genome sequences and the development of new molecular techniques has accelerated research in different fields of investigation, such as:

- *New vaccines.*
- *New drug targets and drug development.*
- *Identification of correlates of protective immunity.*
- *Mechanisms of intracellular survival and pathogenesis.*
- *Genome-driven research.*

Bacilli that are considered to be members of the genus *Mycobacterium* have the following characteristics; aerobic, non-motile, non-sporeforming, rod-shaped and are 0.2-0.7 x 1.0-10.0 μm in size. Mycobacteria belong to the *Actinomycetale* family and, therefore, are closely related to nocardia, rhodococci and corynebacteria. One of the main characteristics of the genus *Mycobacterium* is a complex cell envelope, which is responsible for its acid-fast staining properties with the Zeihl-Neelsen stain (Figure 1.1).

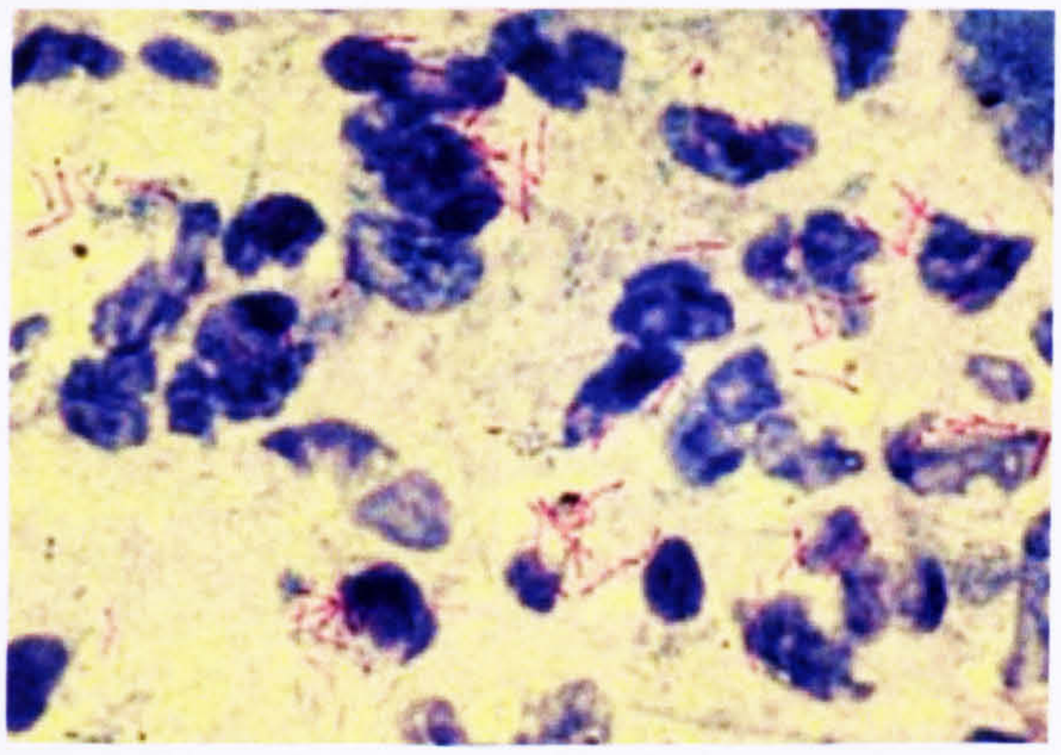


Figure 1.1 **Acid-fast *M. tuberculosis* (red-stained) visualised using the Zeihl-Neelsen stain**

Members of the *Mycobacterium* genus can be sub-grouped according to their growth characteristics i.e. fast-growing and slow-growing (Table 1.1). The slow-growers include most of the major human and animal pathogens, whereas the fast growers, other than *M. chelonae* which has been shown to be a turtle pathogen (Parish & Stoker, 1999) include mainly non-pathogenic species.

Table 1.1 *Mycobacterium spp.*

Slow-growers	Fast-growers
Mycobacterium africanum	*Mycobacterium aurum*
Mycobacterium asiaticum	*Mycobacterium chelonae*
Mycobacterium avium	*Mycobacterium duvali*
Mycobacterium bovis	*Mycobacterium flavescens*
Mycobacterium celatim	*Mycobacterium fortuitum complex*
Mycobacterium farcinogenes	*Mycobacterium gadium*
Mycobacterium gastri	*Mycobacterium gilvum*
Mycobacterium genavense	*Mycobacterium komossense*
Mycobacterium gordonae	*Mycobacterium neoaurum*
Mycobacterium haemophilum	*Mycobacterium phlei*
Mycobacterium interjectum	*Mycobacterium senegalense*
Mycobacterium intermedium	*Mycobacterium smegmatis*
Mycobacterium intracellulare	*Mycobacterium thermoresistibile*
Mycobacterium kansasii	
Mycobacterium leprae	
Mycobacterium malmoense	
Mycobacterium marinum	
Mycobacterium microti	
Mycobacterium nonchromogenicum	
Mycobacterium parafortuitum	
Mycobacterium scrofulaceum	
Mycobacterium shimoidei	
Mycobacterium simiae	
Mycobacterium szulgai	
Mycobacterium terrae	
Mycobacterium triviale	
Mycobacterium tuberculosis	
Mycobacterium ulcerans	
Mycobacterium xenopi	

1.2 Epidemiology of tuberculosis

M. tuberculosis has long been a major cause of mortality worldwide since the time of Robert Koch and no decrease has been observed in its prevalence since. The implementation of good sanitation and chemotherapy has drastically reduced TB in developed countries, but has had no major impact on the global problem (Bloom & Murray, 1992).

Estimations suggest that about a third of the world's population, i.e. approximately 1.7 billion people are infected with *M. tuberculosis* with approximately 0.5 million of these cases being children under 15 years of age (Enarson & Murray, 1996) (Figure 1.2). The lifetime risk of developing active TB is 7 %. Individuals suffering from acquired immune deficiency syndrome (AIDS) however, have an increased risk of 7 % *per* annum. Human immunodeficiency virus (HIV) infection is the greatest known risk factor for the progression of latent TB infection into active TB, the cause of 32 % of deaths among HIV-positive individuals (Narain *et al.*, 1992). The frequency of MDR-TB in HIV patients is increasing, problems in drug absorption, especially RIF, in these patients accelerating the emergence of drug resistance (Narain *et al.*, 1992).

Several studies indicate that resistance to various anti-tubercular agents results from alterations to chromosomal targets (Cole, 1994). Thus, MDR does not stem from the acquisition of transposable elements or plasmids like in other bacteria, but by mutations of the antibiotic target (Fang *et al.*, 1999) or pro-drug activating mechanisms. The inadequate prescription of chemotherapy and poor compliance to drug protocols has selected for

mutations. The advent of high incidence rates of HIV infection and the onset of AIDS, and the decline of socio-economic standards has contributed to the re-emergence of TB.

Current data suggests that, under prevailing treatment practices, it appears that MDR-TB will remain a localised problem, rather than becoming a global obstacle to tuberculosis control. (Dye & Espinal, 2001). These data also indicate that, at present, MDR-TB strains are not in the process of replacing drug-susceptible strains. Statistical analysis of data illustrate that the ratio of resistance to susceptibility is greatest in the cases of INH and streptomycin. RIF resistance is also strongly associated with MDR-TB as it is one of the front-line agents against TB. The conclusions of the study suggested that TB resistance to INH, or to INH plus RIF, is likely to remain low in most parts of the world, with trends showing that only 4 of the countries out of 25 studied, reported INH resistance. However, it is somewhat confusing that MDR-TB is prevalent in the Baltic States, India, Russia and China.

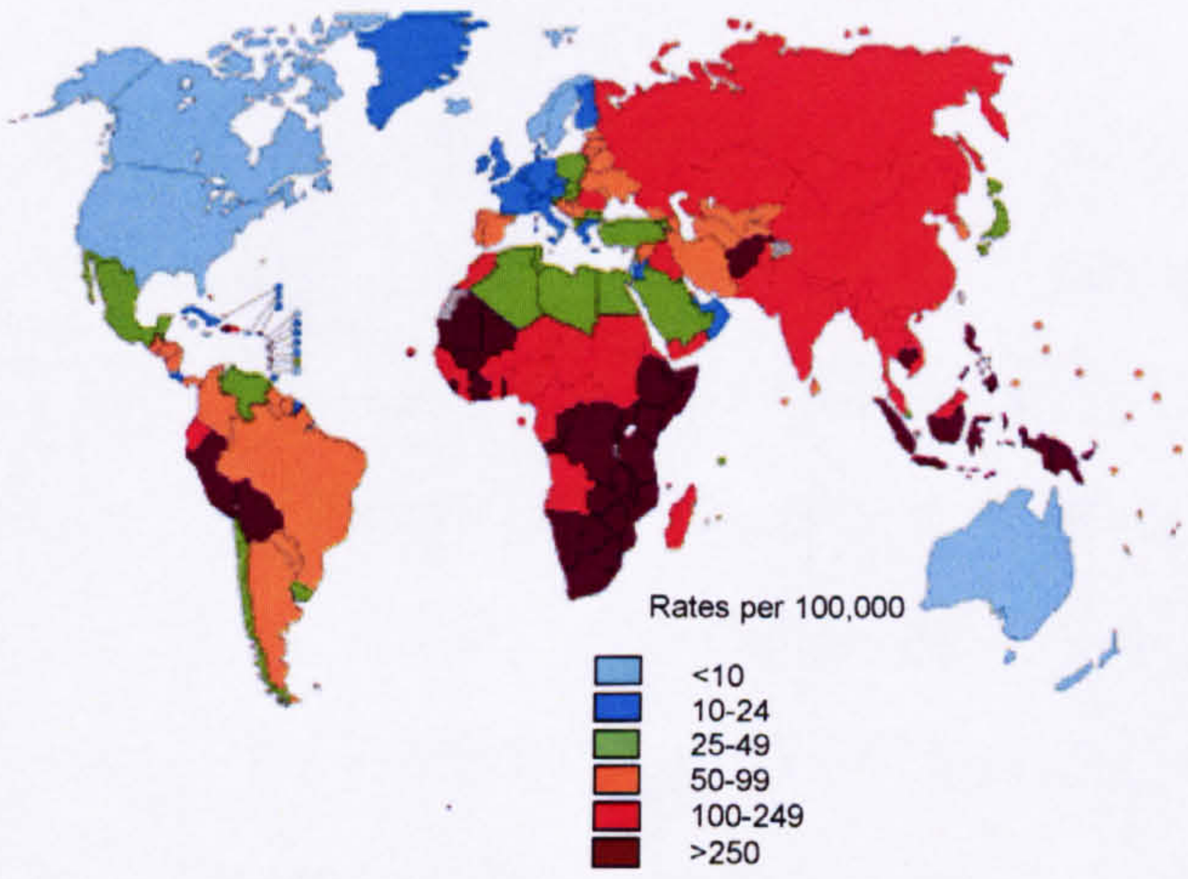


Figure 1.2 Global map showing the estimated new cases of TB in 1997, adapted from WHO, 1997

In the U.S. the rate of infection had been in constant decline, from 84,304 cases in 1953 to 22,225 in 1984. However, between 1985 and 1992 a dramatic increase in cases was reported. The reason for these increases have been linked to a number of possible factors, namely the HIV/AIDS epidemic, increased immigration from high-prevalence countries and the deterioration of the primary health care system to combat TB. In a later study the Centre for Disease Control (CDC) recorded 93,449 cases of TB through the period 1993-96 with 1,457 of these cases being identified as MDR-TB (CDC, 1995). Data from 1998 showed that 13 % of all new cases in the USA are resistant to front line drugs, with 1.6 % being resistant to both INH and RIF (CDC, 1995). Problems associated with diagnosis have been rectified by the implementation of BACTEC to quickly identify *M. tuberculosis* and the broad use of rapid methods for drug susceptibility testing leading to expanded use of preventive therapy and stronger infection control measures. Increased federal resources for state and local TB eradication programs have improved the situation, but problems still occur due to the introduction and spread of MDR-TB (CDC, 1995).

In India about 13 million people, equating to 1.5 % of the total population, suffer from active pulmonary tuberculosis, showing no change since independence. The highest rate of incidence was observed in 1991 with 1,555,353 new cases, a steady decline has been observed since (Figure 1.3) (Corbett *et al.*, 2003; WHO, 2003). The Revised National TB Control Program (RNTCP), configured by the Government of India in 1993 and implemented in 1997, introduced "directly observed therapy short course" (DOTS) and put TB control high on the public health agenda. This resulted in the allocation of more resources for TB control, improved laboratory diagnosis and the adoption of DOTS with standardised drug regimens and reporting methods. TB remains the leading cause of death in India, killing close to 500,000 people each year. India has about 2 million new cases of TB each year, far more than

any other country, and accounts for nearly one-third of the global burden of TB. The treatment success rate for patients registered in 2000 was 84 %, with the estimated percentage of adults 15-49 years old with HIV and TB being recorded as 4 % with a further 3.4 % of new cases being identified as MDR-TB (Dye *et al.*, 2002; WHO, 2003).

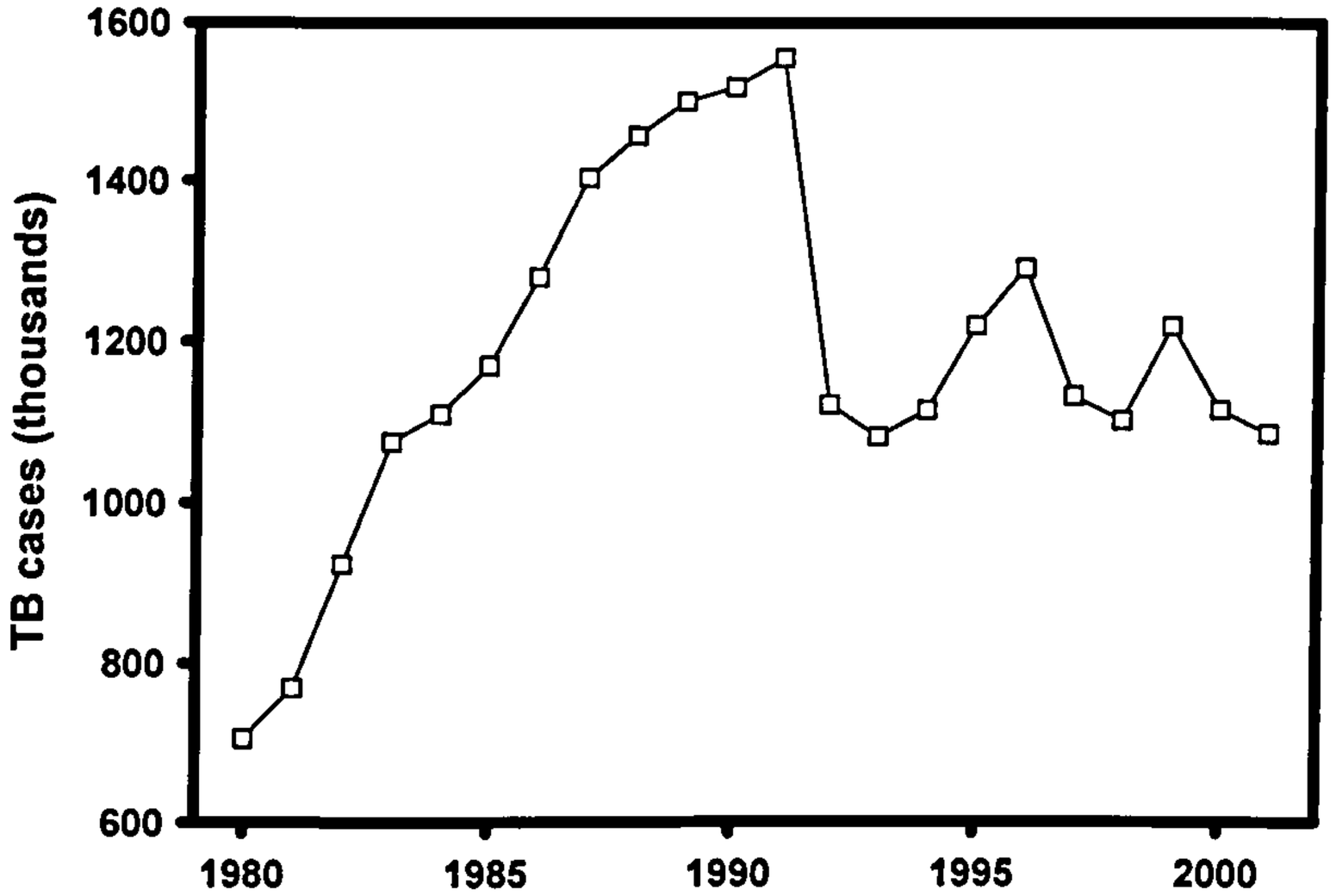


Figure 1.3 Incidence rates of new active pulmonary *M. tuberculosis* infections in India from 1980-2001 adapted from WHO, 2003

1.3 Genome project

Since its discovery in clinical isolates from New York's Trudeau Sanatorium in 1905, the *M. tuberculosis* H37Rv strain has been the strain of choice for TB researchers. This is primarily due to the fact that it has retained full virulence in animal models, is still susceptible to anti-TB agents and amenable to genetic manipulation (Cole *et al.*, 1998). Mycobacteria have been shown to contain a high GC content compared to that of other Gram-positive bacteria with *M. tuberculosis* H37Rv possessing a GC content of 65.6 % (Cole *et al.*, 1998; Hatfull, 1996;

Parish & Stoker, 1999). The *M. tuberculosis* H37Rv genome project was completed in 1998. The single circular chromosome consists of 4.4 Mb (4,411,529 bp), 4,049 genes with 3,924 open reading frames originally identified in the genome accounting for 91 % of its total potential coding capacity. By using various database comparative search techniques, 40 % of the proteins predicted have been assigned a function, with a further 44 % being assigned a putative function. It was shown that 16 % of the remaining genes encoded proteins of no known function or similarity to other proteins. One feature of the genome is the maintenance of a large number of genes apparently devoted to lipid metabolism; around 250 genes have been assigned thus, five-fold more than in *E. coli*. Such a high number is consistent with the diversity of lipids produced by *M. tuberculosis* ranging from simple palmitic acid and tuberculostearic acid, through to the more complex and characteristic long-chain mycolic acids. *M. tuberculosis* also contains a wide variety of polyketide synthases (PKS).

The elucidation of the genome sequence of *M. tuberculosis* H37Rv has provided new avenues for the next generation of TB research. Genomic data is being used for a diverse range of purposes including the deduction of gene function by bioinformatics and reverse genetics. Together they should continue to increase our understanding of *M. tuberculosis* pathogenicity and hopefully prime new leads for chemotherapy and immunoprophylaxis (Cole *et al.*, 1998).

The complete genome of the *M. tuberculosis* clinical strain CDC1551 is also available (Fleischmann *et al.*, 2002).

1.4 The mycobacterial cell wall

There is an unusually high content of lipids in mycobacteria, comprising approximately 60 % of the dry weight, much higher than in other Gram-positive (0.5 %) and Gram-negative (3 %)

bacteria (Smith, 1982). The cell envelopes of mycobacteria exhibit low-permeability due to the high proportion of lipids and associated complex lipids. These are organised as a physical barrier to a variety of agents. This permeability barrier is thought to play a crucial role in the intrinsic resistance of *M. tuberculosis* to antibiotics, chemical injury and lytic enzymes.

The first model of the mycobacterial cell wall incorporating an outer lipid bilyar was proposed by Minnikin (1982). This has been modified by McNeil and Brennan (1991) to incorporate complex lipids and further modified to incorporate novel insights into peptidoglycan (PG) organisation (Dmitriev *et al.*, 2000). Minnikin *et al.* (2002) combined all the theories as illustrated in Figure 1.4.

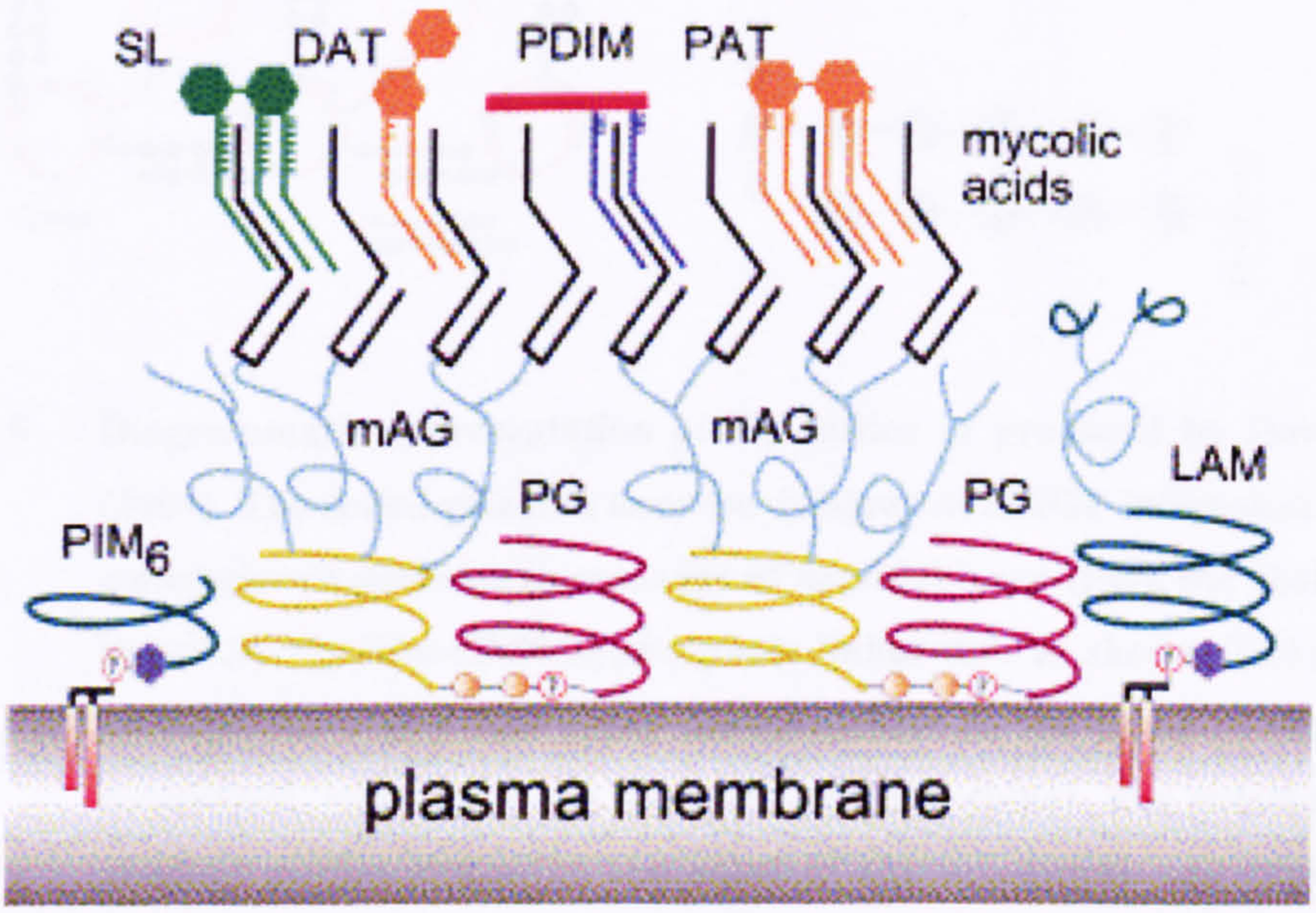


Figure 1.4 **The mycobacterial cell wall as proposed by Minnikin *et al.* (2002) modified from (Dmitriev *et al.*, 2000). Sulphated tetra-acyl trehalose (SL), diacyl trehalose (DAT), phthiocerol dimycocerosate (PDIM), pentaacyl trehalose (PAT), phosphoinositol hexamannoside (PIM_6), mycolyl arabinogalactan (mAG), peptidoglycan (PG) and lipoarabinomannan (LAM)**

In all of these models the cell bound mycolic acid residues which form the inner leaflet of the other membrane are esterified to a special branched arabinofuranose motif. This motif decorates the non-reducing end of a linear arabinan chain which is, in turn, linked to a linear galactan chain, together known as arabinogalactan. This whole assembly is attached to the glycan moiety of the peptidoglycan *via* a unique Rha-GlcNAc-phosphate linker. This covalenty linked structure is known as the mycolyl-arabinogalactan-peptidoglycan (mAGP) complex which has been the subject of much study in recent years.

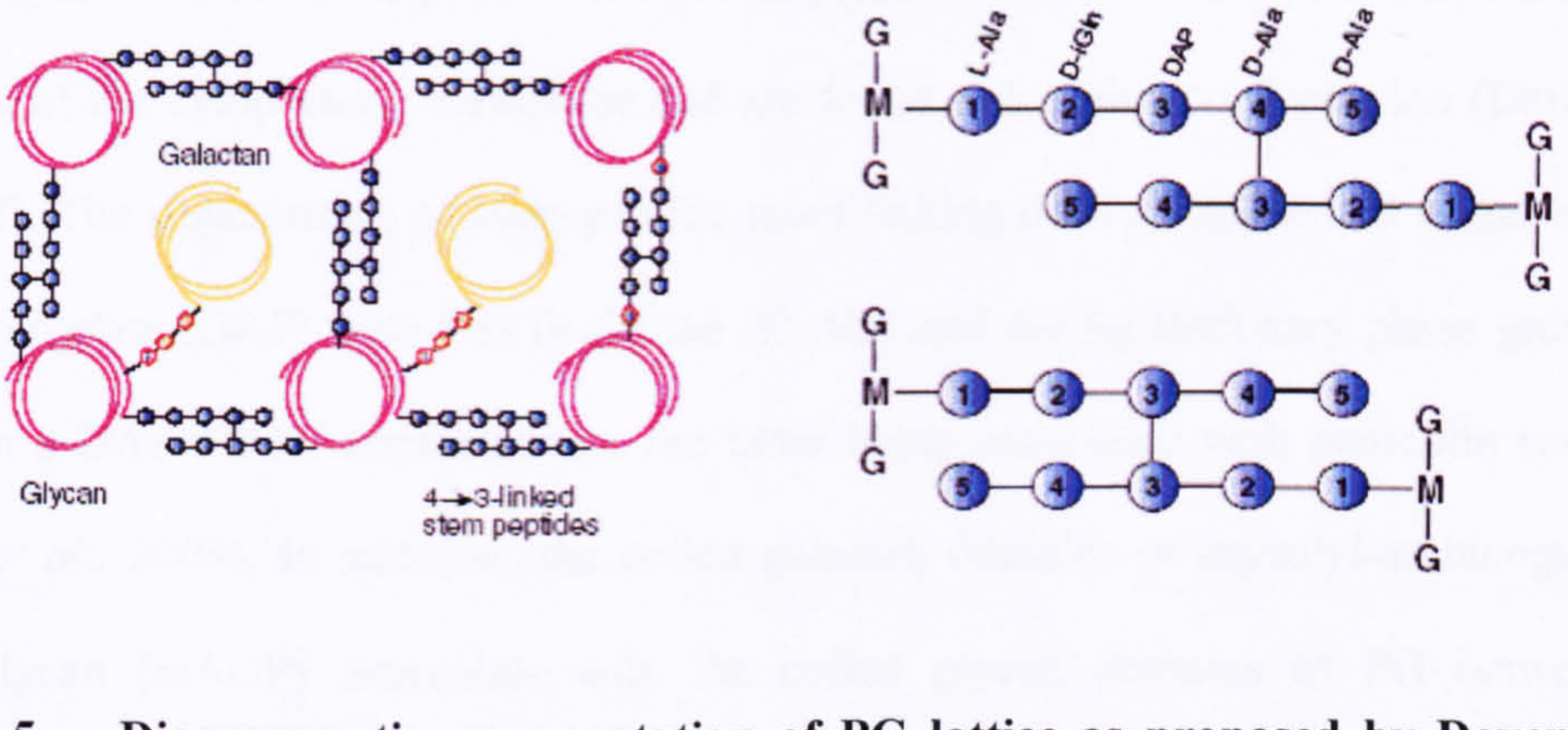


Figure 1.5 **Diagrammatic representation of PG lattice as proposed by Dover *et al.*, (2004). The coiled galactan domains (yellow) of mAGP intercalate with the coiled glycan domains (burgundy) of peptidoglycan (PG), the chains being linked by the Rha-GlcNAc-phosphate linker unit as shown. The glycan is arranged in a grid-like array perpendicular to the plane of the plasma membrane, cross-linked by the stem peptides (blue). The galactan is also oriented perpendicular to the plane of the membrane and between the stem peptide cross-links. The figure on the right represents the organisation of stem-peptide cross-linking during exponential phase (4→3 linked) (upper) and stationary phase (3→3 linked) (lower) growth, the latter being associated with penicillin resistance. G and M represent the N-acetylglucosamine and muramic acid residues of the glycan respectively. (L-Ala (L-alanine), D-iGln (D-glutamate), DAP (*meso*-diaminopimelate), D-Ala (D-alanine))**

The PG layer consists of a highly cross-linked polymer of amino acids and amino sugars. The glycan is thought to be arranged in a grid-like array perpendicular to the plane of the plasma membrane, cross-linked by the stem peptides. Earlier models depicting the structural organisation of the mycobacterial cell wall assume PG and galactan strands to run in parallel to the cytoplasmic membrane forming several horizontal layers beneath perpendicularly oriented mycolic acids (Dmitriev *et al.*, 2000). Dmitriev *et al.* (2000) re-evaluated the current chemical, biochemical and electron microscopical data and proposed a fundamentally distinct principle of the physical organisation and biosynthesis of the mycobacterial cell wall skeleton. According to this new concept, the cross-linked glycan strands which all run perpendicular to the plane of the cytoplasmic membrane and are found in a coiled conformation (Dmitriev *et al.*, 2000). The organisation of stem-peptide cross-linking during exponential phase is *meso*-diaminopimelate (DAP) linked to D-alanine (D-Ala) and during stationary phase growth the link is in a DAP→DAP conformation, the latter being associated with penicillin resistance (Dover *et al.*, 2004). In addition, the coiled galactan domains of mycolyl-arabinogalactan-peptidoglycan (mAGP) intercalate with the coiled glycan domains of PG between the stempeptide cross-links, the chains being linked by the Rha-GlcNAc-phosphate linker unit (Dmitriev *et al.*, 2000). Mycolic acids are esterfied to the terminal arabinose motif of arabinogalactan (AG). This structural entity is known as the mAGP complex. An assortment of surface glycolipids complement the mAGP complex to provide an outer lipid bilayer, this configuration as mentioned earlier has been speculated to provide a protective barrier to drug permeation (McNeil & Brennan, 1991). The mycobacterial cell wall also contains a second unique molecule, termed lipoarabinomannan (LAM). It is speculated that LAM is linked to the interaction of the pathogen with the host, possibly leading to resistance mechanisms associated with macrophage killing.

1.4.1 Peptidoglycan

From 1950 to 1970 there was a spate of investigations into the mAGP complex of *M. tuberculosis,* but due to insolubility problems work was hampered and structural definition didn't arise until the 1980s when Petit & Lederer (1984) predicted mycobacterial PG to consist of alternating units of N-acetylglucosamine (GlcNAc) and modified muramic acid (Mur) residues. The accepted structure of the PG unit of the cell wall of *M. tuberculosis* is illustrated in Figure 1.6, N-glycoyl-MurNAc is linked to the 4-position of GlcNAc and to the stem polypeptide, which participates in PG cross-linking (Goffin & Ghuysen, 2002). The stem polypeptide consists of L-alanine (L-Ala), D-glutamate (D-iGln), *meso*-diaminopimelate (DAP) and finally two D-alanine (D-Ala) residues. The basic pathway for the biosynthesis of PG is similar in most bacteria with all the major enzymes being represented in the *M. tuberculosis* H37Rv genome Figure 1.7 and Table 1.2. In *E. coli* the pathway is encoded by two major clusters, designated *mra* and *mrb*. In *M. tuberculosis* the *mra* cluster also contains genes thought to be involved in the synthesis of other aspects of PG biosynthesis. These genes encode high and low molecular weight penicillin binding proteins (PBPs) (van Heijenoort, 2001). The immediate precursor to polymerised PG in *E. coli* is Lipid II, a polyisoprenoid-bound GlcNAc-MurNgly disaccharide (Higashi *et al.*, 1967).

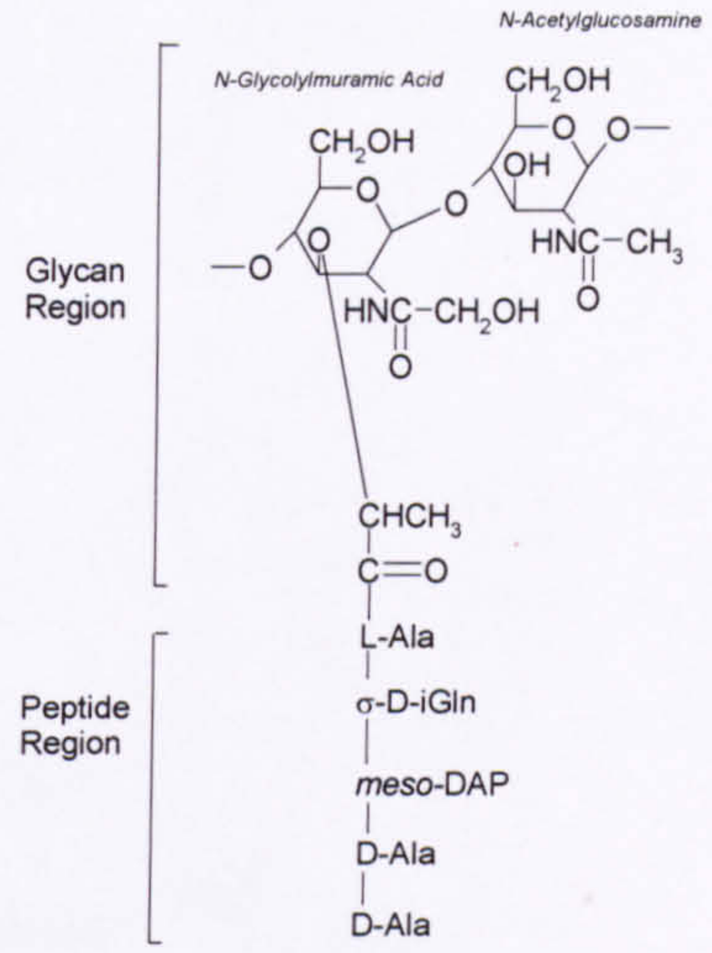


Figure 1.6 **Structure of the basic PG unit of the mycobacterial cell wall. A singular disaccharide subunit of PG with N-acetylglucosamine bound *via* C4 to N-glycolymuramic acid and subsequently *via* its C3 to the stem polypeptide, which participates in PG cross-linking. The stem polypeptide consists of a L-alanine (L-Ala), D-glutamate (σ-D-iGln), meso-diaminopimelate (DAP) and finally two D-alanine (D-Ala) residue.**

Table 1.2 **PG biosynthetic enzymes found in *M. tuberculosis* (Belanger & Inamine, 2000)**

Known Genes	Function	*M. tuberculosis* gene
alr	Alanine racemase	Rv3423c
ddl	D-Alanine-D-alanine ligase	Rv2981c
murI	Glutamate racemase	Rv1338
glmU	GlcNAc-1-P uridyltransferase	Rv1018c
murA (murZ)	UDP-GlcNAc enoylpyruvyl transferase	Rv1315
murB	UDP-GlcNAc enoylpyruvy reductase	Rv0482
murC	L-Alanine-adding enzyme	Rv2152c
murD	D-Glutamate-adding enzyme	Rv2155c
murE	DAP-adding enzyme	Rv2158c
murF	D-Alanine-D-alanine-adding enzyme	Rv2157c
murX	Translocase	Rv2156c
murG	Transferase	Rv2153c
ponA	Transglycosylase/transpeptidase	Rv0050
ponA'	Transglycosylase/transpeptidase	Rv3682
pbpA	Transglycosylase/transpeptidase	Rv0016c
pbpB::ftsI	Transglycosylase/transpeptidase	Rv2163

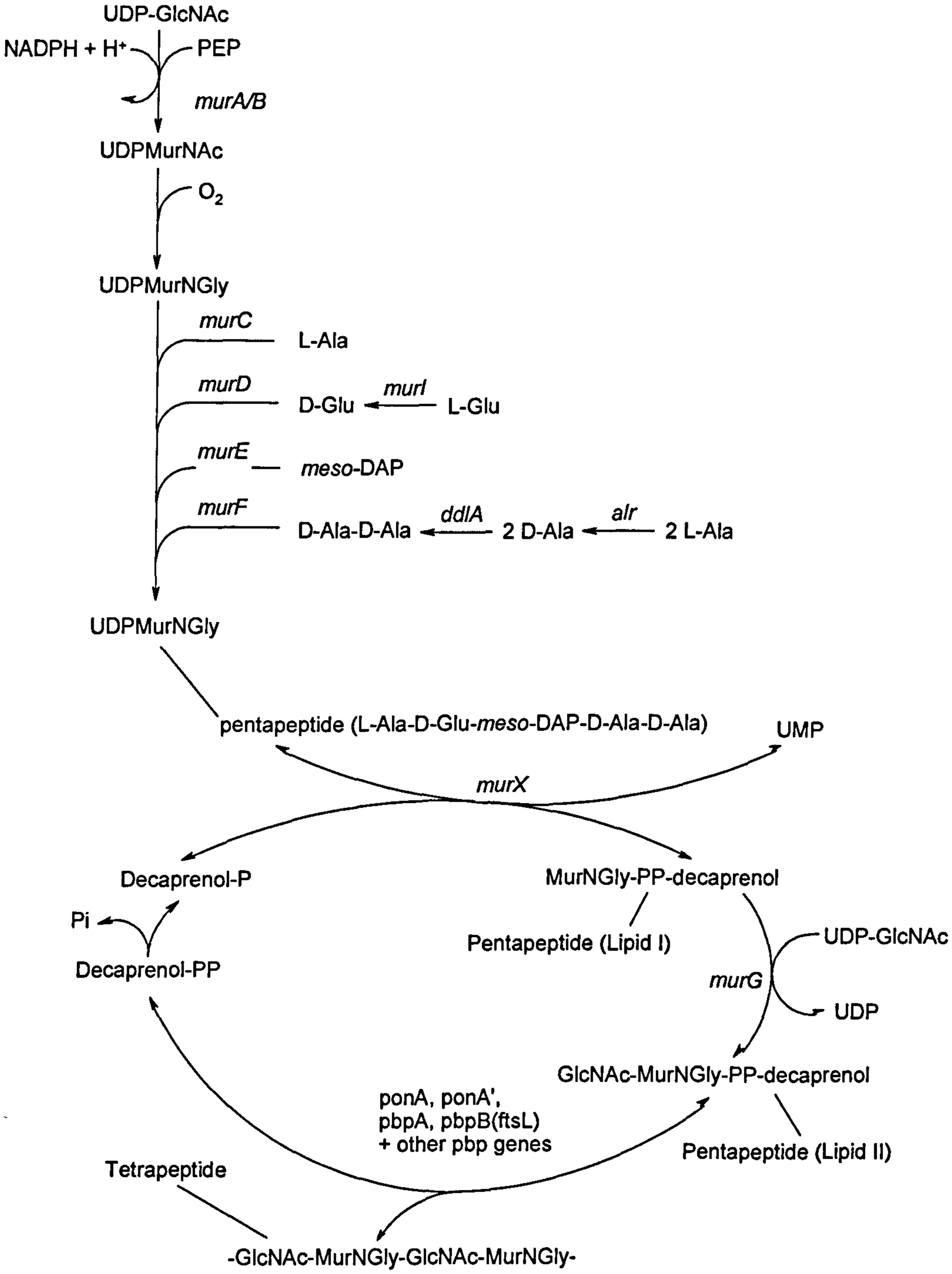


Figure 1.7 Pathway of mycobacterial PG biosynthesis

1.4.2 Linker unit and arabinogalactan

The linker unit (LU) of AG is believed to be the Achilles heel of the complex since the whole mAGP is attached to PG *via* this unique unit (McNeil *et al.*, 1990). The galactan region of AG is linked to the C-6 of some of the MurNGly residues *via* the LU. The LU consists of a diglycosyl-P bridge, α-L-Rha*p*-(1→3)-α-D-GlcNAc-(1→P) (Figure 1.8).

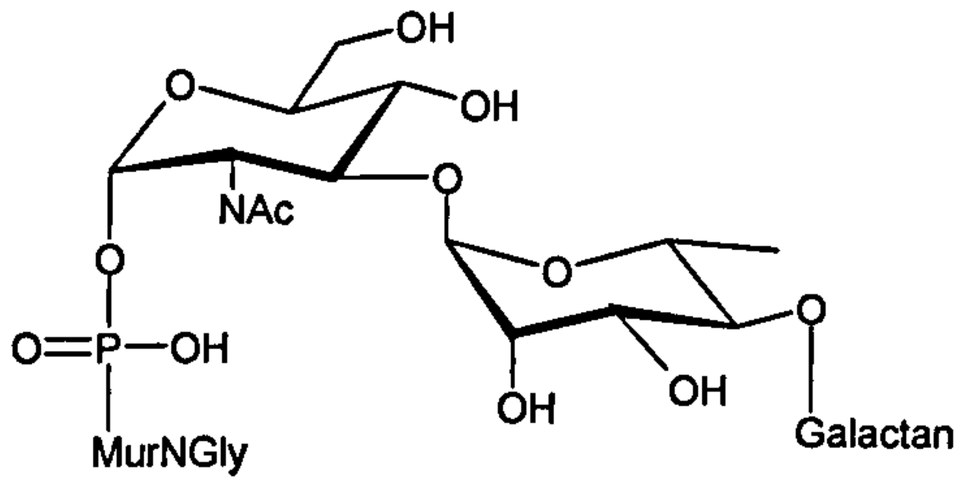


Figure 1.8 Linker unit of the mAGP complex

There was much speculation about the structure of the galactan segment of the AG, whether it was 1→4-linked Gal*p* or 1→5-linked Gal*f* (Vilkas *et al.*, 1973). Recent studies using gas chromatography-mass spectrometry (GS-MS) and fast-atom bombardment mass spectrometry (FAB-MS) established that:

(i) The arabinose (Ara) and galactose (Gal) residues are both in the furanose (*f*) ring form

(ii) The galactan region of AG is linked to the C-6 of some of the MurNGly residues of PG *via* a special diglycosyl-P bridge, α-**L**-Rha*p*-(1→3)-**D**-GlcNAc-(1→P) (Figure 1.8)

(iii) The galactan region consists of a linear alternating 5- and 6-linked β-**D**-Gal*f* residues

(iv) The arabinan chains are attached to C-5 of some of the 6-linked Gal*f* residues and approximately 2-3 arabinan chains are attached to each galactan strand

(v) The arabinan chains are mostly composed of α-5-linked α-**D**-Ara*f* residues with branching introduced at 3,5-α-**D**-Ara*f* residues

(vi) The non-reducing termini of arabinan consists of a hexa-saccharide motif [β-**D**-Ara*f*-(1→2)-α-**D**-Ara*f*]$_2$-3,5-α-**D**-Ara*f*-(1→5)-α-**D**-Ara*f*

(vii) The mycolic acids are located in clusters of four on the terminal hexaarabinofuranoside with only two-thirds of these being mycolated

These linkages are shown in Figure 1.9, which shows the full structure of the mAGP complex.

1.4.2.1 Linker unit and arabinogalactan biosynthesis

LU assembly is dependent on the transfer to GlcNAc-1-phosphate from UDP-GlcNAc to the lipid carrier decaprenyl monophosphate catalysed by *rfe* (Mikusova *et al.*, 2000). RmlA to RmlD have been characterised as an α-D-glucose-1-phosphate thymidylyltransferase, dTDP-D-glucose 4,6-dehydratase, dTDP-4-keto-6-deoxy-D-glucose 3,5 epimerase and dTDP-Rha synthase, respectively (Ma *et al.*, 2002; Ma *et al.*, 2001).

The inactivation of *rmlD* in *M. smegmatis* was found to be lethal in the absence of a rescue plasmid carrying a functional copy of *rmlD* (Ma *et al.*, 2002), thereby demonstrating that dTDP-Rha is an essential component for mycobacterial growth and enzymes involved in its synthesis as novel chemotherapeutic targets. In *E. coli wbbL* has been identified as the rhammosyltransferase involved in lipopolysaccharide biosynthesis (Rubires *et al.*, 1997).

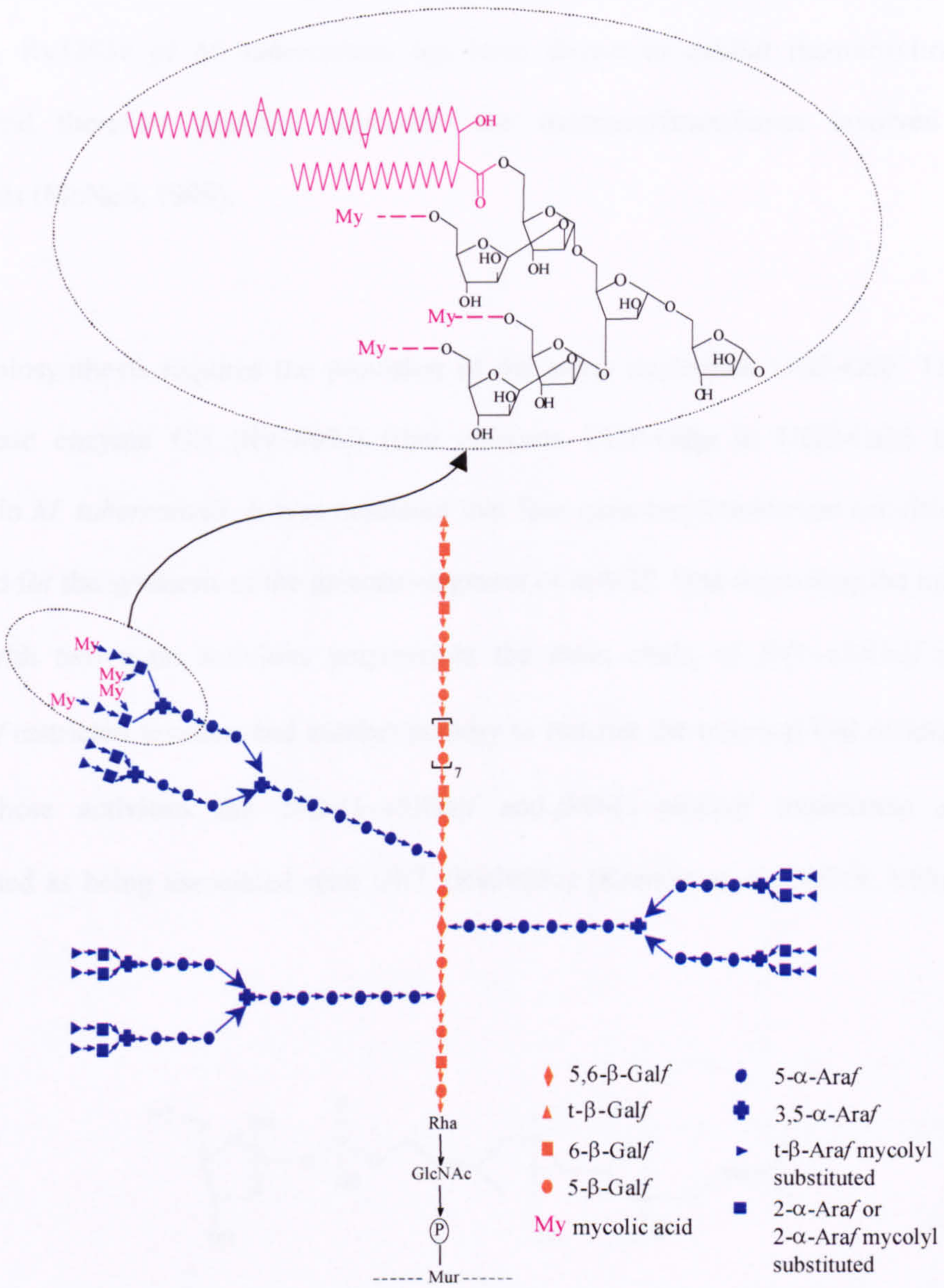


Figure 1.9 **Proposed structure of the mAGP complex of the mycobacterial cell wall. The expanded region shows the structure of the terminal hexaarabinofuranoside attachment site for the esterification of mycolic acids. Adapted from Besra *et al.*, 1995; Daffe *et al.*, 1990; McNeil & Brennan, 1991**

Similarily, Rv3265c of *M. tuberculosis* has been shown to exhibit rhamnosyltransferase activity and therefore possibly represents the rhamnosyltransferase involved in LU biosynthesis (McNeil, 1999).

Galactan biosynthesis requires the provision of the sugar nucleotide UDP-Gal*f*. The UDP-Gal*p* mutase enzyme Glf (Rv3809c) (that converts UDP-Gal*p* to UDP-Gal*f*) has been identified in *M. tuberculosis*. It was predicted that four galactosyltransferase activities would be required for the synthesis of the galactan-segment of mAGP. One depositing the initial Gal*f* residue, with two main activities polymerises the main chain of β-(1→5)Gal*f* and β-D-(1→6)Gal*f* restricted residues and another activity to transfer the terminal Gal residues. With two of these activities the β-D-(1→5)Gal*f* and β-D-(1→6)Gal*f* transferase activities demonstrated as being associated with GlfT (Rv3808c) (Kremer *et al.*, 2001b; Mikusova *et al.*, 2000).

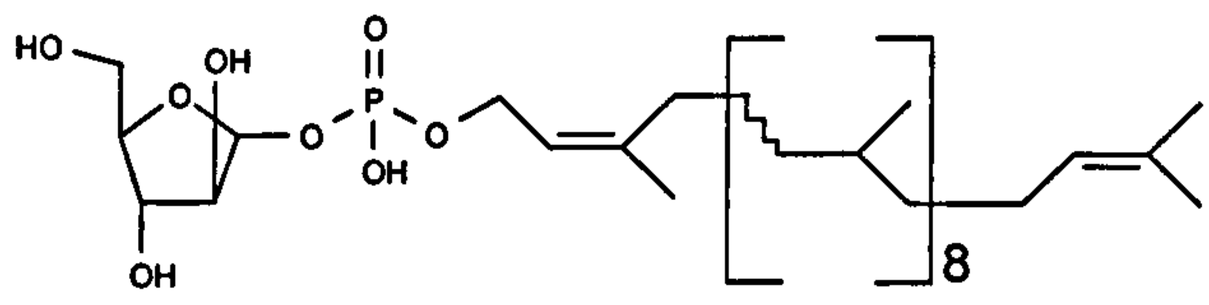


Figure 1.10 Structure of DPA, the donor of Ara*f* units of AG and LAM (Lee *et al.*, 1998; Mikusova *et al.*, 1996)

Biochemical studies using radiolabelled substrates and analysis of internal and external regions of [^{14}C]arabinan, suggested that decaprenyl-monophosphoprenyl-Ara*f* (DPA) is the major, and perhaps only donor of arabinosyl residues in mycobacteria (Xin *et al.*, 1997). The enzymes that direct the formation of DPA in *M. tuberculosis* have not yet been identified.

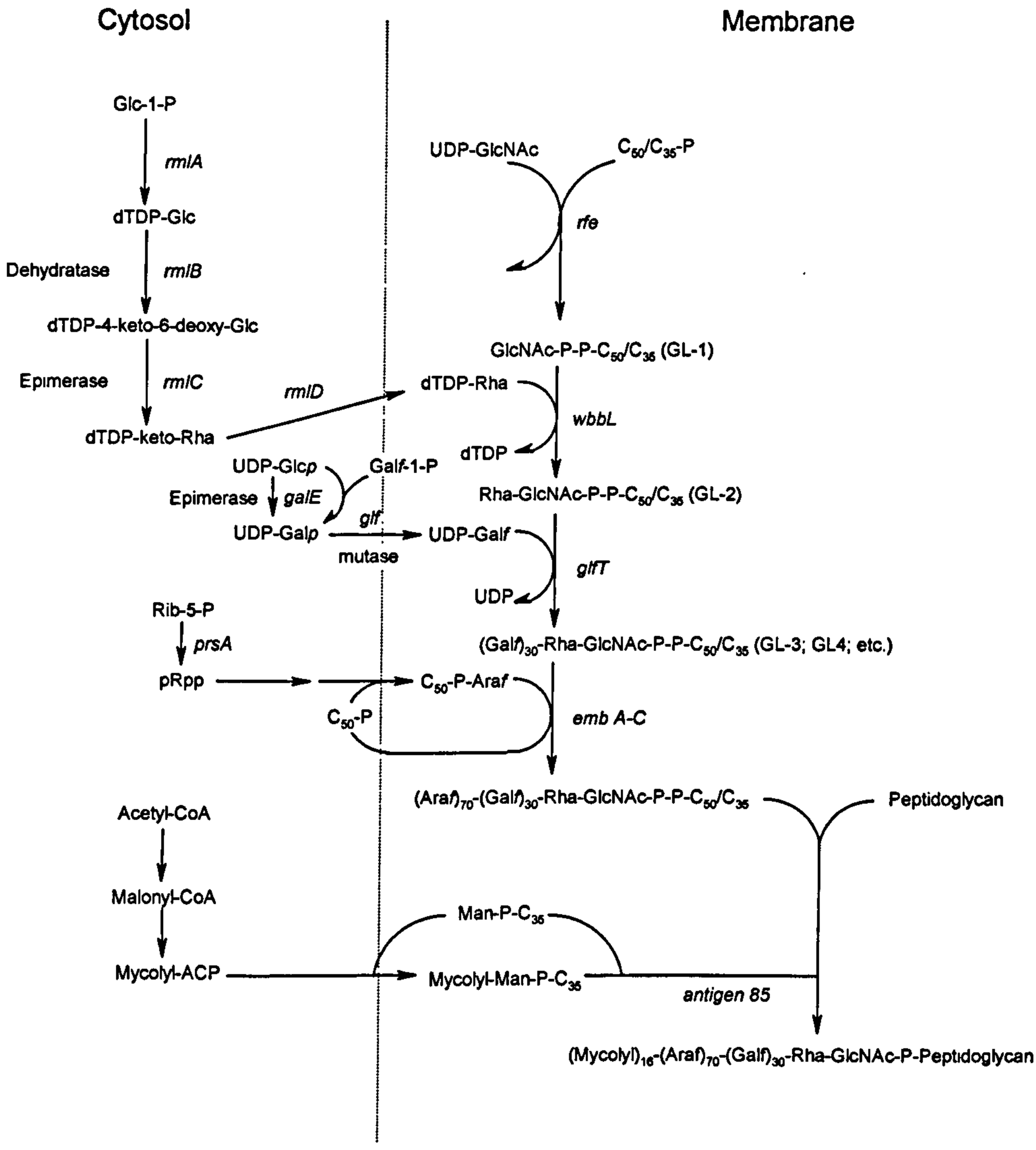


Figure 1.11 Pathway for the biosynthesis of the mAGP complex of *M. tuberculosis* adapted from Baulard *et al.* (1999)

The primary cellular targets of EMB were described as EmbAB in *Mycobacterium avium* and EmbCAB in *M. tuberculosis* (Belanger *et al.*, 1996; Telenti *et al.*, 1997). It was shown that they were responsible for the polymerisation of arabinose into the AG complex (Belanger *et al.*, 1996) and LAM (Khoo *et al.*, 1996). Disruption of either *embA* and *embB* in *M. smegmatis* resulted in the impairment of the synthesis of β-D-Ara*f*-1→2-β-D-Ara*f*

disaccharide to the 3 position of the 3,5-linked Ara*f* residue resulting in a linear terminal motif (Escuyer *et al.*, 2001). Therefore EmbA and EmbB are crucial for the formation of the hexaarabinofuranosyl motif of AG, a important structure for the esterification of mycolic acids to AG (Escuyer *et al.*, 2001). Recently, it was shown that the synthesis of LAM, but not AG, ceases after inactivation of *embC* in *M. smegmatis* by insertional mutagenesis. LAM synthesis was also restored upon complementation with the *embC* wild-type gene. Thus the Emb proteins are capable of differential recognition of the galactan or mannan acceptors prior to appropriate arabinosylation (Zhang *et al.*, 2003).

1.4.3 Mycolic acids

Mycolic acids are high molecular weight (C_{60} – C_{90}) α-alkyl, β-hydroxy fatty acids. In *M. tuberculosis* they comprise approximately 40-60 % of the dry weight of the bacterium. As well as being found esterified to AG, mycolic acids are also present in lipids that are extractable with organic solvents. These are mainly in the form of trehalose 6,6'-dimycolate (TDM) and trehalose monomycolate (TMM) and account for 6 % of the lipid population. Mycolic acids possess at least two chiral centres at positions α and β to the carboxylic acid group (Figure 1.12).

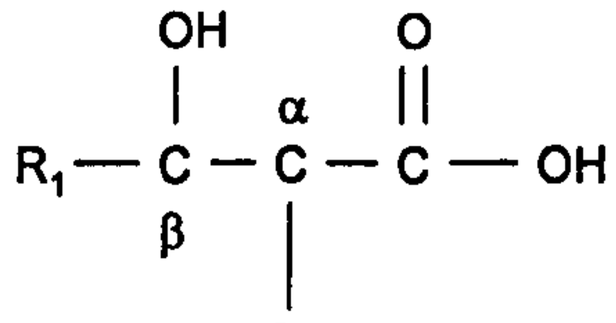


Figure 1.12 Chemical structure of mycolic acids. Where R_1 is a meromycolate chain containing 50-56 carbons and R_2 is a shorter α-branch containing 22-26 carbons

Although the basic mycolate structure is well conserved there are a number of modifications at two mid-chain positions that identify different classes of mycolic acids as shown in Figure 1.13 for *M. tuberculosis* (Kremer *et al.*, 2000a).

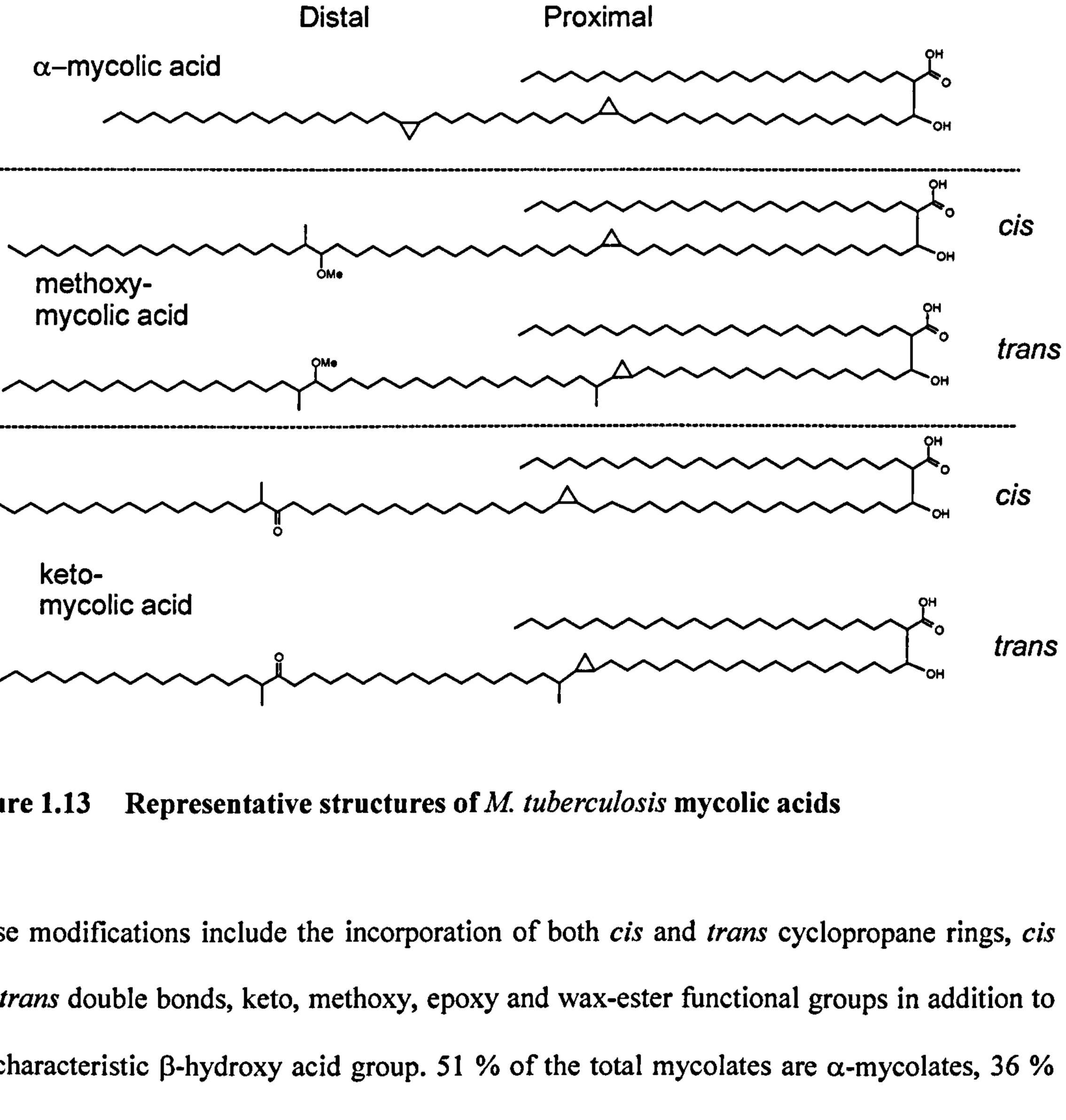


Figure 1.13 Representative structures of *M. tuberculosis* mycolic acids

These modifications include the incorporation of both *cis* and *trans* cyclopropane rings, *cis* and *trans* double bonds, keto, methoxy, epoxy and wax-ester functional groups in addition to the characteristic β-hydroxy acid group. 51 % of the total mycolates are α-mycolates, 36 % are methoxymycolates, and 13 % are ketomycolates (Yuan *et al.*, 1997). α-Mycolic acids contain no oxygen functionality groups at the proximal or distal positions (Minnikin, 1982). The polar oxygen containing functions are found at the distal position with non-polar cyclopropane modifications found at the proximal position keto and methoxymycolates have also been shown to possess *trans*-cyclopropane and *cis*-cyclopropane rings. Cyclopropane

rings and double bonds occurring in a *trans* conformation also possess an adjacent methyl group (Watanabe *et al.*, 2001; Watanabe *et al.*, 2002). Studies into mycolic acid composition revealed that α-mycolates possess 76 to 82 carbons, whereas keto- and methoxymycolates contain 84 to 89 and 83 to 90 carbons, respectively (Watanabe *et al.*, 2001; Watanabe *et al.*, 2002; Yuan *et al.*, 1997). Both the α- and methoxymycolates have only *cis*-cyclopropyl groups at the proximal position, while 17 % of the cyclopropyl groups at the proximal position in ketomycolates are *trans* (Liu *et al.*, 1996). The molecular organisation of these mycolic acid residues may play an important role in nutrient uptake into the bacterium and confer resistance to a wide range of antibacterial agents (Brennan & Nikaido, 1995).

1.4.3.1 Mycolic acid biosynthesis

Mycolic acid biosynthesis can be divided into sequential steps; i) *de novo* fatty acid biosynthesis, ii) elongation, iii) modification of the meromycolate, iv) Claisen condensation and reduction. The synthesis of fatty acids involves the extension of the growing alkyl chain by a two-carbon unit *via* the repetition of a four-part cycle of reactions. The first reaction in each cycle is a condensation reaction in which a malonate residue is decarboxylated and undergoes a condensation reaction with a thioester-linked fatty acyl chain (Heath & Rock, 2002). This reaction achieves the extension of the acyl chain and the remaining three reactions involved in fatty acid biosynthesis remove the resulting β-keto-group of the product and return the chain to its aliphatic form. Firstly, the β-keto product is reduced to form a β-hydroxy intermediate which is then dehydrated to form an enoyl product which is subsequently reduced. The saturated aliphatic substrate is then passed on for the next round of chain extension (Figure 1.14).

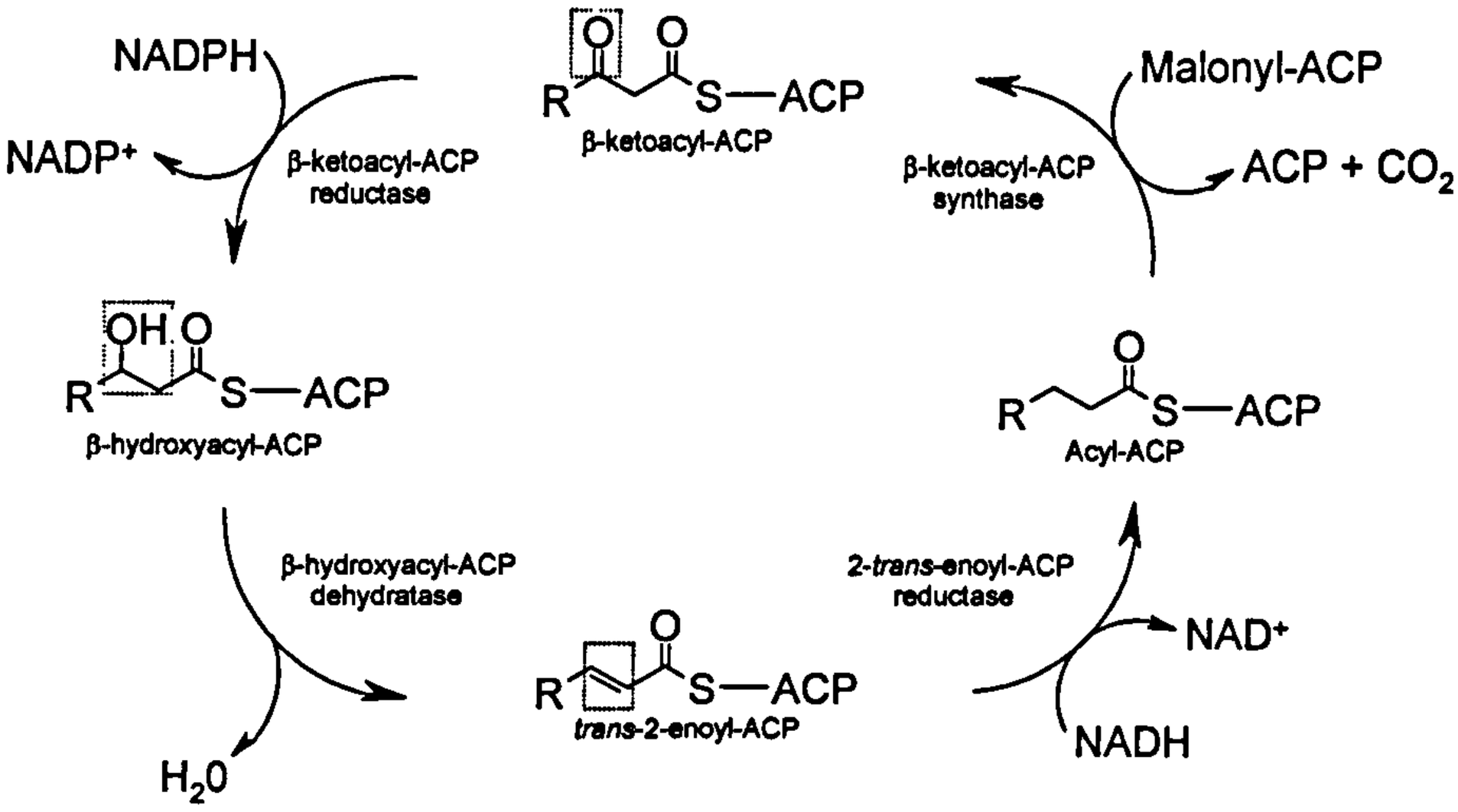


Figure 1.14 Fatty acid biosynthesis. The fatty acid cycle is initiated by decarboxylation of a malonate residue and undergoes a condensation reaction with the thioester of a growing fatty acyl chain. The β-keto product is reduced to form a β-hydroxy intermediate which is then dehydrated to form an enoyl product which is subsequently reduced. The saturated aliphatic substrate is then passed on for the next round of chain extension. Red areas indicate the area of activity

Two types of fatty acid synthases (FAS) are known, both containing similar enzymatic functions but differing in their organisation. A single gene encodes fatty acid synthase-I (FAS-I), commonly found in mammals and the active homodimeric protein possesses all the necessary functions to perform *de novo* fatty acid biosynthesis (Howard, 1968; Smith *et al.*, 2003) from acetyl-CoA and malonyl-CoA. Most bacteria and plants, however, utilise fatty acid synthase-II (FAS-II) in which the growing fatty acyl chain is shuttled between the active sites of disassociated enzymes as an acyl thioester of a small and highly acidic acyl carrier protein (ACP). Mycobacteria are unusual in that they contain both FAS-I and FAS-II (Figure 1.15) (Table 1.3) (Kremer *et al.*, 2000a). FAS-I encoded by *fas*, is responsible for *de novo* fatty acid biosynthesis, producing a bimodal distribution of C_{16}-C_{26} fatty acids, mainly

thioesters of CoA. FAS-II of *M. tuberculosis* is similar to other bacterial FAS-II synthases except for its primer specificity (Choi *et al.*, 2000b). The *M. tuberculosis* FAS-II extends relatively long acyl-chain thioesters of AcpM, a C-terminally-extended homologue of other bacterial ACPs, to provide long-chain meromycolates leading ultimately to mature mycolic acids (Asselineau *et al.*, 2002; Barry *et al.*, 1998; Kremer *et al.*, 2000a).

Table 1.3 Genes involved in fatty acid biosynthesis in *M. tuberculosis*

Known Gene	*M. tuberculosis* Homologue	Function
	Rv0033	Possible acyl carrier protein
fabG4	Rv0242c	3-oxoacyl-ACP reductase
fabH	Rv0533c	β-Ketoacyl-ACP synthase III
fabD2	Rv0649	Malonyl CoA:ACP transacylase
accD3	Rv0904c	Acetyl/propionyl CoA carboxylase β-subunit
	Rv1344	Possible acyl carrier protein
fabG2	Rv1350	3-oxoacyl-ACP reductase
fabG1 (MabA)	Rv1483	3-oxoacyl-ACP reductase
inhA	Rv1484	Enoyl-ACP reductase
tesB1	Rv1618	Thioesterase II
	Rv1722	Possible biotin carboxylase
fabG3	Rv2002	3-oxoacyl-ACP reductase
fabD	Rv2243	Malonyl CoA:ACP transacylase
acpM	Rv2244	Acyl carrier protein
kasA	Rv2245	β-Ketoacyl-ACP synthase
kasB	Rv2246	β-Ketoacyl-ACP synthase
accD6	Rv2247	Acetyl/propionyl CoA carboxylase β-subunit
acpS	Rv2523c	CoA:apo-ACP pantethienephosphotransferase
fas	Rv2524c	Fatty acid synthase I
tesB2	Rv2605c	Thioesterase II
fabG5	Rv2766c	3-oxoacyl-ACP reductase
	Rv3221c	Possible biotin carboxyl carrier
accD5	Rv3280	Acetyl/propionyl CoA carboxylase β-subunit
accA3	Rv3285	Acetyl/propionyl CoA carboxylase β-subunit
	Rv3472	Possible acyl carrier protein
accD4	Rv3799c	Acetyl/propionyl CoA carboxylase

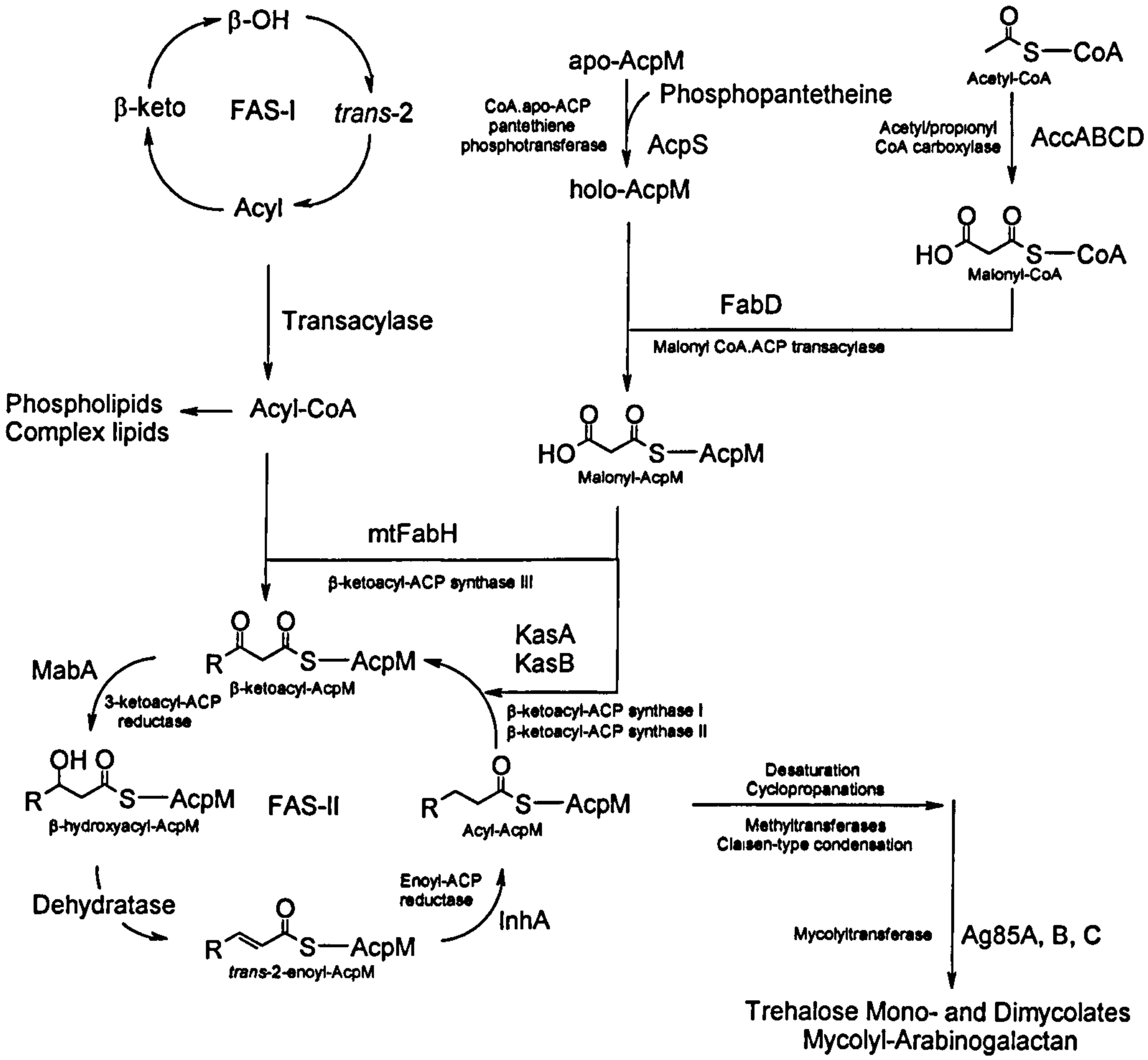


Figure 1.15 **A schematic illustration of *M. tuberculosis* mycolic acid biosynthesis. The product of FAS-I (C_{16}) is transacylated to CoA and utilised by mtFabH in the initiation of FAS-II. Malonyl-AcpM produced by FabD is used as the C2 donor for elongation. FAS-II synthesises the long aliphatic acyl-chains which are further modified to produce the mature meromycolate chains of mycolic acids**

The β-ketoacyl-ACP synthase III, mtFabH (Rv0533c) (Choi *et al.*, 2000b) has been proposed as the pivotal link between FAS-I and FAS-II mtFabH elongates acyl-CoA primers derived from FAS-I forming β-ketoacyl-AcpM thioester products through condensation with malonyl-AcpM. In the proposed reaction mechanism, mtFabH is initially transacylated with the

incoming acyl-CoA forming a thioester with the active site cysteine residue (C122), its CoA moiety being released. The docking of malonyl-AcpM is followed by decarboxylation of the malonate residue. The resulting carbanion attacks the acyl-enzyme thioester linkage, thereby initiating the condensation reaction (Heath & Rock, 2002; Scarsdale *et al.*, 2001) (Figure 1.16).

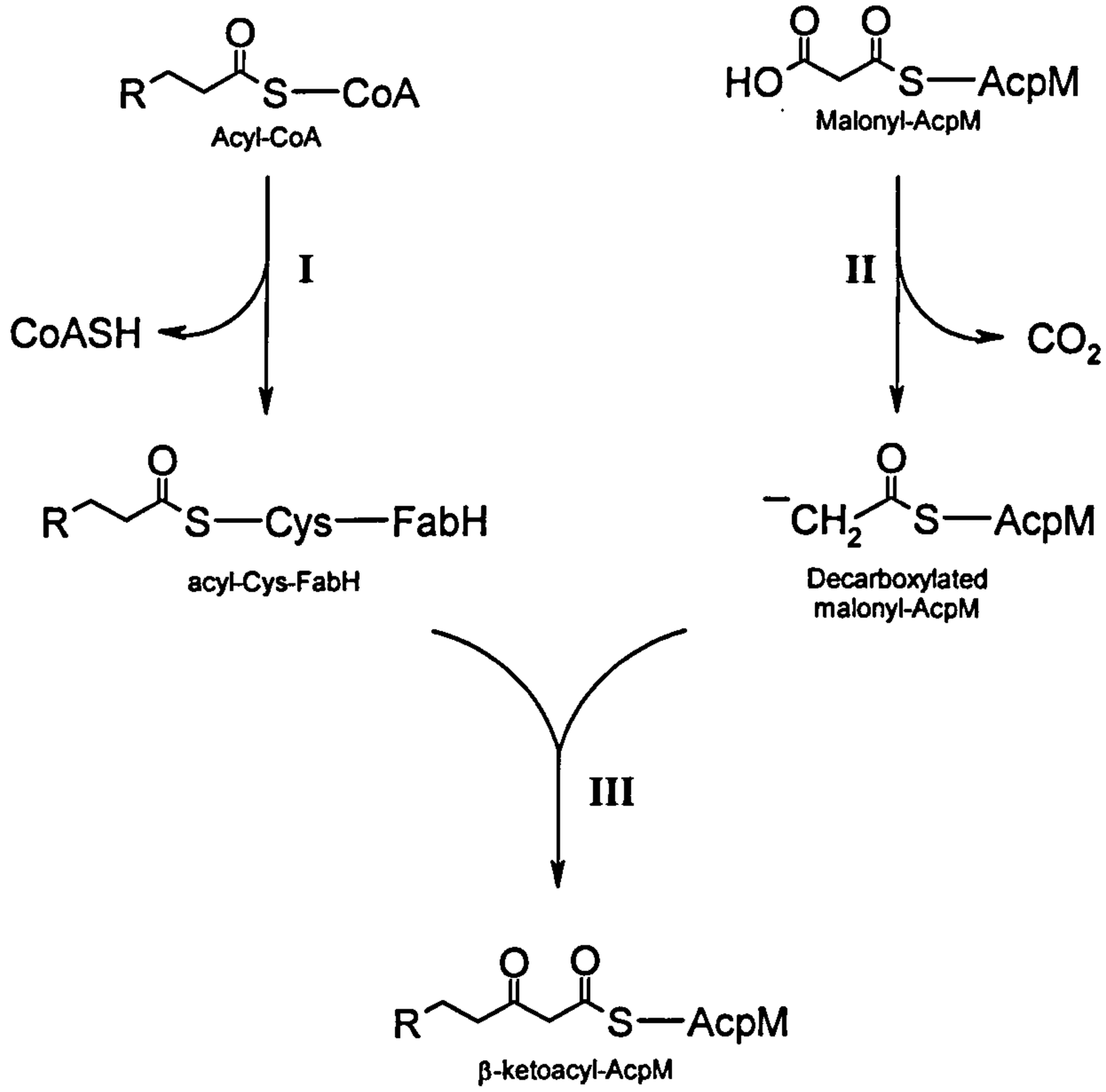


Figure 1.16 β-Ketoacyl-ACP synthase III activity of mtFabH. (I) The initial transacylation of the acyl-CoA moiety to the cysteine of the catalytic triad followed by (II) decarboxylation of the donor unit of the malonyl-AcpM. The condensation of the two subunits (III) form the β-ketoacyl AcpM product

The reduction of the characteristic β-keto group to a hydroxyl function is carried out by the β-ketoacyl-ACP-reductase, MabA (Rv1483). *mabA* has been cloned and expressed in *E. coli* and its NADPH-dependent β-keto-reductase activity biochemically characterised (Banerjee *et al.*, 1998; Marrakchi *et al.*, 2002). The recent crystal structure of MabA has led to the recognition of structural differences that may distinguish *M. tuberculosis* MabA and several homologues suggested to have an affinity for long-chain substrates from most of the other known β-ketoacyl-ACP reductases (Marrakchi *et al.*, 2002). In *M. tuberculosis, mabA* lies adjacent to *inhA* (Rv1484) (Figure 1.17), which encodes an NADH-specific 2-*trans*-enoyl-ACP reductase involved in FAS-II (Dessen *et al.*, 1995).

Figure 1.17 Genomic organisation of open reading frames involved in long-chain fatty acid biosynthesis

A detailed biochemical study demonstrated that InhA prefers long-chain substrates consistent with its involvement in long chain fatty acid biosynthesis (Quemard *et al.*, 1995). The gene encoding the dehydratase function converting the β-hydroxyacyl-AcpM product of MabA into the enoyl-AcpM substrate for InhA has not yet been identified in *M. tuberculosis*. The subsequent rounds of acyl extension in FAS-II are thought to be initiated by the highly homologous β-ketoacyl-ACP synthases, KasA (Rv2245) and KasB (Rv2246). Both enzymes have been expressed in *E. coli*, purified and characterised (Kremer *et al.*, 2002c; Schaeffer *et al.*, 2001). Both extend acyl-ACP thioesters rather than acyl-CoA, condensing them with malonyl-ACP (Kremer *et al.*, 2002c; Schaeffer *et al.*, 2001). Both enzymes prefer acyl-ACPs

greater than 16 carbons long consistent with their proposed role in FAS-II biosynthesis. Kremer *et al.* (2002) also demonstrated that over-expression of *kasA* from *M. tuberculosis* in both *M. smegmatis* and *M. chelonae* led to a decrease in the amount of shorter α'-mycolates with an overall increase in α-mycolates (Kremer *et al.*, 2002c). The KasAB enzymes of *M. smegmatis* are very similar to those of *M. tuberculosis* and over-expression of *M. smegmatis kasA* produced the same effect suggesting that the balance of α'- and α-mycolates is probably achieved by regulation of the amounts of the KAS proteins and associated enzymes (Kremer *et al.*, 2002c). In *M. tuberculosis*, KasAB are clustered with a series of genes implicated in fatty acid biosynthesis, which presumably form an operon. These include *acpM* and *fabD*, the latter encoding a malonyl-CoA acyltransferase responsible for the formation of the malonyl-AcpM which is used as a substrate by mtFabH, KasA and KasB (Kremer *et al.*, 2001a) (Figure 1.17). Recently it was shown that *M. marinum* mutants with a transposon insertion in *kasB* grew poorly in macrophages, although growth *in vitro* was unaffected (Gao *et al.*, 2003). Detailed chemical analysis of mycolic acids from the *kasB* mutant found that they were 2-4 carbons shorter than mycolic acids isolated from wild type *M. marinum*. The defect was localised to the proximal portion of the meromycolate chain. The *kasB* mutant also showed a significant reduction in the abundance of ketomycolates, with a slight compensatory increase in both α- and methoxy mycolates. Despite these small changes in mycolate length and composition, the *kasB* mutants exhibited strikingly altered cell wall permeability, with increased susceptibility to lipophilic antibiotics and host antimicrobial molecules. Complementation experiments with *M. tuberculosis kasB* reverted the phenotype to wild-type, whereas KasA complementation did not (Gao *et al.*, 2003).

1.4.3.2 Meromycolate modification

A range of monounsaturated and diunsaturated long chain fatty acids were extracted from *M. smegmatis* that appeared to be related to the mero-chain of α-mycolates (Sathyamoorthy *et al.*, 1985). Similar meromycolate-like C_{48}-C_{56} fatty acids were synthesised using a palmitate precusor in an *M. tuberculosis*-derived cell free system and [^{14}C]-S-adenosyl-L-methionine (SAM) as a substrate (Qureshi *et al.*, 1984). Recently, Barry and colleagues have provided strong evidence supporting the hypothesis that the meromycolate chain is modified during elongation prior to its condensation with the α-alkyl chain (Yuan *et al.*, 1998b). They observed that a range of long chain fatty acid residues derived from heat-inactivated extracts of *M. smegmatis* act as acceptors of radiolabelled methyl groups from SAM. Labelling was strongly inhibited by an antiserum that specifically recognises AcpM implicating the protein as the acyl chain carrier throughout meromycolate elongation and modification (Yuan *et al.*, 1998b). Analysis of the *M. tuberculosis* genome has identified three potential desaturases DesA1, DesA2 and DesA3 (Cole *et al.*, 1998; Jackson *et al.*, 1997). These desaturases are speculated to form the double bonds of the meromycolate precursor allowing for the production of species-specific modification to form the range of mycolic acids (Jackson *et al.*, 1997). Cyclopropanation is a common modification in mycolic acids from slow-growing pathogenic mycobacteria, whereas fast-growing species (e.g. *M. smegmatis*) do not produce cyclopropanated mycolates, but produce large amounts of unsaturated mycolic acids (Minnikin, 1982). The overproduction of a distal cyclopropanating encoded by *cma1* in *M. tuberculosis* was expressed in *M. smegmatis* resulting in hydrogen peroxide resistance being observed, thus suggesting that this is an important adaptation of pathogenic mycobacteria to oxidative stress (Yuan *et al.*, 1995). The related *cmaA2* was shown to *cis*-cyclopropanate the proximal position of the meromycolate chain (George *et al.*, 1995). Recently, it was

demonstrated that the *cmaA2* is actually required for the synthesis of *trans*-cyclopropane rings of both ketomycolates and methoxymycolates, defining *cmaA2* as a mycolic acid proximal *trans*-cyclopropane synthase (Glickman *et al.*, 2001). The systematic deletion of the cyclopropane modification enzymes in *M. tuberculosis* identified PcaA, a homologue of *cmaA2*, as the enzyme involved in the synthesis of the proximal *cis* cyclopropyl group of mycolic acids (Glickman, 2003) (Figure 1.18).

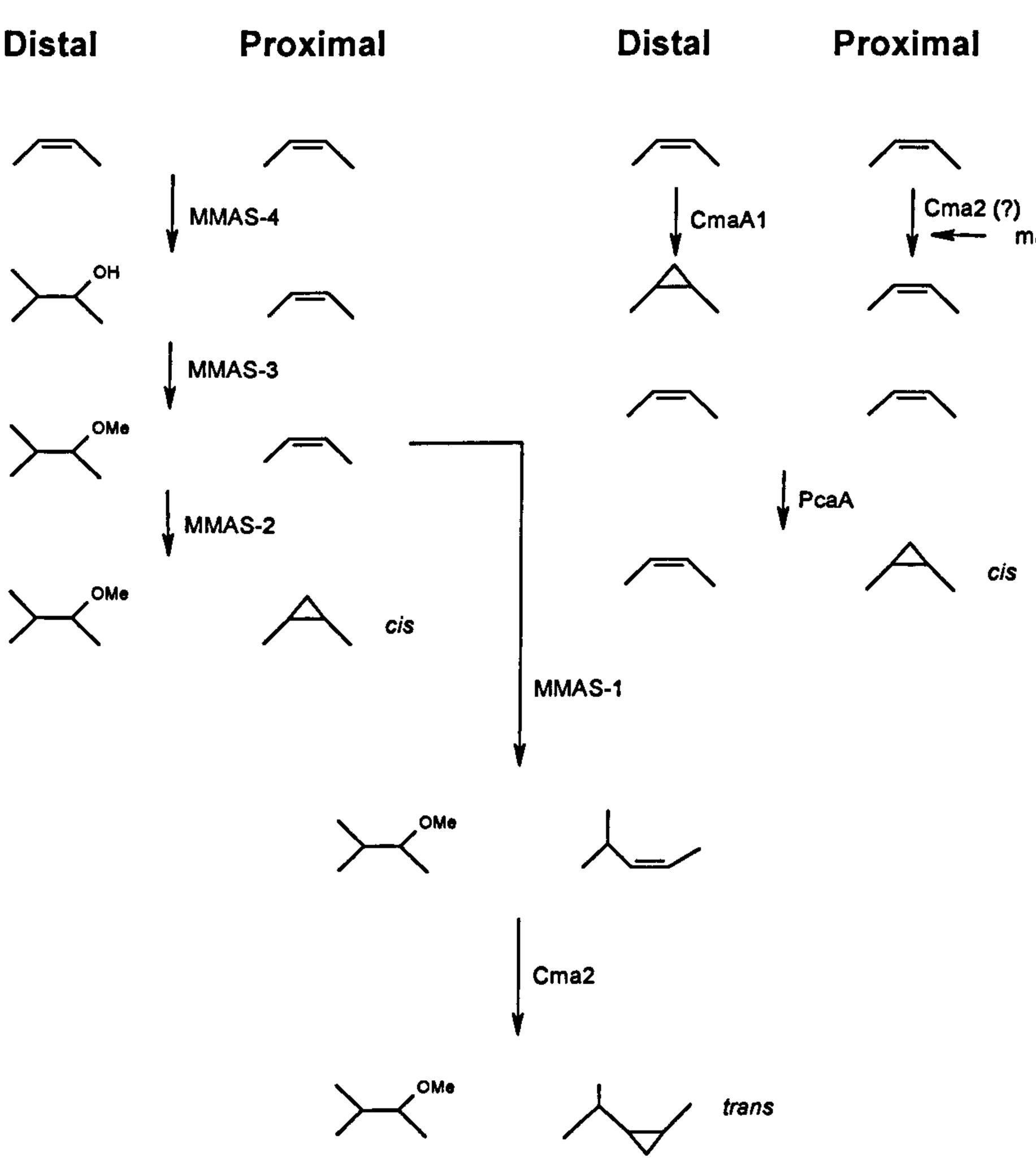


Figure 1.18 Meromycolate modification in *M. tuberculosis*. The proposed pathway and involvement of genes in the biosynthesis of methoxymycolates in *M. tuberculosis* adapted from Glickman, (2003); Huang *et al.*, (2002); Yuan & Barry, (1996).

Through the use of a *cma1* gene probe, other meromycolate modifying enzymes were identified as *mma1-4* encoding MMAS 1-4 (Yuan & Barry, 1996). It was subsequently shown that both the *cma* and *mma* genes have a high degree of homology and share a highly-conserved SAM binding region. The MMAS proteins are suggested to be involved in the biosynthesis of methoxymycolates, with MMAS-4 producing a distal hydroxyl group with a methyl branch on the adjacent carbon atom. MMAS-3 appears to *O*-methylate the hydroxyl group to form a methoxy group. MMAS-2 apparently possess proximal *cis*-cyclopropanation activity (Yuan & Barry, 1996) (Figure 1.18). Although the function of MMAS-1 was not obvious, upon over-expression in *M. tuberculosis* it was suggested that MMAS-1 regulated a branch point between *cis*- and *trans*-cyclopropane containing oxygenated mycolates (Yuan *et al.*, 1997). The over-expression of MMAS-3 in *M. tuberculosis* led to the total replacement of ketomycolates with methoxymycolates. This led to numerous effects, such as increased permeability properties of the cell wall consistent with the hypersensitivity to both ampicillin and RIF, impaired growth at low temperature and the inability to replicate in macrophage THP-1 cells (Yuan *et al.*, 1998a). A strain of *M. tuberculosis* devoid of *mma4* activity was unable to synthesise oxygenated mycolates and was unable to grow in THP-1 cells (Dinadayala *et al.*, 2003). As there are no reported strains of *M. smegmatis* producing ketomycolates the existence of an *mma4* homologue is unlikely, upon over-expression of the *M. tuberculosis mma4* gene in *M. smegmatis* it was expected that ketomycolates would be observed, but this wasn't the case (Yuan *et al.*, 1998a). Glickman *et al.* (2001) recently used gene deletion studies of CMAS-2 which clarified its role in mycolic acid modification. The deletion mutant no longer produced *trans* cyclopropane-containing mycolic acids suggesting that the previous assignment of CMAS-2 as a *cis* cyclopropane synthase was incorrect, this later data indicates that it is indeed a *trans* cyclopropane synthase. This result indicates that further studies of deletion are required for full determination of meromycolate metabolism, as many inconsistencies are observed in comparisons of over-expression data against deletion

results. Although there is no supporting experimental evidence, the coordination of meromycolate elongation and modification in multi-enzyme complexes seems a likely route to efficient meromycolate biosynthesis. Therefore the production of *cis*-cyclopropane rings in *M. smegmatis* over-expressing *cma2* may reflect a mis-incorporation of CMAS-2 in such a complex (Dover *et al.*, 2004).

1.4.3.3 Claisen condensation type reactions

The isolation of a mycolate-deficient mutant of *M. smegmatis* that accumulates novel long-chain fatty acids that correspond in structure to incomplete meromycolates provided strong evidence that the meromycolate chain and α-branch are synthesised separately and fused by condensation (Liu & Nikaido, 1999). The mechanism for this carbon-carbon bond formation used in their biosynthesis is likely to occur *via* a decarboxylating Claisen-type condensation reaction (Heath & Rock, 2002). Various studies have proposed possible pathways for this type of reaction but with conflicting mechanisms, some involving an intermediate decarboxylation step (Figure 1.19, Theory I) (Minnikin, 1982) and others proceeding directly without a decarboxylation step (Figure 1.19, Theory II) (Lee *et al.*, 1997).

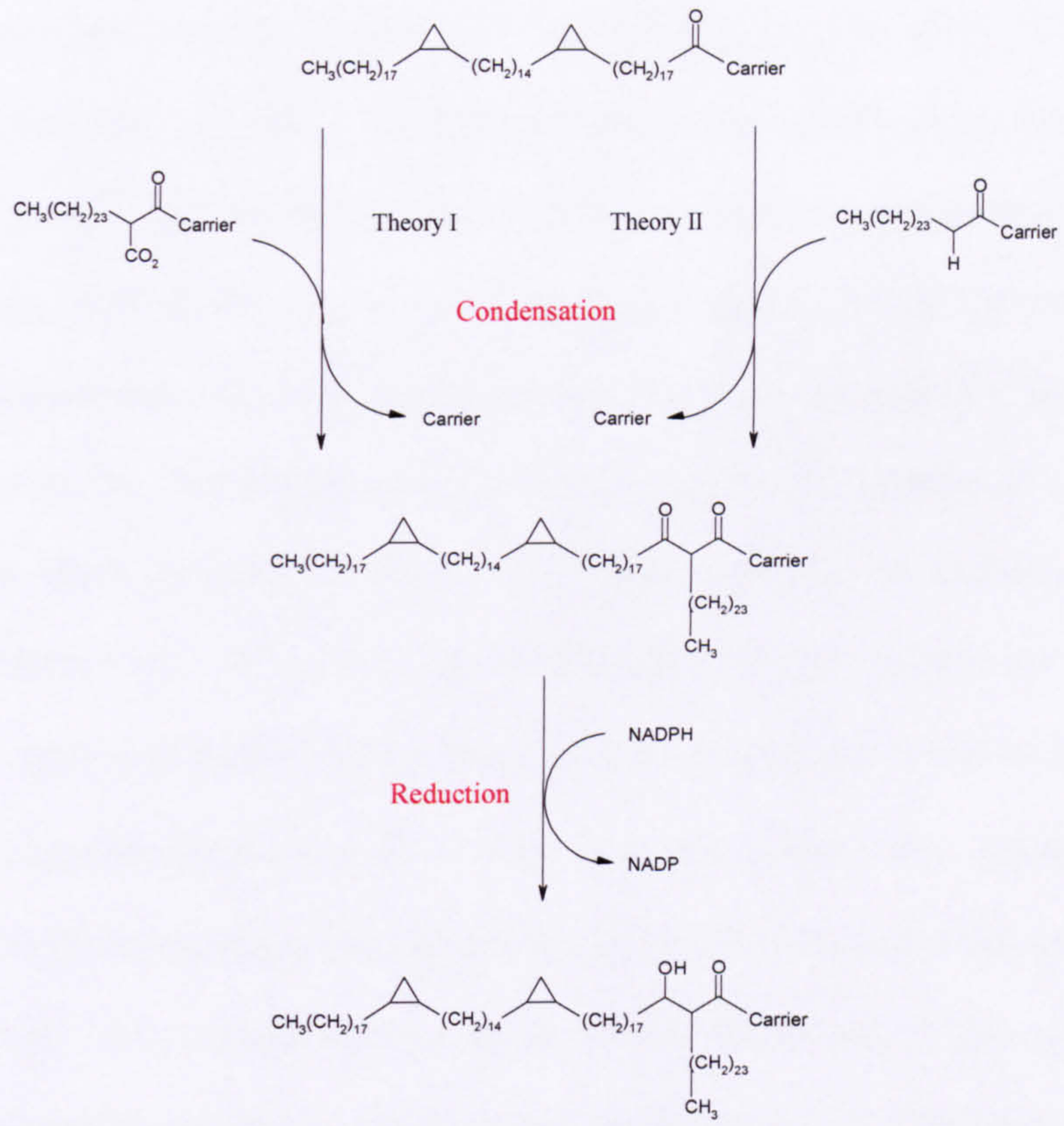


Figure 1.19 Claisen condensation step proposed in the production of mycolic acids.

1.4.3.4 Deposition and export of mycolic acids

The deposition and export of mycolates to the cell wall has been partially elucidated. A mycolated polyprenyl phosphate-linked carrier described as Myc-PL (6-*O*-mycolyl-β-D-mannopyranosyl-1-monophosphoheptaprenol) from *M. smegmatis* (Besra *et al.*, 1994) (Figure 1.11) has been implicated in the final stages of mAG synthesis. The introduction of [^{14}C]Myc-PL into a cell wall preparation from *M. smegmatis* led to the transfer of [^{14}C]mycolates into extractable TMM and TDM as well as insoluble mycolates esterified to AG (Besra *et al.*, 1994). Belisle *et al.* (1997) further demonstrated that three members of the *M. tuberculosis*

antigen 85 complex, Ag85A, Ag85B and Ag85C2 (encoded by *fbpA*, *fbpB* and *fbpC2* respectively) were able to catalyse mycolyltransferase reactions involving TDM, TMM and mycolates of the cell wall AG (Belisle *et al.*, 1997). Heterologous expression of FbpA, FbpC1 and FbpC2 was performed in *E. coli* and recombinant proteins purified. In contrast to FbpA and FbpC2, recombinant FbpC1 did not possess *in vitro* mycolyltransferase activity and was not recognised by two monoclonal antibodies to the native Ag85 (Kremer *et al.*, 2002a). To aid functional characterisation of FbpC1 the crystal structure was determined to 1.7 Å resolution (Wilson *et al.*, 2003). It was shown that FbpC1, like the Ag85 components Ag85B and Ag85C2, folds as an α/β hydrolase, but it does not contain any of the catalytic elements required for mycolyltransferase activity. Moreover, the absence of a recognisable α,α'-trehalose monomycolate-binding site and the failure to detect an active site suggest that the function of FbpC1 is of a non-enzymatic nature and that FbpC1 may in fact represent a new family of non-catalytic alpha/beta hydrolases that might have a role in host tissue attachment, whereby ligands may include the serum protein fibronectin and small sugars (Wilson *et al.*, 2004).

1.4.4 Lipoarabinomannan

LAM and its truncated form lipomannan (LM) are thought to be the primary plasma membrane-associated lipopolysaccharides in mycobacteria (Lee *et al.*, 1996). LAM is present in the envelopes of pathogenic mycobacteria, such as *M. tuberculosis* and *M. leprae*, the vaccine strain *M. bovis* BCG and opportunistic strains, such as *M. avium*, *M. kansasii*, *M. fortuitum* and *M. chelonae* and finally the fast-growing non-pathogenic strain *M. smegmatis* (Besra & Chatterjee, 1994; Nigou *et al.*, 2003). The biological and immunological properties

of LAM are significant and have been studied extensively. Mycobacterial LAMs are lipoglycans comprised of three distinctive domains, a mannosyl-phosphatidyl-*myo*-inositol (MPI) anchor, a polysaccharide backbone and capping moieties. The structure of the MPI-anchor is illustrated in Figure 1.20. The anchor structure is heterogeneous, with variations in the number and location of fatty acid residues (Nigou *et al.*, 1997). Its acylation with a combination of palmitic, stearic and tuberculostearic (10-methyl-octadecanoic) acids facilitates the interaction of the anchor and the plasma membrane (Guerardel *et al.*, 2002). It has also been suggested that mycobacteria can escape intramacrophagic destruction by adapting the lipids of the MPI-anchor to reduce stimulation of the host cytokine response (Nigou *et al.*, 1997).

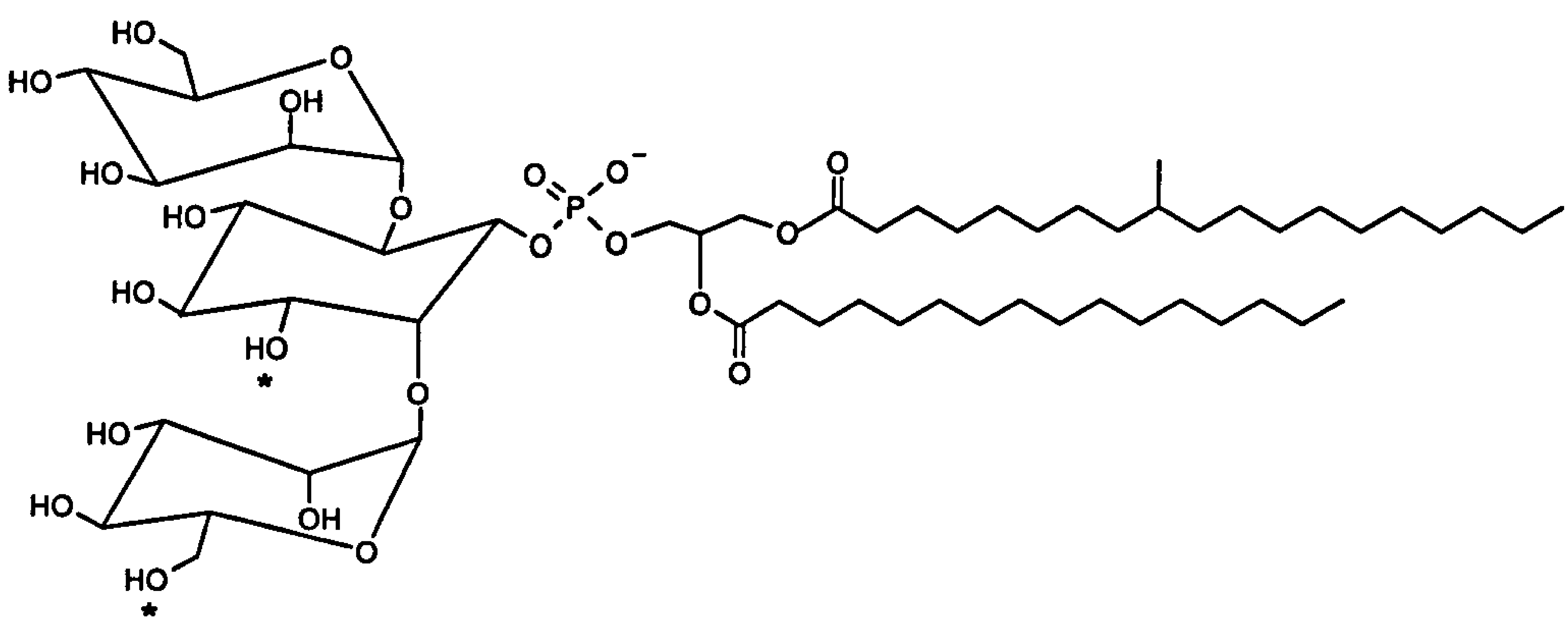


Figure 1.20 Structure of PIM_2 illustrating the MPI-anchor common to LAM, LM and PIMs. The central *myo*-inositol moiety is linked to Man*p* residues at position 2 and 6 and to a glycerol moiety *via* a phosphate group linked to carbon 1. Two of the potential acylation sites are occupied by tuberculostearic (10-methyl-octadecanoic) and palmitic acid residues. Other potential acylation sites are marked with an asterisk (*)

The polysaccharide backbone is composed of two homopolysaccharides, D-mannan and D-arabinan. The structure of this backbone is highly conserved amongst different mycobacterial species. The mannan core is composed of around 30-35 mannose (Man) residues in the pyranose (*p*) form. These residues form a linear α-(1→6)-Man*p* backbone with various single α-(1→2)-Man*p* side chains. The lipoglycans of *M. tuberculosis* H37Rv, *M. tuberculosis* H37Ra and *M. bovis* BCG all have very similar mannan cores with 60-70 % of the backbone possessing side-chains (Nigou *et al.*, 2002b) (Figure 1.21). The arabinan domain comprises around 60 arabinofuranose residues (Ara*f*) linked in a linear α-(1→5)-Ara*f* fashion, with branching occurring *via* the 3-position of some residues. These lateral branches occur in two arrangements, a linear tetraarabinofuranoside (Ara*f*-(β1→2)-Ara*f*-(α1→5)-Ara*f*-(α1→5)-Ara*f*-(α1→) and a bi-antennary hexaarabinofuranoside ([Ara*f*-(β1→2)-Ara*f*-(α1→]$_2$ →3 and →5)-Ara*f*-(α1→5)- Ara*f*-(α1→) (Chatterjee *et al.*, 1991) (Figure 1.22).

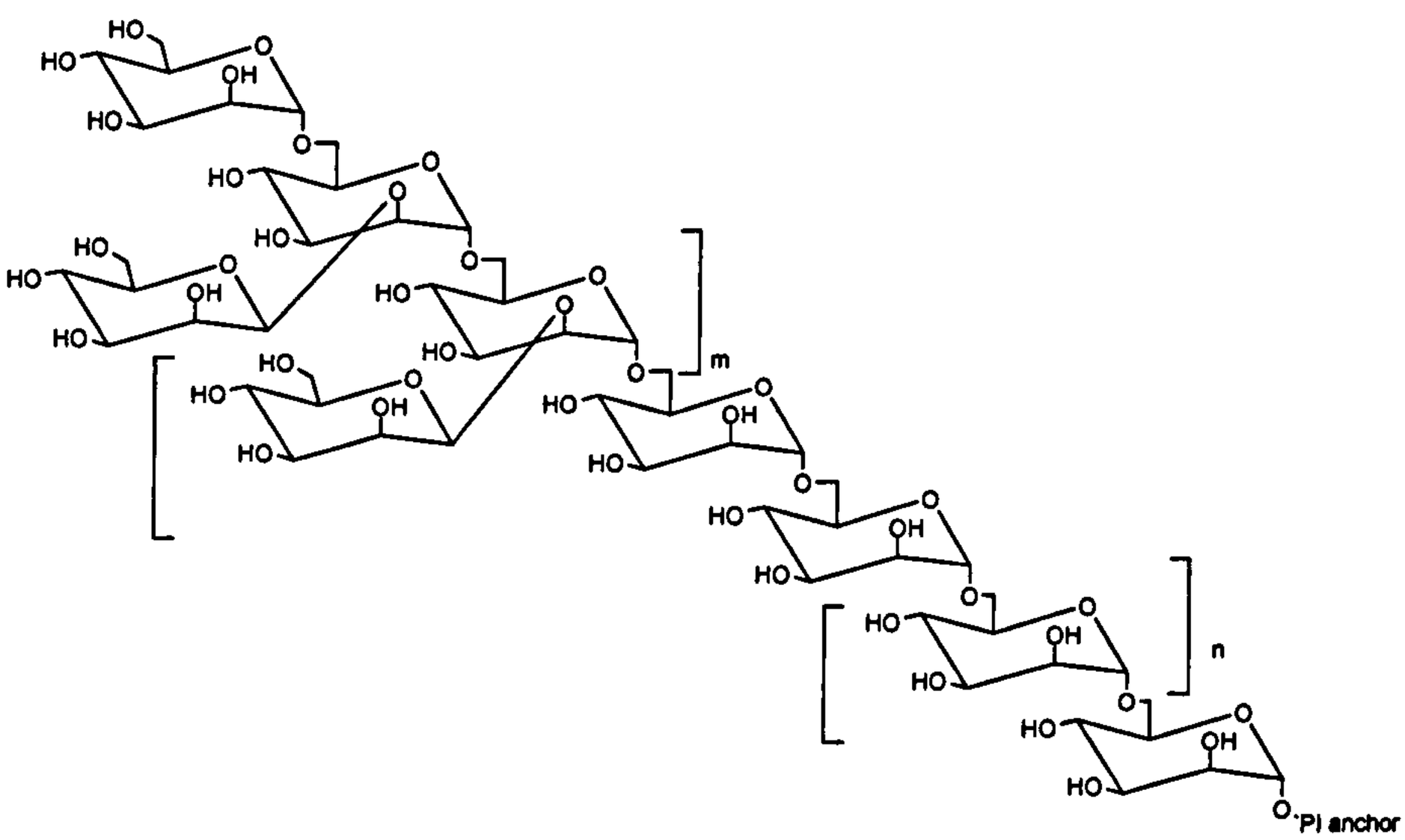


Figure 1.21 **Mycobacterial mannan core composed of a linear α-(1→6)-Man*p* backbone with single α-(1→2)-Man*p* side chains (n and m = 7, 8)**

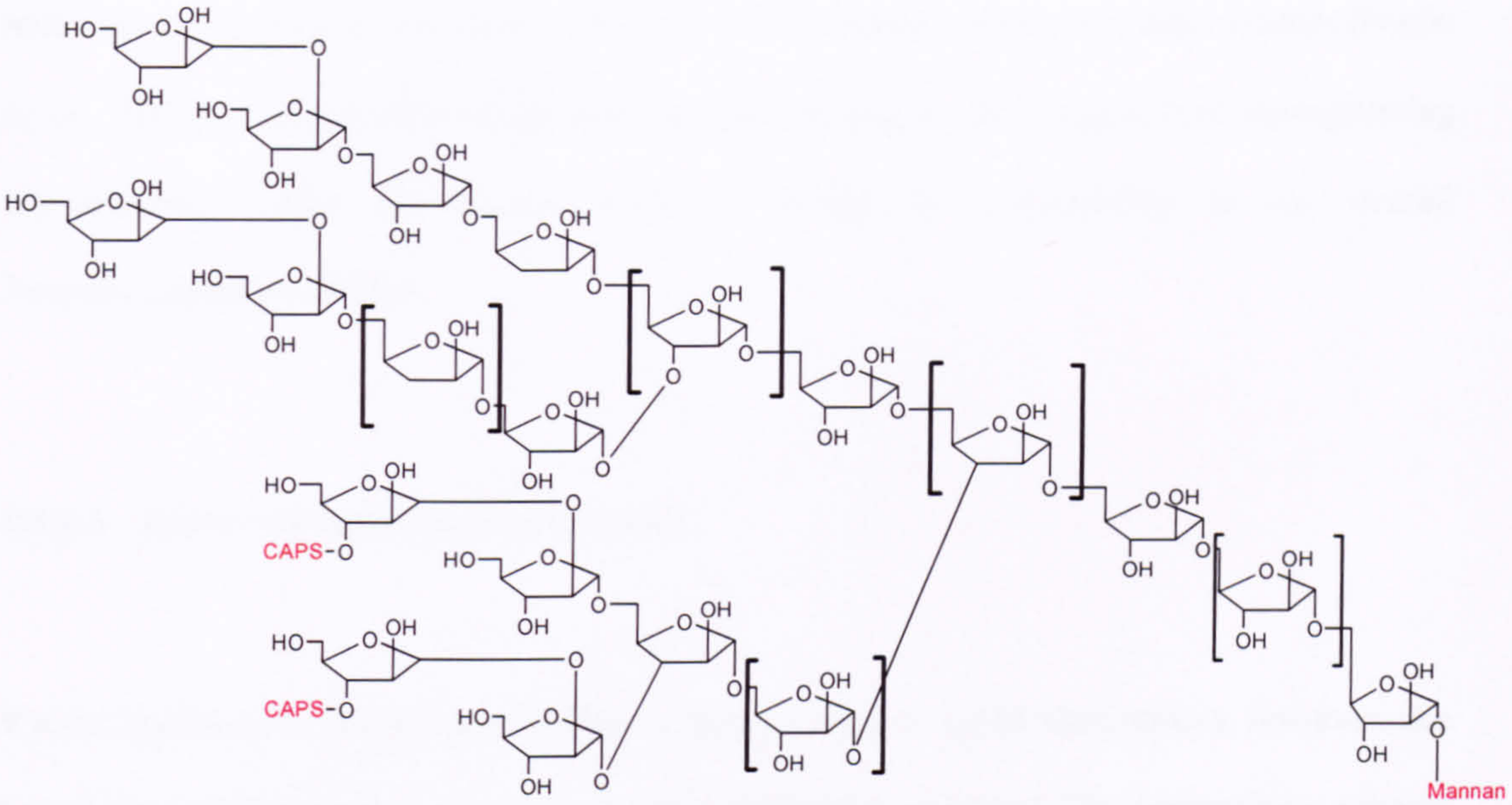


Figure 1.22 Arabinan segment of LAM. Showing the linkages of the arabinose units in the following configuration (Ara*f*-(β1→2)-Ara*f*-(α1→5)-Ara*f*-(α1→5)-Ara*f*-(α1→) and ([Ara*f*-(β1→2)-Ara*f*-(α1→]$_2$ →3 and →5)-Ara*f*-(α1→5)-(Ara*f*-(α1→), the positions of the caps and mannan are indicated

The terminal residues of the arabinan domains are decorated with various 'caps'. Two capping motifs have been identified for LAM (Nigou *et al.*, 2003). Mannooligosaccharides cap the LAM of slow-growing mycobacteria, such as *M. tuberculosis* (Chatterjee *et al.*, 1992) which have now been termed as ManLAM. Phosphoinositide units cap LAM from fast-growing mycobacteria, such as *M. smegmatis*. The LAM of *M. chelonae* has been shown to be devoid of any caps and has been termed AraLAM (Guerardel *et al.*, 2002). ManLAM and phospho-*myo*-inositol-capped LAM (PILAM) exhibit a broad spectrum of immunomodulatory activities (Chatterjee & Khoo, 1998; Nigou *et al.*, 2002b; Strohmeier & Fenton, 1999). The main immune response activities of ManLAM derive from its ability to inhibit the activation of macrophages and to inhibit the production of the Th1 pro-inflammatory cytokines IL-12

and TNF-α. *M. tuberculosis* ManLAM is also able to induce apoptosis in macrophages (Nigou *et al.*, 2003). Consequently, ManLAM has been linked to the longevity of slow-growing mycobacteria within the human reservoir through its contribution to the overall immunosuppressive effect.

1.4.4.1 Lipoarabinomannan biosynthesis

Recent studies have elucidated a number of the key steps in LAM biosynthesis, however, our knowledge of its complete synthesis is still somewhat fragmented. The biosynthetic scheme for the biosynthesis of this membrane molecule has been proposed as PI → PIM → LM → LAM (Belanger & Inamine, 2000; Besra *et al.*, 1997; Nigou *et al.*, 2003) and is illustrated in Figure 1.23. Biosynthesis of the PI anchor involves a number of essential genes that have been identified in *M. tuberculosis*. *Myo*-inositol synthesis is achieved through (Nigou *et al.*, 2003), an inositol-1-phosphate synthase in *M. tuberculosis* (INO1, Rv0046c) (Bachhawat & Mande, 1999).

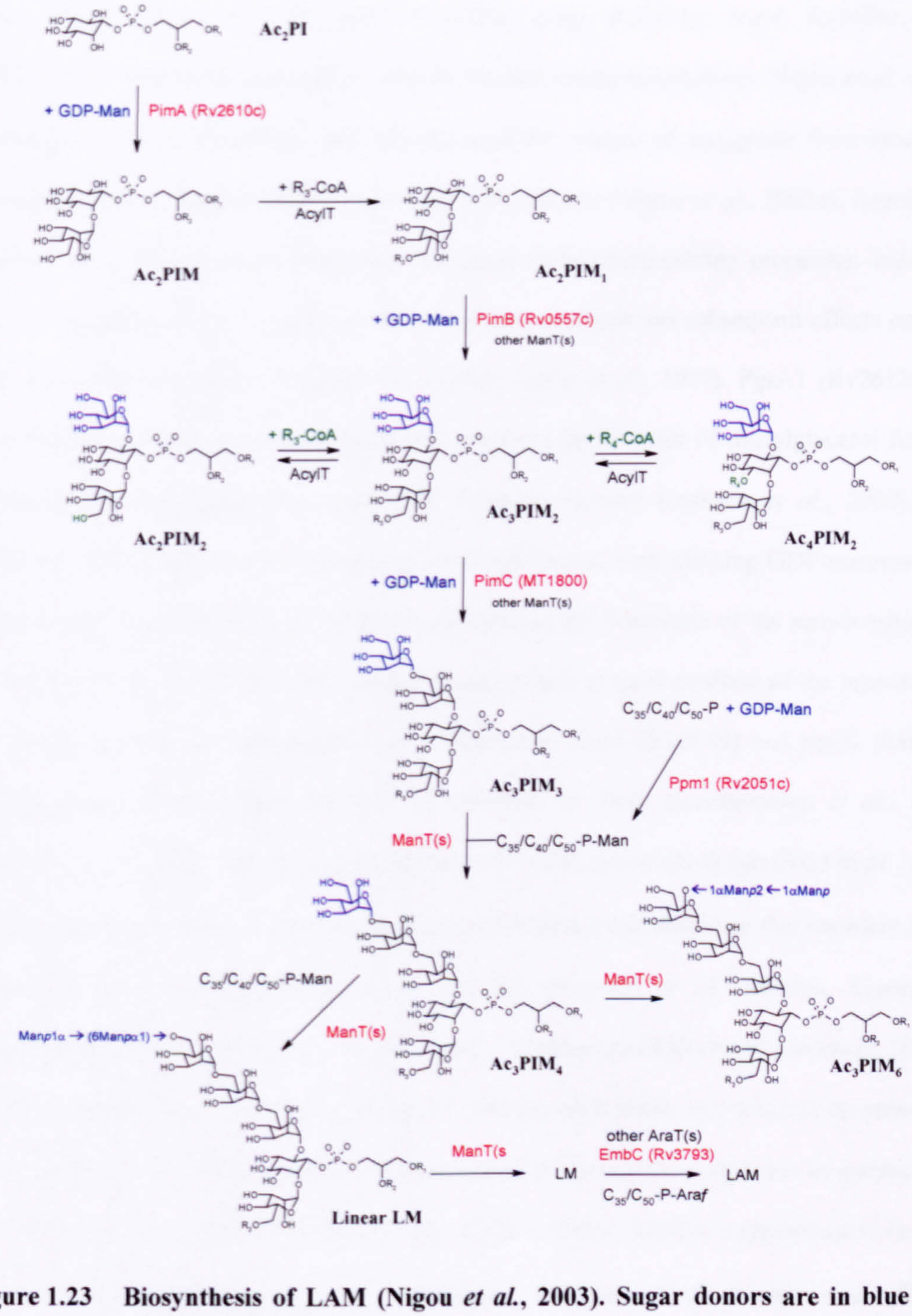


Figure 1.23 **Biosynthesis of LAM (Nigou *et al.*, 2003). Sugar donors are in blue and identified glycosyltransferases in red. AcylT, acyltransferase; ManT(s), mannosyltransferase(s); AraT(s), arabinosyltransferase(s); C_{35}/C_{50}, polyprenols; C_{35}/C_{50}-P-Man, polyprenol-monophosphorylmannose; C_{35}/C_{50}-P-Ara*f*, polyprenol-monophosphoryl-β-D-Ara*f*; R_n, Fatty acid group**

Four related ORF's (Rv3137; *suhB*, Rv2701c; *cysQ*, Rv2131c; *impA*, Rv1604) in *M. tuberculosis* have been suggested to encode inositol monophosphatases. Nigou *et al.* (2003) characterised *suhB* (Rv2701c) and demonstrated the release of phosphate from inositol-1-phosphate and, to a lesser degree other polyol phosphates (Nigou *et al.*, 2002a). Recently an isolated *impA* mutant of *M. smegmatis* possessed altered permeability properties, indicating that the synthesis of the PI anchor and its upstream synthesis has subsequent effects on PIM, LM and LAM biosynthesis (Nigou *et al.*, 2003; Parish *et al.*, 1997). PgsA1 (Rv2612c) has been identified as an essential enzyme that catalyses the transfer of diacylglycerol from its activated cytosine diphosphate nucleotide donor to inositol (Jackson *et al.*, 2000). PIM synthesis is initiated by two distinct mannosyltransferases, both utilising GDP-mannose as a sugar donor. Mannose residues are first transferred to the 2-position of the *myo*-inositol ring of PI to form PIM_1, which is then further mannosylated at the 6-position of the *myo*-inositol ring yielding PIM_2. Recent studies have shown that *pimA* (Rv2610) and *pimB* (Rv0557) encode mannosyltransferases involved in synthesis of PIM_2 (Kordulakova *et al.*, 2002; Schaeffer *et al.*, 1999). The recently identified *pimC* of *M. tuberculosis* (gmt7625 in *M. bovis*), has been characterised as a GDP-Man-dependent α-mannosyltransferase that transfers Man*p* from GDP-Man to Ac_3PIM_2 to form Ac_3PIM_3 (Kremer *et al.*, 2002b). Subsequent mannosyltransferases utilise the sugar donor polyprenolmonophosporylmannose (PPM), which is synthesised *via* PPM (Gurcha *et al.*, 2002) to form linear α(1→6) LM by extending PIM_3 (Besra *et al.*, 1997). Mature LM is produced by further mannosylation to produce the α(1→2) branches of Man*p*. The biosynthesis of the arabinan domain is speculated to be very similar to that of AG using EmbC (Chatterjee & Khoo, 1998; Zhang *et al.*, 2003). Furthermore, complementation of an *embC* mutant with a hybrid gene encoding the N-terminus of EmbC and the C-terminus of EmbB resulted in LAM with a lower molecular weight than wild-type LAM. Structural studies involving enzyme digestion and mass

spectrometry analyses revealed that the arabinan of the 'LAM' formed in the complemented strain was a hybrid of AG and LAM (Zhang *et al.*, 2003). Finally, the synthesis of the mannose caps, a crucial feature of LAM, is still unresolved (Nigou *et al.*, 2003).

1.4.5 Other cell envelope lipids

A number of the complex lipids of the cell envelope are key virulence factors and their structures are based upon very characteristic methyl-branched long-chain acids and alcohols (Minnikin *et al.*, 2002). A variety of glycosylated forms of mycolic acids exist in the outer layer of the cell wall and are composed of mainly trehalose monomycolates (TMM) and dimycolates (TDM) (Figure 1.24). TMM is involved in the transfer of mycolic acids into the cell wall (Belisle *et al.*, 1997). TDM has been implicated in granuloma formation (Yamagami *et al.*, 2001). *M. tuberculosis* is characterised by a family of acylated trehalose molecules all based upon the basic structure of a diacyl trehalose unit with subsequent acylation at other sites to form the tri- (TAT) and penta- (PAT) forms of acylated trehaloses (Figure 1.24) (Besra *et al.*, 1992).

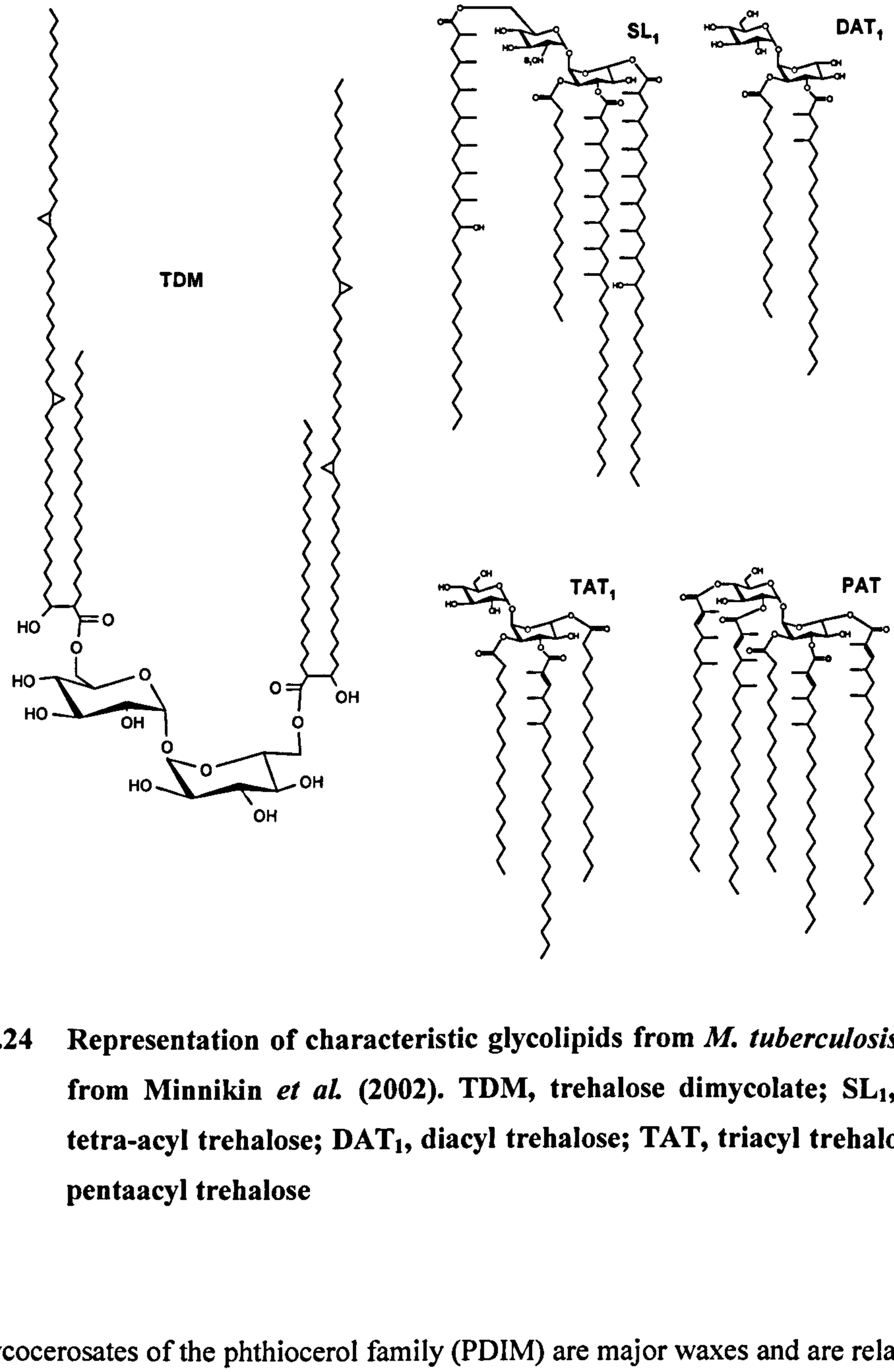


Figure 1.24 Representation of characteristic glycolipids from *M. tuberculosis* adapted from Minnikin *et al.* (2002). TDM, trehalose dimycolate; SL_1, sulfated tetra-acyl trehalose; DAT_1, diacyl trehalose; TAT, triacyl trehalose; PAT, pentaacyl trehalose

The dimycocerosates of the phthiocerol family (PDIM) are major waxes and are related to the phenolic glycolipids (Figure 1.25), which are found in *M. tuberculosis* strains and other related pathogenic mycobacteria (Minnikin *et al.*, 2002). PDIM of *M. tuberculosis* is composed of two mycocerosic acid chains attached to the phthiocerol backbone (Figure 1.25). These complex waxes are distinguished by their range of multimethyl-branched fatty acids

(and have been proposed to covalently interact with the mycolic acids of the outer membrane (Daffe & Draper, 1998).

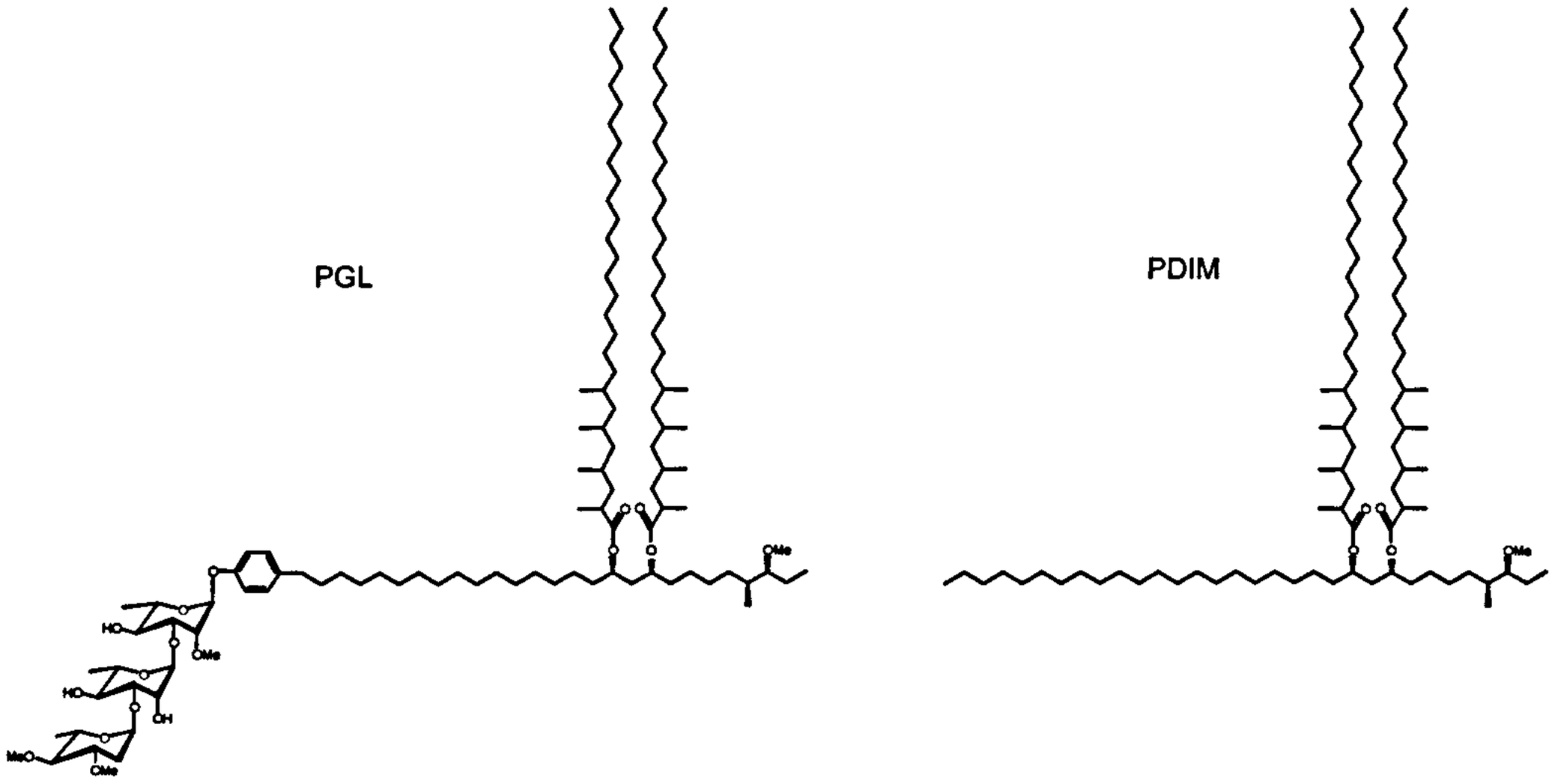


Figure 1.25 Structure of *M. tuberculosis* phenolic glycolipids and PDIM (Besra & Chatterjee, 1994; Watanabe *et al.*, 1994)

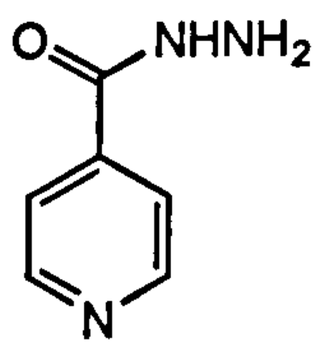
The biosynthesis of these complex lipids by a unique family of PKS (Minnikin *et al.*, 2002) was reviewed recently (Brennan, 2003).

1.5 Old and new drugs targeting *M. tuberculosis*

In 1914 Selman A. Waksman began a systematic screening of soil bacteria and fungi. In 1939, he discovered the marked inhibitory effect of certain 'fungi', especially actinomycetes, on bacterial growth. In 1940, he and his team were able to isolate an effective anti-TB antibiotic, actinomycin, unfortunately it was shown to be toxic in humans and animals. The first success came when streptomycin was purified from *Streptomyces griseus*, and showed maximum inhibition of *M. tuberculosis* with relatively low toxicity in animal tests. November 20, 1944,

was a historic date in the fight against TB because on this day the antibiotic was first administered to a critically ill TB patient with great success. A rapid succession of anti-TB drugs appeared in the following years. These were important because with streptomycin monotherapy, resistant mutants began to appear within a few months, endangering the success of antibiotic therapy. However, it was soon demonstrated that this problem could be overcome with the combination of two or three drugs. Currently there are five major antibiotics that are essential in the treatment of *M. tuberculosis*, RIF, INH, PZA, EMB and streptomycin. With the emergence of MDR-TB, new avenues for treatment are required.

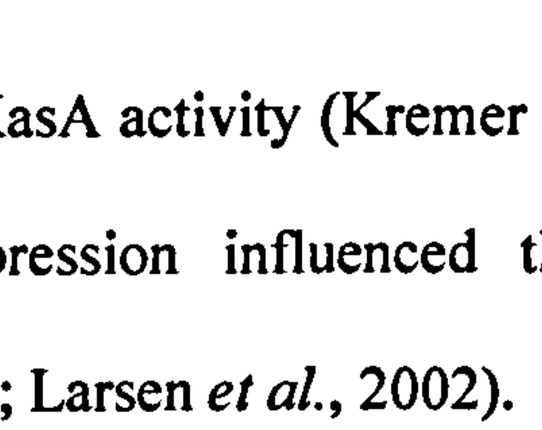
1.5.1 Isoniazid (INH)



INH is a synthetic, bactericidal agent that is used as a front-line drug against TB. INH was first used as an anti-tubercular agent in 1952, almost 50 years after its discovery it remains as one of the first lines of defence against TB (Heifets, 1994). Susceptible strains have an MIC of less than 0.05 μg/ml. Even though INH has been widely used against TB, little is known about the bacterial targets and mode of action of INH. In 1954 Middlebrook discovered that INH-resistant organisms had decreased catalase activity (Middlebrook, 1954). Recent genetic studies identified a catalase/peroxidase (*katG*) which restored susceptibility to INH-resistant strains of *M. tuberculosis*. This study also showed that 50-70 % of INH-resistant strains contained at least one mutation in KatG (Zhang *et al.*, 1992). Therefore, it was suggested that INH is a prodrug and requires activation by KatG. The mode of action of INH was originally

thought to be related to cell wall disruption. It was shown to be a direct inhibitor of mycolic acid biosynthesis (Winder 1982). The mode of action of INH is still being studied but it is widely accepted that INH targets fatty acid biosynthesis, primarily the enoyl-ACP reductase (InhA) (Marrakchi *et al.*, 2000). Activation of INH allows its product to mimic NADH and bind to InhA (Bernadou *et al.*, 2001). Resistant strains of INH have shown A→G transversion at position 280, resulting in a mutation Ser94→Ala, which alters the affinity of INH to NADH binding, thus resulting in resistance (Pantano *et al.*, 2002). As over-expression of *kasA* correlated with increased resistance to INH and the identification of single amino acid mutations in kasA conferring resistance to INH, led Slayden *et al.* to suggest KasA as a potential target for INH (Slayden *et al.*, 2000). However, it has been recently demonstrated by *in vitro* assays using purified InhA and KasA that KatG-activated INH, triclosan, and diazaborine inhibit InhA, but not KasA activity (Kremer *et al.*, 2003). It was also shown that neither *kasA* nor *kasB* over-expression influenced the MIC of INH against several mycobacteria (Kremer *et al.*, 2000b; Larsen *et al.*, 2002).

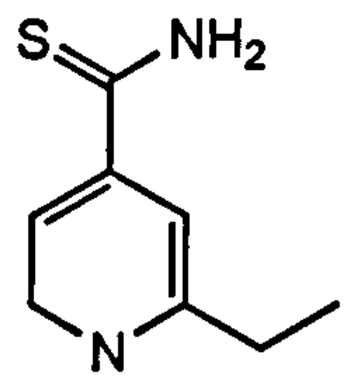
1.5.2 Triclosan



Triclosan (5-chloro-2-[2,4 dichlorophenoxy] phenol]) has also been shown to be an effective inhibitor of fatty acid biosynthesis (Kuo *et al.*, 2003; Marcinkeviciene *et al.*, 2001; McLeod *et al.*, 2001; Roberts *et al.*, 2003; Sivaraman *et al.*, 2003; Waller *et al.*, 2003). Three *M. smegmatis* mutants selected for resistance to triclosan all possessed mutations in InhA (McMurry *et al.*, 1999). Two of the mutations also expressed INH resistance (McMurry *et al.*,

1999). It was demonstrated that InhA, like its *E. coli* homolog FabI, is a target for triclosan (Kuo *et al.*, 2003; McMurry *et al.*, 1998; McMurry *et al.*, 1999; Parikh *et al.*, 2000). It is proposed to bind preferentially to the Enzyme-NAD^+ binary complex, in which the triclosan phenolic hydroxyl interacts with the hydroxyl group of Lys158 (Parikh *et al.*, 2000).

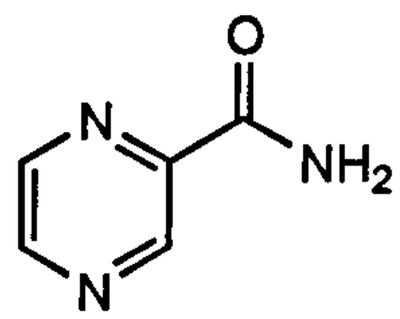
1.5.3 Ethionamide (ETH)



ETH is believed to have a similar mode of action as INH due to their structural similarity (Winder *et al.*, 1971). However, significant differences in their mode of action have been observed. Mutations in KatG were found not to confer resistance to ETH, therefore an additional enzyme is required for the activation of ETH into an activated form (Baulard *et al.*, 2000; DeBarber *et al.*, 2000). Recently, it was shown that ETH is activated by an enzyme with homology to monooxygenases, EthA (Rv3854c) (Baulard *et al.*, 2000; DeBarber *et al.*, 2000). When the neighbouring open reading frame, *ethR* (Rv3855), which is homologous to members of the TetR family of transcriptional repressors, was over-expressed in *M. smegmatis*, ETH resistance was observed (Baulard *et al.*, 2000; DeBarber *et al.*, 2000). In addition, chromosomal inactivation of this gene by transposition led to ETH hypersensitivity and EthR was proposed to regulate the production of EthA, which subsequently activates ETH (Baulard *et al.*, 2000). When the *ethA* promoter region was fused with a *lacZ* reporter gene over-expression of *ethR* in *trans* was found to cause a strong inhibition of *ethA* expression, independently of the presence of ETH in the culture media (Engohang-Ndong *et al.*, 2004).

Vannelli *et al.* (2002) establish that EthA is an FAD-containing enzyme that oxidizes ETH to the corresponding S-oxide. The S-oxide, which has a similar biological activity as ETH, is further oxidized by EthA to 2-ethyl-4-amidopyridine, presumably *via* the unstable doubly oxidized sulfinic acid intermediate (Vannelli *et al.*, 2002). It was found that purified EthA displays a remarkably low activity with the antitubercular prodrug ETH (Fraaije *et al.*, 2004).

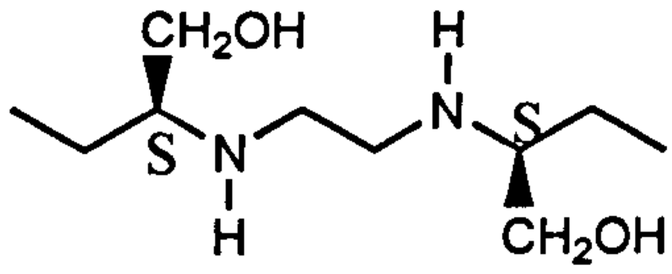
1.5.4 Pyrazinamide (PZA)



The antitubercular activity of PZA was first detected in 1952 but its description was overshadowed by the use of INH. It is a structural analogue of nicotinamide that targets semi-dormant tubercle bacilli under acidic conditions (Tatar, 1974). In the acidic environment of phagosomes, *M. tuberculosis* produces pyrazinamidase (Pzase), which converts PZA into pyrazinoic acid, the active compound which kills *M. tuberculosis*. Notably, *M. tuberculosis* isolates that show resistance to PZA have lost Pzase activity (Huang *et al.*, 2003). The gene that encodes pyrazinamidase in *M. tuberculosis* is *pncA* (Huang *et al.*, 2003). In clinical studies 72 % of PZA^R isolates have *pncA* mutations (Huang *et al.*, 2003). Point mutations and substitutions of Cys138 with Ser, Gln141 with Pro, Asp63 with His, and also deletions of G nucleotide at position 162 and 288 have resulted in defective Pzase activity (Bamaga *et al.*, 2002). Naturally-resistant strains of *M. bovis* lack Pzase activity (Raynaud *et al.*, 1999). To date the cellular target for PZA has not been identified although the apparent similarity of PZA to nicotinamide suggests that enzymes involved in pyridine nucleotide biosynthesis or fatty acid biosynthesis are possible targets (Rattan *et al.*, 1998). Recently, it has been found that FAS-I from *M. avium*, *M. bovis* BCG and *M. tuberculosis* confers resistance to 5-Chloro-

PZA when present on multi-copy plasmids in *M. smegmatis* (Zimhony *et al.*, 2000). PZA was also shown to inhibit FAS-I in *M. tuberculosis* therefore implying FAS-I as a primary target of PZA (Zimhony *et al.*, 2000).

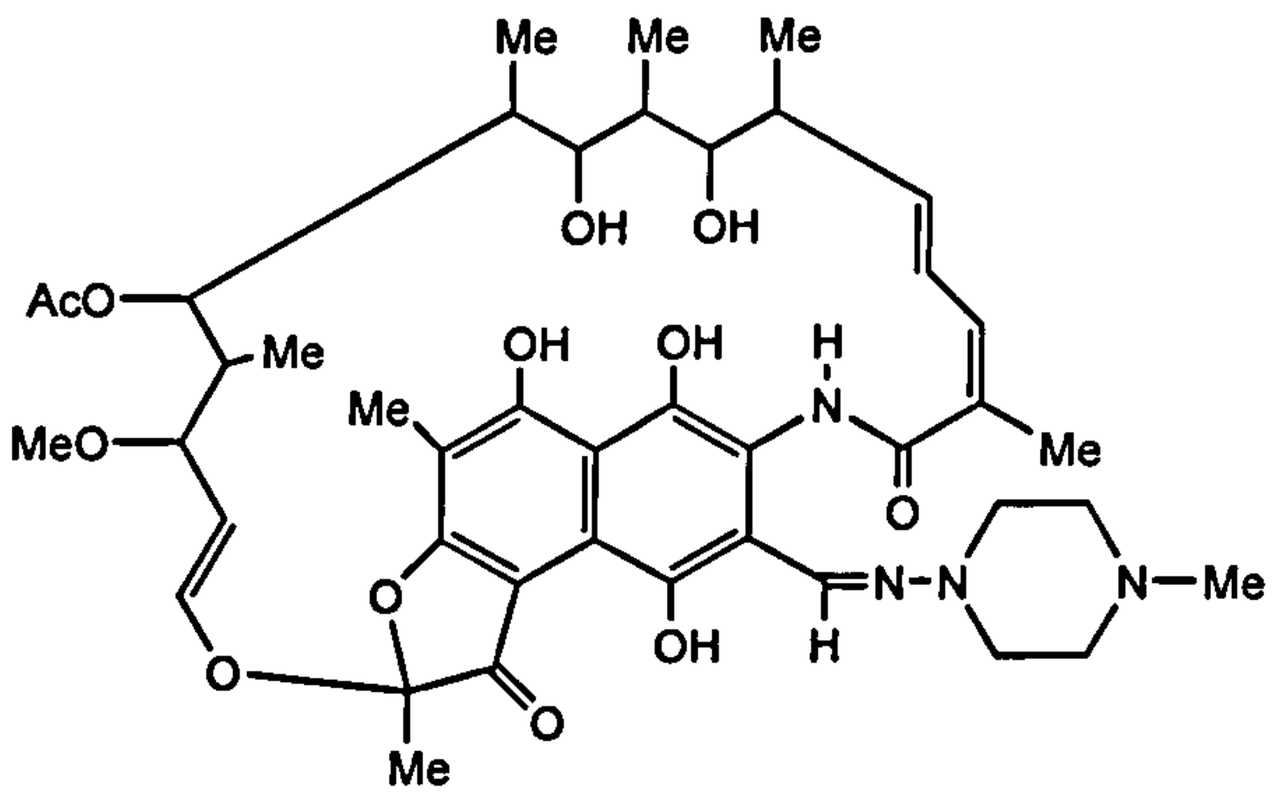
1.5.5 Ethambutol (EMB)



EMB ([S,S']-2,2'-[ethylenediimino])di-1-butanol is one of the front line drugs used against TB. It has a bacteriostatic effect that is thought to be important in preventing the emergence of resistance to the front-line agents. Early studies reported that EMB inhibited incorporation of mycolic acids into the cell wall (Mikusova *et al.*, 1995). The cellular target for EMB was sought using drug resistance, *via* target over-expression by a plasmid vector. This strategy led to the cloning of the *M. avium emb* region which, on over-expression in *M. smegmatis* rendered the otherwise susceptible host resistant to EMB. This data suggests that the main target of EMB in *M. tuberculosis* may be *emb* which encode arabinosyltransferases involved in AG and LAM biosynthesis (Belanger *et al.*, 1996). More recent genetic approaches in relation to *M. tuberculosis* have revealed the products of three homologous open reading frames designated *embC, embA, embB* as EMB targets (Telenti *et al.*, 1997). The EmbCAB proteins are believed to be membrane proteins involved in the synthesis of various motifs of AG and LAM (Escuyer *et al.*, 2001; Telenti *et al.*, 1997). Recently it has been shown that synthesis of LAM, but not AG, ceases after inactivation of *embC* in *M. smegmatis* with LAM synthesis being restored upon complementation with *embC* (Section 1.4.2.1) (Zhang *et al.*, 2003). Mutations in codon 306 of *embB* producing Met to Ile or Val were found in 13 of 28

EMB-resistant strains. Strains with Met306→Leu or Met306→Val were more resistant to EMB than those with Met306→Ile substitutions, consistent with the notion that amino acid substitutions in EmbB alter drug-protein interactions and thereby lead to resistance (Sreevatsan *et al.*, 1997). In another study 70 % of EMB^R isolates all had mutations in the EmbB protein.

1.5.6 Rifampin (RIF)



Rifampin, a lipophilic ansamycin, was first introduced as a theraputic agent against TB in 1967 after being discovered as a product of *Streptomyces mediterranei* (Clark & Wallace, 1967). It was shown to be highly active, diffusing rapidly across the hydrophobic cell envelope. In most strains of *M. tuberculosis* resistant to RIF, mutations were found in a short region of 27 codons situated near the centre of *rpoB*, which encodes the RNA polymerase β–sub-unit (Taniguchi *et al.*, 1996). RNA polymerase is an oligomer of four sub-units designated α, β, β' and σ (encoded by *rpo*A, *rpo*B, *rpo*C and *rpo*D). Double mutations and deletions were observed with 70 % of all mutations of clinical resistant *M. tuberculosis* strains possessing mutations at His526 and Ser531, similar to those observed in *E. coli* (Taniguchi *et al.*, 1996) (Figure 1.26).

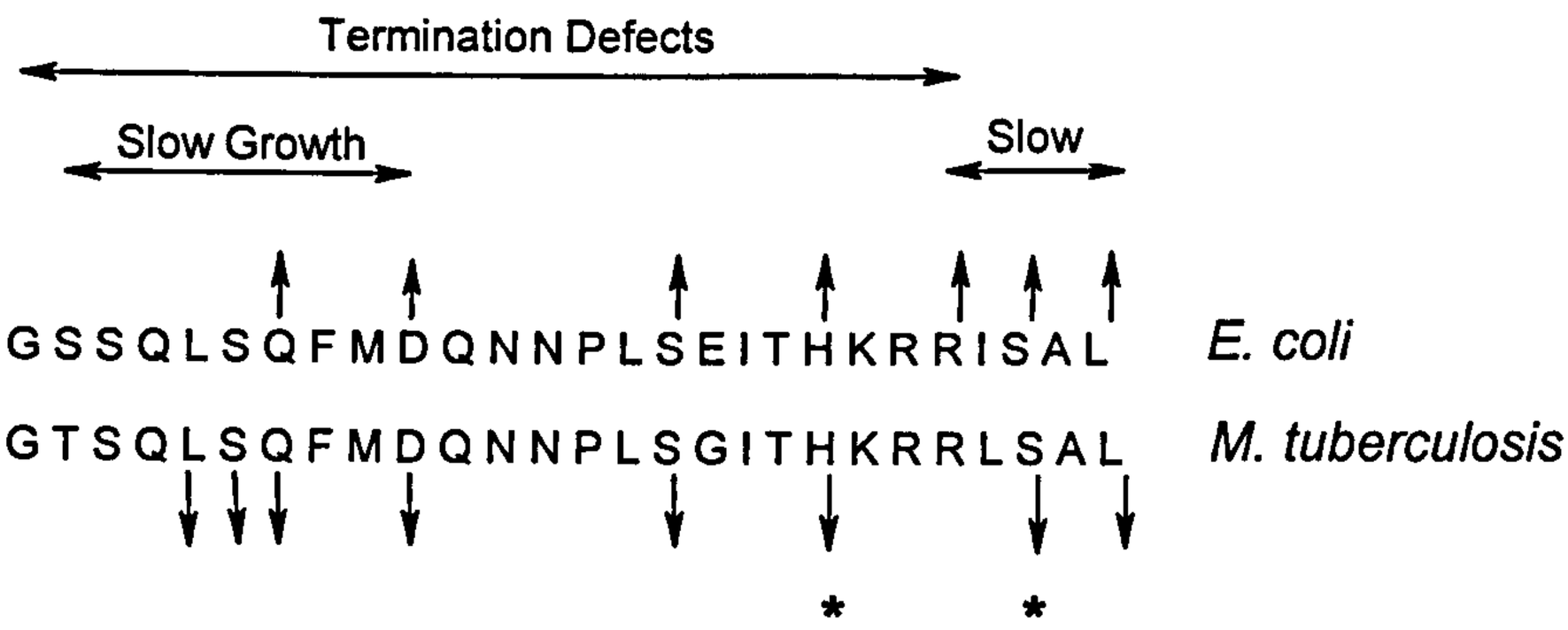
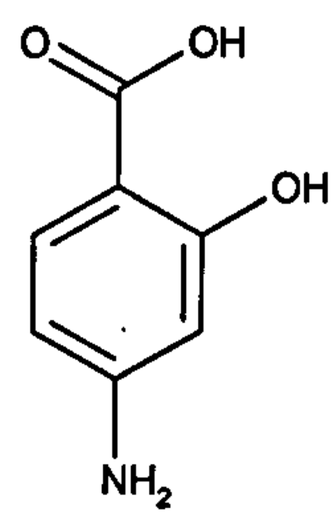


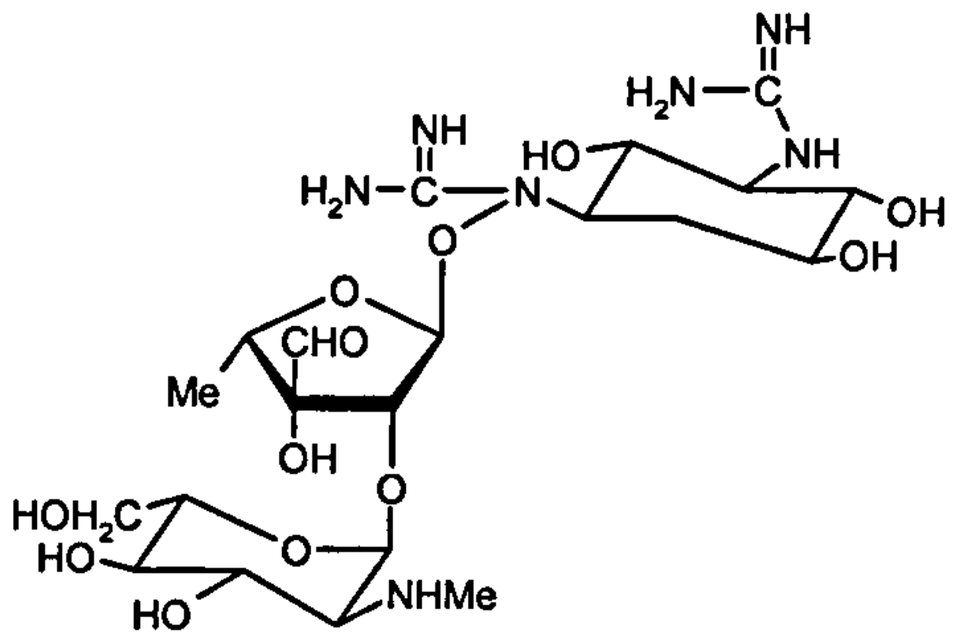
Figure 1.26 Region of *rpoB* gene involved in RIF resistance

1.5.7 *para*-Aminosalicylic acid



para-Aminosalicylic acid displays high activity against *M. tuberculosis* but, not against other mycobacterial strains and bacteria in general. It is rarely used except in the case of MDR-TB. It has been suggested that it inhibits salicylate-dependent biosynthesis of iron-chelating mycobactins (Ratledge & Dover, 2000).

1.5.8 Streptomycin

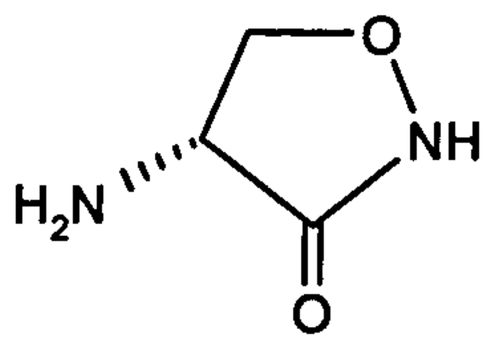


Streptomycin was the first drug to be used effectively against TB in 1944 after its discovery by Selman Waksman in 1943 (Davies & Yew, 2003). Resistance to streptomycin was observed within a few months of its first use but the coupling of two or more anti-tubercular drugs solved the problem at the time. Due to its toxicity streptomycin was more or less excluded from front-line therapy, as less toxic and more active drugs were being discovered (Girling, 1989). The site of action of streptomycin is the small 30S subunit of the bacterial ribosome, which comprises the ribosome protein S12 and the 16S rRNA (Finken *et al.*, 1993). Resistance to streptomycin has been investigated in a number of different bacterial systems resulting in three known mechanisms of resistance being established.

The first mechanism involves mis-sense mutation in the *rpsL* gene, which encodes the S12 protein in *M. tuberculosis* (Heym *et al.*, 1994; Musser, 1995; Ramaswamy & Musser, 1998). These mutations effectively change the binding site or the conformation of the ribosome so that streptomycin is ineffective. In about 70 % of resistant clinical isolates an A→G transition occurred in codon 43 replacing Lys with Arg (Heym *et al.*, 1994). This same mutation was also observed in plant and other bacteria showing streptomycin resistance (Galili *et al.*, 1989). These substitutions have also been observed in *M. tuberculosis rpsL* but these mutations are less frequent. The second mechanism of resistance involves mutations in the *rrs* gene,

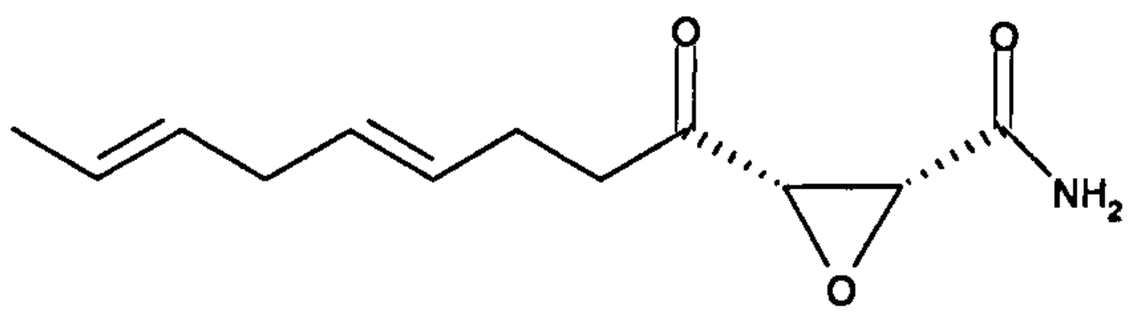
encoding 16S rRNA. A number of nucleotides substitutions have been implicated in this resistance mechanism but most abundant are at the 904 and 530 animo acids. Most common is the mutation of cytosine 904 to adenine or guanosine, or adenine 905 to guanosine (Heym *et al.*, 1994). Little is known about the third and final mechanism of resistance, in these resistant isolates neither *rrs* nor *rpsL* are mutated.

1.5.9 D-Cycloserine



D-Cycloserine a structural analogue of D-alanine, is used as a second-line agent in chemotherapy (Caceres *et al.*, 1997). It inhibits biosynthesis of the mAGP complex and has been shown to target *alrA*, which encodes the D-alanine racemase involved in PG biosynthesis (Figure 1.7). Resistance has been correlated when a single transversion (G→T) located in the *alrA* promoter region (Caceres *et al.*, 1997).

1.5.10 Cerulenin



Cerulenin (*(2R)(3S)-2,3-epoxy-4-oxo-7,10-dodecadienoly-amide*) is a fungal product isolated from *Caephalosporium caerulens* (Omura, 1976) which has been shown to be an irreversible

inhibitor of *E. coli* FabB and FabF (Magnuson *et al.*, 1993). Cerulenin covalently modifies the KAS active site. The crystal structures of FabF (Moche *et al.*, 1999) and FabB (Price *et al.*, 2001) co-crystallised with cerulenen provided evidence that cerulenin forms a covalent adduct with the active site cysteine residue. Incubation of KAS with acyl-ACP protects the enzyme from cerulenin therefore supporting the theory that a covalent adduct occurs with the key cysteine residue in the fatty acyl-binding site as seen in the *E. coli* co-crystallisation studies and therefore creates an irreversible interaction (Figure 1.27) (Moche *et al.*, 1999).

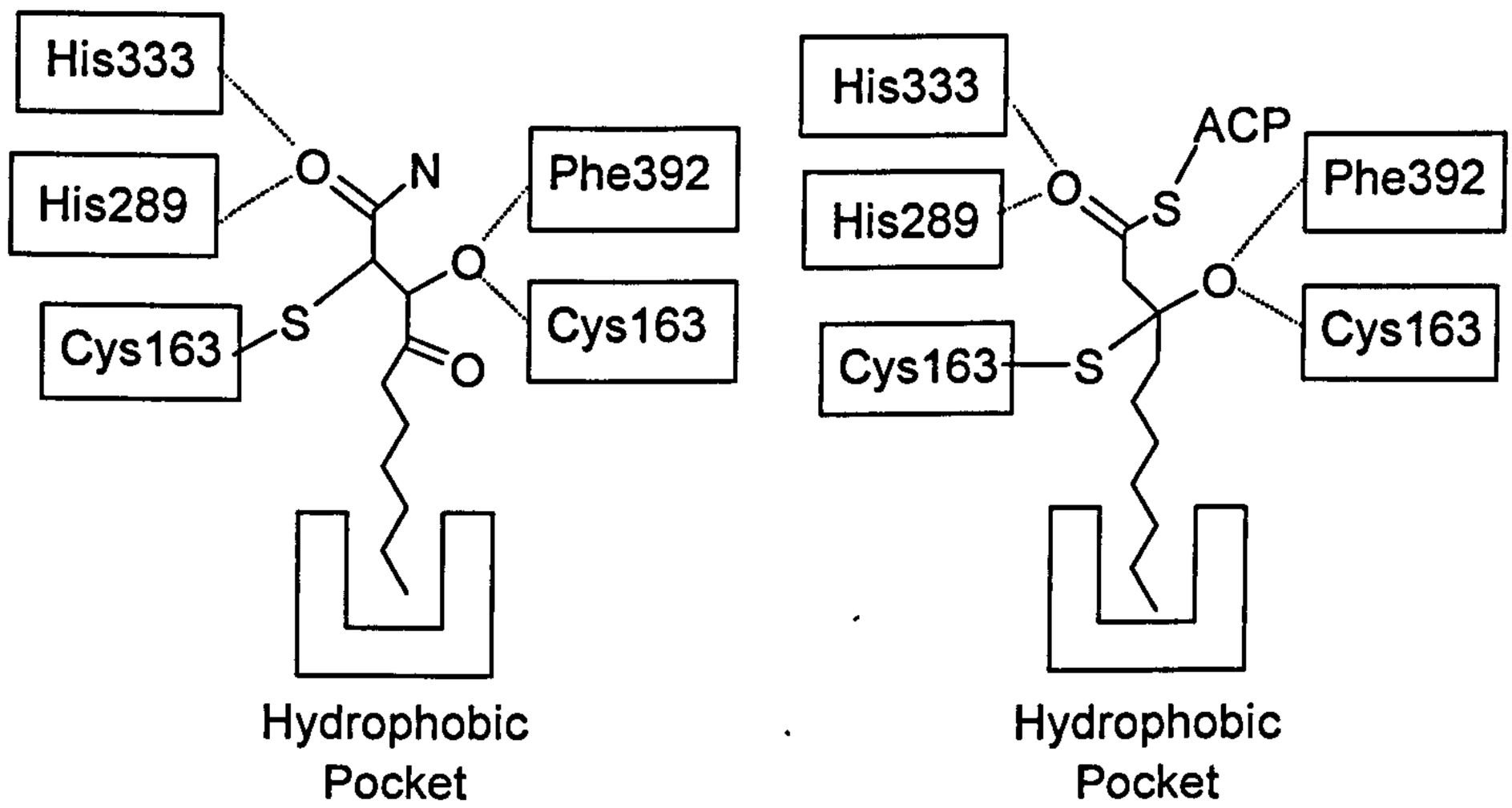
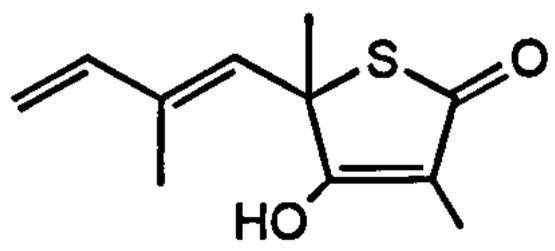


Figure 1.27 **Interaction of FabF and cerulenin (left) and condenstation transition state of acyl-ACP (right) adapted from Price *et al.* (2001). Cerulenin mimics the condensation transition state and spans the two halves of the active site. The oxygen of cerulenin lies in the oxyanion hole formed by the Cys163 and Phe392. The Cys163 rotates and forms a thioester bond with the carbon at position 2 carbon and the acyl-chain of cerulenin mimics the acyl chain of the donor and accommodates the hydrophobic pocket**

The tail of cerulenin occupies a long hydrophobic cavity, which normally contains the growing acyl chain of the natural substrate. This may explain why cerulenin does not inhibit

the *E. coli* β-ketoacyl ACP synthase III (FabH), as this acetyl-CoA-primed enzyme lacks a fatty acyl-ACP binding site. The IC_{50} values of the purified *E. coli* KAS proteins were reported as 3 and 20 μM for FabB and FabF respectively (Price *et al.*, 2001). Recently, it has been shown that the IC_{50} value of cerulenin against KasA is 0.67 μM (Kremer *et al.*, 2002c). Cerulenin has been shown to be an effective inhibitor of KAS proteins but it has also been suggested that cerulenin inhibits the mammalian FAS-I enzymes and could therefore be toxic (Vance *et al.*, 1972).

1.5.11 Thiolactomycin



Thiolactomycin (*(4S)(2E, 5E)–2,4,6–trimethyl-3-hydroxy-2,5,7-octatriene-4-thiolide*) (TLM) is a unique thiolactone antibiotic isolated originally from a *Nocardia sp* (Oishi *et al.*, 1982) isolated from soil. It has proven to be an effective inhibitor of FAS-II, but does not inhibit FAS-I (Slayden *et al.*, 1996). This is a very attractive property as the drug should be able to be used as a chemotherapeutic agent against bacterial infections if it does not have any adverse effects on the human FAS-I. TLM and related thiolactones have been shown to be potent inhibitors of FAS-II and mycolic acid biosynthesis in mycobacteria (Slayden *et al.*, 1996), as well as of chloroplast type II fatty acid synthesis in *Plasmodium falciparum* (Waller *et al.*, 2003). The use of *M. smegmatis* cell extracts demonstrated that TLM specifically inhibited the mycobacterial FAS-II but not the multifunctional FAS-I (Slayden *et al.*, 1996). The *in vivo* and *in vitro* data and knowledge of the mechanism of TLM resistance in *E. coli* suggests that two distinct TLM targets exist in mycobacteria, the β-ketoacyl-acyl carrier protein synthases

involved in the FAS-II-driven elongation steps leading to the synthesis of meromycolates (Slayden *et al.*, 1996). The thiolactone ring of TLM mimics the bent conformation of thiomalonate. The oxygen forms a hydrogen bond with His-298 and His-333, and the carbons of the malonate are mimicked by the carbon ring of TLM. The oxygen of TLM points out of the active site tunnel that would be occupied by the pantetheine arm of the malonyl-ACP substrate (Figure 1.28) (Price *et al.*, 2001). The *E. coli* strain CDM5 was derived as a TLM-resistant mutant but remained sensitive to cerulenin. The fatty acid synthase activity in extracts from strain CDM5 was sensitive to TLM inhibition therefore the TLM resistance was attributed to a different allele *emrB*. Disruption of the *emrB* gene converted strain CDM5 to a TLM-sensitive strain, and the over-expression of the *emrAB* operon conferred TLM resistance to sensitive strains. Thus, activation of the Emr efflux pump is the mechanism for TLM resistance in strain CDM5 (Furukawa *et al.*, 1993). A similar strain with an efflux KO mutation remained resistant to TLM. Further investigation of this strain revealed another missense mutation in the *fabB* gene (T1168G) that directed the expression of a mutant protein, FabB Phe390→Val. Structural modeling predicted that the CG2 methyl group of the valine side chain interfered with the positioning of the C11 methyl on the isoprenoid side chain of TLM in the binary complex, whereas the absence of a bulky methyl group on the leucine side chain permitted TLM binding (Jackowski *et al.*, 2002).

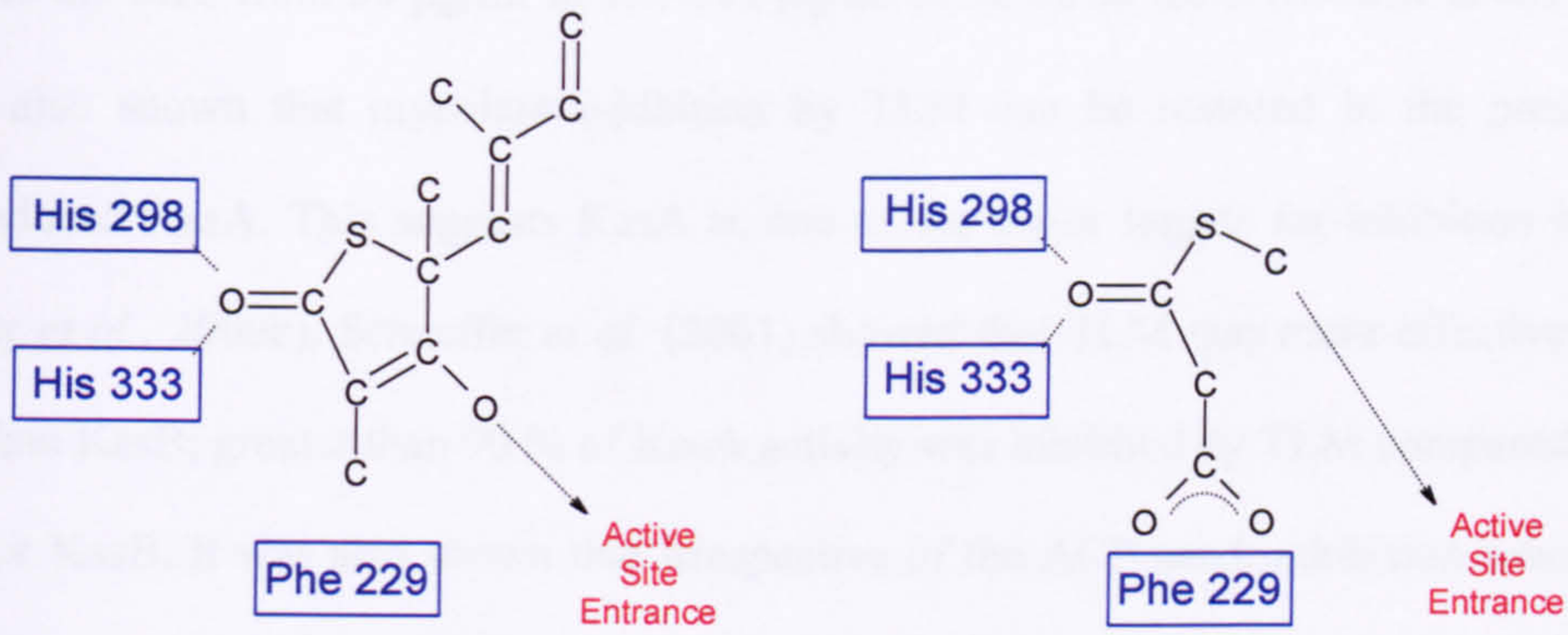
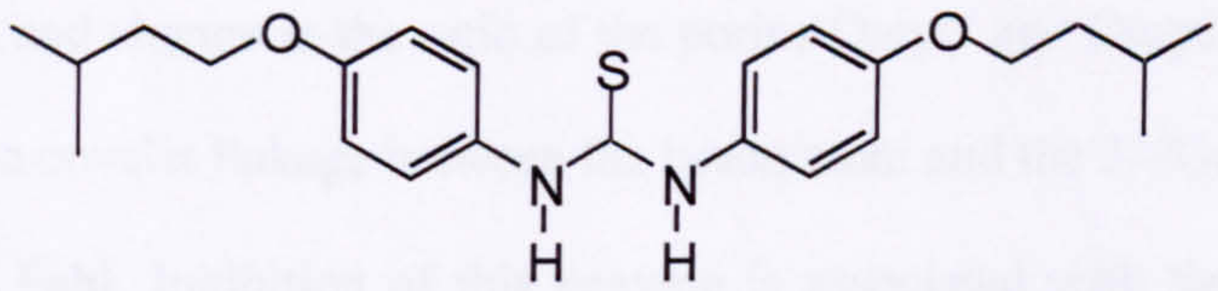


Figure 1.28 Schematic diagram illustrating how TLM mimics substrates in the active site of FabB. The thiolactone ring of TLM mimics the bent conformation of thiomalonate. The oxygen forms a hydrogen bond with His-298 and His-333, and the carbons of the malonate are mimicked by the carbon ring of TLM. The oxygen of TLM points out of the active site tunnel that would be occupied by the pantetheine arm of the malonyl-ACP substrate

In bacterial type II systems, as typified by that of *E. coli*, the first step is catalysed by the condensation of acetyl-CoA and malonyl-ACP by β-ketoacyl-ACP synthase III (FabH). Recently, Choi *et al.,* (2000b) identified an open reading frame in *M. tuberculosis* (Rv0553c), now termed mtFabH that contained the Cys-His-Asn catalytic triad signature (Choi *et al.*, 2000b). In contrast to *E. coli* FabH, mtFabH preferred long-chain acyl-CoA substrates, rather than an acetyl-CoA primer. In addition, mtFabH was sensitive to TLM and resistant to cerulenin *in vitro*. Interestingly, *mtfabH* over-expression in *M. bovis* BCG did not confer TLM resistance, suggesting that mtFabH may not be the primary target of TLM *in vivo* and surprisingly led to several unusual changes in the lipid composition of the cell wall. TLM is effective against these enzymes as they possess a malonyl-ACP binding site, which is targeted by TLM. It has been shown that the interaction between KAS and TLM is reversible (Price *et*

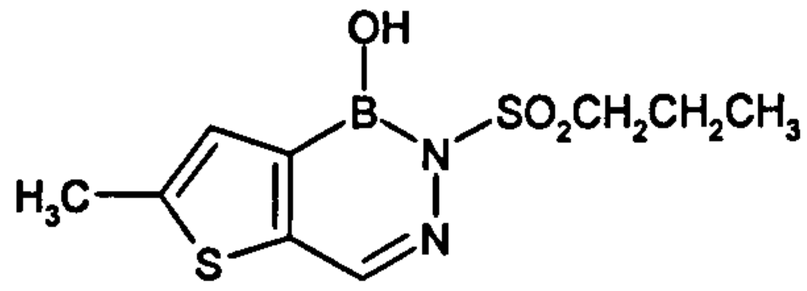
al., 2001). Preliminary experiments have shown that the over-production of KasA and KasB increases the MIC from 30 μg/ml to 150-200 μg/ml in *M. bovis* BCG (Kremer *et al.*, 2000b). It was also shown that mycolate inhibition by TLM can be restored in the presence of overproduced KasA. This suggests KasA as one of the major targets for inhibition by TLM (Kremer *et al.*, 2000b). Schaeffer *et al.* (2001) showed that TLM was more effective against KasA than KasB; greater than 90 % of KasA activity was inhibited by TLM compared to only 65 % for KasB. It was also shown that irrespective of the ACP used, inhibition was similar, suggesting that the mechanism for inhibition for each enzyme is similar and does not involve any interaction with the ACP but still inhibits the decarboxylation (Schaeffer *et al.*, 2001).

1.5.12 Isoxyl



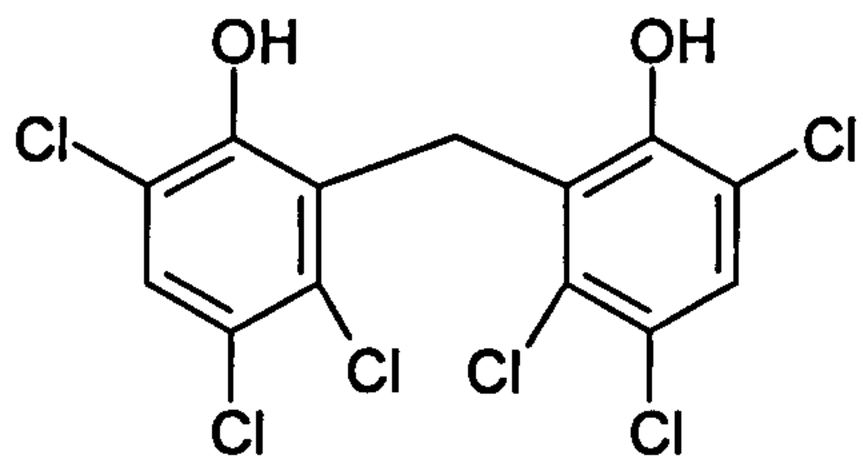
Isoxyl was first used in the clinical treatment of tuberculosis in the 1960s (Winder, 1982). Phetsuksiri *et al.* (1999) demonstrated that isoxyl has potent activity against *M. aurum* A^+, *M. avium* and *M. tuberculosis.* In addition, 12 clinical isolates of *M. tuberculosis* from different geographical areas and with various drug susceptibility profiles were found to be highly sensitive to isoxyl. In the same study a panel of 28 isoxyl analogues were synthesised with simple alkyl and alkoxy substituents which displayed various degrees of activity against *M. tuberculosis*. Isoxyl's mode of action was thought to be similar to INH and ETH in that it inhibits both fatty acid and mycolic acid biosynthesis (Phetsuksiri *et al.*, 1999). However, recent evidence suggests that isoxyl is able to interfere with oleic acid and tuberculostearic acid synthesis by inhibiting the acyl-CoA desaturase DesA3 (Phetsuksiri *et al.*, 2003).

1.5.13 Diazoborine



Diazoborine (DZB) is a class of boron-containing compounds that are only used experimentally, due to the inherent toxicity of boron. DZB have been shown to inhibit polysaccharide biosynthesis in *E. coli.* This effect is thought to be due to inhibition of fatty acid synthesis (Grassberger *et al.*, 1984) by blocking the active site of FabI (the NADH-dependent enoyl-ACP reductase) (Bergler *et al.*, 1994; Turnowsky *et al.*, 1989). Incubating a temperature-sensitive *E. coli envM* mutant at the nonpermissive temperature caused effects on the cells similar to those caused by treatment with diazaborine, i.e., inhibition of fatty acid, phospholipid, and lipopolysaccharide biosynthesis, induction of a 28,000-dalton inner membrane protein, and change in the ratio of the porins OmpC and OmpF (Turnowsky *et al.*, 1989). DZB forms a covalet linkage between the boron atom and the 2'-hydroxyl of the ribose cofactor in *E. coli* FabI. Inhibition of this enzyme is associated with the orientation of the inhibitor within the channel that would normally allow the phosphopanthene of ACP to access the active site (Baldock *et al.*, 1996). Homologues of the FabI of *M. tuberculosis,* such as InhA should be sensitive to DZB but further work is required to analyse this hypothesis. Although the MIC value of DZB for *M. smegmatis* (pMV261) was significantly higher (MIC = 100 μg/ml) compared with that of other drugs, *M. smegmatis* carrying pMV261-*inhA* showed a reproducible 2-fold increase in resistance against this drug over the strain carrying the empty vector. This suggests that InhA may be a target of DZB, in agreement with the fact that DZB inhibits the *E. coli* enoyl-ACP reductase encoded by the *fabI* gene (Kremer *et al.*, 2003).

1.5.14 Hexachlorophene



The biphenyl inhibitor, hexachlorophene, has been shown to inhibit FabI of *Haemophilus influenzae* with an IC_{50} of 2.5 +/- 0.4 μM. Steady-state inhibition patterns did not allow the mode of inhibition to be unambiguously determined, but binding kinetics suggested that free enzyme, enzyme-NAD^+, and enzyme-crotonyl-CoA complexes have similar affinity for hexachlorophene. When the enzyme-NADH complex was mixed with hexachlorophene, concentration-independent fluorescence quenching at 480_{nm} was observed, suggesting at least partial competition between NADH and hexachlorophene for the same binding site (Marcinkeviciene *et al.*, 2001).

1.5.15 Summary of antimycobacterial agents and their targets

Various antitubercular drugs of historical importance have proven to be effective inhibitors of *M. tuberculosis* cell envelope biosynthesis, but with the onset of MDR-TB, investigation of novel cellular targets is required. Table 1.4 summarises the current knowledge of several mycobacterial agents, their proposed targets and proposed activators.

Table 1.4 MIC for several inhibitors against various mycobacteria, na = not available.

Inhibitor	*M. tuberculosis* H37Rv MIC (μg/ml)	*M. smegmatis* MIC (μg/ml)	*M. bovis* BCG MIC (μg/ml)	Known/proposed Target(s)	Proposed Activator
INH	0.1	5	0.05	InhA	KatG
TRC	7.5	5	na	InhA	
ETH	10	20	10	InhA	EthA
PZA	20	25	25	FAS-I	PncA
EMB	4	0.5	2	EmbCAB	
RIF	0.5	17	0.025	*rpoB*	
Para-Amino	na	na	na	MbtA	
STR	2	1	0.25	*rpsL, rrs*	
D-Cyc	na	75	na	*alrA*	
CER	3 - 6.25	0.5	1.5	KasA, KasB	
TLM	5	75	30	KasA, KasB, mtFabH	
ISO	2.5	600	1	DesA3	EthA
DZB	50	100	na	InhA	
HEX	na	2	na	InhA	

1.6 Project aims

The major aims of the project are to investigate a number of enzymes involved in mycolic acid biosynthesis in *M. tuberculosis* by biochemical characterisation. More specifically we aimed to establish methods for the effective expression and purification and biochemical characterisation of:

- mtFabH – β-ketoacyl-ACP synthase III, the pivitol link between FAS-I and FAS-II
- KasA – β-ketoacyl-ACP synthase I, involved in meromycolate extension *via* extension of acyl chains with malonyl-AcpM

- AcpM – Meromycolate extension ACP, essential carrier for lipid metabolism
- AccD4 – Acyl-CoA carboxylase proposed to be one of the enzyme involved in carboxylation of acyl chains immediately before mycolic acid condensation
- Pks13 – Polyketide synthase 13, proposed to be the Claisen condensation enzyme involved in mycolic acid biosynthesis

The biochemical properties of the purified proteins will be assessed *in vitro* along with the development of assays to investigate partial reactions, such as the decarboxylation and transacylation activites of β-ketoacyl-ACP synthases. In addition, their substrate specificity, targeting by various inhibitors and amino acid function will be examined by site-directed mutagenesis. Finally, the biochemical properties of Pks13 will be assessed as a possible Claisen-condensation synthase involved in the production of mature mycolic acids.

CHAPTER 2

β-Ketoacyl–ACP synthase III (mtFabH)

2.1 Introduction

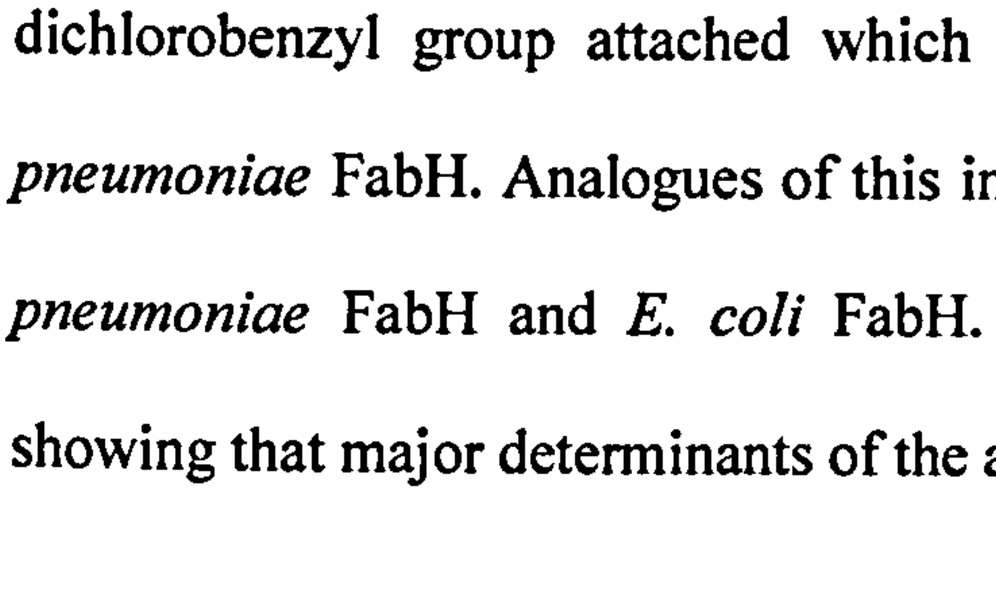
The increase in drug-resistant pathogenic bacteria has created an urgent demand for new antibiotics. Among the more attractive targets for the development of new antibacterial compounds are the enzymes of fatty acid biosynthesis. Although a number of potent inhibitors of microbial fatty acid biosynthesis have been discovered, few of these are clinically useful drugs. Several of these fatty acid biosynthesis inhibitors have potential as lead compounds in the development of new antibacterials (Campbell & Cronan, 2001). β-Ketoacyl-acyl carrier protein synthase (KAS) III catalyses the first condensing step of the FAS-II reaction in plants and bacteria, using acetyl CoA and malonyl-ACP as substrates in most bacterial and plant systems (Abbadi *et al.*, 2000; Clough *et al.*, 1992; Dehesh *et al.*, 2001; Revill *et al.*, 2001; Tsay *et al.*, 1992). The biochemical basis for the production of branched-chain fatty acids by Gram-positive bacteria was analysed using *E. coli* and *Bacillus subtilis* FabH proteins. Reconstitution of a complete round of fatty acid synthesis *in vitro* with purified *E. coli* proteins showed that ecFabH was the only enzyme incapable of using branched-chain substrates. Expression of either bsFabH1 or bsFabH2 in *E. coli* resulted in the appearance of a branched-chain 17-carbon fatty acid. Thus, the substrate specificity of FabH is an important determinant of branched-chain fatty acid production (Choi *et al.*, 2000a). It was also observed in *Streptomyces glaucescens* that FabH is responsible for initiating both straight- and branched-chain fatty acid biosynthesis in Streptomyces (Han *et al.*, 1998).

The accumulation of malonyl-ACP following the inhibition of a reconstituted fatty acid synthase system by acyl-ACP implicated ecFabH as a target for acyl-ACP regulation (Heath & Rock, 1996a). Kinetic analysis showed that acyl-ACP inhibition was mixed with respect to acetyl-CoA, but was competitive with malonyl-ACP, indicating that acyl-ACPs decrease

FabH activity by binding to either the free enzyme or the acyl-enzyme intermediate, thus supporting the theory that the inhibition of chain initiation at the β-ketoacyl-ACP synthase III step contributes to the attenuation of fatty acid biosynthesis by acyl-ACP (Heath & Rock, 1996b).

An open reading frame in the *M. tuberculosis* genome (Rv0533c) now termed *mtfabH*, encodes a 335 amino acid protein with a molecular mass of 34872.5 Da and has an amino acid identity of 37.3 % with ecFabH and contains the Cys-His-Asn catalytic triad signature characteristic of KASIII enzymes. The purified recombinant mtFabH clearly preferred long-chain acyl-CoA substrates rather than acyl-ACP primers but did not utilise acetyl-CoA as a primer, unlike ecFabH (Choi *et al.*, 2000b). mtFabH catalyses a Claisen-type condensation of long chain acyl-CoA substrates, such as myristoyl-CoA (C_{14}-CoA) and malonyl-AcpM (Scarsdale *et al.*, 2001), consistent with a role for mtFabH as the pivotal link between the FAS-I and FAS-II fatty acid elongation systems in *M. tuberculosis* (Choi *et al.*, 2000b). The elucidation of the structure of FabH is important for the understanding of its regulation by feedback inhibition and its interaction with drugs. The crystal structure of ecFabH was determined to 1.8 Å resolution using a 12-site selenium multiwavelength anomalous dispersion experiment. The active site was shown to be formed by the convergence of two α-helices and is accessed *via* a narrow hydrophobic tunnel (Qiu *et al.*, 1999). Hydrogen-bonding networks that include two tightly bound water molecules fix the positions of His244 and Asn274, which are critical for the decarboxylation and condensation reactions. A His244→Ala mutation did not affect the transacylation reaction suggesting that His244 has only a minor influence on the nucleophilicity of Cys112. In ecFabH, these histidine and asparagine residues are both required for the decarboxylation of malonyl-ACP. The nucleophilicity of the active-site cysteine is enhanced by the α-helix dipole effect, and an

oxyanion hole promotes the formation of the tetrahedral transition state (Davies *et al.*, 2000). The first co-crystal structure of a bacterial FabH condensing enzyme and a small molecule inhibitor was reported by Daines *et al.* (2003). In *S. pneumoniae* it was identified that indole analogues were potential inhibitors of FabH (Figure 2.1). This indole subunit has a 2,6-dichlorobenzyl group attached which was shown to block the active site entrance in *S. pneumoniae* FabH. Analogues of this inhibitor were constructed and assessed against both *S. pneumoniae* FabH and *E. coli* FabH. Also in this study they co-crystallised acetyl-CoA showing that major determinants of the adenine-binding site were Trp32 and Arg151.



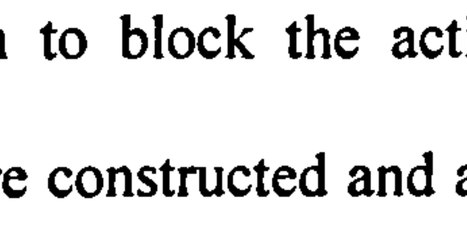
	IC_{50}
	0.040 μM
—$(CH_2)_4CONHSO_4Ph$	0.046 μM
—(4-OH)Ph	0.053 μM

Figure 2.1 Indole analogues that inhibit *S. pneumoniae* FabH, IC_{50} concentrations indicated to the right

The Cys-His-Asn catalytic triad present in other FabH molecules is conserved in *S. pneumoniae* FabH (Khandekar *et al.*, 2001). Dynamic light scattering, size exclusion chromatography, and mass spectrometry results indicated that selenomethionyl *S. pneumoniae* FabH is a noncovalent homodimer (Khandekar *et al.*, 2000). *In silico* docking of the NMR derived solution structure of ACP with the crystal structure of ecFabH showed that ACP docked to a positively charged/hydrophobic patch adjacent to the active site tunnel on ecFabH, which included a conserved arginine (Arg249) that was required for ACP docking. Kinetic analysis and direct binding studies between ecFabH and ACP confirmed the

identification of Arg249 as critical for ecFabH-ACP interaction (Zhang *et al.*, 2001). In the ecFabH structure, this active site tunnel is blocked by a phenylalanine residue, which constrains specificity to acetyl-CoA, whereas in mtFabH, this residue is replaced by a threonine, which permits binding of longer acyl chains. This same channel in mtFabH is capped by an α-helix formed adjacent to a 4-amino acid sequence insertion, which limits the bound acyl chain length to 16 carbons (Scarsdale *et al.*, 2001).

KAS enzymes are targets for two natural products, thiolactomycin (TLM) and cerulenin (CER). TLM mimics malonyl-ACP in the ecFabB active site. It forms strong hydrogen bond interactions with the two catalytic histidines, and the unsaturated alkyl side chain interacts with a small hydrophobic pocket stabilised by π stacking interactions. CER binding mimics the condensation transition state. The ecFabB His333→Asn protein was prepared to convert the FabB His-His-Cys active site triad into the FabH His-Asn-Cys configuration to test the importance of the two His residues in TLM and CER binding. This mutant was significantly more resistant to both antibiotics and had an affinity for TLM an order of magnitude less than the wild-type enzyme, illustrating that the two-histidine active site architecture is critical to protein-antibiotic interaction (Price *et al.*, 2001). Purified mtFabH was sensitive to TLM and resistant to CER in an *in vitro* assay. However, mtFabH over-expression in *M. bovis* BCG did not confer TLM resistance, suggesting that mtFabH may not be the primary target of TLM inhibition *in vivo* leading to several changes in the lipid composition of the bacilli upon over-expression of FabH in the presence of TLM (Choi *et al.*, 2000b).

Although TLM does not inhibit *Plasmodium falciparum* KASIII (IC_{50} > 330 μM), three structurally-similar substituted 1,2-dithiole-3-one compounds did inhibit pfKASIII with IC_{50}

values between 0.53 μM and 10.4 μM. These compounds also inhibited growth in culture (Prigge *et al.*, 2003).

2.2 mtFabH mechanism

mtFabH performs the same type of condensation reaction observed with other β-ketoacyl-ACP synthase enzymes. The elongation condensing enzymes catalyse a three step ping pong reaction (Heath & Rock, 2002). The first substrate is acyl-CoA, which is covalently transferred to the active site cysteine. The active site Cys in *E. coli* FabH is situated at the NH_2-terminus of an α-helix where the microenvironment of the pH is affected by the α-helix dipole effect and reduces the pH by 1.6 units. This in turn allows the Cys to remain in the thiolate form and thus the effective transfer of the acyl chain. The first product, reduced CoA, is released. Malonyl-ACP then binds and the decarboxylation of the malonyl moiety promotes the condensation of the two acyl chains. The final product is the released. The main difference between FabH enzymes and other β-ketoacyl-ACP synthase is that their first substrate is an acyl-CoA (usually acetyl-CoA) instead of acyl-ACP and CoASH is released (Figure 2.2).

The pantotheinyl arm of acyl-CoA or acyl-ACP binds in a narrow tube leading to the active site from the surface (Davies *et al.*, 2000; Qiu *et al.*, 2001; Qiu *et al.*, 1999). As the tube is linked by hydrophobic residues it does not interact strongly with the pantothenal arm. This is consistent with the step wise mechanism of FabH as the CoA pantotheinyl arm must be removed after depositing the acyl chain at the active site cysteine to allow malonyl-ACP to interact *via* the same channel (Davies *et al.*, 2000). The ADP moiety of bound acyl-CoA is sandwiched between the side-chains of Trp32 and Arg151 and hydrogen bonded to the

carbonyl of Arg151 and Thr28 in *E. coli*. This allows for the pantotheinyl arm to be inserted and the transfer of the acyl chain to the thiol of Cys112 (Qiu *et al.*, 2001) (Figure 2.3a). Binding of ACP to FabH has been examined in detail (Zhang *et al.*, 2001). ACP and FabH are both highly negatively charged but there is a patch of positive and hydrophobic residues near the entrance to the FabH active site channel. *In silico* modeling and docking experiments carried out with mutants indicated that Arg249 of ecFabH is critical for the electrostatic interaction with Glu41 on the surface of ACP (Zhang *et al.*, 2001). This predicted docking arrangement allows for the pantothenal arm of the ACP to insert the malonate into the active site.

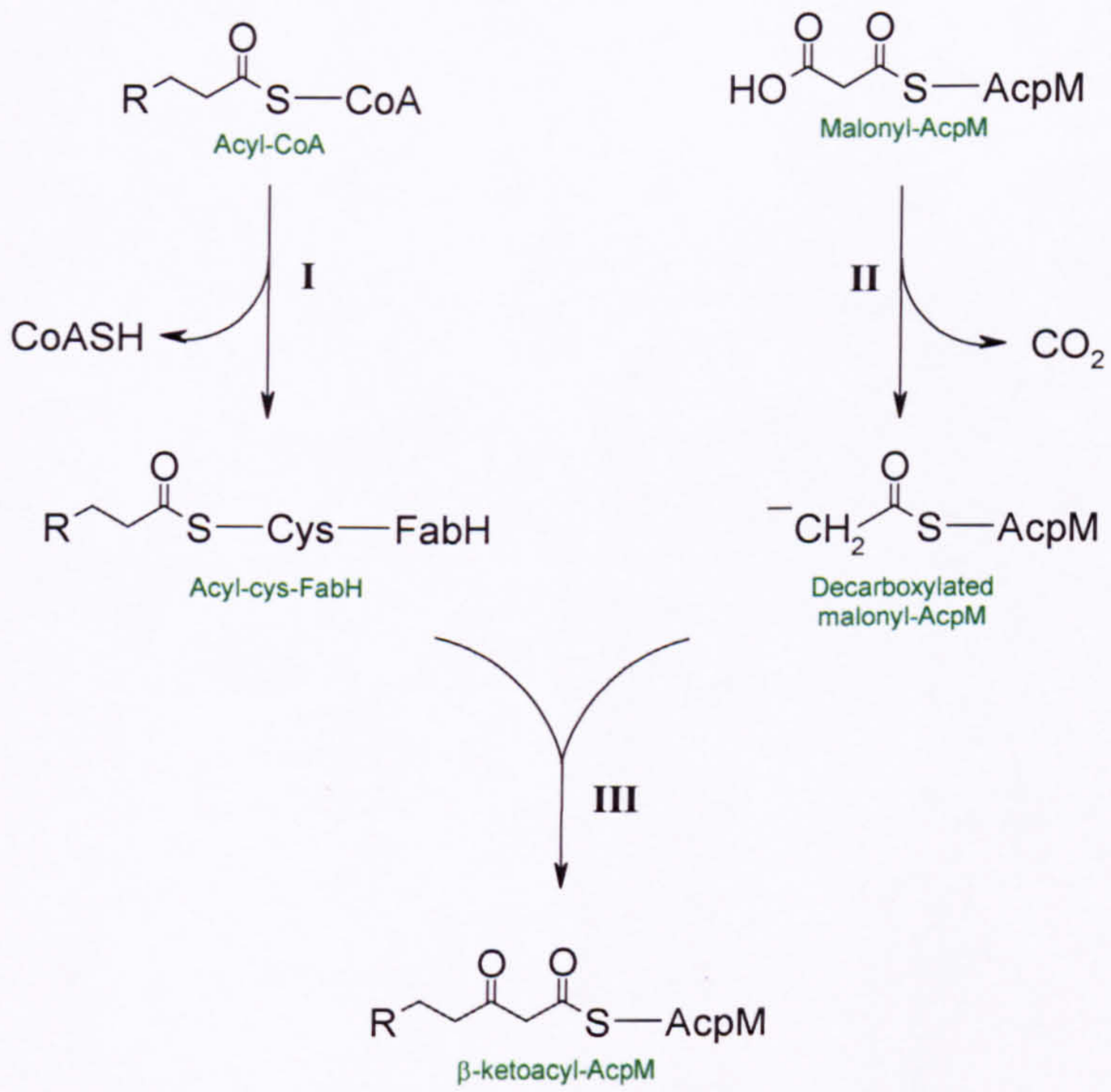


Figure 2.2 **Proposed condensation reaction mechanism performed by mtFabH. (I) The initial acyl-CoA acylation of the catalytic triad cysteine is followed by (II) the decarboxylation of the donor unit of malonyl-AcpM. The condensation of the two subunits (III) form the β-ketoacyl AcpM product**

Figure 2.4 shows the proximity of acyl-CoA binding residues to the entrance to the active site. In *E. coli* FabH, His244 and Asn274 of the active site are both required for malonyl-ACP decarboxylation. Mutation of either of these residues to alanine almost eliminates decarboxylation activity (Davies *et al.*, 2000). The nitrogens of these two residues form strong hydrogen bonds with the malonyl thioester oxygen to stabilise its negative charge formed during catalysis (Figure 2.3b).

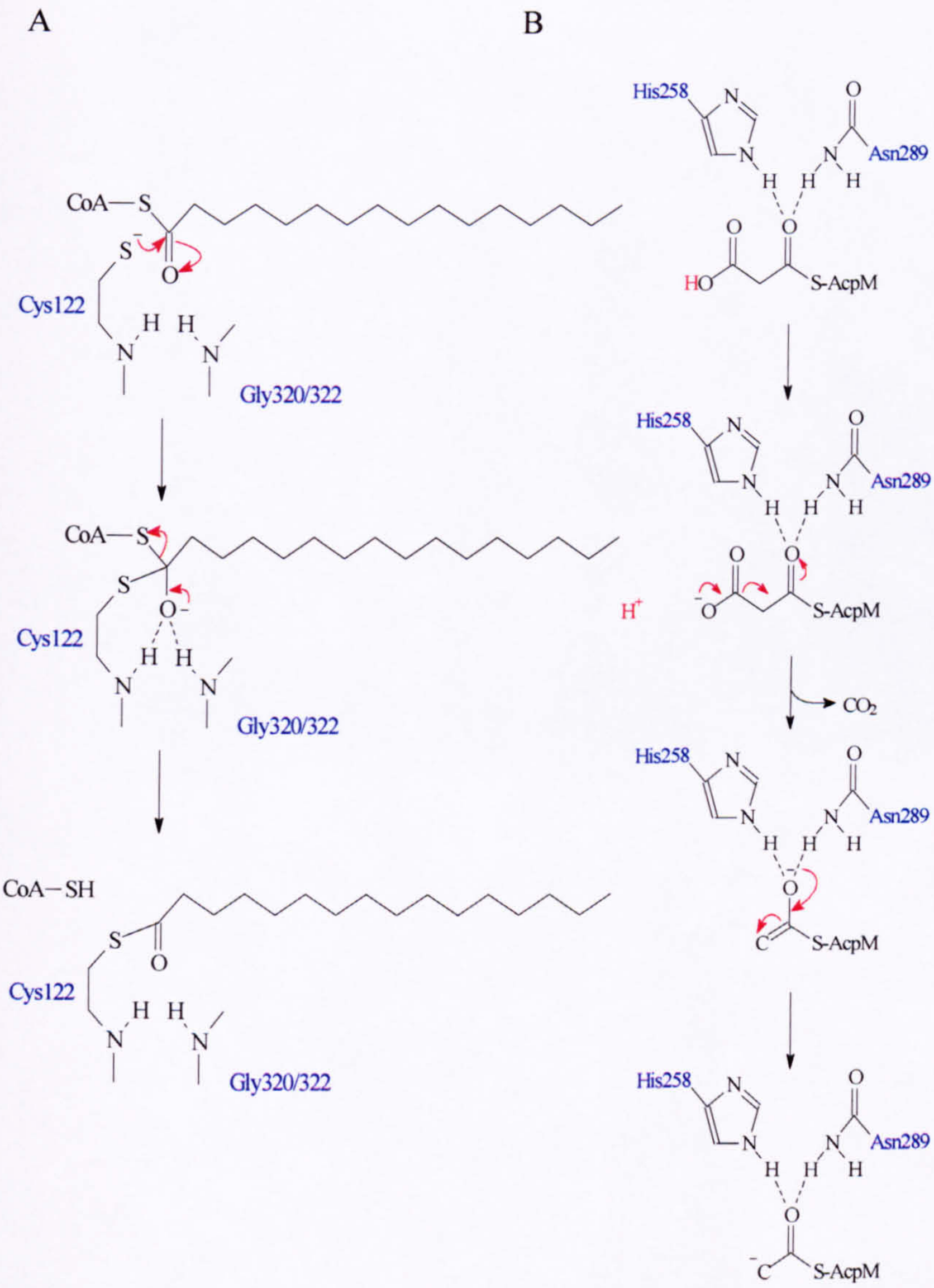


Figure 2.3 Proposed acylation (A) and decarboxylation (B) mechanisms of mtFabH adapted from (Davies *et al.*, 2000).

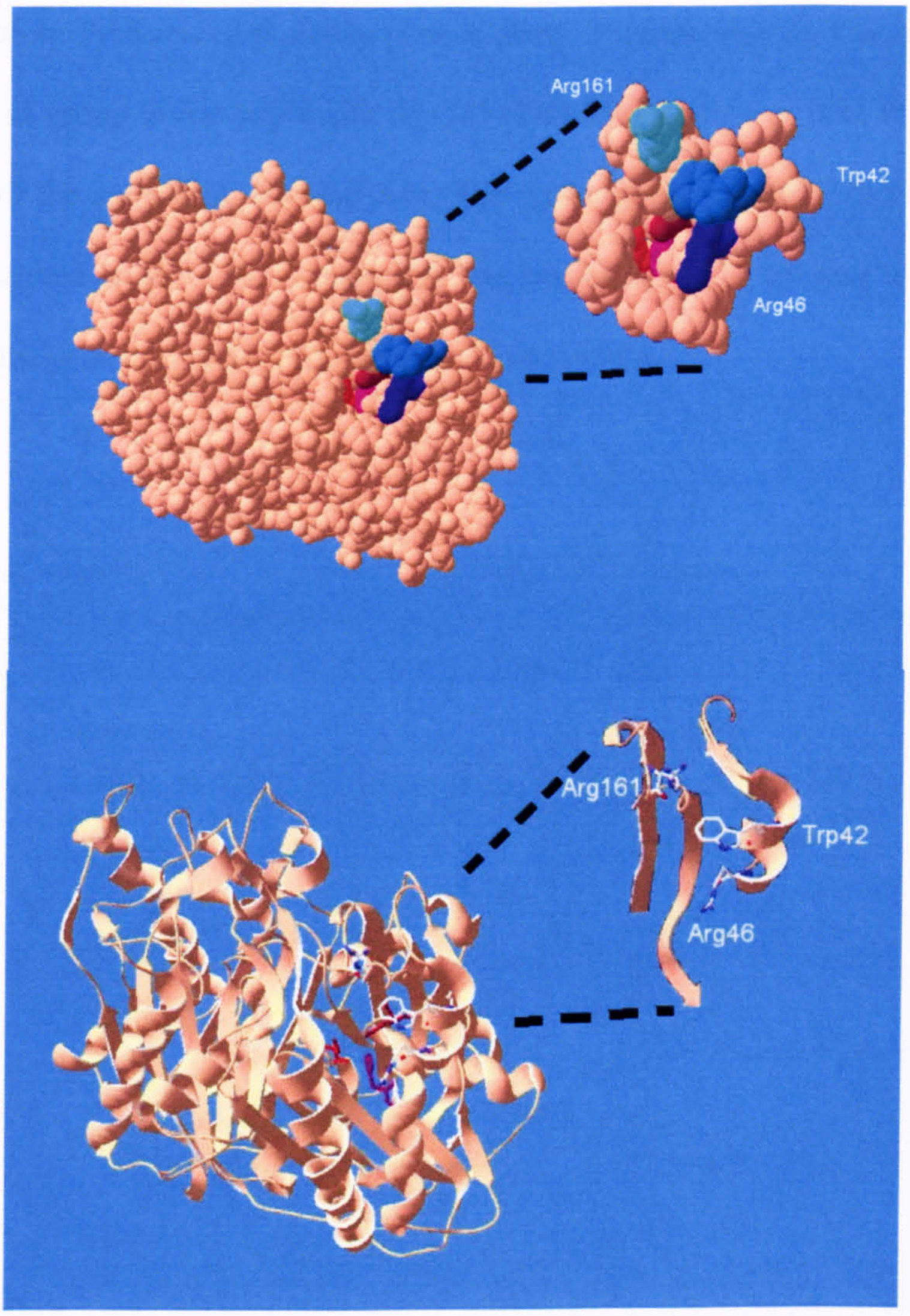


Figure 2.4 **Representation of mtFabH showing the binding site of acyl-CoA (blues) and the active site (Reds). The ADP moiety of bound acyl-CoA is sandwiched between the sidechains of Trp42 and Arg161 and hydrogen bonded to the carbonyl of Arg161 and Thr28 along with interations with Arg46. This allows for the pantotheinyl arm to be inserted and the transfer of the acyl chain to the thiol of Cys112 (122) (Red). This figure was generated using the crystal structure coordinates (1M1M, Scarsdale *et al.*(2001)) and the Swiss-PdbViewer Deep view program.**

During the condensation reaction, the thioester oxygen atom of the malonyl-ACP interacts with the His244 and Asn274 during the enolisation to facilitate the formation of the carbanion

at the C2 position. Protons are probably pushed away from the leaving CoASH and CO_2 by the proximity of group to the active site Phe205. This intermediate is held in position by its interaction with the active site His and Asn. The nucleophilic carbonion formed attacks the acyl group attached to the Cys112 releasing the β-ketoacyl-ACP product which is in turn released from the protein (Figure 2.5) (Heath & Rock, 2002).

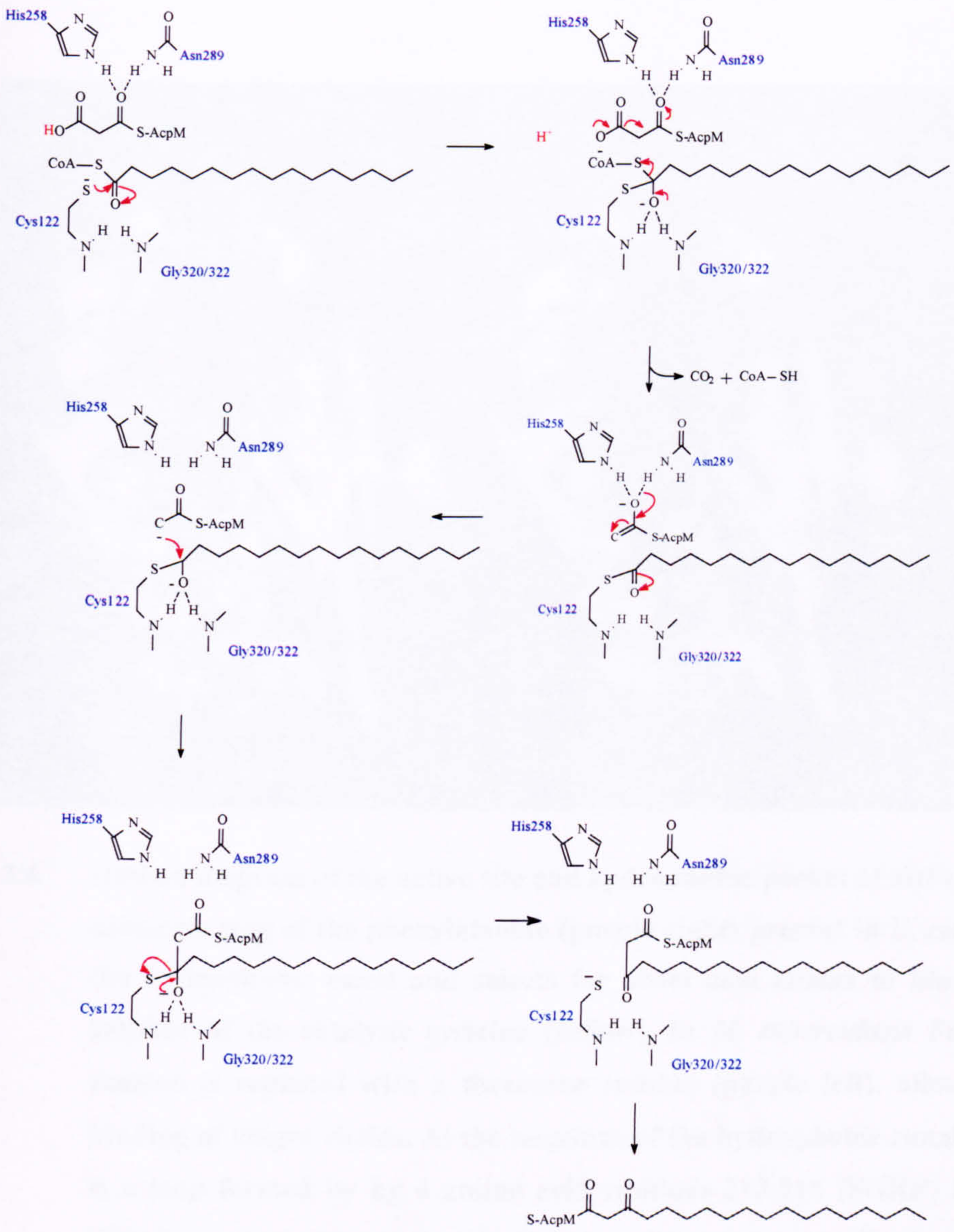


Figure 2.5 mtFabH reaction scheme adapted from (Davies *et al.*, 2000; Heath & Rock, 2002)

Structural analysis of the *E. coli* and *M. tuberculosis* FabH proteins revealed that the side chain of residue 97 protruded into the hydrophobic canal. The ecFabH has a phenyalanine residue which blocks the canal whereas the mtFabH protein has a less bulky threonine residue; it is believed that this residue affects the substrate chain length of the proteins, changing the proteins specificity (Figure 2.6).

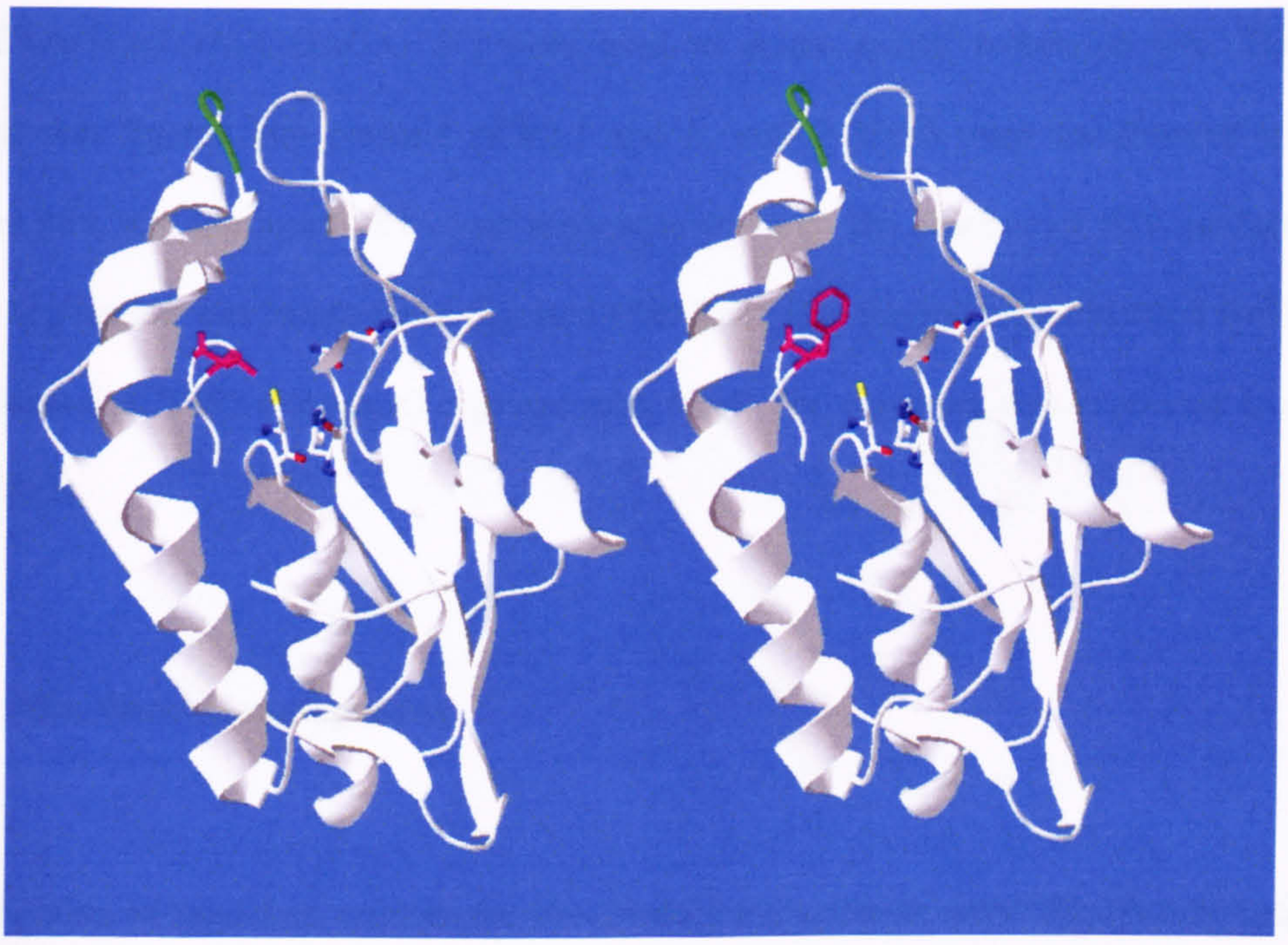


Figure 2.6 **Ribbon diagram of the active site and hydrophobic pocket of mtFabH. The aromatic ring of the phenylalanine (purple right) present in *E. coli* blocks the hydrophobic canal and selects for short acyl chains to bind to the sulphur of the catalytic cysteine (yellow). In *M. tuberculosis* FabH this residue is replaced with a threonine residue (purple left), allowing the binding of longer chains. At the terminus of the hydrophobic canal (green) is a loop formed by by 4 amino acid residues 213-216 (SGRP) (Green). This figure was generated using the crystal structure coordinates (1M1M, Scarsdale *et al.*(2001)) and the Swiss-PdbViewer Deep view program.**

2.3 Materials and methods

2.3.1 Cloning

Dr Laurent Kremer from Institut Pasteur de Lille, France, performed the molecular cloning of mtFabH. The *mtfabH* orf was amplified from genomic DNA isolated from *M. tuberculosis*. PCR primers (5'-AGCTAGGAATTCGCGCGCTCAACCC-3' and 5'-GGACTAGTATGACGGAGA- 3') were used to create a *bfa*I restriction site. The PCR product was ligated into plasmid pCR2.1 and *E. coli* INVaF9 were transformed with the ligation mix, and the insert DNA was sequenced to verify the absence of PCR artifacts. The resulting plasmid was isolated and digested with *Bfa*I. This fragment was isolated and ligated into plasmid pET28b (Novagen) digested with *Nde*I. The construct was expressed in *E. coli* C41 (DE3).

2.3.2 Site-directed mutagenesis

Identification of candidate residues for site-directed mutagenesis were identified by sequence alignment and analysis of the crystal structure of ecFabH. The ORF was 37.3 % identical and 46.8 % similar to *E. coli* FabH and 33.4 % identical and 42.9 % similar to *Bacillus subtilis* FabH2 (Figure 2.7).

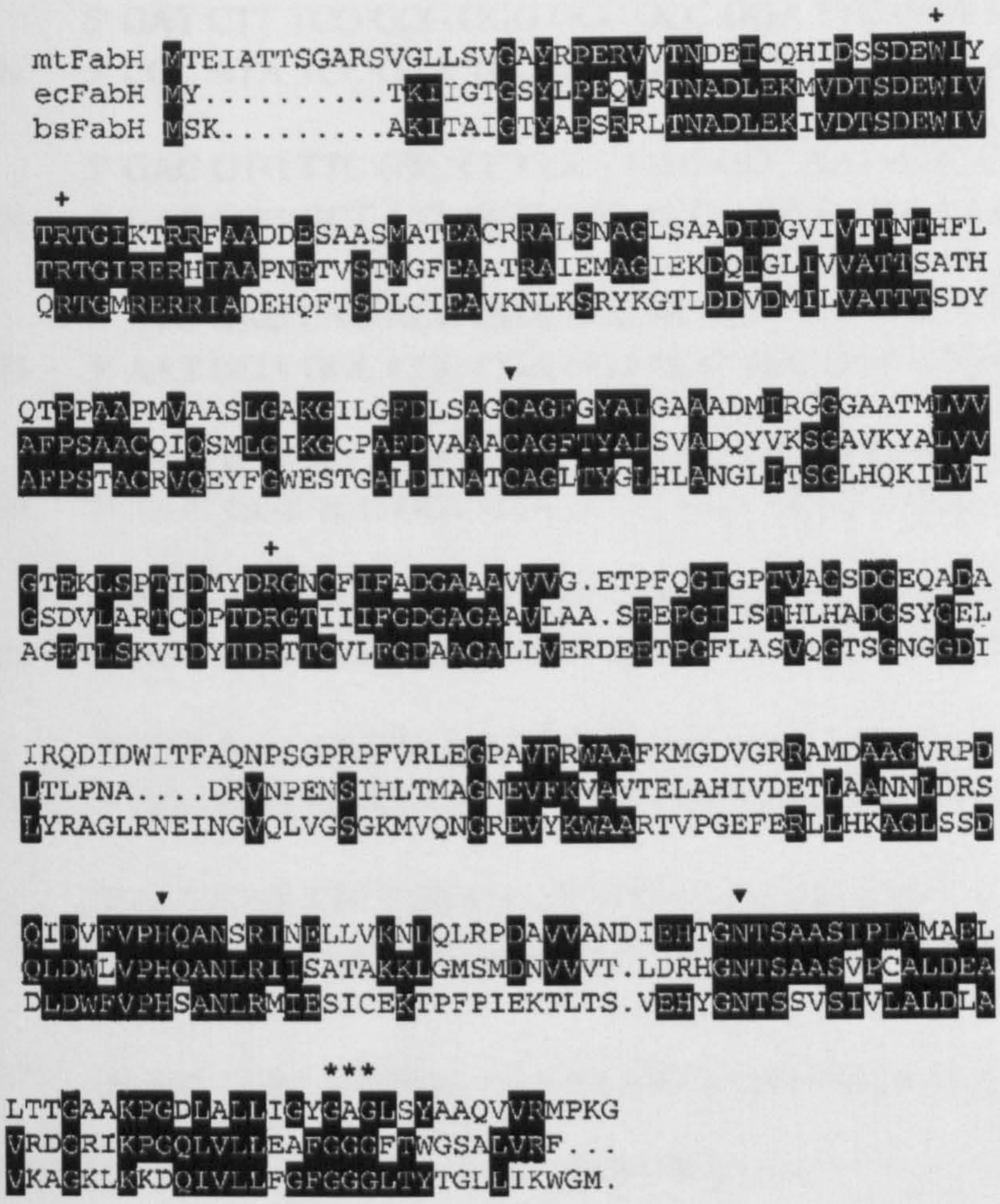


Figure 2.7 **Comparison of *M. tuberculosis* FabH (*mtFabH*) with *E.coli* FabH (*ecFabH*) and *B. subtilis* FabH2 (*bsFabH*). Primary amino acid sequence of Gram-negative (*E. coli*) and Gram-positive (*B. subtilis*) FabH was compared with the predicted protein sequence of mtFabH. Identical amino acids common to any two of the FabH proteins are shaded. mtFabH possesses the hallmark features of β-ketoacyl-ACP synthase III enzymes including the Cys-His-Asn active site triad (*arrowhead*), the oxyanion hole (*), and residues involved in acyl-CoA binding (+).**

A series of site-directed mutants were obtained using pET28a-*mtfabH* as the template for the QuikChange Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer's instructions with the following primers:

C122A	UP	5' GAT CTT TCG GCG GGG GCC GCC GGA TTC GGA TAT GCG 3'
	DWN	5' CGC ATA TCC GAA TCC GGC GGC CCC CGC CGA AAG ATC 3'
H258A	UP	5' GAC GTG TTC GTC CCT GCT CAG GCC AAT AGC CGC ATC 3'
	DWN	5' GAT GCG GCT ATT GGC CTG AGC AGG GAC GAA CAC GTC 3'
N289A	UP	5' ATC GAG CAC ACC GGA GCC ACC TCG GCG GCC TCC ATT 3'
	DWN	5' AAT GGA GGC CGC CGA GGT GGC TCC GGT GTG CTC GAT 3'
W42A	UP	5' GAC TCG TCC GAC GAG GCG ATC TAC ACC CGA ACC GCC 3'
	DWN	5' GGC GGT TCG GGT GTA GAT CGC CTC GTC GGA CGA GTC 3'
R46A	UP	5' GAG TGG ATC TAC ACC GCA ACC GGC ATC AAG ACC CGC 3'
	DWN	5' GCG GGT CTT GAT GCC GGT TGC GGT GTA GAT CCA CTC 3'
R161A	UP	5' ATA GAC ATG TAC GAC GCC GGC AAC TGC TTC ATC TTC 3'
	DWN	5' GAA GAT GAA GCA GTT GCC GGC GTC GTA CAT GTC TAT 3'
T97F	UP	5' ACC CAT TTC CTG CAA TTC CCG CCG GCC GCC CCA ATG 3'
	DWN	5' CAT TGG GGC GGC CGG CGG GAA TTG CAG GAA ATG GGT 3'

Using the R161A plasmid as the template, we were able to generate the double mutations R161A/R46A and R161A/W42A. The DNA was sequenced to assess the construct.

2.3.3 Stratagene Quikchange site-directed mutagenesis

The kit was used to make point mutations of pET28-*mtfabH* using the primers above. The PCR reaction and recipe follow, after the PCR reaction the methylated double stranded template DNA is removed by digestion with *DpnI*. Supercompetent XL1-Blue *E. coli* cells were transformed by heat shock with the digest mixture and plated on selective media containing kanamycin.

5 µl	10x Reaction buffer
1 µl	dsDNA template (pET28-*mtfabH*)
1 µl	5' Mutant primer
1 µl	3' Mutant primer
1 µl	dNTP mix
40 µl	ddH_2O
1 µl	*PfuTurbo* DNA polymerase

95°C	30 seconds	
95°C	30 seconds	16 cycles
55°C	1 minute	
68°C	1 minute/kb	

2.3.4 Purification protocol

E. coli C41 (DE3) expression host were transformed with pET28-*mtfabH*. The cells were cultured at 37°C to an $A_{600\,nm}$ = 0.6 and then recombinant gene expression was induced with 1 mM IPTG. The cultures were grown for a further 4 h at 37°C followed by harvesting by centrifugation. The cell pellets were resuspended in buffer (20 mM phosphate buffer, pH 7.9, 0.5 M NaCl, 10 mM imidazole) containing DNase and RNase. Bacteria were disrupted by sonication and the resulting lysate centrifuged at 27,000 x *g* for 60 min at 4°C. Hi-trap Chelating Sepharose™ fast flow matrix columns (1 ml or 5 ml) equilibrated with 0.02 M phosphate buffer and 0.5 M NaCl pH 7.4. Columns were charged using 2 column volumes of 0.1 M nickel chloride and washed with 4 column volumes of 0.02 M phosphate buffer and 0.5 M NaCl pH 7.4 to re-equilibriate the column and remove excess nickel. The clarified supernatant was applied to the column (maximum of 10 ml for a 1 ml column and 50 ml for a 5 ml column) and non-absorbed material collected for analysis. Proteins were eluted using varying concentrations of imidazole (10 mM – 500 mM). Fractions were collected and protein concentration determined by BCA (Pierce), 5 µg of each protein fraction was analysed by 12 % SDS-PAGE. Purified protein samples were dialysed against 20 mM Tris-HCl pH 7.6, 50

mM NaCl containing 1 mM β-mercaptoethanol and concentrated when required using a Centricon YM-10 filter unit (Millipore), and was stored in 50 % glycerol at -80°C. Routinely, yields of approximately 20 mg pure recombinant protein were recovered from a 2 L culture by Ni^{2+}-chelate affinity chromatography. Pure mtFabH were assessed by visualisation on SDS-PAGE (12 %).

2.3.5 Assays and assay development

2.3.5.1 MtFabH full Assay

The filter disc assay developed for ecFabH (Heath & Rock, 1996b) was modified to measure the activity of mtFabH with long-chain radiolabelled acyl-CoA. The assays contained 50 μM ACP/AcpM, 1 mM β-mercaptoethanol, 0.1 M sodium phosphate buffer, pH 7.0, 50 μM malonyl-CoA, 100,000 cpm [2-^{14}C] malonyl-CoA (specific activity 55 Ci/mole), 12.5 μM substrate (acetyl-CoA, butyryl-CoA, hexanoyl-CoA octanoyl-CoA, decanoyl-CoA, lauroyl-CoA, myristoyl-CoA, palmitoyl-CoA, stearoyl-CoA, or arachidoyl-CoA) and mtFabD (0.3 mg of protein) in a final volume of 50 ml. The mtFabD protein was added to generate the malonyl-ACP substrate for the reaction. A mixture of ACP, 1 mM β-mercaptoethanol, and the buffer was incubated at 37°C for 30 min to ensure complete reduction of ACP, and then the remaining components (except mtFabH) were added. The mixture was then dispensed in aliquots into the assay tubes and any additions of antibiotics/analogues made prior to the initiation by the addition of mtFabH. The reaction mixture was incubated at 37°C for 40 min. The reaction was quenched by the addition of 1 ml 100 mM K_2HPO_4, 100 mM KCl, 30 % THF and 5 mg $NaBH_4$ converting the β-keto-acyl-AcpM product into a acyl-1,3-diol (Figure 2.8). The radiolabelled fatty acids were extracted three times with 1 ml water saturated

toluene. Organic solvent extracts were pooled and washed with an equal volume of water. The solvent was then transferred to scintillation vials and dried, 5 ml EcoScint A was added prior to counting.

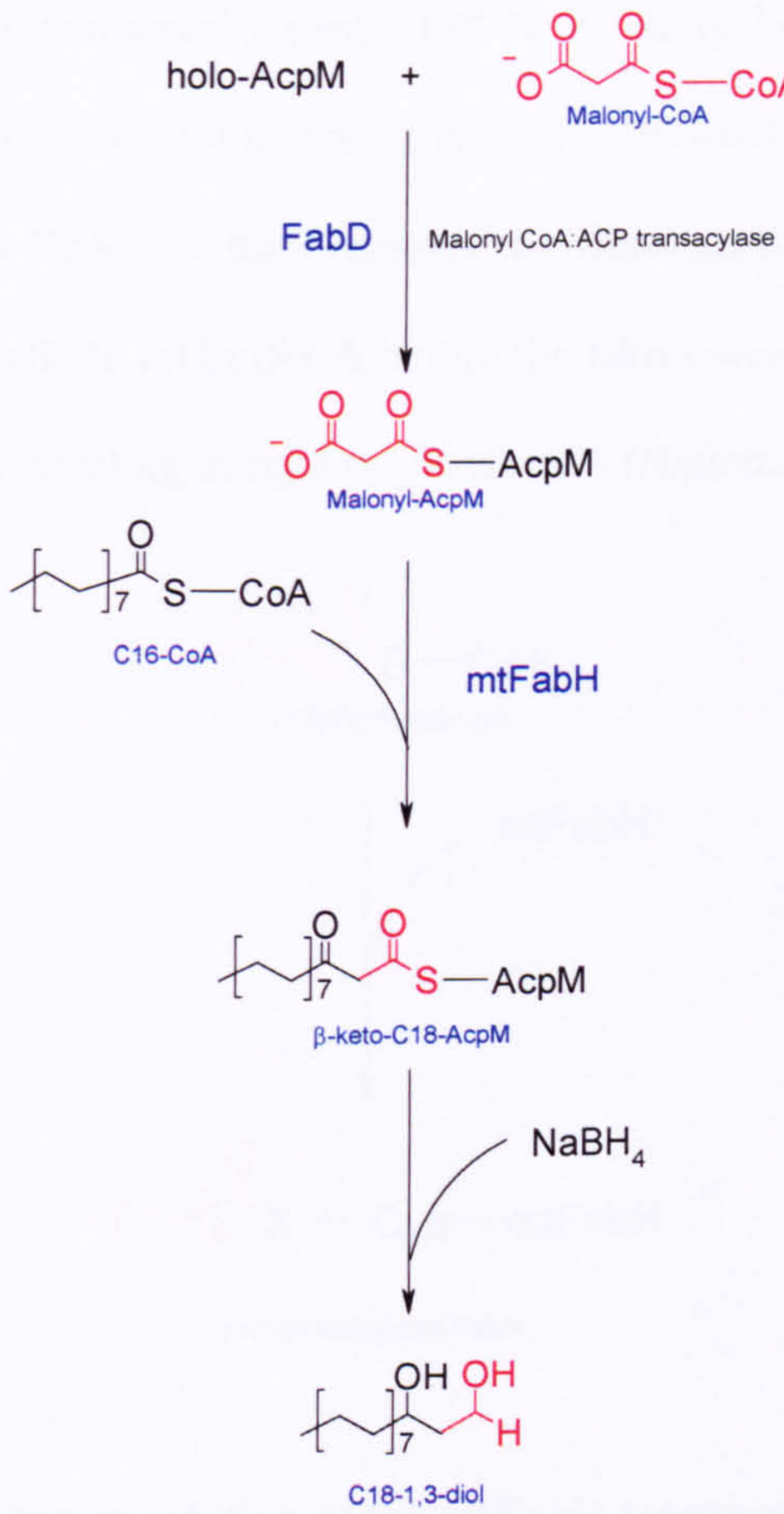


Figure 2.8 Schematic representation of the full mtFabH radiolabelled assay

2.3.5.2 Transacylation assay

The acylation assay for KAS I (McGuire *et al.*, 2001) was adapted for mtFabH. The acyl transferase assay measures the rate of formation of cys-acyl thioesters using a radiolabelled acyl-CoA. Acyl transfer was assessed using [^{3}H]-myristoyl-CoA.

The final reaction mixture consisted of 50 mM Tris-HCl pH 7.5 (Buffer A), 800 pmol [^{3}H]-myristoyl-CoA (1.66 kBq), and 1 μg mtFabH/mutant mtFabH to a final volume of 60 μl. The transfer was carried out at ambient temperature for 10 min. 200 μl of Ni^{2+} charged chelating sepharose™ fast flow chromatography media (50 % $^{v}/_{v}$ slurry in buffer A) was added and incubation continued for a further 5 minutes. The slurry was diluted by adding 300 ml buffer A and unbound myristoyl-CoA was then removed by filtration (cellulose nitrate, 0.45 μm). The beads were washed with 20 ml buffer A before the filters were removed to a scintillation vial for liquid scintillation counting using 5 ml EcoScintA (National Diagnostics).

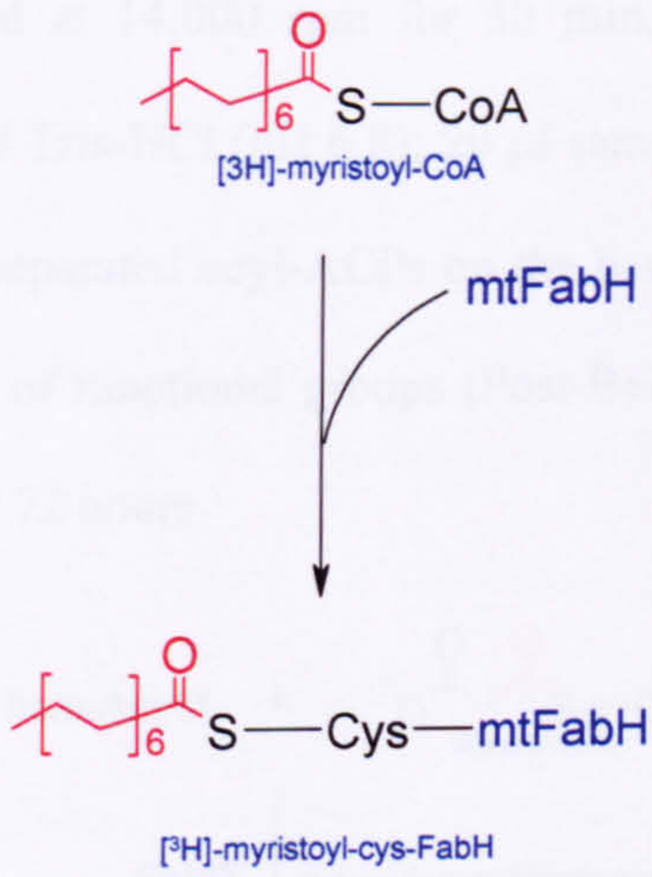


Figure 2.9 Schematic representation of the mtFabH transacylation assay

2.3.5.3 Malonyl-AcpM decarboxylation

The method used by McGuire *et al.* (2001) was adapted to measure the decarboxylation of malonyl-ACP by mtFabH. Here, malonyl-ACP was synthesised from malonyl-CoA and holo-ACPs *in situ* using a malonyl-CoA:ACP transferase (MCAT), FabD.

The final reaction mixture used was as follows. The appropriate ACP and MCAT pre-incubated with dithiothreitol (DTT) for 15-60 min on ice. Synthesis of malonyl-ACP was accomplished using 25 mM K_2HPO_4 (pH 8), 12 μM ACP, [1-^{14}C]malonyl-CoA (1.666kBq), 1 mM DTT, and 5 μg of FabD per assay. After 5 min at 25°C, the reaction mixture, designated A, was transferred to ice. (ii) For a 100 μl decarboxylation reaction, 83 μl of A was aliquoted to an Eppendorf tube with 17 μl (4.25 μg monomer) of mtFabH wild-type or mutant protein. After 10 min at 25°C, the reaction was stopped by the addition of 800 μl of cold 10 % (w/v) trichloroacetic acid (4°C) and 50 μl of 2 mg/ml BSA and incubated on ice for 30 min on ice. The solution was centrifuged at 14,000 rpm for 30 min, and the pelleted proteins were resuspended in 200 μl 50 mM Tris-HCl (pH 6.8); 20 μl samples were run on 1 M urea PAGE (13.3 % acrylamide), which separated acyl-ACPs on the basis of the chain length of the acyl substitution and the presence of functional groups (Post-Beittenmiller et al., 1991). The urea gel was dried and exposed for 72 hours.

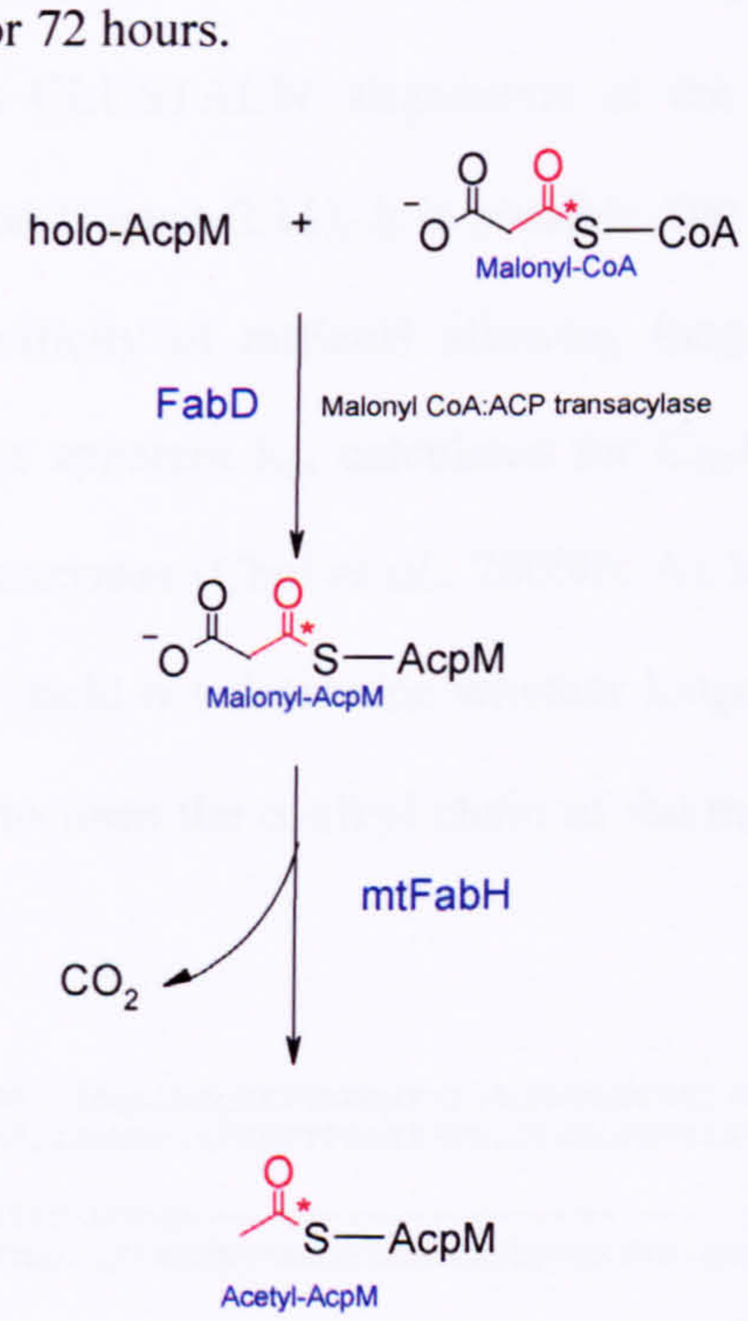


Figure 2.10 Schematic representation of the mtFabH decarboxylation radiolabelled assay

2.4 Results

2.4.1 ACP preference of β-ketoacyl ACP synthase III activity of mtFabH

The genome of *M. tuberculosis* H37Rv contains three putative ACP genes. One of these, AcpM, lies within a cluster of genes encoding several of the components of the FAS-II system (Kremer *et al.*, 2001) and has been suggested as the carrier for the growing meromycolate chain during FAS-II (Besra *et al.*, 1994). When we compared the rate of malonate incorporation into β-ketoacyl-ACP using both AcpM and the *E. coli* ACP (ecACP) we observed a shift in the acyl chain length specificity for mtFabH. When ecACP was used the optimal acyl-CoA substrate was C_{12}-CoA. Both C_8- and C_{10}-CoAs were efficiently extended but larger CoAs represented poor primers. However when AcpM was used as the carrier, malonate incorporation continued to rise with acyl chain length with an apparent optimum with octadecanoyl-CoA. As CLUSTALW alignments of the two ACPs show that AcpM carries a C-terminal extension (Figure 2.11), it is possible that this region may modulate the acyl-CoA chain length specificity of mtFabH allowing longer acyl chains to bind more efficiently (Figure 2.12). The apparent K_m calculated for C_{12}-CoA was 2.23 mM, which is consistent with other FabH enzymes (Choi *et al.*, 2000b). As longer chain length acyl-CoAs were not available to us we could not determine whether longer primers, emulating the C_{22}-C_{24} FAS-I products thought to form the α-alkyl chain of the mature mycolates, are extended by mtFabH.

```
ACP     ----MSTIEERVKKIIGEQLGVKQEEVTNNASFVEDLGADSLDTVELVMALEEEFDTEIP 56
AcpM    VPVTQEEIIAGIAEIIEEVTGIEPSEITPEKSFVDDLDIDSLSMVEIAVQTEDKYGVKIP 60

ACP     DEEAEKITTVQAAIDYINGHQA---------------------------------- 78
AcpM    DEDLAGLRTVGDVVAYIQKLEEENPEAAQALRAKIESENPDAVANVQARLEAESK 115
```

Figure 2.11 CLUSTALW alignments of the *E. coli* ACP and *M. tuberculosis* AcpM, showing the C-terminal extension

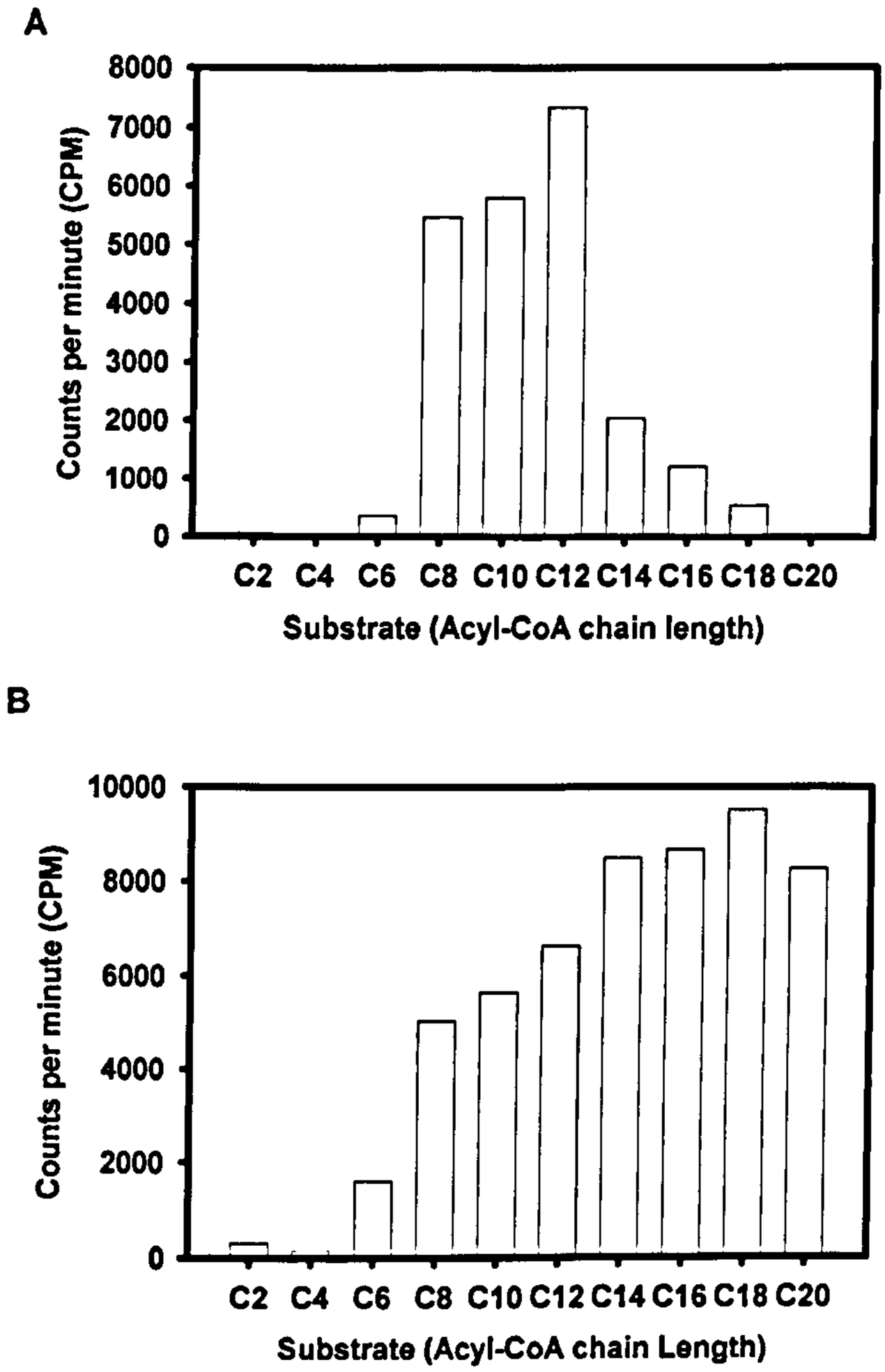


Figure 2.12 **Substrate specificity analysis of mtFabH, *E. coli* ACP vs *M. tuberculosis* AcpM. Comparison of *E. coli* ACP and AcpM as substrates for mtFabH with regards to substrate specificity using acyl-CoA chain lengths ranging from C_2-C_{20}. The mtFabH assay was performed in triplicate and activity assessed by scintillation counting. The method is described in detail section 2.3.5.1.**

2.4.2 β-Ketoacyl ACP synthase III activity of mtFabH muteins

The availability of the mtFabH crystal structure allowed us to target specific residues for site-directed mutagenesis thus facilitating a biochemical analysis of their individual contribution to KAS-III activity. Three groups of amino acid residues were chosen for site-directed mutagenesis in order to determine their contribution to the KAS activity of mtFabH. These could be subdivided according to the predicted functions of the residue derived from our knowledge of the crystal structure of the enzyme; (i) the residues contributing the catalytic triad (Cys121, His258 and Asn289), (ii) those implicated in acyl-CoA binding (Trp42, Arg46 and Arg161) and (iii) those apparently defining the acyl chain length specificity of the enzyme (Thr97 and Δ213-216).

E. coli C41 (DE3) expression host transformed with pET28-*mtfabH* were cultured at 37°C to an $A_{600\ nm}$ = 0.6 and then recombinant gene expression was induced with 1 mM IPTG. The cultures were grown for a further 4 h at 37°C followed by harvesting by centrifugation. The cell pellets were resuspended in buffer and then the bacteria were disrupted by sonication and the resulting lysate centrifuged at 27,000 x *g* for 60 min at 4°C. The clarified supernatant was applied to a Hi-trap Chelating Sepharose™ (1 ml). Fractions of purified mutated and wild type mtFabH were visualised by 12 % SDS-gel electrophoresis and Coomassie blue staining (Figure 2.13). Over-expression of these alleles in *E. coli* C41 (DE3) yielded large quantities of His_6-tagged mutein (~10 mg/l culture). Purified proteins were dialysed into 20 mM Tris-HCl pH 7.9, 50 mM NaCl. Pure protein was stored in 20 % glycerol and stored at -80°C.

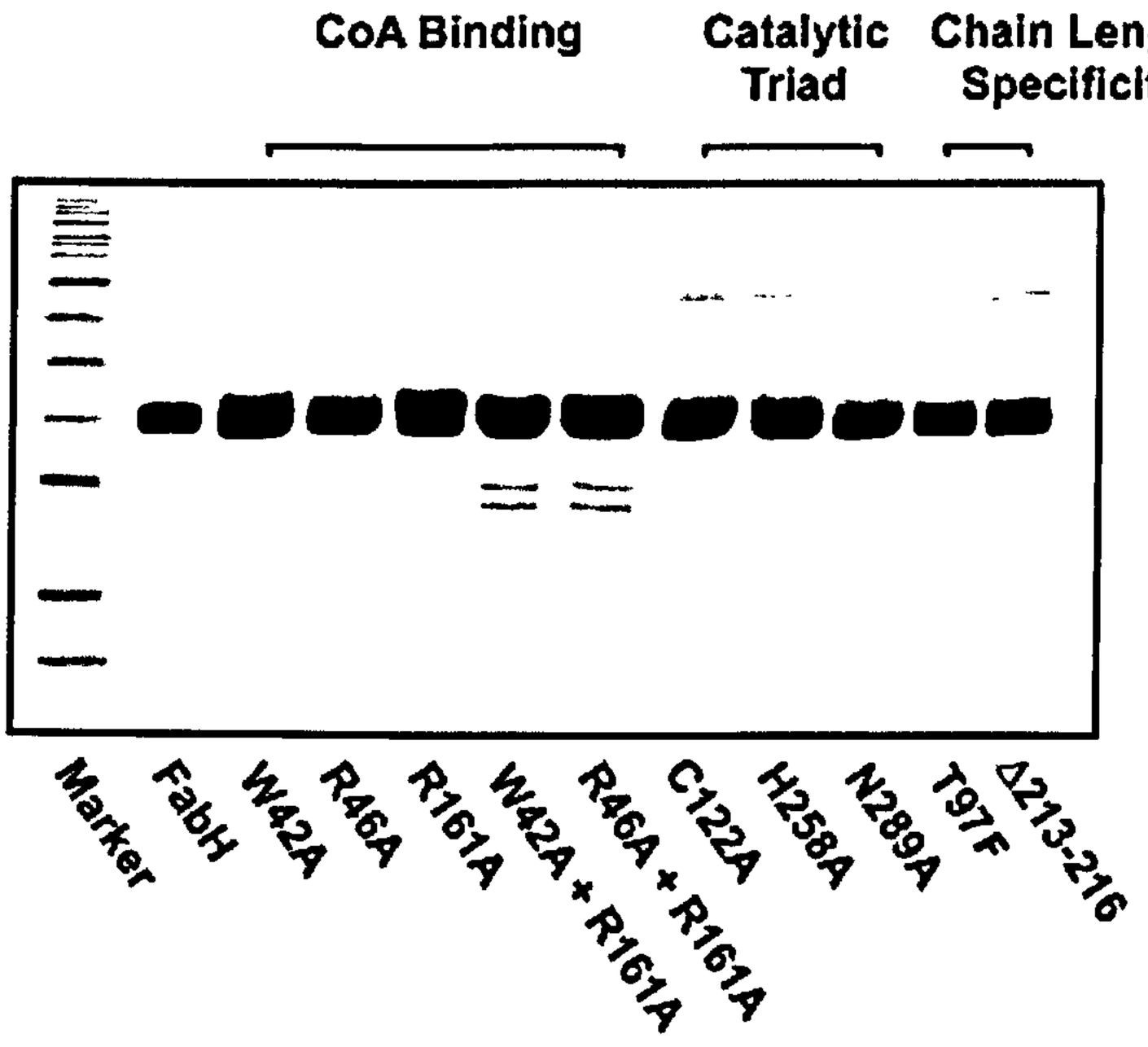


Figure 2.13 **SDS-PAGE gel of mutant protein purity. Produced by over-expression *in E. coli* C41 (DE3), 4 hours 1 mM IPTG and purification by His-trap chelating sepharose (1 ml). Fractions of purified mutated and wild type mtFabH were visualised by 12 % SDS-gel electrophoresis and coomassie blue staining**

The β-ketoacyl-ACP synthase III activity of each mutant protein was determined (Figure 2.14). As expected, the mutations involving the catalytic triad residues completely inactivated the enzyme confirming that these residues are vital for the condensation reaction to occur. The single substitutions Trp42→Ala, Arg46→Ala and Arg161→Ala reduced activity to 44, 13 and 76 % of mtFabH activity, respectively suggesting a non-essential role in catalysis. This is wholly consistent with their predicted function of contributing to the binding of acyl CoA primers (Zhang *et al.*, 2001). The two double muteins Trp42Arg161→Ala and Arg46Arg161→Ala showed further reductions in activity to 0.5 and 0.6 %, respectively. To date we have not been able to construct a mutant containing substitutions at both codons 42 and 46.

A Thr97→Phe substitution was created in an attempt to emulate the hydrophobic acyl-binding canal of ecFabH (Scarsdale *et al.*, 2001), which is partially blocked by the acyl moiety of the Phe residue, and thus alter the substrate specificity of mtFabH. Unfortunately this mutation inactivated the enzyme completely, with none of the acyl-CoA chain primers (C_2-C_{20}) being elongated. The end of this hydrophobic canal is 'capped' by residues 213-216 at the surface of the protein. We investigated the possibility that removing these residues in order to analyse the effect of the mutation on substrate specificity. However, as with the Thr97→Phe mutation, this mutant also exhibited no activity with any of the acyl (C_2-C_{20})-CoA primers used.

With our collaborator J. C. Sacchettini we are at present producing crystals of the double mutant proteins to evaluate structural changes and to shed light on their effects in relation to CoA binding.

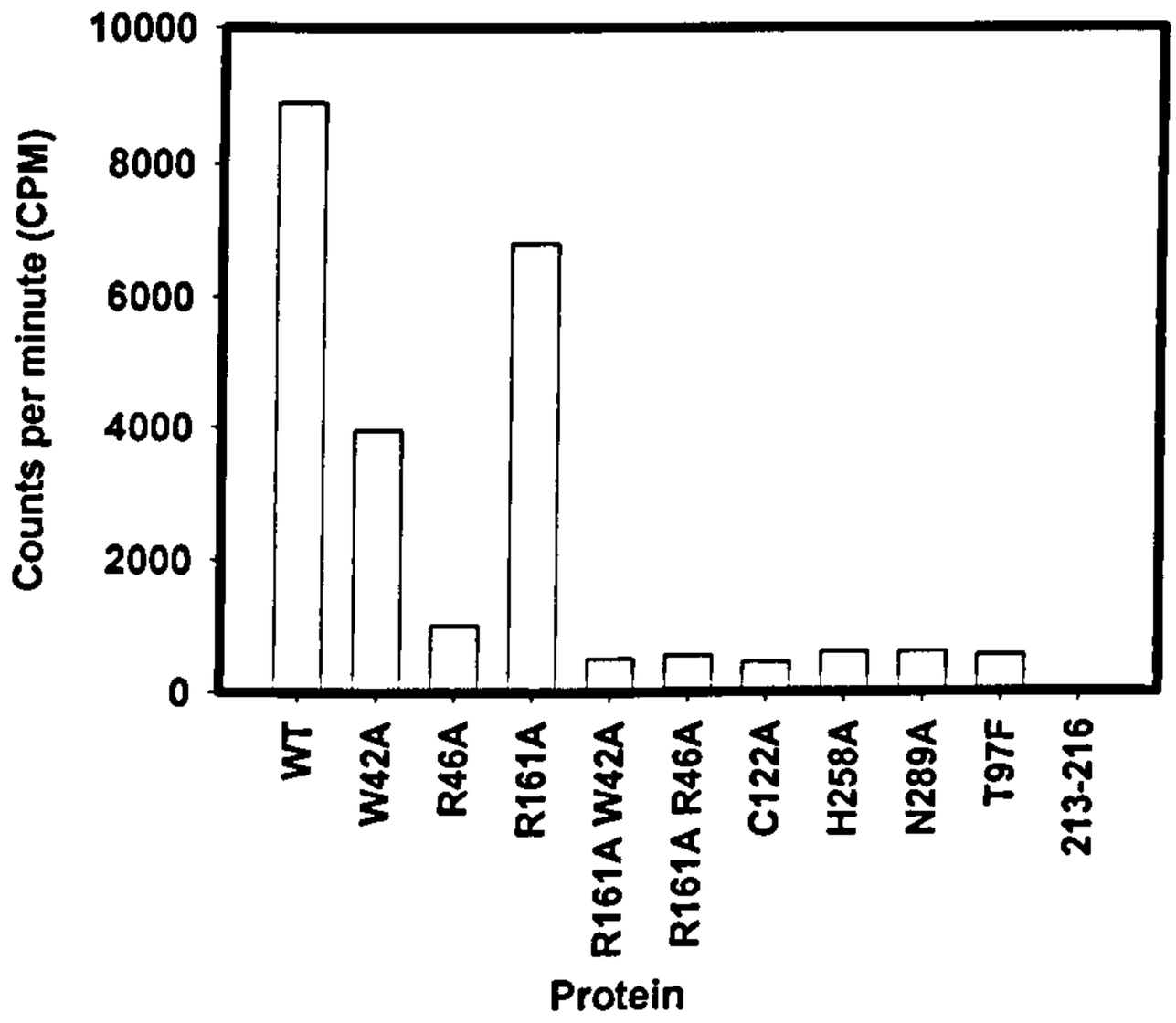


Figure 2.14 **β-ketoacyl ACP synthase III activity of mutant mtFabH. 0.5 μg of mtFabH/mutein were assayed in triplicate. Palmitoyl-CoA was used as an acyl primer and malonyl-AcpM, which was provided *in situ* by FabD.**

2.4.3 Transacylation activity of mutant mtFabH proteins

The contribution of a particular amino acid residue to the overall KAS-III activity can be more accurately defined by considering its performance in the component half-reactions. In the first transacylation reaction, the acyl chain is transferred from its acyl-CoA carrier to the active site residue Cys122 thus forming an enzyme-linked thioester. The subsequent decarboxylation of malonyl-AcpM produces a carbanion that initiates the condensation reaction by displacing the thioester and extending the ACP-bound primer.

We analysed the transacylase activity of mtFabH and its various mutants by following the transfer of [^{3}H]-C_{14}-CoA to FabH. As expected, no transacylase activity was detected with the Cys122→Ala substituted mutein, consistent with the loss of the active site sulphydryl group that forms the acyl thioester (Figure 2.15). Both of the other active site residues, His258→Ala and Asn289→Ala, were capable of transacylation but the reaction proceeded at a reduced rate suggesting that these residues are not essential for, but contribute, to the transfer of the acyl moiety. Somewhat surprisingly, these mutations had a more profound effect on acyl transfer than did individual substitutions in residues implicated in acyl-CoA binding (Trp42→Ala, Arg46→Ala and Arg161→Ala) and even the double-arginine mutant. The other double mutation, Trp42Arg161→Ala demonstrated very low transacylation activity, presumably due to a poor affinity for the acyl-CoA substrate. Together these data suggest that Trp42 and Arg161 both play a major role in acyl-CoA binding and thus influence the transacylation activity of the enzyme. The effect of the His258 and Asn289 alanine substitutions on the transacylation activity is less clear, although modification of active site geometry affecting acyl-CoA binding may occur, the possibility that these two polar residues may influence the binding of the CoA moiety cannot be ruled out. Again the Δ213-216 and

Thr97→Phe mutants were totally inactive in this assay system suggesting deleterious structural changes around the acyl-binding canal.

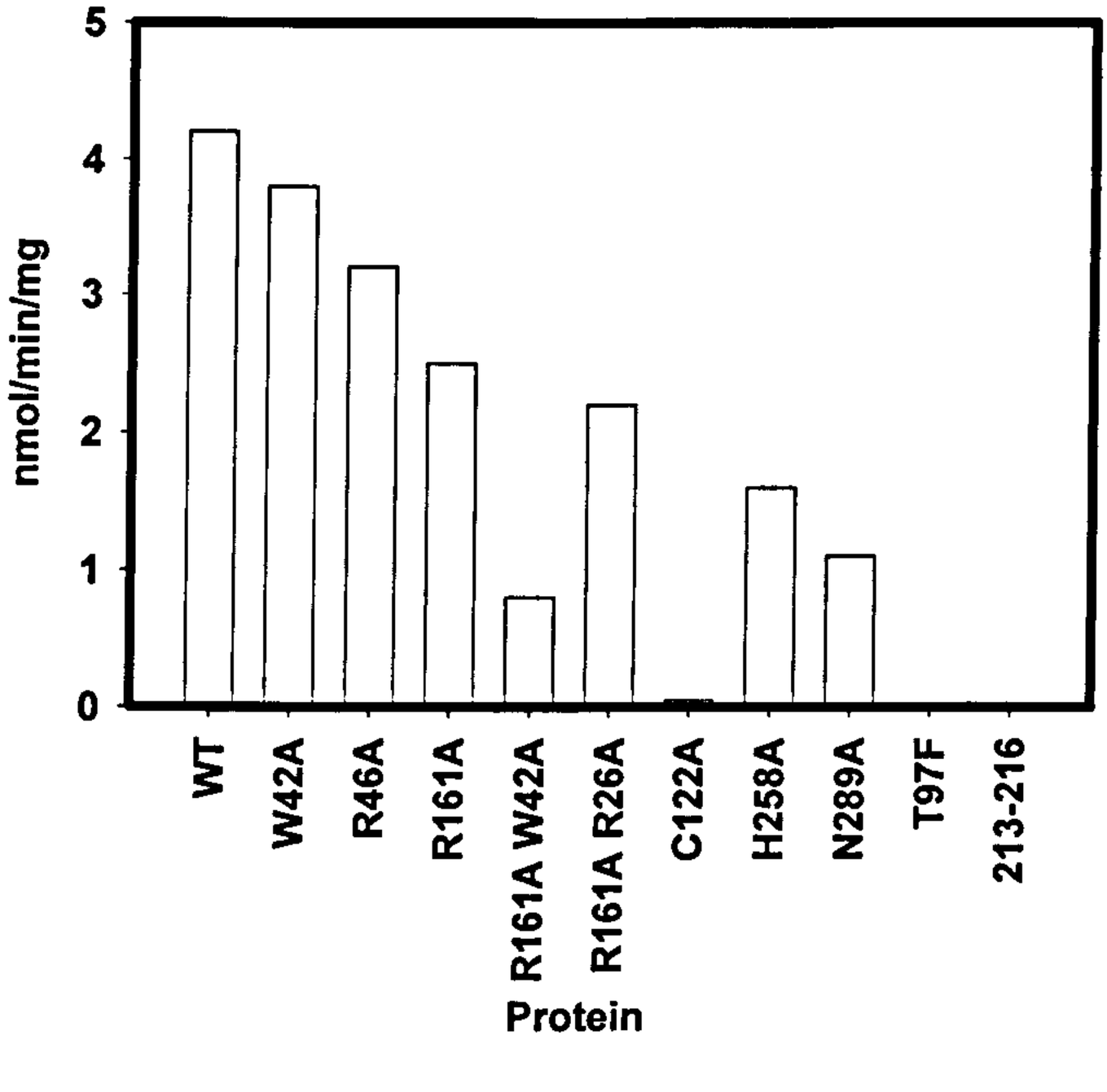


Figure 2.15 **Transacylation activity of mtFabH muteins. Transfer of [^{3}H]-myristoyl-CoA to the active site cysteine was followed for 15 mins. His$_6$-tagged protein was sequestered using Ni^{2+} charged sepharose beeds and washed to remove unbound label. Activity was quantified by liquid scintillation counting**

2.4.4 Decarboxylation of malonyl-ACP by wild-type and mutant mtFabH proteins

Conformationally-sensitive, urea-containing SDS-PAGE gels have previously been used to separate variously substituted acyl-ACPs and was applied to analyse the ability of the mtFabH muteins to decarboxylate the donor substrate malonyl-AcpM to acetyl-AcpM. All the muteins

in this study exhibited decreased malonyl-AcpM decarboxylation activity. Both ACP or AcpM were used with the wild-type mtFabH but no change in activity was observed. The catalytic triad residues His258 and Asn289 had previously been implicated in the decarboxylation of the malonyl-AcpM. Consistent with this suggestion, the decarboxylation activity of His258→Ala was completely abrogated whereas that of the Asn289→Ala mutein was markedly reduced. This observation strongly suggests that His258 is essential for this half-reaction and that substitution of Asn289 merely influences the rate of the reaction. Again both Thr97→Phe and the deletion mutation were inactive. All of the other mutant enzymes retain their decarboxylation activity but the reaction occurs at significantly reduced rates similar to that of the Asn298→Ala mutant. The reason for this reduced activity is not known but it is possible that even minor changes to the active site geometry brought about by these substitutions could significantly reduce the reaction rate.

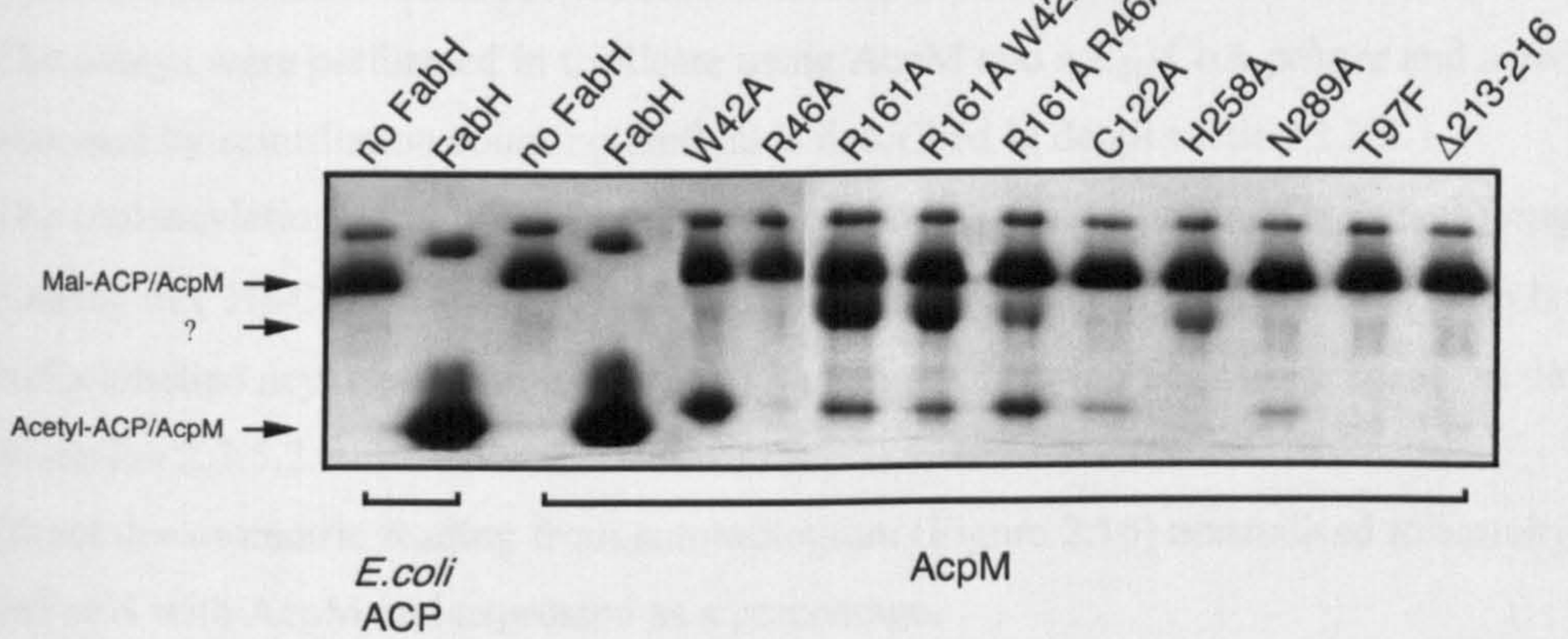


Figure 2.16 **Malonyl-ACP decarboxylation activity of mtFabH muteins. 4.25 μg of the muteins was incubated with radiolabelled malonyl-ACP for 10 min at 25°C. After resolution of the ACP species by conformationally sensitive 13.3 % acrylamide, 1 M urea gels, results were visualised by autoradiography using Kodak X-omat AR film for 3 days**

2.4.5 Summary of mtFabH mutein activities

Table 2.1 Activity of *mt*FabH muteins in β-ketoacyl-ACP synthase and part-reaction assays.

Substitution	mtFabH[a] nmol/min/mg	Transacylation[b] nmol/min/mg	Malonyl-AcpM decarboxylation (%)[c]
*mt*FabH	194.44	4.17	100.00
Trp42→Ala	41.92	3.72	30.08
Arg46→Ala	14.23	3.26	0.79
Arg161→Ala	18.63	2.49	8.67
Trp42 Arg161→Ala	0.46	0.75	5.09
Arg46 Arg161→Ala	0.61	2.15	9.68
Cys122→Ala	0.48	0.04	2.35
His258→Ala	0.89	1.63	0.68
Asn289→Ala	0.99	0.99	2.24
Thr97→Ala	0.91	0.02	0.84
Δ213-216	0.60	0.00	0.97

[a] The assays were performed in triplicate using AcpM and a C_{16}-CoA primer and activity assessed by scintillation counting method is described in detail section 2.3.5.1.

[b] The transacylation activity of *mt*FabH and its muteins was determined by quantifying the binding of [^{3}H]-C_{14}-CoA to 1 μg protein by liquid scintillation counting of protein-bound radio-labelled acyl chains captured on Ni^{2+}-charged chelating Sepharose beads, as detailed in section 2.3.5.2.

c Direct densitometric reading from autoradiogram (Figure 2.16) normalised to activity of mtFabH with AcpM and expressed as a percentage.

2.4.6 Analysis of compound sensitivities of wild-type mtFabH

The susceptibility of the KASIII activity of mtFabH to various antibiotics were assessed by quantifying the effect of addition of each compound over a range of concentrations through their incorporation into the β-Ketoacyl-ACP synthase III assay. As previously observed (Choi *et al.*, 2000b) mtFabH is susceptible to thiolactomycin (IC_{50} = 5 μg/ml), triclosan (IC_{50} = 20 μg/ml) and hexachlorophene (IC_{50} = 6 μg/ml). No inhibition was seen with any of the other members of our panel antibiotics/inhibitors used in this study (Table 2.2).

Table 2.2 Antibiotic sensitivity study of mtFabH. IC_{50} calculations were done using the direct assay method and assessed by scintillation counting method using palmitoyl-CoA.

Antibiotic	(IC_{50} = μg/ml)
TLM	5
HEX	6
TRI	20
CER	>50
ISO	>50
4-PY	>50
P-ACID	>50
DIAZ	>50

2.5 Discussion

The genome of *M. tuberculosis* H37Rv encodes several structurally and functionally related β-ketoacyl-ACP synthase enzymes/domains (Cole *et al.*, 1998). These represent an important family of condensing enzymes that play key roles in the biosynthesis of fatty acids, polyketides and mycolic acids in mycobacteria (Cole *et al.*, 1998; Kremer *et al.*, 2000b). In terms of fatty acid biosynthesis, *M. tuberculosis* H37Rv utilises FAS-I (*fas*, Rv2524c) for *de novo* fatty acid synthesis and three KAS enzymes (Cole *et al.*, 1998) encoded by *mtfabH*, *kasA* and *kasB* (Choi *et al.*, 2000b; Kremer *et al.*, 2002c; Mdluli *et al.*, 1998; Schaeffer *et al.*, 2001) for the further elongation of the acyl chain. The latter pair occupy neighboring loci (Rv2245-Rv2246) which are immediately downstream of *acpM* and are also clustered with *fabD* and *accD6* consistent with a role in FAS-II (Cole *et al.*, 1998; Kremer *et al.*, 2000a). Similarly, the FAS-II employed by *E. coli* for *de novo* synthesis of fatty acids contains three KAS enzymes. FabH initiates this synthesis by condensation of an acetyl-CoA primer with malonyl-ACP (Campbell & Cronan, 2001; Tsay *et al.*, 1992). In contrast, FabB and FabF utilise acyl-ACP primers for condensation with malonyl-ACP (Campbell & Cronan, 2001; Magnuson *et al.*, 1995). Choi *et al.* (2000b) showed that mtFabH corresponds to the *E. coli* KAS-III enzyme (FabH), although its substrate specificity varies, preferring C_{12}-CoA and C_{14}-CoA rather than acetyl-CoA substrates. Thus, mtFabH is envisaged as a pivotal link between FAS-I and FAS-II, condensing the CoA-bound products of FAS-I with malonyl-AcpM, which following a complete cycle of FAS-II, provides the acyl-AcpM primers for further extension, presumably by KasA/B (Choi *et al.*, 2000b). Consistent with this hypothesis KasA and KasB show a clear preference for AcpM-bound primers over acyl-CoA (4 % activity with C_{16}-CoA of that with C_{16}-AcpM) (Kremer *et al.*, 2002c; Schaeffer *et al.*, 2001).

In this study, we sought to analyse the influence of the ACP partner on catalysis and to confirm the roles of several key residues suggested by the recent determination of the *mt*FabH crystal structure. Analysis of mtFabH-catalysed malonate incorporation into β-ketoacyl-ACP using either AcpM or ecACP demonstrated an ACP-specific shift in the acyl-CoA chain length preference. The optimal acyl-CoA primer determined for *ec*ACP was C_{12}-CoA with both C_8- and C_{10}-CoAs being efficiently extended and longer acyl-CoAs representing very poor substrates. In contrast, when AcpM was used as the acyl carrier, malonate incorporation continued to increase with the acyl-CoA chain length to an apparent optimum with C_{18}-CoA. Sequence alignment of AcpM with *ec*ACP revealed that the former bears a 33 amino acid-extended carboxy terminus (Wong *et al.*, 2002). It has been suggested that this extension may have an effect on protein:protein interactions and may also interact with the bound acyl chain (Wong *et al.*, 2002). Molecular docking of *E. coli* ACP with ecFabH has been modelled (Zhang *et al.*, 2001) and, consequently, amino acid residues contributing to this interaction were identified and confirmed by mutagenesis. However, many of the important residues are not conserved in AcpM, therefore suggesting that the KAS-III:ACP interaction differs between the mycobacterial and the *E. coli* counterparts. One possibility is that the carboxy terminal tail of AcpM interacts with a surface region of *mt*FabH to induce conformational changes that might alter its substrate specificity. However, whether interaction with this C-terminal extension may manipulate the hydrophobic canal of mtFabH to allow successful interaction with longer acyl-CoA primers remains to be demonstrated.

Analysis of muteins showed that Ala substitutions of catalytic triad residues (Cys122, His258, Asn289) completely abrogated KAS-III activity, thus confirming these as vital residues for the condensation reaction. The single substitutions of the acyl-CoA binding residues (Trp42→Ala, Arg46→Ala and Arg161→Ala) showed reduced KAS-III activity, suggesting a

non-essential role for these residues in catalysis, consistent with their predicted function in binding of acyl-CoA primers (Scarsdale *et al.*, 2001). The two double muteins (Trp42Arg161→Ala and Arg46Arg161→Ala), however, showed a more profound effect with almost complete loss of activity. We propose that the singly-mutated proteins interact with acyl-CoA substrate with reduced affinity. The combination of amino acid substitutions in the double mutants leads to loss of KAS-III activity, suggesting that at least two of these three residues are required for efficient interaction of acyl-CoA and mtFabH.

The Thr97→Phe substitution was created in an attempt to emulate the hydrophobic acyl-binding canal of ecFabH, which is partially blocked by the side chain of a Phe residue (Scarsdale *et al.*, 2001). We envisaged that this might alter the substrate specificity of *mt*FabH limiting the enzyme to the shorter primers C_2- and C_4-CoA. Unexpectedly, this substitution completely abolished the enzyme activity. Although this insertion of a Phe residue into the hydrophobic canal may have caused deleterious changes to the structure of the protein inactivating mtFabH, we can speculate that the acyl-CoA preference of mtFabH may be governed by the degree of hydrophobic interaction between the incoming acyl chain and the hydrophobic binding canal. In such a scenario, longer acyl chains may be preferred by virtue of a more extensive hydrophobic interaction. Introduction of the Phe residue may have precluded such an extensive interaction thus explaining the inactivity of the mutein. Consistent with this hypothesis, our results show that mtFabH possesses weak KAS-III activity with a C_6-CoA primer; which may represent the shortest primer with an adequate K_M to allow detectable activity.

Previous studies suggested that the hydrophobic canal capping residues (Δ213-216) may limit the capacity of mtFabH to elongate acyl-CoA substrates bearing C_{16}- and shorter acyl chains

(Scarsdale *et al.*, 2001). However, deletion of this peptidic domain did not allow the extension of longer acyl-CoA primers, and as with the Thr97→Phe mutation, no activity was detected with any of the acyl (C_2-C_{20})-CoA primers. It remains possible that this mutation profoundly affects the overall folding of the protein.

The contribution of a particular amino acid residue to the overall KAS activity can be more accurately defined by considering its performance in the component half-reactions (McGuire *et al.*, 2001). In the first transacylation reaction, the acyl chain is transferred from its acyl-CoA carrier to the active site residue Cys122, thus forming an enzyme-linked thioester (Figure 2.4A). The subsequent decarboxylation of malonyl-AcpM produces a carbanion that initiates the condensation reaction by displacing the thioester and extending the ACP-bound primer (Figure 2.4B). No transacylation activity was observed with the Cys122→Ala substituted mutein, consistent with the loss of the active site sulphydryl group that forms the acyl thioester. In contrast, both His258→Ala and Asn289→Ala muteins catalyse the transacylation reaction, although to a reduced rate compared to the mtFabH. This suggests that these residues are not essential, but do influence the transfer of the acyl moiety. Somewhat surprisingly, these mutations had a more profound effect on acyl transfer than did individual substitutions in residues previously implicated in acyl-CoA binding. The transacylation activity of muteins substituted at Trp42, Arg46 and Arg161 mirrored the effects seen in the KAS-III assay, thus confirming that these residues participate in acyl-CoA binding. Although modification of active site geometry may affect acyl-CoA binding, His258 and Asn289 are polar residues, which may also influence the binding of the CoA moiety or aid the transfer of the acyl chain to the active site Cys. Interestingly, Ala substitution of *E. coli* FabB at His289 also resulted in reduced transfer of C_{14} acyl chains to the active site Cys (McGuire *et al.*, 2001). As the mtFabH His258→Ala mutein retains activity, it is unlikely that this active site residue

functions as a general base abstracting a proton from Cys122 to promote acyl transfer as has been shown for chalcone synthase (Jez & Noel, 2000). Rather, the stabilisation of the Cys122 thiolate ion, which facilitates transacylation, appears to be promoted *via* the α-helix dipole effect as in ecFabH (Davies *et al.*, 2000).

In ecFabH, active site residues His244 and Asn274 are both required for decarboxylation of malonyl-ACP. The side-chain nitrogen atoms of these two residues form strong hydrogen bonds with the thioester carbonyl oxygen, effectively stabilising the negative charge it gains during catalysis. Mutation of either of these residues almost eliminated decarboxylation activity (Davies *et al.*, 2000). The reduced ability of some of our mtFabH muteins to decarboxylate malonyl-AcpM implicated several residues in promoting efficient decarboxylation. Our results show that the His258→Ala and Asn289→Ala substitutions almost completely abrogated the decarboxylation activity of mtFabH. These observations strongly suggest that His258 is essential for this half-reaction and that substitution of Asn289 merely influences the rate of the reaction. Nearly all of the other mutant enzymes retain their decarboxylation activity but the reaction occurs at significantly reduced rates similar to that of the Asn298→Ala mutant. The reason for this reduced activity is not known but it is possible that changes to the active site geometry brought about by these substitutions could significantly reduce the reaction rate. Surprisingly, Arg46→Ala also drastically reduces the decarboxylation activity of mtFabH. In the absence of structural information regarding the mode of binding of AcpM by mtFabH, it remains possible that this residue plays a major role in the binding of AcpM, possibly interacting with the phosphopantetheine moiety of AcpM. Maybe incorporating interactions with such a residue may improve the efficacy of novel inhibitors of these enzymes. Interestingly, some recovery of malonyl-AcpM decarboxylation

activity relative to Arg46→Ala is seen in the Arg46Arg161→Ala mutant, possibly indicating the importance of local charge distribution to efficient decarboxylation.

The increasing incidence of MDR-TB intensifies the urgent need for novel targets and leads for anti-TB drug development. With this in mind, we examined the specific effect of several inhibitors of fatty acid synthesis on mtFabH. As previously observed (Choi *et al.*, 2000b; Jackowski *et al.*, 1989) ecFabH and mtFabH are both susceptible to TLM inhibition (IC_{50} 5 μg/ml), the data observed in this study supports these observations with identical IC_{50} values being observed for the *mt*FabH. Our *in vitro* demonstration of CER resistance is consistent with current hypothesis; the Cys-His-Asn catalytic triad of the type III KAS (FabH) enzymes correlating with resistance and Cys-His-His catalytic triad of FabB/F and KasA/B enzymes correlating with sensitivity to cerulenin (Choi *et al.*, 2000b). Interestingly TRC also inhibited the mtFabH *in vitro* (IC_{50} 20 μg/ml) as no increased resistance against TRC in a strain overproducing mtFabH, suggesting that mtFabH is not a target *in vivo* (result provide by Dr L. Kremer, Institut Pasteur de Lille).

Previous investigations had demonstrated *M. tuberculosis* InhA, and its *E. coli* homolog FabI, as targets of TRC (Kuo *et al.*, 2003; McMurry *et al.*, 1998; McMurry *et al.*, 1999; Parikh *et al.*, 2000). Likewise, the structurally-related HEX, another FabI inhibitor (Marcinkeviciene *et al.*, 2001), also inhibited mtFabH with an IC_{50} value of 6 μg/ml. Taken together these data may suggest some cross over in terms of their modes of action with mtFabH and its FAS-II partner, InhA. The inhibition of mtFabH by TRC and HEX is exciting and further investigations are required to understand their modes of action and may prompt the development of new lead compounds as anti-mycobacterial drugs.

The first part of any new study would be that identification of AcpM specific binding residues, this would in turn allow a site-directed mutagenesis study of decarboxylation activity in relation to substrate binding *via* AcpM. Residues required for protein:protein interactions are attractive targets for investigation as these could in turn become targets of synthesised inhibitors and antibiotic analogues. The co-crystallisation of mtFabH with both malonyl-AcpM and *E. coli* ACP would aid the elucidation of the molecular basis for the substrate specificity changeswe presume to be dependent upon the extended regions of carrier proteins such as AcpM. This would indeed show the structural effects that the ACP has on the activated protein and in turn open new areas of interest with regards to inhibition, as this would hopefully identify residues that can be manipulated to allow the binding of blocker substrates or hinder substrate-binding altogether. The investigation of the function of specific amino acid residues in this study along with the identification of an mtFabH inhibitor may also aid the development of future antibiotics.

CHAPTER 3

β-Ketoacyl–ACP synthase I (KasA)

3.1 Introduction

β-Ketoacyl ACP synthases (KAS) form new carbon-carbon bonds in a Claisen style condensation reaction performed in three steps. Firstly, transfer of an acyl-group from ACP to the enzyme, secondly decarboxylation of the elongation substrate, and finally condensation of the acyl primer with the elongation substrate. A number of key residues of *E. coli* KAS I (FabB) have been implicated in these partial reactions Cys163, His298, Asp306, Glu309, Lys328, and His333 (McGuire *et al.*, 2001). Mutation of these residues revealed that the active site Cys163 is not required for decarboxylation, whereas His298 and His333 are essential. Neither of the histidine residues are essential for increasing the nucleophilicity of Cys163 to enable transfer of the acyl-group. Maintenance of the structural integrity of the active site by Asp306 and Glu309 is required for decarboxylation, but not for acylation (McGuire *et al.*, 2000; McGuire *et al.*, 2001). Another facet of catalysis revealed by mutational analysis was that the acylation activity of KAS I is inhibited by free ACP at physiological concentrations (McGuire *et al.*, 2001). Differences in the inhibitory response by individually mutated KAS I proteins indicate that interaction of free ACP with Cys163, Asp306, Glu309, Lys328, and His333 might form a sensitive regulatory mechanism for the transfer of acyl groups (McGuire *et al.*, 2001; Olsen *et al.*, 2001). A detailed comparison of the active site of KAS enzymes using three-dimensional templates reveals differences in both the geometry and the catalytic role of key residues (Dawe *et al.*, 2003). The template based on the catalytic cysteine and two histidines in KAS I and II is totally specific for this family. However, the role of the histidine residues in catalysis differs between KAS I, II, III and thiolase of chalcone synthases (Heath & Rock, 2002). The strucuture of KAS I Cys163→Ser mutant with covalently bound C_{10} and C_{12} acyl-ACP was elucidated to 2.40 and 1.85 Å resolution, respectively (Olsen *et al.*, 2001). In both structures the oxo group was orientated to

make hydrogen bonds to the main amides of Ser163 and Phe392. The acyl chain folds into an extended U-shape in the hydrophobic canal due to sidechain of Phe201, which points into the acyl binding pocket, thereby forcing the acyl chain to bend (Olsen *et al.*, 2001). KAS I extracts of spinach leaves were highly active, elongating acyl-ACP's carrying from C_2 to C_{14}, with C_6-ACP being the most effective substrate. Interestingly, C_{16}-ACP was far less effective and C_{18}-ACP almost inactive (Shimakata & Stumpf, 1983). KAS I has been shown to be sensitive to the ACP released during the acyl transfase reaction, with 50 % inhibition occurring with 0.17 μM ACP which is close to the physiological concentration of holo-ACP (0.13 μM) (McGuire *et al.*, 2000). The KAS I dimer revealed an asymmetric binding of the two substrates (acyl-ACP and malonyl-ACP).

The analysis of mutations of KAS I active site histidine residues involved in malonyl-ACP decarboxylation suggested that one donates a proton to the malonyl thioester group, and the other abstracts a proton from the leaving group (Olsen *et al.*, 2001). Decarboxylation does not take place readily unless an acyl group is positioned at the active-site cysteine residue (Witkowski *et al.*, 1999). The crystal structure of the complex of *E. coli* KAS I, and cerulenin revealed that the inhibitor is bound in a hydrophobic pocket formed at the dimer interface. Cerulenin is covalently attached to the active site cysteine through its C2 carbon atom (Figure 1.27). The location of the inhibitor in the active site is not optimal, and there is thus room for improvement through structure-based drug design. TLM appears to mimic malonyl-ACP in the FabB active site by forming strong hydrogen bonds with the two catalytic histidines. The unsaturated alkyl side chain interacts with a small hydrophobic pocket stabilised by π stacking interactions (Moche *et al.*, 1999). A FabB His333→Asn mutation was prepared to convert the FabB His-His-Cys active site triad into the FabH His-Asn-Cys configuration to test the importance of the two His residues in TLM and cerulenin binding. The mutant was

significantly more resistant to both antibiotics than the wild-type, illustrating that the two-histidine active site architecture is critical to protein-antibiotic interactions (Price *et al.*, 2001).

Mycolic acids are produced by successive rounds of elongation catalysed by FAS-II as described in Chapter 1. A key feature in the elongation process is the condensation of a two-carbon unit from malonyl-acyl-carrier protein (ACP) to a growing acyl-ACP chain catalysed by a β-ketoacyl-ACP synthase A (KasA). KasA (Rv2245) from *M. tuberculosis* is a 416 amino acid residue-containing enzyme of 43316 Da that elongates meromycolate precusors. It utilises prefers C_{16}-ACP rather than short-chain acyl-ACP primers, unlike *E. coli* and other bacterial KAS enzymes that utilise C_4-ACP (Magnuson *et al.*, 1993; Schaeffer *et al.*, 2001). Interestingly, the activities of KasA and KasB both increased with use of *M. tuberculosis* AcpM, suggesting structural differences between mycobacterial AcpM and *E. coli* ACP may affect their recognition by these enzymes (Schaeffer *et al.*, 2001).

The design and synthesis of several TLM derivatives have led to the recognition of compounds more potent both *in vitro* against fatty acid and mycolic acid biosynthesis and *in vivo* against *M. tuberculosis* (Senior *et al.*, 2003). Genotypic analysis of resistance to INH in *M. tuberculosis* showed that mutations in KasA were present in 16 out of 160 strains (10 %). However, Gly312→Ser was also present in 6 of 32 (19 %) susceptible strains (Lee *et al.*, 1999). Mdluli *et al.* (1998) have shown that INH treatment of *M. tuberculosis* inhibits mycolic acid synthesis and is accompanied by a marked up-regulation of both AcpM and KasA, which are linked genetically. They also proposed KasA to be an INH target in *M. tuberculosis*, on the basis of the detection of an interaction between INH and AcpM. A recent study suggests that KasA belongs to FAS-II and participates in mycolic acid biosynthesis (Kremer *et al.*, 2000b). Therefore due to the importance of the FAS-II elongation process and the possible

participation of KasA in relation to INH inhibition (Mdluli *et al.*, 1998), a number of studies were initiated to determine the role of KasA in mycolic acid biosynthesis. Firstly, the *in vivo* influence of KasA on mycolate production (α versus α′). Secondly, analysis *in vitro* of its overall effect on FAS-II activity within an enriched cytosolic protein fraction, and thirdly, its precise enzymic activity in a classical condensation assay using highly purified recombinant KasA, as well as its sensitivity to cerulenin, a well-known KAS inhibitor (Campbell & Cronan, 2001). Our studies demonstrate that INH does not target KasA as observed *in vitro*, consistent with InhA being the primary target of INH (Kremer *et al.*, 2003). KasA was assayed *in vitro* (Kremer *et al.*, 2002c) in the presence of various drugs and activated INH did not inhibit KasA activity, whereas CER was shown to be an effective inhibitor which demonstrated that INH does not target KasA (Kremer *et al.*, 2003).

The focus of the following study was to investigate the enzymological properties of KasA and analyse the role of particular residues by site-directed mutagenesis.

3.2 Materials and methods

3.2.1 Cloning

KasA was cloned according to the conditions for PCR reported in Table 3.1. The *kasA* gene was amplified from genomic DNA isolated from *M. tuberculosis* H37Rv. PCR primers were created to allow the restriction of the PCR product followed by ligation into an appropriately digested expression vector.

Table 3.1 (A) PCR recipe and (B) PCR conditions for *kasA*

(A)	
2 µl	5′ Cloning Primer (100 pmol/ml)
2 µl	3′ Cloning Primer (100 pmol/ml)
1 µl	Genomic DNA (H37Rv) (1 µg/µl)
2 µl	dNTP's (25 mM each type)
10 µl	Thermostable Buffer (10x)
8 µl	DMSO (optional)
2 µl	$MgSO_4$ (optional) (100 mM)
1 µl	Deep Vent ™ DNA polymerase (2000 U/ml)
72 µl	ddH_2O
100 µl	Final Volume

(B)		
95°C	5 min	35 cycles
84°C	∞	
94°C	1 min	
68°C	1 min 20 sec	
72°C	10 min	
4°C	∞	

Cloning primers used in the PCR reactions can be found in Appendix 1. The PCR product was electrophoresed on 1 % agarose. DNA was visualised using EtBr. The appropriate DNA fragment was excised from the gel and extracted using the Qiagen Gel extraction kit. The PCR product purified using the Qiagen Gel extraction kit, the cut DNA fragment was again extracted from the gel was digested. The concentration of the cut PCR product and similarly digested plasmid DNA was assessed by DNA electrophoresis and EtBr staining. The cut PCR product was ligated with the cut vector. The ligation mix was used to transformed *E. coli* TOP 10 cells. Screening was performed by double digestion (pET28a, pET23b, pQE60) or PCR screening with orientation-specific primers and the lower cloning primer (pVV16, pSD26) followed by analysis by electrophoresis on 1 % agarose and staining with EtBr. The DNA fidelity of the constructs were verified by sequencing.

3.2.2 Cloning of *M. smegmatis* kasA and expression in *M. smegmatis*

PCR amplification of *kasA* was performed using *M. smegmatis* genomic DNA in the presence of the upstream primer 5′-TGA CCA GGC CTT CCA CTG CCA ACG GCG GTT AC-3′ and the downstream primer 5′-GGA ATT CCG CTC TGG CCG TCG AGC AGC TTC TTC-3′ under the conditions in Table 3.3. The downstream primer contains an *Eco*RI restriction site (underlined). The 1407 bp fragment corresponding to the *kasA* coding sequence was then digested with *Eco*RI and ligated with the previously *Mlu*NI/*Eco*RI-restricted pMV261. The resulting plasmid, designated pMV261-*kasA*_MSG, was used to transform *M. smegmatis*. A further construct pMV261-*kasA*_mtb from Kremer *et al.* (2000b) was used to transform *M. smegmatis*.

3.2.3 Expression and purification of *M. tuberculosis* AcpM and mtFabD (malonyl-CoA :ACP transacylase) in *E. coli*

Both pET28a-*acpM* and pET28a-*mtfabD* constructs by Kremer *et al.* (2001a) were utilised in the production of the recombinant proteins. The vectors were used to transformed *E. coli* C41(DE3) and purification of recombinant holo-AcpM, C_{16}-AcpM and mtFabD was achieved using conditions described previously in Kremer *et al.* (2001a). An overnight culture of *E. coli* C41 (DE3) carrying pET28a-*acpM* was used to inoculate 1 L of LB broth supplemented with 25 μg/ml kanamycin and incubated at 37°C until $OD_{600\,nm}$ was 0.75. The culture was then induced with 1 mM IPTG. Growth was continued overnight at 16°C, and cells were harvested by centrifugation. The pellet was resuspended in 20 mM phosphate buffer (pH 7.4), 0.5 M NaCl, 50 mM imidazole, containing DNase, RNase. Bacteria were disrupted by passing twice

through a French pressure cell, and the resulting extract was clarified by centrifuged at 27,000 × *g* for 60 min at 4 °C. The supernatant was collected and applied onto a Ni^{2+}-charged His-Trap column (1 ml). The column was extensively washed with phosphate buffer (pH 7.4), 0.5 M NaCl, 50 mM imidazole and eluted with a stepwise gradient of imidazole (50-500 mM). 1 ml fractions were collected, and the presence of AcpM was detected by electropheresis on 15 % SDS-PAGE and staining with coomassie blue stain. Fractions containing pure AcpM were pooled, dialyzed against 50 mM Tris-HCl (pH 7.4), 50 mM NaCl, 2 mM EDTA, 10 % glycerol, and stored at -20°C. The *M. tuberculosis* holo-AcpM obtained above was further purified using a Thiopropyl-Sepharose 6B column equilibrated with binding buffer (0.1 M Tris-HCl (pH 7.5), 0.5 M NaCl). The products obtained from the Ni^{2+}-charged His-Trap column were loaded onto the column and after several washes with binding buffer; holo-AcpM was eluted using binding buffer containing 50 mM β-mercaptoethanol. Purified holo-AcpM was dialyzed against 50 mM Tris-HCl (pH 7.4), 50 mM NaCl, 2 mM EDTA, 10 % glycerol and stored at -20°C. pET28a-*mtfabD* was used to transform strain *E. coli* C41 (DE3) for overproduction of the recombinant protein. Recombinant mtFabD was purified using the same conditions as described above for AcpM, dialyzed against 50 mM Tris-HCl (pH 7.4), 50 mM NaCl, 2 mM EDTA, 10 % glycerol, and stored at -20 °C. The purity of mtFabD was evaluated by electrophoresis on 12 % SDS-PAGE and staining with Coomassie blue stain.

3.2.4 Preparation of cytosolic enzyme fraction

M. smegmatis (30 g [wet weight]) was washed and resuspended in a buffer (30 ml) containing 100 mM potassium phosphate (pH 7.0), 1 mM EDTA, 5 mM dithiothreitol, 5 mM $MgCl_2$, and 2 mM phenylmethylsulfonyl fluoride at 4°C and subjected to probe sonication (1-cm probe; Soniprep 150; MSE Ltd., Crawley, Sussex, United Kingdom) for 10 min in 10x 60-s pulses

with 90-s cooling intervals between pulses. The whole sonicate was centrifuged at 27,000 *x g* for 30 min at 4°C, and the supernatant was recentrifuged at 100,000 *x g* for 1 h to yield the soluble pale yellow cytosolic enzyme fraction with a typical protein concentration of 8 to 10 mg/ml and containing FAS-I and FAS-II enzyme activities. The 40-80 % saturated ammonium sulphate precipitate, containing FAS-I and FAS-II activities, was collected after centrifugation at 27,000 *x g* and dissolved in 3 ml of buffer A (100 mM potassium phosphate buffer, pH 7.0, 1 mM EDTA 1 mM dithiothreitol 5 mM $MgCl_2$) and dialysed overnight against 2 L buffer A.

3.2.5 FAS-I and FAS-II assays

The standard reaction mixture for the incorporation of radioactivity from [2-^{14}C]malonyl-CoA into C_{16} to C_{24} fatty acids catalysed by FAS-I was composed as follows: 100 mM potassium phosphate pH 7.0, 5 mM EDTA, 5 mM dithiothreitol, 300 μM acetyl-CoA, 100 μM NADPH, 100 μM NADH, 1 μM flavin mononucleotide, 500 μM α-cyclodextrin, 20 μM malonyl-CoA, 100,000 cpm of [2-^{14}C] malonyl-CoA, and 100 μl of the cytosolic enzyme preparation (1 to 2 mg of protein) in a total volume of 500 μl. Similarly, the standard reaction mixture for incorporation of radioactivity from [2-^{14}C]malonyl-CoA into C_{24} to C_{30} fatty acids catalysed by FAS-II contained the following: 100 mM potassium phosphate buffer (pH 7.0), 5 mM EDTA, 5 mM dithiothreitol, 100 μM palmitoyl-CoA, 140 μM NADPH, 140 μM NADH, 180 μg of ACP, 40 μM malonyl-CoA, 200,000 cpm of [2-^{14}C]malonyl-CoA, and 100 μl of the cytosolic enzyme preparation (1 to 2 mg of protein) in a total volume of 500 μl.

In both the FAS-I and FAS-II assays, reactions were performed in triplicate at 37°C for 30 min and terminated by the addition of 500 µl of 20 % potassium hydroxide in 50 % methanol at 100°C for 30 min. Following acidification with 300 µl of 6 M HCl, the resultant ^{14}C-labelled fatty acids were extracted three times with petroleum ether. The organic extracts were pooled, washed once with an equal volume of water, and dried in a scintillation vial prior to counting.

3.2.6 KasA assay

Mycobacterial KasA assays were performed according to the condensation assays developed for *E. coli* KAS enzymes (Garwin *et al.*, 1980). Initially, assay components were mixed together in a batch fashion. The amounts stated correspond to those of a single reaction, which are scaled up proportionately to the number of assays performed. Holo-AcpM (40 µg) was incubated on ice for 30 min with β-mercaptoethanol (0.5 µmol) in a total volume of 40 µl. [2-^{14}C]malonyl-CoA (6.78 nmol, 1.66 kBq; Amersham), mtFabD (40 ng) and 25 µl of 1 M potassium phosphate buffer, pH 7.0, were added. The reaction mixture was held at 37°C for 20 min to allow the mtFabD-catalysed transacylation of holo-AcpM using [2-^{14}C]malonyl-CoA to reach equilibrium. C_{16}-AcpM (22.5 µg) was added to obtain a final volume of 89 µl. The reaction mixture was then dispensed into 1.5 ml microcentrifuge tubes and a 10 µl aliquot of KasA (0.25 µg) was added. The reaction was held at 37°C for 1 h, after which the reaction was quenched by the addition of 2 ml of a $NaBH_4$ reducing solution (5 mg/ml $NaBH_4$ in 0.1 M K_2HPO_4, 0.4 M KCl and 30 % (v/v) THF) to convert the β-keto-C_{18}-AcpM into C_{18}-1,3-diol. This reaction was held at 37°C for at least 1 h. The completely reduced product was extracted twice with 2 ml of water-saturated toluene. The combined organic phases from both

extractions were pooled and washed with 2 ml of toluene-saturated water. The organic layer was removed and dried under a stream of nitrogen in a scintillation vial. The $^{14}C_{18}$-1,3-diol product was then quantified by liquid scintillation counting using 5 ml of EcoScintA (National Diagnostics, Hull, U.K.).

3.2.7 Site-directed mutagenesis

3.2.7.1 QuikChange site-directed mutagenesis

The QuikChange site-directed mutagenesis kit is used to make point mutations, switch amino acids, and delete or insert single or multiple amino acids. It utilises a *PfuTurbo* DNA polymerase to replicate both plasmid strands with high fidelity and without displacing the mutant oligonucleotide primers. The basic procedure utilises a supercoiled double-stranded DNA vector with the insert of interest and two complementary synthetic oligonucleotide primers containing the desired mutation. The oligonucleotide primers are extended during temperature cycling by *PfuTurbo* DNA polymerase using the vector as a template, generating a mutated plasmid containing staggered nicks. The *Dpn* I endonuclease (target sequence: 5′-Gm_6ATC-3′) treatement of the PCR reaction digests the parental DNA template and selects for the mutation-containing DNA which can be selected for as *per* the normal plasmid.

Possible interesting functional amino acid residues for site-directed mutagenesis were identified by sequence alignments and structural analysis of KasA homologues. The QuikChange site-directed mutagenesis kit was used to make point mutations of pET28-*kasA* using the primers described in Table 3.2. and the PCR reaction and recipes as shown below in Table 3.3.

Table 3.2 Site directed mutagenesis primers for *kasA* catalytic triad mutations. Mutagenic substitutions are highlighted in bold

Cys171→Ala	UP	5'	CCC CGG TGT CGG CC**G** **C**TT CGT CGG GCT CGG 3'
	DWN	3'	GGG GCC ACA GCC GG**C** **G**AA GCA GCC CGA GCC 5'
His311→Ala	UP	5'	GAC CAC GTC AAC GCG **GC**C GGC ACG GCG ACG 3'
	DWN	3'	CTG GTG CAG TTG CGC **CG**C CCG TGC CGC TGC 5'
Lys340→Ala	UP	5'	GCG GTG TAC GCG CCG **GC**G TCT GCG CTG GGC CAC 3'
	DWN	3'	CGC CAC ATG CGC GGC **CG**C AGA CGC GAC CCG GAG 5'
His345→Ala	UP	5'	GAA GTC TGC GCT GGG C**GC** CTC GAT CGG CGC G 3'
	DWN	3'	CTT CAG ACG CGA CCC G**CG** GAG CTA GCC GCG C 5'

Table 3.3 PCR recipe (A), PCR conditions (B) for QuikChange site-directed mutagenesis of KasA

(A)	
5 µl	10x Reaction buffer
1 µl	dsDNA template (pET28-*kasA*)
1 µl	5' Mutant primer (100 pmol/ml)
1 µl	3' Mutant primer (100 pmol/ml)
1 µl	dNTP mix (25mM each type)
40 µl	ddH_2O
1 µl	*PfuTurbo* DNA polymerase (2500 U/ml)

(B)		
95°C	30 seconds	
95°C	30 seconds	16 cycles
55°C	1 minute	
68°C	1 minute/Kb	

After the PCR reaction, the methylated double stranded template DNA was removed by digestion with *DpnI*. The digest mixture was transformed into supercompetent *E. coli* XL1-Blue cells by heat shock and plated on selective media containing the relavent antibiotic.

3.2.7.2 Megaprimer mutagenesis

The mega primer method involved amplification of a "mega primer" which contained the required point mutation. Using the upper cloning primer and the lower mutagenic primer, a standard PCR reaction was used to create the "megaprimer" (Sarkar & Sommer, 1990). In cases where the PCR reaction did not yield significant quantities of product, the opposite primers were used, i.e. the lower cloning primer and the upper mutagenic primer. The products from this "Phase I" PCR were separated on 1 % agarose, and visualised with EtBr staining. DNA was extracted from an excised block of agarose using the Qiagen Gel extraction kit. The PCR mix in Table 3.3 was modified for the "Phase II" reaction. The extracted mega primer (24 µl) and 1 µl of the opposite cloning primer was used in the reaction. Electrophoresis of the Phase II PCR products confirmed the generation of full-length *kasA* alleles, which were purified and digested with *EcoRI* and *NheI* and assessed by DNA electropheresis. The cut DNA fragment was again extracted from the gel, and ligated with cut pET28a vector. The ligation reaction was carried out at 16°C overnight and the products were used to transform *E. coli* Top 10 cells. Transformants were cultured in liquid media containing kanamycin overnight at 37°C and plasmid DNA extracted using the Qiagen Miniprep Plasmid extraction kit. Screening for recombinant plasmids was performed by double digestion with *EcoRI* and *NheI* and analysis by electrophoreses on 1 % agarose and visualisation with EtBr. The DNA insertions were sequenced to verify the absence of PCR anomalies and to verify the fidelity of the construct. The primers used in the mega primer method with mutations (bold) and cloning indicated in Table 3.4.

Table 3.4 Site directed mutagenesis primers for *kasA* further mutations. Mutagenic substitutions are highlighted in bold

Asp66→Asn	UP	5'	GGG GGT CAC CTC AAG **A**AT CCG GTC GAC AGC 3'
	DWN	3'	CCC CCA GTG GAG TTC **T**TA GGC CAG CTG TCG 5'
Gly115→Ala	UP	5'	GTT GTT GTC GGC ACC G**C**T CTA GGT GGA GCC 3'
	DWN	3'	CAA CAA CAG CCG TGG C**G**A GAT CCA CCT CGG 5'
Lys116→Ala	UP	5'	GTT GTC GGC ACC GGT **GC**A GGT GGA GCC GAG 3'
	DWN	3'	CAA CAG CCG TGG CCA **CG**T CCA CCT CGG CTC 5'
Ala119→Trp	UP	5'	GGC ACC GGT CTA GGT GGA **TGG** GAG AGG ATT GTC GAG 3'
	DWN	3'	CCG TGG CCA GAT CCA CCT **ACC** CTC TCC TAA CAG CTC 5'
Arg121→Lys	UP	5'	CTA GGT GGA GCC GAG A**A**G ATT GTC GAG AGC 3'
	DWN	3'	GAT CCA CCT CGG CTC T**T**C TAA CAG CTC TCG 5'
Ala141→Trp	UP	5'	GTG TCC CCG CTG **TGG** GTT CAG ATG ATC ATG 3'
	DWN	3'	CAC AGG GGC GAC **ACC** CAA GTC TAC TAG TAC 5'
Pro206→Ala	UP	5'	CCC ATC GAG GCG CTG **G**CC ATC GCG GCG TTC 3'
	DWN	3'	GGG TAG CTC CGC GAC **C**GG TAG CGC CGC AAG 5'
Phe210→Ala	UP	5'	CTG CCC ATC GCG GCG **GC**C TCC ATG ATG CGG 3'
	DWN	3'	GAC GGG TAG CGC CGC **CG**G AGG TAC TAC GCC 5'
Val142→Ala	UP	5'	GTG TCC CCG CTG GCC G**C**T CAG ATG ATC ATG 3'
	DWN	3'	CAC AGG GGC GAC CGG C**G**A GTC TAC TAG TAC 5'
Gly269→Ser	UP	5'	CGA TTG CTG GGT GCC **A**GT ATC ACC TCG GAC 3'
	DWN	3'	GCT AAC GAC CCA CGG **T**CA TAG TGG AGC CTG 5'
Gly312→Ser	UP	5'	CAC GTC AAC GCG CAC **A**GC ACG GCG ACG CCT 3'
	DWN	3'	GTG CAG TTG CGC GTG **T**CG TGC CGC TGC GGA 5'
Gly387→Asp	UP	5'	GAC GTC GTC GCC G**A**C GAA CCG CGC TAT GGC 3'
	DWN	3'	CTG CAG CAG CGG C**T**G CTT GGC GCG ATA CCG 5'
Phe404→Ala	UP	5'	GTC AAC AAC TCG TTC GGG G**C**C GGC GGC CAC AAT GTG 3'
	DWN	3'	CAG TTG TTG AGC AAG CCC C**G**G CCG CCG GTG TTA CAC 5'
Phe413→Leu	UP	5'	CAC AAT GTG GCG CTT GCC TT**G** GGG CGT TAC 3'
	DWN	3'	GTG TTA CAC CGC GAA CGG AA**C** CCC GCA ATG 5'
Asp231→Ala	UP	5'	GCC TCC CGC CCG TTC G**C**C AGG GAC CGC GAC 3'
	DWN	3'	CGG AGG GCG GGC AAG C**G**G TCC CTG GCG CTG 5'
Thr315→Ala	UP	5'	GCG CAC GGC ACG GCG **G**CG CCT ATC GGC GAC 3'
	DWN	3'	CGC GTG CCG TGC CGC **C**GC GGA TAG CCG CTG 5'
Phe413→Leu	DWN	5'	GAT CGA TCG AAT TCT TAG TAA CGC CC**C** AAG GCA AGC
			GCC ACA TTG TG 3'
Phe402→Val	DWN	5'	GAT CGA TCG AAT TCT TAG TAA CGC CCG AAG GCA AGC
			GCC ACA TTG TGG CCG CCG AAC CCG **TG**C GAG TTG TTG ACT GC 3'
Phe404→Val	DWN	5'	GAT CGA TCG AAT TCT TAG TAA CGC CCG AAG GCA AGC
			GCC ACA TTG TGG CCG CCG **TG**C CCG AAC GAG TTG TTG ACT GC 3'

3.2.8 FAMEs and MAMEs analysis

For the expression study, *M. smegmatis* mc^2155 and *M. chelonae* were transformed by electroporation with the pMV261 constructs and cultured to mid-log phase in 50 ml of Sauton's medium containing 50 μg/ml hygromycin, 50 μCi of [1,2-^{14}C]acetate (50–62 mCi/mmol, Amersham Pharmacia Biotech) was added to the cultures followed by further incubated at 37°C for 4-6 hrs for *M. smegmatis* mc^2155 and 30°C for 8 hrs for *M. chelonae*.

The [^{14}C]-labelled cells were harvested by centrifugation and washed successively with PBS. The [^{14}C]-labelled cells were then subjected to alkaline hydrolysis using 15 % aqueous tetrabutylammonium hydroxide at 100°C overnight. 4 ml of CH_2Cl_2, 300 μl of CH_3I and 2 ml of water was added to the reaction mixture followed by mixing for 30 min. The upper aqueous phase was discarded, and the lower, organic phase was washed twice with water and evaporated to dryness. Methyl esters were re-dissolved in diethyl-ether, and then the solution evaporated to dryness. The final residue was dissolved in 200 μl of CH_2Cl_2 and the incorporation [1,2-^{14}C]acetate was quantified by liquid scintillation counting of a 5 μl of the resulting solution of FAMEs and MAMEs. Approximately 100,000 cpm of the FAMEs/MAMEs mix were subjected to TLC using silica gel plates (5735 silica gel 60F254; Merck, Darmstadt, Germany), developed in petroleum ether-acetone (95:5). Autoradiograms were produced by overnight exposure of Kodak X-Omat AR film to reveal [^{14}C]-labelled fatty acid and mycolic acid methyl esters.

3.2.9 Anaylsis of recombinant protein solubility

The pET28a-*kasA* and pET23b-*kasA* plasmids were used to transform the expression host *E. coli* C41 (DE3). However pQE60-*kasA* was expressed in *E. coli* M15 (pREP4). The pVV16-*kasA* and pSD26-*kasA* vectors were expressed in *M. smegmatis* mc^2155. An overnight pre-culture of the transformed cells was used to inoculate 1 litre of Terrific broth (*E. coli*) or 1 litre of Sautons broth (*M. smegmatis*), which was incubated at 37°C until $A_{600\ nm}$ = 0.6 and then recombinant protein production was induced with 1 mM isopropyl-β-D-thiogalactoside (*E. coli*) or 0.02 % acetamide (*M. smegmatis* pSD26-*kasA*). The *E. coli* cultures were grown either for a further 4 h at 37°C or 16 h at 16°C, and then harvested by centrifugation. The *E. coli* cell pellets were resuspended in 4 ml/g of cells (wet weight) in 20 mM phosphate buffer, pH 7.9, 0.5 M NaCl, 10 mM imidazole containing DNase and RNase, whereas the *M. smegmatis* cell pellets were resuspended in 4 ml/g of cells (wet weight) in 50 mM MOPS, pH 7.9, 50 mM MgCl, 5 mM β-mercaptoethanol and 10 mM imidazole. Bacteria were disrupted by sonicaton or French press and the resulting extracts clarified by centrifuged at 27,000 x *g* for 60 min at 4°C. The supernatants were stored on ice and the pellet was resuspended in buffer containing 8 M urea, and then centrifuged at 27,000 x *g* for 60 min at 4°C. Protein concentrations were estimated by BCA assay kit (Pierce) and purity assessed by electrophoresis of 25 μg of protein on 12 % SDS-PAGE, visualised by Coomaisse blue-stain.

3.2.10 Purification of KasA and muteins

Purification of protein extracts was performed using Ni^{2+} Hi-trap Chelating Sepharose™ fast flow matrix columns (1 ml or 5 ml). Columns were charged using 2 column volumes of 0.1 M nickel chloride and washed with 4 column volumes of 0.02 M sodium phosphate and 0.5 M

NaCl pH 7.4 to remove excess nickel chloride to equilibriate the column. The clarified lysate was applied to the column (maximum of 10 ml for a 1 ml column and 50 ml for a 5 ml column) and non-absorbed material collected for post-column analysis. Bound proteins were eluted using a step gradient of imidazole concentrations (10 mM – 500 mM). Fractions were collected and protein concentration determinded by BCA (Pierce), 5 μg of each protein fraction was analysed by 12 % SDS-PAGE. Purified protein samples were dialysed against 20 mM Tris-HCl pH 7.6, 50 mM NaCl containing 1 mM β-mercaptoethanol.

3.3 Results

The initial work on the biochemistry of KasA was done in conjunction with Dr. Lynn Dover and Dr. Laurent Kremer in the first year of my PhD program, which I further developed through identification and construction of additional site-directed mutants of KasA.

3.3.1 Over-expression of KasA from *M. tuberculosis* is associated with a decrease in α'-mycolate production

To further clarify whether KasA participates in meromycolate elongation, *M. smegmatis* was transformed with pMV261-*kasA* from *M. tuberculosis* H37Rv. The effect of KasA overproduction on mycolic acid composition was analysed by TLC after *in vivo* labelling of mycobacterial FAMEs and MAMEs with [1,2-^{14}C]acetate. TLC analysis revealed a significant reduction of the α'-subclass in *M. smegmatis* over-expressing mtKasA (pMV261-*kasA*_MTB) when compared with the control strain carrying pMV261 (Figure 3.1A). In a compensatory

fashion, over-expression of KasA led to an increase of *α*- and epoxymycolates (Figure 3.1A). The ratio of α' to long chain (combined α and epoxy) was 24 and 76 % in the control strain, whereas the ratio was 10 and 90 % in the strain overproducing KasA. To determine whether this effect was restricted only to *M. smegmatis* as a host or extended to other mycobacterial species, *M. chelonae* was transformed either with the control plasmid pMV261 or with pMV261-*kasA*_MTB. As shown in Figure 3.1B, over-expression of mtKasA in *M. chelonae* resulted in an increase of the α/α' ratio, the strain carrying pMV261 being approximately 1:2, and that carrying pMV261-*kasA* approximately 1:3.5. These results suggest that over-expression of the *M. tuberculosis* KasA protein consistently increases the α/α' ratio in different mycobacterial species, suggesting that KasA participates in the *in vivo* elongation process of the meromycolic chain. Although, α'-mycolic acids can be considered as precursors of the α-mycolic acids, it is not known why these precursors accumulate in some mycobacterial species, such as *M. smegmatis*. Possible explanations could be that KasA activity may be deficient in these species due to the absence of *kasA*, which seems unlikely, or possibly due to a defect in KasA expression or regulation.

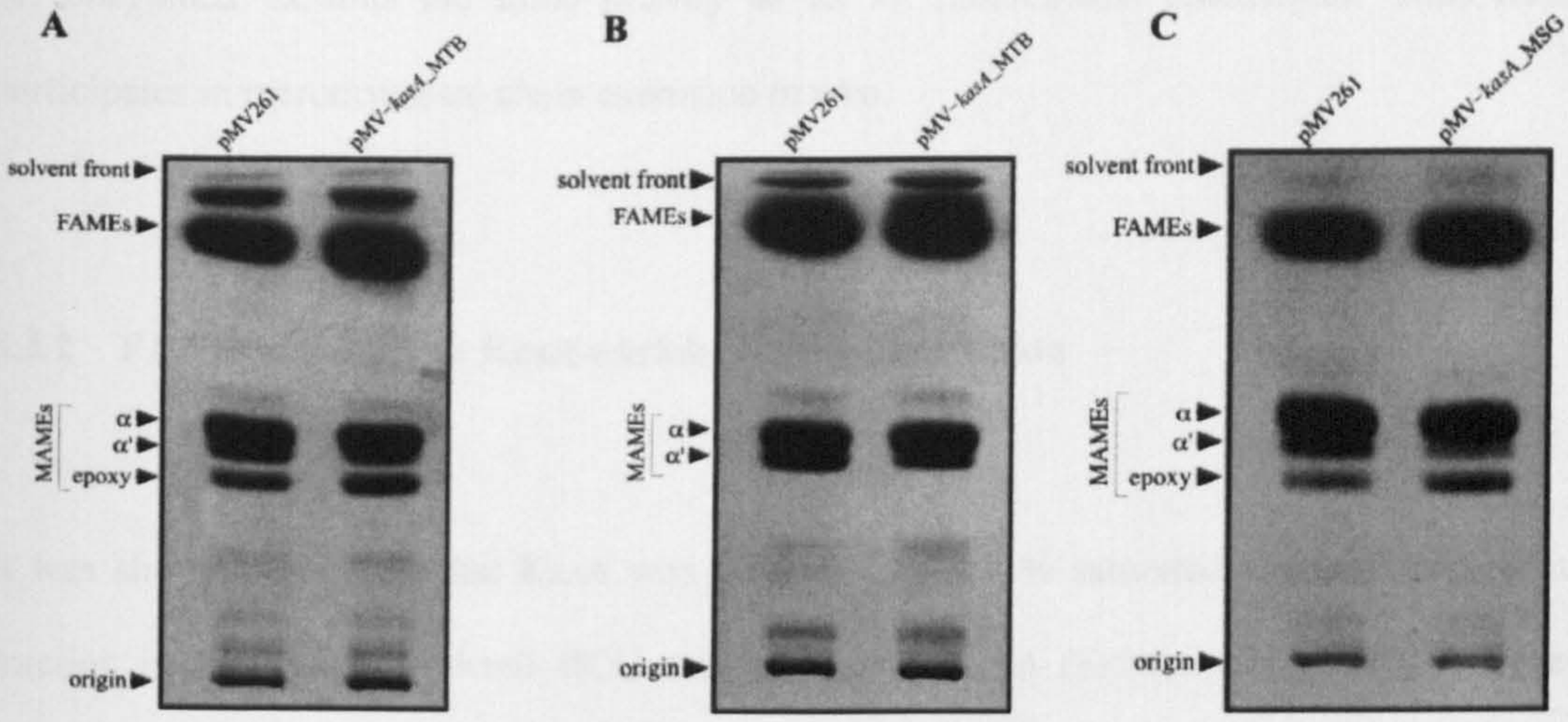


Figure 3.1 **Normal-phase TLC of FAMEs and MAMEs in mycobacteria over-producing KasA from *M. tuberculosis* or *M. smegmatis*. Visualisation of the incorporation of [^{14}C]acetate into lipids to produce radiolabelling FAMEs and MAMEs. The radio-labelled lipids were extracted with 15 % tetrabutylammonium hydroxide at 100°C overnight. The corresponding fatty acid methyl esters (*FAMEs*), α-mycolates, α'-mycolates and epoxy-mycolates were prepared and subjected to TLC (5735 silica gel 60F_{254}, Merck), followed by exposure to Kodak X-Omat film. (A) *M. smegmatis* harbouring pMV261 containing KasA from *M. tuberculosis* (B) *M. chelonae* harbouring pMV261 containing KasA from *M. tuberculosis* (C) *M. smegmatis* harbouring pMV261 containing KasA from *M. smegmatis***

The genomic organisation of this operon is well conserved in *M. smegmatis* and with a high homology of KasA in both species (Kremer *et al.*, 2002c), it seemed reasonable to assume that over-expression of *kasA* from *M. smegmatis* would have the same effect as over-expression of the *M. tuberculosis* gene. As shown by TLC analysis in Figure 3.1C, *M. smegmatis* over-expressing its homologous gene (pMV261-*kasA*_MSG) presents a significant reduction of the α'-subclass when compared with the control strain. This suggests that the gene product from

M. smegmatis exhibits the same activity as its *M. tuberculosis* counterpart. Thus KasA participates in meromycolate chain extension *in vivo*.

3.3.2 FAS-II activity in a KasA-enriched cytosolic fraction

It was shown previously that KasA was present in 40–80 % saturated ammonium sulphate fraction isolated from *M. bovis* BCG over-expressing KasA (Kremer *et al.*, 2000b). When enriched fractions derived from *M. smegmatis* pMV261 were assayed for FAS-I and FAS-II activity, the optimum reaction conditions were obtained in the presence of C_{14}-CoA and C_{16}-CoA, presumably following transacylation to AcpM, mtFabH converts the acyl-CoA followed by a complete cycle of FAS-II into acyl-AcpM, the preferred substrates for further elongation *via* KasA and FAS-II. It is clear long-chain acyl-ACP substrates (C_{18} or C_{20}) do not enter FAS-II as efficiently, and almost no activity was seen with short-chain acyl-ACPs (Figure 3.2). The FAS-II elongation activity was also assessed with the enriched cytosolic fraction isolated from *M. smegmatis* (pMV261-*kasA*_MTB). As shown in Figure 3.2, although the substrate specificity was conserved, the specific activity was enhanced significantly regardless of the primer used. This indicates that KasA produced in *M. smegmatis* is active enzymically and represents a key enzyme within FAS-II. To characterize the compounds synthesised by FAS-II, FAMEs were analysed by C_{18} reversed-phase TLC. The inset (Figure 3.2) illustrates the chain-length distribution of the FAMEs produced for primers C_{12}-CoA/ACP and C_{14}-CoA/ACP and shows that the yield decreases sharply with increasing chain length, an observation also made by others (Marrakchi *et al.*, 2000).

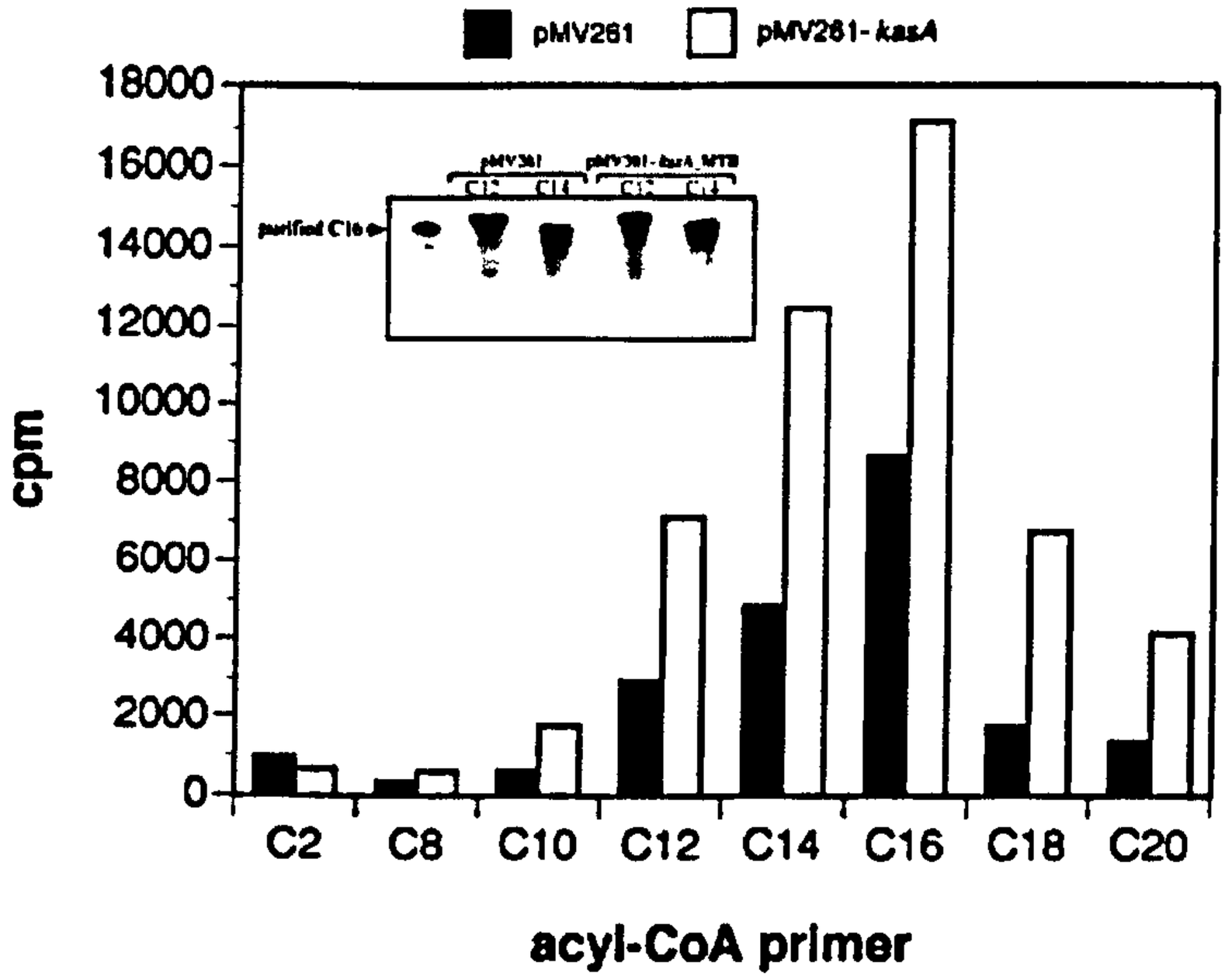


Figure 3.2 **FAS-II enzymatic assay analysis of *M. smegmatis* extracts containing over-expressed *M. tuberculosis* KasA. The FAS-II assay was performed using different primers (C_2–C_{20}-CoA) *E. coli* ACP and the enriched soluble-cytosolic fraction from *M. smegmatis* harbouring either pMV261 or pMV261-*kasA*_MTB. The inset shows the chain length distribution of FAMEs synthesised from C_{12}-CoA and C_{14}-CoA in the FAS-II assay. Radiolabelled FAMEs (7000 c.p.m.) were loaded on to C_{18} reverse-phase TLC in each lane, developed in $CHCl_3/CH_3OH$ (2 : 3, v/v), and the products were revealed by exposure to Kodak X-Omat film after a period of 4 days. Methyl [1-^{14}C]palmitate was used as a reference standard.**

.

In light of the recent results the primer-specific activity profile observed in this assay system is wholly consistent with the substrate specificity of pure recombinant mtFabH with malonyl-ACP and strengthens the argument that mtFabH links FAS of mycobacteria (Figure 2.12). Repetition of this experiment using AcpM instead of *E. coli* ACP would be interesting, maybe longer acyl-CoA primers could be elongated efficiently by FAS-II.

3.3.3 KasA activity and biochemical analysis of active site mutants

To determine directly the nature of the enzyme activity of KasA, it was a requirement that the key recombinant proteins holo-AcpM, C_{16}-AcpM, mtFabD and KasA were produced in *E. coli* to develop an *in vitro* elongation–condensation assay (Kremer *et al.*, 2002c). We were able to obtain the pure KasA enzyme by loading the clarified soluble lysate onto a Ni^{2+}-charged His-Trap column (Amersham). KasA was obtained as a pure protein migrating at approximately 45 kDa, which is in agreement with the theoretical molecular mass.

When incubated together with other key components, KasA elongated C_{16}-AcpM *via* a condensation reaction with [2-^{14}C]malonyl-AcpM, which was generated *in situ* by the action of mtFabD on holo-AcpM and [2-^{14}C]malonyl-CoA. Within the linear range of a KasA dose–response curve, we observed a specific activity of 96 $nmol^{-1} \cdot min^{-1} \cdot mg^{-1}$ for the elongation–condensation reaction (Figure 4.2A). To test the specificity of KasA with regard to acylated primers (acyl-CoA versus acyl-AcpM), we compared its activity with C_{16}-CoA and C_{16}-AcpM. When C_{16}-CoA was incorporated (in equimolar concentrations to C_{16}-AcpM) into the condensation assay, product formation was reduced dramatically (Figure 4.2A) to 4 % of that observed with C_{16}-AcpM. This is consistent with our current understanding of mycobacterial FAS-II and the notion that KasA is a KAS involved in meromycolate extension, which utilises acyl-AcpM bound precursors.

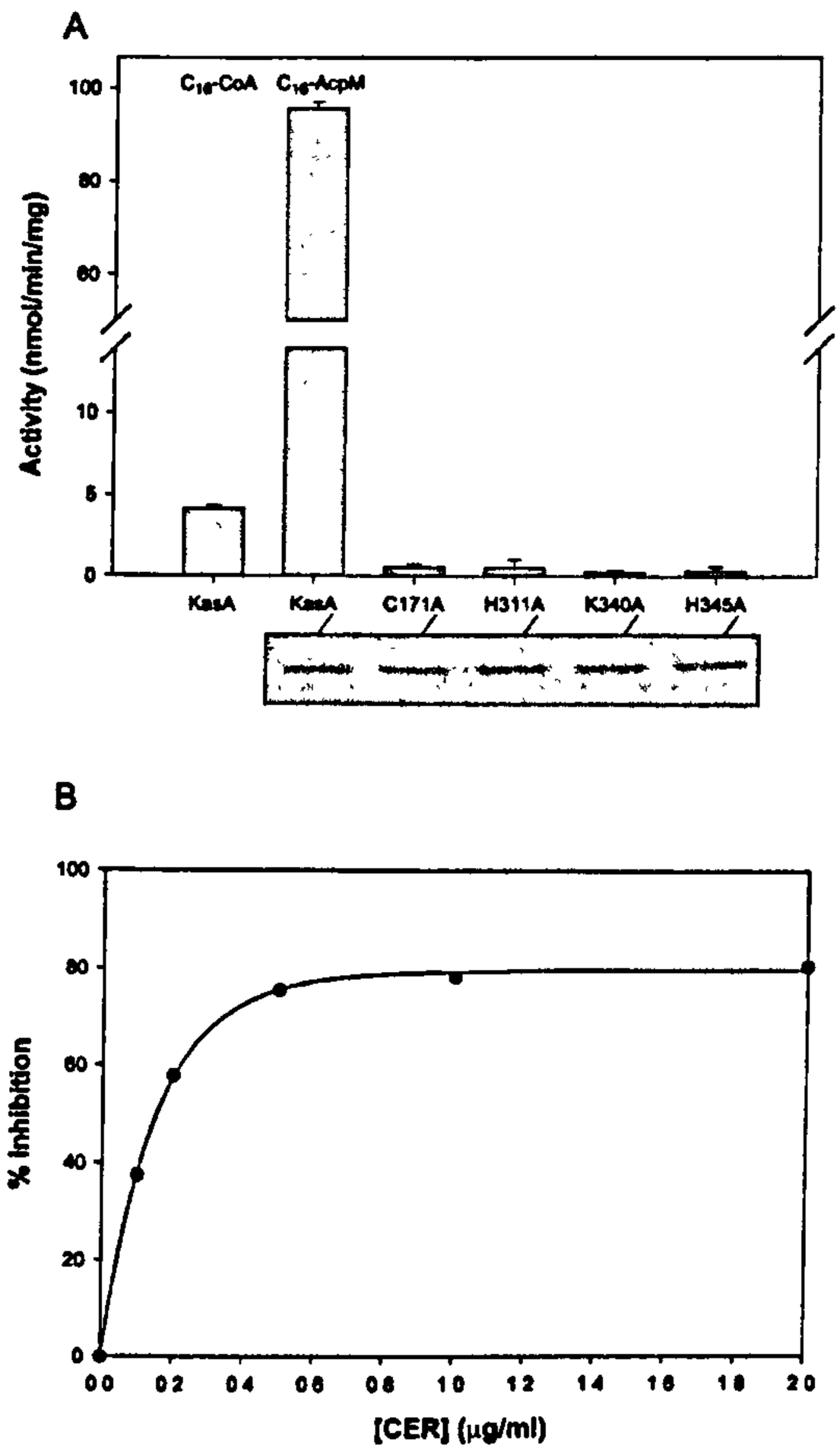


Figure 3.3 **(A) KasA active site mutant analysis. *kasA* and its various mutants were expressed as His-tagged recombinant proteins in *E. coli* C41 (DE3) and purified as described in the materials and methods section. They were assayed for KAS activity as described in the section 4.2.2. The condensation rates for KasA were tested with equimolar amounts of either C_{16}-AcpM or C_{16}-CoA as acyl sources. The activity of each of the mutants was assayed using C_{16}-AcpM. (B) The condensation rates for KasA using C_{16}-AcpM as a substrate in the normal KasA assay were determined in the presence of increasing concentrations of cerulenin.**

Previously it was shown that cerulenin exhibits potent *in vivo* anti-mycobacterial activity (Kremer *et al.*, 2000b). Although over-expression of *kasA* in *M. bovis* BCG did not generate an increased resistance against cerulenin, it was suggested that the *in vivo* anti-mycobacterial

activity of cerulenin was primarily *via* inhibition of *de novo* fatty acid synthase catalysed by FAS-I (Kremer *et al.*, 2000b). Earlier studies have shown that cerulenin is a potent inhibitor of both FAS-I and FAS-II (Omura, 1976). In addition, the *in vitro* inhibition of *E. coli* KAS FabB and FabF by cerulenin has also been well documented for cerulenin (Campbell & Cronan, 2001). It was thus reasonable to assume that cerulenin may also inhibit the β-ketoacyl-AcpM synthase activity of KasA. Increasing concentrations of cerulenin were added to the *in vitro* condensation assay. As shown in Figure 3.3B, a marked dose–response inhibition of KasA activity was observed with cerulenin providing an IC_{50} value of 0.15 μg/ml (0.67 μM). The IC_{50} values of the purified *E. coli* Kas proteins were reported as 3 and 20 μM for FabB and FabF, respectively (Price *et al.*, 2001), suggesting that the mycobacterial condensing enzyme is more susceptible to cerulenin inhibition *in vitro* than the related *E. coli* enzymes.

Siggaard-Anderson has identified a number of important amino acid residues in the *E. coli* FabB protein (Siggaard-Andersen, 1993). For instance, Cys163 constitutes one of the active-site catalytic residues. Other residues, such as His298, Lys328 and His333 were also identified and when replaced by Ala, decarboxylation and the overall elongation activity were abolished completely suggesting the importance of equivalent amino acids in catalysis. The recently presented structural model of the *M. tuberculosis* KasA generated by homology modelling (Kremer *et al.*, 2000b) based on the determined X-ray structure of the *E. coli* FabF protein by Huang *et al.* (1998) allowed the identification of these residues in KasA. The three-dimensional structural model showed Cys171, His311, Lys340 and His345 as likely candidates for the catalytic residues of the active site of KasA (Kremer *et al.*, 2000b). Site-directed mutagenesis was used to determine the role of these four residues in KasA activity by replacing them individually with alanine. The mutated KasA proteins were purified and CD-

spectral analysis were performed to confirm that the overall secondary structure of the mutant proteins did not differ significantly from that of the wild-type enzyme.

KasA activity of all the mutant proteins (Cys171, His311, Lys340 and His345 to Ala) was abolished (Figure 3.3A), confirming the importance of these amino acids in KasA activity. It is likely that the His311→Ala, Lys340→Ala and His345→Ala muteins were unable to decarboxylate the malonyl-AcpM and therefore in the absence of the C2 carbanion the condensation event would not be completed. The intermediate is proposed to be held in position by its interaction with His311 and His345 of the active site *via* interactions with the oxygen of the malonyl-AcpM. In the case of the muteins this interaction is likely reduced thus abrogating activity. It has been proposed that the equivalent amino acids facilitate the transacylation event in the *E. coli* homolog FabB (McGuire *et al.*, 2001). Therefore not only may these muteins be deficient in decarboxylation activity but may also have reduce transacylation activity. The transacylation of C_{16}-AcpM to the active site cysteine would not be able to occur in a Cys171→Ala mutein as the thio group required for the transfer would not be present when substituted for an alanine. Although this mutein might be able to decarboxylate malonyl-AcpM, but it could not condense the decarboxylate malonyl-AcpM with the acyl-protein intermediate as it would not be present.

3.3.4 Sequence derived KasA mutational analysis

To gain a more detailed understanding of KasA we initiated a more extensive mutagenesis study. We hoped that a simultaneous biochemical and crystal structure determination would provide further insight. The low yields of KasA muteins means that we required a more

effective purification or expression system as highly concentrated protein preparations are required for the crystallography study. As for mtFabH it was decided that four major areas of the enzyme properties would be assessed by site-direceted mutagenesis; substrate specificity, catalytic triad, protein-protein interaction and antibiotic sensitivity (Table 3.5).

Table 3.5 KasA targets for site-directed mutagenesis, listed by class.

Substrate Specificity/Binding	Catalytic Triad	Other Active Site	Protein-protein Interaction	Antibiotic Sensitivity
Gly115 Lys116 Ala119 Ala141 Val142 Pro206 Phe210 Asp231 Thr315	Cys171 His311 His345	Lys340	Asp231 Thr315	Asp66 Arg121 Gly269 His311 Gly312 His345 Gly387 Phe402 Phe404 Phe413

The crystal structure of mtKasA has not yet been solved and therefore the identification residues to be mutated would be identified by sequence alignments and the literature on homologs of KasA. The development of a series of site-directed mutants of KasA depended greatly on previous work done on *E. coli* FabB, and FabF.

The sequence alignments of related KAS proteins was compiled to allow for the identification of amino acids which were potential candidates for mutagenesis and have been previously shown to be important residues (Figure 3.4). The majority of the mutations selected were based on previous studies with *E. coli* FabB (Olsen *et al.*, 2001), with several key amino acids being identified, Val134, Gly107, Val135, Pro202 and Phe201(Olsen *et al.*, 1999). Residues Gly107, Val134, Val135 and Phe201 line the cavity of the active site and may play a role in

the substrate specificity, therefore it was assumed that these residues play a major role in transacylation of the acyl chain of acyl-AcpM to the active site Cys. The active site in KASI dimer are accessible at the bottom of a narrow tunnel defined with Pro202 being located at the entrance. The possibility that mutation of this residue affects the kinetics of the catalysis is interesting and requires further investigation. Asp223 and Thr307 of FabF were identified as parts of the binding site of the acyl-ACP substrate (Huang *et al.*, 1998). Thr307 together with the N-terminal end of the preceding α-helix form a binding site at the top of the binding cavity for the hydroxyl and phosphate of the phosphopantetheine group of the acyl-ACP. Some distance from the active site pocket is the Asp223 residue which is located on the surface of the enzyme and could be involved in the interaction with ACP (Huang *et al.*, 1998). Therefore it was decided to analyse these mutant enzymes with substitutions at these residues for overall KAS, transacylation and decarboxylation activites.

```
ecFabB      ------------------------------------------------------------MKRA
pfFabB      MNIRCRKEKKKKKKKNINHVNSKRYAFIKRGIPGISKNYFKGFKLYNSREMKNLCETSRV
scFabB      ---------------------------------------------MSPRHAAGPAPGERAA
mtKasA      ---------------------------------------------MSQPSTANG-GFPSV
mbKasI      ---------------------------------------------MQPSTANG-GFPSV
msgKasA     ---------------------------------------------MTRPSTANG-GYPSV
mlKasI      -------------------------------------MIARLMTTSPELVTGKAFPNV

ecFabB      VITGLGIVSSIGNNQQEVLASLREGRSGITFSQELKDSGMRS----HVWGNVKLDTTGLI
pfFabB      VCTGVGVVTGLGIGIEHFWNNIINGYTSIDKITKFDITGMSCGIGSEIKKSDFNPSDYYT
scFabB      VVTGLGLTTSLGGDVPSTWRALLDGGCGVERV-DFGEPAGPAQV--YLAAPAAVDPGTVL
mtKasA      VVTAVTATTSISPDIESTWKGLLAGESGIHALEDEFVTKWDLAV--KIGGHLKDPVDSHM
mbKasI      VVTAVTATTSISPDIESTWKGLLAGESGIHALEDEFVTKWDLAV--KIGGHLKDPVDSHM
msgKasA     VVTAVTATTSIAPDIESTWKGLLAGESGIRVLEDEFVTKWDLPV--RIGGHLVDNIDDHM
mlKasI      VVTGIAMTTALATDAETTWKLLLDNQSGIRMLDDPFIEEFNLPV--RIGGHLLEEFDHQL

ecFabB      DRKVVRFMSDASIYAFLSMEQAIADAGLSPEAYQ--NNPRVGLIAGSGGGSPRFQVFGAD
pfFabB      NKKDVNRNDDCTHYAVAATRLALDDAKLNLEKL---DKDKTGTIIGSGIGGLRFLEKEMK
scFabB      SSAKAAHCDRSAQFALVAAREAVRDAGFPDPSALAGDGSRVAVVVGVGLGGLTSVLEQDH
mtKasA      GRLDMRRMSYVQRMGKLLGGQLWESAGSPEV-----DPDRFAVVVGTGLGGAERIVESYD
mbKasI      GRLDMRRMSYVQRMGKLLGGQLWESAGSPEV-----DPDRFAVVVGTGLGGAERIVESYD
msgKasA     TRLDMRRMSYVQRMSKFLSKQLWENAGAPEV-----DPDRFAVVIGTGLGGGEKIVETYD
mlKasI      TRVELRRMGYLQRMSTVLSRRLWENAGSPEV-----DTNRLMVSIGTGLGSAEELVFSYD

ecFabB      AMRGPRGLKAVGPYVVTKAMASGVSACLATPFKIHGVNYSISSACATSAHCIGNAVEQIQ
pfFabB      TMY-EKGHKRITPYLIPAMIANTPSGYVSIENNIRGISLGMLSACATSGNTIGEAYRYIK
scFabB      RLR-TQGAGRVSPRTIPVMLPNHAAAEVGLMVGAKAGVHAPVSACAAGAEALAQALGMIR
mtKasA      LMN-AGGPRKVSPLAVQMIMPNGAAAVIGLQLGARAGVMTPVSACSSGSEAIAHAWRQIV
mbKasI      LMN-AGGPRKVSPLAVQMIMPNGAAAVIGLQLGARAGVMTPVSACSSGSEAIAHAWRQIV
msgKasA     AMN-EGGPRKVSPLAVQMIMPNGAAAVVGLELGARAGVITPVSACSSGSEAIAHAWRQIV
mlKasI      DMR-ARGMKAVSPLAVQKYMPNGAAAAVGLEHHAKAGVMTPVSACASGSEAIAHAWQQIV

ecFabB      LGKQDIVFAGGGEELCWEMA-CEFDAMGALSTKYNDTPEKASRTYDAHRDGFVIAGGGGM
pfFabB      YKEYDVMICGGTEASITPISFAGFNSLKALCTGYNDNPKKGCRPFDLKRSGFVMGEGSGI
scFabB      DGRADIVVAGGTEAALHPLVLAGFARLR-ALSRRHDDPKGASRPFDADRDGFVMGEGAGM
mtKasA      MGDADVAVCGGVEGPIEALPIAAFSMMR-AMSTRNDEPERASRPFDKDRDGFVFGEAGAL
mbKasI      MGDADVAVCGGVEGPIEALPIAAFSMMR-AMSTRNDEPERASRPFDKDRDGFVFGEAGAL
msgKasA     MGDADFAVCGGVEGGIEALPIAAFSMMR-AMSTRNDDPQGASRPFDKDRDGFVFGEAGAM
mlKasI      LGEADSAICGGVETKIEAVPIAGFSQMRIVMSTKNDNPAGACRPFDRDRDGFVFGEAGAL

ecFabB      VVVEELEHALARGAHIYAEIVGYGATSDGADMVAP--SGEGAVRCMKMAMHGVD---TPI
pfFabB      LILESYEHAIKRNAPIYGEIISYSSECDAYHITAPEPNGKGLTNSIHKALKNANININDV
scFabB      LVVESAAHAAARGARVHGRLTGAGITNDSHHVAQPAPGGPGCAAAVDAALRDAGLVPEQI
mtKasA      MLIETEEHAKARGAKPLARLLGAGITSDAFHMVAPAADGVRAGRAMTRSLELAGLSPADI
mbKasI      MLIETEEHAKARGAKPLARLLGAGITSDAFHMVAPAADGVRAGRAMTRSLELAGLSPADI
msgKasA     MIIETEEHAKARGAKPLARLMGAGITSDAFHMVAPAADGVRAGQAMKRAMETAGLDPKDI
mlKasI      MLIETEDSAKARSANILARIMGASITSDGFHMVAPDPNGERAGHAIARAVHLAGLSPSDI

ecFabB      DYLNSHGTSTPVGDVKELAAIREVFGD--KSPAISATKAMTGHSLGAAGVQEAIYSLLML
pfFabB      KYINAHGTSTNLNDKIETKVFKNVFKDHAYKLYISSTKSMTGHCIGAAGAIESIVCLKTM
scFabB      QHVNAHATATPLGDLGEAQALHSVFAKGVDDVTVSATKGAFGHTLGAAGAIEAVLTVLAL
mtKasA      DHVNAHGTATPIGDAAEANAIRVAGCD---QAAVYAPKSALGHSIGAVGALESVLTVLTL
mbKasI      DHVNAHGTATPIGDAAEANAIRVAGCD---QAAVYAPKSALGHSIGAVGALESVLTVLTL
msgKasA     DHVNAHATATPIGDTAEANALRVAGVE---HAAVYAPKSALGHSIGAVGALESILTVLAL
mlKasI      DHVNAHATGTQVGDLAEAKAINKALCN--NRPAVYAPKSALGHSVGAVGAVESILTVLAL

ecFabB      EHGFIAPSINIEELDEQAAGLNIVTETT---DRELTTVMSNSFGFGGTNATLVMRKLKD
pfFabB      QTNIIPPTINYEYKDPDC-DLNYTPNKYIHAKENIDISLNTNLGFGGHNTALLFKKIVK
scFabB      RDRTAPPACSLERVDPEI-GLNVVGRDPATLPCGPLAAVSTSMGFGGHNVALAFADAS-
mtKasA      RDGVIPPTLNYETPDPEI-DLDVVAGEPR--YGDYRYAVNNSFGFGGHNVALAFGRY--
mbKasI      RDGVIPPTLNYETPDPEI-DLDVVAGEPR--YGDYRYAVNNSFGFGGHNVALAFGRY--
msgKasA     RDGVIPPTLNYETPDPEI-DLDIVAGEPR--YGEYKYAINNSFGFGGHNVALAFGRY--
mlKasI      RDQVIPPTLNLVNLDPDI-DLDVVAGKPR--PGDYRYAVNNSFGFGGHNVAIAFGCY-
```

Figure 3.4 **Comparison of *M. tuberculosis* KasA (mtKasA) with *E.coli* FabB (ecfabB), *Plasmodium falciparum* FabB (pfFabB), *Streptomyces coelicolor* A3(2) FabB (scFabB), *M. bovis* KasI (mbKasI), *M. smegmatis* KasA (msgKasA) and *M. leprae* KasI (mlKasI). Primary amino acid sequences of all the KasA homologues were compared with the predicted protein sequence of KasA. KasA possesses the hallmark features of synthase I including the Cys-His-His and Lys of the active site (*Red*), new mutation sites identified are indicated (yellow)**

INH resistance has been observed in clinical isolates from Singapore with residues Arg121→Lys, Gly269→Ser, Gly387→Asp and Gly312→Ser involved in INH resistance (Lee *et al.*, 1999). Furthermore these substitutions plus Asp66→Asn and Phe413→Leu were identified to confer INH resistance *in vitro* (Mdluli *et al.*, 1998). Gly269→Ser and Phe413→Leu were single mutations and no other alterations were observed in KatG, InhA or AhpC. The mis-sense Phe404→Ala Phe402→Val and Phe404→Val mutations in KAS of *E. coli* lead to resistance to TLM with the IC_{50} increasing from wild-type FabB of 20 μM to 200 μM (Jackowski *et al.*, 2002).

Table 3.6 Summary of proposed mutations of KasA and their homologs; A, *E.coli* FabB; B, *E. coli* FabF; C, clinical isolates of KasA

Proposed Function	*M. tuberculosis* KasA Residue	Homolog Residue	
Substrate Specificity/Binding	Gly115	Gly107	A
	Lys116	Gly108	A
	Ala119	Pro111	A
	Ala141	Val134	A
	Val142	Val135	A
	Pro206	Pro202	A
	Phe210	Phe201	A
	Asp231	Asp223	B
	Thr315	Thr307	B
	Lys340	Lys328	A
Catalytic Triad	Cys171	Cys163	A
	His311	His328	A
	His345	His333	A
Protein-protein Interaction	Asp231	Asp223	B
	Thr315	Thr307	B
Antibiotic Sensitivity	Asp66		C
	Arg121		C
	Gly269		C
	His311	His328	A
	Gly312		C
	His345	His333	A
	Gly387		C
	Phe402	Phe390	A
	Phe404	Phe392	A
	Phe413		C

The additional generation of these mutants of KasA was attempted using the Stratagene QuikChange site-directed mutagenesis Kit. Despite our initial four success' with the catalytic triad mutants using this method the production of further mutations was not successful after several attempts. Manipulation of the standard protocol was attempted. Firstly the annealing and melting temperatures were investigated to see if products were produced under different PCR conditions. The melting temperature was increased from 95°C to 97°C and it was also reduced to 94°C, the annealing temperature was changed over a 10°C range from 63°C to 73°C but with no mutants being observed it was decided that we should increase the number of cycles in the PCR reaction from 16 to 20, 24 and 28. Yet again with no success our attention was redirected towards the recipe of the PCR. Three areas were assessed in the recipe of the PCR reaction; template concentration, primer concentration and DMSO/$MgSO_4$ additions. The template was diluted 1:20 and 1:50 and tested at the range of conditions stated above, no mutants were observed so the primer concentration was change by the addition of extra 1 µl of the 100 pmol/µl primer. With still no mutations we finally turned our attentions to the additions of DMSO and $MgSO_4$, the $MgSO_4$ was added to the standard reaction in increasing concentrations with and without DMSO but still no mutants were generated. The main disadvantage with the QuikChange site-directed mutagenesis method was that we were unable to find out exactly at what point the method was failing therefore it was decided that other methods of site-directed mutagenesis should be tried.

After all the unsuccessful attempts at producing mutants using the QuikChange site-directed mutagenesis Kit a new method was utilised. It required two separate PCR reactions followed by digestion, ligation with the vector and screening, known as the megaprimer method (Figure 3.5). The mega primer method involved cloning of a "mega primer" which contained the required point mutation. The PCR conditions were as stated in section 4.2.7.2, the further

extension of the megaprimer to the full length gene allowed the identification of PCR conditions that have been successful (Figure 3.6). The advantage of this system was that it allowed the identification of the problems in the production of the mutation, and therefore allowed for development of the technique to achieve our goal. Due to the two primers required for the QuikChange site-directed mutagenesis method the mega primer preparation could be performed in both directions. In some cases the PCR reaction did not yield significant quantities of product so the opposite primers were used, i.e. the upper cloning primer and the lower mutagenic primer. This method allowed for the production of either the front or the back segment of the gene containing the mutation. In some cases neither of the fragments would exend to the full length of the gene. In these situations both mega primers were used in the Phase II reaction. A number of the mutants didn't extend in either direction so new primers were designed. With regards to the mutants at the terminus of the gene (Phe402, Phe404, Phe413) it was difficult to see any change in the length of the mega primer after the phase II PCR, therefore new extended primers were designed which included the mutation, the end of the gene and the cloning site (Phe402, Phe404, Phe413).

The cut DNA fragments were then ligated with cut pET28a vector. In some cases the fragment wouldn't ligate into the vector. Therefore it was necessary to vary the usual ligation reaction which was carried out at 16°C overnight. So other conditions were tried including storage on ice overnight and a step gradient of temperatures starting at 4°C increasing to 20°C over 16 hours.

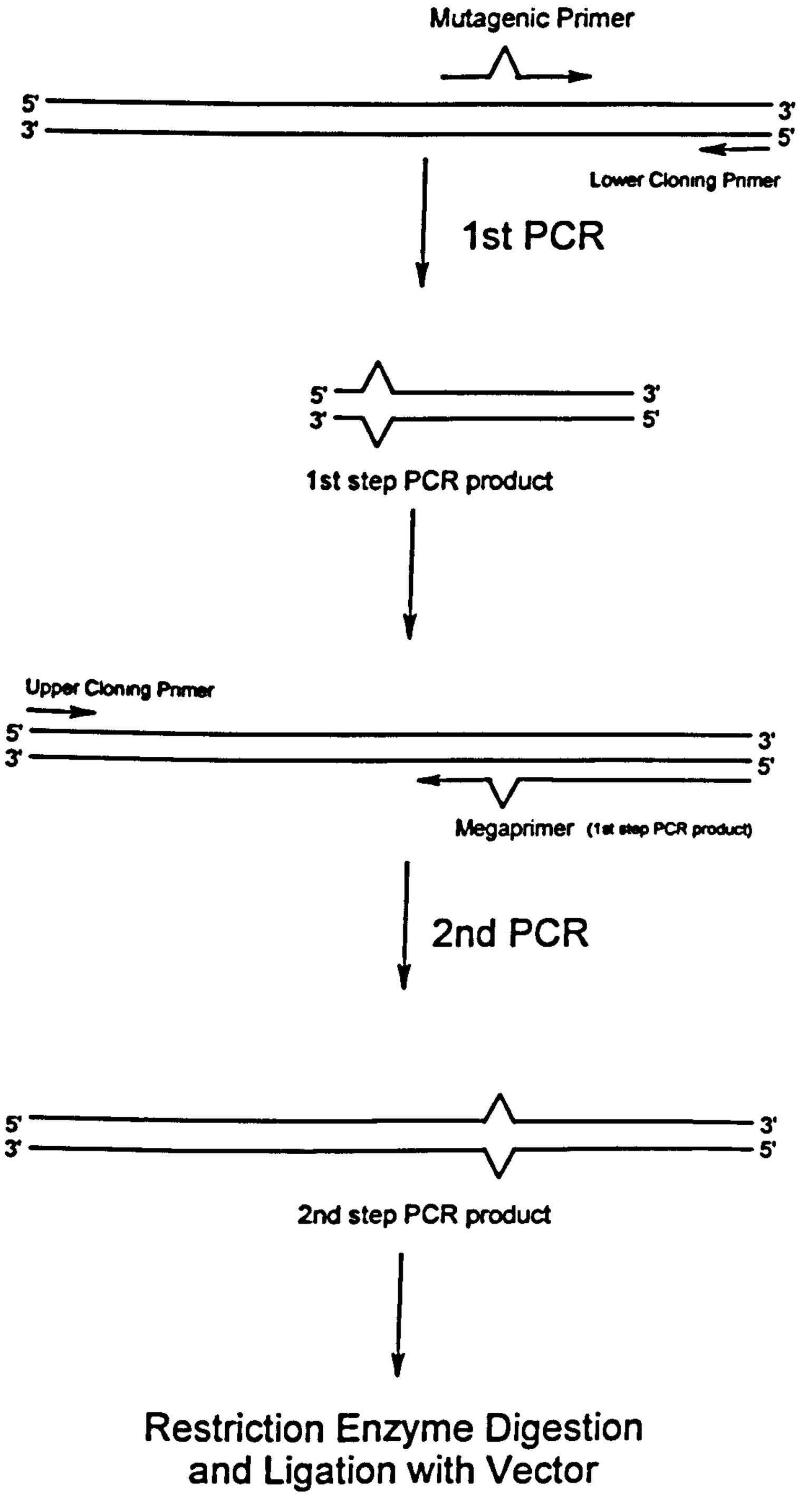


Figure 3.5 **Schematic representation of the Megaprimer method used in the production of site-directed mutants of KasA adapted from Sarkar & Sommer (1990)**

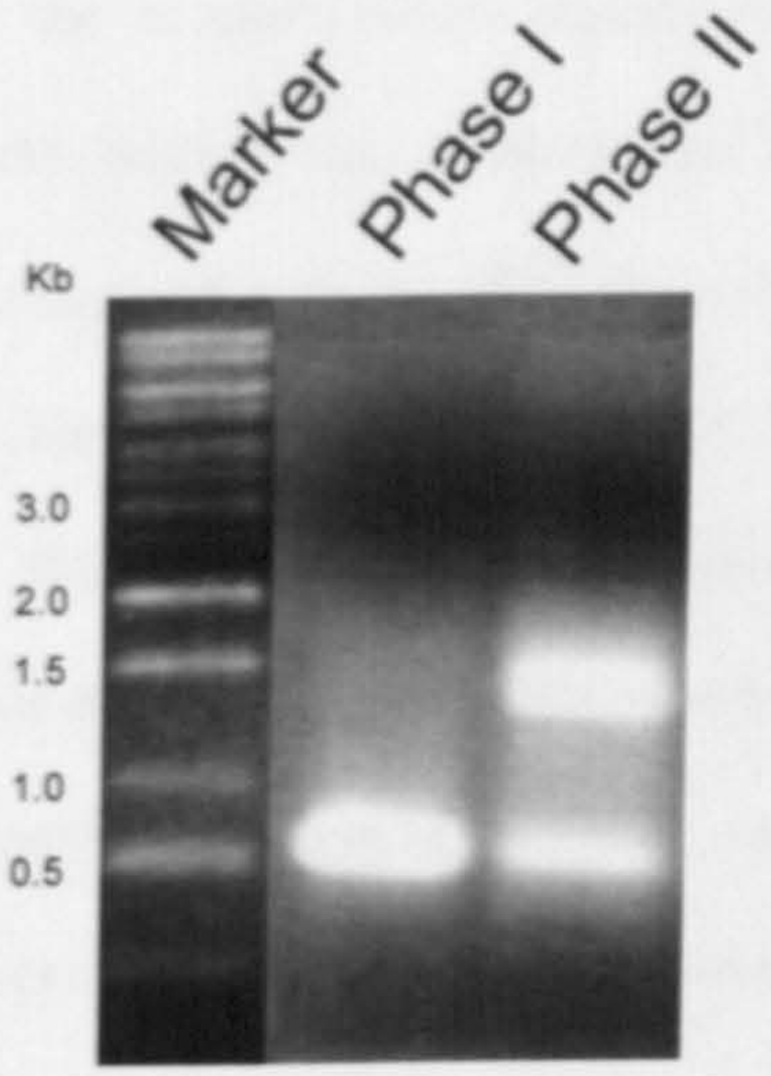


Figure 3.6 Agarose gel of phase I and phase II site-directed mutagenesis of *kasA*

After all the problems of PCR and ligation, not all the proposed mutations were able to be constructed. The following mutations were created; Gly115→Ala, Arg121→Lys, Ala141→Trp, Val142→Ala, Pro206→Ala, Phe210→Ala, Thr315→ala, Gly269→Ser, Gly312→Ser, Gly387→Asp, Phe402→Val, Phe404→Val and Phe413→Lys.

3.3.5 Protein purification and further analysis of KasA and muteins

Along side the production of the mutant KasA plasmids it was essential that protein was obtainable for analysis and in large enough quantities for analysis by circular dichroism and, more importantly, by crystallography. The purification of the new mutant, non-denatured proteins by Ni^{2+} Hi-trap Chelating Sepharose™ fast flow matrix column yielded less protein than previously observed with the catalytic triad muteins. So the attention of the study reverted to the wild-type KasA in order to try to develop a method to increase the yield and

solubility of the wild-type as the mutants would theoretically react the same. Therefore analysis of the solubility of the protein was implemented as *per* the solubility analysis protocol, *E. coli* C41 (DE3) transformed with the pET28b-*kasA* was culture and induced with 0.1-1.0 mM IPTG for 4 hours. The cells were harvested and lysed by sonication. The SDS-PAGE analysis demonstrated that the majority of the recombinant KasA protein was insoluble, shown by the presence of a dominant 45 kDa protein in the crude and urea-treated samples. Production of KasA muteins at lower temperature, i.e. 16°C was attempted. Lower temperatures would slow the overall protein production and may favour correct folding reducing aggregation of intermediates therefore theorectically increasing its solubility. Unfortunately this still produced poor protein recovery. *E. coli* C41 (DE3) transformed with pET28-*kasA* was used to prepare KasA His-tagged recombinant proteins. The expression was performed in triplicate and each lysed by French pressure cell and clarified by centrifugation, the first clarified supernatant was run as normal through the column, the second was run in the presence of 500 mM arginine and finally the last was run in the presence of 2 M urea. The SDS-PAGE analysis demonstrated that the majority of the recombinant KasA protein was insoluble, shown by the presence of a dominant 45 kDa protein in the urea purified KasA (Figure 3.7). Protein concentrations were estimated by BCA assay and their enzyme activity analysed by assay as *per* section 3.2.6. As previously observed the recombinant protein could be purified in small quantities and remained active. However, to develop a structural biology investigation larger yields were required. Our investigation thus far produced low yields of soluble protein but a band corresponding to a 45 kDa protein, consistent with the predicted molecular weight of the recombinant KasA was evident in insoluble aggregates solubilised by urea treatment. On attempted refolding, this protein did not exhibit KasA activity (Figure 3.7). The development of other purification methods including Q-sepharose (anion exchange) and phenyl-sepharose (Hydrophobic interaction) purification did not yield significant quantities of KasA.

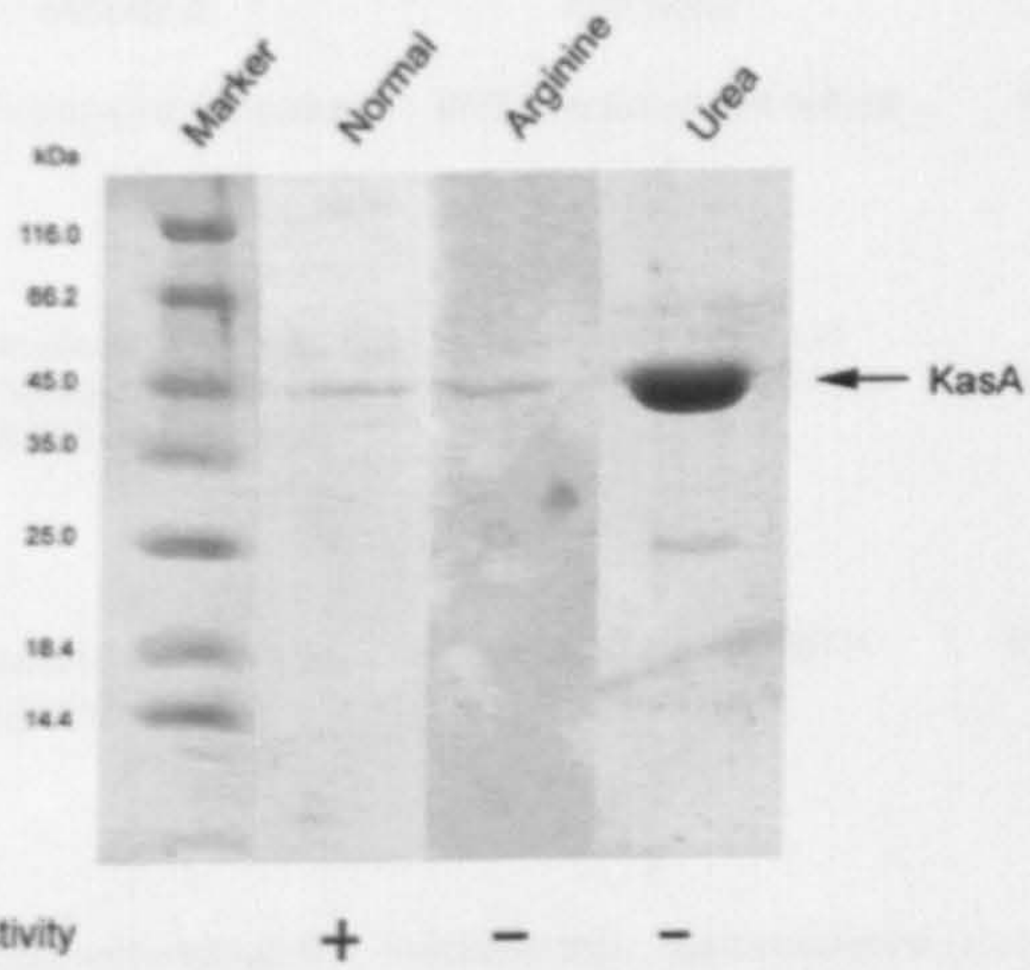


Figure 3.7 **SDS-PAGE (12 %) of purified KasA, arginine and urea. KasA Preparation His-tagged recombinant proteins expressed in *E. coli* C41 (DE3) were expressed as described in the materials and methods section. Cellular lysates were centrifugation at 27,000 x *g* followed by Ni^{2+} column purification (Normal), this was repeated but with the presence of 500mM arginine (Arginine), a further preperation resuspension in 0.02 M Sodium Phosphate and 0.5 M NaCl pH 7.4 containing 2 M urea followed by centrifugation at 27,000 x *g* (Urea) was analysed. KasA activity shown below.**

As the high levels of protein were purified using the urea-based method, it was decided to investigate the refolding of this denatured protein. Schaeffer *et al.,* (2001) purified *M. tuberculosis* KasA *via* Ni^{2+} chelating column system followed by refolding using an 8 M urea dialysis refolding method, whereas Saini *et al.*, (2002) refolded a *M. tuberculosis* histidine protein kinase (DevS) by a matrix-assisted method. Both methods were adapted to try to refold our KasA His-tagged recombinant protein after expression in *E. coli* C41 (DE3) (Figure 3.8). These methods were developed to include arginine which was added to mediate native structure formation and to reducing aggregation.

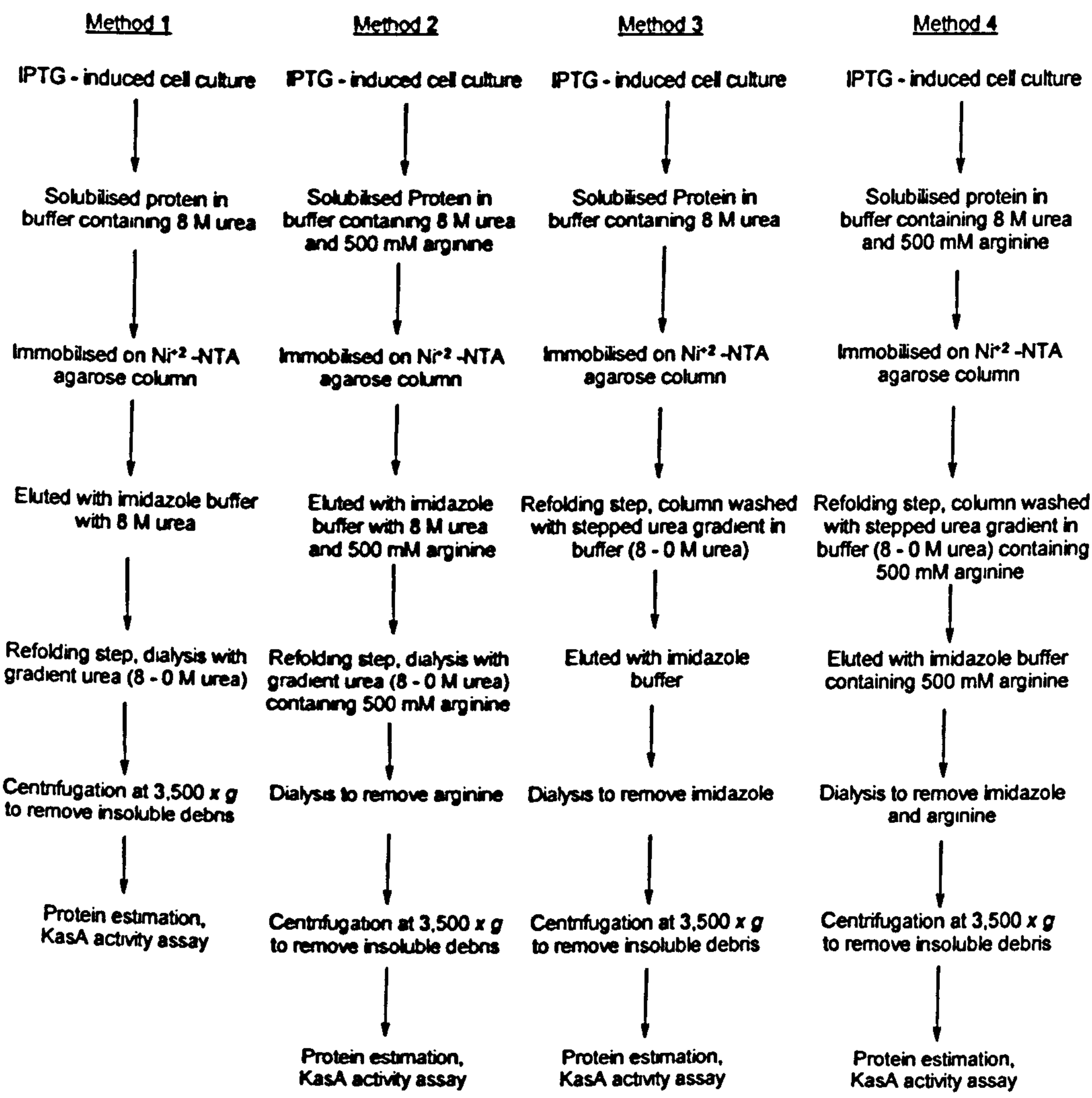


Figure 3.8 **Methods for the refolding of KasA. Two methods of purification and refolding were investigated, one in solution phase and the other bound to a Ni^{+2}-chelating agarose bound matrix. Methods 2 and 4 are further modified by the addition of 500 mM arginine**

All four preparation protocols yielded protein in large quantities similar to that observed in Figure 3.7. When these purified proteins were assayed using the *in vitro* KasA assay, no

catalytic activity was exhibited. The development of the pET28 expression system had now become exhausted and it was decided that other expression systems may prove more fruitful.

Four expression vectors pET23b, pVV16, pSD26 and pQE60 were considered. pET23b was used to utilise the C-terminal His-tag as opposed to the pET28a constructs which contain a N-terminal His-tag. The change in the position of the His-tag may aid purification *via* the Ni^{+2} Hi-trap system and therefore yield greater quatities of enzymatically active protein. The pQE60 expression system utilises an *E. coli* M15 (pREP4) expression host. It has previously been shown that *E. coli* FabB and FabF are both soluble in this system in contrast to the pET expression systems (Edwards *et al.*, 1997). Due to the similarity between KasA and FabB it is feasible that KasA will behave in a similar manner to FabB. The pVV16 expression system utilises *M. smegmatis* as the expression host, but also produces C-terminal His-tagged proteins. This system is not inducible and as the vector is constitutively expressed *via* the hsp60 promoter. The downfall of this expression system is that it is based on a low copy number plasmid therefore only small yields of protein might be obtained. The pSD26 expression system, however, also contains a C-terminal His-tag but is based on an acetamide inducible promoter for use in *M. smegmatis*. This vector has all the same properties as the pVV16 system but affords a higher yield of protein.

3.3.5.1 pET23b-*kasA* analysis

pET23b production of the C-terminal His-tag KasA was first assessed for solubility by first culturing the *E. coli* C41 (DE3) containing the plasmid to an OD $A_{600\ nm}$ = 0.6 followed with induction with IPTG at varying conditions as *per* section 3.2.9. The cells were harvested

followed by sonication and clarification by centrifugation at 27,000 *x g*. The pellet was resuspended in buffer containing 8 M urea. The samples were then visualised by 12 % SDS-PAGE (Figure 3.9).

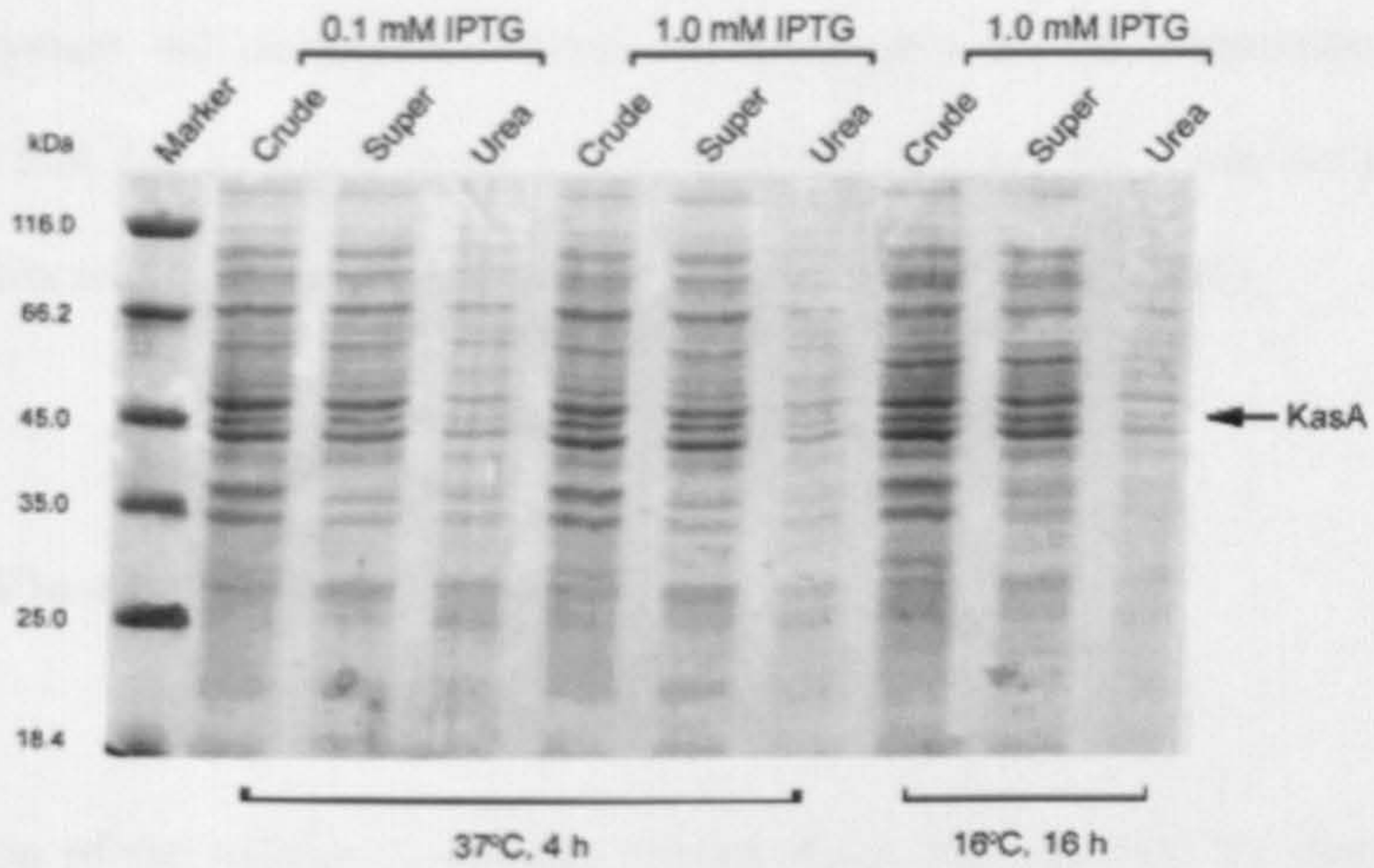


Figure 3.9 **Analysis of recombinant KasA production and solubility. *E. coli* C41 (DE3) pET23b-*kasA* was cultured to O.D.$_{600nm}$ = 0.7 at 37°C when recombinant protein production was induced by addition of IPTG. Different concentrations and induction temperatures were assessed for the production of soluble protein. Induction at 37°C was carried out for a 4 hour period using IPTG at either 0.1 mM or 1 mM. Some cultures were cooled to 16°C before protein production was induced by addition of IPTG to 1mM for a period of 16 h. Crude lysates (crude) were centrifuged at 27,000 x *g* for 30 min at 4°C and the pelleted insoluble material was dissolved in buffered 8M Urea (urea) before comparison with the supernatant and crude samples on 12 % SDS-PAGE.**

It was proposed that changing the position of the His-tag may affect the purification properties of the wild-type KasA *via* the Ni^{+2} Hi-trap system and therefore produce greater quantities of

enzymatically active protein. The supernatant fractions were taken and purified as *per* the materials and methods (section 3.2.10). As for the pET28 system the protein failed to bind to the column and was observed in the flow through or eluting early in the imidazole concentration gradient (10-20 mM) washes prior to the elution washes, therefore no protein was effectively purified *via* this system. After the two failed attempts using the *E. coli* expression system we decided to utilise *M. smegmatis* as the expression host. It was hypothesised that due to KasA being a mycobacterial enzyme there was the possibility that over-production within *M. smegmatis* would produce more soluble protein.

3.3.5.2 pVV16-*kasA* analysis

The production of the soluble C-terminal His-tag KasA was assessed by first culturing *M. smegmatis* containing the pVV16-*kasA* plasmid to an OD $A_{600\ nm}$ = 1.2. The cells were harvested followed by sonication and clarification by centrifugation at 27,000 *x g*. The pellet was resuspended in buffer containing 8 M urea and the samples visualised by Coomaisse blue-staining after electrophoresis on 12 % SDS-PAGE gels (Figure 3.10). This system is not inducible and as the vector is constitutively expressed *via* the hsp60 promoter different expression conditions were unavailable. The downfall of this expression system is that it is based on a low copy number plasmid therefore only small yields of protein were observed. As observed on SDS-PAGE there was relatively high expression of a protein with a molecular weight of 45 kDa consistent with pure recombinant KasA as predicted by the protein sequence. Purification was required in this system to assess levels of expression due to the host being *M. smegmatis* a similar 45 kDa would natively be expressed possibly masking any results. Therefore the clarified lysate was applied to a Ni^{+2} charged Hi-trap column. The protein bound much better to the column and was purified to high degree after elution;

visualisation by Coomaisse blue-staining after electrophoresis on 12 % SDS-PAGE gels showed that this system yielded large quantities of recombinant KasA (Figure 3.11).

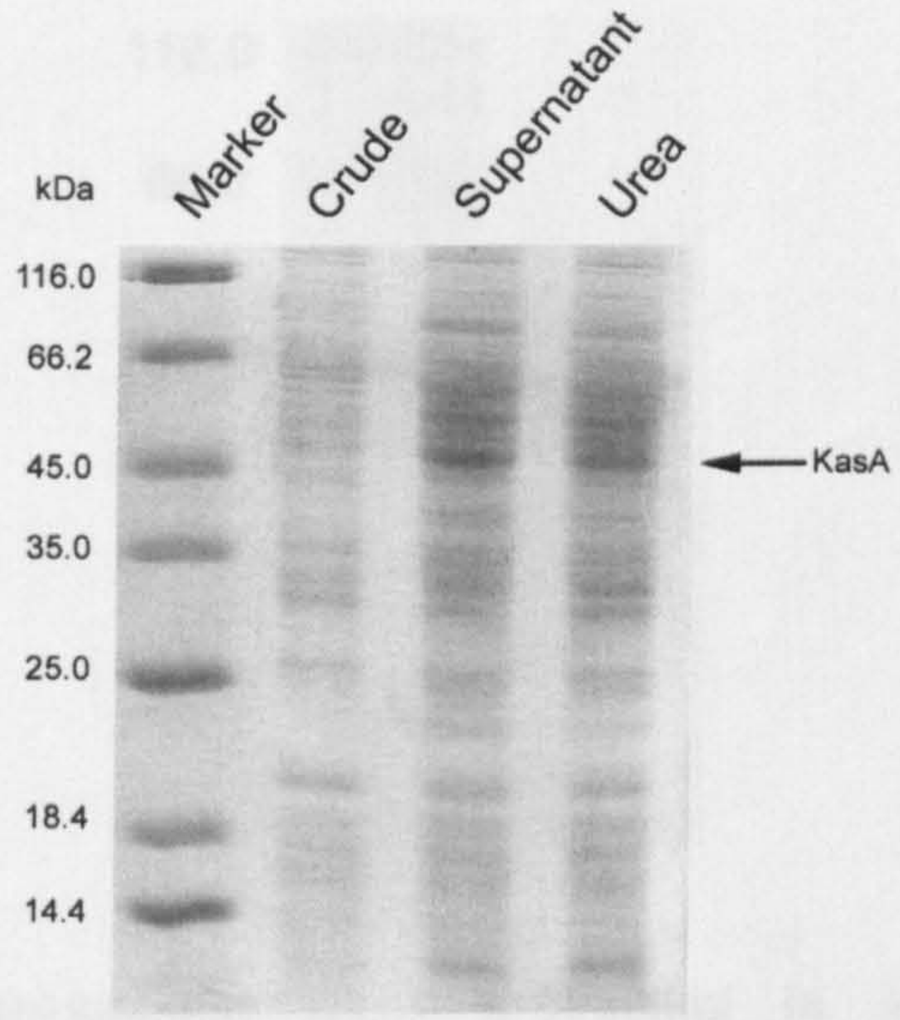


Figure 3.10 **Analysis of recombinant KasA production and solubility. *M. smegmatis* mc²155 pVV16-*kasA* was cultured to O.D.$_{600nm}$ = 0.7 at 37°C of the C-terminal His-tagged recombinant proteins prepared as described in the materials and methods section. Crude lysates (crude) were centrifuged at 27,000 x *g* for 30 min at 4°C and the pelleted insoluble material was dissolved in buffered 8M Urea (urea) before comparison with the supernatant and crude samples on 12 % SDS-PAGE.**

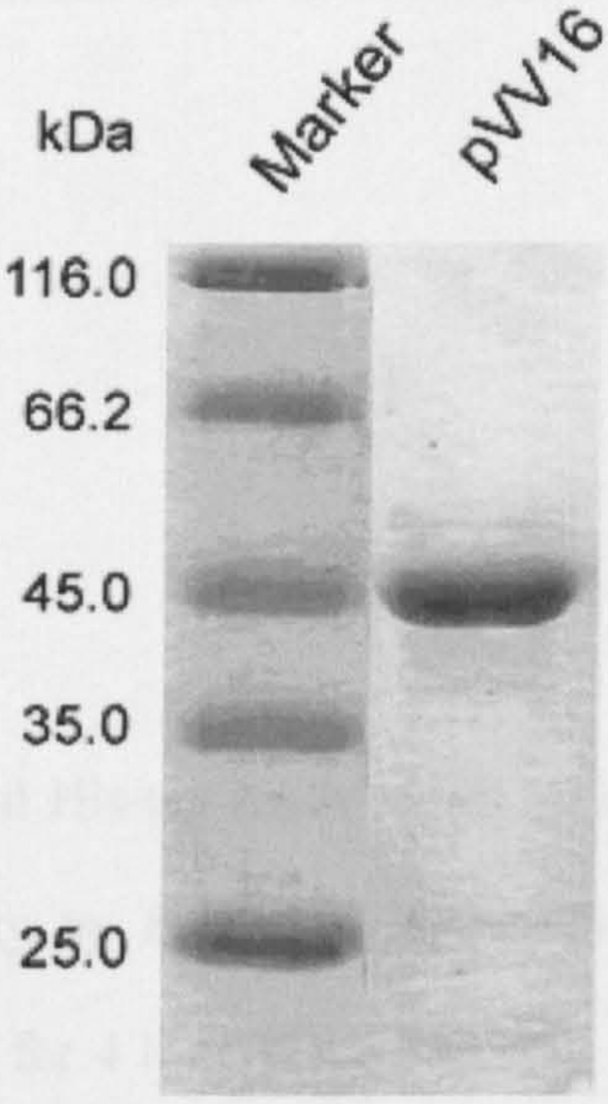


Figure 3.11 **Purified recombinant KasA production in *M. smegmatis* mc²155. *M. smegmatis* mc²155 pVV16-*kasA* were grown to O.D.$_{600nm}$ = 0.7 at 37°C and pelleted by centrifugation. Followed by lysis by sonication and clarification by centrifugation at 27,000 x *g*. Clarified lysates were and subjected to Ni^{2+}-chelating agarose chromatography and elution in imidazole (10-500 mM imidazole).**

The protein concentration of the pVV16-*kasA* produced recombinant KasA was estimated by BCA assay kit (Pierce). *In vitro* radio-chemical assaying of this recombinant protein in the KasA assay unfortunately did not produce any activity, therefore suggesting that this enzyme preparation method is uneffective for recombinant KasA. The system as stated previously is constitutively expressed *via* the hsp60 promoter therefore it was proposed that an inducible expression vector may change expression levels and possibly the overall folding and therefore the activity of the protein. The pSD26 expression vector is inducible *via* additions of actemamide. Over-production within *M. smegmatis* might produce more soluble protein, as observed in the pVV16 system. But allowing the production of protein by induction and

therefore allowing it to be produced in an active form due to a more susceptible cellular environment.

3.3.5.3 pSD26-*kasA* analysis

The production of the C-terminal His-tag KasA using *M. smegmatis* containing the pSD26-*kasA* was assessed for solubility by first culturing to an OD $A_{600\ nm}$ = 0.6 followed by induction with 0.2 % acetamide for 4 h at 37°C. The cells were harvested by centrifugation followed by sonication and clarification by centrifugation at 27,000 *x g*. The pellet was resuspended in buffer containing 8 M urea and the samples visualised by Coomaisse blue-staining after electrophoresis on 12 % SDS-PAGE gels (Figure 3.12). As observed on the SDS-PAGE there was relatively high expression of a protein with a molecular weight of 45 kDa in comparison with uninduced celluar lysate consistent with the production of recombinant KasA. As with the pVV16 system purification was required due to the host *M. smegmatis* having a similar 45 kDa natively expressed protein which may possibly mask results. Therefore the clarified lysate (Super) was applied to a Ni^{+2} charged Hi-trap column. The protein bound much like the two pET expression systems and was not purified; visualisation by Coomaisse blue-staining after electrophoresis on 12 % SDS-PAGE gels showed that this system yielded large quantities of contaminating protein along with the recombinant KasA (Figure 3.13). The protein purification also showed little affinity as observed with the pET systems with a protein with a molecular weight of 45 kDa consistent with pure recombinant KasA as predicted by the protein sequence being eluted in 20 mM imidazole therefore not allowing for effective purification.

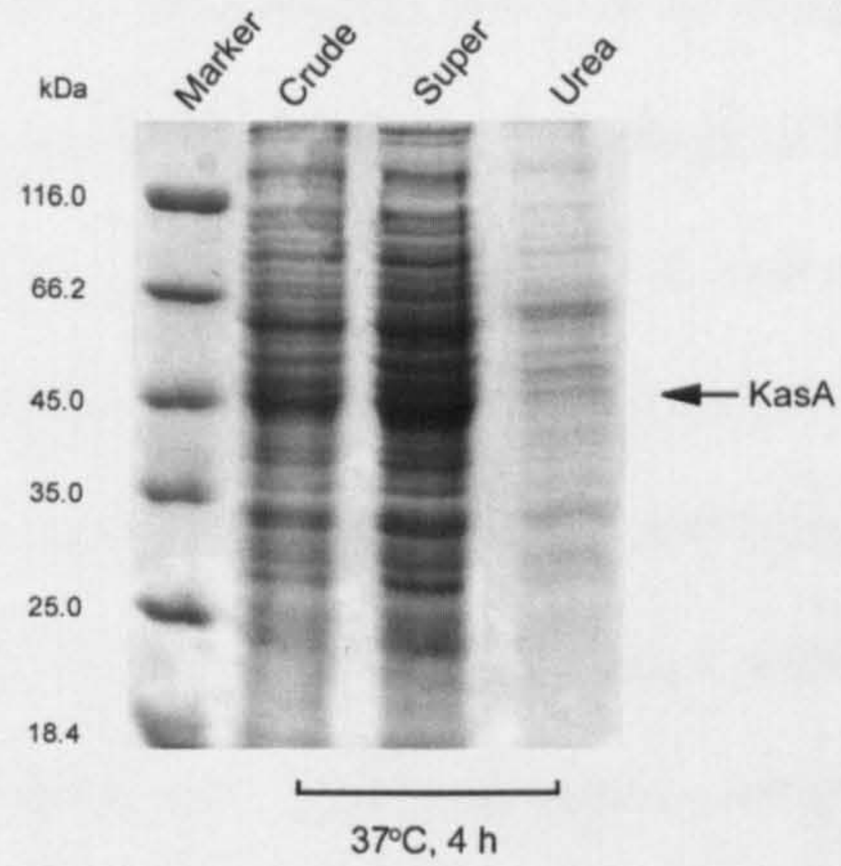


Figure 3.12 **Solubility analysis of pSD26-*kasA* expressed in *M. smegmatis* mc^{2}155. C-terminal His-tagged recombinant proteins expressed were expressed as described in the materials and methods section. Cellular lysates were sampled (Crude) then subjected to centrifugation at 27,000 x *g* (Super), followed by pellet resuspension in 0.02 M MOPS , pH 7.9, 50 mM MgCl, 5 mM β-mercaptoethanol and 10 mM imidazole containing 8 M urea followed by centrifugation at 27,000 x *g* (Urea)**

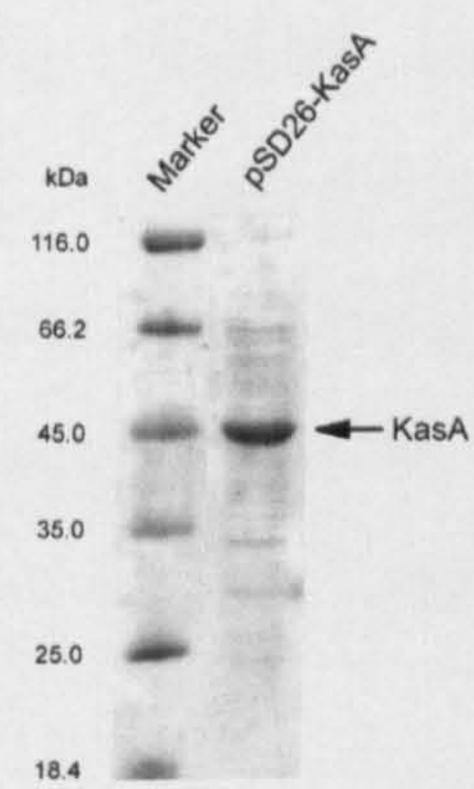


Figure 3.13 **Purified recombinant KasA production in *M. smegmatis* mc^{2}155. *M. smegmatis* mc^{2}155 pSD26-*kasA* were grown to O.D.$_{600nm}$ = 0.7 at 37°C followed by induction with 0.2 % acetamide and pelleted by centrifugation. Followed by lysis by sonication and clarification by centrifugation at 27,000 x *g*. Clarified lysates were and subjected to Ni^{2+}-chelating agarose chromatography and elution in imidazole (10-500 mM imidazole).**

Unfortunately the preparation of recombinant KasA in *M. smegmatis* failed again to produce an active protein. Further investigation into the homologs of KasA showed that the *E. coli* FabB and FabF were shown to be soluble from another *E. coli* expression system. The pQE60 expression system utilises an *E. coli* M15 (pREP4) expression host and had been shown to be an effective preparation system for *E. coli* FabB and FabF (Edwards *et al.*, 1997). Due to the similarity between KasA and FabB it is feasible that KasA will behave in a similar manner to FabB, therefore as before in the *E. coli* expression systems solubility and purification potential were assessed.

3.3.5.4 pQE60-*kasA* analysis

pQE60 production of the C-terminal His-tag KasA was first assessed for solubility by culturing the *E. coli* M15 (pREP4) containing the plasmid to an OD $A_{600\ nm}$ = 0.6 followed with induction with IPTG at varying conditions as *per* section 3.2.9. The cells were harvested followed by sonication and clarification by centrifugation at 27,000 *x g*. The pellet was resuspended in buffer containing 8 M urea. The samples were subjected to electrophoresis on 12 % SDS-PAGE and visualised by staining with Coomaisse blue-staining (Figure 3.14).

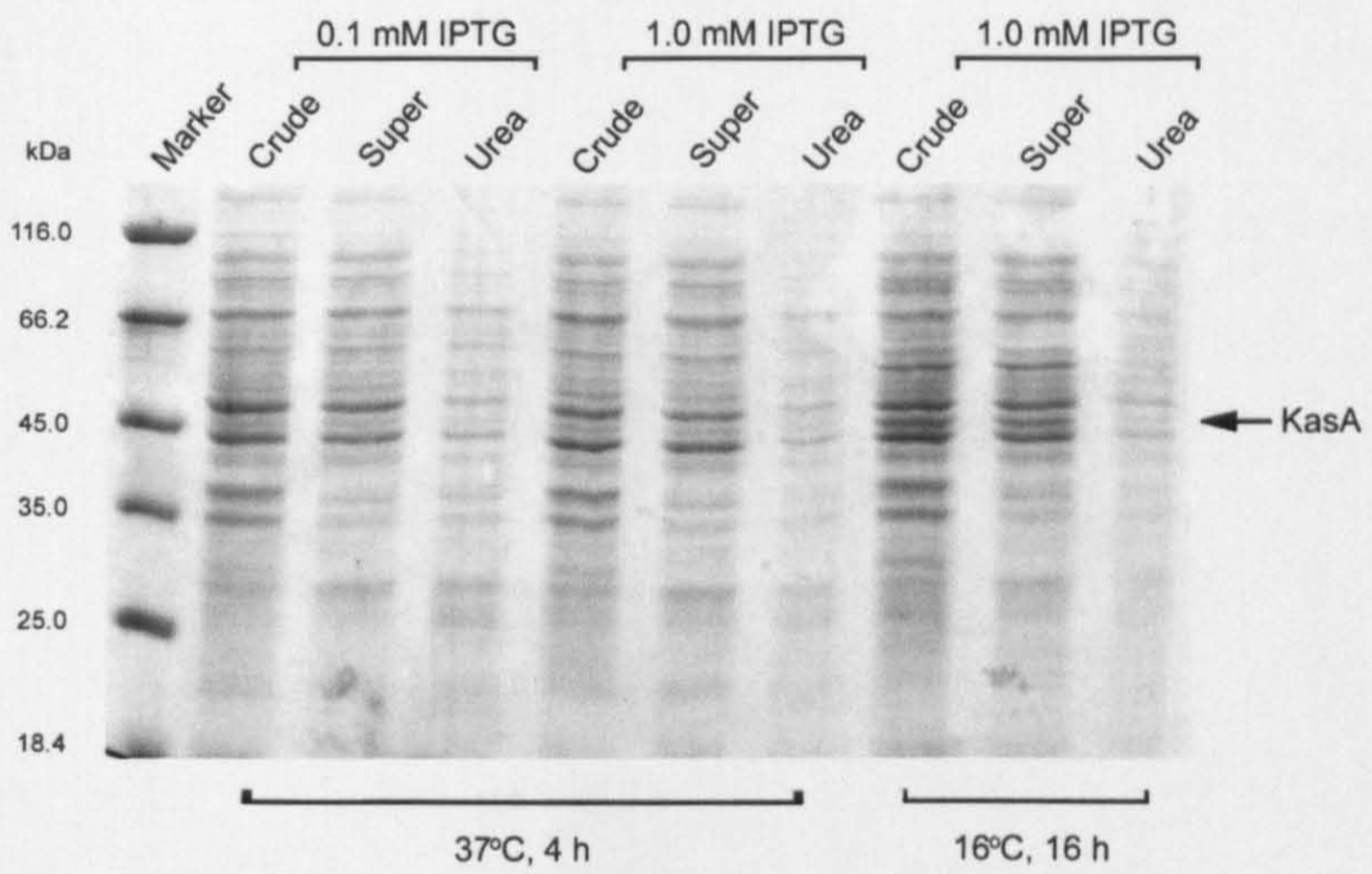


Figure 3.14 **Analysis of recombinant KasA production and solubility. *E. coli* M15 (pREP4) pQE60-*kasA* was cultured to O.D.$_{600nm}$ = 0.7 at 37°C when recombinant protein production was induced by addition of IPTG. Different concentrations and induction temperatures were assessed for the production of soluble protein. Induction at 37°C was carried out for a 4 hour period using IPTG at either 0.1 mM or 1 mM. Some cultures were cooled to 16°C before protein production was induced by addition of IPTG to 1mM for a period of 16 h. Crude lysates (crude) were centrifuged at 27,000 x *g* for 30 min at 4°C and the pelleted insoluble material was dissolved in buffered 8M Urea (urea) before comparison with the supernatant and crude samples on 12 % SDS-PAGE.**

In contrast to the expression profile we observed with the pET vectors (Figure 3.9) no significant IPTG-dependent induction of a protein with a molecular weight consistent with recombinant KasA at 37ºC was observed. However, significant quantities of such a protein were observed in the cultures induced with 1 mM IPTG at 16ºC for 16 h. Purification of this protein was implemented using a Ni^{+2} charged Hi-trap column as described as *per* materials and methods (section 3.2.10). The purified protein was dialysed against 20 mM phosphate

buffer, pH 7.9, 0.5 M NaCl. Final protein preparations were visualised on 12 % SDS-PAGE (Figure 3.15).

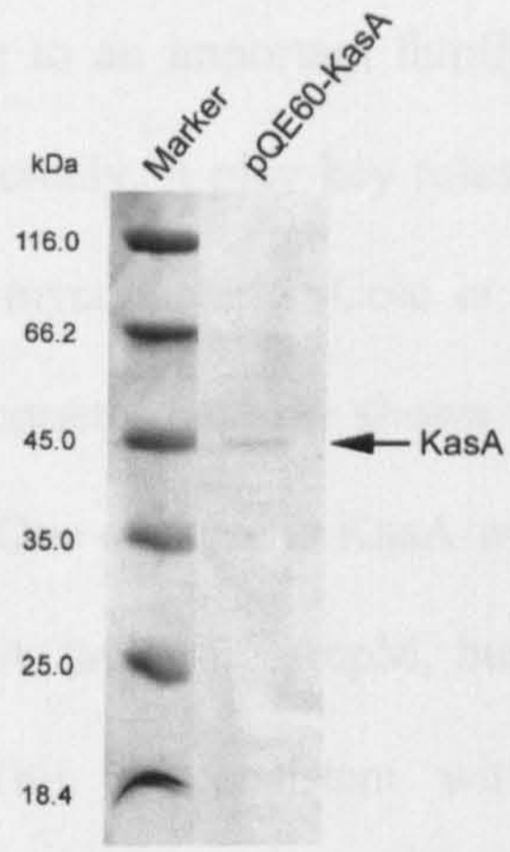


Figure 3.15 Purified recombinant KasA production in *E. coli* M15 (pREP4). *E. coli* M15 (pREP4) pQE60-*kasA* were grown to O.D.$_{600nm}$ = 0.7 at 37°C and induced with 1 mM IPTG followed by pelleted by centrifugation. Followed by lysis by sonication and clarification by centrifugation at 27,000 x *g*. Clarified lysates were and subjected to Ni^{2+}-chelating agarose chromatography and elution in imidazole (10-500 mM imidazole).

The pure protein was of insufficient quantities for assaying therefore further preparation would be required to produce more recombinant *kasA* from the pQE60 expression system.

3.4 Discussion

KAS of the FAS-II system belong to an important family of condensing enzymes that are related both structurally and functionally. It play key roles in the biosynthesis of fatty acids, polyketides and mycolic acids in mycobacteria (Cole *et al.*, 1998; Kremer *et al.*, 2000b). Analysis of the *M. tuberculosis* genome database shows that there are only three potential KAS enzymes (Cole *et al.*, 1998). One of these is KasA which has a substrate specificity for acyl-AcpM or acyl-CoA with KasA favouring AcpM, but still active with acyl-CoA (4 % activity of that with AcpM). This is consistent with the current understanding of mycobacterial FAS-II and that KasA is involved in the extension of acyl-AcpM as a part of FAS-II involved in meromycolate extension, which utilises AcpM-bound precursors (Kremer *et al.*, 2002c).

Over-expression of *kasA* from *M. tuberculosis* was associated with a decrease in α′-mycolate production and an increase in α-mycolates, both in *M. smegmatis* and in *M. chelonae*, suggesting that KasA is able to elongate the short α′-mero-chain into full-length meromycolates (Kremer *et al.*, 2002c). Overproduction of *M. smegmatis* KasA in *M. smegmatis* resulted in similar changes in the α/α′ ratio. Therefore the absence of *kasA* cannot be inferred by the presence of α′-mycolates in some mycobacterial species, or by an alteration of KasA activity *via* point mutations, since it was very similar to KasA from *M. tuberculosis*. The α′ composition may be simply due to different regulation levels of *kasA* expression and/or of other genes involved in the mycolic acid biosynthetic pathway in *M. smegmatis* (Kremer *et al.*, 2002c). Analysis of the FAS-II activity of the cytosolic fraction of *M. smegmatis* using a set of acyl primers ranging from C_2- to C_{20}-CoA, observed that FAS-II

activity was optimal with C_{14}/C_{16}-acyl-CoA/AcpM precursors. By comparing the FAS-II activity with a recombinant strain overproducing KasA, it was shown that although the same substrate preference was followed, the specific activity was enhanced significantly. This supports the notion that KasA is part of the mycobacterial FAS-II system. Recombinant KasA purified under non-denaturating conditions was assayed and found to catalyse the transfer of a two-carbon unit from [2-^{14}C]malonyl-ACP to C_{16}-AcpM, thus confirming KasA as a condensing enzyme involved in FAS-II and meromycolate extension. To improve our understanding of the catalytic site of KasA, Cys171, His311, Lys340 and His345 were mutated to alanine. All mutations impeded the enzymic activity, suggesting that these residues play a critical role in condensation activity mediated by KasA.

The described expression system and culture conditions that allow the generation of soluble KasA mutants were published by Kremer *et al.*, (2002c). The Cys171, His311, Lys340, His345 and wild-type KasA produced varying amounts of protein and yields of soluble were low.

The production of the subsequent mutants showed little soluble protein for some and no soluble protein for others, therefore halting protein production until a more effective method was developed. Refolding of the insoluble proteins was attempted using several methods but to little success. Despite using techniques identified by Schaeffer *et al.* (2001) to produce soluble active protein, in our hands the wild-type protein remained insoluble and inactive after several attempts at refolding.

Mutations in KasA (mutated at Asp66→Asn, Gly269→Ser, Gly312→Ser and Phe413→Leu), which also correlate with low levels of INH resistance, were reported to inhibit the formation of a trimolecular complex consisting of KasA, INH and AcpM (Mdluli *et al.*, 1998). Interestingly, these mutations do not reside within the catalytic site, suggesting that these mutations may rather influence the degree of binding of the acyl-AcpM or the stability of the KasA dimer. Clearly, the generation of substrates, such as C_{16}-AcpM and purified KasA, now allows the possibility to study key interactions between KasA and long-chain acyl-ACPs, in addition to how these mutations affect the KasA–INH–AcpM complex (Kremer *et al.*, 2000b), but the ongoing problem of efficient and consistent production of mutated proteins remains an issue. Other residues of interest like Gly107, Val134, Val135 and Phe201 which line the cavity of the active site and may play a role in the substrate specificity of the acyl-chain, therefore it was assumed that these residues play a major role in transacylation of the acyl-AcpM to the active site Cys and pose viable candidates for investigation (Olsen *et al.*, 1999). The active site in the FabB dimer is accessible at the bottom of a narrow tunnel defined by a number of residues, but Pro202 was shown to be present at the entrance. The possibility that this mutation will have an effect on the kinetics of catalysis is interesting and requires further investigation. Asp231 and Thr315 were identified in FabF to be part of the binding site of the acyl-ACP substrate (Huang *et al.*, 1998). Thr307 together with the N-terminus of the preceding α-helix form a binding site at the top of the binding cavity for the hydroxyl and phosphate of the phosphopantetheine group of the acyl-ACP. Some distance from the active site pocket is the Asp223 residue which is located on the surface of the enzyme and could be involved in the interaction with ACP (Huang *et al.*, 1998). This study set out the groundwork for the further investigation of significant amino acid residues within the KasA enzyme, upon elucidation of a viable purification protocol.

The type II KAS enzymes are commonly separated into two classes based on their sensitivity to cerulenin. In *E. coli*, cerulenin binds irreversibly to KAS proteins (Campbell & Cronan, 2001). The mechanism of inhibition involves covalent cross-linking of the KAS active-site cysteine to cerulenin (Chapter 1.5.11). The crystal structures of the FabB and FabF active sites (Huang *et al.*, 1998; Moche *et al.*, 2001; Olsen *et al.*, 1999; Price *et al.*, 2001) show that a hydrophobic binding pocket is present and occupied by the inhibitor. Structural analyses of KAS→cerulenin complexes suggest that cerulenin mimics the transition state of the condensation reaction effectively (Price *et al.*, 2001). Cerulenin-sensitive KAS enzymes have a catalytic triad, which in FabB consists of Cys163, His298 and His333 (Price *et al.*, 2001). Protein alignments among FabB, FabF and KasA reveal a full conservation of these three residues (Kremer *et al.*, 2000b), which may explain the sensitivity of KasA to cerulenin. In contrast, mtFabH, another KAS, has been shown to be resistant to cerulenin (Choi *et al.*, 2000b). It was shown previously that cerulenin exhibits potent *in vivo* anti-mycobacterial activity (Kremer *et al.*, 2000b). Although over-expression of KasA in *M. bovis* BCG did not generate an increased resistance against cerulenin, it was suggested that the *in vivo* anti-mycobacterial activity of cerulenin was primarily *via* inhibition of earlier events catalysed by *de novo* FAS-I, thus masking inhibition of FAS-II. The *in vitro* inhibition of the *E. coli*, FabB and FabF, has also been well documented for cerulenin (Campbell & Cronan, 2001). Therefore it was reasonable to assume that cerulenin may also inhibit the β-ketoacyl-AcpM synthase activity of KasA. Increasing concentrations of cerulenin were added to the *in vitro* condensation assay. A marked dose-response inhibition of KasA activity was observed with cerulenin providing an IC_{50} value of 0.15 μg/ml (0.67 μM). The IC_{50} values of purified *E. coli* KAS proteins were reported as 3 and 20 μM for FabB and FabF respectively (Price *et al.*, 2001), suggesting that the mycobacterial condensing enzyme is more susceptible to cerulenin inhibition *in vitro* than the related *E. coli* enzymes.

Recent studies by Schaeffer *et al.* (2001) and Kremer *et al.* (2002b) contribute to the understanding of the condensation–reaction mechanism mediated by KasA leading to mycolic acids *via* FAS-II and now allows for the development of new anti-mycobacterial agents that target this key condensation step in *M. tuberculosis*.

Two new possible methods for further investigation of recombinant KasA production have recently been elucidated, firstly the use of quiescent-cells. The quiescent-cell expression system is a radical alternative to conventional fermentation for protein overproduction in *E. coli* (Rowe *et al.*, 1999). It is dependent on the controlled over-expression of a small RNA called Rcd in *hns* mutant strains to generate non-growing, quiescent cells which are not nutrient limited. Quiescent cells no longer produce biomass and have their metabolic resources channelled toward the expression of plasmid-based genes. The biosynthetic capacity of the system has been demonstrated in their ability to express chloramphenicol acetyltransferase to more than 40 % of total cell protein (Rowe & Summers, 1999). Therefore quiescent cells may provide an ideal environment for the expression of KasA. Toxicity problems are negated by this system and with extremely high levels of protein expressed it would hopefully produce some active recombinant KasA. The second method would be the use of chaperones. Takara Bio Inc. have produce a set of five chaperone plasmids developed by HSP Research Institute which have been designed to enable efficient expression of multiple molecular chaperones that are known to work in cooperation as aids of protein folding processes. The have shown that these chaperones have the ability to reduce inclusion body formation and allow for the effective recovery of soluble folded protein. Development of such a system in the production of recombinant KasA would be very beneficial as the majority of expressed protein in any of the systems tested in this study had high levels of insoluble protein.

The following outlines the future work on KasA when the problems of expression, solubility and purification have been solved:

- KasA assay analysis of mutated proteins
- Chain length specificity with a range of acyl-AcpM substrates (C_8, C_{12}, C_{16}, C_{20}, C_{24} and C_{26})
- Decarboxylation activity of wild-type and mutated proteins with malonyl-AcpM
- Transacylation of C_{16}-AcpM
- Antibiotic sensitivity of wild type KasA

CHAPTER 4

Acyl-CoA carboxylases

(AccD)

4.1 Introduction

The *Corynebacterianeae* represent a distinct and unusual group within Gram-positive bacteria, with the most prominent members being the human pathogens *M. tuberculosis* and *M. leprae* (Brennan & Nikaido, 1995; Stackebrandt *et al.*, 1997). In addition, non-pathogenic bacteria belong to this taxon, such as *Corynebacterium glutamicum*, which is used in the industrial production of amino acids (Eggeling *et al.*, 2001). A common feature to all these bacteria is that they possess unusual lipids, for instance mycolic acids (Brennan & Nikaido, 1995).

Mycolic acids are long chain α-alkyl-β-hydroxylated fatty acids (R-CH[OH]-CH[R']-COOH), where R represents the meromycolate chain consisting in *M. tuberculosis* of up to 56 carbons possessing additional structural modifications, and R' a shorter aliphatic branch possessing 22-26 carbons. These two chains are then condensed together *via* a specialised Claisen condensation enzyme, followed by reduction to yield mature mycolic acids (Kremer *et al.*, 2000a). In contrast, mycolic acids from *Corynebacterium* species, like *C. glutamicum*, represent the simplest form of these lipids, whereby two C_{16} fatty acids condense together followed by reduction to afford a C_{32} mycolic acid. These α-branched, β-hydroxy fatty acids are found, primarily, as esters of the non-reducing arabinan terminus of arabinogalactan (McNeil *et al.*, 1991). In addition, mycolic acids can also been found as extractable "free" lipids within the cell wall, mainly linked to glucose and trehalose (Minnikin, 1982). These mycolic acids and their derivatives are believed to play a crucial role in the architecture of the cell envelope (Brennan & Nikaido, 1995).

Interestingly, *M. tuberculosis* is characterised by an exceptionally high number of additional lipids and glycolipids, which are thought to aid in the persistence of the bacterium thereby

fuelling the promise of the identification of new drug targets in the context of tuberculosis (Kremer & Besra, 2002). The rich diversity of lipids present in *M. tuberculosis* is reflected at the genomic level by possessing a panoply of genes involved in lipid biosynthesis (Cole *et al.*, 1998), some of them constituting sets of paralogous genes. For instance, *M. tuberculosis* has 35 *fadD* genes, annotated as fatty-acid-CoA ligases, 16 *pks* genes annotated as PKS, and 6 *accD* genes annotated as acyl-CoA carboxylases, whilst their detailed function in lipid biosynthesis and relevance for persistence is only now emerging (Cole *et al.*, 1998; Glickman & Jacobs, 2001). Importantly, *Corynebacterium* species are considered the archetype within *Corynebacterianeae*, including *M. tuberculosis* due to a low frequency of gene duplications and structural alterations giving rise to a strong conservation at the genomic level within this sub-group of the *Corynebacterianeae* (Nakamura *et al.*, 2003). Together with their simpler chemical composition - as outlined above for mycolic acids - it can therefore be assumed that species of *Corynebacterium* have at their disposal just the core set of genes and reactions characteristic for the *Corynebacterianeae*. Indeed, *C. glutamicum* possesses just 3 *fadD*, 1 *pks* and 4 *accD* genes (Kalinowski *et al.*, 2003). As a result comparative studies using *C. glutamicum* have been employed in understanding the role of several *M. tuberculosis* proteins, for instance Ppm1/D2 in lipoarabinomannan (LAM) biosynthesis (Gibson *et al.*, 2003) and the "antigen 85" mycolyltransferases in cell wall mycolylation (Belisle *et al.*, 1997; Brand *et al.*, 2003; De Sousa-D'Auria *et al.*, 2003; Kacem *et al.*, 2004).

The group of 4 *accD* genes in *C. glutamicum* is interesting, and little is known of the function of these genes in *Corynebacterianeae*. These genes encode polypeptides with similarities to the β-subunits of acetyl-/propionyl-CoA carboxylases, and basically only one of them would be sufficient for the carboxylation of acetyl-CoA to provide malonyl-CoA for fatty acid

biosynthesis. This suggests additional and specific carboxylations, probably involved in the synthesis of unusual lipids.

In this context the still controversial issue of mycolic acid synthesis is very attractive to consider carboxylation reactions. Cell-free extracts of *Corynebacterium* have been shown to utilise [1-^{14}C]-labelled palmitic acid (Shimakata *et al.*, 1984; Shimakata *et al.*, 1985), with the newly synthesised mycolic acid exclusively labelled at C-1 and C-3 (Walker *et al.*, 1973). The Claisen condensation reaction was hypothesised to involve a carboxylation step since it was inhibited by avidin, an inhibitor of biotin dependent enzymes in extracts of *C. diphtheriae* which produced the putative precursor of mycolic acids, 2-tetradecyl-3-oxo-octadecanoic acid (Ahibo-Coffy *et al.*, 1978; Walker *et al.*, 1973). In contrast, avidin had no effect on the Claisen condensation reaction in a cell-free extract of *C. matruchotii* that synthesises mature mycolic acids (Shimakata *et al.*, 1984; Shimakata *et al.*, 1985). What seems to support this last observation came from the incorporation of [2,2-^{2}H] palmitic acid in whole cells of *C. matruchotii* (Lee *et al.*, 1997). During the studies carried out in this thesis a polyketide synthase from *C. glutamicum* (*Cg-pks*), the equivalent of *M. tuberculosis pks13*, was identified which apparently plays a key role in mycolic acid biosynthesis (Portevin *et al.*, 2004).

The studies in this thesis are centered on the 4 *accD* genes located downstream of *Cg-pks*. As a result this study presents a systematic study on lipid and mycolic acid biosynthesis based on defined *C. glutamicum* mutants, focusing on the relevance of the *accD* genes along with a phylogenomic analyses of these genes in *Corynebacterium* and *Mycobacterium* species whose genomes are established (Cerdeno-Tarraga *et al.*, 2003; Cole *et al.*, 1998; Cole *et al.*, 2001; Garnier *et al.*, 2003; Kalinowski *et al.*, 2003).

4.2 Materials and methods

4.2.1 Bacterial strains and growth conditions

Escherichia coli DH5αmcr and *C. glutamicum* ATCC 13032 (the wild type strain, and referred for the remainder of the text as *C. glutamicum*) were grown in Luria-Bertani (LB) broth (Difco) at 37°C and 30°C, respectively. The mutants generated in this study were grown on complex medium LBHIS (5 g tryptone, 5 g NaCl, 2.5 g yeast extract, 18.5 g brain heart infusion (Difco), and 90.1 g sorbitol *per* litre). Kanamycin and ampicillin were used at a concentration of 50 µg/ml. The minimal medium CGXII was used for *C. glutamicum* (Keilhauer *et al.*, 1993) with 30 mg/l protocatechuic acid included as a chelating agent. Samples for lipid analyses were prepared by harvesting cells at an OD 10-15, followed by a saline wash and freeze drying. Growth comparisons of the four *accD*-mutants were performed on un-supplemented CGXII and LBHIS plates, as well as LBHIS plates supplemented with 0.03 % Na-oleate (w/v), 0.03 % Tween 40 (w/v), 0.04 % Brij-35 (w/v) and 0.03 % butter hydrolyzate (w/v).

4.2.2 Construction of plasmids

To enable chromosomal inactivation of the four *acc* genes of *C. glutamicum* internal fragments were amplified by PCR, and blunt end ligated using the Sure Cloning Kit (Amersham) into the *Sma*I-site of the non-replicative vector pk18mob (Schafer *et al.*, 1994). The primers used were: accD1MuF10 (5'-GCA TGT GCA GGT GGC AAC GC-3'), accD1MuR11 (5'-GGT AAT CTT TGG AAC GGT TGC-3'), accD2MuF30 (5'-GTC ACG TGT ACT CCC CT-3'), accD2MuR31 (5'-CAA GCG AAT ACG AGG TC-3'),

accD3MuF40 (5'-GTT GTA GGC GTC GCA GAT AC-3'), accD3MuR41 (5'-GCG TCC TCT GAA GAA GAG-3'), paccD4intfor (5'-TGG GGT TCA TCT GGG CAT CTC AC-3'), paccD4intrev (5'-TGC CCC CAA CGT TTC CAT AAT CTC-3'). The inactivation vectors derived were termed pK18mobaccD1-int, pK18mobaccD2-int, pK18mobaccD3-int and pK18mobaccD4-int.

The in-frame deletion of *Cg-pks* in *C. glutamicum* was achieved with pK19mob*sacB*Δpks. Cross-over PCR was used to enable one-step integration of fragments containing upstream and downstream sequences of *Cg-pks* into pK19mob*sacB*. In the first PCR-round two separate amplification products were generated by using the primer pairs pCipks (5'-TGT TTA AGT TTA GTG GAT GGG AGT CGC CGC ATT GAT GAG ATT TC-3′), with pCopks (5′-GGA ATT CGA CAG CGG AAG CTG ACG ACG -3′) and pNopks (5′-GGA ATT CCG TTG GCA CTG CAC ACG GTG-3′), with pNipks-2 (5′-CCC ATC CAC TAA ACT TAA ACA CTT CTG ATC CGA CGA TTG GCT CTG-3′). Both PCR products subjected to electrophoresed on 1 % agarose and DNA visualised using EtBr. The appropriate DNA fragment was excised from the gel and extracted using the Qiagen Gel extraction kit. The PCR product purified using the Qiagen Gel extraction kit was used as a template to amplify with primers pCopks and pNopks a fragment devoid of *Cg-pks* that was treated with *Eco*RI and ligated with *Eco*RI-cleaved pK19mobsacB to afford pK19mobsacBΔpks. The inserts in constructs used in this study were verified by sequencing.

4.2.3 Genomic mutations

The non-replicative integration and deletion vectors were usually introduced *via* electroporation but conjugation was used when electroporation failed. Conjugation used *E.*

coli S-17-1 as the donor, and its sensitivity to nalidixic acid (50 µg/ml) was used after plating for counterselection. Selection of recombinant *C. glutamicum* strains for the integration of the four pK18mob-int vectors used 15 µg/ml kanamycin. The correct integration of sequences into the chromosome and absence of sequences, respectively, in the resulting recombinants was verified by PCR using two different primer pairs and by Southern analysis.

To achieve deletion of *Cg-pks*, plasmid pK19mobsacBΔpks was introduced into *C. glutamicum* by electroporation. Selection for resistance to kanamycin in at least 20 independent assays yielded regularly a number of clones indicating integration of the vector in the chromosome by homologous recombination. At least 100 clones derived from independent electroporation experiments were subjected to the subsequent selection for the second homologous recombination event. In this round the presence of *sacB* (together with addition of 10 % sucrose to the medium) resulted in a positive selection of clones where vector sequences were lost. Small colonies picked after 10 days were analysed for deletion of *Cg-pks*. Of 60 colonies analysed, 19 exhibited the loss of *Cg-pks*. One of these was further analysed by Southern-blot analysis and the resulting strain, *C. glutamicum*Δpks was used in all subsequent studies.

4.2.4 Southern blot analysis

Genomic DNA was extracted from *accD1-D4* mutants and the *Cg-pks* mutant and cleaved with *Bst*EII. The resulting fragments were separated on a 1 % agarose gel and blotted onto a Nytran NY13N nitrocellulose membrane, with subsequent washings according to standard protocols. Detection was carried out with fragments of the *acc*-genes as probes which were labelled with digoxygenin (Roche Dig Labeling and Detection Kit). For *accD1*, *accD2*, *accD3*

and *accD4* the *Sac*I-*Bgl*II, *Pvu*II-*Sgr*AI, *Cla*I-*Stu*I and *Sfu*I-*Pvu*I fragments were used, respectively. For the *C. glutamicum*Δpks analysis the 645 bp-fragment used was generated by PCR with primers pNopks (5'-GGA ATT CCG TTG GCA CTG ACA ACG GTG-3') and ppksACPintrev (5'-CAA CAT CGC GAG AGG AAA GG-3').

4.2.5 Extraction and analysis of [^{14}C]-labelled lipids

LBHIS (5 ml) was inoculated with a single colony of the *accD1, accD2, accD3, accD4* and *Cg-pks* mutant, respectively, and shaken at 120 rpm overnight at 30°C. An aliquot (1 ml) of this culture was used to inoculate 5 ml CGIII and grown again overnight. This second pre-culture was used to inoculate 5 ml of CGXII to give a starting OD_{600nm} of about 0.2. Cells were incubated at 30°C until an OD_{600nm} of 0.4 at which point 5 µCi/ml of [^{14}C]acetate (62 mCi/mmol, Amersham) was added followed by overnight incubation with shaking at room temperature. Cells were harvested by centrifugation followed by wash in a saline wash and freeze-dried.

Free lipids were extracted by two consecutive extractions with 2 ml of $CHCl_3/CH_3OH/H_2O$ (10:10:3, v/v/v) for 3 h at 50°C. These lipidic extracts were combined with 1.75 ml of $CHCl_3$ and 0.75 ml of water, mixed and centrifuged. The lower organic phase was recovered, washed twice with 2 ml of $CHCl_3/CH_3OH/H_2O$ (3:47:48, v/v/v) and the resulting organic phase dried and resuspended in 200 µl of $CHCl_3/CH_3OH/H_2O$ (10:10:3, v/v/v). An aliquot (20,000 cpm) from each strain was subjected to thin-layer chromatography (TLC) using silica gel plates (5735 silica gel $60F_{254}$, Merck), developed in $CHCl_3/CH_3OH/H_2O$ (60:16:2, v/v/v). Autoradiograms were produced by 2-3 day exposure to Kodak X-Omat AR film to reveal [^{14}C]-labelled lipids and compared to known standards (Puech *et al.*, 2000).

The bound lipids from the de-lipidated extracts were released by the addition of 2 ml of 5 % aqueous solution of tetra-butyl ammonium hydroxide (TBAH), followed by overnight incubation at 95°C. After cooling, water (2 ml), CH_2Cl_2 (4 ml) and CH_3I (500 µl) were added and mixed thoroughly for 30 min. The lower organic phase was recovered following centrifugation and washed three times with water (4 ml), dried and resuspended in diethyl ether (4 ml). After centrifugation the clear supernatant was again dried and resuspended in CH_2Cl_2 (200 µl). An aliquot (10,000 cpm) from each strain was subjected to TLC using silica gel plates (5735 silica gel $60F_{254}$, Merck), and developed in petroleum ether/acetone (95:5, v/v). Autoradiograms were produced by overnight exposure to Kodak X-Omat AR film to reveal [^{14}C]-labelled fatty acid methyl esters (FAMES) and mycolic acid methyl esters (MAMES) and compared to known standards (Lee *et al.*, 1997).

4.2.6 Acyl carboxylation assay

To assay acyl carboxylase activity we adapted the method of Rainwater *et al.* (1982). Briefly, a clarified lysate of the appropriate strain was prepared by resuspending the bacterial cultures to a density of 0.5 g wet weight/ml in 0.1 M potassium phosphate buffer pH 8.0 with 5 mM 2-mercaptoethanol. The cell paste was lysed by two passages through a French pressure cell (19000 p.s.i., cell pre-chilled to 4°C) and the lysate clarified by centrifugation at 27000 x *g* for 30 minutes at 4°C. The supernatant was removed and stored on ice until assayed. Assay mixtures (100 µl) contained 100 mM potassium phosphate pH 8.0, 5 mM ATP, 5 mM $MgCl_2$, 2 % DMSO, 0.9 mM acyl substrate (either acetyl-CoA, palmitoyl-CoA or palmitic acid), 65 mg bovine serum albumin, 0.5 µCi NaH[^{14}C]CO_3 (0.1 Ci/mol; AP Biotech CFA421) and 250 µg total protein from the lysate. All of the components of the assay excluding the enzyme

extract were pre-mixed, and then mixed with the enzyme extract to initiate the reaction, which was held at 30°C for 30 minutes. The reaction was quenched by the addition of 50 µl of concentrated HCl. This mixture was then held at 95°C under a stream of air to remove non fixed [^{14}C]CO_2. After evaporation to dryness, the residue was dissolved in 1 ml of water and added to 10 ml EcoscintA (National Diagnostics) prior to liquid scintillation counting.

4.2.7 Acetyl-CoA carboxylation assay

To assay for acetyl-CoA dependent malonyl-CoA formation, cells grown on complex medium CGIII were washed twice with 0.9 % NaCl, resuspended in 60 mM Tris-HCl pH 7.2 and disrupted as described above. Assay mixtures (250 µl) contained 60 mM Tris-HCl pH 7.2, 65 mM $KHCO_3$, 1 mM ATP, 1.5 mM $MgCl_2$ and 135 µl of cell extract. The mixture was pre-incubated at 30 °C for 1 minute, and the reaction initiated by adding acetyl-CoA (Lithium salt, Roche) to give a final concentration of 2 mM and further incubated at 30°C. Samples were taken at the given time intervals and mixed immediately with 5 µl of 30 % perchloric acid to quench the reaction. The assay tubes were centrifuged subsequently (14000 x g, 4°C, 5 minutes), 50 µl of the supernatant withdrawn, neutralised with 12.5 µl Na_2CO_3 and analysed *via* HPLC. The acyl-CoA synthesis was monitored by reversed-phase chromatography using a column (125 x 4 mm) filled with LiChrospher 1000 RP 18-EC-5µ (Merck KGaA, Darmstadt, Germany) on an Agilent 1100 Series HPLC (Palo Alto, California, USA). Samples (12 µl) were automatically injected, and separated by an increasing gradient consisting of 50 mM Na-phosphate buffer pH 5.0 containing 2 % acetonitrile, increasing up to 26 % acetonitrile in the same buffer at a flow rate of 0.3 ml/min, and monitored at 254 nm. Reactions were compared to known standards of acetyl-CoA, malonyl-CoA, and CoA in concentrations of 0.1 to 1 mM dissolved in 200 mM Na-phosphate buffer pH 3.0.

4.3 Results

4.3.1 The *accD* genes of *Corynebacterianeae*

From the recent Himar1-based transposon mutagenesis in *M. tuberculosis* (Sassetti *et al.*, 2003) *accD4* and *accD6* were found to be essential genes, whereas *accD1* and *accD3* are non-essential. In terms of *accD1* (previously *dtsR1*) of *C. glutamicum* an involvement in L-glutamate formation was demonstrated which is related to fatty acid synthesis (Kimura *et al.*, 1997). To address the problem on functional similarities of the carboxyltransferase proteins we performed a phylogenomic analysis of the *accD* genes present in the genomes of the sequenced *Corynebacterianeae* (Cerdeno-Tarraga *et al.*, 2003; Cole *et al.*, 1998; Cole *et al.*, 2001; Garnier *et al.*, 2003; Kalinowski *et al.*, 2003; Nishio *et al.*, 2003).

In fatty acid biosynthesis, malonyl-CoA is required for each elongation step, which is generated by carboxylation of acetyl-CoA. The process is catalysed by an enzyme consisting of a biotinylated α-subunit and a β-subunit, the latter being the actual carboxyltransferase. In *E. coli* and *Bacillus subtilis* each of these subunits consists of two polypeptides, whereas in all *Corynebacterianeae* analysed, we have found that the α- and the β-subunits consist of a single polypeptide. *C. glutamicum* has four putative carboxyltransferase β-subunits (*accD* genes) its close relatives *C. efficiens* and *C. diphteriae* possess five, whereas *M. tuberculosis, M. bovis,* as well as *M. marinum* possess as many as 6 carboxyltransferases. The sequence identities of the polypeptides within one organism are also in part exceptional. For instance, AccD1 and AccD2 of *M. tuberculosis* have identities of 49 %, and even the most distant paralogous, which are AccD1 and AccD3 still share 22 % identity. The primary structures of 35

polypeptides were analysed using ClustalW (Pearson, 1990), with the results presented in Figure 4.1.

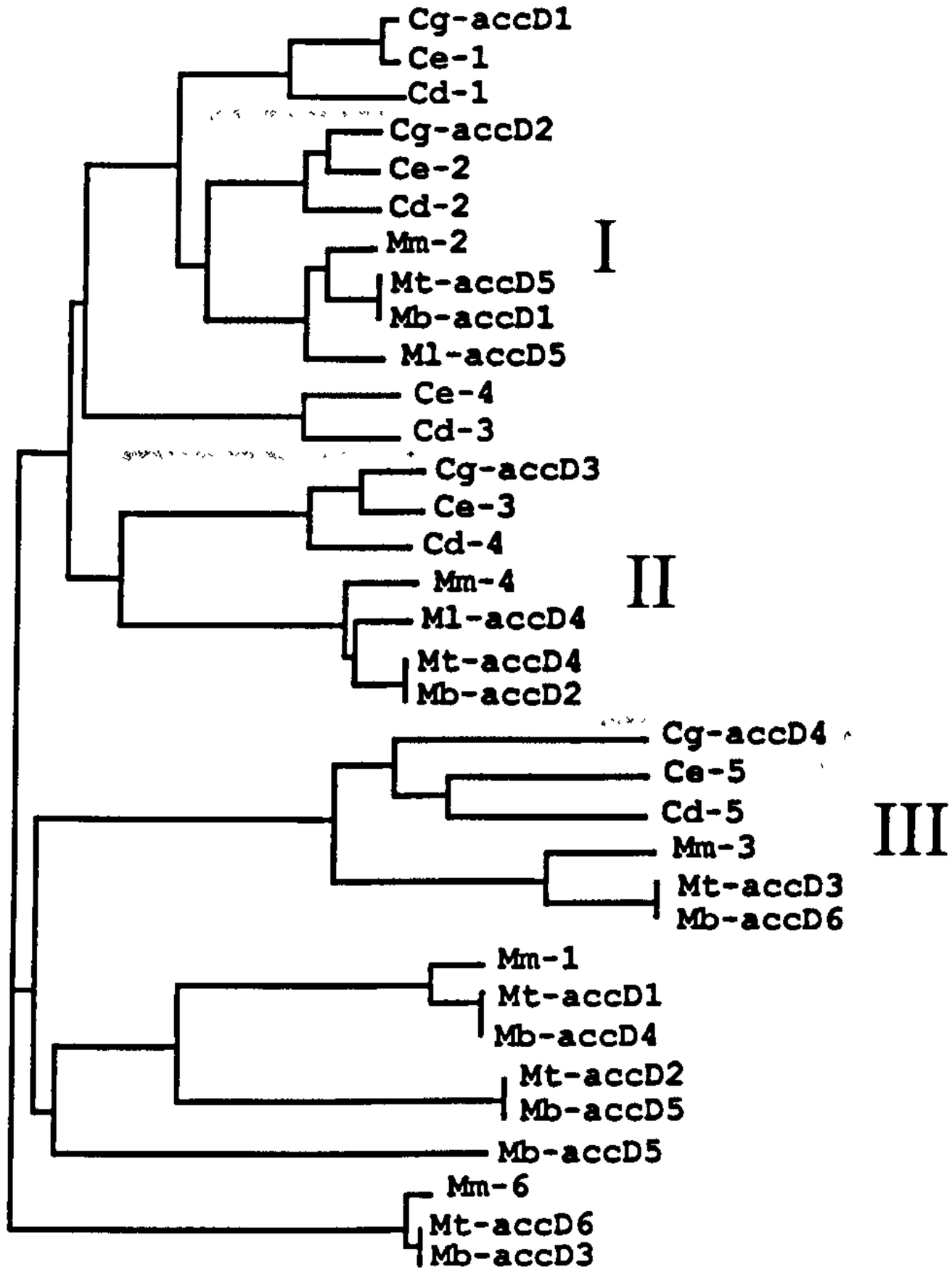


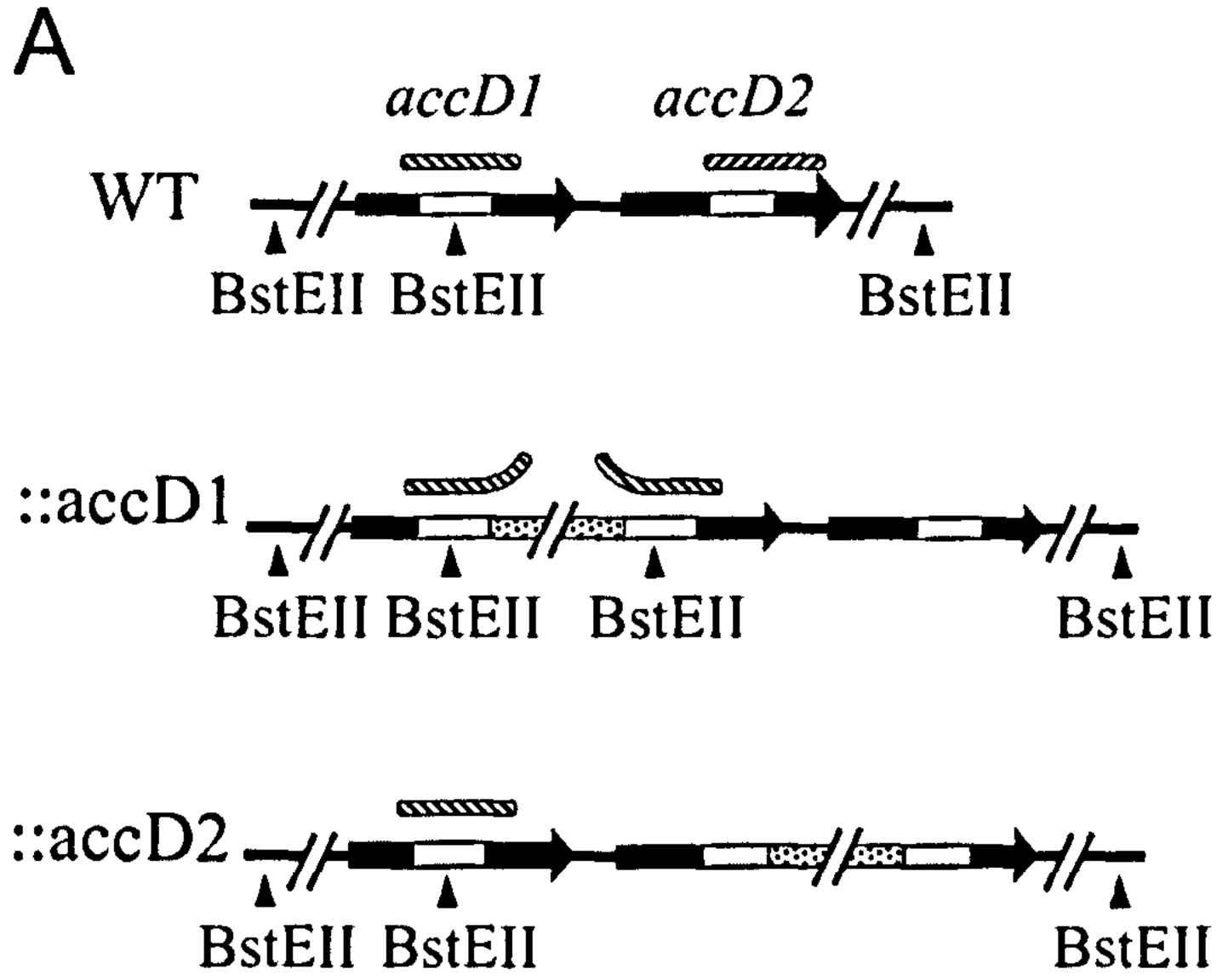
Figure 4.1 **Phylogenomic analysis of mycobacterial and corynebacterial *acc* genes. Groups of orthologous carboxyltransferases are boxed in grey and labelled I-III. The organism key is as follows: Cg, *C. glutamicum*; Ce, *C. efficiens*; Cd, *C. diptheriae*; Mt, *M. tuberculosis*; Mb, *M. bovis*; Mm, *M. marinum*; Ml, *M. leprae*. Diagram provided by Lothar Eggeling, (Gande *et al.*, 2004)**

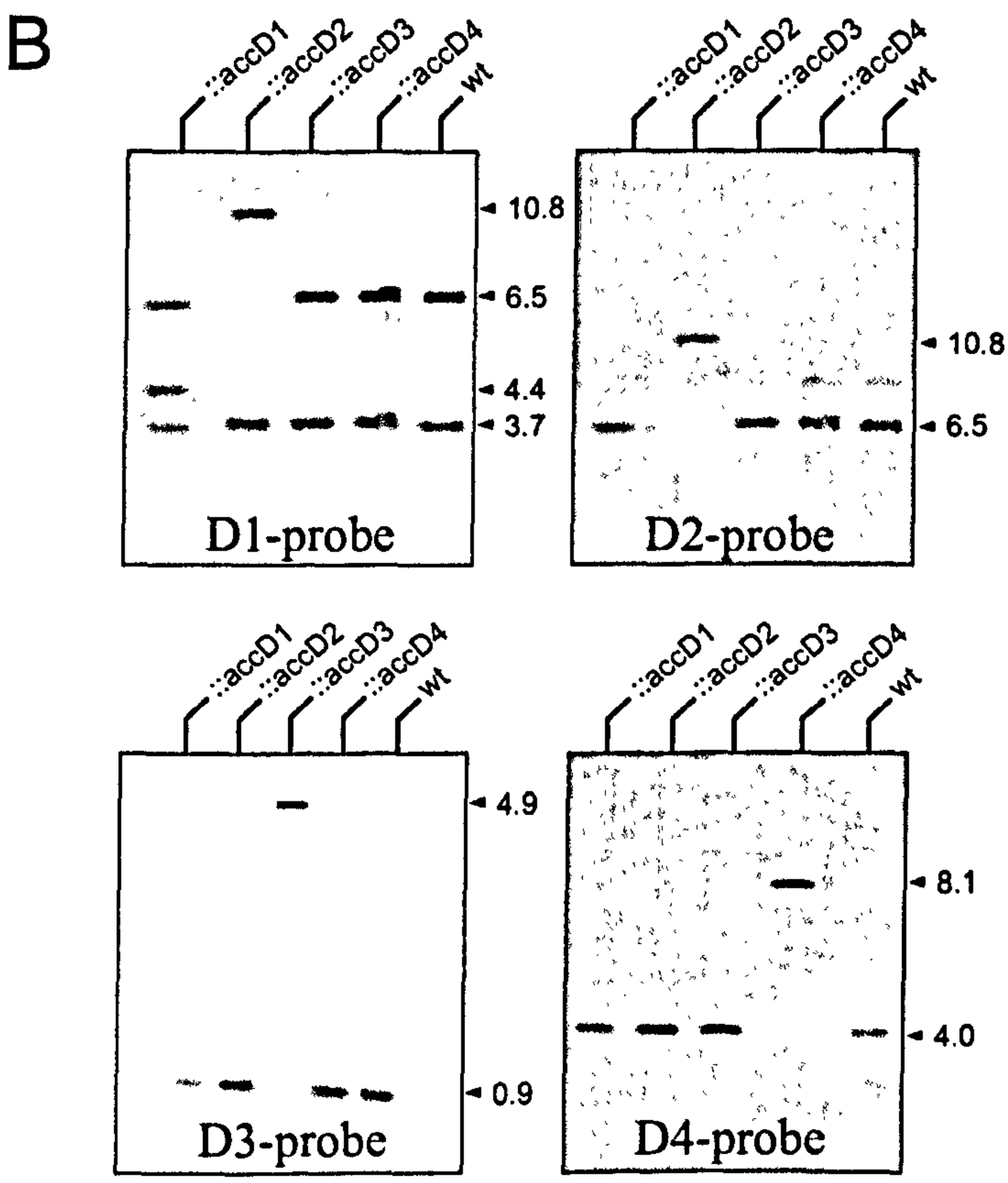
There are several remarkable features. Firstly, there are two distinct groups (I and II) of closely related proteins, where each organism analysed is represented only once with one carboxyltransferase. Group III is clearly distant from groups I and II, whereas the latter two are more closely related. The distinct clustering of these groups suggests that members of

these groups are orthologues in each group with an identical function. Secondly, there are three additional clusters, where *M. bovis* and *M. tuberculosis* are represented once. Also, *M. marinum* is present in two of these clusters. These results indicate that these carboxyltransferases serve more specialised functions within the mycobacterial branch of *Corynebacterianeae*, whereas that of clusters I and II, represent fundamental core functions specific for lipid biosynthesis in all bacteria. In addition, *M. leprae* possesses 3 pseudogenes sharing in part identities with the other carboxyltransferases (Cole *et al.*, 2001). However, the fact that in *M. leprae* the two single carboxyltransferases fully retained are members of group I and II, strengthens the view that these members encode essential functions in lipid biosynthesis, as is the case with mycobacterial mycolic acid biosynthesis (Kremer *et al.*, 2000a). Since, carboxyltransferase activity requires an activated carboxy group derived from a biotinylated α-subunit we searched the genomes for the corresponding proteins. The three *Corynebacterium* species, as well as *M. leprae* have only one such subunit, whereas *M. tuberculosis, M. bovis* and *M. marinum* possess three.

4.3.2 Inactivation of the corynebacterial β-subunit genes

To systematically investigate the function of the *accD* genes each of them was inactivated in *C. glutamicum*. For this purpose we constructed non-replicative vectors to enable their specific disruption by use of internal fragments of *accD1*, *accD2*, *accD3*, and *accD4*, respectively (Figure 4.2A). After transformation kanamycin resistant clones were assayed for vector integration by PCR with two different primer sets.


A
accD1
accD2
WT
BstEII BstEII
BstEII
::accD1
BstEII BstEII BstEII
BstEII
::accD2
BstEII BstEII
BstEII


B
::accD1
::accD2
::accD3
::accD4
wt
10.8
6.5
4.4
3.7
D1-probe
::accD1
::accD2
::accD3
::accD4
wt
10.8
6.5
D2-probe
::accD1
::accD2
::accD3
::accD4
wt
4.9
0.9
D3-probe
::accD1
::accD2
::accD3
::accD4
wt
8.1
4.0
D4-probe

Their correct chromosomal integration was further confirmed by a scrupulous Southern-blot analysis, with the final result shown for each mutant in Figure 4.2B. The three chromosomal fragments derived from the *accD1* inactivation mutant hybridised with the *accD1*-probe (Figure 4.2B, upper left panel), which is in agreement with the correct integration and inactivation of *accD1*. The fact that a shift occurs with the same *accD1*-probe with DNA prepared from the *accD2* mutant is due to the immediate proximity of both *accD1* and *accD2* separated by 344 bp (Figure 4.2A). One of these genes probably originated by a recent gene duplication event since the paralogues share more than 74 % identical nucleotides over 75 % of their length. With the *accD2*, *accD3* and *accD4*-probes the chromosomal fragment of each respective mutant carrying the inactivation vector was shifted to the expected fragment increased by the vector length (3.99 – 4.35 kb) (Figure 4.2B). These results confirm the specific inactivation of each *accD* gene and the integrity of the gene locus in the mutants generated.

For control, the four inactivation mutants were made competent and transformed with either pEC7-*accD1*, *accD2*, *accD3*, or *accD4*, respectively. Resulting recombinants were confirmed by plasmid preparations and PCR analysis to have the chromosomal locus disrupted with an intact copy of the respective gene plasmid encoded. Growth of these complemented mutants was undistinguishable from that of wild type *C. glutamicum* either on LBHIS or minimal medium CGXII. This makes it unlikely that the *accD* disruptions result in any polar effects.

4.3.3 Phenotypic characterisation of mutants

As shown in Figure 4.3 the *accD4* mutant (compare left and right panels) did not exhibit a phenotype, whereas growth of the *accD1*, *accD2* and *accD3* mutants were strongly reduced on salt medium CGXII (upper row). The worst growth was in fact seen with the *accD1* inactivation mutant, and the colonies visible in Figure 4.3 are due to revertants.

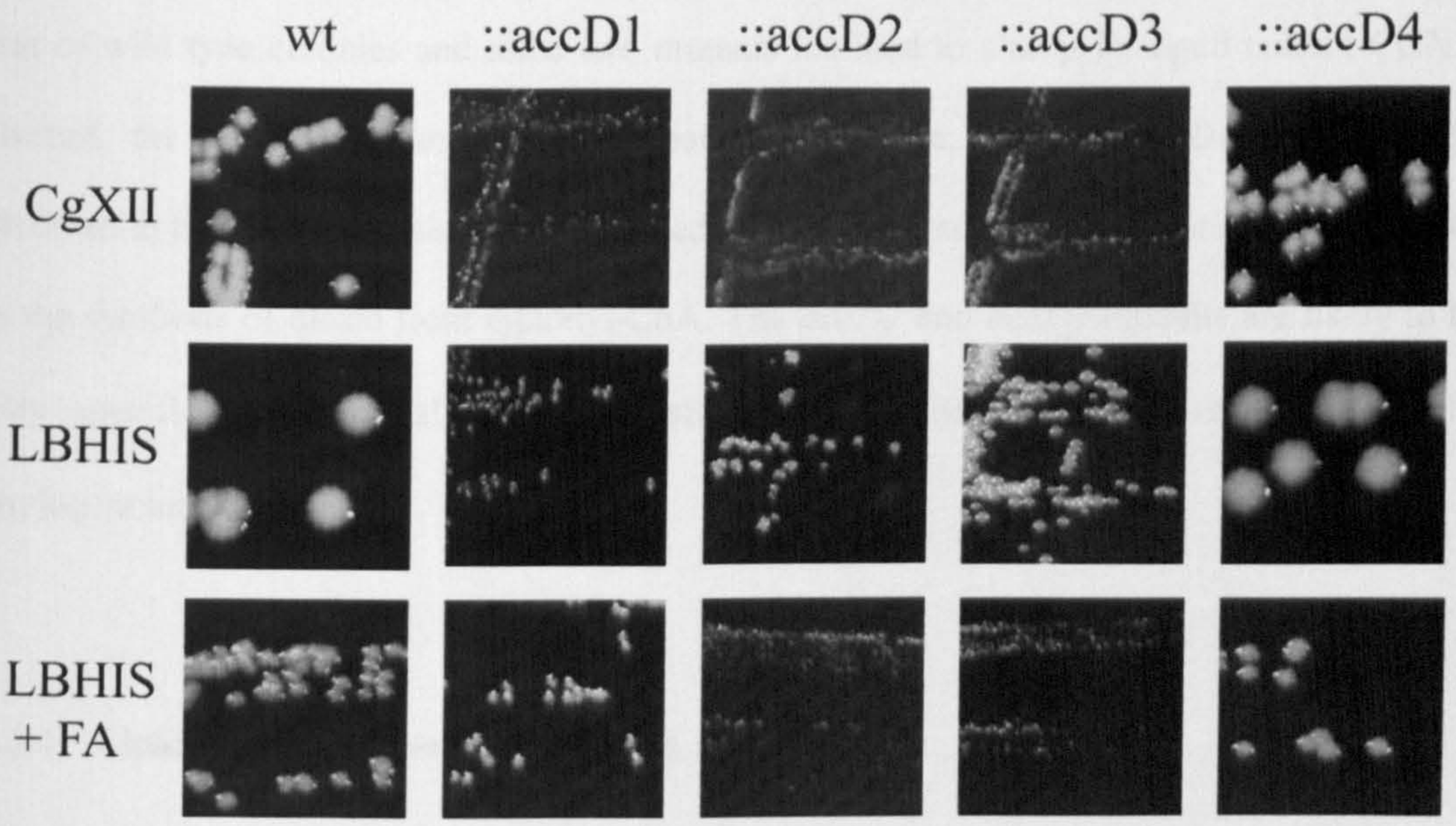


Figure 4.3 **Growth of *accD* mutants of *C. glutamicum*. The *accD* mutants of *C. glutamicum* and the wild type (wt) were grown for three days at 30°C. The plates used were salt medium with glucose (CGXII), complex medium (LBHIS), and complex medium supplemented with Na-oleate (300 mg/l), and butter hydrolysate (300 mg/l) together with the detergents Tween 40 (300 mg/l) and Brij-35 (400 mg/l)**

Also on complex medium LBHIS growth of the three mutants was still impaired illustrating that these mutants are not rescued by specific components supplied from yeast extract and tryptone, which are present in LBHIS. However, when this complex medium was supplemented with Na-oleate together with butter hydrolysate (and detergents for emulsification) growth of the *accD1* mutant was markedly improved, whereas the *accD2* and *accD3* mutants were still impaired, indicating a direct relation of *accD1* with fatty acid biosynthesis, and confirming previous observations with a *dtsR1* deletion mutant (Kimura *et al.*, 1997). The colony surface of both the *accD2* and the *accD3* mutants were rougher than that of wild type colonies and these two mutants inclined to clump in liquid culture CGXII. Overall, the *accD4* mutation has no apparent phenotype, whereas *accD1* appears to be involved in lipid biosynthesis, as it is rescued by Na-oleate suggesting that it may be involved in the synthesis of oleate from malonyl-CoA. The *accD2* and *accD3* mutants are likely to be very specific and essential carboxyltransferase β-subunits, as also illustrated from the phylogenomic analysis.

4.3.4 Lipid analysis of the *accD* mutants

In order to relate the phenotypic changes of the *accD* mutants to their cellular composition, the mutants and wild type bacteria were grown on salt medium CGXII for 4 hours and then labelled with [^{14}C]-acetate. After further incubation the cells were harvested, dried and analysed for their lipid composition. The free lipids were extracted with chloroform-methanol-water and analysed by TLC as shown in Figure 4.4A. The *accD1* and *accD4* mutation did not change considerably the relative amounts of phospholipids, trehalose monomycolates (TMM) and dimycolates (TDM), and glucose monomycolates (GMM). However, the consequences of the *accD2* and *accD3* mutations were dramatic. Surprisingly, both mutations resulted in the

complete loss of "free" and extractable mycolates (TMM, TDM and GMM), without effecting phospholipid synthesis, and the possible accumulation of an unknown intermediate. The extracted lipids were also hydrolysed and their methyl esters separated by TLC confirming that the synthesis of fatty acids was not affected in the *accD2* and *accD3* mutants.

The remaining cell wall bound lipids were analysed by hydrolysis and the preparation of methyl esters. The profile of the extracted FAMES and MAMES is shown in Figure 4.4B. In the *accD2* and *accD3* mutant cell wall bound mycolic acids were absent, whereas the relative amounts of fatty acids and mycolic acids in the *accD1* and *accD4* mutants were comparable. These results illustrate the specific involvement of both *accD2* and *accD3* in mycolic acid synthesis, and that neither of these two genes themselves can complement each other, nor *accD1* or *accD4*.

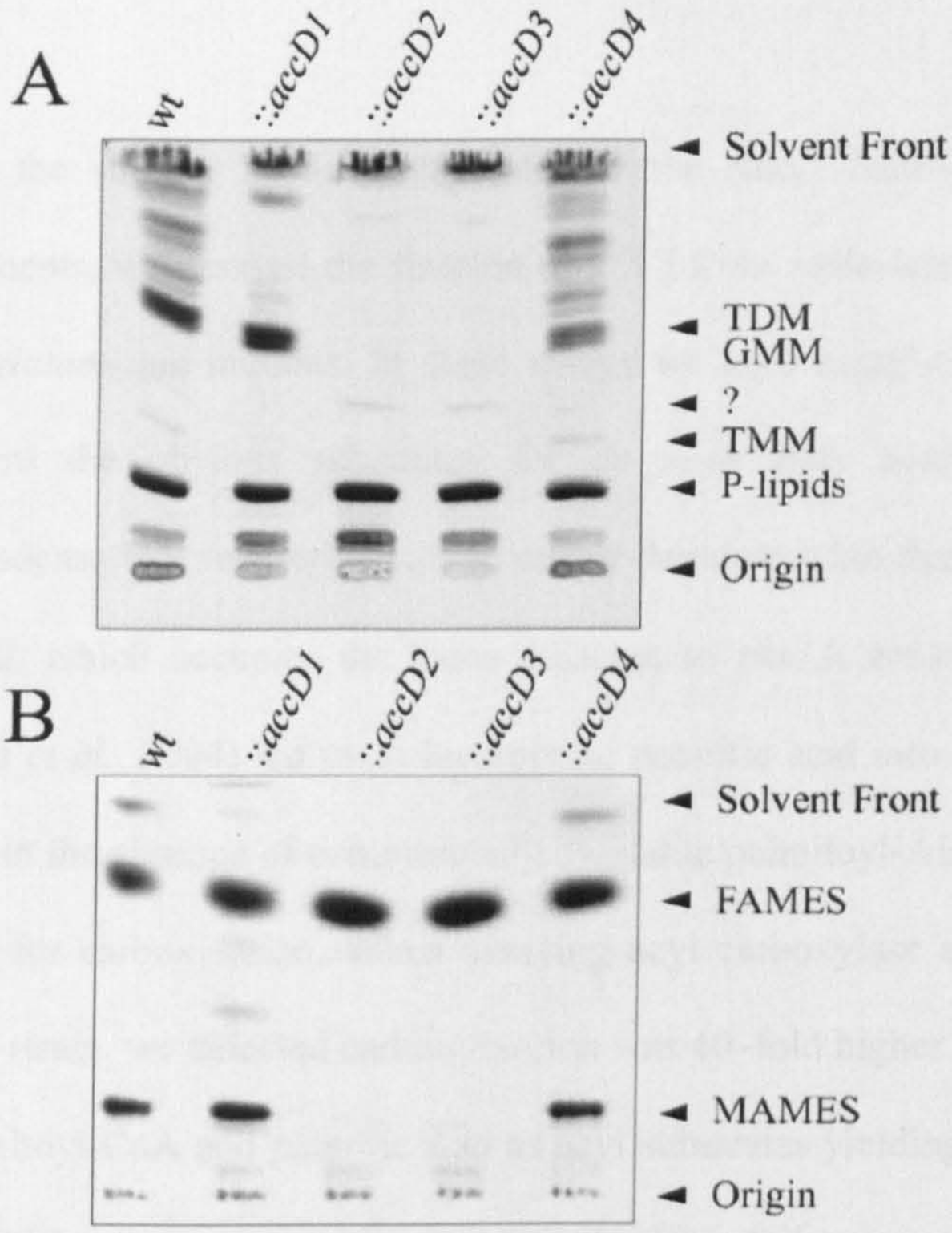
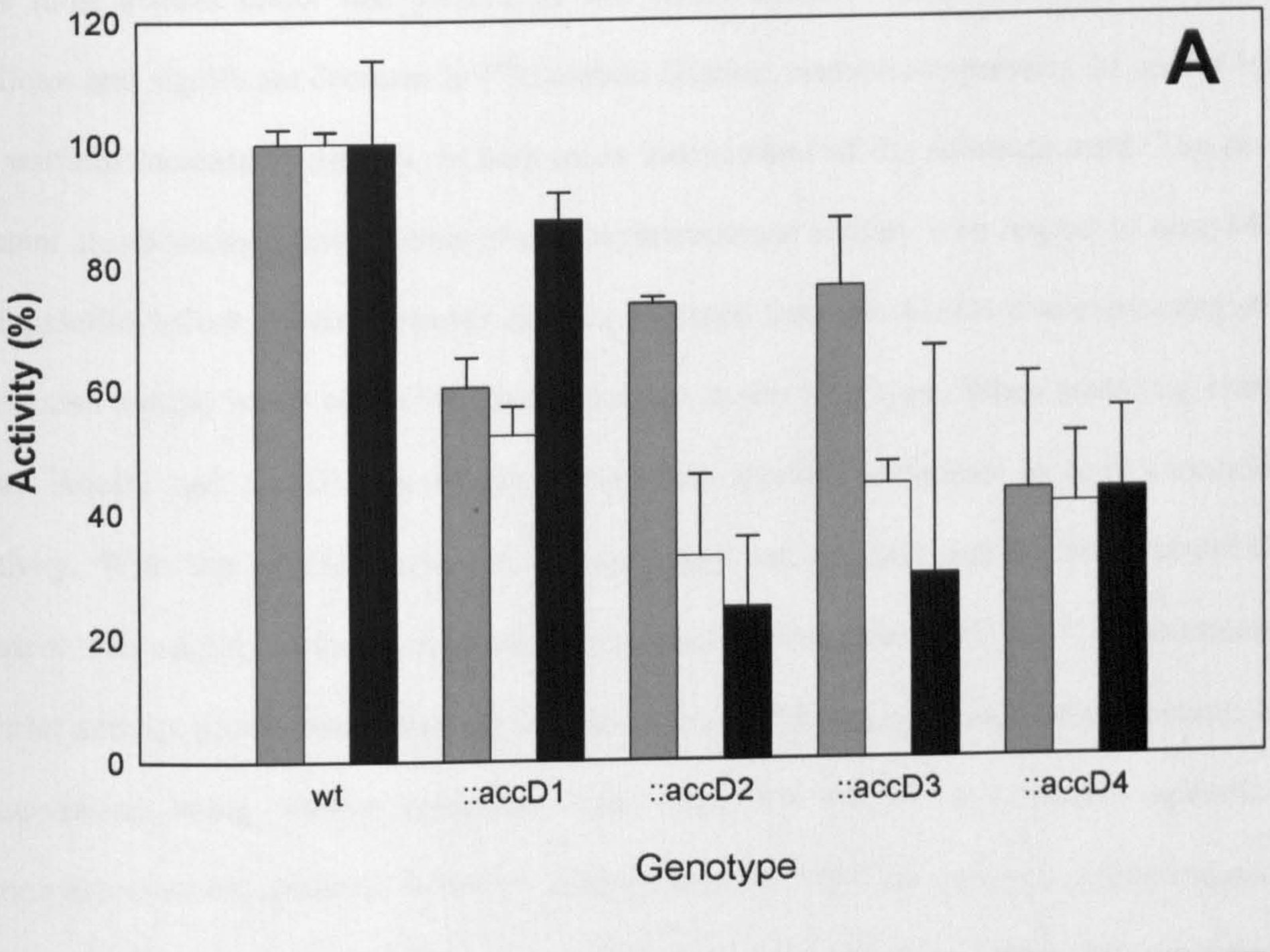
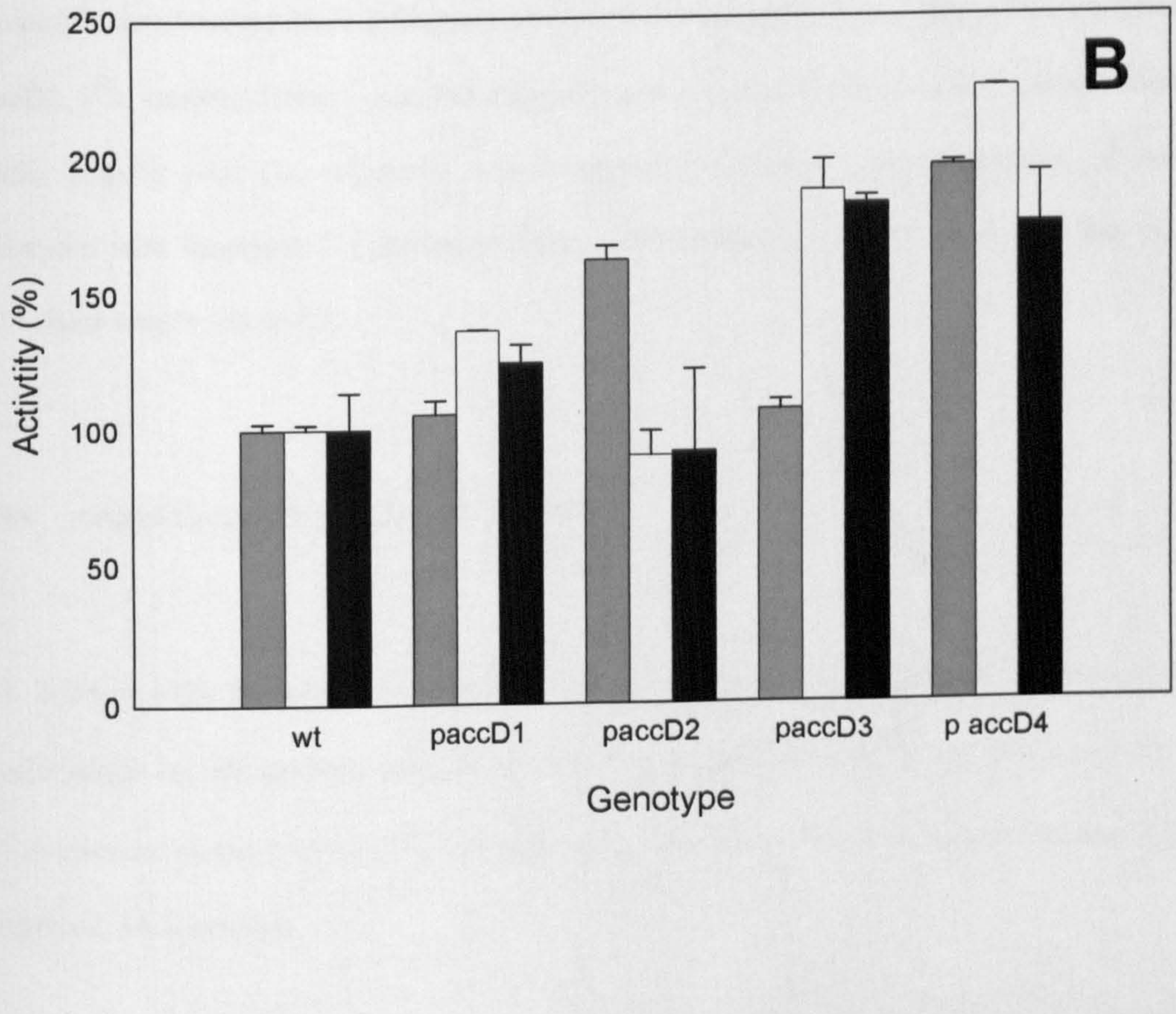


Figure 4.4 **Lipid analysis of *accD* mutants of *C. glutamicum*. A) Free lipids were extracted and an aliquot (20,000 cpm) from each strain subjected to TLC analysis using silica gel plates (5735 silica gel $60F_{254}$, Merck), developed in $CHCl_3/CH_3OH/H_2O$ (60:16:2, v/v/v). Autoradiograms were produced by 2-3 day exposure to Kodak X-Omat AR film to reveal [^{14}C]-labelled lipids and compared to known standards (TMM, TDM, and GMM, (Puech *et al.*, 2000)). B) The bound lipids from the de-lipidated extracts were released by the addition TBAH followed by overnight incubation at 95°C. Following preparation of methyl esters an aliquot (10,000 cpm) from each strain was subjected to TLC analysis using silica gel plates (5735 silica gel $60F_{254}$, Merck), developed in petroleum ether/acetone (95:5, v/v). Autoradiograms were produced by overnight exposure to Kodak X-Omat AR film to reveal [^{14}C]-labelled FAMES and MAMES, and compared to known standards**

4.3.5 Carboxylation reactions in the *accD* mutants

In order to assess the validity of the assignment of the AccD family members as acyl carboxylase components, we assayed the fixation of [^{14}C] from radio-labelled bicarbonate in extracts of the *C. glutamicum* mutants. In these assays we used acetyl-CoA and palmitoyl-CoA as these were the obvious substrates for *de novo* fatty acid biosynthesis and corynomycolate condensation, respectively. The recent demonstration that the product of *M. tuberculosis fadD32*, which occupies the locus adjacent to *pks13*, activates fatty acids by adenylation (Trivedi *et al.*, 2004) led us to incorporate palmitic acid into our experiments to determine whether, in the absence of commercially available palmitoyl-ADP, it may generate a suitable substrate for carboxylation. When assaying acyl carboxylase activity in clarified lysates of the parent strain, we detected carbon fixation was 40–fold higher when using acetyl-CoA than with palmitoyl-CoA and palmitic acid as acyl substrates yielding specific activities of 2.45, 0.06 and 0.06 nmol min^{-1} mg^{-1}, respectively. The following reaction rates were normalised to these values. Basically, with each *accD* mutant and overexpressing strain, a significant influence on carbon fixation was apparent (Figure 4.5).


A
120
100
80
60
40
20
0
Activity (%)
wt
::accD1
::accD2
::accD3
::accD4
Genotype


B
250
200
150
100
50
0
Activtity (%)
wt
paccD1
paccD2
paccD3
p accD4
Genotype

The most distinct effect was present for the *accD4* mutant where inactivation resulted in uniform and significant decrease in [^{14}C]carbon fixation, and over-expression of *accD4* led to an uniform increase in activity, in both cases independent of the substrate used. The *accD1* mutant also possessed lower levels of carboxyltransferase activity with respect to acetyl-CoA and palmitoyl-CoA effects, however extracts prepared from the AccD1 overexpressing strain possessed similar levels of activity in comparison to the wild type. When analysing extracts from AccD2 and AccD3, we observed substrate specific variations in acyl carboxylase activity. With the *accD2* inactivation mutant, acyl carboxylase activity with acetyl-CoA present was slightly reduced, whereas it was most pronounced with the C_{16} substrates. A similar activity profile was observed with the extract of the *accD3* inactivation mutant; both observations being wholly consistent with roles for AccD2 and AccD3 specific to corynomycolate biosynthesis. When we assayed extracts from bacteria overexpressing *accD2* or *accD3*, we observed clear differences in their substrate preference. The extract enriched in AccD3, [^{14}C]carbon fixation detected using acetyl-CoA was similar to the wild-type extract, whilst activity with C_{16} substrates was increased. In contrast, over-expression of *accD2* coincided with increased [^{14}C]carbon fixation when using acetyl-CoA rather than the longer acyl chain length substrates.

4.3.6 Acetyl-CoA carboxylation by AccD1

We inferred from the growth response of the *accD1* inactivation mutant (Figure 4.3) that *accD1* might encode the β-subunit of the carboxylase transfering the carboxyl group from the β-subunit onto acetyl-CoA to yield malonyl-CoA. Therefore a HPLC assay was used to follow malonyl-CoA formation.

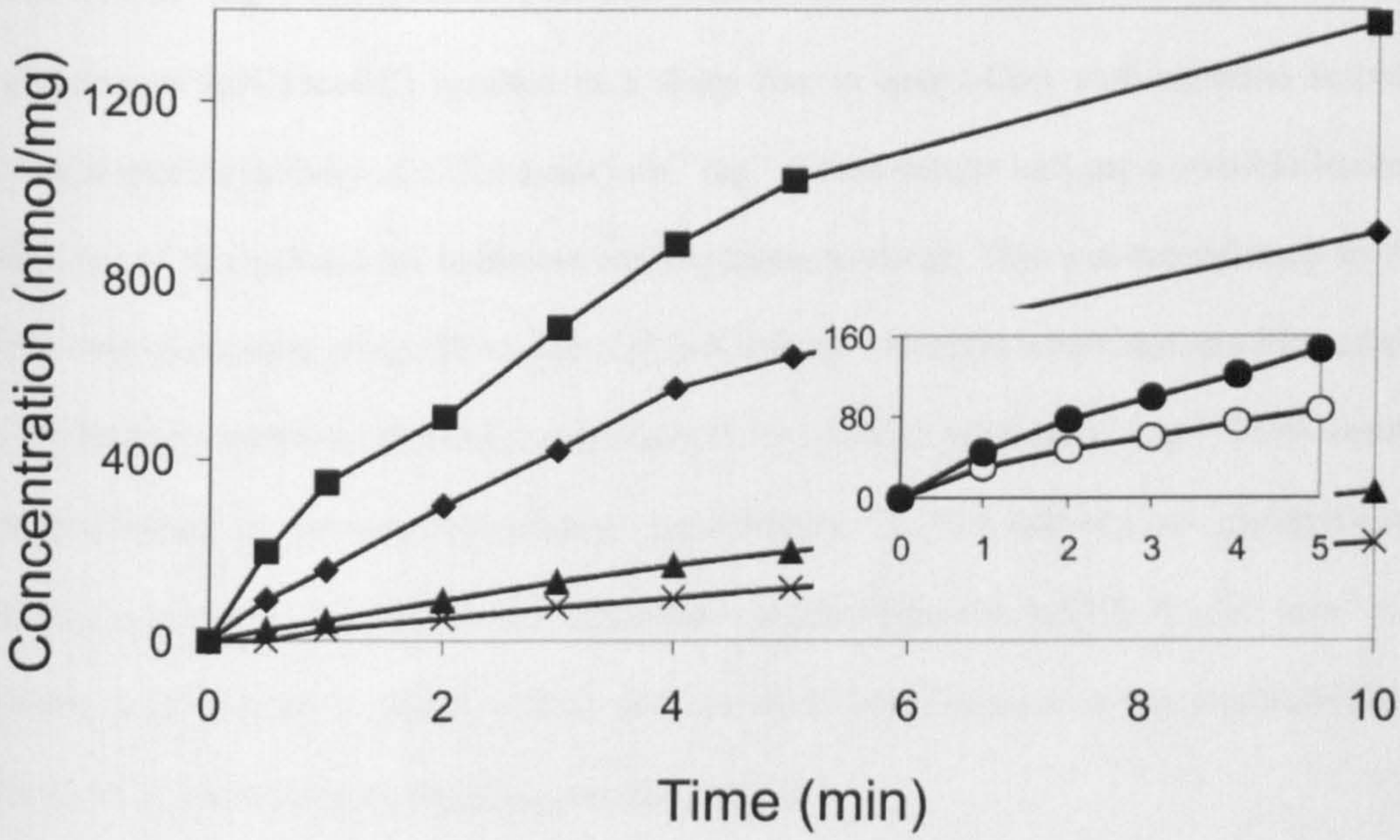


Figure 4.6 **Malonyl-CoA formation with extracts of recombinant *C. glutamicum*. Malonyl-CoA synthesis shown in nmol *per* mg protein of recombinant *C. glutamicum* strains derived from the wild type strain and with the genotypes Δ*pyc* pJC1*accBC* pEC7*accD1* (■), Δ*pyc* pJC1*accBC* pEC7 (◆), Δ*pyc* pJC1 pEC7*accD1* (▲), and Δ*pyc* as control (X) after incubation at 30°C at given time intervals. The inset shows comparable analyses using a separate *accD1* overexpressing strain pVWEx2*accD1* (●) compared to control with empty vector pVWEx2 (○).**

As shown in Figure 4.6, use of extracts derived from wild type *C. glutamicum* results in a constant increase in malonyl-CoA formation over time that is dependent upon acetyl-CoA and ATP. The derived specific activities were 23.1 nmol min^{-1} mg^{-1}, which is in the range of that determined for *M. tuberculosis* and *M. bovis* (Rainwater & Kolattukudy, 1982). When *accD1* was overexpressed (pJC1accD1), the specific activity increased 2-fold to 40.7 nmol min^{-1} mg^{-1}. This significant increase was confirmed in a separate experiment using a separate construct (pVWEx2accD1), see inset in Fig. 4.6, where the specific activity was increased from 15.1 to

27.2 nmol min^{-1} mg^{-1}. Unexpectedly, the over-expression of the biotin-containing α-subunit of *C. glutamicum* (pJC1accBC) resulted in a sharp rise in acetyl-CoA carboxylation activity, yielding a specific activity of 132.8 nmol min^{-1} mg^{-1}. These results indicate a possible limiting availability of this subunit for sufficient carboxylation to occur. This was reconfirmed by the mutual over-expression of *accBC* and *accD1* (pJC1accBC paccD1) where the specific activity was additionally increased to yield a specific activity of 180.3 nmol min^{-1} mg^{-1}. These results were confirmed in several independent experiments. A low activity of malonyl-CoA formation was observed in the *accD1* inactivation mutant (specific activity 8 nmol min^{-1} mg^{-1}), where poor growth in liquid culture and malonyl-CoA detection at the threshold level prevented a more detailed quantification in this mutant.

In summary, the acyl carboxylase assays described above for AccD1-4 demonstrate that these enzymes possess carboxylase activity, with AccD3 (and possibly AccD2) preferring long-chain substrates consistent with their involvement in mycolic acid biosynthesis, and that AccD1 uses acetyl-CoA as a substrate.

4.3.7 The *pks* locus of *Corynebacterianeae*

Inspection of the chromosomal organisation of the *accD* genes in *C. glutamicum* revealed the localisation of *accD3* close to the single *Cg-pks* gene in this bacterium (Figure 4.7). The *Cg-pks* has high identity to *pks13* of *M. tuberculosis* and is located in a locus strictly conserved in all *Mycobacteria* and *Corynebacteria* species (Figure 4.7), as it is also in *Rhodococcus* (Portevin *et al.*, 2004). It consists of at least six genes, all of them transcribed in the same direction and extending over more than 12 kb.

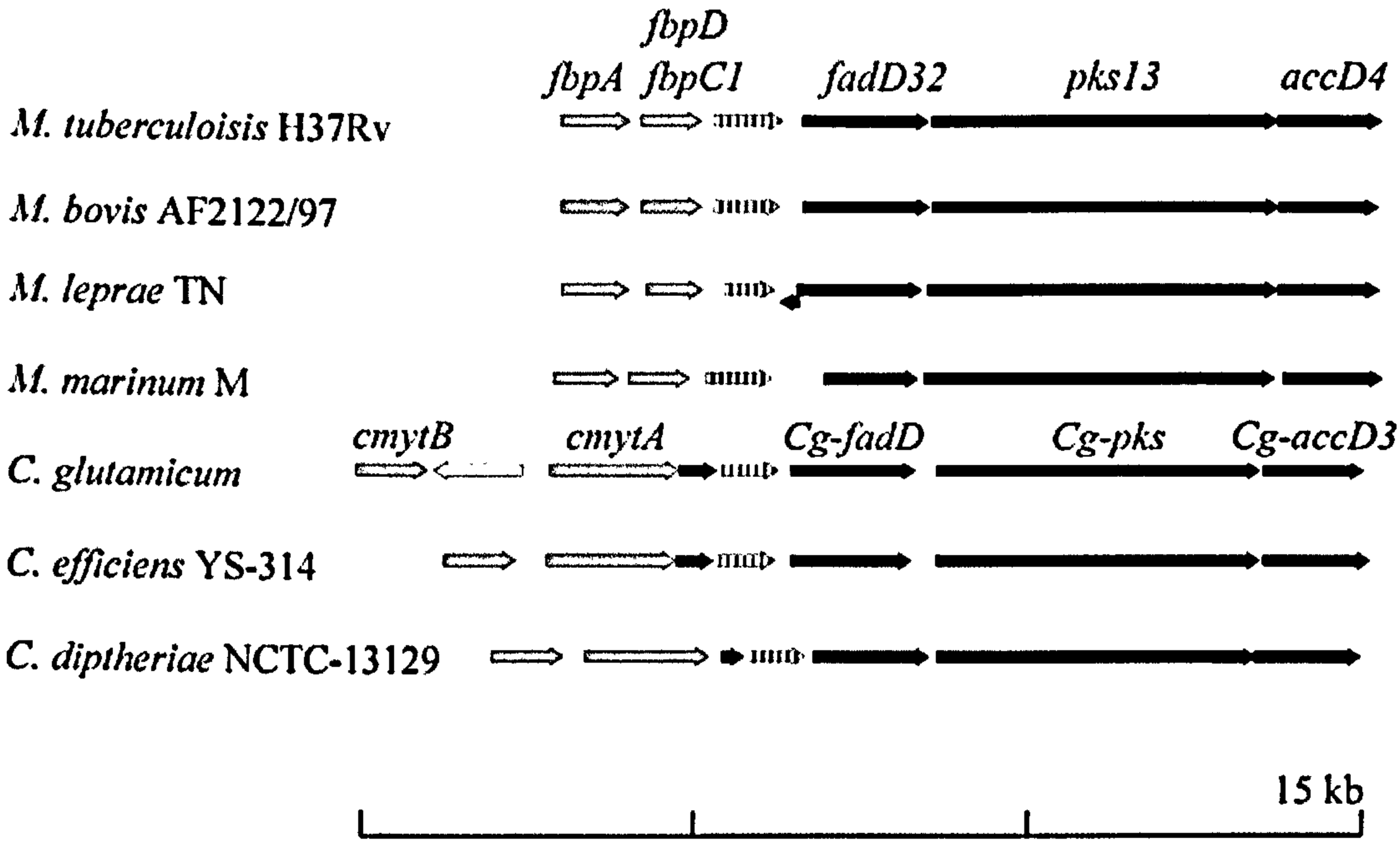


Figure 4.7 The *pks* locus of mycolic acid containing bacteria responsible for mycolic acid biosynthesis

In *M. tuberculosis*, the first two genes are *fbpA* and *fbpC1*, which encode mycolyltransferases, respectively (Belisle *et al.*, 1997). These mycolyltransferases are followed by an essential gene (Sassetti *et al.*, 2003), whose product is predicted to be anchored in the membrane by one transmembrane spanning helix. The remaining segment (αα 35–336) is probably directed towards the periplasm and is predicted to possess an esterase acitivity. Further, located downstream is *fadD32*, recently shown to activate long chain fatty acids as acyl-adenylates (Trivedi *et al.*, 2004), followed by polyketide synthase *pks13*. The gene locus is completed with the β-chain of the acyl carboxylase *accD4* (*accD3* in *C. glutamicum*). The overall organisation of the entire locus in all species analysed is almost identical, although a slight difference is apparent in the mycolyltransferase region (grey arrows in Figure 4.7). In *C. glutamicum* a transposase is located between *cmytA* and *cmytB*. The strong conservation of this entire locus, the known mycolyltransferase activities in *M. tuberculosis* (Belisle *et al.*,

1997) and *C. glutamicum* (Brand *et al.*, 2003; De Sousa-D'Auria *et al.*, 2003), the absence of mycolic acids in *accD3* and *Cg-pks* mutants (see below and (Portevin *et al.*, 2004) and the acyl-AMP ligase identified (Trivedi *et al.*, 2004) suggests a fundamental and essential role of these genes in mycolic acid biosynthesis and translocation of mature mycolic acids to the cell wall and the outer lipid layer.

4.3.8 Deletion of *pks13*

In order to examine the phenotype of a *Cg-pks* mutant, whilst preserving the entire gene locus we attempted to delete the *pks* gene. For this purpose the upstream and downstream sequences adjacent to *Cg-pks* where fused, with the in-frame deletion in the construct verified by sequencing. The construct was used in two rounds of positive selection. The original clones obtained after 3 days were all shown to have the wild type locus restored. This indicates a strong disadvantage of cells deleted of *pks*. However, when we analysed the small colonies that appeared after 10 days, these had lost *Cg-pks*. One of these clones was chosen, verified by PCR (Figure 4.8) and Southern-blot analysis confirmed a *Cg-pks* deletion.

This strain, *C. glutamicum*Δpks, exhibited poor growth and a rough colony surface. Lipid analysis revealed that *C. glutamicum*Δpks, is devoid of extractable mycolic acids (Figure 4.8) and bound cell wall mycolic acids, although fatty acids are still present in phospholipids. Since the *Cg-pks* deletion has virtually the same consequences on mycolic acid synthesis as *accD2* and *accD3*, this illustrates that these genes are equally important and possibly act in concert in the final assembly of mature mycolic acids.

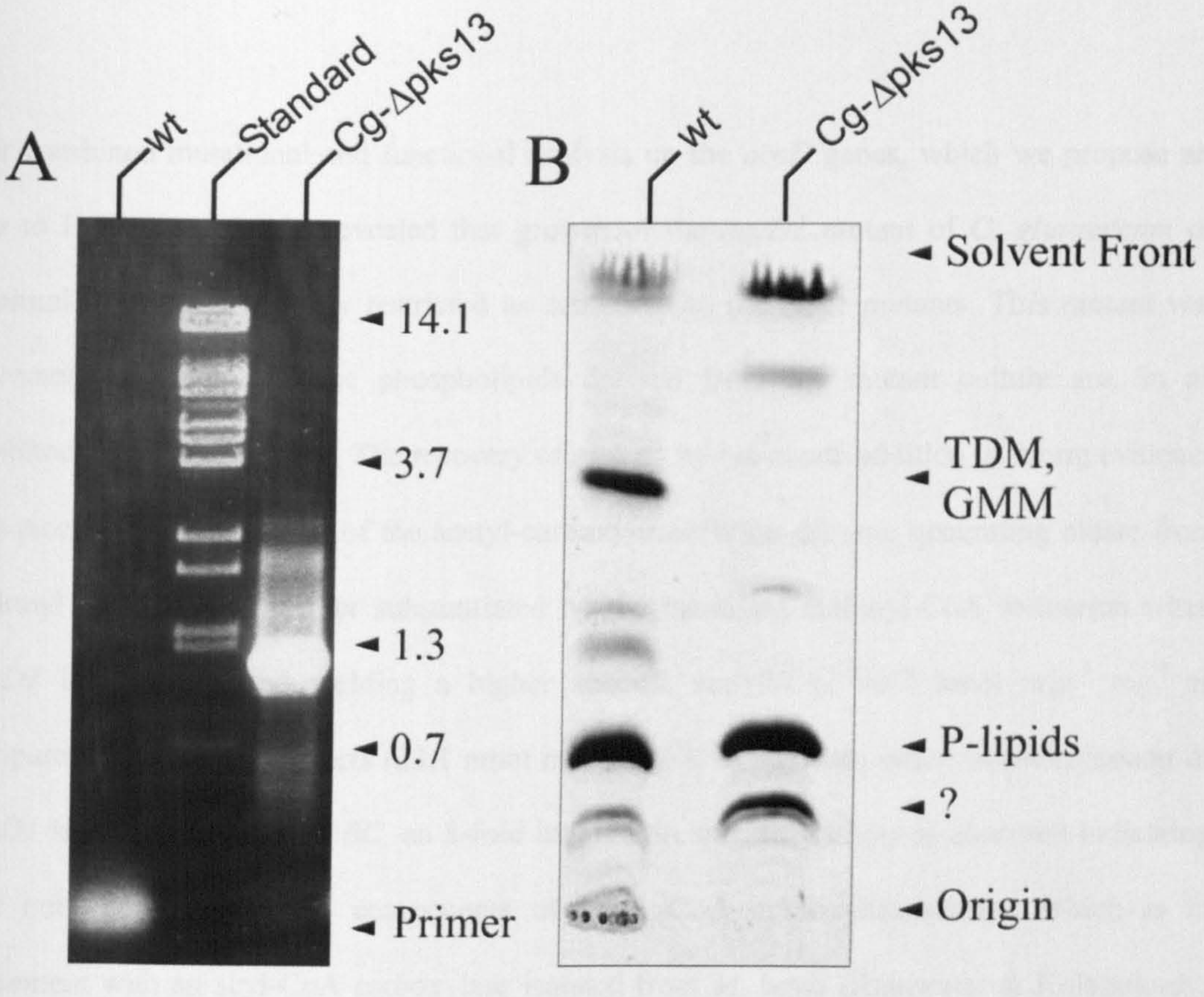


Figure 4.8 **Analysis of the *pks* deletion mutant of *C. glutamicum*. A) The panel shows the PCR analysis of the *Cg-pks* locus of the wild type (wt) and the in-frame deletion mutant (*C. glutamicum*Δpks). A standard is applied in the middle (standard), with the molecular weights given on the right. Whereas, with the primers annealing upstream and downstream of *Cg-pks* the expected amplification product of 1.11 kb is obtained with chromosomal DNA from *C. glutamicum*Δpks, no product can be formed under the conditions used with wild type DNA due to the length of *Cg-pks*. B) TLC analysis of extractable lipids from *C. glutamicum* (wt) and the deletion mutant (*C. glutamicum*Δpks), illustrate the absence of the dominant lipids, TDM and GMM, upon *Cg-pks* deletion**

4.4 Discussion

Our combined mutational and functional analysis on the *accD* genes, which we propose are key to lipid biosynthesis, revealed that growth of the *accD1* mutant of *C. glutamicum* on minimal medium is severely restricted as compared to the other mutants. This mutant was extremely unstable, and the phospholipids derived from the mutant culture are, in all likelihood, due to revertants. The recovery of growth by Na-oleate addition is strong evidence that AccD1 is a constituent of the acetyl-carboxyltransferase enzyme generating oleate from malonyl-CoA. This is further substantiated by the increased malonyl-CoA formation when *accD1* is overexpressed yielding a higher specific activity of 40.7 nmol min^{-1} mg^{-1} as compared to wild type extracts (23.1 nmol min^{-1} mg^{-1}). In addition, when over-expression of *accD1* is combined with *accBC,* an 8-fold increase in specific activity is observed indicating that both polypeptides are components of acetyl-CoA carboxyltransferase, which is in agreement with an acyl-CoA carboxylase isolated from *M. bovis* (Rainwater & Kolattukudy, 1982), and a propionyl-CoA carboxylase from *M. smegmatis* (Haase *et al.*, 1982), both of which are composed of 2 individual polypeptides, with one being biotinylated.

Extracts of the *accD4* mutant possessed lower levels of acyl carboxylase activity and extracts from the overexpressing *accD4* strain possessed increased levels of acyl carboxylase activity, suggesting that AccD4 is a carboxylase enzyme. However, the absence of a phenotype for the AccD4 mutants, as well as their unaltered lipid profiles led us to conclude that AccD4 is not directly involved in lipid biosynthesis. In sharp contrast are the AccD2 and AccD3 mutants where the lipid analyses revealed that both proteins are essential for corynomycolate biosynthesis. We cannot be sure whether the general carboxylation assays reported here represent an indication of relative acyl-carboxylase activity or if it simply reflects the half-life

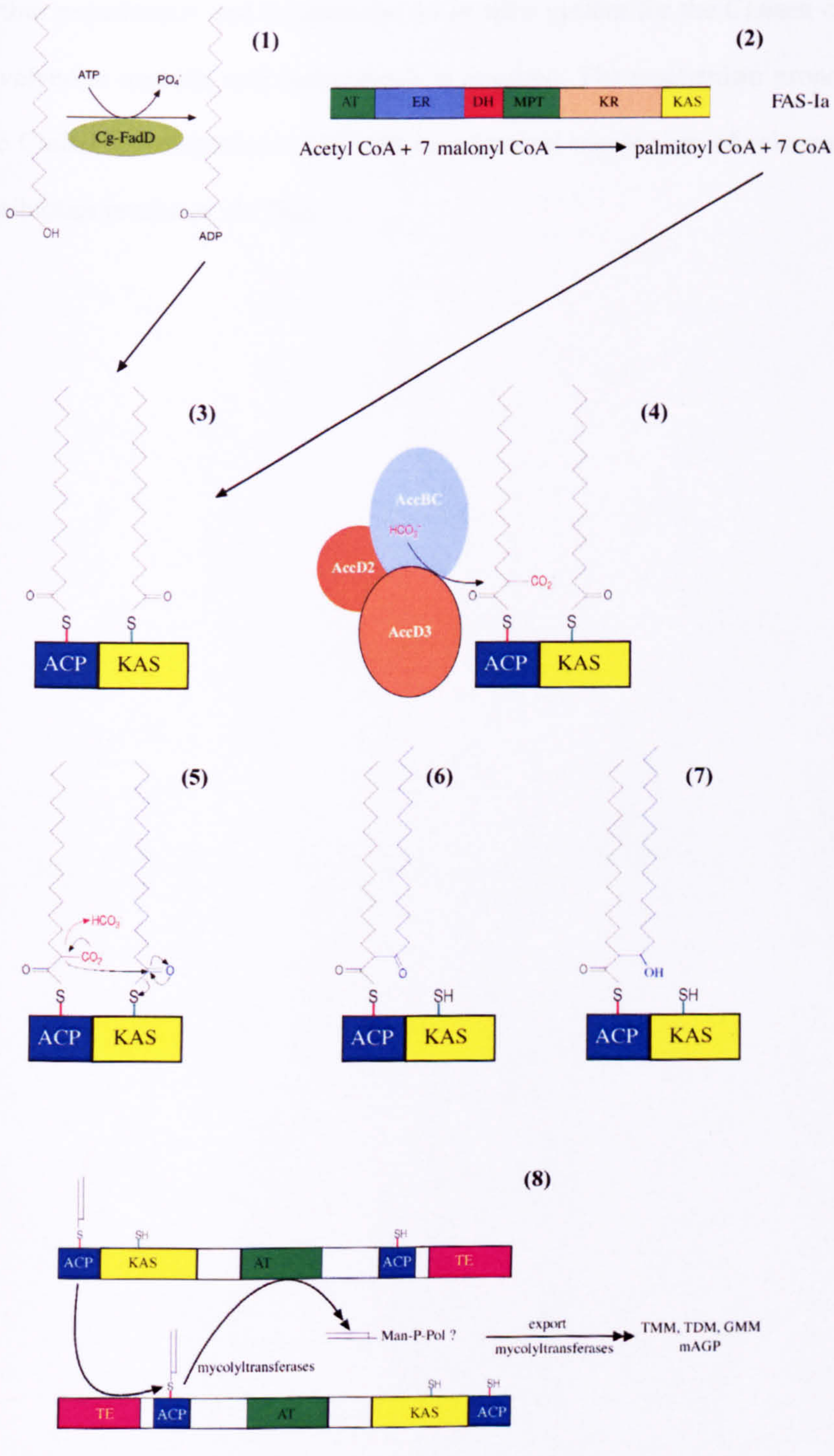
of the carboxylated products, which we presume, are readily decarboxylated in a Claisen-type condensation reaction. However, the assays revealed a profound decrease in carboxylation activity for the *accD2* and *accD3* inactivation mutants with respect to the C_{16} substrates. The data would appear wholly consistent with the proposed function of AccD2 and AccD3 in the later stages of corynomycolate biosynthesis. As expected, over-expression of *accD3* led to a similar increase in acyl-carboxylase activity with both C_{16} substrates, whereas when using acetyl-CoA, carboxylation activity was comparable to extracts from the parent strain. Intriguingly, over-expression of *accD2* led to an increase in the acyl-carboxylase activity with acetyl-CoA only. Maybe the peculiar substrate specificities implied by the above *in vitro* data indicate the formation of heterooligomeric assemblies whose composition defines substrate specificity. However, as both AccD2 and AccD3 are required for corynomycolate biosynthesis we can rule out any complementation by AccD1 and may speculate that AccD2 and AccD3 operate as a heterodimer. Furthermore, our *in vitro* analyses of *accD2* and *accD3* over-expression are highly suggestive that AccD3 might define the strict substrate specificity of such a complex. Sequence analysis of the AccD paralogues support this hypothesis as AccD2 is more closely related to AccD1 than AccD3.

The *M. tuberculosis* antigen 85 complex, which possess mycolyltransferase activity are located outside of the cytoplasm in the cell envelope and often found in the culture medium (Belisle *et al.*, 1997) as observed for *C. glutamicum* (Brand *et al.*, 2003; De Sousa-D'Auria *et al.*, 2003). The localisation of these enzymes suggests that they use mature mycolic acids once they are synthesised to transfer them to their final destination (Belisle *et al.*, 1997; Brand *et al.*, 2003; De Sousa-D'Auria *et al.*, 2003; McNeil *et al.*, 1991). The mycolyltransferase locus is slightly different in the related *Corynebacterium* species (Figure 4.7). CmytA in *C. glutamicum* is fused to a larger polypeptide with a domain of unknown function. In addition a

small open reading frame is present downstream (166 αα). A distinct possibility exists that *cmytA* and *cmytB* have different specificities as they are conserved in *Corynebacterianeae*. Interestingly, downstream of the mycolyltransferase locus is a gene essential for *M. tuberculosis* (Sassetti *et al.*, 2003). The Pro-rich protein encoded is predicted to be membrane anchored with the remaining segment directed towards the periplasm, and exhibiting identity with an acetylxylan esterase from *Penicillium purpurogenum* possessing the characteristic serine esterase motif. A tempting hypothesis is that this particular gene may be responsible for the first steps of mature mycolic acid transfer from a putative carrier molecule, such as ACP/CoA or from an intermediate polyprenol carrier (Besra *et al.*, 1994) to possibly either trehalose or to the cell wall arabinogalactan. This seems meaningful as it is located in close conjunction with the mycolyltransferases and together with the *fadD-pks-accD3* locus, and warrants further investigation as a possible specific mycolyltransferase. Another structural characteristic is the gene arrangement of *pks* and its downstream *accD3* gene. In the three *Corynebacterium* species analysed the genes are separated by less than 20 nucleotides, however both genes overlap by 4 nucleotides in *M. tuberculosis*, *M. bovis* and *M. leprae*. Such, intimate structural organisation has also been observed for other important genes in the *Corynebacterianea*, like for instance Ppm1, where even protein fusion has occurred (Baulard *et al.*, 2003; Gibson *et al.*, 2003; Gurcha *et al.*, 2002).

The β-ketoacyl synthases catalyse the formation of new carbon-carbon bonds by condensation of a variety of acyl-chain precursors with an elongation substrate, usually malonyl or methylmalonyl residues, that are covalently attached in a thioester linkage to an acyl carrier protein (ACP). The data obtained in this paper, along with the structural characteristics of the *fadD-pks-accD3* locus, and recent findings on FadD32 (Trivedi *et al.*, 2004) and *pks13* (Portevin *et al.*, 2004) yields three distinct features, which are very likely to represent the

basis of discrete steps in the mechanism of mycolic acid biosynthesis. (i) FadD, recently shown to constitute a new class of fatty acyl-AMP ligases, activates long-chain fatty acids as acyl-adenylates, which are loaded onto multifunctional PKS for further chain extension (Trivedi *et al.*, 2004), (ii) PKS possess a β-ketoacyl synthase domain with 2 phosphopantetheinyl binding sites and a thioesterase domain, and (iii) the carboxyltransferase activity of the Acc enzymes, with AccD2 and AccD3 in *C. glutamicum* directly linked to mycolic acid biosynthesis. An explanation for the two phosphopantetheinyl arms is that one site is occupied by the acyl-chain forming the mero-chain and the other by the acyl-chain resulting in the incoming α-branch. The role of AccD3 as shown in these studies would be to generate an activated carboxylated acyl-derivative which would then react with the bound acyl-chain to form a 3-oxo intermediate which would be reduced to form mature mycolic acids (Figure 4.9). The experimental data as well the phylogenomic analysis indicate that in addition to AccD3 in *C. glutamicum*, a second Acc protein (AccD2) is also apparently required for mycolic acid biosynthesis. One possibility is that after the fixation of carbon by AccBC, this is not directly transferred *via* AccD3 (AccD4 in *M. tuberculosis*) but transmitted *via* AccD2 (AccD5 in *M. tuberculosis*). The carboxylating transferase might fall into a group of multienzyme complexes, such as transcarboxylases from *Propionibacterium shermanii* (Hall *et al.*, 2003) which is a 1.2 MDa multienzyme complex that couples two carboxylation reactions, transferring CO_2^- from methylmalonyl-CoA to pyruvate, yielding propionyl-CoA and oxaloacetate. Cg-AccD3 also exhibits a high degree of identity with the 12S subunit of this enzyme. Shuttling *via* intermediates between different catalytic subunits provides a complex and intriguing mechanism for regulation and modulation of this specificity.


(1)
ATP
PO4
Cg-FadD
O
OH
O
ADP
(2)
AT
ER
DH
MPT
KR
KAS
FAS-Ia
Acetyl CoA + 7 malonyl CoA
palmitoyl CoA + 7 CoA
(3)
S
S
ACP
KAS
(4)
AccBC
HCO3-
AccD2
AccD3
CO2
(5)
HCO3
CO2
(6)
(7)
OH
SH
(8)
TE
Man-P-Pol ?
export
mycolyltransferases
TMM, TDM, GMM
mAGP

Clearly further experiments and in particular an *in vitro* system for the Claisen condensation reaction involved in mycolic acid biosynthesis is required. The mechanism proposed (Figure 4.9) for the Claisen condensation could well be a general mechanism of relevance for other lipids or antibiotics produced *via* PKS.

CHAPTER 5

Polyketide Synthase (Pks13)

5.1 Introduction

Polyketides form a large family of natural products found in bacteria, fungi and plants, which includes many clinically-important drugs such as tetracycline, daunorubicin, erythromycin (antibacterial), rapamycin (immunosuppressant) and lovastatin (anti-hypercholesterolemic). Their unparalleled range of biological activities drives current research into polyketide biosynthesis and enormous commercial value as promising candidates for future drug discovery. Also the structure, mechanism and catalytic reactivity of PKS synthases provide an opportunity to further investigate the molecular mechanisms of enzyme catalysis, molecular recognition and protein–protein interactions (Shen, 2003). Both PKS and FAS employ repeated Claisen-type condensations with acyl residues, typically derived from acetyl-CoA and malonyl-CoA. In contrast with fatty acid biosynthesis, the assembly of polyketides does not always employ a complete cycle (β-ketoacyl reduction, dehydration and enoyl reduction) after each chain extension (Child & Shoolingin-Jordan, 1998).

There are three types of bacterial PKS known. Type I and II PKS utilise ACP domains to activate acyl-CoA substrates and allow the binding and manipulation of polyketide intermediates, whereas type III PKS, are not ACP-dependent acting directly on the acyl-CoA substrates. Type I PKS are multifunctional enzymes that are organised into modules that are further sub-divided into active regions. An example of a type I PKS product is 6-deoxyerythromycin B synthases (DEBS) which is involved in the biosynthesis of reduced polyketides product, such as erythromycin A (Reeves *et al.*, 2001). Type II PKSs are multienzyme complexes that carry only a single module, for instance *Streptomyces glaucescens* tetracenomycin C synthesis. The genes *tcmJ/K/L/M/N* encode 5 modules of the tetracenomycin C PKS. These have been shown to be highly specific. In the absence of any

one module significantly diminished activity is observed, implicating formation of a multienzyme complex (Shen & Hutchinson, 1993). Type III PKS are also characterised as chalcone synthase-like PKS, they are homodimeric enzymes that are essentially condensing enzymes. Recent literature on polyketide biosynthesis suggests that PKS have much greater diversity in both mechanism and structure than the current type I, II and III paradigms (Shen, 2003). In some incidences type I PKS lack a domain which is essential for complete function like in the biosynthesis of leinamycin. The biosynthetic gene cluster from *Streptomyces atroolivaceus* S-140 lacks the acyl transferase (AT) domain, whose missing activity instead is provided in *trans* by a discrete protein, LnmG. This PKS has the majority of the type I PKS properties but also has type II characteristics therefore in some incidences there may be call for sub-typing to occur (Cheng *et al.*, 2003).

In *M. tuberculosis* several families of polyketide and non-ribosomal peptide synthases exist (Figure 5.1), for instance, *ppsABCDE* and *mas*, which are involved in the biosynthesis of PDIM, an important factor in pathogenicity (Betts *et al.*, 2002; Camacho *et al.*, 2001). Lipids derived from *in vivo* labeling experiments using [1-^{14}C]propionic acid showed that *pks2* mutant of *M. tuberculosis* was incapable of producing hepta- and octamethyl phthioceranic acids and hydroxyphthioceranic acids, the major acyl constituents of sulfolipids (Sirakova *et al.*, 2002; Sirakova *et al.*, 2001).

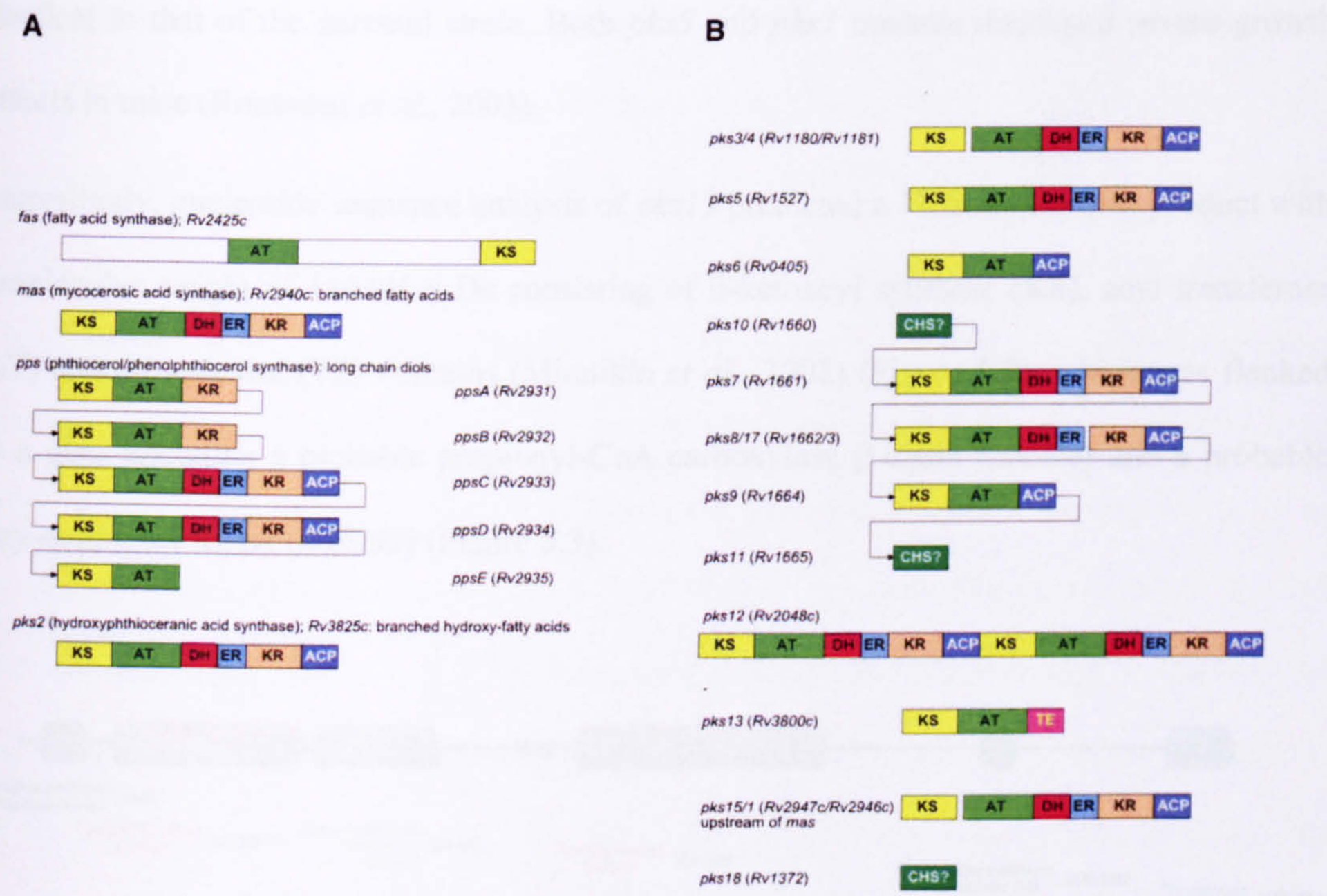


Figure 5.1 Domain organisation among *pks* families from *M. tuberculosis*. (A) Known *pks* genes (B) *pks* genes of unkown function. The domain organisation consists of the subunits β-ketoacyl synthase (KS), acyl transferase (AT), dehydratase (DH), enoyl reductase (ER), β-ketoreductase (KR), acyl carrier protein (ACP), thioesterase (TE), and chalcone synthase-like (CHS?).

Interestingly labelling studies also showed in both *pks2* and *pks12* mutants using [1-^{14}C]propionate that the mutants were deficient in the synthesis PDIM. Disruption of *msl5* (*pks8* plus *pks17*) in *M. tuberculosis* H37Rv generated a mutant incapable of producing monomethyl-branched C_{16}-C_{20} fatty acids that are minor constituents of acyltrehaloses and sulfolipids (Dubey *et al.*, 2003). Two further *mas*-like PKS genes *pks5* and *pks7* were disrupted in *M. tuberculosis*. Disruption of *pks7* resulted in a mutant deficient in PDIM production. However, the cell envelope composition of the *pks5* mutant was found to be

identical to that of the parental strain. Both *pks5* and *pks7* mutants displayed severe growth defects in mice (Rousseau *et al.*, 2003).

Interestingly, nucleotide sequence analysis of *pks13* predicted a 1733 amino acid product with a molecular weight of 186445.6 Da consisting of β-ketoacyl synthase (KS), acyl transferase (AT) and thioesterase (TE) domains (Minnikin *et al.*, 2002) (Figure 5.2), which was flanked by a gene encoding a probable propionyl-CoA carboxylase β chain (*accD4*) and a probable fatty-acid-CoA ligase (*fadD32*) (Figure 5.3).

Figure 5.2 Domain organisation of Pks13 of *M. tuberculosis*

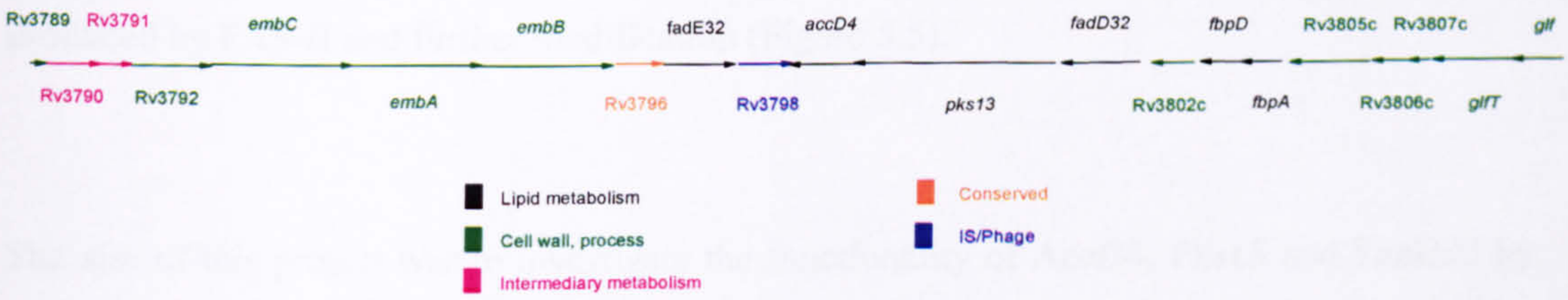


Figure 5.3 *pks13* region of *M. tuberculosis* H37Rv chromosome

The locus occupied by *pks13* lies in a 48.5 kb region in which several key genes involved in cell wall assembly and biosynthesis are located. The modular organisation and genetic linkage with *accD4* and *fadD32* suggest that Pks13 is involved in lipid biosynthesis and their location within this region suggests that Pks13 is involved in the biosynthesis of lipids of the cell wall.

The cluster of ORF, Rv3789-3792 are all involved in cellular metabolism and possibly involved in transmembrane functions. Alongside these ORFs lies the *embCAB* cluster, these genes are predicted to encode mycobacterial arabinosyltransferases involved in the synthesis of AG and LAM (Telenti *et al.*, 1997). FadE32 has been predicted to encode an acyl-CoA dehydrogenase. FbpA has been characterised and shown to encode a mycolyltransferase implicated in the deposition of mycolates allowing the synthesis of TMM and TDM cord factor (Kremer *et al.*, 2002a). FbpD has recently been shown to be inactive in mycolyltransferase activity (Kremer *et al.*, 2002a). *glfT* has been shown to encode a novel UDP-galactofuranosyltransferase. This enzyme possesses dual functionality in performing both (1→5) and (1→6) galactofuranosyltransferase reactions implicating its involvement in AG biosynthesis (Kremer *et al.*, 2001b). The clustering of functionally related genes within genomes isn't uncommon therefore hypothetically it is possible that the cluster of *fadD32* to *accD4* is involved in cell wall biosynthesis. We propose that *pks13* is involved the Claisen condensation of the meromycolate chain to the α-chain forming the α-alkyl β-hydroxy fatty acid (mycolic acids) (Figure 5.4) from the fatty acids from FAS-I and the meromycolates produced by FAS-II and further modification (Figure 5.5).

The aim of this project was to investigate the functionality of AccD4, Pks13 and FabD32 by further characterisation of the enzymes and elucidate their function as proposed in Figure 5.4. Although it is not clear how the covalently sequestered biosynthetic intermediates are transferred from one enzymatic complex to another it was assumed that it would be either be by an ACP or CoA adduct.

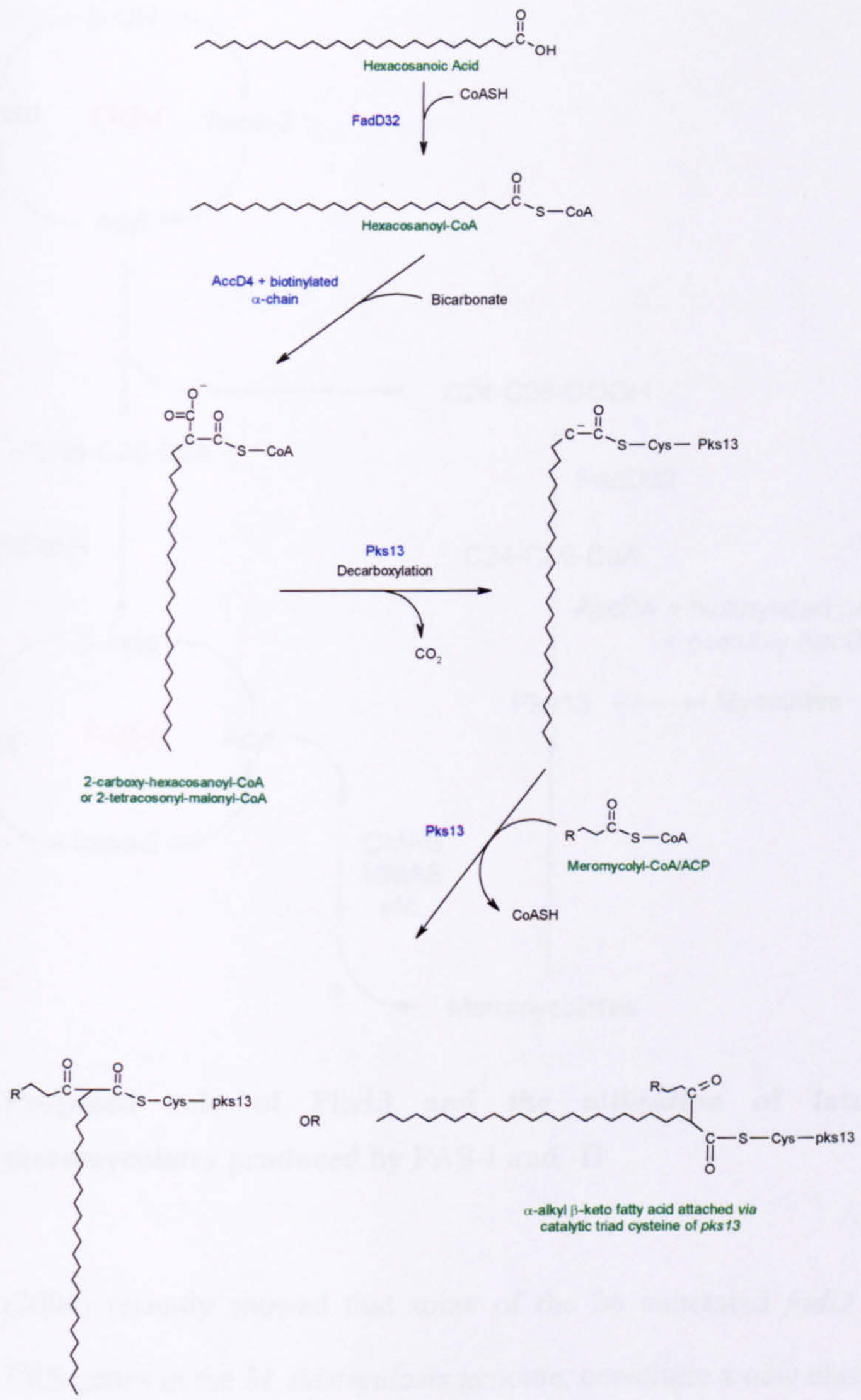


Figure 5.4 Schematic representation of the proposed mycolyl condensation activity of Pks13

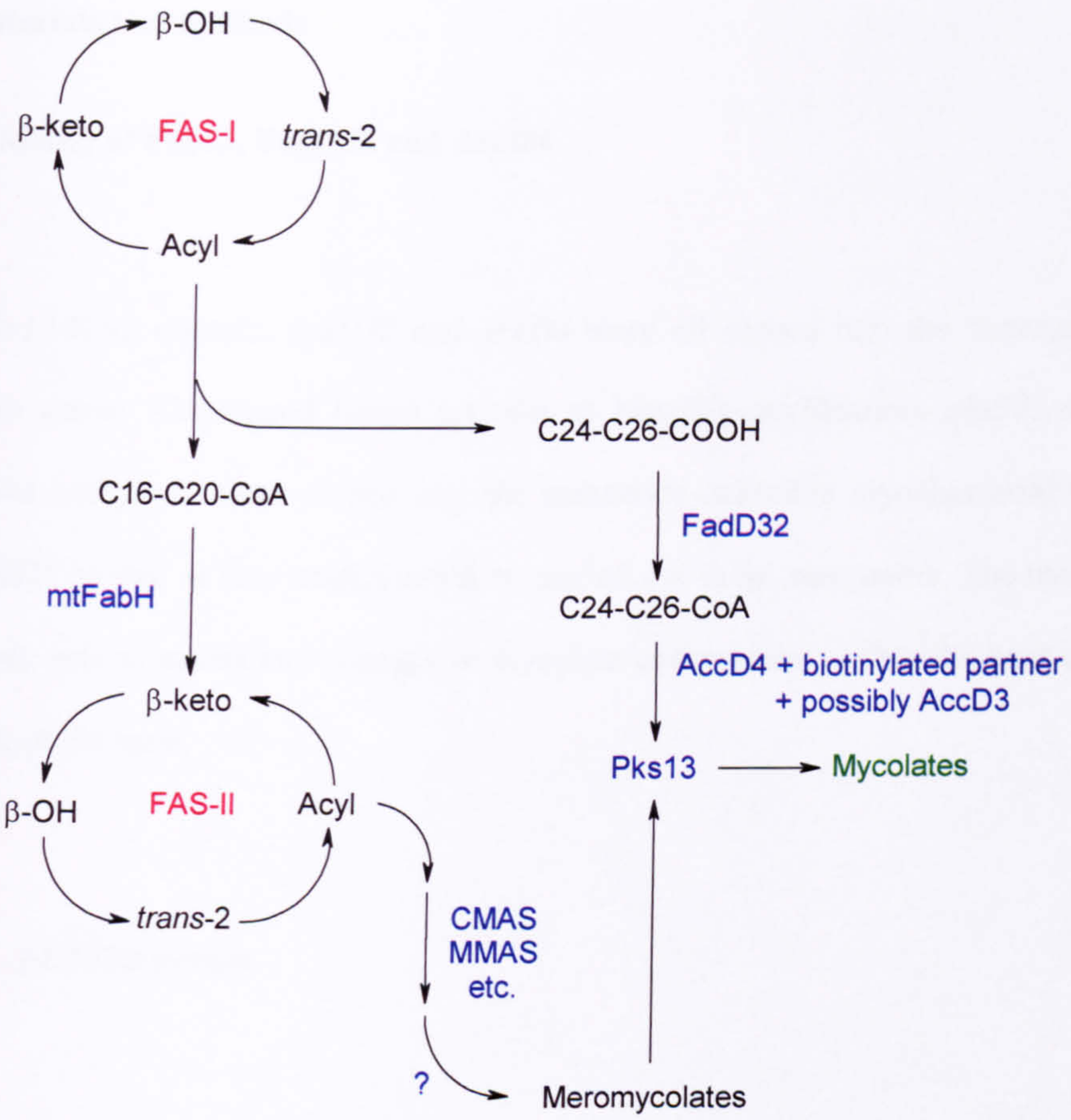


Figure 5.5 Proposed role of Pks13 and the utilisation of fatty acids and meromycolates produced by FAS-I and -II

Trivedi *et al.* (2004) recently showed that some of the 36 annotated *fadD* genes, located adjacent to the PKS genes in the *M. tuberculosis* genome, constitute a new class of long-chain fatty acyl-AMP ligases. These proteins activate long-chain fatty acids as acyl-adenylates, which are then transferred to the multifunctional PKSs for further chain extension. This mode of activation and transfer of fatty acids is contrary to the previously described universal mechanism involving the formation of acyl-CoA thioesters (Trivedi *et al.*, 2004). The identification of the acyl-carrier was essential as studies on Pks13 using acyl-CoA and AcpM substrates might indicate negative results.

5.2 Materials and methods

5.2.1 Cloning of Pks13, FadD32 and AccD4

pks13, *pks13*-KAS domain, *fadD32* and *accD4* were all cloned into the expression vector pET23b to create His_6-tagged fusion proteins to simplify purification. *pks13*, *pks13*-KAS domain and anti-*pks13* were cloned into the acetamide-inducible mycobacterial expression vector pSD26 so that *in vivo* studies could be carried out in *M. smegmatis*. The main purpose of the study was to assess any changes in mycolate composition within the host upon over-expression of the gene.

5.2.1.1 pET23b system

Pks13 and the KAS domain were cloned using the recipe in Table 5.1. The conditions for the PCR reaction are found in Table 5.2. For the KAS domain the normal upper primer was used but a different lower primer was used. The DNA encoding *pks13* was amplified from *M. tuberculosis* H37Rv genomic DNA. PCR primers upper (5'-GATCGATCCATATGGCTGACGTAGCGGATTC-3') and Full lower (5'-GACTGATCAAGCTTCTGCTTGCCTACCTCACTTG- 3') *kas* lower (5'-GATCGATCAAGCTTGGCGGGTTCGGCGGCCGCC-3') were used to create *NdeI* and *HindIII* restriction sites (underlined), respectively. The PCR product was separated from primers and nucleotides by electrophoresis on 1 % agarose, and after visualisation with EtBr, fragments consistent with the size of the gene were extracted using the Qiagen Gel Extraction Kit. The PCR product was then digested with *NdeI* and *HindIII* and yield was assessed by DNA electrophoresis, along side a similar digest of pET23b. The cut DNA fragment was

again extracted from the gel, and ligated with the cut vector and transformed into *E. coli* Top Ten. Plasmid DNA was isolated from transformants showing ampcillin resistance. Screening was performed by double digest with and analysis on 1 % agarose and visualisation by staining with EtBr. The integrity of the DNA construct was confirmed by nucleotide sequencing.

Table 5.1 PCR recipe

Volume	
2 µl	5′ Cloning Primer (100 pmol/ml)
2 µl	3′ Cloning Primer (100 pmol/ml)
1 µl	Genomic DNA (1 µg/µl)
2 µl	dNTP's (25 mM each type)
10 µl	Thermostable Buffer (10x)
8 µl	DMSO
2 µl	$MgSO_4$ (100 mM)
1 µl	Deep Vent ™ DNA polymerase (2000 U/ml)
72 µl	ddH_2O
Final volume of 100 µl with the addition of ddH_2O	

Table 5.2 PCR conditions for Pks13

Full Pks13			KAS Domain		
94°C	3 min		94°C	3 min	
94°C	30 sec	} 35 cycles	94°C	30 sec	} 35 cycles
68°C	10 min		68°C	3 min	
68°C	10 min		68°C	10 min	
4°C	∞		4°C	∞	

accD4 and *fadD32* were cloned by the same method using the recipe in Table 5.1 and the following primers, under the PCR conditions in Table 5.3. Both DNA fragments were

separated from primers and nucleotides by electrophoresis on 1 % agarose, and after staining with EtBr, relevant bands excised from the gel and extracted using the Qiagen Gel Extraction Kit. The PCR fragment was then digested with both restriction enzymes incorporated in the primer and assessed by DNA electrophoresis, along side a similar digest of pET23b. The cut DNA fragment was again extracted from the gel, and ligated with the cut vector and transformed into *E. coli* Top Ten. Screening was performed by double digest with and analysis on 1 % agarose. The DNA was then sequenced to assess the viability of the construct.

AccD4	Upper	5' GATCGATCCATATGACCGTCACCGAGCCGG 3'
	Lower	5' GATCGATCAAGCTTCTAGACCGGGATCAGGCCG 3'
FadD32	Upper	5' GATCGATCCATATGTTTGTGACAGGAGAGAGTGGG 3'
	Lower	5' GATCGATCAAGCTTGTCCGAAGTGGCGAAGACCGTC 3'

Table 5.3 PCR conditions for AccD4 and FadD32

AccD4			FadD32		
94°C	3 min		94°C	3 min	
94°C	30 sec	35 cycles	94°C	30 sec	35 cycles
68°C	4 min		68°C	3 min	
68°C	10 min		68°C	10 min	
4°C	∞		4°C	∞	

5.2.1.2 pSD26 system

Pks13 and KAS domain were cloned using the recipe and conditions as for the pET23b system Table 5.1 and Table 5.2. For the KAS domain the normal upper primer was used but with a different lower primer. The *pks13* gene was amplified from *M. tuberculosis* H37Rv genomic DNA using the following primers to clone the genes of interest and create *BamHI* restriction sites upper (5'-GATCGATCGGATCCATGGCTGACGTAGCGGATTC-3') and

lower (5'-GACTGATCGGATCCCTGCTTGCCTACCTCACTTG- 3') Kas lower (5'-GATCGATCGGATCCGGCGGGTTCGGCGGCCGCC-3'). The PCR product was separated from primers and nucleotides by electrophoresis on 1 % agarose, and after visualisation with EtBr, was extracted using the Qiagen Gel Extraction Kit. The PCR fragment was then digested with *BamHI* and assessed by DNA electropheresis, alongside a similar digest of pSD26. The cut DNA fragment was again extracted from the gel, and ligated into the cut vector and transformed into *E. coli* Top Ten. Screening was performed using an orientation PCR method. Normal PCR conditions, an orientation primer (5'-ATGCCCGAGGTAGTTTTTATCCATGG-3') and the lower primer were used to access the plasmid. The pSD26-anti*pks13* was created *via* the normal cloning regime but with the gene ligated in the incorrect orientation. The DNA was then sequenced to verify the construct.

5.2.2 Culturing and purification

An induction protocol yielding good quantities of soluble protein was sought by varying the concentration of IPTG and the induction temperature used. *E. coli* C41 (DE3) expression host were transformed with the pET23b-*pks13* and pET23b-*pks13*KAS constructs. The cells were cultured at 37°C to an $A_{600\ nm}$ = 0.6 and then recombinant gene expression was induced with varying concentrations of IPTG (0.1 mM, 0.5 mM and 1.0 mM). The cultures were grown for a further 4 h at 37°C. Another culture was cooled to 16°C before 1mM IPTG was added and was grown for a further 16 h at 16°C and harvested by centrifugation. The cell pellets were resuspended in buffer (20 mM phosphate buffer, pH 7.9, 0.5 M NaCl, 10 mM imidazole) containing DNase and RNase. Bacteria were disrupted by sonication and the resulting lysate centrifuged at 27,000 x *g* for 60 min at 4°C (Table 5.4). Samples were analysed by

electrophoresis on 8 % SDS-PAGE followed by staining with Coomassie brilliant blue staining.

Table 5.4 Induction Protocol for pks13

IPTG (mM)	Time (hours)	Temp (°C)
0.1	4	37
0.5	4	37
1.0	4	37
1.0	16	16

Solubility of the recombinant protein was assessed by comparing the amount of the protein seen in the crude and clarified lysates. If the protein was soluble then purification was performed. Hi-trap Chelating Sepharose™ fast flow matrix columns (1 ml or 5 ml) were washed with 4 column volumes of Hi-Trap running buffer (0.02 M Sodium Phosphate, 0.5 M NaCl pH7.4). The column was charged using 2 column volumes of 0.1 M nickel chloride. The column was then washed with 4 column volumes of running buffer to equilibriate the column and remove excess nickel. The clarified lysate was applied to the column (maximum of 10 ml for a 1 ml column and 50 ml for a 5 ml column) and non-absorbed material collected for analysis. Absorbed proteins were eluted using varying concentrations of imidazole (10 mM – 500 mM). Fractions were collected and analysed by SDS-PAGE following protein level assessment. Pure protein samples were dialysed against large volumes (100-200 fraction volumes) of dialysis buffer (20 mM Tris-HCl pH 7.6, 50 mM NaCl, 1 mM β-mercaptoethanol)

5.2.3 FAMEs and MAMEs analysis

For the expression study *M. smegmatis* mc^2155 were transformed by electroporation with the pSD26-based constructs and cultured to mid-log phase in 50 ml of Sauton's medium containing 50 μg/ml hygromycin, 50 μCi of [1,2-^{14}C]acetate (50–62 mCi/mmol, Amersham Pharmacia Biotech) was added to the cultures followed by acetamide at various concentrations, then incubated at 37°C for 4 hrs.

Investigation of the effects of TLM on the production of fatty acid methyl esters (FAMEs) and mycolic acid methyl esters (MAMEs) was performed using *M. smegmatis* mc^2155 transformed by electroporation with pSD26 and the pSD26-*pks13* construct. As before, they were cultured to mid-log phase in 50 ml of Sauton's medium containing 50 μg/ml hygromycin, the cultures were induced with 0.2 % acetamide followed by incubation at 37°C for 4 hrs. 50 μCi of [1,2-^{14}C]acetate (50–62 mCi/mmol, Amersham Pharmacia Biotech) was added along with various concentrations of TLM, followed by a further incubation at 37°C for 4 hrs.

The [^{14}C]-labelled cells were harvested by centrifugation and washed successively with 0.9 % aqueous NaCl and water. The [^{14}C]-labelled cells were then subjected to alkaline hydrolysis using 15 % aqueous tetrabutylammonium hydroxide at 100°C overnight. 4 ml of CH_2Cl_2, 300 μl of CH_3I and 2 ml of water was added to the reaction and mixed for a further 30 min. The upper aqueous phase was discarded, and the lower, organic phase was washed twice with water and evaporated to dryness. Methyl esters were re-dissolved in diethyl-ether, and then the solution evaporated to dryness. The final residue was dissolved in 200 μl of CH_2Cl_2 and

the incorporation [1,2-^{14}C]acetate was quantified by liquid scintillation counting of a 5 µl of the resulting solution of FAMEs and MAMEs. Approximately 100,000 cpm of the FAMEs/MAMEs mix were subjected to TLC using silica gel plates (5735 silica gel 60F254; Merck, Darmstadt, Germany), developed in petroleum ether-acetone (95:5). Autoradiograms were produced by overnight exposure of Kodak X-Omat AR film to reveal [^{14}C]-labelled fatty acid and mycolic acid methyl esters.

5.2.4 Growth Curve and induction of Pks13 in *M. smegmatis*

Investigation of the effects of Pks13 on *M. smegmatis* growth was analysed along with the expression of the plasmid. *M. smegmatis* mc^{2}155 transformed by electroporation with the empty pSD26 vector and the pSD26-*pks13* construct. They were cultured to OD_{600} = 0.2 in 100 ml of Sauton's medium containing 50 µg/ml hygromycin, the cultures were induced with 0.2 % acetamide followed by incubation at 37°C. Regular OD readings were taken along with 1 ml samples every 2-4 hrs. Samples were harvested by centrifugation and resuspensed in 50 mM MOPS pH 7.9, 50 mM MgCl, 5 mM β-mercaptoethanol. Cells were lysed by sonication and debris removed by centrifugation (27,000 x *g*). Clarified lysate were visualised on 10 % SDS-PAGE.

5.2.5 Minimum inhibition concentration analysis (MIC)

The MIC of TLM was performed using *M. smegmatis* mc^{2}155 transformed by electroporation with the pSD26 and the pSD26-*pks13* construct. Cultures were grown to OD_{600} = 0.4 in 50 ml of Sauton's medium containing 50 µg/ml hygromycin, they were induced with 0.2 %

acetamide followed by further incubation at 37°C for 4 hrs. Serial dilutions ($10^{-1} \rightarrow 10^{-6}$) of the cultures were plated on LB agar containing 50 μg/ml hygromycin, with and without 0.2 % acetamide and increasing concentrations of TLM followed by incubation at 37°C for 3-4 days.

5.2.6 Lipid Analysis

5.2.6.1 Culturing and labelling

Investigation of the effects of TLM on lipid production was performed using *M. smegmatis* mc^2155 transformed by electroporation with the pSD26 and the pSD26-*pks13* construct. 4 cultures (2x pSD26, 2x pSD26-*pks13*) were grown to mid-log phase in 50 ml of Sauton's medium containing 50 μg/ml hygromycin, induced with 0.2 % acetamide followed by incubation at 37°C for 4 hrs. 50 μCi of [1,2-^{14}C]acetate (50–62 mCi/mmol, Amersham Pharmacia Biotech) was added to each culture along with 15 μg/ml of TLM, followed by a further incubation at 37°C for 4 hrs. Radiolabelled cells were harvested by centrifugation.

5.2.6.2 Extraction

Cells were transferred to glass PTFE-capped tubes and resuspended in 2 ml MeOH-0.3 % NaCl (100:10 v/v) and 1 ml petroleum ether (high boiling point), followed by mixing on a rotator for 15 minutes. The cellular suspension was centrifuged at 3,500 x *g* for 5 minutes. The upper petroleum ether layer was removed and a further 1 ml of petroleum ether added, followed by a further mixing on the rotator for 15 minutes. Again the petroleum ether was removed and pooled with the previous extraction and evaporated under nitrogen using a

heathing block (<37°C). This extract known as the "non-polar lipids" was dried and resuspended in 200 µl dichloromethane prior to use. 20 µl of the non-polar lipids were dried in scintillation vials and counted after the addition of 10 ml of EcoscintA with 25,000 cpm of the sample run on TLC using systems A-D. The methanolic-saline extract (lower layer) was heat treated in a boiling-water bath for 30 minutes and allowed to cool to room temperature. chloroform-methanol-0.3 % NaCl (2.3 ml; 9:10:3) was added, followed by mixing for 60 minutes. The solvent extract and biomass were separated by centrifugation at 3,500 x *g* and retained. A further 2 extractions were performed using chloroform-methanol-0.3 % NaCl (750 µl; 5:10:4) followed by mixing for 30 minutes, and centrifugation at 3,500 x *g* with extracts combined. The combined extracts were then mixed with chloroform (1.3 ml) and 0.3 % NaCl (1.3 ml) followed by mixing for 5 minutes. The fraction was centrifuged at 3,500 x *g* for 10 minutes to separate the lower-organic layer and the upper aqueous layer discarded. The lower layer was evaporated to dryness yielding the "polar lipids" which were resuspended in 200 µl of chloroform/methanol (2:1) prior to analysis. The polar lipids (20 µl) were dried in scintillation vials and counted after the addition of 10 ml of EcoscintA, with 25,000 cpm of the sample was run on each TLC using systems D and E.

5.2.6.3 Analysis of the non-polar and polar lipids by TLC

The non-polar and polar lipid fractions are further analysed by two-dimensional TLC using the developing systems in Table 4.5, using aluminium backed silica gel plates (5735 silica gel 60F254; Merck, Darmstadt, Germany). Autoradiograms were produced by 2-3 day exposure of Kodak X-Omat AR film to reveal [^{14}C]-labelled lipids.

Table 5.5 Polar and non-polar TLC developing systems

System	Direction 1	No. Runs
A	Petroleum ether/ethyl acetate (98:2)	3
B	Petroleum ether/acetone (92:8)	3
C	Chloroform/methanol (96:4)	1
D	Chloroform/methanol/water (100:14:0.8)	1
E	Chloroform/methanol/water (60:30:6)	1

System	Direction 2	No. Runs
A	Petroleum ether/acetone (98:2)	1
B	Toluene/acetone (80:20)	1
C	Toluene/acetone (95:5)	1
D	Chloroform/acetone/methanol/water (50:60:2.5:3)	1
E	Chloroform/acetic acid (glacial)/methanol/water (40:25:3:6)	1

5.2.7 Construction of a *pks13* knockout mutant in *M. smegmatis*

Regions of DNA flanking *pks13* corresponding to 1 kb upstream and downstream of the ORF were amplified from the *M. smegmatis* mc^2155 genome using the recipe (Table 5.1) but using *M. smegmatis* mc^2155 genomic DNA and conditions in Table 5.6 along with the following cloning primers;

Upstream 5' primer (5'-GATCGATCACTAGTTGTGCCGAAGCCGGGCTCAC-3')
SpeI

Upstream 3' primer (5'-GATCGATCGATATCATGTGCGCAGCCACTCCCGC-3')
EcoRV

Downstream 5' primer (5'-GATCGATCTCTAGAAGGCGCTCAACCGCATCGAGGCTC-3')
XbaI

Downstream 3' primer (5'-GATCGATCCTTAAGTCGCGCCCGACAACACCAT-3')
AflII

The PCR products were cloned into *SmaI*-cut pUC18. The recombinant plasmids were then digested *SpeI / EcoRV* for the upstream and *XbaI / AflII* for the downstream and the *pks13* flanking DNA were purified by extraction from agarose gels. The uptstream region was first ligated into *SpeI / EcoRV* cut pJSC347 KO vector and transformed into *E. coli* HB101.

Screening by restriction digestion with *SpeI / EcoRV* was used to identify constructs containing the upstream fragment. The construct was then digested with *XbaI / AflII* to allow the ligation of the downstream fragment. Screening of plasmid DNA extracted from transformants of *E. coli* HB101 derived from the ligation products by two separate double digests yielded the 1 kb upstream (*SpeI / EcoRV*) and the 1 kb downstream fragments (*XbaI / AflII*).

Table 5.6 PCR conditions for *pks13* knockout fragments

1 kb Fragments		
94°C	3 min	
94°C	30 sec	35 cycles
68°C	1 min 20 sec	
68°C	10 min	
4°C	∞	

As it was possibile that a *pks13* KO might be lethal, pSD26-*pks13* construct could be used as an inducible survival vector. The KO pJSC347 vector was sent to Dr Aproova Bhatt (Albert Einstein Institute, New York). Who then packaged the vector into phage and prepared high titre of phage suspensions. Cultures of *M. smegmatis* and *M. smegmatis* pSD26-*pks13* were grown until OD_{600nm} = 0.8-1.0. The pellet was washed twice with MP buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 10 mM $MgSO_4$, 2 mM $CaCl_2$). The cells were then resuspended in 0.5 ml of MP buffer and sterile culturing tubes followed by addition of 250 µl of the phage (10^{10} PFU/ml), the suspension was incubated at 37°C for 3 hours, 3 ml of 7H9 media (7H9, OADC, Tween) was added to the cells follwed by growth for 3 hours at 37°C. The cells were then harvested and washed once in 10 ml of 7H9 media before resuspension in 0.5 ml. The culture was spread on plates (7H10, OADC, Hyg 75 µg/ml) with and without acetamide (0.2 %) and Tween (0.5 %). Screening for KO-*pks13* was performed by PCR using the 5' upstream

primer and a 3' hygromycin cassette primer. The PCR should yield a fragment size of approximately 1.5 kb with screening confirmed by southern blotting as described in the materials and methods (Appendices).

5.3 Results

5.3.1 Cloning and expression of Pks13, AccD4 and FadD32

5.3.1.1 pET23b system

The cloning and preparation of the gene or gene fragment required for the construction of the pET23b construct allowed for the utilisation of the C-terminal His-tag aiding purification of recombinant protein *via* the Ni^{+2} Hi-trap system.

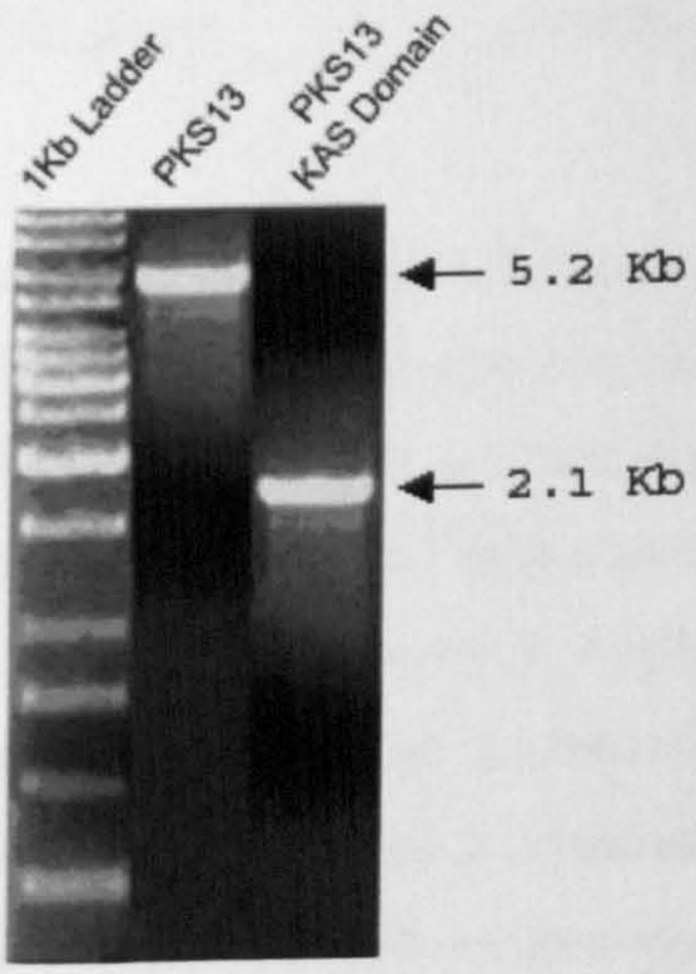


Figure 5.6 DNA electrophoresis (1 % agarose) of pET23b *pks13* PCR product.

pET23b constructs of *pks13* and *pks13*KAS domain were created using PCR products from H37Rv template DNA in specific PCR reactions (Figure 5.6) followed by digestion and ligation into the empty vector. Screening was performed by double digest followed by electrophoresis of the cut plasmid on 1 % agarose and staining with EtBr.

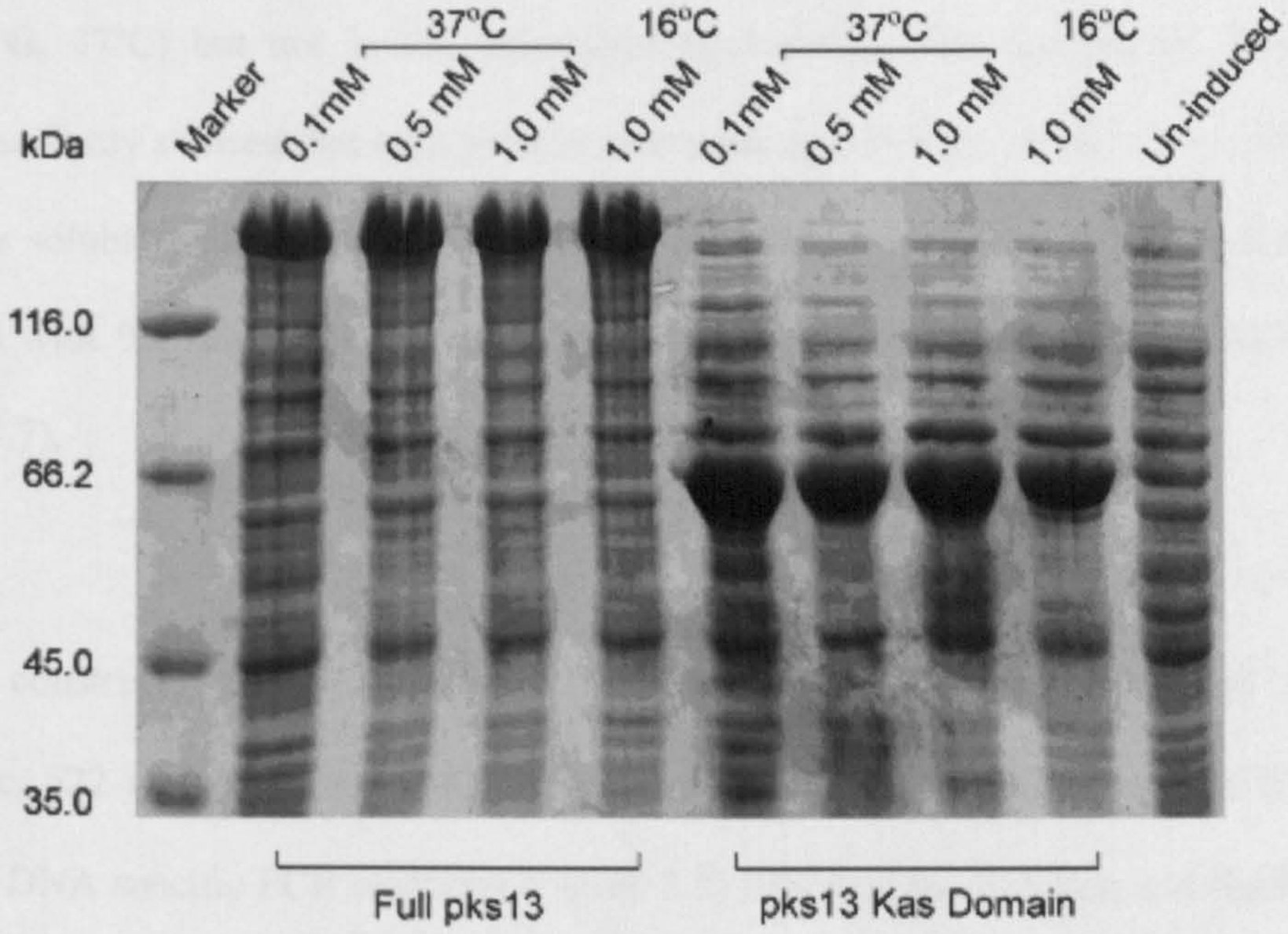


Figure 5.7 **Solubility and expression analysis of Pks13 and Pks13-KAS by 8% SDS-PAGE. *E. coli* C41 (DE3) transformed with the pET23b constructs were culture in 50 ml LB broth to OD_{600} = 0.5. 3 cultures induced with varying IPTG concentrations (0.1 mM, 0.5 mM and 1.0 mM) whereas another was cooled to 16°C and induced with 1 mM IPTG for 16 h. A further uninduced culture was used as a control. Cells were harvested by centrifugation and resuspended in 3 volumes of 20 mM sodium phosphate pH7.9, 500 mM NaCl. The cell suspension was lysed using a French Pressure cell and cell debris removed by centrifugation at 27,000 x *g*. Equal amounts of protein were added to the gels after protein concentration determination using a BCA assay kit (Pierce)**

E. coli C41 (DE3) were transformed with the pET23b constructs and analysed for the expression and solubility. Pks13 and Pks13KAS were analysed by varying conditions of induction as stated in section 5.2.2. Analysis of extracts on 12 % SDS-PAGE demonstrated novel polypeptides in the order of ~180 kDa and 60 kDa in molecular weight consistent with the predicted molecular weight of Pks13 and Pks13KAS were observed in induced cultures (1 mM IPTG, 37°C) but not in the uninduced equivalents. The analysis of the expression conditions firstly showed that both proteins are expressed in high yields and secondly that the protein is soluble under all the conditions of expression tested. The optimal conditions were observed with 0.1 mM IPTG for both the full Pks13 protein and the Pks13-KAS domain (Figure 5.7).

pET23b constructs of *fadD32* (Rv3801c, 637 amino acids, 69232.0 KDa) and *accD4* (Rv3799c, 522 amino acids, 56647.7 Kda) were created using PCR products from H37Rv template DNA specific PCR reactions (Figure 5.8) followed by digestion and ligation into the empty vector. Screening was performed by double digest followed by electrophoresis of the cut plasmid on 1 % agarose and staining with EtBr.

Figure 5.8 DNA electrophoresis of pET23b *accD4* and *fadD32* PCR product.

An expression study and extensive analysis of the solubility of both proteins was carried out as before, but FadD32 was insoluble under all conditions tested. AccD4 was expressed well with high yields being observed in the crude lysate but clarification of the crude lysate by centrifugation showed that the majority of the protein was insoluble when induced at 37°C for 4 hours. The solubility problem was overcome by induction with 1 mM IPTG at 16°C for 16 hours (Figure 4.10). The crude lysate contain large yields of the recombinant protein and after clarification soluble protein was still visible.

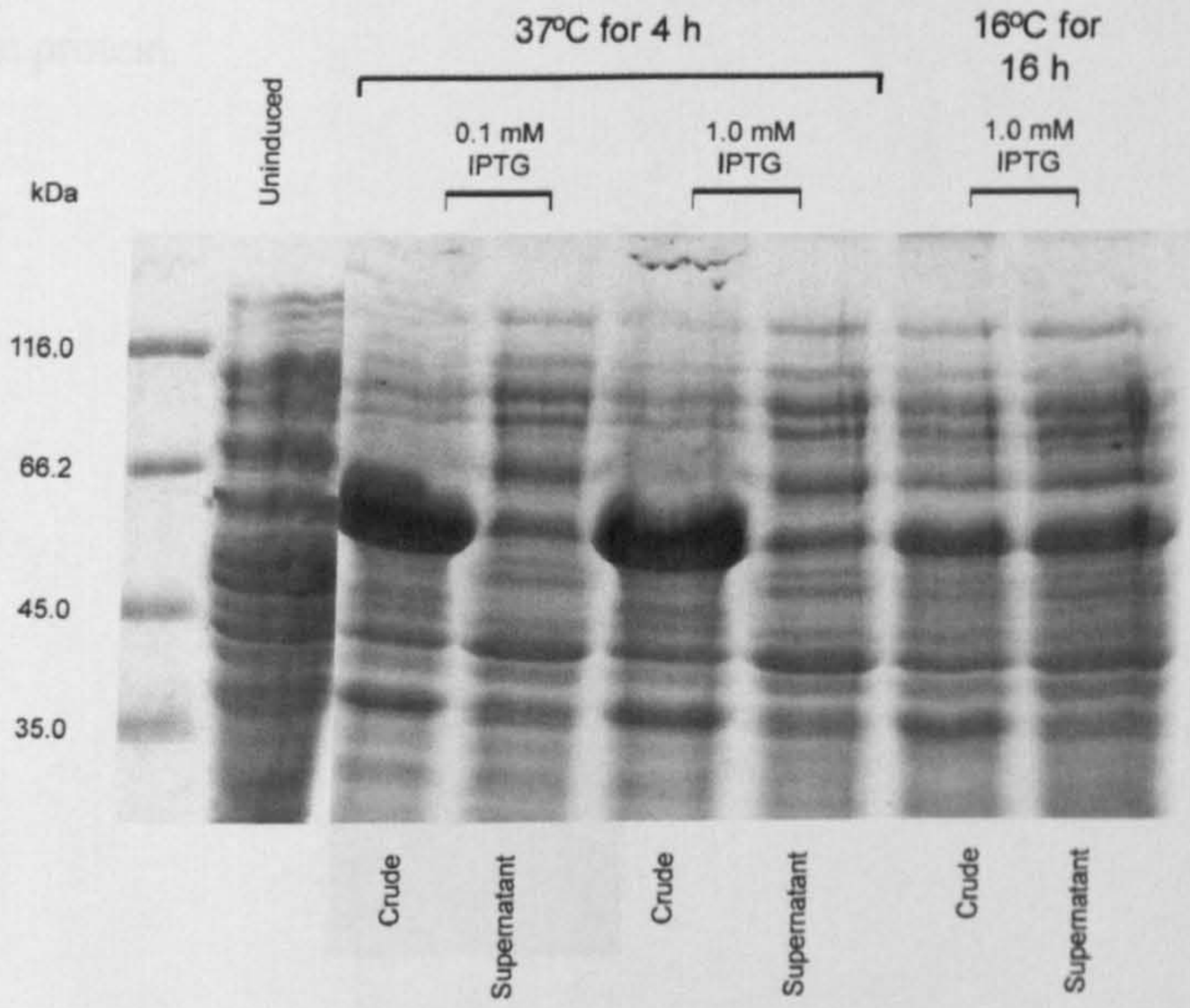


Figure 5.9 **AccD4 expression study. *E. coli* C41 (DE3) transformed with the pET23b-*accD4* construct were culture in 50 ml LB broth to OD_{600} = 0.5. 3 cultures induced with varying IPTG concentrations (0.1 mM and 1.0 mM) whereas another was cooled to 16°C and induced with 1 mM IPTG for 16 h. A further uninduced culture was used as a control. Cells were harvested by centrifugation and resuspended in 3 volumes of 20 mM sodium phosphate pH7.9, 500 mM NaCl. The cell suspensions were lysed by French press and samples taken (crude) followed by removal of the cell debris by centrifugation at 27,000 x *g* (Supernatant). Equal volumes of protein were added to the gels after protein concentration determination using a BCA assay kit (Pierce)**

5.3.1.2 pSD26 system

The successful cloning and preparation of the gene or gene fragment required for the production of the pSD26 expression systems was necessary for the over-expression of the *M. tuberculosis* protein in *M. smegmatis.* The pSD26 expression system also contains a C-terminal His-tag which would aid purification *via* the Ni^{+2} Hi-trap system and is also based on an acetamide inducible promoter for use in *M. smegmatis* allowing for the controlled expression of the protein.

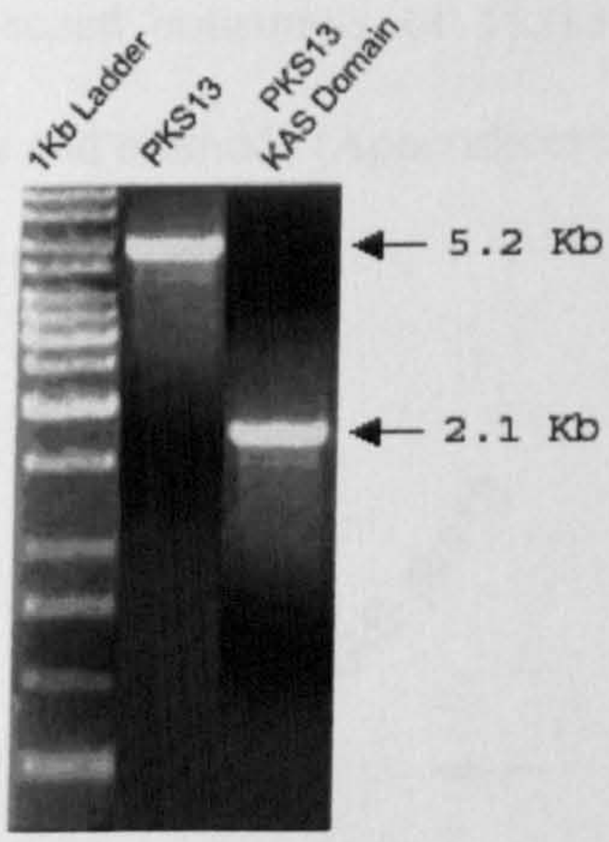


Figure 5.10 DNA electrophoresis (1 % agarose) of pSD26 *pks13* PCR products.

pSD26 constructs of *pks13* and *pks13*KAS domain were created using PCR products from H37Rv template DNA in specific PCR reactions (Figure 5.10) followed by digestion and ligation into the empty vector. Screening was performed by double digest followed by electrophoresis of the cut plasmid on 1 % agarose and staining with EtBr.

M. smegmatis mc^2155 were transformed with the fully sequenced vectors pSD26-*pks13*, pSD26-anti-*pks13*, pSD26-Kas-*pks13* and empty pSD26 for *in vivo* analysis. No significant change in the colony morphology was observed when the gene expression was induced with 0.2 % acetamide on 7H10 agar containing Hyg^{50}.

5.3.2 Purification

Recombinant proteins were synthesised for purification using transformants of *E. coli* C41 (DE3) containing the pET23b-based constructs of Pks13 and Kas-Pks13. Proteins were purified as described in materials and methods (Appendices). SDS-PAGE was used to analyse the protein to assess their purity (Figure 5.11).

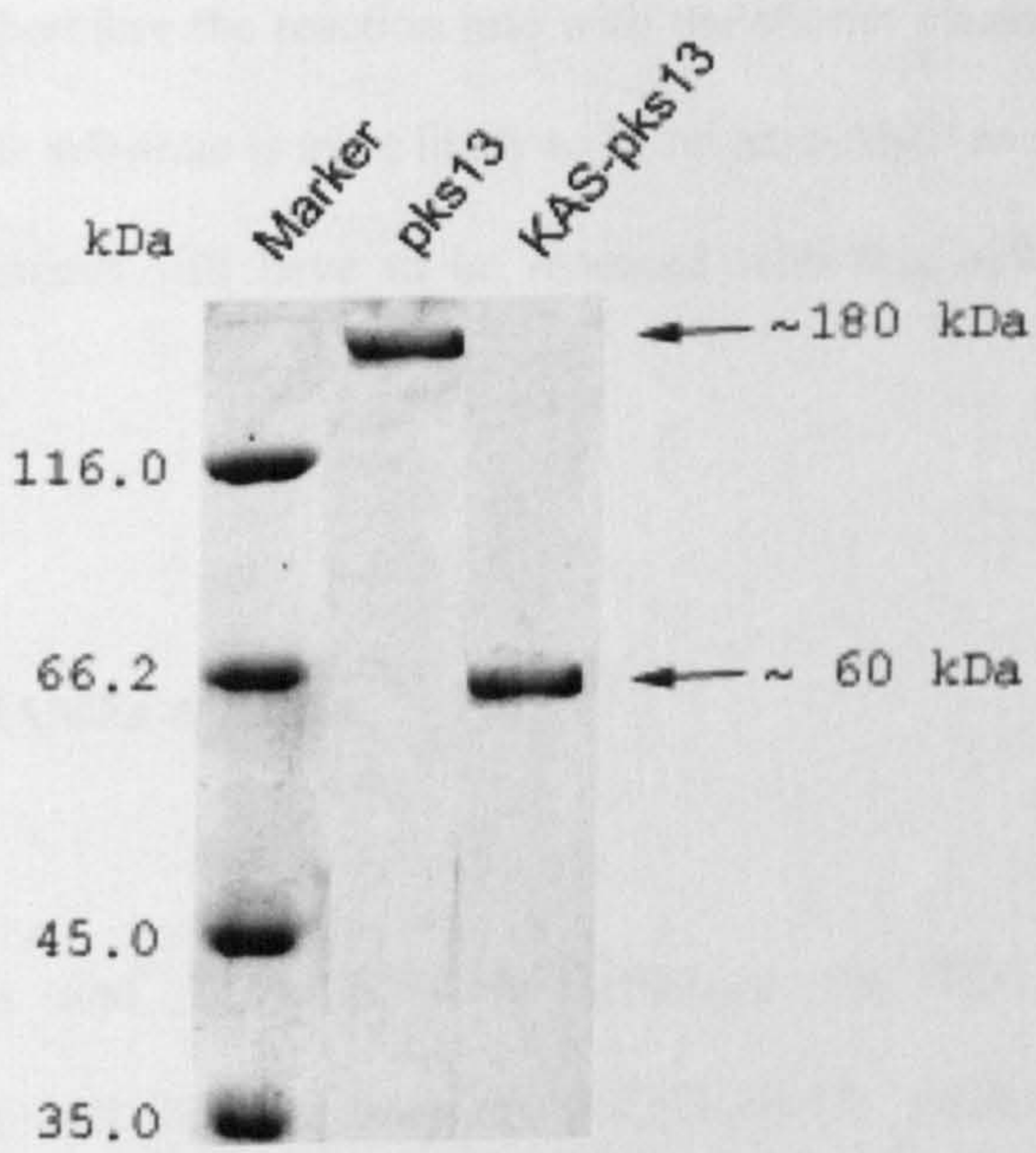


Figure 5.11 Purification of *pks13* and *Kas-pks13* by Ni2+ chelating sepharose column (1 ml) and visualised by 8 % SDS-PAGE

Recombinant protein with molecular weights of 180 kDa and 60 kDa consistent with the predicted molecular weigth of Pks13 and Pks13KAS were purified. The production of this recombinant protein would allow for *in vitro* biochemical analysis, but first the production of the substrates *via* FadD32 and AccD4 are required.

A single polypeptide of ~57 kDa consistent with the predicted molecular weight of AccD4 was visualised on 12 % SDS-PAGE after the clarified lysate was subjected to liquid chromatography using Ni^{2+} chelating His-trap columns. This procedure yielded approximately 3.0 mg of pure AccD. Purified proteins were dialysed 3 times for 4 hours against 100 volumes of 20 mM Tris-HCl, 50 mM NaCl, pH 7.5 then flash frozen in liquid nitrogen and stored at -80°C. Attempts at radiobiochemical assays involving pure AccD4, palmitoyl-CoA and radio-labelled bicarbonate proved to be ineffective. AccD4 is thought to utilise long-chain acyl-CoA's (C_{24}-C_{26}) and therefore the reaction rate with the shorter chains may be undetectable. As previously stated the substrate is most likely to be an acyl-AMP as shown in a recent study therefore these experiments will have to be repeated with this newly identified substrate (Trivedi *et al.*, 2004).

5.3.3 FAMEs and MAMEs Analysis

Radiolabelled FAMEs and MAMES were separated by TLC and visualised by autoradiography using cell lines expressing pSD26-*pks13*, pSD26-anti-*pks13*, pSD26-*pks13*KAS and pSD26 to assess *in vivo* overproduction of Pks13 on FAMES and MAMES. After expression and extraction (section 5.2.3) the FAMEs and MAMEs TLC showed that little or no change when the plasmids were over-expressed (Figure 5.12). In cells bearing the

antisense construct it was hypothesised that this mRNA produced by the expression of the antisense vector would bind the mRNA of the *M. smegmatis* and therefore cease Pks13 production changing the profile of the FAMEs and MAMEs, unfortunately no change was observed on the TLC. The control samples showed that the vector had no effect on lipid synthesis upon induction. Hypothetically the over-expression of *pks13* should increase the production of mycolates if indeed involved in the Claisen condensation. The study of the over-expression of the *pks13* had little effect on the composition of the MAMEs, as did the over-expression of the KAS domain therefore this isn't a rate limiting step in mycolate biosynthesis.

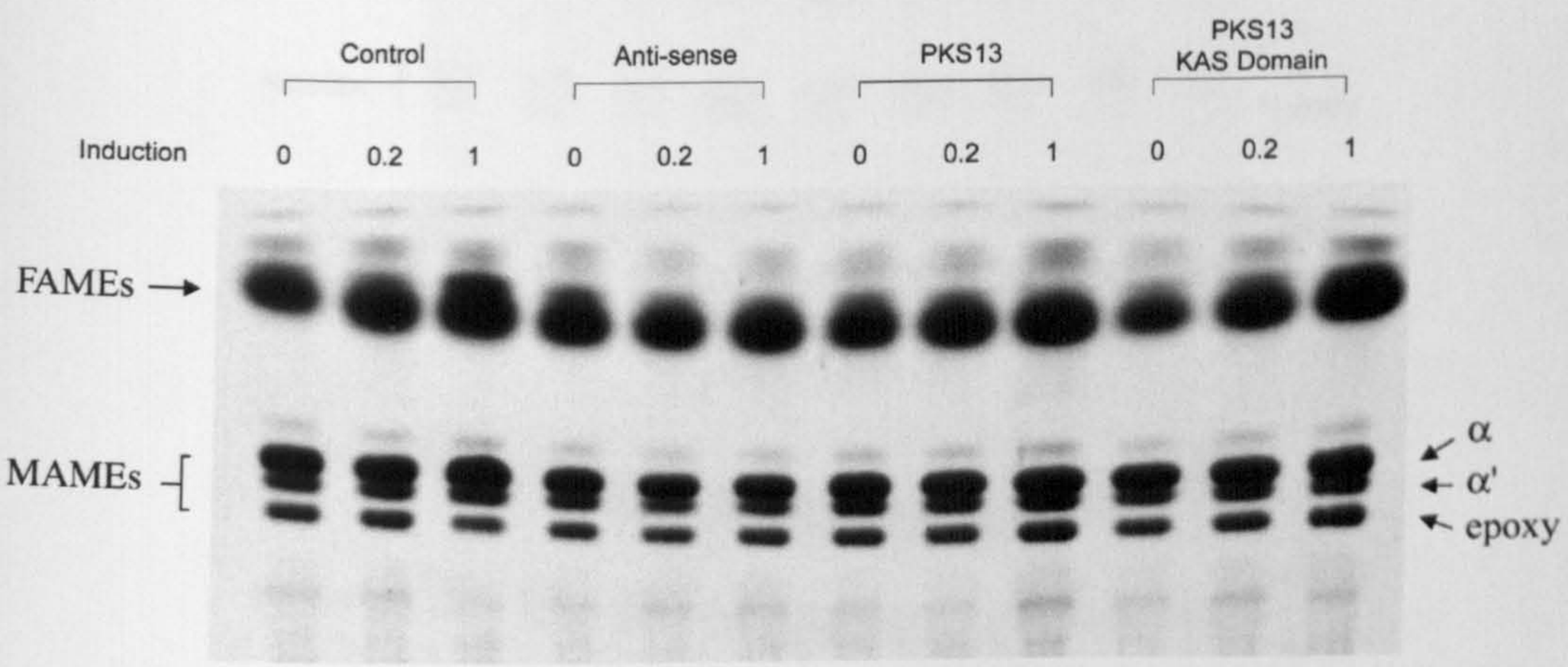


Figure 5.12 **The effect of over-expression of Pks13 constructs in *M. smegmatis* visualised by the incorporation of [^{14}C]acetate into lipids to produce FAMEs and MAMEs in the presence of increasing concentrations of acetamide. The reaction was terminated by the addition of 15 % tetrabutylammonium hydroxide at 100°C overnight. The corresponding FAMEs, α-mycolates, α'-mycolates and epoxy-mycolates (MAMEs) were isolated, subjected to TLC (5735 silica gel 60F_{254}, Merck), and exposed to a Kodak X-Omat film.**

Further analysis of the FAMEs and MAMEs composition of *M. smegmatis* extracts was assessed in the presence of TLM as it was theorised that Pks13 over-expression would overcome the effects of TLM by increasing the mycolate composition of the cell wall thus making the wall less permeable to TLM. Radiolabelled FAMEs and MAMES were separated by TLC and visualised by autoradiography using cell lines expressing pSD26-*pks13* and pSD26 to assess *in vivo* effects of Pks13 on FAMES and MAMES and resistance to TLM. After expression and extraction (section 5.2.3) the TLC was exposed to film and developed.

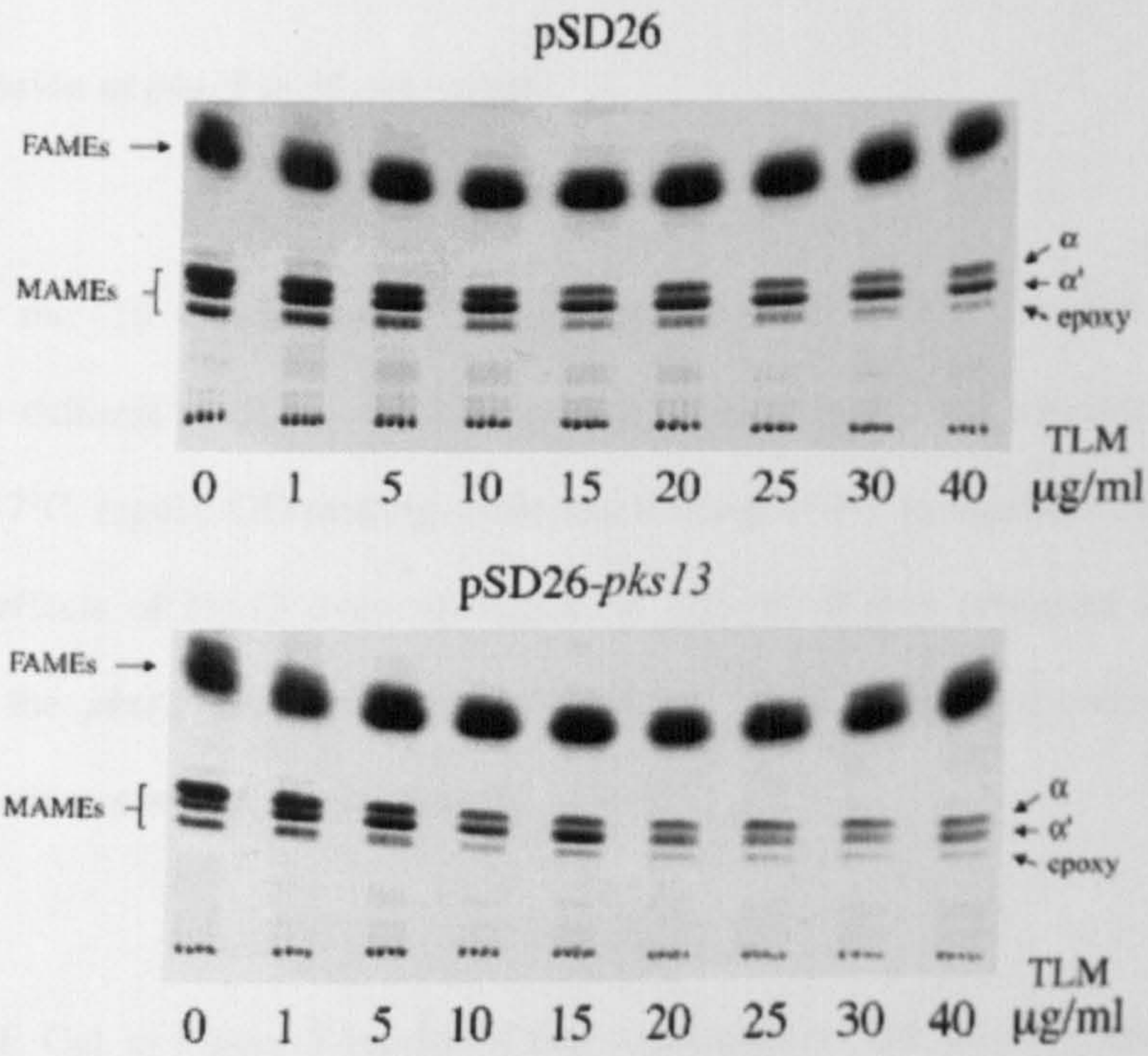


Figure 5.13 The effect of over-expression of Pks13 construct in *M. smegmatis* in the presence of increasing concentrations of TLM. Followed by visualisation by the incorporation of [^{14}C]acetate into lipids to produce FAMEs and MAMEs. The reaction was terminated by the addition of 15 % tetrabutylammonium hydroxide at 100°C overnight. The corresponding FAMEs, α-mycolates, α'-mycolates and epoxy-mycolates (MAMEs) were isolated, subjected to TLC (5735 silica gel 60F_{254}, Merck), and exposed to a Kodak X-Omat film.

Interestingly, the overproduction of Pks13 in the presence of TLM changed the lipid profile more significantly than the WT strain. The MAMEs profile of the WT is as previously observed with increasing concentration of TLM (Kremer *et al.*, 2000b) with less epoxy- and α-mycolates being observed as the TLM concentration increases. The over-expression of Pks13 reduced the overall MAMEs observed with increasing concentration of TLM, resulting in the over-expression strain becoming more sensitive to mycolate inhibition. This would suggest that the over-expression of Pks13 increases the sensitivity of *M. smegmatis* to TLM.

5.3.4 Expression of *pks13* in *M. smegmatis*

M. smegmatis mc^2155 transformed by electroporation with pSD26 and the pSD26-*pks13* construct were cultured to OD_{600} = 0.2. Upon induction with 0.2 % acetamide followed by incubation at 37°C, regular OD readings were taken along with 1 ml samples every 2-4 hours to assess the effects of Pks13 over-expression on growth. It was proposed that the over-expression of the *pks13* gene from *M. tuberculosis* in *M. smegmatis* may increase the synthesis of mycolates and accelerate growth.

The SDS-PAGE Gel in Figure 5.14 shows that expression of the pSD26-*pks13* within *M. smegmatis* throughout the time course experiment produced large quantities of protein.

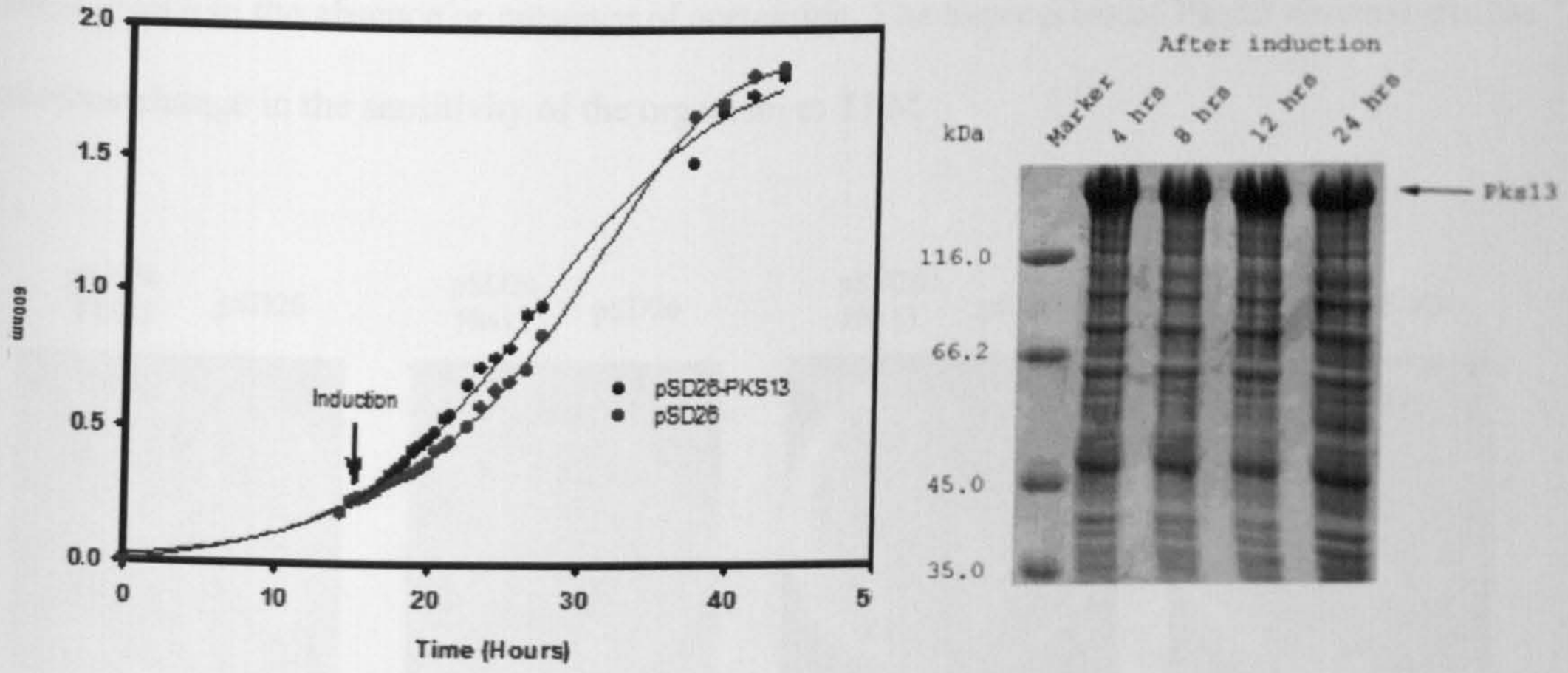


Figure 5.14 **Growth curve showing the effect of Pks13 on *M. smegmatis* growth. *M. smegmatis* mc^{2}155 transformed by electroporation with pSD26 and the pSD26-*pks13* construct were cultured to OD_{600} = 0.2 in Sauton's liquid medium containing 50 μg/ml hygromycin, the cultures were induced with 0.2 % acetamide followed by incubation at 37°C**

5.3.5 TLM MIC analysis of *M. smegmatis* over-expressing *pks13*

The MIC of TLM against *M. smegmatis* mc^{2}155 transformed by electroporation with pSD26 and pSD26-*pks13* was performed to assess the effect of over production of Pks13 with regards to TLM resistance/sensitivity. Cultures were grown as *per* the section 4.2.5 and plated on LB agar containing 50 μg/ml hygromycin, with and without 0.2 % acetamide and increasing concentrations of TLM followed by incubation at 37ºC for 3-4 days. It was hypothesised that the increased expression of Pks13 would infact increase the resistance of *M. smegmatis* to TLM by increasing the production of mycolates. Figure 5.15 shows that the *M. smegmatis* transformed with either vector were able to grow on plates containing up to 30 μg/ml TLM

when grown in the absence or presence of acetamide. The expression of Pks13 showed that no obvious change in the sensitivity of the organism to TLM.

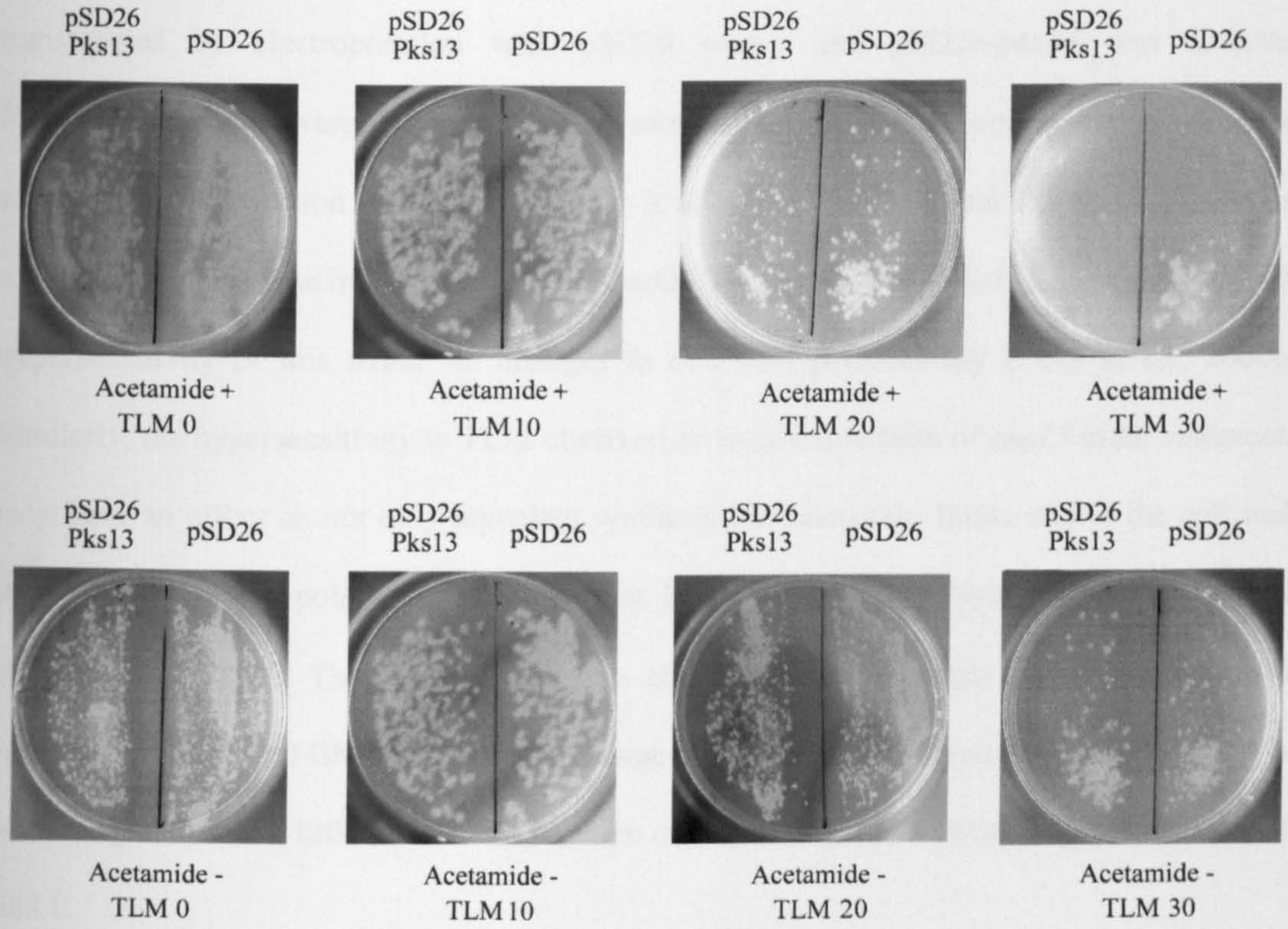


Figure 5.15 MIC of TLM using *M. smegmatis* mc²155 transformed by electroporation with pSD26 and pSD26-*pks13*. Cultures were grown to OD_{600} = 0.4 in 50 ml of Sauton's medium containing 50 μg/ml hygromycin, they were induced with 0.2 % acetamide followed by further incubation at 37°C for 4 hrs. Serial dilutions (results shown are the 10^{-3} dilution) of the cultures were plated on LB agar containing 50 μg/ml hygromycin, with and without 0.2 % acetamide and increasing concentrations of TLM followed by incubation at 37°C for 3-4 days.

5.3.6 Lipid analysis of *M. smegmatis* over-expressing Pks13 in the presence of TLM

Investigation of the effects of TLM on overall lipid production using *M. smegmatis* mc^2155 transformed by electroporation with pSD26 vector and pSD26-*pks13* was assessed. Radiolabelled cells were produced as *per* section 4.2.6.1 and non-polar and polar lipids extracted as *per* section 4.2.6.1. Previously it was hypothesized that the up regulation of multiacylated trehalose in *M. bovis* BCG on *mtfabH* over-expression led to the observed TLM hypersensitivity of this strain *via* changes in cell wall permeability (Choi *et al.*, 2000b). Similarly, the hypersensitivity to TLM observed on over-expression of *pks13* in *M. smegmatis* may have an effect on not only mycolate synthesis but also other lipids within the cell wall. Analysis of the non-polar lipids showed that little change in the lipid profiles upon over-expression of Pks13. The system D profiles changed in the presence of TLM showing an increase in the overall GMM but with no change being observed between the two strains. Also with the polar lipids, little or no changes were observed on TLC's developed with systems D and E.

5.3.7 Deletion of *pks13* in *M. smegmatis*

The over-expression of *pks13* which resulted in the reduction on the overall MAMEs observed in the presence of increasing concentration of TLM resulted in the over-expression strain becoming more sensitive to mycolate inhibition. This indicates that futher experiments especially a knockout of the gene would prove more useful in understanding the role Pks13 plays in the production of mycolates. Regions of DNA flanking *pks13* corresponding to 1 kb upstream and downstream of the ORF were amplified from the *M. smegmatis* mc^2155 genome

DNA and cloned into the pJSC347 vector, screening was performed by double digest followed by electrophoresis on 1 % agarose and visualisation by staining with EtBr (Figure 5.16).

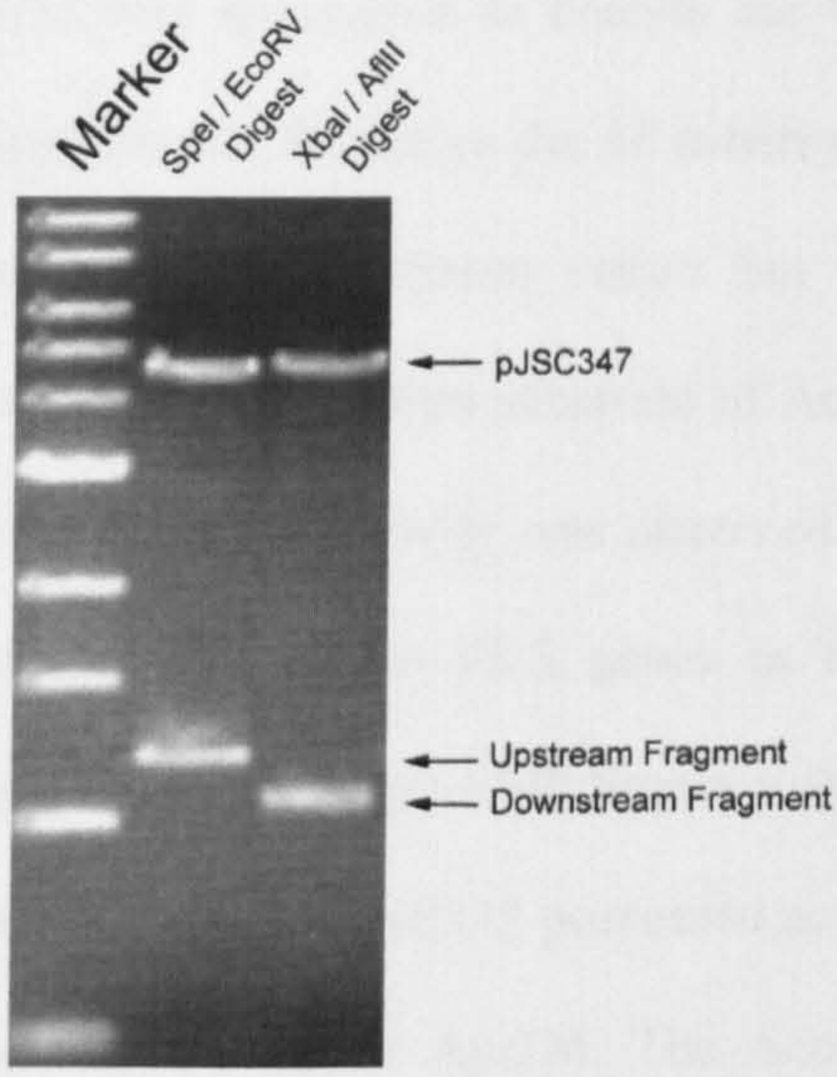


Figure 5.16 Double digested pJSC347-*pks13*KO. *SpeI / EcoRV* digest yielding the upstream fragment and *XbaI / AflII* for the downstream. The upstream fragment was slightly larger than the downstream fragment as the region cloned was required to contain a larger proportion of the upstream gene so as not to inactivate not only the Pks13 but also the AccD4.

The construct was successfully packaged into the phage but as of yet any screening of possible knockouts has not been achieved.

5.4 Discussion

The alignment study of the *M. tuberculosis pks13* region and the *Corynebacterium diphtheriae pks13* homologue region showed a highly conserved region (Dover *et al.*, 2004). In the previous chapter *pks13* was speculated to encode the Claisen condensation enzyme required for mycolic acid biosynthesis. Therefore the *M. tuberculosis* genes in this study were successfully cloned into the relevant expression vector but the insolubility problems of FadD32 halted the *in vitro* testing. The prefered substrate of AccD4 was proposed to be C_{24}-C_{26}-CoA and may be the reason why no activity was observed with shorter acyl-CoAs. The observation that FadD's located adjacent to PKS genes in the *M. tuberculosis* genome, constituted a new class of long-chain fatty acyl-AMP ligases was timely (Trivedi *et al.*, 2004). More specifically, they demonstrated that FadD32 possessed acyl-AMP substrate specificity, which would suggest this as the carrier for AccD4. The AccD4 *in vitro* assay could not proceed further without FadD32 and the acyl-AMP's.

PKS present an exciting area for the controlled manipulation of complex natural product structure. In particular, regioselective modification of polyketide structure is possible by alterations in either intracellular acyl-CoA pools or, more commonly, by manipulation of acyl transferases that act as the primary gatekeepers for building blocks (Liou & Khosla, 2003). A computational protocol has been used to predict the domain organisation and substrate specificity for PKS clusters from various microbial genomes. A searchable computerized database (PKSDB) was produced that would serve as a valuable tool for identification of polyketide products biosynthesised by uncharacterised PKS clusters. This database can also provide guidelines for rational design of experiments to engineer novel polyketides (Yadav *et al.*, 2003).

A mutant strain of *M. smegmatis* defective in the biosynthesis of mycolic acids was recently isolated (Liu & Nikaido, 1999). This mutant failed to synthesise full-length mycolic acids and accumulated a series of long chain β-hydroxymeromycolates. Wang *et al.* (2000) further characterised this mutant and revealed that the meromycolates were covalenty attached to AG at the 5-OH position of the terminal arabinofuranosyl residues, therefore demonstrating that mutants devoid of mycolic acids could be produced. However, the permeability barrier of this mutant was significantly compromised (Liu & Nikaido, 1999). The attempt at down-regulating Pks13 activity using an 'antisense' transcript of *M. tuberculosis pks13* in *M. smegmatis* in this study did not produce similar changes to lipid profiles as seen with the mutant strain. The experiment carried out here was not optimised; ideally an antisense transcript of the native gene should be used to maximise homology and promote the formation of a duplex with the *pks13* mRNA *in vivo*. Furthermore, such constructs are often more effective if the antisense construct contains the reverse complement of the ribosome binding site. Despite these inadequacies, the experiment was warranted given that a mis-oriented construct was discovered during recombinant plasmid screening.

The alternative strategy, in which Pks13-was overproduced *in vivo,* resulted in no noticeable changes in the mycolic acid or lipid compositions of *M. smegmatis* mc^2155 but this is an acceptable result as the recombinant protein may only facilitate mycolic acid production but is tightly regulated by limiting substrates. The increased sensitivity of *M. smegmatis* mc^2155 to TLM in the presence of over expressed Pks13 could in turn be explained by the same theory. Over-production of Pks13 may bind up small quantities of either the meromycolate or the α-branch subunits which would not be produced in vast quantities therefore slowing mycolate production enough to increase the sensitivity of the organism to TLM. As with MMAS and CMAS over-production studies, the results could differ from the KO studies therefore for true

investigation of the *M. tuberculosis* Pks13's involvement in mycolic acid biosynthesis a disruption or KO mutation is required. As stated previously it is believed that Pks13 is essential therefore a survival vector is required.

As an exciting and new area of work all the basic biochemical and physiological properties of *pks13* are still to be established. With the production of *M. bovis* BCG and *M. smegmatis* KOs, *in vivo* investigation into Pks13 can take place, ranging from FAMES and MAMES analysis to more complex mycolate separation and characterisation. Determination of the crystal structure will in turn, allow us to compute possible binding sites and therefore identify amino acid residues which are essential for catalysis and can be studied by site-directed mutagenesis. Finally antibiotic sensitivity can be assessed and new drug analogues can be designed from the structural and biochemical data achieved from the work. Pks13 looks like a very exciting target for new drugs, as its possible role in the production of mycolates is key to cell survival and as no mycobacteria have been shown to survive without their mycolates other than the mutant mentioned above, therefore it's potentially an excellent one.

CHAPTER 6

Summary

The mycolic acids are the dominant feature of the *Mycobacterium tuberculosis* cell wall, providing the basis for its lipid-rich permeability barrier. Throughout this thesis a number of fundamentally important proteins involved in the biosynthesis of these characteristic α-alkyl, β-hydroxy fatty acids have been analysed to allow a greater understanding of their properties and functions. Mycolic acids are thought to be formed by the Claisen-type condensation of a long meromycolic acid and a shorter C_{24}-C_{26} fatty acid. These component fatty acids are produced *via* a combination of type I and II fatty acid synthase (FAS) systems. The C_{16}-C_{24} fatty acyl products of FAS-I are elongated by FAS-II to form meromycolic acids. The β-ketoacyl ACP synthase (KAS) mtFabH links FAS-I and FAS-II, catalyzing the condensation of acyl-CoA derivatives of FAS-I products with malonyl-AcpM. Two discrete KAS enzymes, which are encoded by *kasA* and *kasB*, located in tandem within a five-gene operon, perform subsequent decarboxylation condensation within FAS-II. Finally, enzymology of the mycolic condensation, in which the meromycolate and C_{24}-C_{26} fatty acid are ligated, was addressed.

Investigation of the acyl-CoA chain length specificity of mtFabH *in vitro* using either *E. coli* ACP (ecACP) or the *M. tuberculosis* AcpM revealed altered substrate specificity. It was observed that the acyl-CoA chain length specificity of mtFabH was similar to that previously described by Choi *et al.* (2000) when ecACP was utilised in the *in vitro* assay. However with *M. tuberculosis* AcpM, which bears a 33 residue carboxy terminal extension relative to ecACP, the enzyme was able to utilise longer acyl-CoA chains. The recent elucidation of the crystal structure of mtFabH facilitated a detailed study of the roles of key amino acid residues in catalysis (Scarsdale *et al.* 2001). Selected residues thought to be involved in acyl-CoA chain length specificity, substrate binding and residues proposed to represent the catalytic triad of the enzyme were mutated. Substitution of residues implicated in acyl-CoA chain length specificity totally abrogated overall KAS activity. Alanine substitution of residues

implicated in acyl-CoA binding led to a reduction in the transacylation activity of the enzyme; a double alanine substitution of Trp42 and Arg161 drastically reduced transacylation activity. On investigation of the proposed catalytic triad residues, Cys122 was confirmed to be an essential residue for transacylation and His258 was essential for malonyl-AcpM decarboxylation, whilst mutation of Asn289 resulting in reduced decarboxylation activity.

In this study, evidence was provided that demonstrated *kasA* from *M. tuberculosis* to encode an enzyme that elongates the meromycolate chain of both *M. smegmatis* and *M. chelonae*, *in vivo*. Investigation of KasA substrate specificity identified the enzyme as a key component of the FAS-II system, which utilises primarily palmitoyl-ACP rather than short-chain acyl-ACP primers. An *in vitro* condensing assay using purified recombinant KasA, palmitoyl-AcpM and malonyl-AcpM, demonstrated that KasA unequivocally possess KAS activity in the presence of the ACP thioesters and is significantly less active using acyl-CoA precursors. Purified KasA was highly sensitive to cerulenin, a well-known inhibitor of KAS activity, which may lead to the development of novel anti-mycobacterial drugs targeting KasA.

Alanine substitutions of key amino acids of KasA allowed the functional analysis of these residues importance in terms of KAS activity of KasA. The mutation of Cys171, His311, Lys340 and His345 completely abrogated the condensation activity of KasA *in vitro*, suggesting that these residues play a critical role in the condensation reaction. Furthermore, analysis of further key residues required the production of further substitutions hence mutant *kasA* allele production was facilitated by the development of a mutagenesis system termed the "mega primer method". This system provided more success than the Stratagene QuikChange system and allowed for the production of 13 of the 21 mutations identified as candidates for investigation. The production of soluble enzymematically-active protein became a major

obstacle in this study, hence the elucidation of a viable protein purification protocol is still required and is vital for the further analysis of key amino acid residues within KasA.

Like *M. tuberculosis*, *C. glutamicum* possess mycolic acids but these corynomycolates are much shorter in chain length (C_{32}). Although the function of a number of genes involved in fatty acid and mycolic acid biosynthesis have been identified, information relevant to the condensation event within these biosynthetic pathways is relatively sparse. Analysis of these Corynebacterianeae genomes showed that they possess a high number of accD genes, whose gene products resemble the β-subunit of the acetyl-CoA carboxylase of *E. coli*, providing the activated intermediate for fatty acid synthesis. In this study four putative *accD* genes found in *C. glutamicum* were analysed. Although growth of the *accD4* mutant remained unchanged, growth of the *accD1* mutant was strongly impaired and partially recovered by the addition of exogenous oleic acid. Over-expression of *accD1* and *accBC*, encoding the carboxylase α-subunit, resulted in an 8-fold increase in malonyl-CoA formation from acetyl-CoA in cell lysates, providing evidence that *accD1* encodes a carboxyltransferase involved in the biosynthesis of malonyl-CoA. Interestingly, fatty acid profiles remained unchanged in both *accD2* and *accD3* mutants. However a complete loss of mycolic acids, either as organic extractable trehalose and glucose mycolates or as cell wall-bound mycolates was observed. These two carboxyltransferases are also retained in all Corynebacterianeae, including *M. leprae*, constituting two distinct groups of orthologs. Furthermore, carboxyl fixation assays, as well as a study of a *Cg-pks* deletion mutant, led us to conclude that *accD2* and *accD3* are key to mycolic acid biosynthesis, thus providing a carboxylated intermediate during condensation of the mero-chain and α-branch directed by the *pks*-encoded polyketide synthase. This study illustrates that the high number of *accD* paralogs have evolved to represent specific variations on the well known basic theme of providing carboxylated intermediates in lipid biosynthesis.

Further analysis of the *pks13* encoded polyketide synthase identified it as a homolog of the *Cg-pks* recognized in the previous study and is the most likely candidate to perform the Claisen condensation event involved in mycolate biosynthesis. Successful expression and purification of Pks13 may allow for the development of *in vitro* assays and therefore validate the results found in the *C. glutamicum* system. The identification and careful characterisation of the key biosynthetic partners of the mycobacterial Pks13, will allow for the production of the high molecular weight fatty acid 2-carboxy-derivatives, which would otherwise be difficult to synthesis chemically.

Overall, the increased understanding of the roles played by these key enzymes of mycolic acid biosynthesis will future development of existing inhibitors and identification of new inhibitors, eventually leading to novel anti-tubercular agents, which might target the production of meromycolates or the final Claisen condensation event to produce the mature mycolate.

CHAPTER 7

References

Abbadi, A., Brummel, M. & Spener, F. (2000). Knockout of the regulatory site of 3-ketoacyl-ACP synthase III enhances short- and medium-chain acyl-ACP synthesis. *Plant J* **24**, 1-9.

Ahibo-Coffy, A., Aurelle, H., Lacave, C., Prome, J. C., Puzo, G. & Savagnac, A. (1978). Isolation, structural studies and chemical synthesis of a 'palmitone lipid' from *Corynebacterium diphtheriae*. *Chem Phys Lipids* **22**, 185-195.

Asselineau, C., Asselineau, J., Laneelle, G. & Laneelle, M. A. (2002). The biosynthesis of mycolic acids by Mycobacteria: current and alternative hypotheses. *Prog Lipid Res* **41**, 501-523.

Bachhawat, N. & Mande, S. C. (1999). Identification of the *INO1* gene of *Mycobacterium tuberculosis* H37Rv reveals a novel class of inositol-1-phosphate synthase enzyme. *J Mol Biol* **291**, 531-536.

Baldock, C., Rafferty, J. B., Sedelnikova, S. E., Baker, P. J. & Stuitje, A. R. (1996). A mechanism of drug action revealed by structure studies of enoyl reductase. *Science* **274**, 2107-2110.

Bamaga, M., Wright, D. J. & Zhang, H. (2002). Selection of in vitro mutants of pyrazinamide-resistant Mycobacterium tuberculosis. *Int J Antimicrob Agents* **20**, 275-281.

Banerjee, A., Sugantino, M., Sacchettini, J. C. & Jacobs, W. R., Jr. (1998). The *mabA* gene from the *inhA* operon of *Mycobacterium tuberculosis* encodes a 3-ketoacyl reductase that fails to confer isoniazid resistance. *Microbiology* **144 (Pt 10)**, 2697-2704.

Barry, C. E., 3rd, Lee, R. E., Mdluli, K., Sampson, A. E., Schroeder, B. G., Slayden, R. A. & Yuan, Y. (1998). Mycolic acids: structure, biosynthesis and physiological functions. *Prog Lipid Res* **37**, 143-179.

Baulard, A. R., Gurcha, S. S., Engohang-Ndong, J., Gouffi, K., Locht, C. & Besra, G. S. (2003). *In vivo* interaction between the polyprenol phosphate mannose synthase Ppm1 and the integral membrane protein Ppm2 from *Mycobacterium smegmatis* revealed by a bacterial two-hybrid system. *J Biol Chem* **278**, 2242-2248.

Baulard, A. R., Betts, J. C., Engohang-Ndong, J., Quan, S., McAdam, R. A., Brennan, P. J., Locht, C. & Besra, G. S. (2000). Activation of the pro-drug ethionamide is regulated in mycobacteria. *J Biol Chem* **275**, 28326-28331.

Belanger, A. E. & Inamine, J. M. (2000). Genetics of Cell Wall Biosynthesis. In *Molecular Genetics of Mycobacteria*, pp. 191-202. Edited by G. F. Hatfull & W. R. Jacobs, Jr. Washington D. C.: ASM Press.

Belanger, A. E., Besra, G. S., Ford, M. E., Mikusova, K., Belisle, J. T., Brennan, P. J. & Inamine, J. M. (1996). The *embAB* genes of *Mycobacterium avium* encode an arabinosyl transferase involved in cell wall arabinan biosynthesis that is the target for the antimycobacterial drug ethambutol. *Proc Natl Acad Sci U S A* **93**, 11919-11924.

Belisle, J. T., Vissa, V. D., Sievert, T., Takayama, K., Brennan, P. J. & Besra, G. S. (1997). Role of the major antigen of *Mycobacterium tuberculosis* in cell wall biogenesis. *Science* **276**, 1420-1422.

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

Bernadou, J., Nguyen, M. & Meunier, B. (2001). [The mechanism of action of isoniazid. A chemical model of activation]. *Ann Pharm Fr* **59**, 331-337.

Besra, G. S. & Chatterjee, D. (1994). Lipids and carbohydrates of *Mycobacterium tuberculosis*. In *Tuberculosis: pathogenesis, protection and control*, pp. 285-306. Edited by B. Bloom. Washington D.C.: American Society for Microbiology.

Besra, G. S., Morehouse, C. B., Rittner, C. M., Waechter, C. J. & Brennan, P. J. (1997). Biosynthesis of mycobacterial lipoarabinomannan. *J Biol Chem* **272**, 18460-18466.

Besra, G. S., Sievert, T., Lee, R. E., Slayden, R. A., Brennan, P. J. & Takayama, K. (1994). Identification of the apparent carrier in mycolic acid synthesis. *Proc Natl Acad Sci U S A* **91**, 12735-12739.

Besra, G. S., Bolton, R. C., McNeil, M. R., Ridell, M., Simpson, K. E., Gluska, J., van Halbeek, H., Brennan, P. J. & Minnikin, D. E. (1992). Structural elucidation of a novel family of acyltrehalose from *Mycobacterium tuberculosis*. *Biochemistry* **31**, 9832-9837.

Betts, J. C., Lukey, P. T., Robb, L. C., McAdam, R. A. & Duncan, K. (2002). Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. *Mol Microbiol* **43**, 717-731.

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

Brand, S., Niehaus, K., Puhler, A. & Kalinowski, J. (2003). Identification and functional analysis of six mycolyltransferase genes of *Corynebacterium glutamicum* ATCC 13032: the genes *cop1*, *cmt1*, and *cmt2* can replace each other in the synthesis of trehalose dicorynomycolate, a component of the mycolic acid layer of the cell envelope. *Arch Microbiol* **180**, 33-44.

Brennan, P. J. (2003). Structure, function, and biogenesis of the cell wall of *Mycobacterium tuberculosis*. *Tuberculosis (Edinb)* **83**, 91-97.

Brennan, P. J. & Nikaido, H. (1995). The envelope of mycobacteria. *Annu Rev Biochem* **64**, 29-63.

Caceres, N. E., Harris, N. B., Wellehan, J. F., Feng, Z., Kapur, V. & Barletta, R. G. (1997). Overexpression of the D-alanine racemase gene confers resistance to D-cycloserine in *Mycobacterium smegmatis*. *J Bacteriol* **179**, 5046-5055.

Camacho, L. R., Constant, P., Raynaud, C., Lanelle, M. A., Triccas, J. A., Gicquel, B., Daffe, M. & Guilhot, C. (2001). Analysis of the phthiocerol dimycocerosate locus of

Mycobacterium tuberculosis. Evidence that this lipid is involved in the cell wall permeability barrier. *J Biol Chem* **276**, 19845-19854.

Campbell, J. W. & Cronan, J. E., Jr. (2001). Bacterial fatty acid biosynthesis: targets for antibacterial drug discovery. *Annu Rev Microbiol* **55**, 305-332.

CDC (1995). Tuberculosis morbidity--United States. *MMWR* **45**, 365-370.

Cerdeno-Tarraga, A. M., Efstratiou, A., Dover, L. G. & other authors (2003). The complete genome sequence and analysis of *Corynebacterium diphtheriae* NCTC13129. *Nucleic Acids Res* **31**, 6516-6523.

Chatterjee, D. & Khoo, K. H. (1998). Mycobacterial lipoarabinomannan: an extraordinary lipoheteroglycan with profound physiological effects. *Glycobiology* **8**, 113-120.

Chatterjee, D., Bozic, C. M., McNeil, M. & Brennan, P. J. (1991). Structural features of the arabinan component of the lipoarabinomannan of *Mycobacterium tuberculosis*. *J Biol Chem* **266**, 9652-9660.

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

Cheng, Y. Q., Tang, G. L. & Shen, B. (2003). Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. *Proc Natl Acad Sci U S A* **100**, 3149-3154.

Child, C. J. & Shoolingin-Jordan, P. M. (1998). Inactivation of the polyketide synthase, 6-methylsalicylic acid synthase, by the specific modification of Cys-204 of the beta-ketoacyl synthase by the fungal mycotoxin cerulenin. *Biochem J* **330 (Pt 2)**, 933-937.

Choi, K. H., Heath, R. J. & Rock, C. O. (2000a). beta-ketoacyl-acyl carrier protein synthase III (FabH) is a determining factor in branched-chain fatty acid biosynthesis. *J Bacteriol* **182**, 365-370.

Choi, K. H., Kremer, L., Besra, G. S. & Rock, C. O. (2000b). Identification and substrate specificity of beta -ketoacyl (acyl carrier protein) synthase III (mtFabH) from *Mycobacterium tuberculosis*. *J Biol Chem* **275**, 28201-28207.

Clark, J. & Wallace, A. (1967). The susceptibility of mycobacteria to rifamide and rifampicin. *Tubercle* **48**, 144-148.

Clough, R. C., Matthis, A. L., Barnum, S. R. & Jaworski, J. G. (1992). Purification and characterization of 3-ketoacyl-acyl carrier protein synthase III from spinach. A condensing enzyme utilizing acetyl-coenzyme A to initiate fatty acid synthesis. *J Biol Chem* **267**, 20992-20998.

Cohn, D. L., Bustreo, F. & Raviglione, M. C. (1997). Drug-resistant tuberculosis: review of the worldwide situation and the WHO/IUATLD Global Surveillance Project. International Union Against Tuberculosis and Lung Disease. *Clin Infect Dis* **24 Suppl 1**, S121-130.

Cole, S. T. (1994). *Mycobacterium tuberculosis*: drug-resistance mechanisms. *Trends Microbiol* **2**, 411-415.

Cole, S. T., Brosch, R., Parkhill, J. & other authors (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537-544.

Cole, S. T., Eiglmeier, K., Parkhill, J. & other authors (2001). Massive gene decay in the leprosy bacillus. *Nature* **409**, 1007-1011.

Corbett, E. L., Watt, C. J., Walker, N., Maher, D., Williams, B. G., Raviglione, M. C. & Dye, C. (2003). The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. *Arch Intern Med* **163**, 1009-1021.

Daffe, M. & Draper, P. (1998). The envelope layers of mycobacteria with reference to their pathogenicity. *Adv Microb Physiol* **39**, 131-203.

Daines, R. A., Pendrak, I., Sham, K. & other authors (2003). First X-ray cocrystal structure of a bacterial FabH condensing enzyme and a small molecule inhibitor achieved using rational design and homology modeling. *J Med Chem* **46**, 5-8.

Davies, C., Heath, R. J., White, S. W. & Rock, C. O. (2000). The 1.8 A crystal structure and active-site architecture of beta-ketoacyl-acyl carrier protein synthase III (FabH) from *Escherichia coli. Structure Fold Des* **8**, 185-195.

Davies, P. D. & Yew, W. W. (2003). Recent developments in the treatment of tuberculosis. *Expert Opin Investig Drugs* **12**, 1297-1312.

Dawe, J. H., Porter, C. T., Thornton, J. M. & Tabor, A. B. (2003). A template search reveals mechanistic similarities and differences in beta-ketoacyl synthases (KAS) and related enzymes. *Proteins* **52**, 427-435.

De Sousa-D'Auria, C., Kacem, R., Puech, V., Tropis, M., Leblon, G., Houssin, C. & Daffe, M. (2003). New insights into the biogenesis of the cell envelope of corynebacteria: identification and functional characterization of five new mycoloyltransferase genes in *Corynebacterium glutamicum. FEMS Microbiol Lett* **224**, 35-44.

DeBarber, A. E., Mdluli, K., Bosman, M., Bekker, L. G. & Barry, C. E., 3rd (2000). Ethionamide activation and sensitivity in multidrug-resistant *Mycobacterium tuberculosis. Proc Natl Acad Sci U S A* **97**, 9677-9682.

Dehesh, K., Tai, H., Edwards, P., Byrne, J. & Jaworski, J. G. (2001). Overexpression of 3-ketoacyl-acyl-carrier protein synthase III in plants reduces the rate of lipid synthesis. *Plant Physiol* **125**, 1103-1114.

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

Dinadayala, P., Laval, F., Raynaud, C., Lemassu, A., Lanelle, M. A., Lanelle, G. & Daffe, M. (2003). Tracking the putative biosynthetic precursors of oxygenated mycolates of

Mycobacterium tuberculosis. Structural analysis of fatty acids of a mutant strain deviod of methoxy- and ketomycolates. *J Biol Chem* **278**, 7310-7319.

Dmitriev, B. A., Ehlers, S., Rietschel, E. T. & Brennan, P. J. (2000). Molecular mechanics of the mycobacterial cell wall: from horizontal layers to vertical scaffolds. *Int J Med Microbiol* **290**, 251–258.

Dover, L. G., Cerdeno-Tarraga, A. M., Pallen, M. J., Parkhill, J. & Besra, G. S. (2004). Comparative cell wall core biosynthesis in the mycolated pathogens, Mycobacterium tuberculosis and Corynebacterium diphtheriae. *FEMS Microbiol Rev* **28**, 225-250.

Dubey, V. S., Sirakova, T. D., Cynamon, M. H. & Kolattukudy, P. E. (2003). Biochemical function of *msl5* (*pks8* plus *pks17*) in *Mycobacterium tuberculosis* H37Rv: biosynthesis of monomethyl branched unsaturated fatty acids. *J Bacteriol* **185**, 4620-4625.

Dye, C. & Espinal, M. A. (2001). Will tuberculosis become resistant to all antibiotics? *Proc R Soc Lond B Biol Sci* **268**, 45-52.

Dye, C., Espinal, M. A., Watt, C. J., Mbiaga, C. & Williams, B. G. (2002). Worldwide incidence of multidrug-resistant tuberculosis. *J Infect Dis* **185**, 1197-1202.

Edwards, P., Nelsen, J. S., Metz, J. G. & Dehesh, K. (1997). Cloning of the *fabF* gene in an expression vector and in vitro characterization of recombinant *fabF* and *fabB* encoded enzymes from *Escherichia coli*. *FEBS Lett* **402**, 62-66.

Eggeling, L., Pfefferle, W. & Sahm, H. (2001). Amino acids. In *Basic Biotechnology*, pp. 281-302. Edited by C. Ratledge & B. Kristiansen: Cambridge University Press.

Enarson, N. D. & Murray, J. F. (1996). Global epidemiology of tuberculosis. In *Tuberculosis*, pp. 57-75. Edited by W. N. Rom & S. Garay. Boston: Little, Brown and Co.

Engohang-Ndong, J., Baillat, D., Aumercier, M., Bellefontaine, F., Besra, G. S., Locht, C. & Baulard, A. R. (2004). EthR, a repressor of the TetR/CamR family implicated in ethionamide resistance in mycobacteria, octamerizes cooperatively on its operator. *Mol Microbiol* **51**, 175-188.

Escuyer, V. E., Lety, M. A., Torrelles, J. B. & other authors (2001). The role of the *embA* and *embB* gene products in the biosynthesis of the terminal hexaarabinofuranosyl motif of *Mycobacterium smegmatis* arabinogalactan. *J Biol Chem* **276**, 48854-48862.

Fang, Z., Doig, C., Rayner, A., Kenna, D. T., Watt, B. & Forbes, K. J. (1999). Molecular evidence for heterogeneity of the multiple-drug-resistant *Mycobacterium tuberculosis* population in Scotland (1990 to 1997). *J Clin Microbiol* **37**, 998-1003.

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

Fleischmann, R. D., Alland, D., Eisen, J. A. & other authors (2002). Whole-genome comparison of Mycobacterium tuberculosis clinical and laboratory strains. *J Bacteriol* **184**, 5479-5490.

Fraaije, M. W., Kamerbeek, N. M., Heidekamp, A. J., Fortin, R. & Janssen, D. B. (2004). The prodrug activator EtaA from *Mycobacterium tuberculosis* is a Baeyer-Villiger monooxygenase. *J Biol Chem* **279**, 3354-3360.

Furukawa, H., Tsay, J. T., Jackowski, S., Takamura, Y. & Rock, C. O. (1993). Thiolactomycin resistance in *Escherichia coli* is associated with the multidrug resistance efflux pump encoded by *emrAB*. *J Bacteriol* **175**, 3723-3729.

Galili, S., Fromm, H., Aviv, D., Edelman, M. & Galun, E. (1989). Ribosomal protein S12 as a site for streptomycin resistance in *Nicotiana* chloroplasts. *Mol Gen Genet* **218**, 289-292.

Gande, R., Gibson, K. J., Brown, A. K. & other authors (2004). Acyl-CoA carboxylases (accD2 and accD3), together with a unique polyketide synthase (Cg-pks), are key to mycolic acid biosynthesis in Corynebacterianeae such as Corynebacterium glutamicum and Mycobacterium tuberculosis. *J Biol Chem* **279**, 44847-44857.

Gao, L. Y., Laval, F., Lawson, E. H., Groger, R. K., Woodruff, A., Morisaki, J. H., Cox, J. S., Daffe, M. & Brown, E. J. (2003). Requirement for *kasB* in *Mycobacterium* mycolic acid biosynthesis, cell wall impermeability and intracellular survival: implications for therapy. *Mol Microbiol* **49**, 1547-1563.

Garnier, T., Eiglmeier, K., Camus, J. C. & other authors (2003). The complete genome sequence of *Mycobacterium bovis*. *Proc Natl Acad Sci U S A* **100**, 7877-7882.

Garwin, J. L., Klages, A. L. & Cronan, J. E., Jr. (1980). Structural, enzymatic, and genetic studies of beta-ketoacyl-acyl carrier protein synthases I and II of *Escherichia coli*. *J Biol Chem* **255**, 11949-11956.

George, K. M., Yuan, Y., Sherman, D. R. & Barry, C. E., 3rd (1995). The biosynthesis of cyclopropanated mycolic acids in *Mycobacterium tuberculosis*. Identification and functional analysis of CMAS-2. *J Biol Chem* **270**, 27292-27298.

Gibson, K. J., Eggeling, L., Maughan, W. N., Krumbach, K., Gurcha, S. S., Nigou, J., Puzo, G., Sahm, H. & Besra, G. S. (2003). Disruption of Cg-Ppm1, a polyprenyl monophosphomannose synthase, and the generation of lipoglycan-less mutants in *Corynebacterium glutamicum*. *J Biol Chem* **278**, 40842-40850.

Girling, D. J. (1989). The Chemotherapy of Tuberculosis. In *The biology of Mycobacteria*, pp. 285-323. Edited by C. Ratledge, J. Stanford & J. M. Grange. London: Academic Press.

Glickman, M. S. (2003). The *mmaA2* gene of *Mycobacterium tuberculosis* encodes the distal cyclopropane synthase of the alpha-mycolic acid. *J Biol Chem* **278**, 7844-7849.

Glickman, M. S. & Jacobs, W. R., Jr. (2001). Microbial pathogenesis of *Mycobacterium tuberculosis*: dawn of a discipline. *Cell* **104**, 477-485.

Glickman, M. S., Cahill, S. M. & Jacobs, W. R., Jr. (2001). The *Mycobacterium tuberculosis cmaA2* gene encodes a mycolic acid trans-cyclopropane synthetase. *J Biol Chem* **276**, 2228-2233.

Goffin, C. & Ghuysen, J. M. (2002). Biochemistry and comparative genomics of SxxK superfamily acyltransferases offer a clue to the mycobacterial paradox: presence of penicillin-susceptible target proteins versus lack of efficiency of penicillin as therapeutic agent. *Microbiol Mol Biol Rev* **66**, 702-738.

Grassberger, M. A., Turnowsky, F. & Hildebrandt, J. (1984). Preparation and antibacterial activities of new 1,2,3-diazaborine derivatives and analogues. *J Med Chem* **27**, 947-953.

Guerardel, Y., Maes, E., Elass, E., Leroy, Y., Timmerman, P., Besra, G. S., Locht, C., Strecker, G. & Kremer, L. (2002). Structural study of lipomannan and lipoarabinomannan from *Mycobacterium chelonae*. Presence of unusual components with alpha 1,3-mannopyranose side chains. *J Biol Chem* **277**, 30635-30648.

Gurcha, S. S., Baulard, A. R., Kremer, L. & other authors (2002). Ppm1, a novel polyprenol monophosphomannose synthase from *Mycobacterium tuberculosis*. *Biochem J* **365**, 441-450.

Haase, F. C., Henrikson, K. P., Treble, D. H. & Allen, S. H. (1982). The subunit structure and function of the propionyl coenzyme A carboxylase of *Mycobacterium smegmatis*. *J Biol Chem* **257**, 11994-11999.

Hall, P. R., Wang, Y. F., Rivera-Hainaj, R. E., Zheng, X., Pustai-Carey, M., Carey, P. R. & Yee, V. C. (2003). Transcarboxylase 12S crystal structure: hexamer assembly and substrate binding to a multienzyme core. *Embo J* **22**, 2334-2347.

Han, L., Lobo, S. & Reynolds, K. A. (1998). Characterization of beta-ketoacyl-acyl carrier protein synthase III from *Streptomyces glaucescens* and its role in initiation of fatty acid biosynthesis. *J Bacteriol* **180**, 4481-4486.

Hatfull, G. F. (1996). The molecular genetics of *Mycobacterium tuberculosis*. *Curr Top Microbiol Immunol* **215**, 29-47.

Heath, R. J. & Rock, C. O. (1996a). Regulation of fatty acid elongation and initiation by acyl-acyl carrier protein in *Escherichia coli*. *J Biol Chem* **271**, 1833-1836.

Heath, R. J. & Rock, C. O. (1996b). Inhibition of beta-ketoacyl-acyl carrier protein synthase III (FabH) by acyl-acyl carrier protein in *Escherichia coli*. *J Biol Chem* **271**, 10996-11000.

Heath, R. J. & Rock, C. O. (2002). The Claisen condensation in biology. *Nat Prod Rep* **19**, 581-596.

Heifets, L. B. (1994). Expedited detection of drug resistance in tuberculosis patients. *Ann Emerg Med* **24**, 457-461.

Heym, B., Honore, N., Truffot-Pernot, C., Banerjee, A., Schurra, C., Jacobs, W. R., Jr., van Embden, J. D., Grosset, J. H. & Cole, S. T. (1994). Implications of multidrug resistance

for the future of short-course chemotherapy of tuberculosis: a molecular study. *Lancet* **344**, 293-298.

Higashi, Y., Strominger, J. L. & Sweeley, C. C. (1967). Structure of a lipid intermediate in cell wall peptidoglycan synthesis: a derivative of a C55 isoprenoid alcohol. *Proc Natl Acad Sci U S A* **57**, 1878-1884.

Howard, C. F., Jr. (1968). *De novo* synthesis and elongation of fatty acids by subcellar fractions of monkey aorta. *J Lipid Res* **9**, 254-261.

Huang, T. S., Lee, S. S., Tu, H. Z., Huang, W. K., Chen, Y. S., Huang, C. K., Wann, S. R., Lin, H. H. & Liu, Y. C. (2003). Correlation between pyrazinamide activity and *pncA* mutations in *Mycobacterium tuberculosis* isolates in Taiwan. *Antimicrob Agents Chemother* **47**, 3672-3673.

Huang, W., Jia, J., Edwards, P., Dehesh, K., Schneider, G. & Lindqvist, Y. (1998). Crystal structure of beta-ketoacyl-acyl carrier protein synthase II from *E.coli* reveals the molecular architecture of condensing enzymes. *Embo J* **17**, 1183-1191.

Jackowski, S., Murphy, C. M., Cronan, J. E., Jr. & Rock, C. O. (1989). Acetoacetyl-acyl carrier protein synthase. A target for the antibiotic thiolactomycin. *J Biol Chem* **264**, 7624-7629.

Jackowski, S., Zhang, Y. M., Price, A. C., White, S. W. & Rock, C. O. (2002). A missense mutation in the *fabB* (beta-ketoacyl-acyl carrier protein synthase I) gene confers tiolactomycin resistance to *Escherichia coli*. *Antimicrob Agents Chemother* **46**, 1246-1252.

Jackson, M., Crick, D. C. & Brennan, P. J. (2000). Phosphatidylinositol is an essential phospholipid of mycobacteria. *J Biol Chem* **275**, 30092-30099.

Jackson, M., Portnoi, D., Catheline, D., Dumail, L., Rauzier, J., Legrand, P. & Gicquel, B. (1997). *Mycobacterium tuberculosis* Des protein: an immunodominant target for the humoral response of tuberculous patients. *Infect Immun* **65**, 2883-2889.

Jez, J. M. & Noel, J. P. (2000). Mechanism of chalcone synthase. pKa of the catalytic cysteine and the role of the conserved histidine in a plant polyketide synthase. *J Biol Chem* **275**, 39640-39646.

Kacem, R., De Sousa-D'Auria, C., Tropis, M., Chami, M., Gounon, P., Leblon, G., Houssin, C. & Daffe, M. (2004). Importance of mycoloyltransferases on the physiology of *Corynebacterium glutamicum*. *Microbiology* **150**, 73-84.

Kalinowski, J., Bathe, B., Bartels, D. & other authors (2003). The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. *J Biotechnol* **104**, 5-25.

Keilhauer, C., Eggeling, L. & Sahm, H. (1993). Isoleucine synthesis in *Corynebacterium glutamicum*: molecular analysis of the *ilvB-ilvN-ilvC* operon. *J Bacteriol* **175**, 5595-5603.

Khandekar, S. S., Konstantinidis, A. K., Silverman, C. & other authors (2000). Expression, purification, and crystallization of the *Escherichia coli* selenomethionyl beta-ketoacyl-acyl carrier protein synthase III. *Biochem Biophys Res Commun* **270**, 100-107.

Khandekar, S. S., Gentry, D. R., Van Aller, G. S. & other authors (2001). Identification, substrate specificity, and inhibition of the *Streptococcus pneumoniae* beta-ketoacyl-acyl carrier protein synthase III (FabH). *J Biol Chem* **276**, 30024-30030.

Khoo, K. H., Douglas, E., Azadi, P., Inamine, J. M., Besra, G. S., Mikusova, K., Brennan, P. J. & Chatterjee, D. (1996). Truncated structural variants of lipoarabinomannan in ethambutol drug-resistant strains of *Mycobacterium smegmatis*. Inhibition of arabinan biosynthesis by ethambutol. *J Biol Chem* **271**, 28682-28690.

Kimura, E., Abe, C., Kawahara, Y., Nakamatsu, T. & Tokuda, H. (1997). A *dtsR* gene-disrupted mutant of *Brevibacterium lactofermentum* requires fatty acids for growth and efficiently produces L-glutamate in the presence of an excess of biotin. *Biochem Biophys Res Commun* **234**, 157-161.

Kordulakova, J., Gilleron, M., Mikusova, K., Puzo, G., Brennan, P. J., Gicquel, B. & Jackson, M. (2002). Definition of the first mannosylation step in phosphatidylinositol mannoside synthesis. PimA is essential for growth of mycobacteria. *J Biol Chem* **277**, 31335-31344.

Kremer, L., Baulard, A. R. & Besra, G. S. (2000a). Genetics of Mycolic Acid Biosynthesis. In *Molecular Genetics of Mycobacteria*, pp. 173-190. Edited by G. F. Hatfull & W. R. Jacobs, Jr. Washington D.C.: ASM Press.

Kremer, L., Maughan, W. N., Wilson, R. A., Dover, L. G. & Besra, G. S. (2002a). The *M. tuberculosis* antigen 85 complex and mycolyltransferase activity. *Lett Appl Microbiol* **34**, 233-237.

Kremer, L., Gurcha, S. S., Bifani, P., Hitchin, P. G., Baulard, A., Morris, H. R., Dell, A., Brennan, P. J. & Besra, G. S. (2002b). Characterization of a putative alpha-mannosyltransferase involved in phosphatidylinositol trimannoside biosynthesis in *Myocbacterium tuberculosis. Biochem J* **363**, 437-447.

Kremer, L., Nampoothiri, K. M., Lesjean, S. & other authors (2001a). Biochemical characterization of acyl carrier protein (AcpM) and malonyl-CoA:AcpM transacylase (mtFabD), two major components of *Mycobacterium tuberculosis* fatty acid synthase II. *J Biol Chem* **276**, 27967-27974.

Kremer, L., Dover, L. G., Carrere, S. & other authors (2002c). Mycolic acid biosynthesis and enzymic characterization of the beta-ketoacyl-ACP synthase A-condensing enzyme from *Mycobacterium tuberculosis. Biochem J* **364**, 423-430.

Kremer, L., Douglas, J. D., Baulard, A. R. & other authors (2000b). Thiolactomycin and related analogues as novel anti-mycobacterial agents targeting KasA and KasB condensing enzymes in *Mycobacterium tuberculosis. J Biol Chem* **275**, 16857-16864.

Kremer, L., Dover, L. G., Morehouse, C. & other authors (2001b). Galactan biosynthesis in *Mycobacterium tuberculosis*. Identification of a bifunctional UDP-galactofuranosyltransferase. *J Biol Chem* **276**, 26430-26440.

Kremer, L., Dover, L. G., Morbidoni, H. R. & other authors (2003). Inhibition of InhA activity, but not KasA activity, induces formation of a KasA-containing complex in mycobacteria. *J Biol Chem* **278**, 20547-20554.

Kremer, L. S. & Besra, G. S. (2002). Current status and future development of antitubercular chemotherapy. *Expert Opin Investig Drugs* **11**, 1033-1049.

Kuo, M. R., Morbidoni, H. R., Alland, D. & other authors (2003). Targeting tuberculosis and malaria through inhibition of Enoyl reductase: compound activity and structural data. *J Biol Chem* **278**, 20851-20859.

Larsen, M. H., Vilcheze, C., Kremer, L. & other authors (2002). Overexpression of inhA, but not kasA, confers resistance to isoniazid and ethionamide in Mycobacterium smegmatis, M. bovis BCG and M. tuberculosis. *Mol Microbiol* **46**, 453-466.

Lee, A. S., Lim, I. H., Tang, L. L., Telenti, A. & Wong, S. Y. (1999). Contribution of kasA analysis to detection of isoniazid-resistant *Mycobacterium tuberculosis* in Singapore. *Antimicrob Agents Chemother* **43**, 2087-2089.

Lee, R. E., Brennan, P. J. & Besra, G. S. (1996). *Mycobacterium tuberculosis* cell envelope. In *Tuberculosis*, pp. 1-27. Edited by T. M. Shinnick: Springer.

Lee, R. E., Brennan, P. J. & Besra, G. S. (1998). Synthesis of beta-D-arabinofuranosyl-1-monophosphoryl polyprenols: examination of their function as mycobacterial arabinosyl transferase donors. *Bioorg Med Chem Lett* **8**, 951-954.

Lee, R. E., Armour, J. W., Takayama, K., Brennan, P. J. & Besra, G. S. (1997). Mycolic acid biosynthesis: definition and targeting of the Claisen condensation step. *Biochim Biophys Acta* **1346**, 275-284.

Liou, G. F. & Khosla, C. (2003). Building-block selectivity of polyketide synthases. *Curr Opin Chem Biol* **7**, 279-284.

Liu, J. & Nikaido, H. (1999). A mutant of *Mycobacterium smegmatis* defective in the biosynthesis of mycolic acids accumulates meromycolates. *Proc Natl Acad Sci U S A* **96**, 4011-4016.

Liu, J., Barry, C. E., 3rd, Besra, G. S. & Nikaido, H. (1996). Mycolic acid structure determines the fluidity of the mycobacterial cell wall. *J Biol Chem* **271**, 29545-29551.

Ma, Y., Pan, F. & McNeil, M. (2002). Formation of dTDP-rhamnose is essential for growth of mycobacteria. *J Bacteriol* **184**, 3392-3395.

Ma, Y., Stern, R. J., Scherman, M. S. & other authors (2001). Drug targeting *Mycobacterium tuberculosis* cell wall synthesis: genetics of dTDP-rhamnose synthetic enzymes and development of a microtiter plate-based screen for inhibitors of conversion of dTDP-glucose to dTDP-rhamnose. *Antimicrob Agents Chemother* **45**, 1407-1416.

Magnuson, K., Carey, M. R. & Cronan, J. E., Jr. (1995). The putative fabJ gene of Escherichia coli fatty acid synthesis is the fabF gene. *J Bacteriol* **177**, 3593-3595.

Magnuson, K., Jackowski, S., Rock, C. O. & Cronan, J. E., Jr. (1993). Regulation of fatty acid biosynthesis in *Escherichia coli*. *Microbiol Rev* **57**, 522-542.

Marcinkeviciene, J., Jiang, W., Kopcho, L. M., Locke, G., Luo, Y. & Copeland, R. A. (2001). Enoyl-ACP reductase (FabI) of *Haemophilus influenzae*: steady-state kinetic mechanism and inhibition by triclosan and hexachlorophene. *Arch Biochem Biophys* **390**, 101-108.

Marrakchi, H., Laneelle, G. & Quemard, A. (2000). InhA, a target of the antituberculous drug isoniazid, is involved in a mycobacterial fatty acid elongation system, FAS-II. *Microbiology* **146 (Pt 2)**, 289-296.

Marrakchi, H., Ducasse, S., Labesse, G., Montrozier, H., Margeat, E., Emorine, L., Charpentier, X., Daffe, M. & Quemard, A. (2002). MabA (FabG1), a *Mycobacterium tuberculosis* protein involved in the long-chain fatty acid elongation system FAS-II. *Microbiology* **148**, 951-960.

McGuire, K. A., McGuire, J. N. & von Wettstein-Knowles, P. (2000). Acyl carrier protein (ACP) inhibition and other differences between beta-ketoacyl synthase (KAS) I and II. *Biochem Soc Trans* **28**, 607-610.

McGuire, K. A., Siggaard-Andersen, M., Bangera, M. G., Olsen, J. G. & von Wettstein-Knowles, P. (2001). beta-Ketoacyl-[acyl carrier protein] synthase I of *Escherichia coli*: aspects of the condensation mechanism revealed by analyses of mutations in the active site pocket. *Biochemistry* **40**, 9836-9845.

McLeod, R., Muench, S. P., Rafferty, J. B. & other authors (2001). Triclosan inhibits the growth of *Plasmodium falciparum* and *Toxoplasma gondii* by inhibition of apicomplexan Fab I. *Int J Parasitol* **31**, 109-113.

McMurry, L. M., Oethinger, M. & Levy, S. B. (1998). Triclosan targets lipid synthesis. *Nature* **394**, 531-532.

McMurry, L. M., McDermott, P. F. & Levy, S. B. (1999). Genetic evidence that InhA of *Mycobacterium smegmatis* is a target for triclosan. *Antimicrob Agents Chemother* **43**, 711-713.

McNeil, M. (1999). In *Genetics of bacterial polysaccarides*, pp. 207-223. Edited by J. B. Goldberg. Washington D.C.: CRC Press.

McNeil, M., Daffe, M. & Brennan, P. J. (1990). Evidence for the nature of the link between the arabinogalactan and peptidoglycan of mycobacterial cell walls. *J Biol Chem* **265**, 18200-18206.

McNeil, M., Daffe, M. & Brennan, P. J. (1991). Location of the mycolyl ester substituents in the cell walls of mycobacteria. *J Biol Chem* **266**, 13217-13223.

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

Mdluli, K., Slayden, R. A., Zhu, Y., Ramaswamy, S., Pan, X., Mead, D., Crane, D. D., Musser, J. M. & Barry, C. E., 3rd (1998). Inhibition of a *Mycobacterium tuberculosis* beta-ketoacyl ACP synthase by isoniazid. *Science* **280**, 1607-1610.

Middlebrook, G. (1954). Isoniazid-resistance and catalase activity of tubercle bacilli; a preliminary report. *Am Rev Tuberc* **69**, 471-472.

Mikusova, K., Slayden, R. A., Besra, G. S. & Brennan, P. J. (1995). Biogenesis of the mycobacterial cell wall and the site of action of ethambutol. *Antimicrob Agents Chemother* **39**, 2484-2489.

Mikusova, K., Mikus, M., Besra, G. S., Hancock, I. & Brennan, P. J. (1996). Biosynthesis of the linkage region of the mycobacterial cell wall. *J Biol Chem* **271**, 7820-7828.

Mikusova, K., Yagi, T., Stern, R., McNeil, M. R., Besra, G. S., Crick, D. C. & Brennan, P. J. (2000). Biosynthesis of the galactan component of the mycobacterial cell wall. *J Biol Chem* **275**, 33890-33897.

Minnikin, D. E. (1982). Lipids: complex lipids, their chemistry, biosynthesis and roles. In *The Biology of the Mycobacteria: Physiology, Identification and Classification*. Edited by C. Ratledge & J. Stanford. London, UK: Academic Press.

Minnikin, D. E., Kremer, L., Dover, L. G. & Besra, G. S. (2002). The methyl-branched fortifications of *Mycobacterium tuberculosis*. *Chem Biol* **9**, 545-553.

Moche, M., Dehesh, K., Edwards, P. & Lindqvist, Y. (2001). The crystal structure of beta-ketoacyl-acyl carrier protein synthase II from *Synechocystis sp.* at 1.54 A resolution and its relationship to other condensing enzymes. *J Mol Biol* **305**, 491-503.

Moche, M., Schneider, G., Edwards, P., Dehesh, K. & Lindqvist, Y. (1999). Structure of the complex between the antibiotic cerulenin and its target, beta-ketoacyl-acyl carrier protein synthase. *J Biol Chem* **274**, 6031-6034.

Musser, J. M. (1995). Antimicrobial agent resistance in mycobacteria: molecular genetic insights. *Clin Microbiol Rev* **8**, 496-514.

Nakamura, Y., Nishio, Y., Ikeo, K. & Gojobori, T. (2003). The genome stability in *Corynebacterium* species due to lack of the recombinational repair system. *Gene* **317**, 149-155.

Narain, J. P., Raviglione, M. C. & Kochi, A. (1992). HIV-associated tuberculosis in developing countries: epidemiology and strategies for prevention. *Tuber Lung Dis* **73**, 311-321.

Nigou, J., Dover, L. G. & Besra, G. S. (2002a). Purification and biochemical characterization of *Mycobacterium tuberculosis* SuhB, an inositol monophosphatase involved in inositol biosynthesis. *Biochemistry* **41**, 4392-4398.

Nigou, J., Gilleron, M. & Puzo, G. (2003). Lipoarabinomannans: from structure to biosynthesis. *Biochimie* **85**, 153-166.

Nigou, J., Gilleron, M., Rojas, M., Garcia, L. F., Thurnher, M. & Puzo, G. (2002b). Mycobacterial lipoarabinomannans: modulators of dendritic cell function and the apoptotic response. *Microbes Infect* **4**, 945-953.

Nigou, J., Gilleron, M., Cahuzac, B., Bounery, J. D., Herold, M., Thurnher, M. & Puzo, G. (1997). The phosphatidyl-*myo*-inositol anchor of the lipoarabinomannans from *Mycobacterium bovis* bacillus Calmette Guerin. Heterogeneity, structure, and role in the regulation of cytokine secretion. *J Biol Chem* **272**, 23094-23103.

Nishio, Y., Nakamura, Y., Kawarabayasi, Y. & other authors (2003). Comparative complete genome sequence analysis of the amino acid replacements responsible for the thermostability of *Corynebacterium efficiens*. *Genome Res* **13**, 1572-1579.

Oishi, H., Noto, T., Sasaki, H., Suzuki, K., Hayashi, T., Okazaki, H., Ando, K. & Sawada, M. (1982). Thiolactomycin, a new antibiotic. I. Taxonomy of the producing organism, fermentation and biological properties. *J Antibiot (Tokyo)* **35**, 391-395.

Olsen, J. G., Kadziola, A., von Wettstein-Knowles, P., Siggaard-Andersen, M. & Larsen, S. (2001). Structures of beta-ketoacyl-acyl carrier protein synthase I complexed with fatty acids elucidate its catalytic machinery. *Structure (Camb)* **9**, 233-243.

Olsen, J. G., Kadziola, A., von Wettstein-Knowles, P., Siggaard-Andersen, M., Lindquist, Y. & Larsen, S. (1999). The X-ray crystal structure of beta-ketoacyl [acyl carrier protein] synthase I. *FEBS Lett* **460**, 46-52.

Omura, S. (1976). The antibiotic cerulenin, a novel tool for biochemistry as an inhibitor of fatty acid synthesis. *Bacteriol Rev* **40**, 681-697.

Pantano, S., Alber, F., Lamba, D. & Carloni, P. (2002). NADH interactions with WT- and S94A-acyl carrier protein reductase from *Mycobacterium tuberculosis*: an *ab initio* study. *Proteins* **47**, 62-68.

Parikh, S. L., Xiao, G. & Tonge, P. J. (2000). Inhibition of InhA, the enoyl reductase from *Mycobacterium tuberculosis*, by triclosan and isoniazid. *Biochemistry* **39**, 7645-7650.

Parish, T. & Stoker, N. G. (1999). Mycobacteria: bugs and bugbears (two steps forward and one step back). *Mol Biotechnol* **13**, 191-200.

Parish, T., Liu, J., Nikaido, H. & Stoker, N. G. (1997). A *Mycobacterium smegmatis* mutant with a defective inositol monophosphate phosphatase gene homolog has altered cell envelope permeability. *J Bacteriol* **179**, 7827-7833.

Pearson, W. R. (1990). Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol* **183**, 63-98.

Petit & Lenderer (1984). In *The mycobacteria sourcebook*, pp. 301-322. Edited by G. P. Kubica & L. G. Wayne. New York: Marcel Dekker.

Phetsuksiri, B., Baulard, A. R., Cooper, A. M., Minnikin, D. E., Douglas, J. D., Besra, G. S. & Brennan, P. J. (1999). Antimycobacterial activities of isoxyl and new derivatives through the inhibition of mycolic acid synthesis. *Antimicrob Agents Chemother* **43**, 1042-1051.

Phetsuksiri, B., Jackson, M., Scherman, H. & other authors (2003). Unique mechanism of action of the thiourea drug isoxyl on *Mycobacterium tuberculosis*. *J Biol Chem* **278**, 53123-53130.

Portevin, D., De Sousa-D'Auria, C., Houssin, C., Grimaldi, C., Chami, M., Daffe, M. & Guilhot, C. (2004). A polyketide synthase catalyzes the last condensation step of mycolic acid biosynthesis in mycobacteria and related organisms. *Proc Natl Acad Sci U S A* **101**, 314-319.

Post-Beittenmiller, D., Jaworski, J. G. & Ohlrogge, J. B. (1991). *In vivo* pools of free and acylated acyl carrier proteins in spinach. Evidence for sites of regulation of fatty acid biosynthesis. *J Biol Chem* **266**, 1858-1865.

Price, A. C., Choi, K. H., Heath, R. J., Li, Z., White, S. W. & Rock, C. O. (2001). Inhibition of beta-ketoacyl-acyl carrier protein synthases by thiolactomycin and cerulenin. Structure and mechanism. *J Biol Chem* **276**, 6551-6559.

Prigge, S. T., He, X., Gerena, L., Waters, N. C. & Reynolds, K. A. (2003). The initiating steps of a type II fatty acid synthase in *Plasmodium falciparum* are catalyzed by pfACP, pfMCAT, and pfKASIII. *Biochemistry* **42**, 1160-1169.

Puech, V., Bayan, N., Salim, K., Leblon, G. & Daffe, M. (2000). Characterization of the *in vivo* acceptors of the mycoloyl residues transferred by the corynebacterial PS1 and the related mycobacterial antigens 85. *Mol Microbiol* **35**, 1026-1041.

Qiu, X., Janson, C. A., Smith, W. W., Head, M., Lonsdale, J. & Konstantinidis, A. K. (2001). Refined structures of beta-ketoacyl-acyl carrier protein synthase III. *J Mol Biol* **307**, 341-356.

Qiu, X., Janson, C. A., Konstantinidis, A. K., Nwagwu, S., Silverman, C., Smith, W. W., Khandekar, S., Lonsdale, J. & Abdel-Meguid, S. S. (1999). Crystal structure of beta-ketoacyl-acyl carrier protein synthase III. A key condensing enzyme in bacterial fatty acid biosynthesis. *J Biol Chem* **274**, 36465-36471.

Quemard, A., Sacchettini, J. C., Dessen, A., Vilcheze, C., Bittman, R., Jacobs, W. R., Jr. & Blanchard, J. S. (1995). Enzymatic characterization of the target for isoniazid in *Mycobacterium tuberculosis*. *Biochemistry* **34**, 8235-8241.

Qureshi, N., Sathyamoorthy, N. & Takayama, K. (1984). Biosynthesis of C30 to C56 fatty acids by an extract of *Mycobacterium tuberculosis* H37Ra. *J Bacteriol* **157**, 46-52.

Rainwater, D. L. & Kolattukudy, P. E. (1982). Isolation and characterization of acyl coenzyme A carboxylases from *Mycobacterium tuberculosis* and *Mycobacterium bovis*, which produce multiple methyl-branched mycocerosic acids. *J Bacteriol* **151**, 905-911.

Ramaswamy, S. & Musser, J. M. (1998). Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. *Tuber Lung Dis* **79**, 3-29.

Ratledge, C. & Dover, L. G. (2000). Iron metabolism in pathogenic bacteria. *Annu Rev Microbiol* **54**, 881-941.

Rattan, A., Kalia, A. & Ahmad, N. (1998). Multidrug-Resistant *Mycobacterium tuberculosis*: Molecular Perspectives. *Emerg Infect Dis* **4**, 195-207.

Raynaud, C., Laneelle, M. A., Senaratne, R. H., Draper, P., Laneelle, G. & Daffe, M. (1999). Mechanisms of pyrazinamide resistance in mycobacteria: importance of lack of uptake in addition to lack of pyrazinamidase activity. *Microbiology* **145 (Pt 6)**, 1359-1367.

Reeves, C. D., Murli, S., Ashley, G. W., Piagentini, M., Hutchinson, C. R. & McDaniel, R. (2001). Alteration of the substrate specificity of a modular polyketide synthase acyltransferase domain through site-specific mutations. *Biochemistry* **40**, 15464-15470.

Revill, W. P., Bibb, M. J., Scheu, A. K., Kieser, H. J. & Hopwood, D. A. (2001). Beta-ketoacyl acyl carrier protein synthase III (FabH) is essential for fatty acid biosynthesis in *Streptomyces coelicolor* A3(2). *J Bacteriol* **183**, 3526-3530.

Roberts, C. W., McLeod, R., Rice, D. W., Ginger, M., Chance, M. L. & Goad, L. J. (2003). Fatty acid and sterol metabolism: potential antimicrobial targets in apicomplexan and trypanosomatid parasitic protozoa. *Mol Biochem Parasitol* **126**, 129-142.

Rousseau, C., Sirakova, T. D., Dubey, V. S., Bordat, Y., Kolattukudy, P. E., Gicquel, B. & Jackson, M. (2003). Virulence attenuation of two Mas-like polyketide synthase mutants of *Mycobacterium tuberculosis*. *Microbiology* **149**, 1837-1847.

Rowe, D. C. & Summers, D. K. (1999). The quiescent-cell expression system for protein synthesis in *Escherichia coli*. *Appl Environ Microbiol* **65**, 2710-2715.

Rubires, X., Saigi, F., Pique, N., Climent, N., Merino, S., Alberti, S., Tomas, J. M. & Regue, M. (1997). A gene (wbbL) from *Serratia marcescens* N28b (O4) complements the rfb-50 mutation of *Escherichia coli* K-12 derivatives. *J Bacteriol* **179**, 7581-7586.

Saini, D. K., Pant, N., Das, T. K. & Tyagi, J. S. (2002). Cloning, overexpression, purification, and matrix-assisted refolding of DevS (Rv 3132c) histidine protein kinase of *Mycobacterium tuberculosis*. *Protein Expr Purif* **25**, 203-208.

Sarkar, G. & Sommer, S. S. (1990). The "megaprimer" method of site-directed mutagenesis. *Biotechniques* **8**, 404-407.

Sassetti, C. M., Boyd, D. H. & Rubin, E. J. (2003). Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* **48**, 77-84.

Sathyamoorthy, N., Qureshi, N. & Takayama, K. (1985). Purification and characterization of C28-55 fatty acids from *Mycobacterium smegmatis*. *Can J Microbiol* **31**, 214-219.

Scarsdale, J. N., Kazanina, G., He, X., Reynolds, K. A. & Wright, H. T. (2001). Crystal structure of the *Mycobacterium tuberculosis* beta-ketoacyl-acyl carrier protein synthase III. *J Biol Chem* **276**, 20516-20522.

Schaeffer, M. L., Agnihotri, G., Volker, C., Kallender, H., Brennan, P. J. & Lonsdale, J. T. (2001). Purification and biochemical characterization of the Mycobacterium tuberculosis beta-ketoacyl-acyl carrier protein synthases KasA and KasB. *J Biol Chem* **276**, 47029-47037.

Schaeffer, M. L., Khoo, K. H., Besra, G. S., Chatterjee, D., Brennan, P. J., Belisle, J. T. & Inamine, J. M. (1999). The *pimB* gene of *Mycobacterium tuberculosis* encodes a mannosyltransferase involved in lipoarabinomannan biosynthesis. *J Biol Chem* **274**, 31625-31631.

Schafer, A., Tauch, A., Jager, W., Kalinowski, J., Thierbach, G. & Puhler, A. (1994). Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene* **145**, 69-73.

Senior, S. J., Illarionov, P. A., Gurcha, S. S., Campbell, I. B., Schaeffer, M. L., Minnikin, D. E. & Besra, G. S. (2003). Biphenyl-based analogues of thiolactomycin, active against *Mycobacterium tuberculosis* mtFabH fatty acid condensing enzyme. *Bioorg Med Chem Lett* **13**, 3685-3688.

Shen, B. (2003). Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Curr Opin Chem Biol* **7**, 285-295.

Shen, B. & Hutchinson, C. R. (1993). Enzymatic synthesis of a bacterial polyketide from acetyl and malonyl coenzyme A. *Science* **262**, 1535-1540.

Shimakata, T. & Stumpf, P. K. (1983). Purification and characterization of beta-ketoacyl-ACP synthetase I from *Spinacia oleracea* leaves. *Arch Biochem Biophys* **220**, 39-45.

Shimakata, T., Iwaki, M. & Kusaka, T. (1984). *In vitro* synthesis of mycolic acids by the fluffy layer fraction of *Bacterionema matruchotii*. *Arch Biochem Biophys* **229**, 329-339.

Shimakata, T., Tsubokura, K., Kusaka, T. & Shizukuishi, K. (1985). Mass-spectrometric identification of trehalose 6-monomycolate synthesized by the cell-free system of *Bacterionema matruchotii*. *Arch Biochem Biophys* **238**, 497-508.

Siggaard-Andersen, M. (1993). Conserved residues in condensing enzyme domains of fatty acid synthases and related sequences. *Protein Seq Data Anal* **5**, 325-335.

Sirakova, T. D., Fitzmaurice, A. M. & Kolattukudy, P. (2002). Regulation of expression of *mas* and *fadD28*, two genes involved in production of dimycocerosyl phthiocerol, a virulence factor of *Mycobacterium tuberculosis*. *J Bacteriol* **184**, 6796-6802.

Sirakova, T. D., Thirumala, A. K., Dubey, V. S., Sprecher, H. & Kolattukudy, P. E. (2001). The *Mycobacterium tuberculosis pks2* gene encodes the synthase for the hepta- and

octamethyl-branched fatty acids required for sulfolipid synthesis. *J Biol Chem* **276**, 16833-16839.

Sivaraman, S., Zwahlen, J., Bell, A. F., Hedstrom, L. & Tonge, P. J. (2003). Structure-activity studies of the inhibition of FabI, the enoyl reductase from *Escherichia coli*, by triclosan: kinetic analysis of mutant FabIs. *Biochemistry* **42**, 4406-4413.

Slayden, R. A., Lee, R. E. & Barry, C. E., 3rd (2000). Isoniazid affects multiple components of the type II fatty acid synthase system of *Mycobacterium tuberculosis*. *Mol Microbiol* **38**, 514-525.

Slayden, R. A., Lee, R. E., Armour, J. W., Cooper, A. M., Orme, I. M., Brennan, P. J. & Besra, G. S. (1996). Antimycobacterial action of thiolactomycin: an inhibitor of fatty acid and mycolic acid synthesis. *Antimicrob Agents Chemother* **40**, 2813-2819.

Smith, D. W. (1982). Mycobacteria. In *Microbiology*. Edited by A. J. Braude, A. C. E. Davies & J. Frierer: W. B. Saunders Co.

Smith, S., Witkowski, A. & Joshi, A. K. (2003). Structural and functional organization of the animal fatty acid synthase. *Prog Lipid Res* **42**, 289-317.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* **98**, 503-517.

Sreevatsan, S., Stockbauer, K. E., Pan, X., Kreiswirth, B. N., Moghazeh, S. L., Jacobs, W. R., Jr., Telenti, A. & Musser, J. M. (1997). Ethambutol resistance in *Mycobacterium tuberculosis*: critical role of *embB* mutations. *Antimicrob Agents Chemother* **41**, 1677-1681.

Stackebrandt, E., Sproer, C., Rainey, F. A., Burghardt, J., Pauker, O. & Hippe, H. (1997). Phylogenetic analysis of the genus *Desulfotomaculum*: evidence for the misclassification of *Desulfotomaculum guttoideum* and description of *Desulfotomaculum orientis* as *Desulfosporosinus orientis* gen. nov., comb. nov. *Int J Syst Bacteriol* **47**, 1134-1139.

Stratton, M. A. & Reed, M. T. (1986). Short-course drug therapy for tuberculosis. *Clin Pharm* **5**, 977-987.

Strohmeier, G. R. & Fenton, M. J. (1999). Roles of lipoarabinomannan in the pathogenesis of tuberculosis. *Microbes Infect* **1**, 709-717.

Taniguchi, H., Aramaki, H., Nikaido, Y., Mizuguchi, Y., Nakamura, M., Koga, T. & Yoshida, S. (1996). Rifampicin resistance and mutation of the *rpoB* gene in *Mycobacterium tuberculosis*. *FEMS Microbiol Lett* **144**, 103-108.

Tatar, J. (1974). [Sensitivity of tubercle bacilli to pyrazinamide determined on the basis of their sensitivity to nicotinamide]. *Gruzlica* **42**, 765-771.

Telenti, A., Philipp, W. J., Sreevatsan, S., Bernasconi, C., Stockbauer, K. E., Wieles, B., Musser, J. M. & Jacobs, W. R., Jr. (1997). The *emb* operon, a gene cluster of *Mycobacterium tuberculosis* involved in resistance to ethambutol. *Nat Med* **3**, 567-570.

Trivedi, O. A., Arora, P., Sridharan, V., Tickoo, R., Mohanty, D. & Gokhale, R. S. (2004). Enzymic activation and transfer of fatty acids as acyl-adenylates in mycobacteria. *Nature* **428**, 441-445.

Tsay, J. T., Oh, W., Larson, T. J., Jackowski, S. & Rock, C. O. (1992). Isolation and characterization of the beta-ketoacyl-acyl carrier protein synthase III gene (*fabH*) from *Escherichia coli* K-12. *J Biol Chem* **267**, 6807-6814.

Turnowsky, F., Fuchs, K., Jeschek, C. & Hogenauer, G. (1989). *envM* genes of *Salmonella typhimurium* and *Escherichia coli*. *J Bacteriol* **171**, 6555-6565.

van Heijenoort, J. (2001). Formation of the glycan chains in the synthesis of bacterial peptidoglycan. *Glycobiology* **11**, 25R-36R.

Vance, D., Goldberg, I., Mitsuhashi, O. & Bloch, K. (1972). Inhibition of fatty acid synthetases by the antibiotic cerulenin. *Biochem Biophys Res Commun* **48**, 649-656.

Vannelli, T. A., Dykman, A. & Ortiz de Montellano, P. R. (2002). The antituberculosis drug ethionamide is activated by a flavoprotein monooxygenase. *J Biol Chem* **277**, 12824-12829.

Vilkas, E., Amar, C., Markovits, J., Vliegenthart, J. F. & Kamerling, J. P. (1973). Occurrence of a galactofuranose disaccharide in immunoadjuvant fractions of *Mycobacterium tuberculosis* (Cell walls and wax D). *Biochim Biophys Acta* **297**, 423-435.

Walker, R. W., Prome, J. C. & Lacave, C. S. (1973). Biosynthesis of mycolic acids. Formation of a C32 beta-keto ester from palmitic acid in a cell-free system of *Corynebacterium diphtheriae*. *Biochim Biophys Acta* **326**, 52-62.

Waller, R. F., Ralph, S. A., Reed, M. B., Su, V., Douglas, J. D., Minnikin, D. E., Cowman, A. F., Besra, G. S. & McFadden, G. I. (2003). A type II pathway for fatty acid biosynthesis presents drug targets in *Plasmodium falciparum*. *Antimicrob Agents Chemother* **47**, 297-301.

Watanabe, M., Yamada, Y., Iguchi, K. & Minnikin, D. E. (1994). Structural elucidation of new phenolic glycolipids from *Mycobacterium tuberculosis*. *Biochim Biophys Acta* **1210**, 174-180.

Watanabe, M., Aoyagi, Y., Ridell, M. & Minnikin, D. E. (2001). Separation and characterization of individual mycolic acids in representative mycobacteria. *Microbiology* **147**, 1825-1837.

Watanabe, M., Aoyagi, Y., Mitome, H., Fujita, T., Naoki, H., Ridell, M. & Minnikin, D. E. (2002). Location of functional groups in mycobacterial meromycolate chains; the recognition of new structural principles in mycolic acids. *Microbiology* **148**, 1881-1902.

WHO (2003). Global TB control. **7**, 1-192.

Wilson, R. A., Maughan, W. N., Kremer, L., Besra, G. S. & Futterer, K. (2004). The structure of *Mycobacterium tuberculosis* MPT51 (FbpC1) defines a new family of non-catalytic alpha/beta hydrolases. *J Mol Biol* **335**, 519-530.

Wilson, R. A., Rai, S., Maughan, W. N., Kremer, L., Kariuki, B. M., Harris, K. D., Wagner, T., Besra, G. S. & Futterer, K. (2003). Crystallization and preliminary X-ray diffraction data of *Mycobacterium tuberculosis* FbpC1 (Rv3803c). *Acta Crystallogr D Biol Crystallogr* **59**, 2303-2305.

Winder (1982). Mode of action of the antimycobacterial Agents and Associated Aspects of the Molecular Biology of the Mycobacteria. In *The Biology of the Mycobacteria*, pp. 354-442. Edited by C. Ratledge & J. Stanford. London: Academic Press.

Winder, F. G., Collins, P. B. & Whelan, D. (1971). Effects of ethionamide and isoxyl on mycolic acid synthesis in *Mycobacterium tuberculosis* BCG. *J Gen Microbiol* **66**, 379-380.

Witkowski, A., Joshi, A. K., Lindqvist, Y. & Smith, S. (1999). Conversion of a beta-ketoacyl synthase to a malonyl decarboxylase by replacement of the active-site cysteine with glutamine. *Biochemistry* **38**, 11643-11650.

Wong, H. C., Liu, G., Zhang, Y. M., Rock, C. O. & Zheng, J. (2002). The solution structure of acyl carrier protein from *Mycobacterium tuberculosis*. *J Biol Chem* **277**, 15874-15880.

Xin, Y., Lee, R. E., Scherman, M. S., Khoo, K. H., Besra, G. S., Brennan, P. J. & McNeil, M. (1997). Characterization of the *in vitro* synthesized arabinan of mycobacterial cell walls. *Biochim Biophys Acta* **1335**, 231-234.

Yadav, G., Gokhale, R. S. & Mohanty, D. (2003). Computational approach for prediction of domain organization and substrate specificity of modular polyketide synthases. *J Mol Biol* **328**, 335-363.

Yamagami, H., Matsumoto, T., Fujiwara, N., Arakawa, T., Kaneda, K., Yano, I. & Kobayashi, K. (2001). Trehalose 6,6'-dimycolate (cord factor) of *Mycobacterium tuberculosis* induces foreign-body- and hypersensitivity-type granulomas in mice. *Infect Immun* **69**, 810-815.

Yuan, Y. & Barry, C. E., 3rd (1996). A common mechanism for the biosynthesis of methoxy and cyclopropyl mycolic acids in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A* **93**, 12828-12833.

Yuan, Y., Zhu, Y., Crane, D. D. & Barry, C. E., 3rd (1998a). The effect of oxygenated mycolic acid composition on cell wall function and macrophage growth in *Mycobacterium tuberculosis*. *Mol Microbiol* **29**, 1449-1458.

Yuan, Y., Lee, R. E., Besra, G. S., Belisle, J. T. & Barry, C. E., 3rd (1995). Identification of a gene involved in the biosynthesis of cyclopropanated mycolic acids in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A* **92**, 6630-6634.

Yuan, Y., Crane, D. C., Musser, J. M., Sreevatsan, S. & Barry, C. E., 3rd (1997). MMAS-1, the branch point between *cis*- and *trans*-cyclopropane-containing oxygenated mycolates in *Mycobacterium tuberculosis*. *J Biol Chem* **272**, 10041-10049.

Yuan, Y., Mead, D., Schroeder, B. G., Zhu, Y. & Barry, C. E., 3rd (1998b). The biosynthesis of mycolic acids in *Mycobacterium tuberculosis*. Enzymatic methyl(ene) transfer to acyl carrier protein bound meromycolic acid *in vitro*. *J Biol Chem* **273**, 21282-21290.

Zhang, N., Torrelles, J. B., McNeil, M. R., Escuyer, V. E., Khoo, K. H., Brennan, P. J. & Chatterjee, D. (2003). The Emb proteins of mycobacteria direct arabinosylation of lipoarabinomannan and arabinogalactan via an N-terminal recognition region and a C-terminal synthetic region. *Mol Microbiol* **50**, 69-76.

Zhang, Y., Heym, B., Allen, B., Young, D. & Cole, S. (1992). The catalase-peroxidase gene and isoniazid resistance of *Mycobacterium tuberculosis*. *Nature* **358**, 591-593.

Zhang, Y. M., Rao, M. S., Heath, R. J., Price, A. C., Olson, A. J., Rock, C. O. & White, S. W. (2001). Identification and analysis of the acyl carrier protein (ACP) docking site on beta-ketoacyl-ACP synthase III. *J Biol Chem* **276**, 8231-8238.

Zimhony, O., Cox, J. S., Welch, J. T., Vilcheze, C. & Jacobs, W. R., Jr. (2000). Pyrazinamide inhibits the eukaryotic-like fatty acid synthetase I (FASI) of *Mycobacterium tuberculosis*. *Nat Med* **6**, 1043-1047.

Appendices

8.1 General materials and methods

8.1.1 Culture media

LB (Luria-Bertani) Broth

	Amount *per* Litre (g)
Bacto-tryptone	10
Bacto-yeast extract	5
NaCl	5

Sauton's Broth (pH 7.2)

	Amount *per* Litre
K_2HPO_4	0.5 g
$MgSO_4$	0.5 g
Asparagine	4 g
Fe Ammonium Citrate	0.05 g
Glycerol	60 ml
Citric Acid	2 g
ZnSO4 (1 % w/v)	0.1 ml
Triton (Tyloxapol) (5 %)	5 ml

LB agar was purchased from Sigma and prepared using manufacturers directions. All media is to be autoclaved at 121°C for 15 minutes.

Terrific Broth

Solution 1	Amount *per* 900 ml
Bacto-tryptone	12 g
Bacto-yeast extract	24 g
Glycerol	4 ml

Solution 2	Amount *per* 100 ml
KH_2PO_4	2.3 g
K_2HPO_4	12.5 g

The above solutions were autoclaved at 121°C for 15 minutes. Solution 2 for terrific broth was added to solution 1 prior to use, with the appropriate antibiotic.

8.1.2 Growth of bacterial strains

Batch culturing of the *E. coli* stains used in the project were carried out at 37°C or 16°C in an orbital incubator at a speed of 180 rpm; length of incubation varied depending on the experiment. In broth and solid medium the following antibiotics were used:

- Kanamycin (25 μg/ml) to select for *E. coli* containing the pET28a plasmid vector.
- Ampicilin (100 μg/ml) to select for *E. coli* containing the pET23b plasmid vector.
- Ampicilin (100 μg/ml) to select for *E. coli* containing the pQE60 plasmid vector.
- Ampicilin (100 μg/ml) to select for *E. coli* containing the pUC18 plasmid vector.
- Hygromycin (200 μg/ml) to select for *E. coli* containing either pSD26 or pVV16 plasmid vectors.
- Hygromycin (50 μg/ml) to select for *M. smegmatis* containing either pSD26 or pVV16 plasmid vectors.

8.1.3 Cell lines

Escherichia coli C41 (DE3)

Escherichia coli TOP 10

Escherichia coli HB101

Escherichia coli XL-1 Blue Super competent (Stratagene)

Escherichia coli M15 (pREP4)

Mycobacterium smegmatis mc^2155

8.1.4 Plasmid extraction

40 ml of LB broth was inoculated with the appropriate glycerol stock. After incubation at 37°C overnight 5 mls of the cells were harvested by centrifugation at 12,500 x *g* for 15 mins at 4°C. Plasmid DNA was prepared using the Qiagen Plasmid purification kit.

8.1.5 Polymerase chain reaction (PCR)

The recipe for PCR was similar for all the genes cloned but slight variations are used in each case to create the best yield. Four PCR mixtures are possible and are shown below.

Each PCR reaction has a final volume of 100 µl made up with H_2O, one reaction contains the following ingredients with the addition of one of the methods making a final volume of 99 µl.

2 µl	5′ Cloning Primer
2 µl	3′ Cloning Primer
1 µl	Genomic DNA
2 µl	dNTP's
10 µl	Thermostable Buffer (10x)

Method 1	Method 2	Method 3	Method 4
0 µl DMSO	8 µl DMSO	0 µl DMSO	8 µl DMSO
0 µl $MgSO_4$	0 µl $MgSO_4$	2 µl $MgSO_4$	2 µl $MgSO_4$
82 µl H_2O	74 µl H_2O	80 µl H_2O	72 µl H_2O

$MgSO_4$ = 100 mM

1 µl of the Vent™ was added prior to starting the PCR followed by gentle mixing.

8.1.6 Gel electrophoresis

8.1.6.1 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

Protein levels were assessed using a bicinchoninic protein assay (BCA) Kit (Pierce). Protein fractions (crude 25 µg, purified 5 µg) were analysed by electrophoresis on 10 - 15 % SDS polyacrylamide gels by the method of Laemmli (1970) using a Hoefer "mighty small" SE200 vertical slab gel system. The gels were made with the following ingredients: -

	Amount required for 2 gels
4 x Resolving gel buffer	3.75 ml
Acrylamide / bis / water mix	11.25 ml
TEMED	30 µl
10 % ammonium persulphate	75 µl

(Resolving gel buffer 1.5 M Tris-HCl, 0.4 % SDS pH 8.8). The acrylamide / bis / water mix contained the following concentrations depending on the percentage acrylamide required, normally percentages used ranged from 10 % to 15 %: -

%	8	10	12	15
Acrylamide (30 %, 37.5:1 acrylamide/bis-acrylamide)	4.0	5.0	6.0	7.5
Water	7.25	6.25	5.25	3.75

The stacking Gel contained the following ingredients: -

	Amount required for 2 gels
4 x stacking gel buffer	1.25 ml
Acrylamide	0.65 ml
Water	3.05 ml
TEMED	15 µl
10 % ammonium persulphate	75 µl

(Stacking gel buffer 0.5 M Tris-HCl, 0.4 % SDS pH 6.8). Protein samples were added to 1/5 of SDS loading buffer (360 mM Tris-HCl, pH 8.8, 9 % (w/v) SDS, 0.9 % (w/v) bromophenol blue, 15 % (w/v) β-mercaptoethanol and 30 % glycerol) and boiled for 5 minutes. Samples were loaded onto the gel and run at 20 mA for 60 minutes. The running buffer (25 mM Tris, 190 mM Glycine, 4 mM SDS) were applied as *per* the manufacturers instructions.

The proteins were visualised by staining the gels in 50 ml of Coomassie brilliant blue R (CBB) stain (5 % glacial acetic acid, 0.025 % CBB (w/v), 50 % methanol) and left for 0.5-1.0 hour on an oscilating platform. Excess stain that was not bound was removed by rinsing in destaining solution (10 % glacial acetic acid in 30 % methanol).

8.1.6.2 Electrophoresis of DNA

DNA fragments from restriction reactions and PCR products were analysed by horizontal slab agarose gel electrophoresis (Hoefer). 1 % (w/v) Agarose gels were prepared by dissolving electrophoresis grade agarose in TAE. DNA samples were prepared by mixing with one-fifth the volume of 5X DNA loading buffer (0.25 % bromophenol blue, 0.25 % xylene cyanol FF, 40 % (w/v) sucrose in water). All agarose gels were run in 1X TAE (40 mM Tris-acetate, 1 mM EDTA) buffer. The samples were run at 100 volts until marker dye migrated the required distance. The gel was then stained with a 0.5 μg/ml solution of ethidium bromide for 20 mins, then viewed by UV transillumination and photographed using a gel documentation system (Bio-Rad). Relevant band sizes were excised from the gel and DNA fragments extracted using the Qiagen Gel extraction kit.

8.1.7 Restriction enzyme digest of DNA

All restriction endonuclease digests were performed under conditions recommended by the manufacturer. For example, the final volume of these reactions equate to 20 μl where the following volumes were added to the reaction: -

2 µl sure-cut buffer (x10)
1 µl restriction endonulease A
1 µl restriction endonulease B
2 ng DNA

DNase and RNase-free water was added to make the volume up to 20 µl. The digest conditions were 37°C for 2 hours and the products then analysed on a 1 % agarose gel.

8.1.8 Transformation of *E. coli* with plasmid DNA

8.1.8.1 Electroporation

To 50 µl of electrocompetent cells, 1 ng of plasmid DNA was added and mixed gently. This was transferred to an electro-cuvette and stored on ice for 5 minutes. Cells were electroporated using a Biorad Pulse controller and Biorad gene pulser with 2 mm cuvette gaps at the following conditions:

Resistance = 200 Ω
25 µFD
2.5 V

Immediately after electroporation, 250 µl of sterile LB broth was added to the cuvette, mixed gently and the incubated at 37°C for 1 hour. The culture was plated out onto LB agar containing the relevant antibiotic and cultured overnight at 37°C.

8.1.8.2 Heat shock

To 50 µl of competent cells, 1 ng of plasmid DNA was added in Falcon 2054 tubes, gently mixed and incubated on ice for 30 minutes. The mixture was heat shocked at 42°C for 45 seconds, and replaced on ice for 5 minutes. To allow the expression host to recover, 250 µl of fresh LB broth was added and incubated at 37°C for 1 hour. The culture was spread on LB agar containing the relevant antibiotic and cultured overnight at 37°C.

8.1.9 Induction

The induction of protein synthesis described in chapters 2, 3 and 5 were performed using IPTG for pET/pQE vectors and acetamide for the pSD26 vector. A single transformant colony from the plates was used to inoculate 5 ml of LB Broth (pET/pQE) and Sauton's broth (pSD26/pVV16) containing the relevant antibiotic and incubated at 37°C overnight.

The overnight culture was added to fresh media containing the relevant antibiotic and incubated at 37°C in an orbital incubator until the OD_{600} was 0.5 → 0.7. Cells were induced by the addition of varying concentrations of IPTG (0.1 mM – 1 mM) or acetamide (0.01 %-0.2 %) followed by incubation for a further 4 hours, in some cases there was call for extended lengths of induction. The cells were harvested by centrifugation at 27,000 x *g* for 10 minutes and the supernatant discarded. Cell pastes were stored at -20°C.

8.1.10 Extraction of recombinant proteins

8.1.10.1 Sonication

The pelleted cells from the induction experiments were re-suspended in 4 volumes of 20 mM phosphate buffer pH 7.5. The cells were lysed by sonication in a Soniprep 150, the suspension was stored on ice during the process and sonicated for 5-8 cycles (*E. coli* 1 cycle = 30 sec on, 30 sec off) (*M. smegmatis* 1 cycle = 60 sec on, 90 sec off). An aliquot of the resulting sample was store at 4°C for analysis and the rest of the cell debris pelleted at 27,000 x *g*. The crude lysate, supernatant and the pellet analysed by SDS-PAGE electrophoresis. In some cases the pellet was resuspended in 20 mM phosphate buffer pH 7.5 containing 8 M urea followed by re-centrifugation.

8.1.10.2 French pressure cell

The French pressure cell was cooled to 4°C for 2 hours before use. Samples were prepared by re-suspension in the appropriate buffer solution and then homogenised in a glass homogeniser. The slurry was then passed through the French press set at 3,000 psi. To ensure complete cell lysis, the homogenate was the passed through the French press again. DNase I (100 µg) was added *per* gram wet weight of original cells and mixed for 5 minutes to reduce the viscosity. The cell debris was pelleted at 27,000 x *g* at 4°C. The crude lysate, supernatant and the pellet were analysed by SDS-PAGE. Again in some cases the pellet was resuspended in 20 mM phosphate buffer pH 7.5 containing 8 M urea followed by re-centrifugation.

8.1.11 Purification

Hi-trap Chelating Sepharose™ fast flow matrix columns (1 ml or 5 ml) were equilibrated by washing with 4x column volumes of Hi-Trap running buffer (0.02 M Sodium Phosphate, 0.5 M NaCl pH7.4). The column was charged using 2 column volumes of 0.1 M nickel chloride. The column was further washed with 4 column volumes of running buffer to re-equilibriate the column and remove excess nickel. The supernatant was applied to the column (maximum of 10 ml for a 1 ml column and 50 ml for a 5 ml column) and non-absorbed material collected for analysis column. Proteins were eluted using increasing concentrations of imidazole (10 mM – 500 mM). Fractions were collected and analysed by SDS-PAGE following protein level assessment using a BCA assay kit (Pierce). Samples were dialysed against the relevant dialysis buffer eg. mtFabH against 20 mM Tris-HCl pH 7.9, 50 mM NaCl, 1 mM β-mercaptoethanol.

8.1.12 Southern blotting

DNA hybridisation was performed as described by (Southern, 1975). After gel electrophoresis the DNA was denatured by washing the agarose gel in denauturation solution (0.5 M NaOH, 1.5 M NaCl) for 45 mins and then neutralised by washing in neutralsiation solution (0.5 M Tris-HCl pH 7.4, 1.5 M NaCl) for 45 minutes. Finally the gel was rinsed in 20x SSC solution (3 M NaCl and 0.3 M trisodium citrate). The DNA was then transferred to Hybond N+ nylon membrane (Amersham) by capillary action. The cling film has a window shape of the gel removed to allow the moisture to be drawn upwards. This blotting procedure was allowed to

proceed overnight. The blot was then dismantled and the nylon membrane air-dried. The membrane was then visualised by DNA hybridisation.

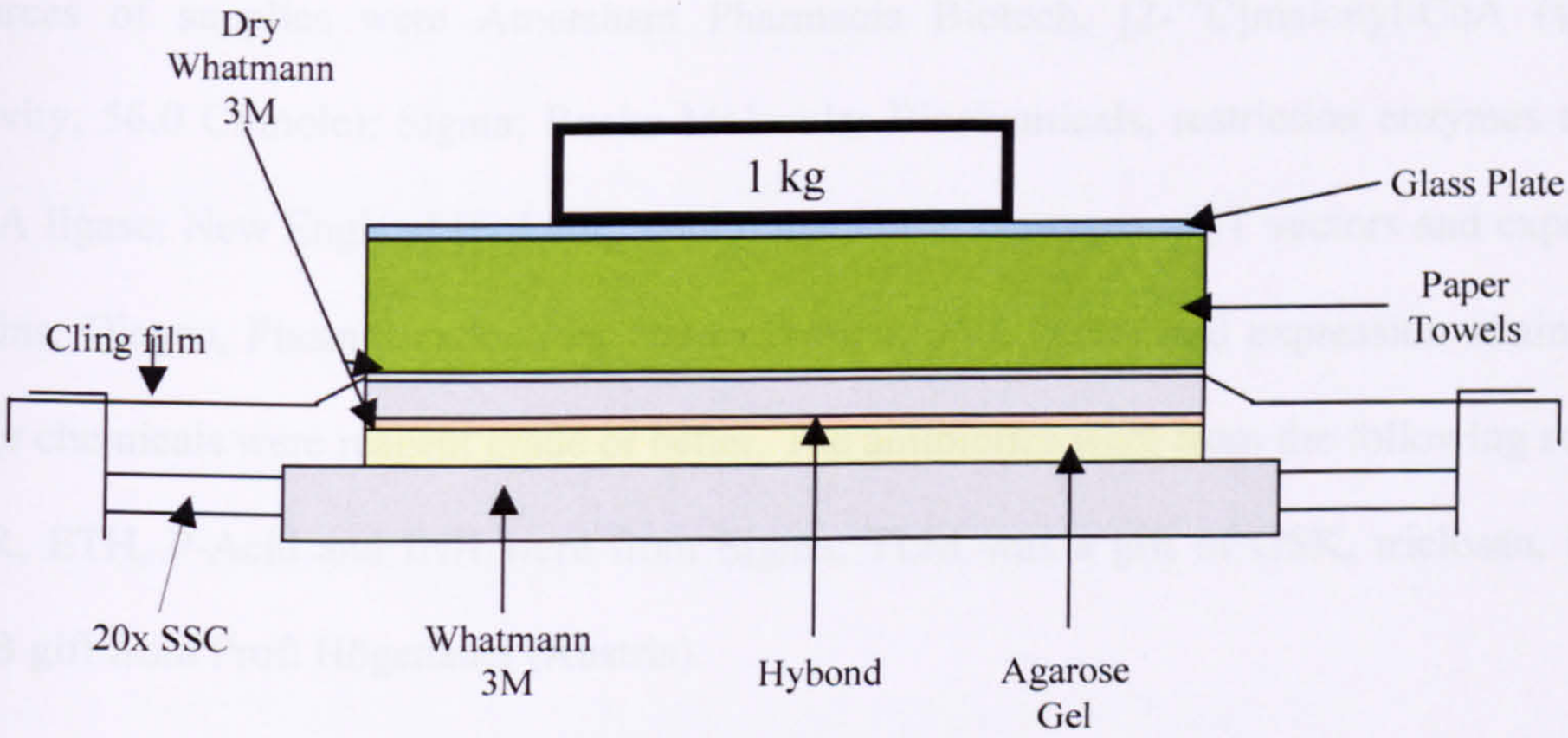


Figure 8.1 Schematic of Southern blotting technique.

8.1.13 DNA hybridisation

The Hybond N+ nylon membranes were overlaid onto wetted nylon supports. Membranes were then placed in glass hybridisation bottles (Hybaid) with 5-10 ml EasyHyb solution (Amersham) depending on the size of the membrane. Pre-hybridisation of membrane easyHyb was done at 42ºC for a minimum of 1 hour in a rotating hybridisation oven (Hybaid). The probes used were labelled with digoxigenin (DIG)-dUTP5' using the PCR DIG probe synthesis kit (Boehringer Mannheim). Denatured labelled DNA probe was added to pre-hybridised membranes and allowed to hybridise in the presence of pre-hybridisation buffer at 50ºC for a minimum of 2 hours. Post-hybridisation washes were performed twice in 2x SSC, 0.1 % SDS for 5 min at room temperature and twice in 0.1x SSC, 0.1 % SDS for 15 min at 68ºC. the hybridised probe was detected using a DIG luminescence detection kit (Boehringer Mannheim) according to the manufacturer's recommendations.

8.1.14 Materials

Sources of supplies were Amersham Pharmacia Biotech, [2-^{14}C]malonyl-CoA (specific activity, 56.0 Ci/mole); Sigma; Roche Molecular Biochemicals, restriction enzymes and T4 DNA ligase; New England BioLabs, Vent polymerase; Novagen, pET vectors and expression strains; Qiagen, Plasmid extraction, Gel extraction, pQE vector and expression strains. All other chemicals were reagent grade or better. The antibiotics were from the following sources: CER, ETH, P-Acid and INH were from Sigma, TLM was a gift of GSK, triclosan, isoxyl, DZB gift from Prof. Högenauer (Austria).

8.2 KasA Cloning Primers

pET28a

Upper (5'-GATCGATCGCTAGCATGAGTCAGCCTTCCACCG-3')
EcoRI

Lower (5'-GACTGATCGCTAGCCTGCTTGCCTACCTCACTTG- 3')
NheI

pET23b

Upper (5'-GATCGATCGCTAGCATGAGTCAGCCTTCCACCG-3')
NheI

Lower (5'-GACTGATCGCGGCCGCGTAACGCCCGAAGGCAAGC- 3')
NotI

pVV16

Upper (5'-GATCGATCAAGCTTATGAGTCAGCCTTCCACCG-3')
HindIII

Lower (5'-GACTGATCAAGCTTGTAACGCCCGAAGGCAAGC- 3')
HindIII

pQE60

Upper (5'-GATCGATCCCATGGATGAGTCAGCCTTCCACCG-3')
NcoI

Lower (5'-GACTGATCAGATCTGTAACGCCCGAAGGCAAGC- 3')
BglII

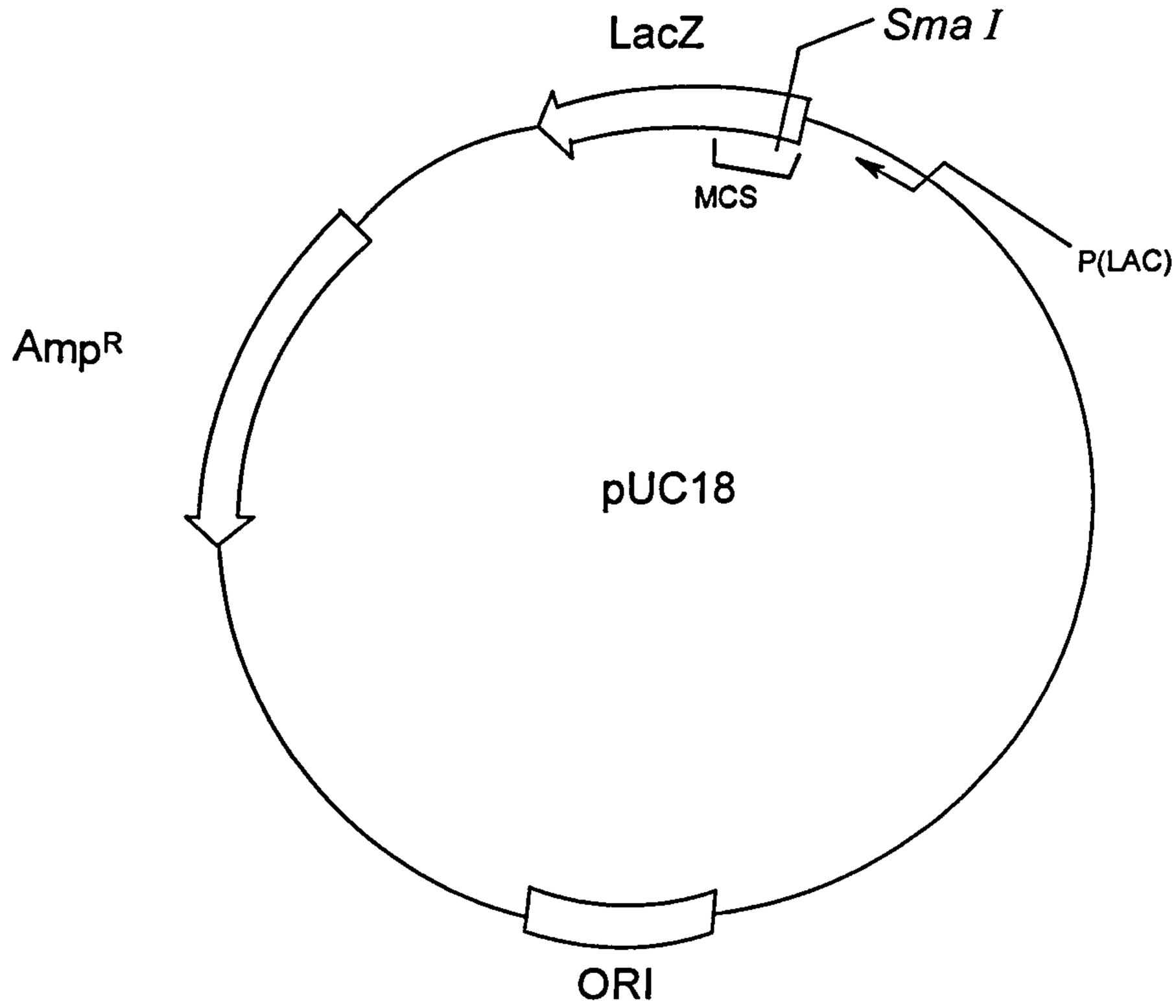
pSD26

Upper (5'-GACTGATCGATATCATGAGTCAGCCTTCCACCG-3')
EcoRV

Lower (5'-GACTGATCGATATCGTAACGCCCGAAGGCAAGC- 3')
EcoRV

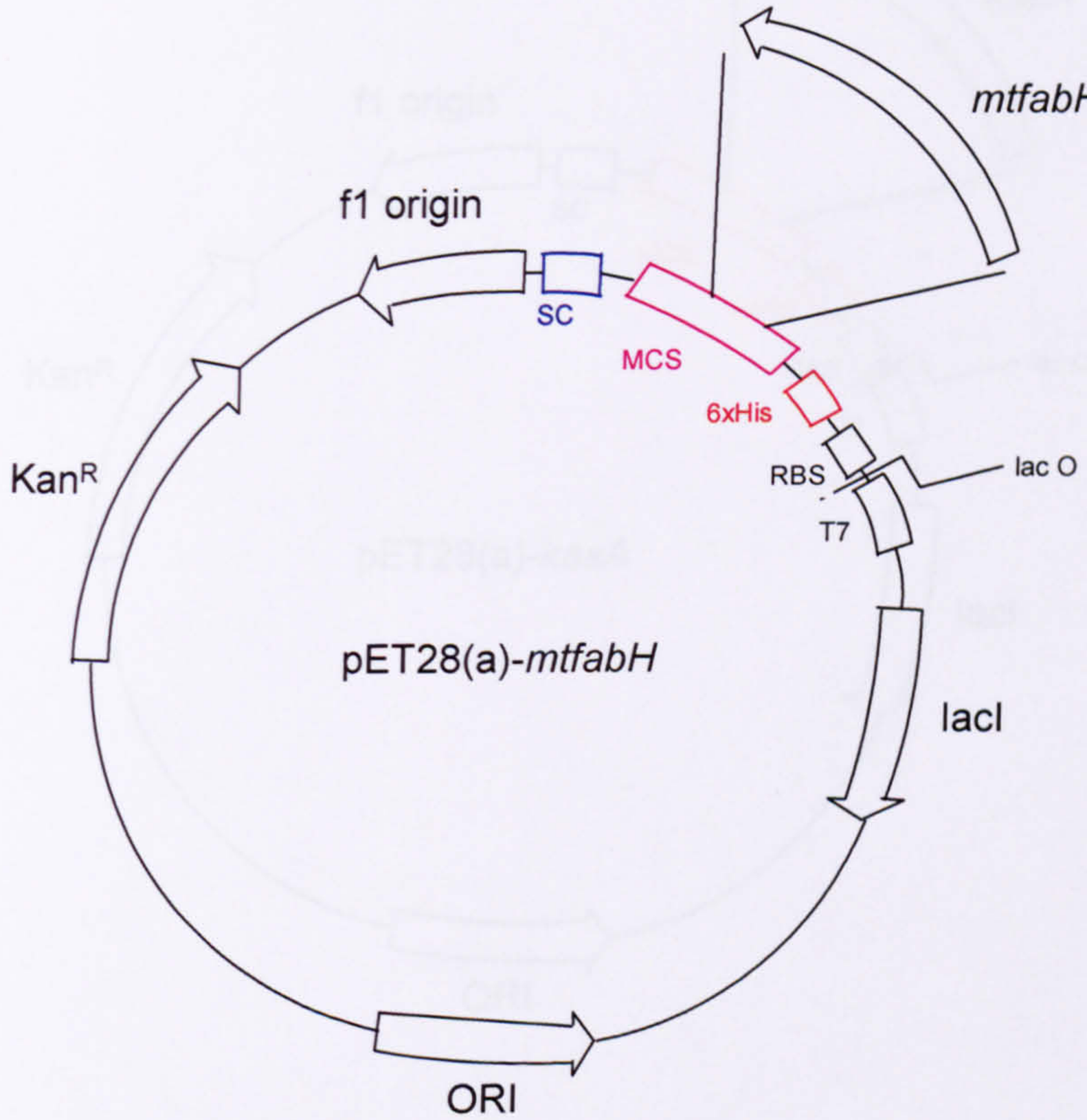
8.3 Plasmid Diagrams

8.3.1 pUC18



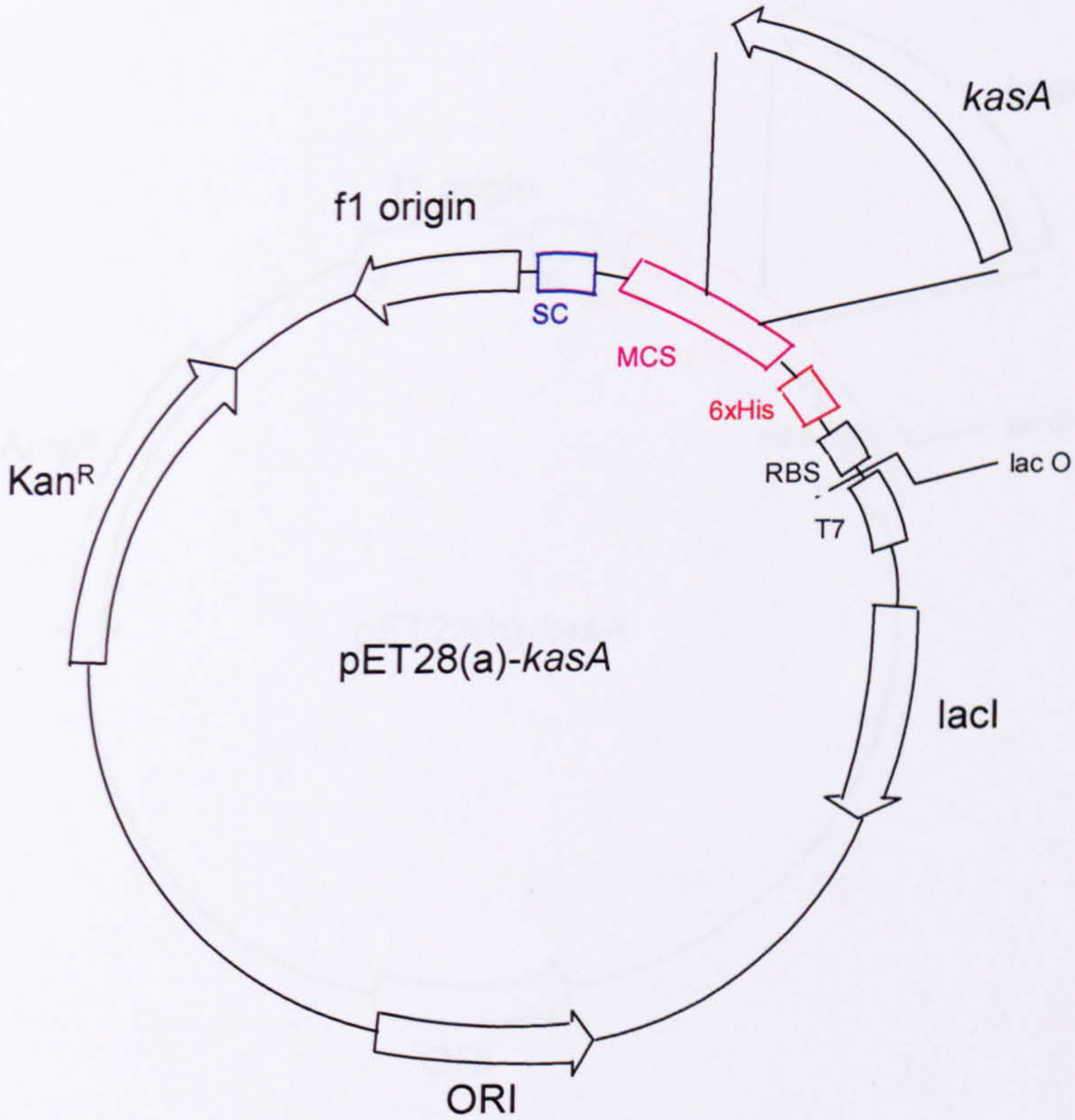
pUC18 vector for cloning. **Plac:** lac promotor, **MCS:** Multiple cloning site, **ORI:** origin of replication (pMB1 replicon), **Amp^R:** ampicillin resistance gene (*Bla* β-lactamase), **LacZ:** encodes the N-terminal fragment of β-galactosidase, **Sma I:** Restriction site.

8.3.2 pET28(a)-*mtFabH*



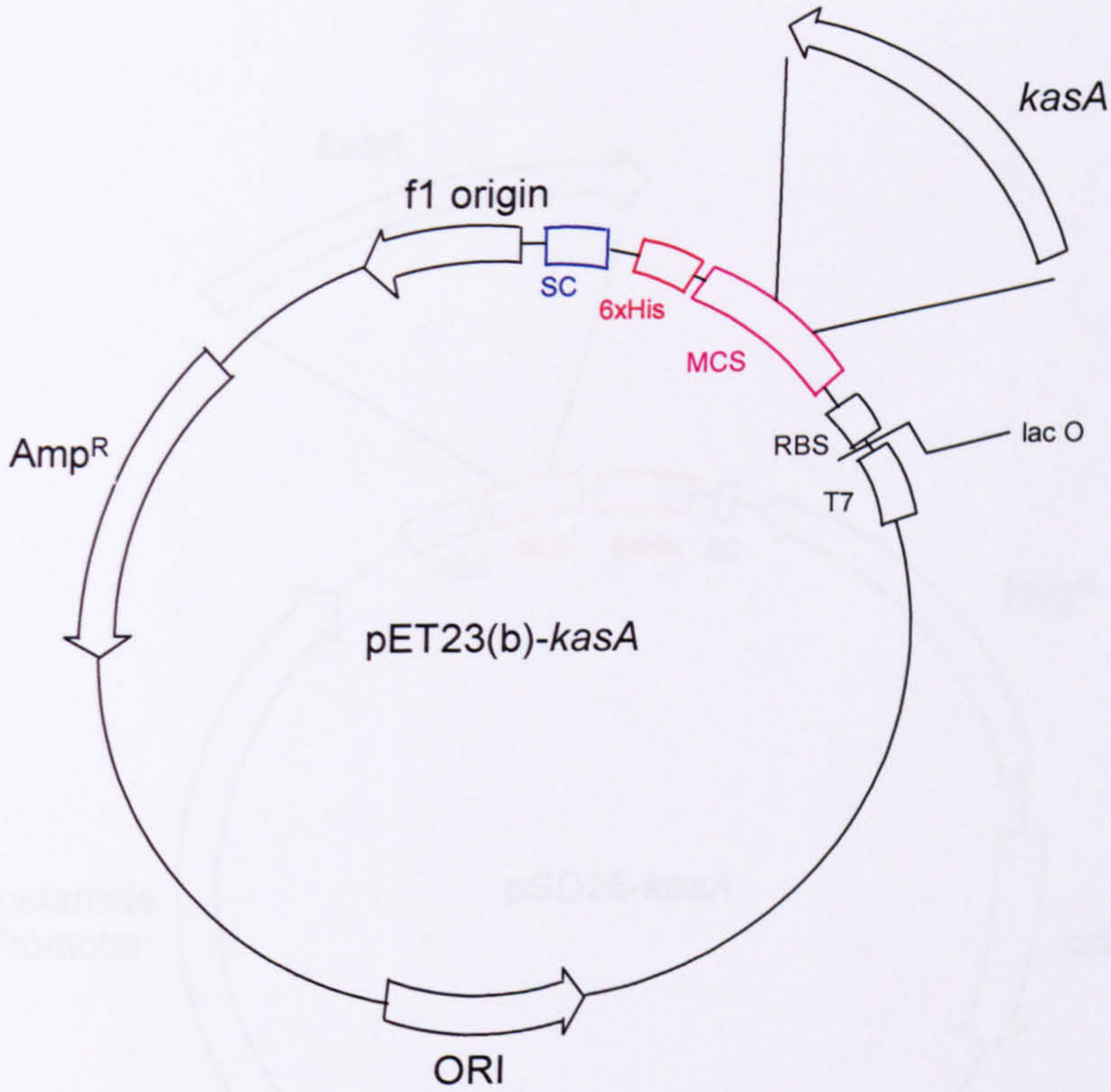
pET28(a) vector for N-terminal 6xHis tag constructs. **T7**: T7 promoter, **lac O**: lac operator, **lac I**: lac regulator, **F1 origin**: Helper Phage origin of replication, **RBS**: ribosome-binding site, **6xHis**: 6xHis tag sequence, **MCS**: Multiple cloning site, **SC:** stop codons in all three reading frames, **ORI**: origin of replication, **Kan^R**: Kanamycin resistance gene (*Neo*: gene from Tn5 encodes an aminoglycoside 3'phosphotransferase (3'APH II) that confers resistance to the antibiotics kanamycin in bacteria). (Novagen, Wisconsin, USA)

8.3.3 pET28(a)-*kasA*



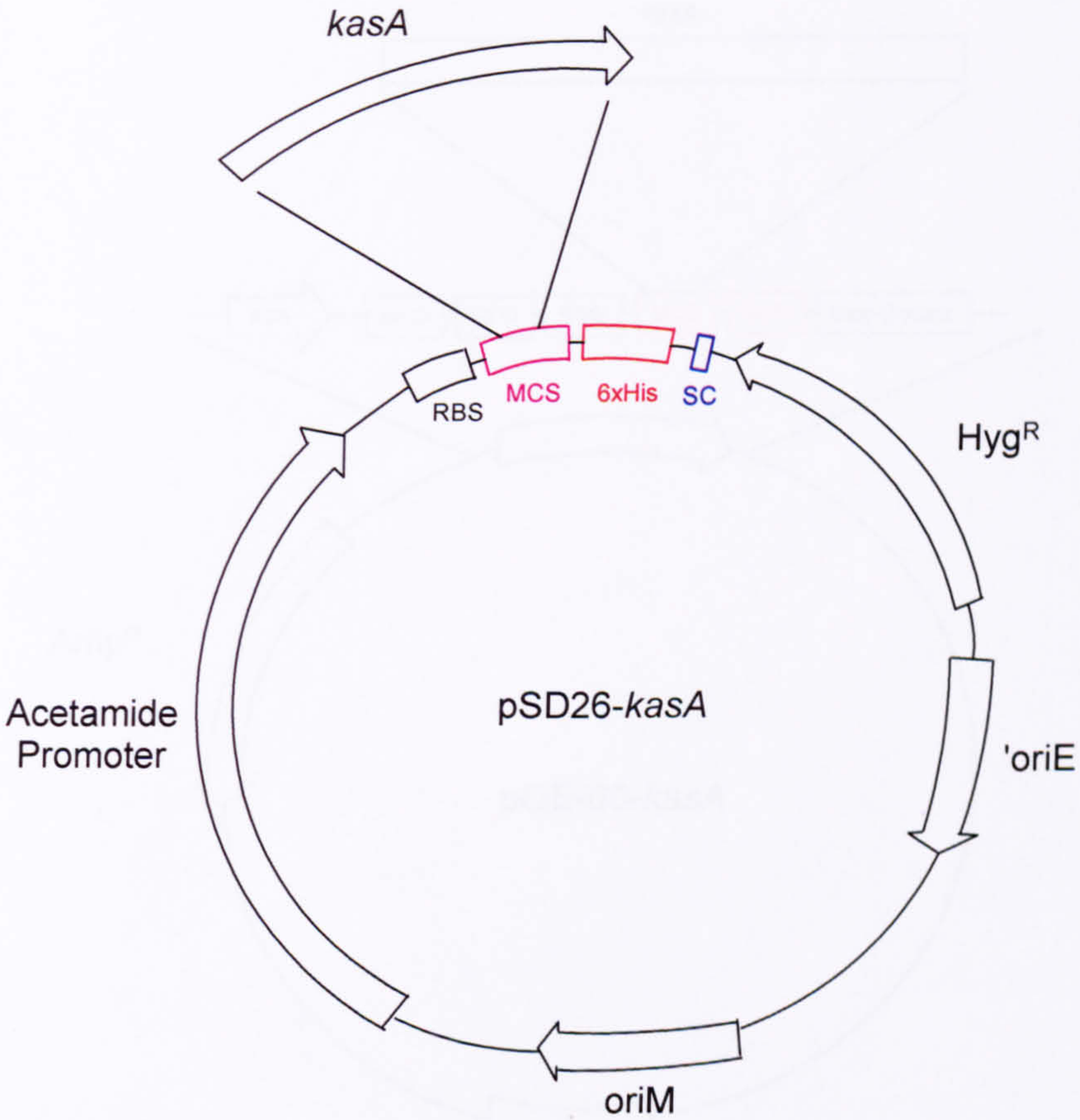
pET28(a) vector for N-terminal 6xHis tag constructs. **T7**: T7 promoter, **lac O**: lac operator, **lac I**: lac regulator, **F1 origin**: Helper Phage origin of replication, **RBS**: ribosome-binding site, **6xHis**: 6xHis tag sequence, **MCS**: Multiple cloning site, **SC:** stop codons in all three reading frames, **ORI**: origin of replication, **Kan^R**: Kanamycin resistance gene (*Neo*: gene from Tn5 encodes an aminoglycoside 3'phosphotransferase (3'APH II) that confers resistance to the antibiotics kanamycin in bacteria). (Novagen, Wisconsin, USA)

8.3.4 pET23(b)-*kasA*



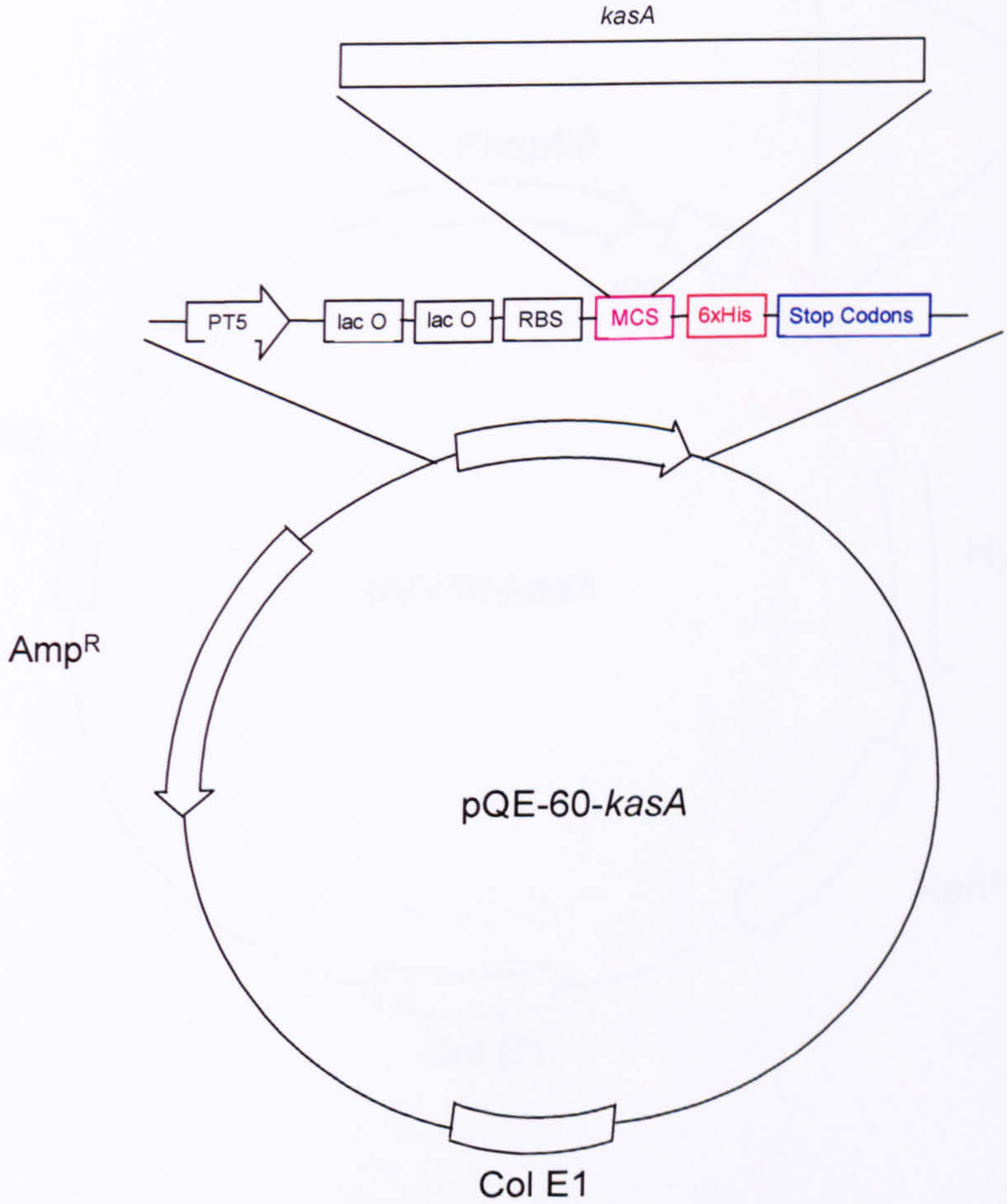
pET23(a) vector for C-terminal 6xHis tag constructs. **T7**: T7 promoter, **lac O**: lac operator, **lac I**: lac regulator, **F1 origin**: Helper Phage origin of replication, **RBS**: ribosome-binding site, **6xHis**: 6xHis tag sequence, **MCS**: Multiple cloning site, **SC:** stop codons in all three reading frames, **ORI**: origin of replication, **Amp^R**: Ampicillin resistance gene (*Bla* β-lactamase) (Novagen, Wisconsin, USA).

8.3.5 pSD26-*kasA*



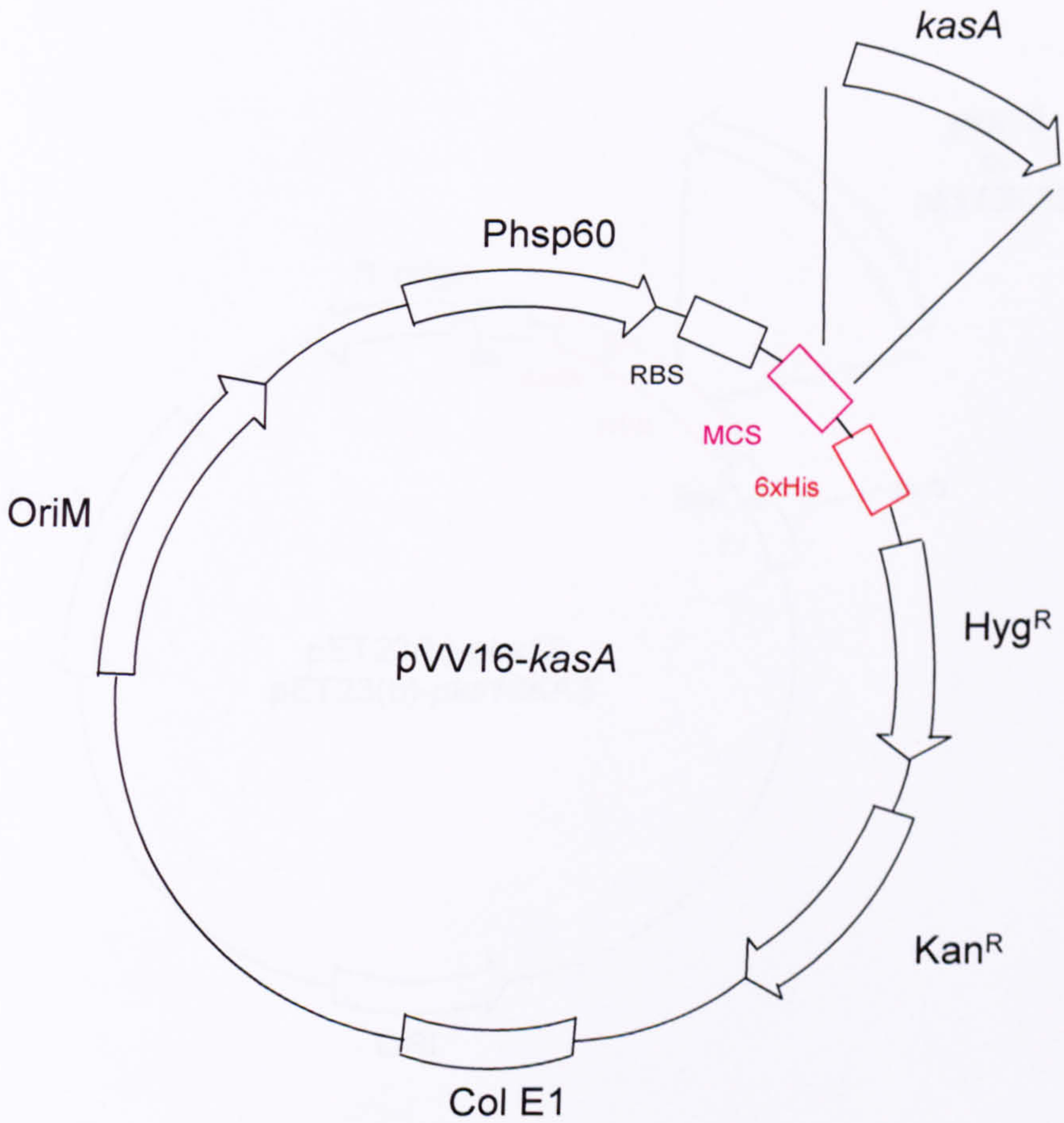
pSD26 vector for C-terminal 6xHis tag constructs. **Acetamide Promoter**: promoter **RBS**: ribosome-binding site, **6xHis**: 6xHis tag sequence, **MCS**: Multiple cloning site, **SC:** stop codons in all three reading frames, **oriM**: mycobacterial origin of replication, **'oriE**: *E. coli* origin of replication **Hyg^R**: hygromycin resistance gene (phosphotransferase).

8.3.6 pQE-60-*kasA*



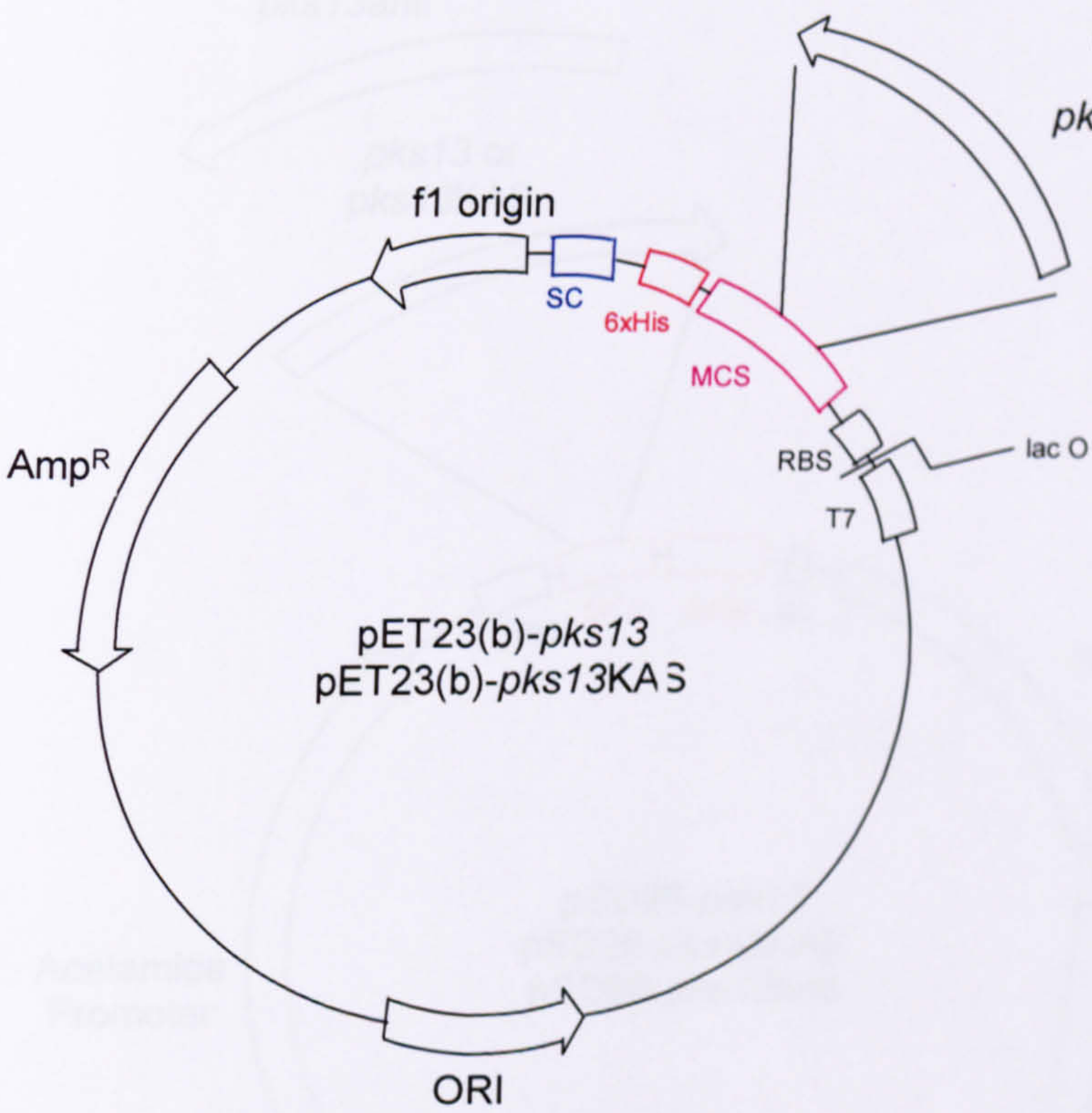
pQE vectors for C-terminal 6xHis tag constructs. **PT5**: T5 promoter, **lac O**: lac operator, **RBS**: ribosome-binding site, **6xHis**: 6xHis tag sequence, **MCS**: Multiple cloning site, **Stop codons**: stop codons in all three reading frames, **Col E1**: Col E1 origin of replication, **Amp^R**: ampicillin resistance gene (*Bla* β-lactamase).

8.3.7 pVV16-*kasA*



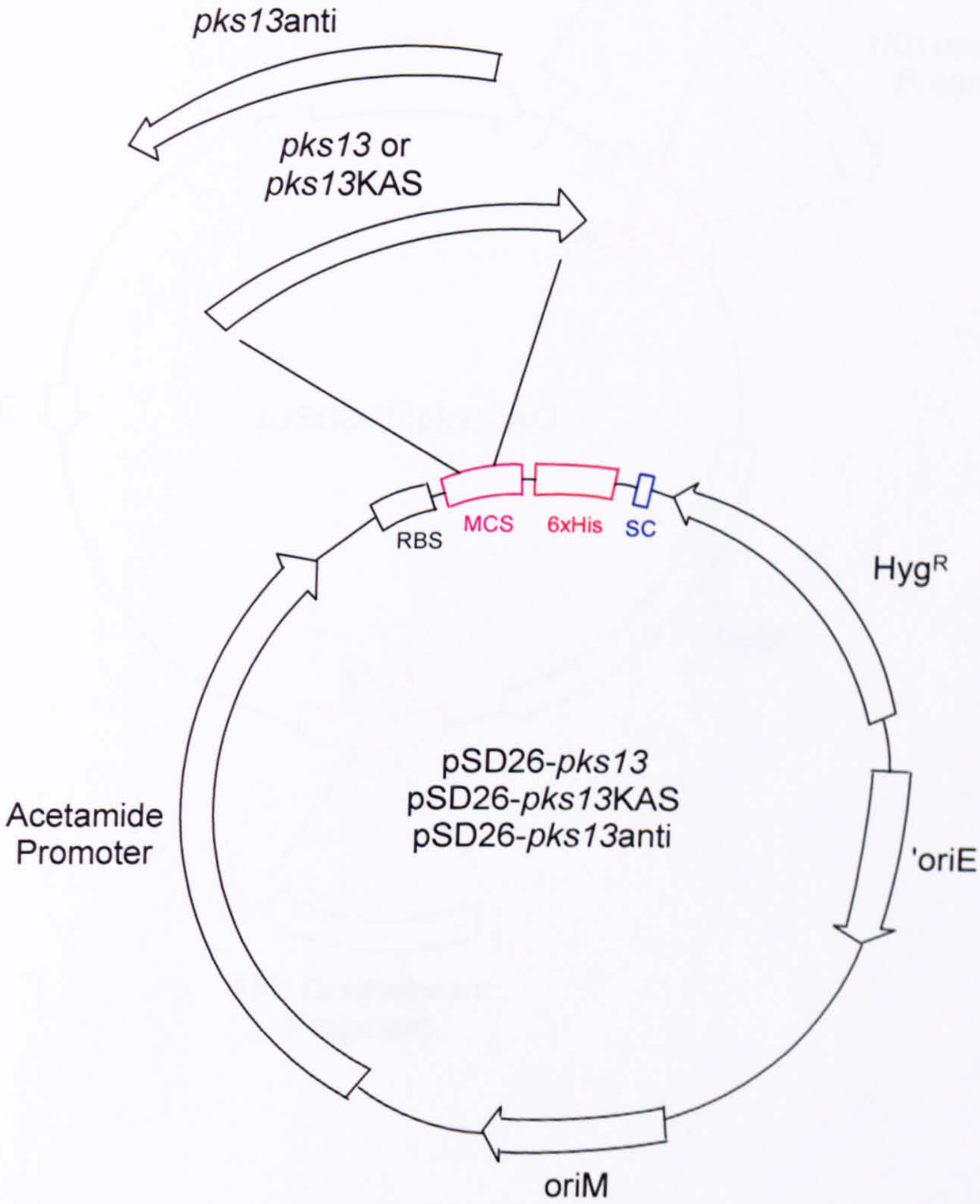
pVV16 vector for C-terminal 6xHis tag constructs in *Mycobacterium*. **Phsp60**: mycobacterial promoter, **OriM**: mycobacterial origin of replication, **RBS**: ribosome-binding site, **6xHis**: 6xHis tag sequence, **MCS**: Multiple cloning site, **Col E1**: Col E1 origin of replication, **Kan^R**: Kanamycin resistance gene (*Neo*: gene from Tn5 encodes an aminoglycoside 3'phosphotransferase (3'APH II) that confers resistance to the antibiotics kanamycin in bacteria), **Hyg^R**: Hygromycin resistance gene (phosphotransferase).

8.3.8 pET23(b)-*pks13* and pET23(b)-*pks13KAS*



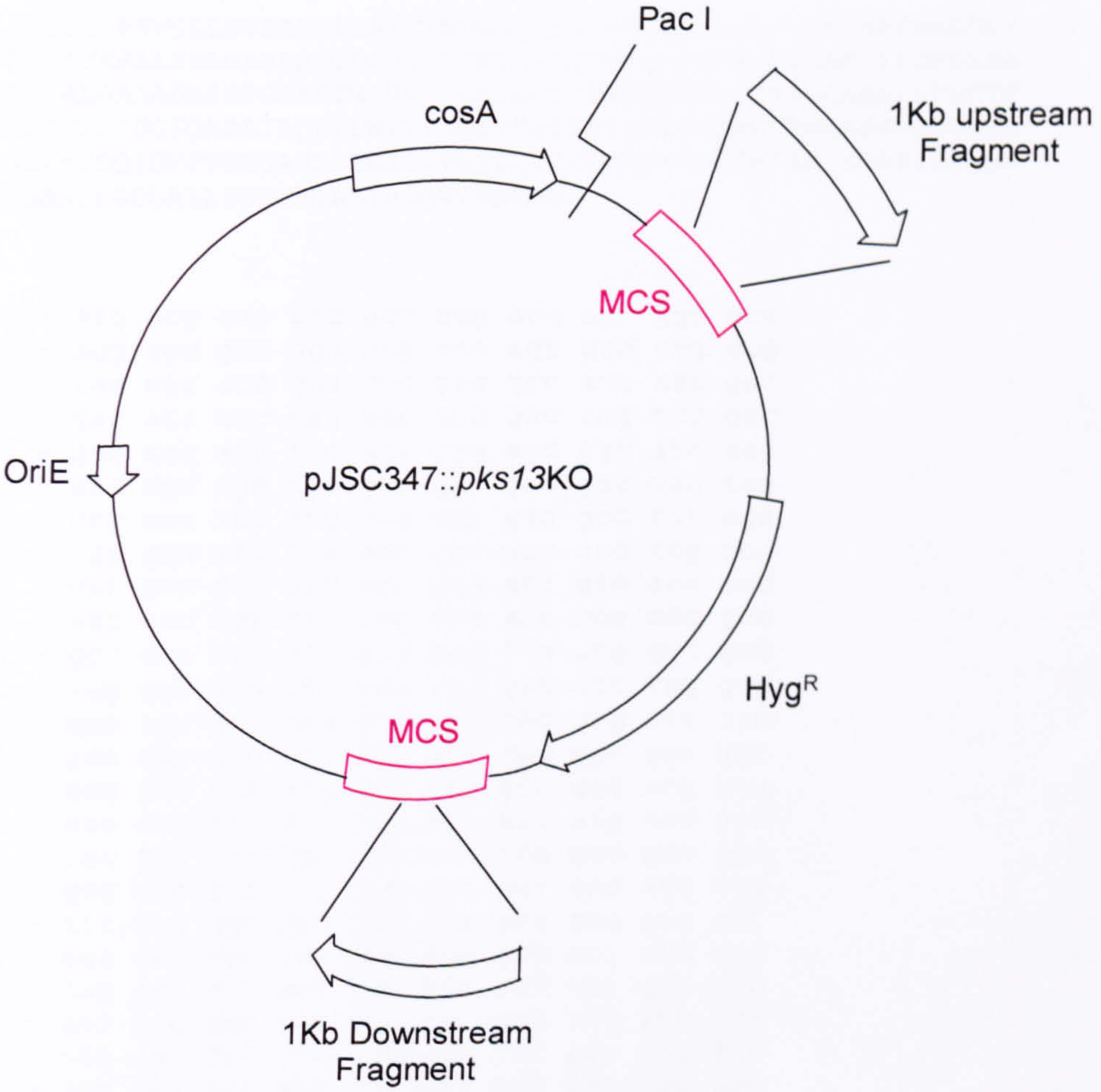
pET23(a) vector for C-terminal 6xHis tag constructs. **T7**: T7 promoter, **lac O**: lac operator, **lac I**: lac regulator, **F1 origin**: Helper Phage origin of replication, **RBS**: ribosome-binding site, **6xHis**: 6xHis tag sequence, **MCS**: Multiple cloning site, **SC:** stop codons in all three reading frames, **ORI**: origin of replication, **Amp^R**: Ampicillin resistance gene (*Bla* β-lactamase).

8.3.9 pSD26-*pks13*, pSD26-*pks13KAS* and pSD26-*pks13Anti*



pSD26 vector for C-terminal 6xHis tag constructs. **Acetamide Promoter**: promoter **RBS**: ribosome-binding site, **6xHis**: 6xHis tag sequence, **MCS**: Multiple cloning site, **SC:** stop codons in all three reading frames, **oriM**: mycobacterial origin of replication, **'oriE**: *E. coli* origin of replication **Hyg^R**: hygromycin resistance gene (phosphotransferase).

8.3.10 pJSC347::*pks13*KO



pJSC347 vector for creation of knock outs. **Pac I**: restriction site for packaging, **MCS**: Multiple cloning site, **OriE**: *E. coli* origin of replication **Hyg^R**: hygromycin resistance gene (phosphotransferase).

8.4 Protein/DNA Sequences

8.4.1 mtFabH (Rv0533c) (1008 bp) (335 aa) (34872.5 da) (Theoretical pI: 4.98)

MTEIATTSGARSVGLLSVGAYRPERVVTNDEICQHIDSSDEWIYTRTGIKTRRFAADDES
AASMATEACRRALSNAGLSAADIDGVIVTTNTHFLQTPPAAPMVAASLGAKGILGFDLSA
GCAGFGYALGAAADMIRGGGAATMLVVGTEKLSPTIDMYDRGNCFIFADGAAAVVVGETP
FQGIGPTVAGSDGEQADAIRQDIDWITFAQNPSGPRPFVRLEGPAVFRWAAFKMGDVGRR
AMDAAGVRPDQIDVFVPHQANSRINELLVKNLQLRPDAVVANDIEHTGNTSAASIPLAMA
ELLTTGAAKPGDLALLIGYGAGLSYAAQVVRMPKG

```
  1 - atg acg gag atc gcc acg acc agc ggc gcc
 31 - agg agc gtc ggg ctg ctc agt gtc ggg gcg
 61 - tac cgg ccc gaa cgc gtg gtc acc aac gac
 91 - gag ata tgc cag cac atc gac tcg tcc gac
121 - gag tgg atc tac acc cga acc ggc atc aag
151 - acc cgc cga ttc gcc gcc gac gac gag tcg
181 - gcg gct tcc atg gcg act gag gcc tgt cga
211 - cgg gca ctg tcg aac gcc ggc ctg tcg gcg
241 - gcc gac atc gat ggc gtg atc gtc acc acc
271 - aac acc cat ttc ctg caa acc ccg ccg gcc
301 - gcc cca atg gtc gcg gcg tcg ctg ggc gcc
331 - aag ggc ata ctc ggg ttc gat ctt tcg gcg
361 - ggg tgc gcc gga ttc gga tat gcg ctt ggc
391 - gca gcg gcc gac atg atc cgg ggc gga ggt
421 - gcg gcc acg atg ctg gtg gtc ggc acg gaa
451 - aaa ctg tcc ccc acg ata gac atg tac gac
481 - cgc ggc aac tgc ttc atc ttc gcc gac ggc
511 - gcg gcc gca gtg gtg gtg ggc gag aca ccg
541 - ttt caa ggc att gga cca acc gtg gcg ggt
571 - agc gac ggc gaa cag gcc gat gcc ata cgg
601 - cag gac atc gac tgg atc act ttc gcc cag
631 - aat ccc agc ggc cca cgc ccg ttt gtg cgg
661 - ctc gaa ggt ccc gcg gtc ttc cgt tgg gca
691 - gcg ttc aaa atg ggc gac gtc ggt cgg cgc
721 - gcg atg gac gcc gcc ggg gtg cga ccc gac
751 - cag ata gac gtg ttc gtc cct cat cag gcc
781 - aat agc cgc atc aac gag ctg ctg gtc aag
811 - aac ctg cag ttg cgg ccc gac gcg gtg gtc
841 - gcc aac gat atc gag cac acc gga aac acc
871 - tcg gcg gcc tcc att ccg ctc gcg atg gcc
901 - gaa tta ctg acg acc ggc gcg gcc aag ccc
931 - ggc gat ctg gcc ctg ttg atc ggc tac ggc
961 - gcc ggt ctg agc tat gcc gcc cag gtg gtg
991 - cga atg ccg aag ggt tga
```

8.4.2 KasA (Rv2245) (1251 bp) (416 aa) (43284.0 da) (Theoretical pI: 5.11)

MSQPSTANGGFPSVVVTAVTATTSISPDIESTWKGLLAGESGIHALEDEFVTKWDLAVKI
GGHLKDPVDSHMGRLDMRRMSYVQRMGKLLGGQLWESAGSPEVDPDRFAVVVGTGLGGAE
RIVESYDLMNAGGPRKVSPLAVQMIMPNGAAAVIGLQLGARAGVMTPVSACSSGSEAIAH
AWRQIVMGDADVAVCGGVEGPIEALPIAAFSMMRAMSTRNDEPERASRPFDKDRDGFVFG
EAGALMLIETEEHAKARGAKPLARLLGAGITSDAFHMVAPAADGVRAGRAMTRSLELAGL
SPADIDHVNAHGTATPIGDAAEANAIRVAGCDQAAVYAPKSALGHSIGAVGALESVLTVL
TLRDGVIPPTLNYETPDPEIDLDVVAGEPRYGDYRYAVNNSFGFGGHNVALAFGRY

```
   1 - gtg agt cag cct tcc acc gct aat ggc ggt
  31 - ttc ccc agc gtt gtg gtg acc gcc gtc aca
  61 - gcg acg acg tcg atc tcg ccg gac atc gag
  91 - agc acg tgg aag ggt ctg ttg gcc ggc gag
 121 - agc ggc atc cac gca ctc gaa gac gag ttc
 151 - gtc acc aag tgg gat cta gcg gtc aag atc
 181 - ggc ggt cac ctc aag gat ccg gtc gac agc
 211 - cac atg ggc cga ctc gac atg cga cgc atg
 241 - tcg tac gtc cag cgg atg ggc aag ttg ctg
 271 - ggc gga cag cta tgg gag tcc gcc ggc agc
 301 - ccg gag gtc gat cca gac cgg ttc gcc gtt
 331 - gtt gtc ggc acc ggt cta ggt gga gcc gag
 361 - agg att gtc gag agc tac gac ctg atg aat
 391 - gcg ggc ggc ccc cgg aag gtg tcc ccg ctg
 421 - gcc gtt cag atg atc atg ccc aac ggt gcc
 451 - gcg gcg gtg atc ggt ctg cag ctt ggg gcc
 481 - cgc gcc ggg gtg atg acc ccg gtg tcg gcc
 511 - tgt tcg tcg ggc tcg gaa gcg atc gcc cac
 541 - gcg tgg cgt cag atc gtg atg ggc gac gcc
 571 - gac gtc gcc gtc tgc ggc ggt gtc gaa gga
 601 - ccc atc gag gcg ctg ccc atc gcg gcg ttc
 631 - tcc atg atg cgg gcc atg tcg acc cgc aac
 661 - gac gag cct gag cgg gcc tcc cgg ccg ttc
 691 - gac aag gac cgc gac ggc ttt gtg ttc ggc
 721 - gag gcc ggt gcg ctg atg ctc atc gag acg
 751 - gag gag cac gcc aaa gcc cgt ggc gcc aag
 781 - ccg ttg gcc cga ttg ctg ggt gcc ggt atc
 811 - acc tcg gac gcc ttt cat atg gtg gcg ccc
 841 - gcg gcc gat ggt gtt cgt gcc ggt agg gcg
 871 - atg act cgc tcg ctg gag ctg gcc ggg ttg
 901 - tcg ccg gcg gac atc gac cac gtc aac gcg
 931 - cac ggc acg gcg acg cct atc ggc gac gcc
 961 - gcg gag gcc aac gcc atc cgc gtc gcc ggt
 991 - tgt gat cag gcc gcg gtg tac gcg ccg aag
1021 - tct gcg ctg ggc cac tcg atc ggc gcg gtc
1051 - ggt gcg ctc gag tcg gtg ctc acg gtg ctg
1081 - acg ctg cgc gac ggc gtc atc ccg ccg acc
1111 - ctg aac tac gag aca ccc gat ccc gag atc
1141 - gac ctt gac gtc gtc gcc ggc gaa ccg cgc
1171 - tat ggc gat tac cgc tac gca gtc aac aac
1201 - tcg ttc ggg ttc ggc ggc cac aat gtg gcg
1231 - ctt gcc ttc ggg cgt tac tga
```

8.4.3 FabD (Rv2243) (909 bp) (302 aa) (30788.2 da) (Theoretical pI: 4.84)

MIALLAPGQGSQTEGMLSPWLQLPGAADQIAAWSKAADLDLARLGTTASTEEITDTAVAQ
PLIVAATLLAHQELARRCVLAGKDVIVAGHSVGEIAAYAIAGVIAADDAVALAATRGAEM
AKACATEPTGMSAVLGGDETEVLSRLEQLDLVPANRNAAGQIVAAGRLTALEKLAEDPPA
KARVRALGVAGAFHTEFMAPALDGFAAAAANIATADPTATLLSNRDGKPVTSAAAAMDTL
VSQLTQPVRWDLCTATLREHTVTAIVEFPPAGTLSGIAKRELRGVPARAVKSPADLDELA
NL

```
  1 - gtg att gcg ttg ctc gca ccc gga cag ggt
 31 - tcg caa acc gag gga atg ttg tcg ccg tgg
 61 - ctt cag ctg ccc ggc gca gcg gac cag atc
 91 - gcg gcg tgg tcg aaa gcc gct gat cta gat
121 - ctt gcc cgg ctg ggc acc acc gcc tcg acc
151 - gag gag atc acc gac acc gcg gtc gcc cag
181 - cca ttg atc gtc gcc gcg act ctg ctg gcc
211 - cac cag gaa ctg gcg cgc cga tgc gtg ctc
241 - gcc ggc aag gac gtc atc gtg gcc ggc cac
271 - tcc gtc ggc gaa atc gcg gcc tac gca atc
301 - gcc ggt gtg ata gcc gcc gac gac gcc gtc
331 - gcg ctg gcc gcc acc cgc ggc gcc gag atg
361 - gcc aag gcc tgc gcc acc gag ccg acc ggc
391 - atg tct gcg gtg ctc ggc ggc gac gag acc
421 - gag gtg ctg agt cgc ctc gag cag ctc gac
451 - ttg gtc ccg gca aac cgc aac gcc gcc ggc
481 - cag atc gtc gct gcc ggc cgg ctg acc gcg
511 - ttg gag aag ctc gcc gaa gac ccg ccg gcc
541 - aag gcg cgg gtg cgt gca ctg ggt gtc gcc
571 - gga gcg ttc cac acc gag ttc atg gcg ccc
601 - gca ctt gac ggc ttt gcg gcg gcc gcg gcc
631 - aac atc gca acc gcc gac ccc acc gcc acg
661 - ctg ctg tcc aac cgc gac ggg aag ccg gtg
691 - aca tcc gcg gcc gcg gcg atg gac acc ctg
721 - gtc tcc cag ctc acc caa ccg gtg cga tgg
751 - gac ctg tgc acc gcg acg ctg cgc gaa cac
781 - aca gtc acg gcg atc gtg gag ttc ccc ccc
811 - gcg ggc acg ctt agc ggt atc gcc aaa cgc
841 - gaa ctt cgg ggg gtt ccg gca cgc gcc gtc
871 - aag tca ccc gca gac ctg gac gag ctg gca
901 - aac cta taa
```

8.4.4 AcpM (Rv2244) (348 bp) (115 aa) (12523.9 da) (Theoretical pI: 4.00)

MPVTQEEIIAGIAEIIEEVTGIEPSEITPEKSFVDDLDIDSLSMVEIAVQTEDKYGVKIP
DEDLAGLRTVGDVVAYIQKLEEENPEAAQALRAKIESENPDAVANVQARLEAESK

```
  1 - gtg cct gtc act cag gaa gaa atc att gcc
 31 - ggt atc gcc gag atc atc gaa gag gta acc
 61 - ggt atc gag ccg tcc gag atc acc ccg gag
 91 - aag tcg ttc gtc gac gac ctg gac atc gac
121 - tcg ctg tcg atg gtc gag atc gcc gtg cag
151 - acc gag gac aag tac ggc gtc aag atc ccc
181 - gac gag gac ctc gcc ggt ctg cgt acc gtc
211 - ggt gac gtt gtc gcc tac atc cag aag ctc
241 - gag gaa gaa aac ccg gag gcg gct cag gcg
271 - ttg cgc gcg aag att gag tcg gag aac ccc
301 - gat gcc gtt gcc aac gtt cag gcg agg ctt
331 - gag gcc gag tcc aag tga
```

8.4.5 Pks13 (Rv3800c) (5202 bp) (1733 aa) (186445.6 da) (Theoretical pI: 4.83)

MADVAESQENAPAERAELTVPEMRQWLRNWVGKAVGKAPDSIDESVPMVELGLSSRDAVA
MAADIEDLTGVTLSVAVAFAHPTIESLATRIIEGEPETDLAGDDAEDWSRTGPAERVDIA
IVGLSTRFPGEMNTPEQTWQALLEGRDGITDLPDGRWSEFLEEPRLAARVAGARTRGGYL
KDIKGFDSEFFAVAKTEADNIDPQQRMALELTWEALEHARIPASSLRGQAVGVYIGSSTN
DYSFLAVSDPTVAHPYAITGTSSSIIANRVSYFYDFHGPSVTIDTACSSSLVAIHQGVQA
LRNGEADVVVAGGVNALITPMVTLGFDEIGAVLAPDGRIKSFSADADGYTRSEGGGMLVL
KRVDDARRDGDAILAVIAGSAVNHDGRSNGLIAPNQDAQADVLRRAYKDAGIDPRTVDYI
EAHGTGTILGDPIEAEALGRVVGRGRPADRPALLGAVKTNVGHLESAAGAASMAKVVLAL
QHDKLPPSINFAGPSPYIDFDAMRLKMITTPTDWPRYGGYALAGVSSFGFGGANAHVVVR
EVLPRDVVEKEPEPEPEPKAAAEPAEAPTLAGHALRFDEFGNIITDSAVAEEPEPELPGV
TEEALRLKEAALEELAAQEVTAPLVPLAVSAFLTSRKKAAAAELADWMQSPEGQASSLES
IGRSLSRRNHGRSRAVVLAHDHDEAIKGLRAVAAGKQAPNVFSVDGPVTTGPVWVLAGFG
AQHRKMGKSLYLRNEVFAAWIEKVDALVQDELGYSVLELILDDAQDYGIETTQVTIFAIQ
IALGELLRHHGAKPAAVIGQSLGEAASAYFAGGLSLRDATRAICSRSHLMGEGEAMLFGE
YIRLMALVEYSADEIREVFSDFPDLEVCVYAAPTQTVIGGPPEQVDAILARAEAEGKFAR
KFATKGASHTSQMDPLLGELTAELQGIKPTSPTCGIFSTVHEGRYIKPGGEPIHDVEYWK
KGLRHSVYFTHGIRNAVDSGHTTFLELAPNPVALMQVALTTADAGLHDAQLIPTLARKQD
EVSSMVSTMAQLYVYGHDLDIRTLFSRASGPQDYANIPPTRFKRKEHWLPAHFSGDGSTY
MPGTHVALPDGRHVWEYAPRDGNVDLAALVRAAAAHVLPDAQLTAAEQRAVPGDGARLVT
TMTRHPGGASVQVHARIDESFTLVYDALVSRAGSESVLPTAVGAATAIAVADGAPVAPET
PAEDADAETLSDSLTTRYMPSGMTRWSPDSGETIAERLGLIVGSAMGYEPEDLPWEVPLI
ELGLDSLMAVRIKNRVEYDFDLPPIQLTAVRDANLYNVEKLIEYAVEHRDEVQQLHEHQK
TQTAEEIARAQAELLHGKVGKTEPVDSEAGVALPSPQNGEQPNPTGPALNVDVPPRDAAE
RVTFATWAIVTGKSPGGIFNELPRLDDEAAAKIAQRLSERAEGPITAEDVLTSSNIEALA
DKVRTYLEAGQIDGFVRTLRARPEAGGKVPVFVFHPAGGSTVVYEPLLGRLPADTPMYGF
ERVEGSIEERAQQYVPKLIEMQGDGPYVLVGWSLGGVLAYACAIGLRRLGKDVRFVGLID
AVRAGEEIPQTKEEIRKRWDRYAAFAEKTFNVTIPAIPYEQLEELDDEGQVRFVLDAVSQ
SGVQIPAGIIEHQRTSYLDNRAIDTAQIQPYDGHVTLYMADRYHDDAIMFEPRYAVRQPD
GGWGEYVSDLEVVPIGGEHIQAIDEPIIAKVGEHMSRALGQIEADRTSEVGKQ

```
   1 - atg gct gac gta gcg gaa tcc cag gag aac
  31 - gcc ccc gcc gaa agg gcc gag cta acg gtc
  61 - ccc gag atg cgc cag tgg ctg cgc aac tgg
  91 - gtg ggt aag gcc gtc gga aag gca ccg gac
 121 - tcg atc gac gaa tcg gtg ccc atg gtg gag
 151 - ctg ggt ctg tcg tcg cgc gat gcc gtc gcg
 181 - atg gcc gcc gac ata gaa gac ctg acc ggg
 211 - gtc acg ctg tcg gtc gcg gtg gcg ttc gcg
 241 - cat ccg acc atc gaa tcg ctg gcc acc cgg
 271 - atc atc gag ggc gag ccg gag acc gac cta
 301 - gcg ggc gat gac gcc gaa gac tgg tcg cgc
 331 - acc ggc ccg gcc gag cgc gtc gac atc gcg
 361 - atc gtg ggc ttg tcc acc cgc ttc ccg ggc
 391 - gag atg aac acc ccc gag cag acc tgg cag
 421 - gcg ctg ctg gaa ggc cgc gac ggg atc acc
 451 - gac ctg ccc gac ggg cgc tgg tcg gaa ttc
 481 - ctc gaa gag ccg cgg ctg gcc gcg cgg gtc
 511 - gcc ggg gcc cgc acc cgg ggc ggc tac ctg
 541 - aag gac atc aag ggc ttc gat tcg gag ttc
 571 - ttc gcg gtg gcc aag acc gaa gcc gac aac
 601 - atc gac ccg cag cag cgg atg gcg ctg gag
 631 - ctg acc tgg gag gcg ctc gag cac gcc cgc
 661 - atc ccg gcg tcg agc ctg cgc ggc cag gcc
 691 - gtc ggt gtg tac atc ggc agc tcc acc aac
 721 - gac tac agc ttc ctg gcg gtg tcg gac ccg
 751 - acg gtc gcg cac ccg tat gcg atc acc ggc
 781 - acc agc agc tcg atc atc gcc aac cgg gtg
 811 - tcc tac ttc tac gac ttc cac gga ccg tcg
 841 - gtc acc att gac acc gcg tgc tcg agt tcg
 871 - ctg gtg gcc atc cac cag ggg gtg cag gcg
 901 - ctg cgc aac ggc gag gcc gac gta gtg gtc
 931 - gcc ggc ggg gtg aac gcg ttg atc aca ccg
 961 - atg gtc acc ctg ggt ttc gac gag atc ggt
 991 - gcg gtg ctg gcg ccc gac ggc cgg atc aag
1021 - tcg ttc tca gcc gac gcc gac ggc tac acc
1051 - cgc tcc gaa ggc ggc ggc atg ctg gtg ctc
1081 - aag cgg gtc gac gac gcc cgc cgc gac ggc
1111 - gac gcg atc ctg gcc gtg atc gcc ggc agc
1141 - gcg gtc aac cac gac ggc cgg tcc aac ggc
1171 - ctg atc gca ccc aac cag gac gcg cag gcc
1201 - gac gtg ctg cgc cgg gcc tac aag gac gcc
1231 - ggc atc gat ccg cgc acc gtc gac tac atc
1261 - gag gcg cac ggc acc ggc acc atc ctc ggc
1291 - gac cca atc gag gcc gag gcg ctg ggc cgg
1321 - gtg gtc ggt agg ggc cgt ccg gcc gat cgg
1351 - ccg gcg ctg ctg ggt gcg gtg aaa acc aac
1381 - gtc ggg cac ctg gaa tcg gcg gcc ggc gcg
1411 - gcc agc atg gcc aag gtg gtg ctg gcg ctg
1441 - cag cac gac aaa ctg ccg ccg tcg atc aac
1471 - ttc gcc ggc ccc agc ccc tac atc gac ttc
1501 - gac gcg atg cgg ttg aag atg atc acc acg
1531 - ccc acc gac tgg ccg cga tac ggc ggc tac
1561 - gcg ctg gcc ggg gtg tcc agc ttc ggc ttc
1591 - ggc ggc gcc aac gcg cac gtg gtg gtg cgc
1621 - gag gtc ctg ccg cgt gac gtg gtg gaa aag
```

1651 - gaa ccg gaa ccc gag ccg gaa ccc aag gcg
1681 - gcc gcc gaa ccc gcc gag gcg ccc acg ttg
1711 - gca ggc cac gcg ctg cgg ttc gac gag ttc
1741 - ggc aac atc atc acc gac tcg gcg gtc gcc
1771 - gaa gag ccg gag ccc gaa ctg ccc gga gtc
1801 - acc gag gag gcg ctg cgg ctc aag gaa gcc
1831 - gcg ttg gaa gag ctt gcg gcc caa gag gtt
1861 - acg gca cca ttg gtc ccg ttg gcg gtg tcg
1891 - gcg ttt ctg acg tcc cgc aag aag gcg gcg
1921 - gcc gcc gag ttg gcg gac tgg atg caa agc
1951 - ccg gaa ggc cag gcc tcc tcg ctg gaa tcg
1981 - atc ggc agg tcg ttg tcg cgg cgc aac cac
2011 - ggc cgt tcc cgc gcg gtg gtg ttg gcc cac
2041 - gac cac gac gag gcc atc aag ggc ctg cgc
2071 - gcg gtc gcc gcg ggc aag cag gcg ccg aac
2101 - gtg ttc agc gtc gac ggg ccg gtg acc acc
2131 - ggc ccg gtc tgg gtg ctc gcc gga ttc ggc
2161 - gcc cag cat cgc aag atg ggc aag agc ctg
2191 - tac ctg cgc aac gag gtg ttc gcg gcg tgg
2221 - atc gag aag gtc gac gcc ctg gtc caa gac
2251 - gag ctg ggc tac tcg gtg ctg gag ctg atc
2281 - ctg gac gac gcg cag gac tac ggc atc gag
2311 - acc acc cag gtc acc atc ttc gcg atc cag
2341 - atc gcg ctg ggt gag ctg ctg cgc cat cac
2371 - ggc gcc aaa ccg gcc gcg gtc atc ggc cag
2401 - tcg ctg ggt gag gcc gcg tcg gcc tac ttc
2431 - gcc ggc ggg ctg tcg ctg cgg gat gcc acc
2461 - cgg gcg atc tgc tcg cgc tcg cac ctg atg
2491 - ggc gag ggt gag gcg atg ctg ttc ggc gag
2521 - tac atc cgg ttg atg gcg ctg gtg gaa tac
2551 - tcc gcc gac gaa atc aga gaa gtg ttc tcc
2581 - gac ttc ccc gat ctg gag gtg tgt gtc tac
2611 - gcc gcg ccc acc cag acg gtc atc ggc ggc
2641 - ccc ccc gag cag gtg gac gcg atc ctt gcc
2671 - cgc gcc gag gcc gag ggc aag ttc gcc cgc
2701 - aaa ttc gcg acc aag ggc gcc agc cac acc
2731 - tcg cag atg gac ccg ctg ctg ggc gag ctc
2761 - acc gcg gag ctg caa ggc atc aag ccg acg
2791 - agc ccg acg tgt ggg atc ttc tcg acg gtg
2821 - cac gag ggc cgc tac atc aaa ccc ggc ggc
2851 - gaa ccc atc cac gac gtc gaa tac tgg aag
2881 - aag ggg ctg cgg cat tcc gtc tac ttc acc
2911 - cac ggc atc cgc aac gcc gtc gac agc ggg
2941 - cac acc acc ttc ctg gag ctg gca ccc aat
2971 - ccg gtg gcg ctg atg cag gtc gcc ctg acc
3001 - acc gcc gat gcc ggg ctg cat gac gcc cag
3031 - ttg atc ccg acg ctg gcc cgc aag caa gac
3061 - gag gtc tcc tcg atg gtc tcg acc atg gcg
3091 - cag ctg tat gtg tac ggc cac gac ctg gac
3121 - ata cgc acg ctg ttt agc cgc gcc agt ggg
3151 - ccg cag gat tac gcg aac att ccg ccg acc
3181 - cgg ttc aag cgc aag gag cac tgg ctg ccc
3211 - gcg cac ttc tcc ggc gac ggc tcg acg tac
3241 - atg ccg ggc acc cat gtc gcc ctg ccg gat
3271 - ggg cgg cac gtc tgg gag tac gcg ccg cgg
3301 - gac ggc aat gtg gac ttg gcc gcg ttg gtc

```
3331 - agg gcc gcc gcc gcc cac gtg ctt ccg gac
3361 - gcg caa ctg acc gcc gcc gag cag cgc gcg
3391 - gtg ccc ggc gac ggc gcc cgg ctg gtg acg
3421 - acg atg acc cgt cac ccc ggc ggc gcc tcg
3451 - gtg cag gtg cac gcc cgc atc gac gag tcc
3481 - ttc acg ctg gtc tac gac gcc ctg gtg tcc
3511 - cga gcg ggg tcc gaa tcg gtg ttg ccc acc
3541 - gcg gtg ggt gcg gcg acg gcg atc gcg gtt
3571 - gcg gac ggg gcg cct gtc gcg ccg gaa acg
3601 - ccc gcc gaa gac gcg gac gcc gag acg ctt
3631 - tcg gac agc ctg acc acc cgt tac atg ccg
3661 - tcc ggc atg acc cga tgg tcg cct gat tcc
3691 - ggt gag acc atc gcc gag cgg ctg ggc ctg
3721 - att gtc ggg tct gcg atg ggc tat gag ccc
3751 - gag gac ctg ccg tgg gag gtg ccg ctg atc
3781 - gag ctt ggc ctg gac tcg ctg atg gcg gtg
3811 - cgc atc aaa aac cgc gtc gag tac gac ttc
3841 - gac ctg cca ccg atc cag ctg acc gcg gtg
3871 - cgc gac gcc aac ctc tac aac gtg gag aag
3901 - ctg atc gaa tac gcg gtc gag cac cgt gac
3931 - gag gtg cag cag ctg cac gag cac cag aaa
3961 - acc cag acc gct gag gag atc gcg cgg gcc
3991 - cag gcc gaa ttg ctg cat ggc aag gtg ggc
4021 - aag acc gag ccg gtc gac tcg gaa gcc ggg
4051 - gtt gcg ctc ccg tcg ccg caa aac ggc gag
4081 - cag cca aac ccg aca ggg ccc gcg ctc aac
4111 - gtc gac gtg ccg ccg cgg gac gct gcc gag
4141 - cgg gtc acc ttc gcc acc tgg gcg atc gtc
4171 - acc ggc aag tcc ccg ggc ggc atc ttc aac
4201 - gag ctg ccc agg ctg gac gac gag gcc gcg
4231 - gcc aag att gcg cag cgg ctt tcc gag cgc
4261 - gcc gaa ggc ccg atc acc gcc gag gac gtg
4291 - ctg acg tcg tcg aac atc gag gcg ctg gcc
4321 - gac aag gtg cgc acg tat ttg gag gcc ggg
4351 - cag atc gat ggg ttc gtc cgc acc ctg cgg
4381 - gcg cgg ccc gaa gca ggc ggg aag gtg ccg
4411 - gtg ttc gtg ttt cat ccg gcc ggc ggc tcg
4441 - acg gtg gtg tac gag ccg ctg ctg ggc cgg
4471 - ctg ccg gcg gac acc cca atg tat ggc ttc
4501 - gaa cgg gtc gag ggg tcg atc gaa gag cgt
4531 - gca cag cag tac gtg ccg aag ctg atc gag
4561 - atg cag ggc gac ggg ccc tat gtc ctg gtg
4591 - ggt tgg tcg ctg ggc ggt gtg ctg gcc tac
4621 - gcg tgc gcg atc ggt ttg cgg cgg ctg ggc
4651 - aag gac gtg cgg ttc gtc ggg ctg atc gac
4681 - gcg gtg cgc gcc ggt gag gag atc ccg cag
4711 - acc aag gag gag atc cgc aag cgc tgg gac
4741 - cgc tac gcc gcc ttc gcc gag aag acg ttc
4771 - aac gtg acc atc ccg gcg atc ccg tac gag
4801 - cag ctc gag gag ctc gac gac gag ggc cag
4831 - gtc cgg ttc gtg ctg gac gcc gtc agc cag
4861 - tcc ggt gtg cag atc ccg gcc ggg atc atc
4891 - gaa cac caa cgc acg tcg tat ctg gac aac
4921 - cgg gcg atc gac acc gcc cag atc cag ccg
4951 - tac gac ggg cat gtc acc ctc tac atg gcc
4981 - gat cgc tac cat gac gac gcg atc atg ttc
```

```
5011 - gag ccc cgc tac gcc gtg cgc cag ccg gac
5041 - ggc ggg tgg ggc gag tac gtt tcc gac ctc
5071 - gag gtg gtg ccg atc ggt ggc gag cac att
5101 - cag gcc atc gac gag ccg atc atc gcc aag
5131 - gtg ggc gaa cac atg agc cgc gcg ttg ggg
5161 - cag atc gag gcc gat cga aca agt gag gta
5191 - ggc aag cag tga
```