Biochemical and Functional Genomic Analysis in Mycobacteria: Effect of Iron Limitation and Mutational Analysis of *pykA* and *ald*

A Thesis
Submitted for the Degree of
Doctor of Philosophy

By

Sivagami sundaram C A S

Department of Animal Sciences
School of Life Sciences
University of Hyderabad
Hyderabad 500 046
India

March 2007

School of Life Sciences
Department of Animal Sciences

STATEMENT

I hereby declare that the work embodied in this dissertation is the result of the investigation carried out by me in the Department of Animal Sciences, School of Life Sciences, University of Hyderabad, Hyderabad, under the supervision of **Dr. Manjula Sritharan.** As a recipient of the Commonwealth Split-Site Doctoral Fellowship, the studies on mycobacterial genetics were done in the lab of Dr. Paul Wheeler, Veterinary Laboratories Agency, UK.

In keeping with the general practice of reporting scientific observations, due acknowledgements have been made wherever the work described is based on the findings of other investigators. Any omission, which might have occurred by oversight or error, is regretted.

March 2007 **Sivagami sundaram C A S**

Statement verified

Dr. Manjula Sritharan
PhD Supervisor

School of Life Sciences
Department of Animal Sciences

CERTIFICATE

Certified that the work embodied in this thesis entitled **'Biochemical and functional genomic analysis in mycobacteria: effects of iron limitation and mutational analysis of *pykA* and *ald*'** has been carried out by Mr. Sivagami sundaram C A S under my supervision and the same has not been submitted elsewhere for a degree.

Dr. Manjula Sritharan
(PhD Supervisor)

Head
Department of Animal Sciences

Dean
School of Life Sciences

CHAPTER 1
REVIEW OF LITERATURE

1.1. The genus *Mycobacterium*

Mycobacteria belong to the Order *Actinomycetales*, Family *Mycobacteriaceae*. The generic name *Mycobacterium* (meaning fungus bacterium) was first introduced by Lehmann and Neumann due to the mould-like growth of the bacillus in liquid medium (Lehmann & Neumann, 1896); before this they were known as Koch's bacillus, as Robert Koch was the to first describe the tuberculosis causing organisms in 1882. Mycobacteria are rod-shaped bacilli that are classified as Gram-positive. Due to lipid-rich cell wall, they are difficult to stain by the Gram stain and can be stained with basic dyes such as carbol fuchsin and cannot be decolourised with 95% ethanol and 3% hydrochloric acid. This unique property is termed "acid-fastness" and is the basis of the Ziehl-Neelsen staining technique for the identification of mycobacteria. Mycobacteria are aerobic, although some of them, for example *M. bovis* are microaerophilic and grow better at lower oxygen tensions. Some mycobacteria are pigmented and based on the production of pigments, they can be classified as scotochromogens (produce yellow pigment in the dark) or photochromogens (produce an orange pigment in the light) or achromogens (do not produce any pigment) (Table 1). The cultivable members of the genus can be divided into two main groups on the basis of growth rate, namely slow and fast growing.

Table 1 - Classification of Non-tuberculous mycobacteria

Mycobacterial species	Runyon Group	Pigment formation
M. kansasii	I	Photochromogens
M. marinum		
M. simiae		
M. scrofulaceum	II	Scotochromogens
M. szulgai		
M. gordonae		
M. avium	III	Nonchromogens
M. intracellularae		
M. ulcerans		
M. fortuitum	IV	Rapid growers
M. chelonae		

1.1.1. *Mycobacterium tuberculosis* and *M. leprae*: two important human pathogens

Mycobacterium tuberculosis and *M. leprae* are two important human pathogens that cause tuberculosis and leprosy respectively. Tuberculosis is the leading cause of death among the curable infectious diseases, killing around 2-3 million people annually (WHO, 2004 & 2005). There was an estimated 8.9 million new cases in 2004, with 3.9 million cases being sputum smear positive (Fig. 1).

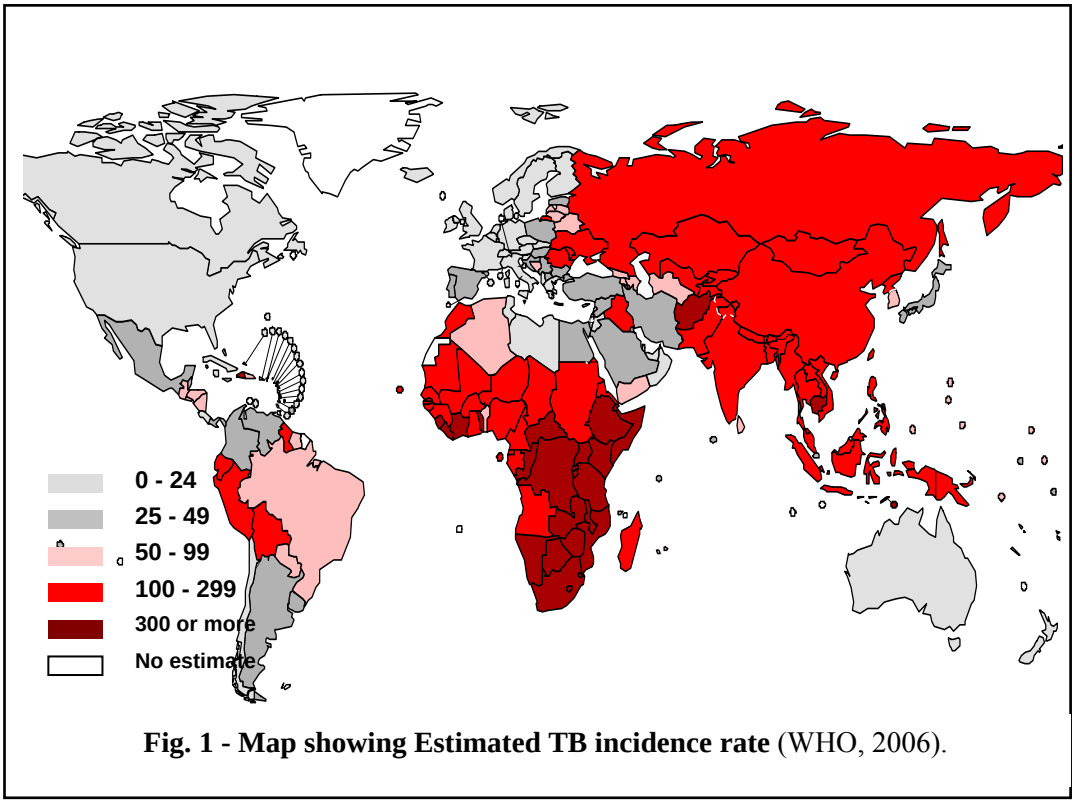


Fig. 1 - Map showing Estimated TB incidence rate (WHO, 2006).

Death due to tuberculosis in HIV infected people increase at least twice that of the person infected only with HIV. This alarming rise led the WHO to declare tuberculosis 'a global emergency' in 1993 (WHO, 1993). In addition, about one third of human population is estimated to suffer from latent tuberculosis, which can be reactivated even after several decades (Glassroth *et al.,* 2004). It is estimated that, between 2000 and 2020 nearly one billion people will be newly infected, the active disease will affect 200 million, and 35 million will die of TB, if control measures are not significantly improved (WHO, 2006).

1.1.2 Control measures

Control measures for tuberculosis include timely diagnosis and chemotherapy. Existing diagnostic methods can detect up to 60% of tuberculosis cases but tuberculosis management in developing countries is difficult the existing diagnostic methods are non-specific and time consuming. The present day diagnosis for human tuberculosis include chest X-ray for identifying the granuloma formation, Mantoux test, AFB testing of sputum smears and biopsies. A positive culture is required to confirm the diagnosis, but this is time consuming. Many patients in endemic areas are never diagnosed and this contributes to the astonishing number of early deaths from tuberculosis (Brodie & Schulger, 2005). Modern diagnostic methods like PCR, ELISA and more rapid ways to detect the positive cultures are being developed.

Tuberculosis can be cured in 95% of patients with active, drug sensitive pulmonary TB (Spigelman & Gillepsie, 2006). Multi-drug therapy with a combination of three frontline drugs, isoniazid, rifampin and pyrazinamide and one or more of the second-tier antibiotics including ethambutol, ethionamide, *p*-aminosalicylic acid, D-cycloserine, streptomycin, capreomycin, kanamycin and thiacetazone considerably reduced the incidence of TB, especially in developed countries. Nevertheless, this decline has been reversed in the past two decades due to several factors including poverty, overcrowding, travel and the synergy between HIV and TB (Corbett *et al.*, 2003). Despite the efforts like the DOTS (Direct Observed Treatment Strategy), numbers of drug resistant cases are on the rise, thereby necessitating newer and better drugs.

Vaccination with *M. bovis* BCG has not been promising. The variable efficacy of the BCG vaccine and the genetic heterogeneity between strains demands that a better vaccine be developed for the prevention of TB.

Mycobacterium leprae causes leprosy, which is a chronic granulomatous infection of the skin and peripheral nerves. The damage to the peripheral nerves results in sensory and motor impairment with characteristic deformities and disability. Since *M. leprae* cannot be cultivated in axenic media, progress has been slow in understanding the pathogen. However, the completion of the genome

sequence of this pathogen showing the extensive reduction and decay of the genome will aid in better understanding (Britton & Lockwood, 2004).

Tuberculosis is not only a disease of humans, but also has a devastating effect on cattle and livestock. Bovine tuberculosis caused by *M. bovis* is a significant public health problem and it causes great economic losses in countries with infected livestock. Despite the control measures, the incidence of bovine tuberculosis in some countries has remained the same or increased due to the presence of endemic wildlife reservoirs. This has been particularly the case in New Zealand, United Kingdom and Republic of Ireland (Olsen *et al.,* 2001; Cousins, 2001). Also, bovine TB remains a significant problem in developing countries; indeed more than 94% of the world population live in countries in which the control of bovine TB is either limited or completely absent (Vordermeier *et al.*, 2006).

1.1.3. *Mycobacterium kansasii*: an opportunistic mycobacterial pathogen

Mycobacterium kansasii is one of the important and frequently encountered non-tuberculous mycobacterial pathogen in humans, in whom it produces a chronic, progressive and caviatary lung disease similar to tuberculosis (Bloch *et al.,* 1998; Bittner *et al.,* 1996). With the increase in the rate of HIV infection, infection rate as high as 300 times as that of normal population has been recorded in Switzerland (Tortoli *et al.,* 1994). Laboratory identification of *M. kansasii* depends upon its growth characteristics such as the photochromogenicity, various biochemical tests like Tween hydrolysis, nitrate reduction and also on the pattern of susceptibility to various anti-tubercular drugs. Although, biochemical tests cannot differentiate the clinical isolates of *M. kansasii*, PCR-RFLP analysis of *hsp65* gene and the differences in the 16S –23S rRNA spacer regions identified seven subtypes (I to VII) of *M. kansasii*. However, types I and II are the most important and frequently isolated form from the specimens, while the rest of the types isolated are mostly from the environment. The heterogenicity may have an important role in its pathogenic, clinical and epidemiological implications (Alcaide *et al.,* 2004).

1.2. Understanding host-pathogen interactions

1.2.1. Unique cell envelope of mycobacteria

The mycobacterial cell envelope is complex and has a lower layer that consists of the plasma membrane surrounded by a lipid and carbohydrate rich cell wall called the cell wall core or the mycolyl–arabinogalactan–peptidoglycan (mAGP) complex, which in turn is encircled by the upper layer consisting of polysaccharide, complex free lipids and proteins (Fig. 2). Upon addition of various solvents, mAGP complex remains as the insoluble residue.

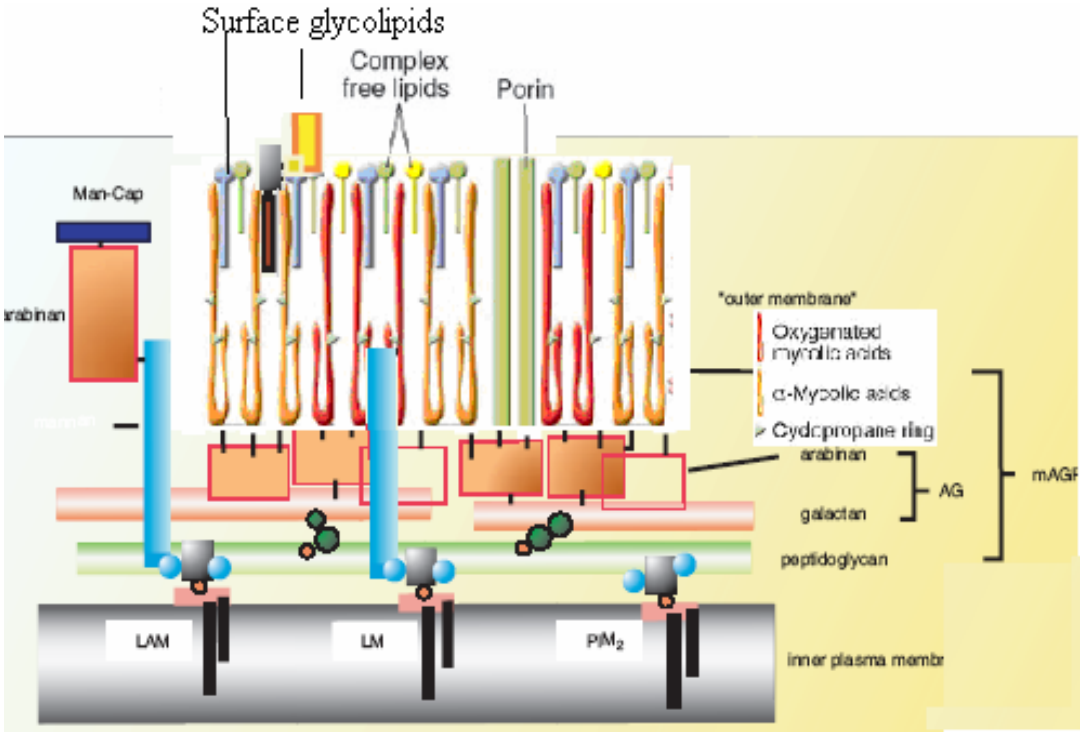


Fig. 2 - Schematic representation of *M. tuberculosis* cell envelope. Three forms of mycolic acids are depicted. α-Mycolates are the most abundant form in *M. tuberculosis* (orange) and has 2 cyclopropane rings (triangles) in *cis* configuration. Oxygenated mycolates (keto- and methoxy-, shown in red) have one cyclopropane ring each that is in either *cis* or *trans* configuration. They are covalently linked to the arabinogalactan layer, which is linked to the peptidoglycan layer. Other lipid complexes in the cell wall include acyl glycolipids (including TDM) and other complex free lipids (e.g., phthiocerol dimycocerosate) as well as sulfolipids (Khasnobis *et al,* 2002; Brennan, 2003; Riley, 2006).

1.2.1.1. Mycolyl–arabinogalactan–peptidoglycan (mAGP) complex

The insoluble cell wall core (formed after the removal of soluble proteins, lipids and carbohydrates) is chemically composed of three covalently linked macromolecules: highly cross-linked peptidoglycan (PG), arabinogalactan (AG) and mycolic acids. The arabinan of AG provides the anchoring point for the outer mycolyl lipid layer and links it to the underlying PG layer *via* a galactan chain to form the mAGP complex.

1.2.1.2. Peptidoglycan (PG)

The PG layer of mycobacteria resembles the gram positive bacteria and consists N-acetyl-β–D-glucosaminyl-(1→4)-N-acetyl-muramic acid units that are cross linked by short peptides chains linked to the acid groups of the muramic acid. The tetrapeptide side chains of PG consist of L-alanyl–D-isoglutaminyl–meso-diaminopimelyl–D-alanine (L-Ala–D-Glu–A_2pm–D-Ala) with the Glu being further amidated. In mycobacteria, some or all of the Mur residues are N-acylated with glycolic acid (MurNGly), and the cross-links include a proportion between two diaminopimelyl (A_2pm) residues as well as between A_2pm and D-Ala (Brennan, 2003).

1.2.1.3. Arabinogalactan (AG)

The PG is covalently linked to the galactan chain through the C-6 of its muramyl residues *via* a diglycosylphosphoryl bridge, L-Rha*p*-GlcNAc- 1-phosphate. Mycobacterial AG comprises of D-galactofuranoses and D-arabinofuranoses, which are extremely rare in nature. Also, AG has a few defined structural motifs rather than repeating units as seen in all other bacterial polysaccharides. The non-reducing end of the arabinan consists of a branched hexa-arabino-furanosyl structure [βAra*f*-(1→2)-αAra*f*] 2-3,5-αAra*f*-(1→5)-αAra*f* and the back-bone of AG chain consists of 5-linked αAra residues punctuated with 3,5-αAra*f* branching. The galactan consists of alternating 5- and 6-linked β-D-galactofuranose (Gal*f*) units and the arabinan chain is attached to the galactan core through C-5 of some of the 6-linked Gal*f* units (Crick *et al.,* 2001). The link between the arabinan and the galactan is not yet defined.

1.2.1.4. Mycolic acids (MA)

The distal ends of the AG are esterified with high molecular weight fatty acids called mycolic acids, seen only in mycobacteria. Mycolic acids are 1-alkyl branched 2-hydroxy fatty acids typically with 70 - 90 carbon atoms. The branch is commonly about 24 carbon atoms and is a simple alkyl chain, but the main chain contains (in *M. tuberculosis*) cyclopropyl, methoxyl or keto and methyl groups. The cyclopropanation as seen in *M. tuberculosis,* not seen in fast-growing mycobacterial species such as *M. smegmatis* has been implicated as one of the virulence factors (Brennan, 2003).

1.2.1.5. Cord factor (TDM)

Cord factor (trehalose 6, 6′- dimycolate, TDM) is a neutral glycolipid. Formation of 'cord' during the growth is one of the characteristic property of *M. tuberculosis* and this was attributed to TDM (Goren & Brennan, 1980).

1.2.1.6. Phosphatidylinositolmannosides (PIMs), Lipomannan (LM) and Lipoarabinomannan (LAM)

PIMs, LM and LAM are major lipoglycans found to be attached non-covalently to the plasma membrane through their phophatidyl-myo-inositol anchor and extend to the exterior of the cell wall. PIMs with two mannoses are common in *M. tuberculosis* while PIMs with five mannoses have also been recovered (Jackson *et al.,* 2000). LM and LAM are extensions of PIM. In LM, the mannan chain is extended and branching occurs and in LAM, additional arabinan is attached. These complex molecules are believed to play an important role in the physiology of the bacterium as well as in the modulation to host responses during infection. LAM and LM present in the cell wall could play a major role in virulence. LM was found to trigger the host pro-inflammatory responses. Thus, the ratio between LAM and LM may be a crucial factor in determining the virulence. LAM is an important modulator of the immune response in both tuberculosis and leprosy. It is found to be the key ligand in mediating interaction between the macrophage, dendritic cells and tubercle bacilli. Several reports suggest that PIM isolated from *M. tuberculosis* are able to induce TNF-α and IL-8 secretion in both human and murine macrophages (Chatterjee & Khoo, 1998; Chatterjee*,* 1997).

Other major components include sulpholipids (SL) phenolic glycolipid (PGL1) and phthiocerol dimycocersate (PDIM). *Mycobacterium leprae* PGL1 was shown to be associated with the resistance to intracellular killing by macrophages and the promotion of phagocytosis. Thus, the major cell wall components not only protect the cell but also enhance its virulence and its survival in the host.

1.3. General metabolism in the intracellular environment of the macrophages

The metabolism of mycobacteria and *M. tuberculosis* in particular is similar to other bacteria in terms of carbohydrate metabolism, energy production and biosynthesis of macromolecules. However, much needs to be understood about the intracellular environment in the macrophages (Fig. 3) where conditions such as

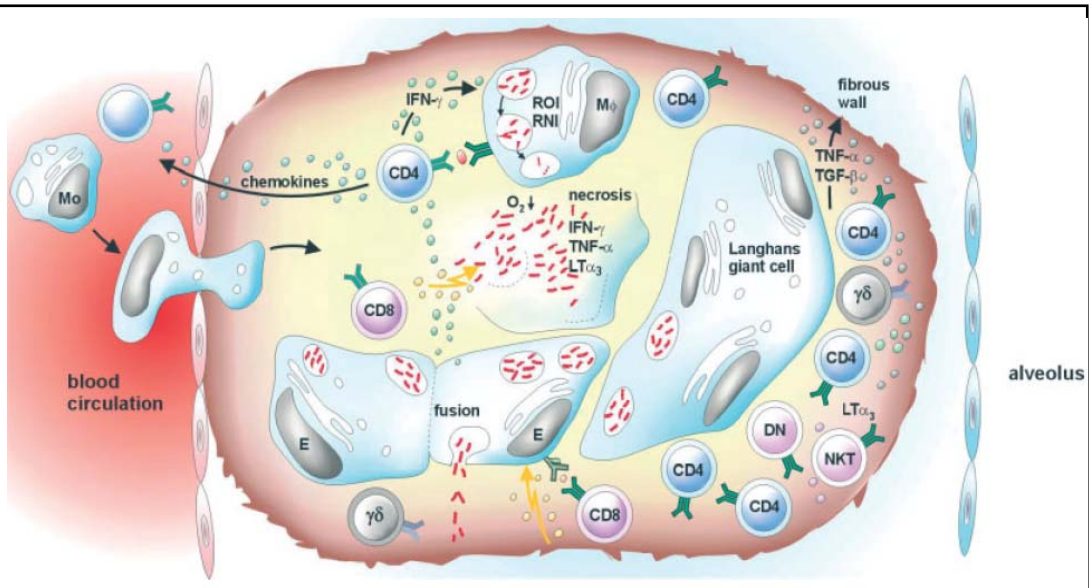


Fig. 3 - Structure of classical tuberculous granuloma and the immune responses (Ulrichs and Kaufmann, 2006), where Mφ = Alveolar macrophages, E = epithelioid cells, ROI = reactive oxygen intermediates, RNI = reactive nitrogen intermediates, TNF-α = Tumor necrosis factor, NKT = Natural killer cells, IFN γ and LTα 3 = lymphotoxin-α 3.

nutrient limitation, low pH, hydrolytic enzymes, reactive oxygen species and reactive nitrogen intermediates, low oxygen tension are said to occur (Wagner *et al.,* 2005). It is not clear whether the delicate balance between the host immune response and pathogen is responsible for the maintenance of the latency or the bacilli shuts down all its metabolic activities and remains inactive as long as strong cellular immune responses prevail in the microenvironment (Ulrichs & Kaufmann, 2006). The nature of metabolic adaptations to the anaerobic non-replicative state is largely

unknown. Studies have indicated that the bacterium can modulate its metabolic pathways and can remain dormant. Metabolism of these bacteria amid the triggers for the reactivation requires greater understanding.

1.4. Carbon metabolism

Pathogenic bacteria have adapted themselves as heterotrophs capable of utilising variety of carbon sources including carbohydrates, lipids, glycolipids and dicarboxylic acids. Catabolism of carbon compounds is very important as they provide energy in the form of reducing equivalents and ATP, as well as essential biosynthetic precursors such as phosphoenolpyruvate (PEP) & acetyl-CoA. The major pathways involving metabolism of carbon sources are lipid metabolism, sugar catabolism, the tricarboxylic acid (TCA) cycle, gluconeogenesis, and anaplerotic pathways.

1.4.1. Lipid metabolism

Mycobacterium tuberculosis possesses along with other mycobacteria, perhaps the most remarkable range of lipid molecules in the prokaryotic world. Genomic analysis showed that almost 250 enzymes representing almost every known biosynthetic pathway and enzyme involved in lipid synthesis and degradation, many of which are usually found in mammals and plants. This remarkable feature shows the importance of lipids in the tubercle bacillus life style and reflects the unusual composition of the cell envelope with array of various lipids in them (explained in detail in Section 1.2.1). This discovery had led to the development of drugs against the enzymes involved in both lipid biosynthesis and degradation (Khasnobis *et al.,* 2002).

Observations made approximately 50 years ago indicated that *M. tuberculosis* shifts from a metabolism that preferentially uses carbohydrates when growing *in vitro* to one that utilizes fatty acids when growing in the infected host (Smith, 2003). However, gluconeogenesis (Section 1.4.3) can be used to generate sugar phosphates from fatty acids.

A look into the biosynthetic pathway showed two types of fatty acid synthase system, FAS-I and FAS-II. The mycobacterial FAS-I system usually seen in eukaryotes catalyzes not only the synthesis of C16 and C18 fatty acids, the normal

products of *de novo* synthesis, but also elongation to produce C24 and C26 fatty acids (Bloch, 1975). The Type II fatty acid synthase of mycobacteria is analogous to other bacterial FAS-II systems, but not capable of *de novo* synthesis from acetate but instead elongates palmitoyl-ACP to fatty acids ranging from 24 to 56 carbons in length (Mdluli *et al*., 1998; Slayden & Barry 2002). The mycobacterial FAS-I produces a bimodal (C14:0-C16:0 to C24:0-C26:0) distribution of acyl-CoA fatty acids. The shorter acyl-CoA fatty acid precursors (C14:0-C16:0) condense with malonyl-ACP and the products are then elongated by the FAS-II system, yielding the long carbon chains (50-60 carbons) characteristic of mycolic acids. The FAS-II system is one of the promising drug target. Indeed, it is already known that the primary target for the action of first line anti-TB drug INH is InhA, which is part of the FAS-II multi-enzyme system. InhA elongates acyl fatty acid precursors yielding the long carbon chain of meromycolate branch of mycolic acids (Schroeder *et al.,* 2002).

1.4.2. Carbohydrate metabolism

Glycolysis, pentose phosphate pathway (PPP) and Entner-Doudoroff

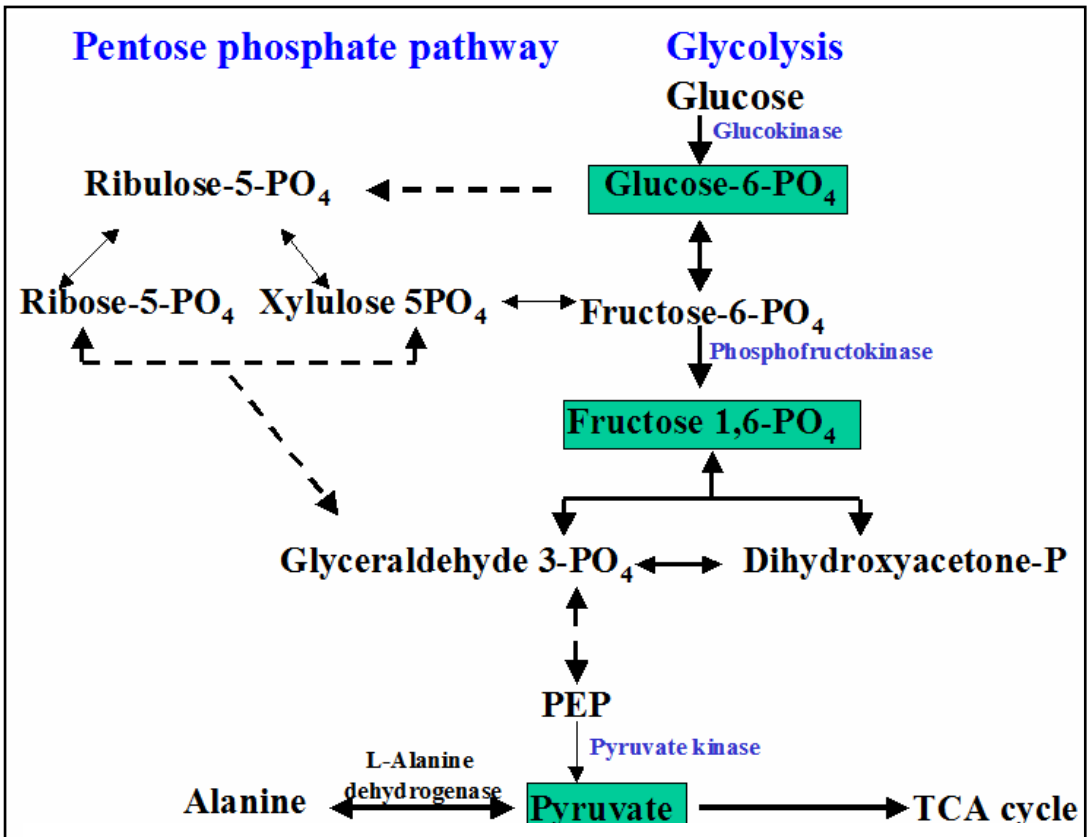


Fig. 4. Major pathways of carbohydrate metabolism in mycobacteria represented with enzymes involved in the rate determining steps (blue).

pathway are the best-characterised pathways of prokaryotes. Based on the genomic analysis, *M. tuberculosis* has an intact glycolytic (but not the other members of *M. tuberculosis* members) and pentose phosphate pathway (Fig. 4), while the Entner-Doudoroff pathway is absent (Wheeler & Ratledge, 1994).

1.4.3. Gluconeogenesis

Several glycolytic intermediates serving as precursors for biosynthetic pathway are generated by gluconeogenesis (Fig. 5) (Anderson & Wood, 1969). The glycolytic and gluconeogenic pathways comprise essentially the same set of enzymes catalysing readily reversible reactions. However, glycolytic conversion of fructose-6-phosphate to fructose-1, 6-bisphosphate, catalysed by phosphofructokinase, is effectively irreversible. The reaction in the gluconeogenic direction is catalysed by fructose-1, 6-bisphosphatase. The entry point into gluconeogenesis varies according to the carbon substrate but is usually PEP (Fig. 5).

1.4.4. The citric acid cycle (TCA cycle)

Acetyl-CoA derived from sugars or fatty acids is further oxidized to CO_2 by the citric acid cycle (Fig. 5), which provides reducing equivalents for respiration-mediated ATP synthesis and precursors for multiple biosynthetic pathways. During the normal operation of TCA cycle, two carbons enter via the condensation of oxaloacetate and acetyl CoA, and two carbons exits via successive decarboxylation. Even though it is the predominant cycle for the oxidation of acetyl CoA in eukaryotes, in prokaryotes sometimes some of the enzymes are missing or replaced by other enzymes catalyzing similar reaction present in the cycle. In *M. tuberculosis,* the absence of a functional α-ketoglutarate dehydrogenase has been found recently. An alternative route of α-ketoglutarate metabolism has been identified, based on the conversion of α-ketoglutarate to succinic semialdehyde by α-ketoglutarate decarboxylase. This is followed by the production of succinate via succinic semialdehyde dehydrogenase (Tian *et al.,* 2005).

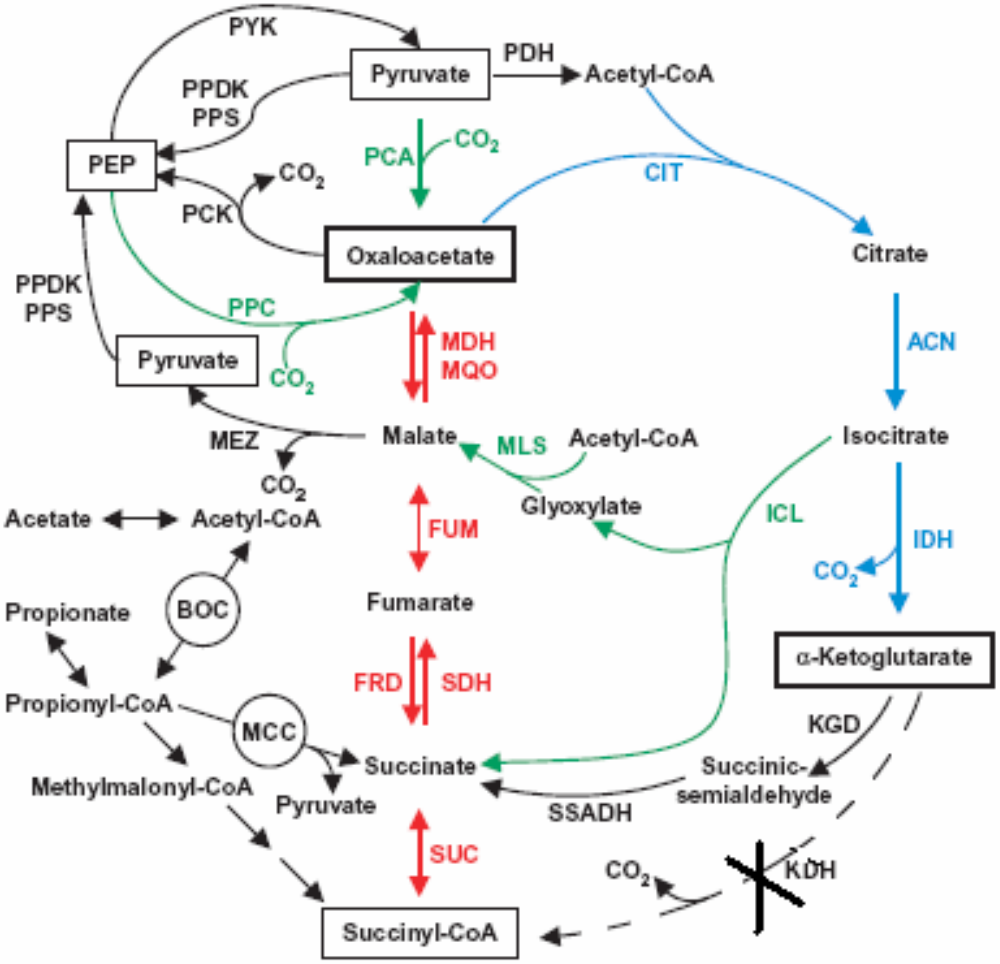


Fig. 5 - Citric acid cycle and related anaplerotic pathways showing the various precursors and their use, with the dotted line indicating the absence of enzyme in *M. tuberculosis*. The bifurcated biosynthetic mode of operation of the citric acid cycle, comprising oxidative (blue) and reductive (red) branches, is employed under aerobic conditions in excess glucose or anaerobically on any carbon substrate, and by bacteria that lack KDH/KGD. Depleted intermediates are replenished via anaplerotic reactions (green), catalysed by PPC or PCA during growth on excess carbohydrate, or by ICL and MLS during growth on fatty acids or limiting carbohydrate. MEZ, malic enzyme; PPC, PEP carboxylase; PCK, PEP carboxykinase; PPS, PEP synthase; PPDK, pyruvate phosphate dikinase; PYK, pyruvate kinase; PCA, pyruvate carboxylase; PDH, pyruvate dehydrogenase; CIT, citrate synthase; ACN, aconitase; IDH, isocitrate dehydrogenase; KGD, α -ketoglutarate decarboxylase; SSADH, succinic semialdehyde dehydrogenase; KDH, α -ketoglutarate dehydrogenase; SUC, succinyl-CoA synthetase; SDH, succinate dehydrogenase; FRD, fumarate reductase; FUM, fumarase; MDH, malate dehydrogenase; MQO, malate:quinine oxidoreductase; ICL, isocitrate lyase; MLS, malate synthase; BOC, β -oxidation cycle; MCC, methylcitrate cycle.

1.4.5. Anaplerosis

Anaplerotic (filling up) reactions are required to replenish citric acid cycle intermediates that are diverted to biosynthetic pathways. Different anaplerotic pathways may be used depending upon the available substrate and the electron acceptors. During growth on fatty acids, anaplerosis involves the glyoxylate cycle, which generates one molecule of malate from two molecules of acetyl CoA via sequential activity of isocitrate lyase (ICL) and malate synthase (Kornberg & Krebs, 1957). Glyoxylate cycle is the only other pathway capable of oxidation of acetyl CoA to CO_2 while simultaneously generating pyruvate and PEP for biosynthetic pathways, which may lead to the possibility that the TCA cycle may act in bifurcated mode (Fig. 5) in *M. tuberculosis* (Tian *et al.,* 2005).

1.5. Glycerol utilisation and intermediary metabolism: role in virulence and tropism

Mycobacteria are unusual in their carbon utilisation, with variation across its members. Glycerol is normally used as the carbon source in axenic media to get copious growth of many mycobacterial species. However, *M. bovis, M. africanum,* and *M. microti* were unable to grow when glycerol was used as a sole carbon source. Indeed, glycerol showed a pronounced effect on the size and appearance of the colony morphology. *Mycobacterium tuberculosis* displayed abundant, ‘eugonic’ and rough crumbly ‘rugose’ colonies while *M. bovis* colonies appear as flat, ‘dysgonic’, moist and glossy ‘smooth’. Morphology was used as a classical method of differentiating bovine and human bacilli. The effect could be due to the effect of glycerol and its metabolic derivatives in the biosynthesis of lipids and cell wall constituents (Keating *et al.,* 2005). Pyruvate, one of the important metabolic intermediate, can condense with glyceraldehyde-3 phosphate in the first step in the synthesis of various lipids (Khasnobis *et al.,* 2002).

SNP analysis of *M. tuberculosis* and *M. bovis* genomes revealed that the genes involved in the uptake and utilisation of glycerol are disrupted in the latter (Fig. 4). They include *glpK, pykA* and *ald. glpK* encodes glycerol kinase, which catalyses the conversion of phosphorylation of glycerol to glycerol-3-phosphate in the presence of Mg^{2+} and ATP. The frame shift mutation at codon 191 in *glpK* of *M.*

bovis and *M. microtii,* results in a truncated non-functional protein. A functional protein is however seen in *M. bovis* AN5, used for the production of tuberculin allowing its growth on glycerol medium. In BCG, there is a 2 bp insertion at codon 191, resulting in an extra codon with respect to *M. tuberculosis* H37Rv (Keating *et al.,* 2005).

pykA encodes pyruvate kinase (PK) which catalyses the conversion of phosphoenol pyruvate to pyruvate, the final step in glycolysis. *pykA* SNP results in the substitution of Glu_{220} by Asp. This highly conserved amino acid has been associated with the cofactor Mg^{2+} and substrate (ADP / ATP, Phosphoenolpyruvate) binding. The *pykA* SNP was found to be absent in *M. bovis* AN5 and BCG. However, SNP was universal among the field isolates of *M. bovis, M. africanum* and *M. microti,* these organisms failed to show any enzyme activity (Keating *et al.,* 2005).

L-Alanine dehydrogenase (Ald) was the first enzyme to be characterized from the culture filtrate of *M. tuberculosis* H37Rv. It is also a predominant antigen and one of the few functional antigens characterized from *M. tuberculosis* H37Rv and not from BCG (Anderson *et al.,* 1992). Ald, like many other extracellular proteins secreted by *M. tuberculosis* H37Rv lacks a signal peptide. Though Ald catalyses the reversible conversion of alanine to pyruvate, the conversion of pyruvate to alanine is preferred at physiological pH (Hutter & Singh, 1998). Over production of Ald during nutrient starvation and under hypoxia has been reported (Rosenkrands *et al.,* 2002).

1.6. Iron acquisition in mycobacteria

1.6.1. Role of iron and bioavailability

Iron is an essential nutrient. It is an important catalyst in numerous biological redox reactions of fundamental importance (Sritharan, 2000) and also serves as a cofactor for many vital enzymes mainly in the electron transport. Though it is the fourth most abundant element on earth, iron is biologically unavailable. Iron, at pH 7 exists as $Fe(OH)_2^+$, which has a solubility of approximately $1x10^{-09}$ M (Chipperfield & Ratledge, 2000), which is low to support growth. In the mammalian host also very little free iron is present as 99.9% of iron is held in protein-bound form.

Intracellularly, it is present in the storage protein ferritin. Extracellularly, it is bound to transferrin in the plasma and lymph and to lactoferrin in leukocytes and in mucosal secretions. The host limits the availability of free iron both by decreasing the absorption from the intestine as well as by increasing the synthesis of transferrin by a phenomenon called nutritional immunity (Kochan, 1976).

1.6.2. Mycobacterial siderophores

Mycobacteria are unique in that they produce two kinds of siderophores the intracellular mycobactins and the extracellular carboxymycobactins / exochelins. Exochelin is peptide in nature while the carboxymycobactin resembles the mycobactin. Based upon the type of siderophore(s) expressed, mycobacteria can be classified into four groups, namely those

1) producing mycobactin and both the types of extracellular siderophores e.g. *M. smegmatis*
2) producing mycobactin and carboxymycobactin, e.g. *M. tuberculosis*
3) producing only the exochelins and no mycobactin, e.g. *M. vaccae*
4) that not produce any siderophores e.g. *M. paratuberculosis*.

Mycobactin produced by mycobacterial species are unique to each species (Snow, 1970) and can be used as taxonomic markers (Table 2). But due to the sophisticated techniques required it is not practical to do so.

1.6.2.1. Intracellular mycobactins

Mycobactins are lipid soluble and are found in the cell wall of mycobacteria. The yield of mycobactin varies from species to species with the maximum reported in *M. smegmatis* which can go up to 10% of the dry weight and the least being *M. kansasii* gives only 0.05% or less. The structure of mycobactin is represented in Fig. 6. Mycobactin belong to the mixed ligand type wherein they have two hydroxamate groups and the third pair being provided by an oxygen atom on the aromatic residue and nitrogen in the oxazoline ring. Mycobactin from different mycobacterial species have the core nucleus that consists of a 2-hydroxyphenyloxazoline moiety linked by

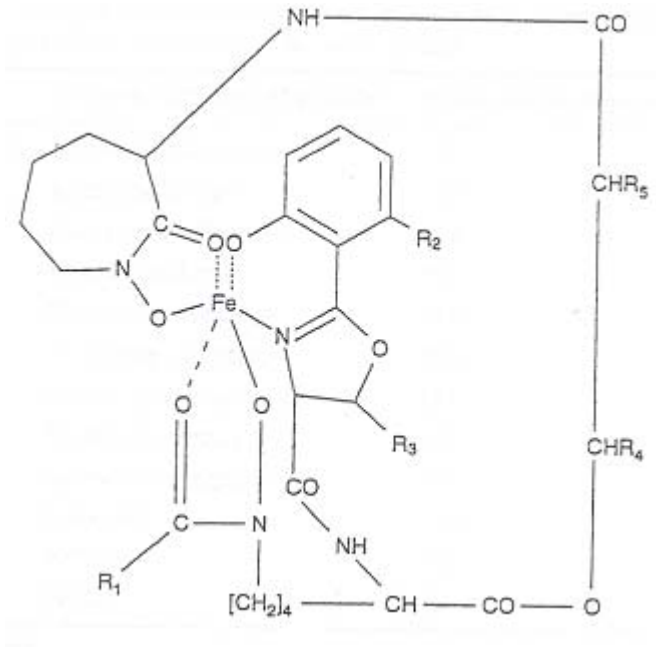


Fig. 6 - Structure of Mycobactin, where R_1 consists of alkenyl groups, R_2, R_3 and R_5 are either H or CH_3, R_4 are different alkyl groups.

an amide bond to an acylated ε-N-hydroxylysine residue. The second ε-N-hydroxylysine is cyclised to form the seven membered lactam and is attached to β-hydroxyacid via an amide bond. This, in turn is connected to the α-carboxyl of the first lysine residue. Within this core, a methyl group may or may not be present at the 6th position of phenolic ring and the 5′ position of the oxazoline. The variation in the structure occurs in the alkyl substituents of the hydroxyacids (R_3 and R_4) and the acyl moiety R_5. In general, the R_5 group is a long chain fatty acid that is unsaturated and with an unusual *cis* double bond conjugated to the carbonyl group. The different substituents in the side chains, accounting for the variation between the different species are explained in Table 2 (Dover & Ratledge, 2000).

Table 2 - Variations among the mycobactins produced by different mycobacterial species

Organism	Mycobactin	Substituents				
		R_1	**R_2**	**R_3**	**R_4**	**R_5**
M. aurum	A	13Δ	CH_3	H	CH_3	H
M. fortuitum	F	17, 11Δ	H	CH_3	H	
M. fortuitum	H	19, 17Δ	CH_3	CH_3	CH_3	H
M. marinum	M	1	H	CH_3	$C_{17}H_{35}$	CH_3
M. marinum	N	2	H	CH_3	$C_{17}H_{35}$	CH_3
M. phlei	P	17*cis*Δ	CH_3	H	C_2H_5	CH_3
M. terrae	R	19Δ	H	H	C_2H_5	CH_3
M. smegmatis	S	17, 15*cis*Δ	H	H	CH_3	H
M. tuberculosis	T	19Δ	H	H	CH_3	H
M. avium	Av	Δ2 alkenyl	H	H	$C_{10}H_{23}$	CH_3
M. intracellulare	Av	Δ2 alkenyl	H	CH_3	Satd. alkyl	CH_3
M. scrofulaceum	Av	Alkenyl	H	H	Satd. alkyl	CH_3
M. paratuberculosis	Av	Δ3 alkenyl	H	CH_3	Satd. alkyl	CH_3
M. paratuberculosis	J	15Δ	H	H	Isopropyl	CH_3
Nocardia asteroids	NA		H	CH_3	Satd. alkyl	CH_3

1.6.2.1.1. Biosynthesis of mycobactins

For the synthesis of mycobactin, basic building blocks include L-lysine that serves as the precursor of both the ε-N-hydroxyllysine moieties. β-hydroxy butyrate is derived from the condensation of acetate or propionate molecules, resulting in the diversification of the mycobactins' structure in the R_4 and R_5 side chain (Tateson, 1970). The salicylate moiety, as seen in *M. smegmatis* was considered to be synthesised by the shikimate pathway, while the 6-methyl salicylate present in *M. tuberculosis* was shown to be synthesized by the condensation of four acetate units by polyketide synthesis (De Voss *et al.,* 1999). No studies have been carried out on the origins of the oxazoline, but serine and threonine are presumed to be the

precursors of the unsubstituted and methyl substituted rings respectively (Snow, 1970).

A cluster of 10 genes (*mbt A-J*) called the *mbt* operon involved in the biosynthesis of mycobactin (Fig. 7) is identified from the *M. tuberculosis* genome. They encode two polyketide synthetases (MbtC and MbtD), three peptide synthetases (MbtB, MbtE and MbtF), and isochorismate synthase (MbtI) (Quadri *et al.,* 1998). MbtF, MbtE and MbtB are highly homologous to the non-ribosomal peptide synthases (NRPSs) seen in the synthesis of yersiniabactin. They are large modular enzymes that activate the amino acids as their acyl adenylates through an activating domain and covalently link them to the enzyme via a phosphopantethenyl thioester attached to a carrier protein domain. The activated amino acids then condense to form a peptide bond, which is catalysed by a condensation domain. The activation domain determines the sequence in which the amino acids are inserted. There are several other domains such as epimerization domains, cyclisation domains and thioesterase domains that are also present in the NRPSs. MbtA activates salicylate as an acyl adenylate that is attached to a phosphophantetheine prosthetic group of MbtB. MbtB activates serine, and then condenses with salicylic acid and cyclises the product to hydroxyphenyloxazoline. MbtE and MbtF then carry out the activation, condensation and act as peptide carrier domains for the donation of the two lysine derivatives. MbtC and MbtD are involved in the production of β-hydroxy butyrate. *mbtG* encodes ornithine and lysine oxygenases which are required for the production of ε-N-hydroxyllysines (Quadri *et al.,* 1998). Rv1347c, previously annotated as aminoglycoside N-acetyltransferase, was recently identified as an essential gene immediately flanking the *mbt* genes, as the missing enzyme in the biosynthesis of mycobactin, which adds a long chain acyl group to N-hydroxyllysine side chain. Since, this is also regulated by iron levels in mycobacteria, it has since been re-annotated as the gene involved in the biosynthesis of mycobactin (Card *et al.,* 2005).

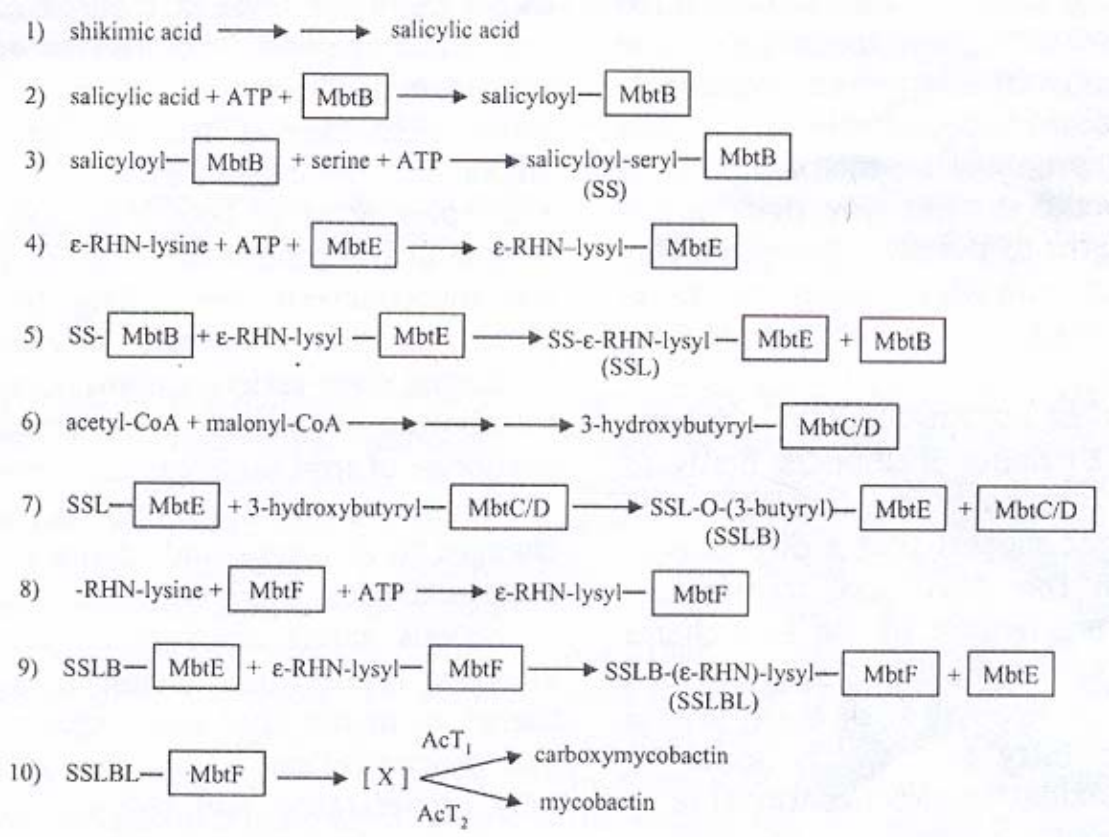


Fig. 7 - Schematic representation of the reactions in the biosynthesis of mycobactins and carboxymycobactins (Ratledge, 2004)**.**

1.6.2.2. Carboxymycobactin

The carboxymycobactins are water-soluble extracellular siderophores, which are structurally related to mycobactins. The lipophilicity of the mycobactins is modified by the shorter carboxyl chain and hence the name 'carboxymycobactin' (Lane *et al.,* 1998). The carboxymycobactins are produced mainly in the pathogenic mycobacterial species (Gobin *et al.,* 1999), but they also occur in small quantities in non-pathogenic mycobacteria (Ratledge & Ewing, 1996). They can solublise iron from transferrin, lactoferrin and ferritin. The synthesis of carboxymycobactin is similar to mycobactin and differs from mycobactin after the formation of the common precursor (X) (Fig. 7). The final step in the synthesis of carboxymycobactin differs as the alkyl chain terminates in a carboxylic acid suggesting dicarboxylic acid may be involved in the final acyl transferase. However,

nothing is known about this final step and awaits further investigation (Ratledge, 2004).

1.6.2.3. Exochelins

Exochelins are water-soluble, ornithine derived siderophores produced by non-pathogenic mycobacteria and well characterised in *M. smegmatis* and *M. neoaurum* (Fig. 8). The exochelins are peptidyl in nature but involve D-amino acids and thus there are no conventional peptide bonds present (Ratledge & Dover, 2000). The coordination center with Fe^{3+} is hexadendate in an octahedral structure involving the three-hydroxamic acid groups, which are from ornithine. The exochelin MS from *M. smegmatis* is a formylated pentapeptide derived from three molecules of δ-N-hydroxyornithine, β-alanine and a threonine (Fig. 8). Exochelin MN from *M. neoaurum* is a hexapeptide with two δ-N-hydroxyhistidines, providing the coordination center for iron chelation and two β-alanine and an ornithine (Sharman *et al.,* 1995a and 1995b).

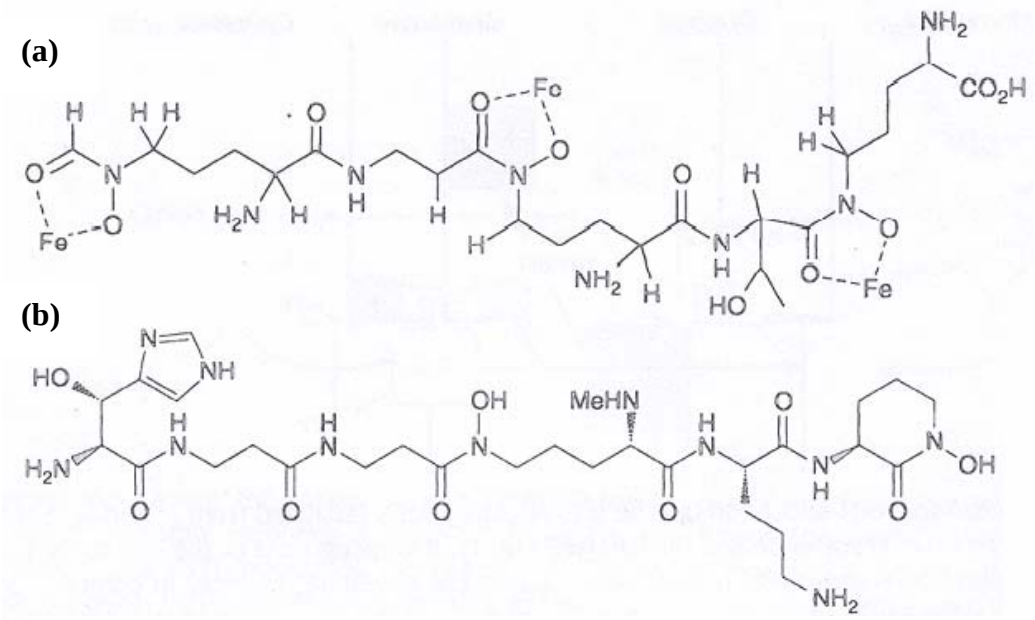


Fig. 8 - Structure of exochelins (a) MS and (b) MN from *M. smegmatis* and *M. neoaurum* respectively (Sharman *et al.,* 1995 a & b).

1.6.2.3.1. Biosynthesis of Exochelin

The genes involved in the synthesis of exochelin MS have been identified and designated as *FxbA, FxbB* and *FxbC*. *FxbA* codes for N-formyltransferase,

which is homologous to phospho-ribosylglycineamide formyltransferases, which attaches the terminal formyl group to the pentapeptide (Yu *et al.,* 1998). The other two proteins are large non-ribosomal synthases involved in the synthesis of various small peptides and form the backbone of the exochelin (Zhu *et al.,* 1998; Fiss *et al.,* 1994). Not much is known about their role in synthesis.

1.6.3. Ferri siderophore uptake in mycobacteria

1.6.3.1. Ferri-exochelin

Uptake of ferri-exochelin has been well studied in *M. smegmatis* (Ratledge, 2004). It is thought to be an active transport as it is inhibited by energy poisons and uncouplers of oxidative phosphorylation. Uptake involves the complete transfer of the molecule along with the metal ligand. Several proteins are involved in the uptake process that includes a 29 kDa ferri-exochelin receptor (Hall *et al.,* 1987). After recognition by the receptor, the ferri-siderophore complex is taken up by the FxuD protein and then transferred through the cytoplasmic membrane proteins FxuA, FxuB and FxuC, which share amino acid sequence homology with FepG, FepC and FepD, which are involved in the uptake of ferri-enterochelin in *E. coli* (Fig. 9) (Ratledge, 2004).

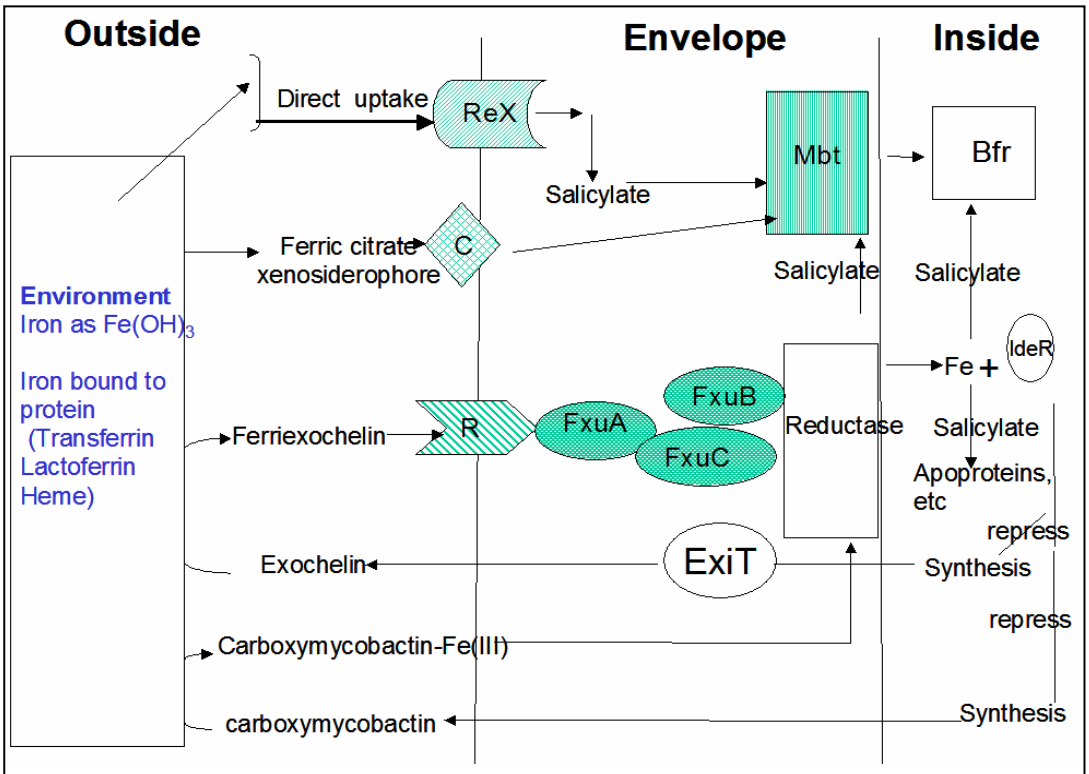


Fig. 9 - Schematic representation of ferri-siderophore uptake, where Rec receptor; Fxu ferriexochelin uptake proteins; ExiT exochelin transporter (Sritharan, 2000).

1.6.3.2. Ferri Carboxymycobactin

The mechanism of uptake of carboxymycobactin is not known completely. It is thought that it might traverse the cell envelope either by diffusion, by virtue of its hydrophobicity (Rodriguez & Smith, 2006), or it is transported via a porin-like molecule. The hypothesis involving porins is supported by the size of the inner diameter of the porin molecule, which is about 2.2 nm, and the diameter of the carboxymycobactin may be equivalent to the ferri-mycobactin *i.e.* 1.1-1.4 nm (Trias & Benz, 1992). Recently, the role of an ABC transporter for the internalization of carboxymycobactin has been reported (Rodriguez & Smith, 2006). The study also speculates the role of mycobactin in carboxymycobactin transport as Fe-carboxymycobactin can be fed exogenously in siderophore biosynthesis mutants. The same study also annotated both Rv1348 and Rv1349, the two genes flanking the mycobactin biosynthesis region, as *irtA* and *irtB* respectively (Rodriguez & Smith, 2006). These genes code for an ABC transporter, which is thought to internalize carboxymycobactin, a system highly similar to the YbtPQ system of *Yersinia pestis* (Fetherson *et al.,* 1999).

ABC (ATP binding cassette) transporters have a highly conserved ATPase domain, which binds and hydrolyses the ATP to provide energy for the transport process. They are classified as exporters and importers, the former also known as permeases. They generally require a substrate binding protein (SBP) as well as a translocator in the cytoplasmic membrane in gram-positive bacteria (Garmony & Titball, 2004). IrtA has been speculated to act as the substrate-binding protein due to its N-terminal domain. Hence, the transport of carboxymycobactin across the cell envelope may occur with the aid of a porin, a typical SBP or the N-terminal region of the IrtA and then translocated into the cytoplasm by IrtAB (Rodriguez & Smith, 2006). This hypothesis needs to be proved as it contradicts the passive transport proposed by Stephenson & Ratledge (1980).

1.6.4. Release of iron from mycobacterial siderophores

According to the hypothesis proposed by Ratledge (2004), Fe^{3+} in mycobactin is reduced to Fe^{2+} by NADP (H)-dependent mycobactin reductases (Brown & Ratledge, 1975). Even though salicylate does not form a stable complex at pH 7, Fe^{2+} is thought to be chelated by salicylate. However, ferrisalicylate may

become stable when mycobactin, ferrimycobactin reductase and ferrichelatase aggregate and help in the transfer of ferrous ion as ferrisalicylate (Fig. 9). It is likely that ferri-carboxymycobactin will be reduced similarly (Dailey & Dailey, 2002). In the case of exochelin a membrane reductase reduces ferri-exochelin and incorporates the ferrous form into apoproteins. If the acceptor molecule is not available, Fe^{3+} from both the exochelin and carboxymycobactin will be sequestered by mycobactin and the iron will be released from the mycobactin.

Mycobacteria are unique in that they produce two types of siderophores, but the inter-relationship and the necessity for two is not yet clear. Mycobactin probably acts as a temporary iron storage molecule when iron becomes sufficient. It has been demonstrated that during iron deprivation mycobacteria represses the synthesis of porphyrin. When there is increase in the iron concentration, the cells may not have the precursor molecules to synthesise heme precursor molecules. This may leads to iron toxicity and increase in the oxidative damage to the bacteria. In order to avoid this, mycobactins are utilised as a temporary iron storage molecule. Mycobactin is found within the cell envelope of the mycobacteria and appears to be located physically next to the cytoplasmic membrane but it is not part of the membrane structure itself nor is it within the cytoplasm (Ratledge *et al.*, 1982, 2004). When iron becomes available to the cell, it derepresses the synthesis of various molecules. Once iron storage proteins are synthesised, iron is mobilised from mycobactin in a controlled manner and balances the requirement for iron by the newly synthesized molecules.

The relationship between the extracellular and intracellular siderophores may be that exochelins acquire the iron from the external environment and deliver it to the intracellular mycobactins (Ratledge, 2004).

1.6.5. Iron-regulated envelope proteins (IREPs)

Several bacteria express iron regulated envelope proteins (IREPs) / Iron regulated membrane proteins (IRMPs) that functions as receptors for ferric siderophores. Mycobacteria also express IREPs as part of the iron acquisition machinery with Hall *et al.* (1987) demonstrating a 29 kDa IREP as a receptor for the ferric-exochelin MS, as a Anti-29 kDa antibodies inhibited the uptake of the ferric exochelins (Dover & Ratledge, 1996). Subsequently, IREPs were also identified in

other mycobacterial species not only under defined lab conditions of established iron status but also under *in vivo* conditions. IREPs were identified from armadillo derived *M. leprae* and *M. avium* recovered from experimentally infected animals (Sritharan & Ratledge, 1990).

1.6.6. Regulation of iron metabolism

The iron acquisition machinery is well understood in *E. coli.* The expression of the siderophores and IRMPs is regulated by the Fur protein (Ferric Uptake Regulator) encoded by *fur* (Fig. 10). The corresponding homologue in gram-positive bacteria is DtxR and IdeR in mycobacteria.

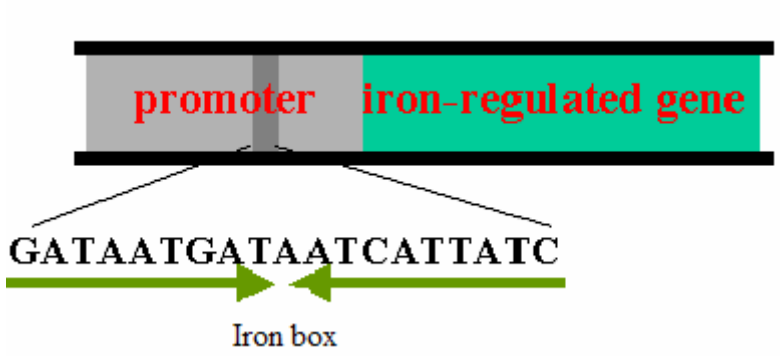


Fig. 10 - Iron regulation by *Fur* in *E. coli*

1.6.6.1. Regulation of iron metabolism in Mycobacteria

From the genome sequence of *M. tuberculosis* four regulators, two belonging to the Fur family (FurA and FurB) and two belonging to DtxR family (IdeR and SirR) were identified. IdeR is recognized as the main regulator of iron metabolism. It is present in both pathogenic as well as non-pathogenic mycobacteria. Of the Fur-like proteins, FurA negatively regulates the expression of *katG* and *furA* is located immediately upstream of the *katG* gene. Both *furA* and *katG* are pseudogenes in *M. leprae.* Fur B is shown to be involved in the zinc ion homeostasis, while the function of Sir R is not yet known.

IdeR, sharing 90% homology with DtxR proteins in the first 180 amino acids, contains an N-terminal region (1-73 residues) containing a helix-turn-helix motif that binds DNA. The dimerisation domain (74-120 residues) bears most of the metal-binding residues and a third domain in the carboxy terminal is free and has a

SH_3-like fold suggesting a possible interaction with other proteins. In addition to iron, IdeR can also bind other divalent metal ions such as Mn, Zn, Co, Ni, and Mg. Four monomers of IdeR form two functional dimers, and metal binding activates the protein's DNA binding ability by causing conformational changes in the DNA binding domain (Pohl *et al.,* 1999). In the presence of iron IdeR binds to a 19 bp (5'-TTAGGTTAGGCTAACCTAA-3') sequence called the IdeR box in the vicinity of promoter region of iron-regulated genes. It is a dual function regulator; under high iron conditions it represses siderophore biosynthesis and induces iron storing bacterioferritin (Bfr) (Gold *et al.,* 2001). Smith and his group showed by microarray analysis four groups of iron regulated genes namely (a) Iron and IdeR induces genes, (b) IdeR-independent, iron repressed genes, (c) Iron and IdeR repressed genes and (d) IdeR-independent iron induced genes (Table 3a, b, c and d respectively).

Table 3a - Iron and IdeR induced genes (Rodriguez *et al.,* 2002)

Rv No. / Gene	**Gene product**
PpiA	Peptidyl-prolyn-*cis-trans* isomerase
0338c	Iron-sulfur protein
BfrA	Bacteriferritin
BfrB	Ferritin- like protein

Table 3b - IdeR independent, iron repressed genes (Rodriguez *et al.,* 2002)

Rv No. / Gene	**Gene Product**
PE-PGRS	PE-PGRS protein
0464c	CHP
0465c	Probable transcriptional regulator (PbsX/Xre family)
ICL	Isocitrate lyase
MmpL5	Conserved large membrane protein
MmpS5	Conserved small membrane protein
PqqE	Coenzyme PQQ synthesis protein E
IldDI	Lactate dehydrogenase (cytochrome)
1169c	PE protein
1184c	CHP
1195	PE protein
1393c	FAD containing mono-oxygenase
1461	CHP
1462	CHP
1463	Probable ABC transporter
1464	NifS like protein
1465	NifU like protein
1466	CHP
1520	Probable glycosyltransferase
2621c	Probable transcriptional regulator
2794c	CH; similar to proteins involved in vibriobactin and enterobactin Synthesis
DesA3	Acyl [ACP] desaturase
3230c	Probable oxygenase
3614c	CHP
EthA	Probable mono-oxygenase

Table 3c - Iron and IdeR repressed genes (Rodriguez *et al.,* 2002)

Rv No. / Gene	Gene product
0116c	CHMP
0282	CHP
0283	CHMP
0284	CHMP
0285	PE subfamily
0286	PPE subfamily
0287	Probable transcriptional regulator; PE subfamily
0288	CHP; ESAT –6 family
0289	CHP
0290	Unknown
0291	Probable secreted protease
0292	Unknown
MmpL4	Cytochrome conserved large membrane protein
mmps4	Cytochrome conserved small membrane protein
YrbE2A	Part of *mce2*operon
1343c	CHP
1344	Probable acyl carrier protein
FadD33	Acyl-coA synthase
FadE14	Acyl-coA dehydrogenase
1347c	Iuc B like protein
1348	YbtP like protein
1349	YbtQ like protein
1519	CHP
HisE	Phophoribosyl-AMP cyclohydrolase
2123	PPE subfamily
MbtH	Mycobactin synthesis
MbtG	Lysine-N-Oxygenase mycobactin synthesis
MbtF	Peptide synthase; mycobactin synthesis
MbtE	Peptide synthase; mycobactin synthesis
MbtD	Polyketide synthase; mycobactin synthesis
MbtC	Polyketide synthase; mycobactin synthesis
MbtB	Peptide synthase; mycobactin synthesis
MbtA	Salicylate-AMP ligase; mycobactin synthesis
LipK	Probable acetyl-hydrolase
MbtI (trpE2)	Isochorismate synthesis; mycobactin synthesis
3402c	EryCIV like protein
3403c	Unknown
3839	Unknown
3840	Probable transcriptional regulator

Table 3d - IdeR independent iron induced genes (Rodriguez *et al.,* 2002)

Rv No. / Gene	Gene product
RplV	50s ribosomal protein L22
IprE	Lipoprotein
AtpE	ATP synthase C chain
katG	Catalase-peroxidase
1943c	CHP
2526	CHP
2549c	Unknown
2550c	Unknown
PE-PGRS	PE-PGRS family
2927c	CHP
3075c	CHP
NuoA	NADH dehydrogenase chain A
NuoB	NADH dehydrogenase chainB
NuoC	NADH dehydrogenase chain C
NuoD	NADH dehydrogenase chain D
NuoH	NADH dehydrogenase chain H
NuoI	NADH dehydrogenase chain I
NuoK	NADH dehydrogenase chain K
NuoL	NADH dehydrogenase chain L
NuoM	NADH dehydrogenase chain M
NuoN	NADH dehydrogenase chain
MtrA	Two-component response regulator
3394	Unknown

1.6.7. Storage of iron

Many bacteria have evolved strategies to store the excess of iron in storage proteins. These iron stores can be used when the external iron supply is restricted (Andrews, 1998). Iron storage proteins produced in bacteria are of three types; the

ferritins, the bacterioferritins and the Dps proteins. The ferritins are present in eukaryotes, while eubacteria contain bacterioferritins and the Dps are present only in prokaryotes. All these three types of proteins share common structural properties and key molecular structure, which enable them to act as iron storage molecules. Bacterioferritins differ from ferritins in that they are heme proteins containing protoporphyrin IX as their prosthetic group. These are composed of 24-mers (ferritins and bacteriferritins) and 12-mers (Dps proteins) forming an approximately spherical protein with the central cavity for Fe^{3+}. Each subunit is folded to form a four α-helix bundle. The large bacterioferritin can accommodate at least 2000-3000 iron atoms per 24-mer while the Dps proteins have a lower storage capacity of 500 iron atoms per 12-mer. In mycobacteria, even though iron can be stored in mycobactin, it must be prevented from premature and unwanted release into the cytoplasm. Crystallographic and biochemical characterisation of an 18.2 kDa protein found to be up-regulated under iron excess was potentially identified as bacterioferritin. This is located within the cytoplasm and donates the iron whenever it is required to all the iron requiring proteins (Pessolani *et al.*, 1994).

1.6.8. Iron and virulence

Iron regulates not only the iron acquisition machinery but also the expression of virulence factors / toxins in several bacterial systems namely diphtheria toxin by *Corynebacterium diphtheriae,* the Shiga toxin by *Shigella* spp., exotoxin A from *Pseudomonas aeruginosa*, haemolytic toxin of *Vibrio cholerae,* vero – cytotoxin of enterohaemorrhagic *Escherichia coli* (Sritharan, 2000). In *M. tuberculosis*, the ability to synthesise mycobactin is recognized as a virulence determinant as demonstrated by De Voss *et al* (2000) as mutants defective in the biosynthesis of mycobactin failed to grow within the macrophages. IdeR mutants showed increased sensitivity to hydrogen peroxide due to the decreased activity of catalase- peroxidase (KatG) and superoxide dismutase (SodA) activity (Dusserget *et al.,* 1996). By the introduction of a dominant Dtx (E175K) iron-dependent repressor into *M. tuberculosis,* it has been proved that the IdeR repressor-controlled events influenced the virulence in murine model of infection (Manabe *et al.,* 2005).

1.7. Mycobacterial genomics

Progress was slow in understanding several aspects of metabolism in mycobacteria as it was difficult to genetically manipulate the bacteria. However, the completion of the genome sequence of *M. tuberculosis* generated the first significant approach for the understanding of these organisms (Cole *et al.,* 1998). The table 4 lists the various mycobacterial members whose genomes have been sequenced or are ongoing.

Table 4 - Genome sequencing projects for various mycobacteria

Name of the species	Web site for Sequence information	Progress as on 30-06-2006
M. tuberculosis	http://genolist.pasteur.fr/TubercuList/	Cole *et al.,* 1998
M. bovis	http://www.sanger.ac.uk/Projects/M.bovis	Garnier *et al.,* 2003
M. bovis BCG	http://www. Pasteur.fr/recherché/unites/ Lgmb/	Available but not published
M. leprae	http://www.sanger.ac.uk/Projects/M_lepra e	Cole *et al.,* 2001
M. avium	http://www.tigr.org/tdb	
M. paratuberculosis	http://www.cbc.umn.edu/ResearchProjects /AGAC/Mptb/Mptbhome.html	To be completed by 2006
M. microti	http://www.sanger.ac.uk/Projects/M_micr oti/	Shotgun in progress
M. marinum	http://www.sanger.ac.uk/Projects/M_mari num/	Finished, But not published
M. ulcerans	http://www. Pasteur.fr/recherché/unites/ Lgmb/	Shotgun in progress
M. smegmatis	http://www.tigr.org/tdb	To be completed by 2006
M. africanum	http://www.sanger.ac.uk/sequencing/Myc obacterium/africanum/	Shotgun in progress
M. canetti	http://www.sanger.ac.uk/sequencing/Myc obacterium/canetti/	Shotgun in progress
M. tuberculosis A1, C, EKAT-4, F11, Haarlem, Peruvian1, Peruvian2, W148	http://www.broad.mit.edu/annotation/micr obes/mycobacterium_tuberculosis/	
M. flavescens PYR-GCK;	http://genome.jgi-psf.org/draft_microbes /mycfl/mycfl.home.html	
Mycobacterium vanbalenii	http://genome.jgi-psf.org/draft_microbes /mycva/mycva.home.html	

1.7.1. The genome of *M. tuberculosis*

The genome of *M. tuberculosis* comprises of 4, 411, 529 bp, contains 3924 genes (Fig. 10) (Cole *et al.,* 1998). The genome has an average G + C content of 65.6 % (Cole, 1999). The general classification of *M. tuberculosis* annotated genes is presented in Table 5. The *M. tuberculosis* genome has some unusual features like the number of genes involved in fatty acid metabolism and the presence of unrelated PE and PPE families of acidic, glycine rich proteins.

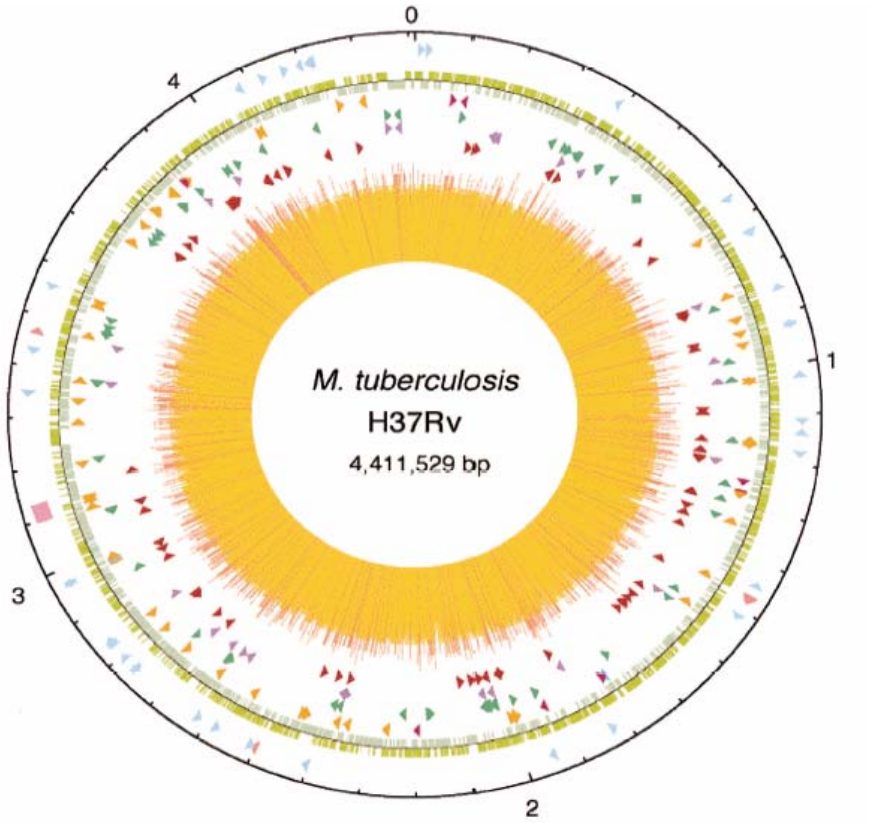


Fig. 11 - Circular map of the chromosome of *M. tuberculosis* H37Rv. The outer circle shows the scale in mega bases, with 0 representing the origin of replication. The first ring from the exterior denotes the positions of stable RNA genes (tRNAs are blue, and others are pink) and the direct-repeat region (pink cube); the second ring shows the coding sequence by strand (clockwise, dark green; anticlockwise, light green); the third ring depicts repetitive DNA (insertion sequences, orange; 13E12 REP family, dark pink; prophage, blue); the fourth ring shows the positions of the PPE family members (green); the fifth ring shows the positions of the PE family members (purple, excluding PGRS); and the sixth ring shows the positions of the PGRS sequences (dark red). The histogram (center) represents the G_C content, with _65% G_C in yellow and _65% G_C in red. (Cole *et al.,*1998).

Table 5 - General classification of annotated *M. tuberculosis* genes

Function	No. of genes annotated	% of total genes	% of total coding capacity
Lipid metabolism	225	5.7	9.3
PE & PPE proteins	167	4.2	7.1
Cell wall & cell process	517	13	15.5
Information pathways	877	22	24.6
Regulatory proteins	188	4.7	4.0
Virulence, detoxification and adaptation	91	2.3	2.4
IS elements and bacteriophages	137	3.4	2.5
Conserved hypothetical function	911	22.9	18.4
Unknown function	607	15.3	9.9
Stable RNAs	50	1.3	0.2

1.7.1.1. Fatty acid metabolism

The information obtained from whole genome analysis has provided valuable insight into the life style of the tubercle bacillus and highlighted the importance of lipid metabolism. Even though the cell envelope of bacillus was made up of a remarkable array of lipids, glycolipids, lipoglycans and polyketides (Daffe & Draper, 1998), it was rather unusual to find numerous genes conferring lipolytic functions. This may be because of the abundance of lipids and sterols, which are as potential substrates by these pathogens (Wheeler & Ratledge, 1994).

1.7.1.2. The PE and PPE gene families

The other major feature of *M. tuberculosis* genome is the sequence repeats referred to as PGRS (polymorphic G + C rich sequence) and MPTR (major polymorphic tandem repeat). The regions corresponding to the PE and PPE family proteins are exceptionally G + C rich, being 80% (Cole, 1999). About 8% of the genome is occupied by these two protein families (Cole, 2002). These proteins are unique for mycobacteria as they show no similarity to any known proteins (Gordon *et al.,* 2002). It has been suggested that both PE and PPE proteins may represent

antigens of immunological relevance. Rv3367 (PE-PGRS) gene product was recognised by pooled sera of TB patients and not from healthy controls (Sherman *et al.,* 2001). Some of these proteins were found to be localised in the cell wall and cell membrane of *M. tuberculosis* (Banu *et al.,* 2002; Delogu & Brennan, 2001). Disruption of Rv1818c in *M. tuberculosis* has reduced bacterial clumping, indicating their role in cell-cell adhesion and phagocytosis of mutant cells is also reduced. Transposon insertion of Rv1818c in *M. bovis* BCG Pasteur affected the ability of this mutant to enter and survive in macrophages, suggesting a role of these proteins in virulence (Brennan *et al.,* 2001). It has also been shown that a *M. marinum* mutant that encodes a protein homologous to the *M. tuberculosis* PE-PGRS family was attenuated for survival in macrophages and granulomas (Ramakrishnan *et al.,* 2000).

1.7.1.3. Other features of *M. tuberculosis* genome

The *M. tuberculosis* sequence contains 56 copies of insertion sequences with many of them inserted into intergenic or non-coding regions (Cole *et al.,* 1998). Two prophages (phiRv1 and phiRv2) were also detected in the chromosome, with phiRv1 corresponding to the RD3 locus that was shown to be absent in the vaccine strain *M. bovis* BCG (Mahairas *et al.,* 1996). The presence of 13 sigma factors, 11 serine-threonine protein kinases and 103 repressors and activators within the genome suggests that the metabolism of the bacillus is regulated in response to environmental stimuli (Gordon *et al.,* 2002). The genes encoding ABC transporters occupy about 2.5% of the genome of *M. tuberculosis*, with many of them showing similarity to transporters involved in the export of drugs (Braibant *et al.,* 2000). This probably contributes to resistance of the tubercle bacilli to many antibiotics. In contrast to *E. coli* and *B. subtilis,* the number of importers encoded by the genome is relatively few, indicating that *M. tuberculosis* can synthesise many essential compounds. The genome sequence also revealed 20 enzyme systems predicted to utilise cytochrome P450 as a cofactor, which is often involved in the degradation of xenobiotics or modified sterols, enabling them to utilise various compounds as sources of carbon and energy (Aoyama *et al.,* 1998). This may have a significant impact on the survival of the bacilli in different environments (Braibant *et al.,* 2000).

1.7.2. Comparative genomics

Comparative sequence analysis of orthologous genes (genes that perform the same function) from different bacteria is the basis of evolutionary relatedness. However, species' phylogenies based on the comparison of single genes are often inconsistent. This is due to the high rate of horizontal gene transfer in bacteria (Sassetti & Rubin, 2002). Comparative genomics presents an attractive tool for evolutionary analysis of strain relatedness, as whole genomes can be examined rather than just individual genes (Gordon *et al.,* 1999). A study of genetic variability within natural populations of pathogens may provide insight into their virulence. Comparison of the genomes of closely related organisms can also be useful for making biological predictions for organisms where traditional genetic techniques are difficult. For example, comparison of the *M. leprae* genome with its close relative, *M. tuberculosis* revealed that this intracellular pathogen had lost more than a quarter of its genes and as a result it seems to have a limited repertoire of metabolic responses (Cole *et al.,* 2001; Cole, 1998). Furthermore, comparative genomic analysis of the virulent and avirulent strains of the same organism presents a powerful approach for the generation of diagnostic tests. Comparative genomics of *M. bovis* and *M. bovis* BCG revealed that the *esat-6* gene was present in *M. bovis* but absent from all the vaccine strains (Gordon *et al.,* 2001; Mahairas *et al.*, 1996). Previously, the Esat-6 protein was shown to be a suitable antigen to differentiate between *M. bovis* infected and BCG vaccinated cattle (Vordermeier *et al.,* 1999).

1.7.3. Comparative genomics of the *M. tuberculosis* complex

The *M. tuberculosis* complex contains 5 pathogenic species that share identical 16S rRNA sequences and over 99.9% nucleotide identity (Sreevatsan *et al.,* 1997; Garnier *et al.,* 2003). They include *M. tuberculosis*, *M. africanum*, *M. microti, M. canetti* and *M. bovis*. The members of *M. tuberculosis* complex differ in terms of their host range, phenotype and virulence for humans (Brosch *et al.,* 2000a).

1.7.3.1. Single Nucleotide Polymorphisms (SNPs)

SNPs occur in the genomes of members of the *M. tuberculosis* complex but at a relatively low level for a bacterium of 1 in every 2000-4000 bp, depending upon the species (Sreevatsan *et al.,* 1997). SNP analysis not only reveals the conservation

of gene sequence across members of the *M. tuberculosis* complex, but also the divergence of *M. bovis* from *M. tuberculosis*. Some of the SNPs results in the phenotypic change e.g. *pncA* gene responsible for pyrazinamide resistance (Scorpio & Zhang, 1996), while most of them are silent.

SNPs can be classified as synonymous (sSNPs) and non-synonymous (ns) SNPs based on their effect on the expression of the gene and on the phenotype. There are 2437 and 2423 SNPs between *M. bovis* and *M. tuberculosis* H37Rv and CDC1551 respectively. Direct comparison of 2504 coding sequences (CDS) of identical length across the three genomes revealed that 1629 and 1656 *M. bovis* CDS are identical in *M. tuberculosis* H37Rv and CDC1551 respectively (Fig. 12).

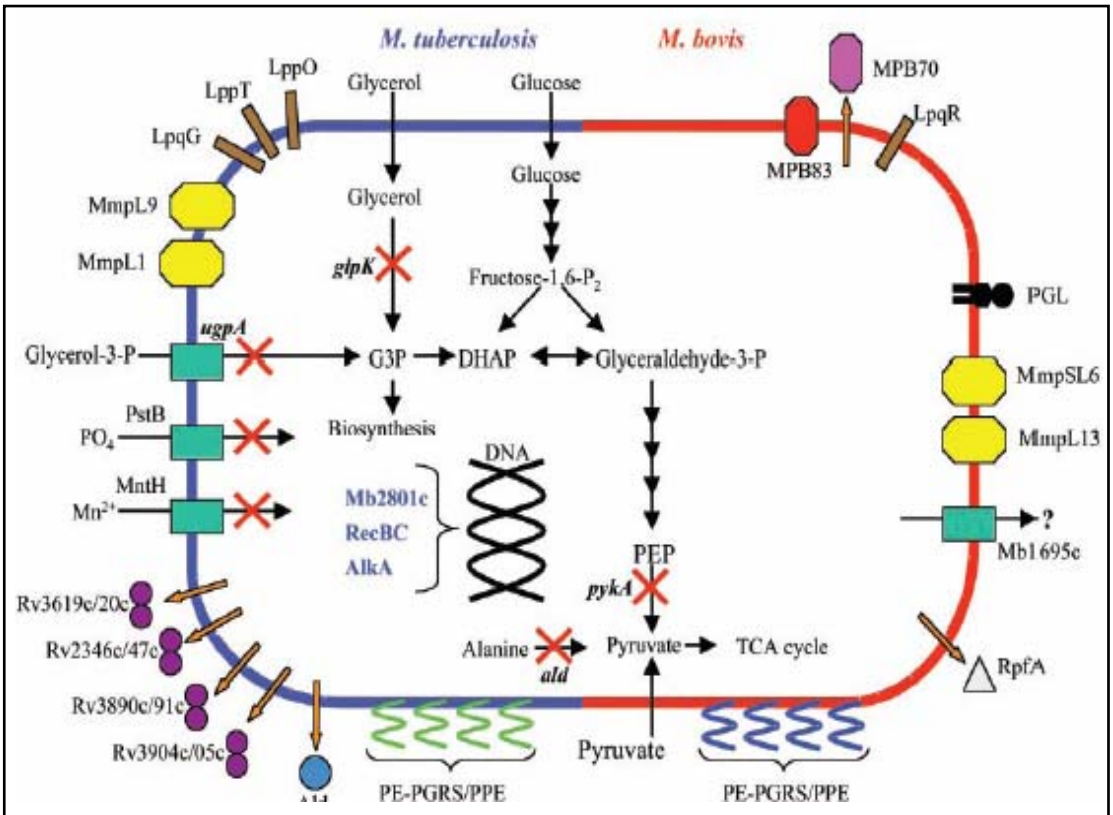


Fig. 12 - Schematic representation of major differences due to SNPs between *M. tuberculosis* and *M. bovis* (Garnier *et al,* 2003)**.**

This compares to 2,082 CDS that show no difference between the two *M. tuberculosis* strains. Across these selected CDS, *M. bovis* showed 506 synonymous and 769 non-synonymous SNPs compared with *M. tuberculosis* H37Rv, with 506

synonymous and 800 non-synonymous SNPs against *M. tuberculosis* CDC1551. The two *M. tuberculosis* strains showed 339 non-synonymous and 241 synonymous SNPs, respectively. The unexpectedly high frequency of non-synonymous to synonymous changes may be a product of the close evolutionary relationship between these strains (Garnier *et al.,* 2003). Comparative genome analysis of *M. bovis* BCG against *M. bovis* (Fig. 13) showed 736 SNPs while it is 2379 against *M. tuberculosis*. 440 nsSNPs and 204 sSNPs were identified between *M. bovis* and *M. bovis* BCG.

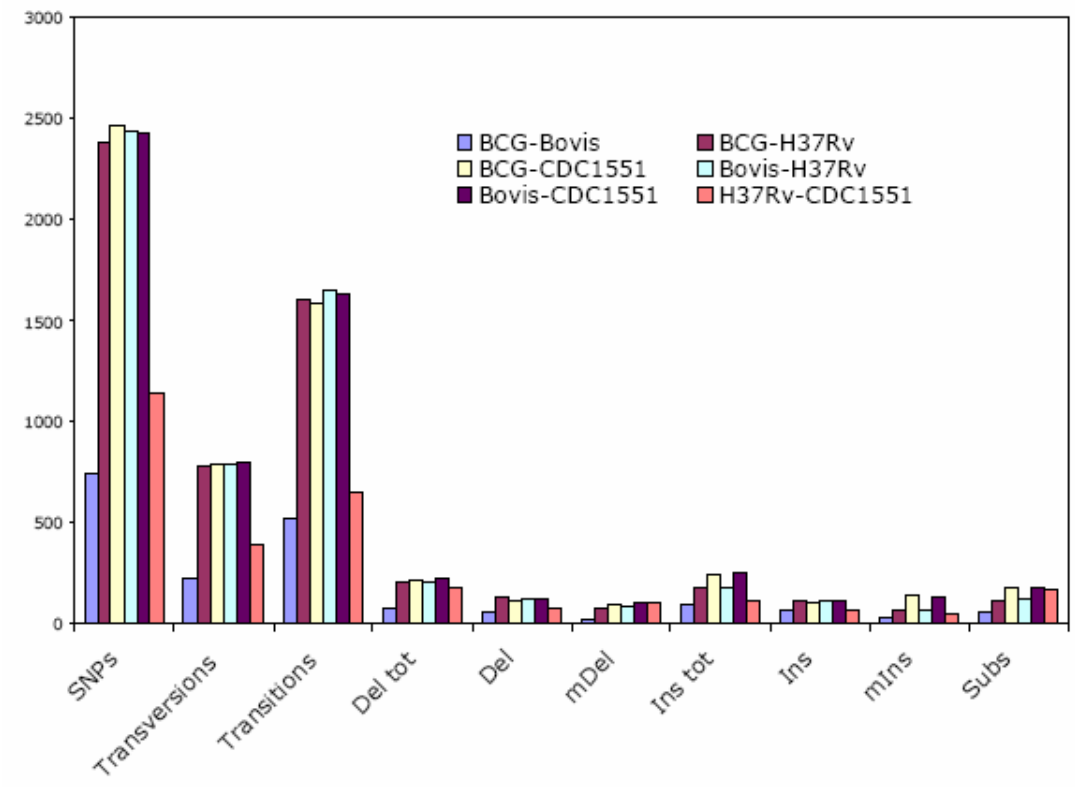


Fig. 13 - Comparative genomics showing SNPs across *M. tuberculosis, M. bovis* and *M. bovis* BCG.

Cluster analysis of various *M. tuberculosis* strains using 212 SNP markers identified six deeply branching, phylogenetically distinct SNP cluster groups (SCGs) and five subgroups. The SCGs were strongly associated with the geographical origin of the *M. tuberculosis* samples and the birthplace of the human hosts. The most ancestral cluster (SCG-1) predominated in patients from the Indian subcontinent,

while SCG-1 and another ancestral cluster (SCG-2) predominated in patients from East Asia, suggesting that *M. tuberculosis* first arose in the Indian subcontinent and spread worldwide through East Asia. Restricted SCG diversity and the prevalence of less ancestral SCGs in indigenous populations in Uganda and Mexico suggested a more recent introduction of *M. tuberculosis* into these regions (Filliol *et al.,* 2006).

1.7.3.2. Large Sequence Polymorphisms (LSPs)

Large sequence polymorphisms (LSPs) are now recognised as evolutionary markers for the *M. tuberculosis* complex (Brosch *et al*., 2002; Mostowy *et al.,* 2005). The deletions fall into two groups: "ancient" and "recent". The ancient deletions are widespread and occurred at different stages in the speciation process, while the recent deletions are rather restricted in distribution.

Comparative genomics has identified at least 18 variable regions (Table. 6) ranging from 0.3 kb to 12.7 kb, which are present in *M. tuberculosis* and not in BCG. RD1 is the only region that is absent from all BCG strains but present in virulent *M. bovis* and *M. tuberculosis* strains (Brosch *et al.,* 2000; Gordon *et al.,* 1999; Mahairas *et al.,* 1996). This region contains 8 genes (Rv3871- Rv3879c), most of which belong to the *esat-6* cluster and genes encoding PE and PPE proteins (Brosch *et al.,* 2000). RD2, another variable region (Rv1978- Rv1988c) is a recent deletion restricted to BCG strains derived since 1927, and includes genes coding for a variety of functions including methyl transferases, permeases, ribonucleotide reductase, a regulatory protein and a secreted protein, namely MPT64. The RD3 locus is a prophage (phiRv1) and RD4 encodes enzymes involved in the biosynthesis of lipopolysaccharides and both are absent from *M. bovis* and *M. bovis* BCG strains. RD5 contains eight ORFs, three of them encode phospholipase C enzymes (*plcA*, *plcB*, *plcC*), and the remaining five encode proteins, belonging to the Esat-6 and PPE families respectively (Gordon *et al.,* 1999). The RD6 region varies with the *M. tuberculosis* complex members and essentially consists of PPE proteins and IS1532. The RD7 region contains one of the 4 *mce* operons that encode invasin-like proteins required for *M. tuberculosis* (Arruda *et al.,* 1993). The effect of the loss of *mce3* on virulence is not known, but it was suggested that the remaining three *mce* operons could balance for any lost activity (Gordon *et al.,* 1999). The

RD8 region contains genes belonging to the ESAT-6 family, PE and PPE families and an *ephA* gene that encodes epoxide hydrolase (Brosch *et al.,* 2000; Gordon *et al.,* 1999). The RD9 contains genes encoding for an export protein, oxidoreductase, and a pre-corrin methyltransferase that is involved in cobalamin biosynthesis (Gordon *et al.,* 1999). The RD10 encompasses the genes encoding for an enoyl CoA hydratase and an aldehyde dehydrogenase. RD 11, 12 and 13 were absent in both *M. bovis* and *M. bovis* BCG, while the RD 14, 15 and 16 were restricted to few members of BCG (Brosch *et al.,* 2002). These regions can be accessed for the potential to differentiate between *M. tuberculosis* and *M. bovis / M. bovis* BCG.

Table 6 - Region of difference and their significance in speciation of *M. tuberculosis* complex

Region deleted	Species	ORF / gene	Function
RD1	All strains of BCG	Rv3871-Rv3879c	PE, PPE, ESAT-6, Various conserved hypothetical
RD2	Some *M. bovis* BCG	Rv1978-Rv1988	Methyltransferases, permease, MPT64, ribonucleotide reductase, conserved hypothetical, membrane and secreted proteins, LysR transcriptional regulator
RD3	*M. africanum, M. microti,* Some *M. bovis,* all *M. bovis* BCG	Rv1573-Rv1876c	PhiRv1 prophage
RD4	Some *M. bovis,* all *M. bovis* BCG	Rv1505c-Rv1516c	Various membrane proteins and enzymes involved exopolysaccharide synthesis
RD5	*M. microti, M. bovis, M. bovis* BCG	Rv2346c-Rv2353c	ESAT-6, PE, PPE family members, phospholipases C
RD6	Variable in all	Rv3425-Rv3428c	PPE proteins, IS 1532
RD7	All except *M. tuberculosis* and *M. canetti*	Rv1964-Rv1977	Various exported and integral membrane proteins, MceP invasions
RD8	*M. microti, M. bovis, M. bovis*	*ephA*-	Epoxide hydrolase, mono-oxygenase, lipoprotein, ESAT-6,

	BCG	*lpqG*	PE, PPE protein family
RD9	All except *M. tuberculosis* and *M. canetti*	*CobL*-Rv2075	Precorrin methylase, oxidoreductase, exported protein
RD10	*M. microti, M. bovis, M. bovis* BCG	Rv0221-Rv0223c	Enoyl CoA hydratase, aldehyde dehydrogenase
RD11	*M. bovis, M. bovis* BCG	Rv2645-Rv2695c	PhiRv2 prophage
RD12	*M. bovis, M. bovis* BCG	*SseC*-Rv3121	Thiosulfate sulfur-transferase, molybdoprotein converting factor, methyltransferase, cytochrome P450.
RD13	*M. bovis, M. bovis* BCG	Rv1255c-Rv1257c	Transcriptional regulator, Cytochrome 450, Dehydrogenase
RD14	Some BCG	Rv1765c-Rv1773c	PE_PGRS, Conserved hypothetical, IclR transcriptional regulator
RD15	Some BCG	Rv0309-Rv0312	Conserved hypothetical
RD16	Some BCG	Rv3400-Rv3405c	Transcriptional regulator, conserved hypothetical, β−phosphoglucomutase
TbD1	Modern *M. tuberculosis*	*mmpS*6, *mmpL6*	Membrane proteins

Though the role of these deletions in strain differentiation is unclear, they can be applied to propose a new evolutionary scenario for the members of the *M. tuberculosis* complex (Brosch *et al.,* 2002). The authors analysed the distribution of the 20 variable regions in a total of 100 strains belonging to *M. tuberculosis*, *M. africanum*, *M. canetti*, *M. microtii*, *M. bovis* and *M. bovis* BCG. Their study showed that *M. bovis* had undergone several deletions compared to *M. tuberculosis* (Fig. 14) (Brosch *et al.,* 2001).

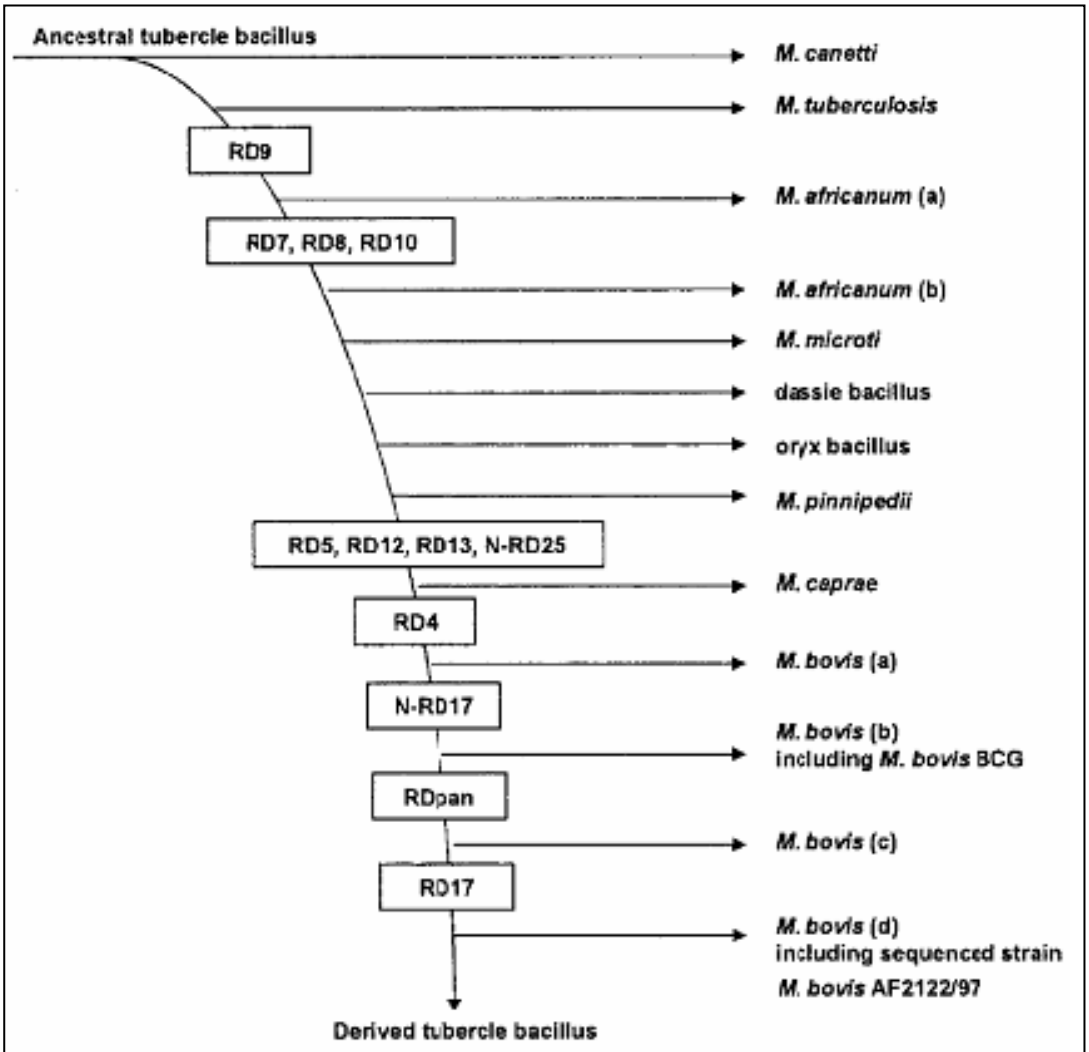


Fig. 14 - Deletion based phylogeny of *M. tuberculosis* complex (Brosch *et al., 2*002)

1.7.4. Functional genomics

Post-genomic research, though it encompasses diverse aspects of modern science includes the two broad areas of bio-informatics and functional genomics. Functional genomics unravels the roles of uncharacterised genes present in a particular genome and involves a number of levels of investigation that have been named transcriptome, proteome, and metabolome (Oliver, 2000). Transcriptomics, which defines the phenotypic state of an organism by its expressed genes (Kellam, 2001) is currently being done by DNA micro arrays for the determination of global gene expression. The gene expression is measured by isolating mRNA, synthesising fluorescently-labelled cDNA and hybridising it to the array.

The term “proteome”, first used in 1995, describes the protein complement of a genome (Wasinger *et al.,* 1995) and deals with variations in translation activity in a cell. The cornerstone of this technology is two-dimensional gel electrophoresis (2-DE), which separates proteins according to their isoelectric points in the first dimension and molecular mass in the second dimension. Due to the tremendous resolving power of the technique, it generates catalogue of spots enabling the creation of databases of all expressed proteins. A modification of this is the 2D-HPLC / MS proteins, in which protein fractions are digested with trypsin, that are separated by HPLC followed by subjecting the separated fractions by reverse phase HPLC coupled with the mass spectrometer. As the expression of protein is directionally proportional to the intensity of peptides, this technique was useful for the overall analysis of proteins expressed at any particular time during growth of an organism.

Metabolomics deals with the entire complement of metabolites in a tissue, organ or organism (Oliver, 2000). As the metabolome of an organism changes according to its physiological, developmental or pathological state, this technique can reveal phenotypes of proteins active in metabolic regulation (Thomas, 2001).

Additionally, gene disruption experiments or transposon mapping can be applied to determine the function. Mutagenesis is a powerful tool to determine gene function by establishing gene requirements under different conditions. This also helps in identifying the potential candidates for the development of vaccines and drugs. But in the *M. tuberculosis* complex, the generation of mutants by homologous recombination has been difficult in practice. Studies based on the mycobacterial genetics had considerably improved our knowledge about the pathogenic process.

1.7.5. Mycobacterial Genetics

Early studies to create mutations in mycobacteria were done with *M. smegmatis.* But genetic manipulation of slow-growing mycobacterial species proved to be difficult due to the slow growth rate, low transformation efficiency and lower rates of homologous recombination with high levels of illegitimate recombination, preventing the inactivation of target gene. An added problem is the relatively

impermeable cell wall, as it acts as a diffusion barrier for the antibiotics resulting in poor selection along with the occurrence of high levels of spontaneous resistant colonies (Morsczeck, 2003). Kanamycin, hygromycin and streptomycin are the most commonly used antibiotics along with gentamycin. However, kanamycin resistance is not favoured in slow growing members as high rate of spontaneous mutation was reported due to the presence of a single copy of rRNA cistron in *M. tuberculosis* while it is two in *M. smegmatis*. But this problem was avoided by complementing the plasmid with a second resistance marker such as the streptomycin resistance cassette. Hygromycin, despite its high cost and lower stability, is the preferred antibiotic selection marker because of the frequency of spontaneous mutation is very low (Smith, 2003).

The identification of essential genes, using allelic exchange or transposon mutagenesis techniques has profound application in establishing the biological significance of the mutated gene(s) as well as the potential drug targets (Parish *et al.,* 2001). Other cloning strategies include directed gene inactivation, global gene inactivation, complementation and anti-sense methods. The numbers of reports with successful manipulation of mycobacterial genes are increasing. A better understanding of the host pathogen interrelation will help with the development of better diagnostic assays and vaccines with increased efficacy.

Objectives of the study

The objectives of this study include the understanding the role of two important basic nutrients, iron and carbon, mainly their acquisition and their probable role in virulence.

1. Iron acquisition machinery in *M. kansasii*.
    (a) To demonstrate the co-ordinated regulation of mycobactin, carboxymycobactin and IREPs by iron.
    (b) Characterisation of mycobactin and carboxymycobactin.
    (c) Identification of IREPs.

2. Carbon metabolism with special reference to glycerol utilisation and study the effect, if any on iron metabolism (This study, using mycobacterial genetics was done in Veterinary Laboratories Agency, Weybridge, UK as a Commonwealth Spilt Site Doctoral Fellow).
    (a) Generation of knock-out and knock-in mutants in *M. tuberculosis* and *M. bovis* with respect to *pykA*, *ald* and both.
    (b) Study the morphological difference of the wild type and mutants using pyruvate and glycerol as carbon source.
    (c) Proteome analysis of mutants to see the global effect of the mutation in carbohydrate metabolism.

CHAPTER 2

MATERIALS AND METHODS

2.1. Sources of chemicals

Culture media Middlebrook 7H9, 7H10 and 7H11 with supplements OADC / ADC were purchased from Becton and Dickinson. LJ was purchased from Himedia. Iron-free chemicals were purchased from Puratronic, Alfa AESAR. Other chemicals were purchased from Sigma, Sisco research laboratories (SRL), Qualigens, Fisons and Merck. All organic solvents including HPLC grade solvents and acids were obtained from Qualigens. The 0.22 µm filters for HPLC samples were obtained from Sartorius. The nitrocellulose and PVDF membranes for western blot analysis were obtained from Amersham Pharmacia. All molecular biology reagents and kits were obtained from Qiagen, New England Biolabs, Promega and Sigma. Primers were obtained from MWG, Germany.

2.2. Studies on the iron acquisition machinery in *Mycobacterium kansasii*

2.2.1. Growth under high and low concentrations of iron

2.2.1.1. Bacterial strains

Growth under iron-regulated conditions was initially established for several mycobacterial species including *M. smegmatis, M. vaccae, M. kansasii* and *M. fortuitum* (National Mycobacterial Repository at JALMA Institute, Agra, India), *M. neoaurum* (ATCC 23071).

2.2.1.2. Establishment of growth under high and low iron conditions

All glassware was made iron free and glass double-distilled water was used for all the experiments. Proskauer and Beck medium was prepared by weighing 5 gm asparagine, 5 gm potassium dihydrogen orthophosphate and 20 mL glycerol in 800 mL of glass double distilled water. The pH was adjusted to 6.8 with 10 N sodium hydroxide and the medium was made up to 1L. Five gm of alumina was added and autoclaved at 121°C at 15 lbs / square inch for 15 minutes. Upon cooling, the media was filtered and aliquoted as 100 mL volumes in 250 mL conical flasks and then re-autoclaved (Hall *et al*., 1987).

To 100 mL of Proskauer and Beck medium, 1 mL of salt solution containing (16.8 mM Mg^{2+}, 5.01 μM Mn^{2+} and 17 μM Zn^{2+}) was added and iron added at a final concentration of 0.02 μg Fe / ml (0.36 μM of Fe) for low iron growth and 8 μg Fe / ml for high iron growth (144 μM of Fe). The fast growing strains were harvested after 7 days and slow growing strains were grown for 14 days. Growth was measured both by absorbance at 600 nm and determination of the cell dry weight.

2.2.1.3. Assay of mycobactin and carboxymycobactin

Mycobactin was extracted as follows. To this cell pellet from 15 mL culture, 750 μL of EtOH was added, left at RT overnight, followed by addition of 100 μL saturated ferric chloride (in ethanol) to the cell-free supernatant. The ferric mycobactin was extracted into chloroform, washed thrice with water and the concentration was determined by reading at 450 nm using ethanol as blank ($E\%_{450nm} = 43$).

The carboxymycobactin was assayed either by (a) Universal CAS assay or (b) reading the absorbance of the ferric form at 450 nm.

a) Universal CAS assay (Schwyn & Neilands, 1987)

Preparation of reagents

Solution A

6 mL of 10 mM HDTMA [Hexadecyltrimethyl ammonium bromide (3.645 g / L)] was pipetted in a standard flask and diluted with 10 mL of distilled water. To this, 1.5 mL of a freshly prepared of 1 mM ferric chloride (0.1622 g / L) in 10 mM HCl and 7.5 mL 2 mM Chrome azurol sulphonate (CAS) (1.211 g / L) were added slowly.

Solution B

4.3 g of anhydrous piperazine was dissolved in a minimum volume of water and 6.25 mL of conc. HCl was added. All the above solutions were made in iron free glassware.

Preparation of CAS assay solution

To prepare working CAS assay solution, solution B was added slowly to solution A and made up to 100 mL with glass double distilled water.

Procedure

Equal amount of culture filtrate and CAS assay solution were mixed and incubated for 30 min at RT. The absorbance was read at 630 nm and the values were expressed as siderophore units per 100 mg dry weight of cells. The un-inoculated medium was used as blank.

b) Absorbance at 450 nm (Ratledge & Ewing, 1996)

To the spent growth medium, saturated aqueous ferric chloride was added drop wise, till it just formed a precipitate, centrifuged at 10,000 rpm for 15 min. To the clear supernatant chloroform was added to extract the ferri-carboxymycobactin and read the absorbance of the clear supernatant at 450 nm ($E\%_{450nm} = 48$).

2.2.2. Structural elucidation of mycobactin and carboxymycobactin from *M. kansasii*

2.2.2.1. Extraction of mycobactin and carboxymycobactin

Mycobactin and carboxymycobactin were extracted essentially as described above from 100 ml cultures. The chloroform extracts were taken in a separating funnel, washed thrice with water to remove the excess ferric chloride and the excess water was removed by passing through anhydrous sodium sulphate. The sample was then rotavaporised and the residue was dissolved in minimal quantity of ethanol and stored for further use.

2.2.2.2. Analysis of carboxymycobactin by IR and NMR (Barclay *et al.*, 1986)

The NEAT infrared spectra were recorded on SHIMADZU FT-IR spectrophotometer model 8300 with polystyrene as reference. ^{1}H NMR 400MHz spectra were recorded on Bruker Avance-400 spectrometer with 2 mg of carboxymycobactin, using chloroform-d as a solvent and TMS as reference ($\delta = 0$ ppm) at 303 K. The chemical shifts are expressed in δ downfield from the signal of internal TMS.

2.2.2.3. Separation by HPLC (Barclay *et al.*, 1986; Lane *et al.*, 1998)

Both mycobactin and carboxymycobactin was separated by reverse phase HPLC on Water's C_{18} column (150 mm x 4.6 mm id) at a flow rate of 1 mL / min for 35 min and monitored continuously at 450 nm. The hydrophobic mycobactin was separated using a gradient (Table 7) of 0 to 100 % with buffer A (0.09% formic acid / 90% acetonitrile) and buffer B (0.1% formic acid in 60% acetonitrile) and 40% MeOH and the carboxymycobactins were separated using a gradient of 0 to 100% with buffer A (0.1% formic acid) and buffer B (0.09% formic acid in 90% acetonitrile). Fractions with maximal absorbance at 450 nm were collected, concentrated by Speed vac system and then subjected to mass spectrometry.

Table 7 - HPLC gradient for purification of mycobactin and carboxymycobactin

Time (min)	Buffer A %	Buffer B %
2	100	0
32	0	100
33	0	100
35	100	0

2.2.2.4. LC-MS / MS analysis of mycobactins and carboxymycobactins (Gobin *et al.*, 1995, Wong *et al.*, 1996)

Mycobactin and carboxymycobactins were subjected to electro spray ionization mass spectrometry (ESI-MS). ESI-MS was conducted in the positive ion mode using LCQ ion trap (Thermo Hemel, Hampstead, Herts, UK) mass spectrometer. The data was acquired in the voltage scan mode range of 400-900 amu at 40 s per decay. Conditions for HPLC were similar to that described for the analytical separation of mycobactins and carboxymycobactins except the sample volume was 50 µL. This was also done with HPLC purified fractions and the data was acquired in the voltage scan mode range of 200-800 amu at 35 s per decay. Further fragmentation was achieved by selective fragmentation of molecular ion peaks.

2.3. Isolation of Iron-regulated envelope proteins (IREPs) from *M. kansasii*: analysis of cell wall and membrane proteins by SDS-PAGE (Hall *et al.,* 1987)

2.3.1. Isolation of cell wall and cell membrane fractions by differential centrifugation.

The cells grown in high and low iron condition were harvested, washed and resuspended in Tris-HCl (0.01 M, pH 7.8) and subjected to sonication (total time 20 min at 30 s pulse with 30 s interval, amplitude 40 Hz, at probe temperature of 4^{o}C in Sonics Vibra cell sonicator). The sonicate was centrifuged at 6000 rpm at 4^{o}C for 15 min to remove cell debris and the supernatant was centrifuged at 18,000 rpm for 30 min at 4^{o}C to pellet the cell wall; the supernatant was then subjected to ultracentrifugation at 45,000 rpm for 2 h 30 min to pellet the cell membrane.

2.3.2. Solubilisation of cell wall and membrane using Triton X-114 (Sinha *et al.,* 2000)

The cell wall and cell membrane pellets were solubilised by adding an appropriate volume of 5% pre-condensed Triton X-114 in Tris-HCl (0.01 M, pH 7.8) and incubating at 4^{o}C overnight with rocking. The two phases were separated by warming to 37^{o}C in a water bath for an hour and then centrifuging at 3000 rpm for 15 min at RT to separate the detergent and aqueous phases. After suitable washes, the proteins were precipitated by adding 5 volumes of acetone and left overnight at -20^{o}C. The protein pellets, obtained by centrifuging at 12,000 rpm for 10 min were dissolved in 4% SDS in Tris-HCl (0.05 M, pH 6.8) and then subjected to SDS-PAGE.

The protein concentration in these samples was estimated either by the modified Lowry method (Markwell *et al.,* 1978) or the bicinchonic acid (BCA) Kit (Sigma).

2.3.3. SDS-PAGE analysis of cell wall and membrane proteins (Laemmli, 1970)

Solutions:

1. Acrylamide and N, N'- bisacrylamide (30:0.8);
2. Resolving gel buffer: Tris-Hcl (1.5M, pH 8.8) with 0.4% SDS;

3. Stacking gel buffer: Tris-Hcl (0.5 M, pH 6.8) with 0.4% SDS;
4. Sample buffer: 0.125 M Tris-HCl (pH 6.8) containing 4% SDS, 20% glycerol and 0.002% bromophenol blue;
5. Running buffer: Tris-glycine buffer (molarity, pH 8.8) with 0.5 %SDS;
6. Staining solution: (0.25% Coomassie brilliant blue in 50% MeOH and 10% acetic acid);
7. Destaining solution: 10% MeOH and 10% acetic acid

The resolving gel was prepared as 5-20% gradient and overlaid with 5% stacking gel. The protein samples were prepared by mixing equal amount of protein with sample buffer, boiled for 10 min, centrifuged at 10,000 rpm for 15 min, to remove any insoluble material and loaded on the gel. Electrophoresis was carried out initially at 100 V till the bromophenol dye entered the resolving gel and then at 125 V. The tracking dye was allowed to run out of the gel, followed by electrophoresis for an additional 30 minutes.

Gels were stained with Coomassie blue for 2 h, followed by removal of the background stain by washing with destaining solution.

2.4. Identification and characterization of a cell membrane associated iron-regulated envelope protein by HPLC

Due to the difficulty in the solubilisation of cell wall pellet, only the cell membrane proteins were analysed by HPLC. The membrane pellet was dissolved overnight at 37^{o}C in 200 µl of acetonitrile, diluted with water to get proteins in 70% acetonitrile, filtered through 0.22 µm filter and then subjected to HPLC (as done for the siderophores) using buffer A (0.1% TFA in 90% acetonitrile) and buffer B (0.1% TFA in 10% acetonitrile) at a flow rate of 0.5 mL / min for 60 min. The eluate was monitored continuously at 280 nm. After a preliminary run of the proteins from high and low iron cells, the peak fraction at 40 min (referred as MK40), seen only in low iron cells was collected, concentrated by acetone precipitation and used for ligand binding studies and mass spectrometry.

2.4.1. Ligand binding studies using spectrofluorimetry

The MK40' was subjected to excitation at 280 nm and the emission spectrum was recorded from 310 nm to 400 nm. Then, carboxymycobactin was added at concentration ranging from an initial concentration of 2.5 μM to a final concentration of 30 μM and the emission spectrum recorded after each addition. This was done till the saturation point was achieved. The data obtained was fitted to Wolf's equation to calculate the binding constant. The intrinsic binding constant (K_b) was determined from the plot of [ligand] /I-I_o *vs.* [ligand] where [ligand] is the concentration of carboxymycobactin and I corresponds to fluorescence intensity at 327 nm at a particular concentration and I_o corresponds to fluorescence intensity at 327 nm in free form. The data were fitted to (1)

$$[ligand] /I-I_o = [ligand] /I-I_o + 1/ K_b (I-I_o) \quad (1)$$

Each set of data, when fitted to the above equation, gave a straight line with a slope of 1/ (I-I_o) and a y-intercept of 1/ K_b (I-I_o). K_b was determined from the ratio of the slope to intercept. An in-house non-linear least square analysis program or the MicroCal Origin software package was used for curve fitting the data (Arounaguiri *et al.,* 2000).

The value of binding constant was obtained from a plot of concentration of ligand *vs.* Ligand concentration/ I-I_o where I_o is the initial fluorescence intensity and I is the fluorescence intensity at the particular concentration. The intercept/slope gives the binding constant value.

2.5. Analysis of protein bands separated by SDS-PAGE by mass spectrometry (Coldham & Woodward, 2005)

All the steps were carried out with gloves to prevent contamination with keratin. The gel was cleaned with ethanol and bands of interest were cut out using a scalpel blade. The cut out gel strip was divided into 4 equal pieces and placed in a 96 well plate. The gel pieces were washed twice with 50 μl of 50% MeOH / 5% acetic acid at RT with shaking. Then the gel pieces were rinsed with 50 μl acetonitrile for 5

min. The gel pieces become opaque, and the excess of acetonitrile was dried completely by freeze-drying. The gel pieces can be stored at -20^{o}C after the wash prior to tryptic digestion. The gel pieces were digested with 20 µl of 100 ng / µl sequencing grade porcine trypsin dissolved in 50 mM ammonium bicarbonate pH 8.0 on ice. The gel pieces were allowed to swell for 30 min on ice. After the removal of excess of trypsin, the gel pieces were covered with sufficient amount of ammonium bicarbonate and incubated overnight at 37^{o}C. The following day the excess of liquid around the gels were transferred to siliconised insert tube and the gels were rinsed with 20 µl ammonium bicarbonate for 10min with shaking. The gel pieces were extracted twice with 20 µl of 0.1% TFA : acetonitrile (50:50) extract by vortexing for 10 min on a plate shaker. The gel extracts were pooled together, and the volume was reduced to 10 µl using Gyrovap at 60^{o}C for 2 h at 0.005 bar. The gel extract was dissolved in 10-20 µl of 0.1% formic acid and analysed by LC-MS (Kinter and Sherman; Promega technical bulletin No. 512).

2.5.1. Mass spectrometry analysis

Tryptic-digested peptides in 0.1% (v/v) formic acid were analysed by capillary-HPLC-MS using a Famos auto sampler, a surveyor HPLC pump equipped with an Accurate flow splitter (100 to 1) and LCQ ion-trap mass spectrometer (ThermoFinnigan, Hemel Hamstead, UK). Tryptic peptides (5 µl) were chromatographed on a Picofrit column and peptides eluted at an estimated flow rate of 400 nL / min with a linear binary gradient of 0.1% formic acid (A) and acetonitrile (B) as explained in Table 8. The Picofrit column (70 mm) was mounted in a steel filter (0.5 µM pore size) assembly to which the electrospray ionisation potential (2.5 KV)

Table 8 - HPLC gradient for analysis of tryptic digested peptides

Time (min)	Buffer A %	Buffer B %
0	100	0
40	60	40
50	40	60
52	100	0
75	100	0

was applied. The Picofrit tip was positioned slightly off axis within 3 mm of the LCQ mass spectrometer heated capillary inlet using a Proton nanospray source. Mass data for each fraction was collected over a mass of 300-2000 using a Big Three acquisition method with data-dependent product ion scanning of 1^{st}, 2^{nd} and 3^{rd} most abundant ions above a threshold trigger of 3×10^5 counts / sec. The mass isolation window and collision energy were set to 4 amu and 35% respectively. The mass spectrometer was tuned to the doubly charged ion (m/z 820.5) derived from synthetic peptide FNPGELLPEAAGPTQV. Sensitivity of the mass spectrometer was assessed by monitoring the retention time and intensity of ions derived from a low level standard prepared from the synthetic peptides MRFA (50 ng / mL; $[M+H]^+$ m/z 524.2); APYELNITSATYQSAI (10 ng/ mL; $[M+H]^{2+}$ m/z 871.5); FNPGELLPEAAGPTQV (10 ng/ mL; $[M+H]^{2+}$ m/z 820.5) and AIQGNVVSTSIHSLLDEG (10 ng / mL; $[M+H]^{2+}$ m/z 827.9).

2.5.2. Bioinformatic data analysis

The SEQUEST algorithm embedded within the Bioworks software package was used for the identification of proteins from tryptic peptide mass spectra. Proteins were identified by comparison of tryptic peptide product ion mass spectra against those generated from *M. tuberculosis* database derived from Tuberculist. The multi-consensus report function was used to assign tryptic peptides to individual proteins and compile rank listings of the proteomes. Identified proteins were ranked in ascending order according to consensus scores and false positives were minimised by filtration against 4 of the 5 following criteria X corr >2.0. DeltCn >0.2, Sp >400, rsp <5, ions >30%.

2.5.3. Assessment of assay performance

The performance of the SCX and reversed phase chromatographic separation and sensitivity of the mass spectrometer was monitored by analysis of specific standards at the beginning of each batch. The SCX procedure was assessed by analysis of the blanks to control for carryover between samples and a standard preparation of

tryptic peptides derived from BSA (25 μg / mL) to ensure reproducible chromatography and sensitivity of the mass spectrometer was assessed by monitoring the retention time and intensity of ions derived from a low level standard prepared from the synthetic peptides (described above Section 2. 2. 6).

2.6. Functional genomics: studies on intermediary metabolism in *M. tuberculosis, M. bovis* and *M. bovis* BCG

2.6.1. Bacterial strains and growth conditions

The mycobacterial strains used in this study include *M. tuberculosis* H37Rv, *M. bovis* Type -35 (VLA culture collection), *M. bovis* BCG (Statens Serum Institute, Denmark). Long term storage of the mycobacterial stocks in 20% glycerol was done at -80^{o}C.

Table 9 shows the different media used for generating mycobacterial mutants. Both *M. tuberculosis* H37Rv and *M. bovis* BCG Pasteur strains were grown in Middlebrook Albumin-Dextrose-Catalase-Tween (MADC-TW), while *M. bovis* was grown in Middlebrook Albumin-Dextrose-Catalase-Tween-Pyruvate (MADC-TWP) medium. Following recombination experiments, these organisms were plated on 7H10 agar (with 0.2% glycerol for *M. tuberculosis* H37Rv and BCG Pasteur mutants or 0.4% sodium pyruvate for *M. bovis* mutants) containing 25 μg / mL kanamycin or 200 μg / mL hygromycin or both. The *E. coli* strains used for cloning include *E. coli* DH5α and *E. coli* DH10β (GIBCO/BRL). *Escherichia coli* strains DH5α and DHB10β were grown at 37°C in LB broth with agitation / LB + G agar supplemented with 25 μg / mL kanamycin or 200 μg / mL hygromycin or both.

Table 9 - Preparation of culture media

Medium	Recipe (per litre)
MADC-TW[a]	4.7 g Middlebrook 7H9, 100 mL Middlebrook Albumin-Dextrose-Catalase enrichment (ADC: 5 g bovine serum albumin Fraction V, 2 g dextrose, 0.003 g catalase in 100 mL water), 2 mL Glycerol, 0.5 mL Tween 80.
MADC-TWP[a]	4.7 g Middlebrook 7H9, 100 mL Middlebrook ADC enrichment, 4.16 g sodium pyruvate[c], 0.5 mL Tween 80.
7H10 agar medium[a]	19 g Middlebrook 7H10, 100 mL Middlebrook Oleic acid-Albumin-Dextrose enrichment (OADC: 0.05 g oleic acid, 5 g bovine serum albumin Fraction V, 2 g dextrose, 0.004 g catalase, 0.85 g sodium chloride in 100 mL water).
7H10 agar for mutants	19 g Middlebrook 7H10, 100 mL Middlebrook OADC enrichment, 2 mL Glycerol (for *M. tuberculosis* H37Rv and *M. bovis* BCG Pasteur mutants) or 4.16 g sodium pyruvate (for *M. bovis* mutants), with 25 µg / mL kanamycin or 200µg / mL hygromycin or both.
7H11 agar for different carbon sources	21 g Middlebrook 7H10, 100 mL Middlebrook OADC enrichment, either with no other carbon sources or 2 mL Glycerol or 4.16 g sodium pyruvate, with 25 µg / mL kanamycin or 200 µg / mL hygromycin or both.
Modified Sauton's medium[d]	4 g asparagine, 2 g citric acid, 0.5 g of K_2HPO_4, 0.5 g of $MgSO_4.7H_2O$ and 0.05 g of ferric ammonium citrate, 4.16 g of sodium pyruvate, 0.025% tylaxapol, pH was adjusted to 7.2 with KOH.
Sauton's medium for growth in different carbon sources[d]	4 g asparagine, 2 g citric acid, 0.5 g of K_2HPO_4, 0.5 g of $MgSO_4.7H_2O$ and 0.05 g of ferric ammonium citrate, the pH was adjusted to 7.2 with KOH. Carbon sources were added to give a final concentration of 40mM of sodium pyruvate, 40mM of Glucose, 0.05% (v / v) for Tween 80 (+0.5% BSA) and 6% glycerol or omitted for negative control.
LB (Luria-Bertani)[b]	10 g bacto-tryptone, 5 g bacto-yeast extract, 5 g NaCl, 15 g bacto agar, 2 g glucose.
S.O.C.	20 g bacto-tryptone, 5 g bacto-yeast extract, 10 mL 1M NaCl, 10 mL 2M $MgSO_4.7H_2O$, 10 mL 2M glucose.

[a]Media was sterilised by autoclaving at 121°C for 15 minutes. ADC or OADC enrichment or antibiotics (where appropriate) were added to the sterile media after it had cooled to 50-55 °C.
[b]Excluded the addition of bacto-agar when preparing LB broth.
[c] Sterilised by using 0.2 µm filters.
[d] Carbon sources and ferric ammonium citrate were not added. Media was sterilised by autoclaving at 121°C for 15 minutes. Filter sterilized ferric ammonium citrate and carbon sources were added.

2.6.2. Isolation of genomic DNA (van Soolingen *et al.*, 1994)

A single colony was inoculated in 10 mL of MADC-TW or MADC-TWP and incubated for 10 days at 37°C. 2 mL of this culture was centrifuged at 12,000 x g for 10 min at 4°C, washed three times in 2 mL of TE buffer (100 mM Tris / HCl, pH 8.0 and 100 mM EDTA), resuspended in 600 µl of TE buffer and heat killed at 80°C for 30 min. All the following steps were performed in a thermomixer (Eppendorf, UK) with continuous shaking. 50 µl of lysozyme (10 mg / mL) was added, vortexed and incubated for 90 min at 37°C. Then 35 µl of 20% SDS and 10 µl of proteinase K (10 mg / mL) were added and incubated for a further 30 min at 65°C. Then 100 µl of 5 M NaCl and 80 µl of CTAB solution (4.1 g NaCl and 10 g N-cetyl-N, N, N-trimethyl ammonium bromide in 100 mL water, prewarmed to 65°C) were added and the mixture was vortexed until the solution became milky. The solution was incubated at 65°C for 30 min. This mixture was then cooled to RT and 750 µL of chloroform / isoamyl alcohol (24:1) was added, and mixed by gentle inversion of the microfuge tube. The aqueous layer was removed by centrifuging at 12,000 x g for 15 min at 4°C followed by the addition of an equal volume of isopropanol and incubating the mixture at -20°C for 30 min to precipitate the DNA. The DNA was pelleted by centrifugation at 12,000 x g for 15 min at 4°C, washed in 1 mL of cold 70% EtOH and resuspended in 50 µl distilled water. DNA concentration was estimated by measuring absorbance at 280 nm.

2.6.3. Agarose gel electrophoresis (Sambrook *et al.*, 1989)

TAE buffer (4.84 g Tris, 1.14 mL glacial acetic acid and 20 mL 0.5 M EDTA in a volume of 1 L, pH 8), agarose (between 0.8 to 1.5 % in TAE buffer), ethidium bromide (added to the gels to give a final concentration of 0.1 µg / mL), gel loading buffer (0.25 % bromophenol blue, 0.25 % xylene cyanol FF, 40 % (w/v) sucrose in water) were used. The DNA samples were prepared by adding gel loading buffer (4:1 - DNA: buffer) and loaded on the gel. The DNA size standards included 1 kb ladder (0.8 µg) or 1 kb plus ladder (0.7 µg). Samples were subjected to electrophoresis

between 40 - 100 V and the DNA was visualised in a UV transilluminator. Quantification of DNA products was determined by comparing the intensity of ethidium bromide incorporation into sample fragments with incorporation into 0.5 μg phage λ DNA

2.6.4. Purification of DNA from agarose gels

DNA fragments to be recovered for further manipulation were run on 1% agarose gels. The required fragment then was excised from the gel and purified using the Sephaglass B Prep Kit (Pharmacia Biotech) as per manufacturer's instructions**.** Briefly, the excised gel was cut into smaller pieces, weighed added 1 mL of gel solubiliser per mg of agarose and incubated at 60°C for 10-15 min until the agarose dissolved. For 1 μg of DNA, 30 μl of Sephaglas BP was added to the dissolved slice, followed by incubation at RT for 10 min. The solution was vortexed gently every min and the Sephaglas / DNA mix was pelleted by centrifugation at 12,000 g for 30 s and the supernatant was discarded. The mix was re-centrifuged to remove any residual liquid and the Sephaglas pellet was washed three times in wash buffer. The amount of wash buffer used was dependent on the volume of the Sephaglas used (16 x the volume of Sephaglas added). The Sephaglas pellet was allowed to air dry for 10 min and re-suspended in 20-40 μl of nuclease free distilled water. The Sephaglas / water mixture was incubated at RT for 5 min with periodic agitation and then spun down at 13,000 rpm for 1 min. The eluted DNA was stored at 4°C until needed.

DNA was concentrated by precipitation with alcohol (Sambrook *et al.,* 1989). To the DNA sample, 1/10 volume of 3 M sodium acetate (pH 5.5) and 2 volumes of absolute EtOH were added and incubated at –20°C for 30 min. The DNA was then pelleted by centrifugation in a microfuge at 12,000 x g for 15 min. The pellet was washed with 70% (v / v) EtOH to remove excess salt and re-suspended in an appropriate volume of nuclease free distilled water.

2.6.5. Polymerase Chain Reaction (PCR)

PCR was performed in a final volume of 50 µl, using 200 ng of genomic DNA / 50 ng of plasmid DNA / 10 ng of PCR products as template DNA. The reaction mixture contained 50 mM KCl, 10 mM Tris-HCl (pH 7.5), 3-7 mM $MgCl_2$, 200 µM of each dNTP, 0.5 U Taq DNA polymerase and 10 pmol of each oligonucleotide primer. Two negative controls, one without template DNA and the other without Taq DNA polymerase were included. The primers, designed by Seqman software analysis are listed in Table 10.

Table 10 - Primers used in this study

Primer	Sequence (5'-3')
ALDFC	AAC TGC AGC ATG CGC GTC GGT ATT CCG ACC
ALDRC	AAC TGC AGT AAC GAG CGG CCG AGA GTC AGG
ALDSF	GTG CAG GAA AGA AGC TCT AC
ALDSF/2	AGG TTG GCG CTT ACC ACC TG
ALDSR	GCT TGC ATG CCT GCA GTA AC
ALDRFF	ATT TCC CGG GTT TCG TTG CGA CCG ACA AAC AC
ALDRFR	AAT TAC TAG TCA AGA TCT GCC CGT GTC GCA GG
ALDLFF	AAC TGC AGC CCA GGT TGG CGC TTA CCA C
ALDLFR	AAC TGC AGC CGG TAC CAA TGC CGC ACA C
ALDSRF	GGC GAT TAA GTT GGG TAA CG
ALDSRF/2	CGA GAT CAT CGC CTG CAA CG
ALDSRR	GTA CGG GCC GGC TTA ATT CC
ALDSLF	AAC TGG CGC AGT TCC TCT GG
ALDSLF/2	CGG CTG TTT CGA AGG CTC AC
ALDSLR	AAG CAG GCA AAT GGC TAA CC
PYKARFF	ATA ATC TAG AGG TCG TGC CGG ACG GTA AAC
PYKARFR	AAT ATC TAG AGG CCG ACG GCA GCT TCA CAG
PYKALFF	ATT TCC CGG GTT GAG CCA ACT CAG GAA ATG
PYKALFR	ATT AAC TAG TGC CCG GGT CGG CAC ACC GGTC
PYKASRF	AAC TGG CGC AGT TCC TCT GG
PYKASRF/2	GTG AGC TGT TGC GCG GTT AC
PYKASRR	AAG CAG GCA AAT GGC TAA CC
PYKASLF	TGA ATC TCT GAG CGG AGG TG
PYKASLF/2	CAT TGA CGC CCT GCT GAT ATT G

PYKASLR	GTA CGG GCC GGC TTA ATT CC
PYKAFBamHI	CTG AGG ATC CAT GAC GAG ACG CGG GAA AAT C
PYKARBamHI	GTC CGG ATC CAC TAG TCT AGA CGT CAT CTT CCC CGA
HYG F	CTT CAC CGA TCC GGA GGA ACT G
HYG R	TTG TGT CAC AGC GGA CCT CTA T
HYGOUT1	AGG GTA CGG GCC GGC TTA ATT C
HYGOUT3	CTT CAC CGA TCC GGA GAT TCC T
KANF	ATA TAC TAG TCG GGC GCG ATC CAG AAG AAC
KANR	AAT AGA TCT GCT TGC CGT CCC GTC AAG TC
PYKADIF	CCC AGT ACT GTG ACG AGA CGC GGG AAA ATC G
PYKADIR	CCG CTT AAT TAA CTA GAC GTC ATC TTC CCC GAT CCG
ALDDIF	AAC TGC AGA TGC GCG TCG GTA TTC CGA CCG
ALDDIR	AAC TGC AGT TAA TTA AAG TAC TCC TCC TTC AGG CCA GCA CGC TGG CGG G
ALDDISF	AAT TGA CAG CAC ACC GCC GTC CAG G
ALDDISF2	CGC GAC CGT TAC GGT TCT AGA C
PYKADISF	AAT GGA CGT CGC CCG AAT GAA C
PYKADISF2	AGA GGT CCC GCT GGT ACA GAA G
PYKADISR	GCT GCA AGG CGA TTA AGT TGG G
KANF6	CTC CTG ATG ATG CAT GGT TAC TCA C
KANR7	ACT CAC CGA GGC AGT TCC ATA GGA T

PCR was done in Biometra thermocycler using heated lid method to avoid loss of sample due to evaporation using reaction parameters, as indicated in the specific results. A final step of 68 to 72°C up to 10 min was carried out to ensure the extension of all products. The products were purified using Mini-Elute PCR purification clean-up kit according to manufacturer's recommendations.

2.6.6. Cloning experiments

a) Restriction digestion of vector and inserts

1 µg of plasmid DNA or PCR products were digested with suitable restriction enzymes in final reaction volumes ranging between 20-50 µl. Digestion of plasmid DNA was carried out for 2-4 h whereas the PCR products were digested for 4 h to overnight incubation based upon the half life of the enzymes and the length of PCR

product. The reactions were stopped either by heat inactivation or by purifying the products with mini-elute reaction clean up kit (QIAGEN).

b) Preparation of the vector

400 ng of digested plasmid was treated with 0.5 µl of shrimp alkaline phosphatase (SAP) in Tris-HCl buffer (20 mM, pH 8.0) with 10 mM $MgCl_2$ for 1 h at 37°C. The dephosphorylated vector was purified through mini elute reaction clean up kit.

c) Ligation of vector and insert DNA

Digested vector and insert DNA were quantified by their absorbance at 280 nm using Biometra quant. Concentration of the restriction digested vector DNA was kept constant as 100 ng and liagation was done using vector and insert DNA in molar ratio of 3:1. The concentration of insert DNA was calculated based on the following formulae:

$$\text{Concentration of the insert (ng)} = \frac{\text{ng of vector x size of insert (kb)}}{\text{Size of vector (kb)}} \text{ x Molar ratio}$$

The reaction mixture in a volume of 20 µl. contained 30 mM Tris-HCl, 10 mM $MgCl_2$, 10 mM dithiothreitiol (DTT), 1 mM ATP and T4 DNA ligase. The ligation was carried out at 16^{o}C overnight. The DNA was precipitated and the washed pellet was dissolved in 5 µl of nuclease-free water at 37^{o}C for 10 min.

2.6.7. Transformation experiments

i) Transformation of *E. coli*

The recombinant plasmid was used for transformation of *E. coli* DH5α by heat shock method or *E. coli* DH10β by electroporation.

a) Heat shock method

Frozen *E. coli* DH5α competent cells (Gibco-BRL) were thawed on ice prior to use. 5 µl of ligation mix or 50 ng of purified plasmid was incubated with 50 µl of competent cells on ice for 30 min in Falcon 2059 15 mL tubes. The cells were then

subjected to heat shock in a 42°C water bath for 45 s and transferred to ice for 2 min. To recover the cells, 450 µl of SOC medium was added to the mixture and cells were incubated at 37°C for 1 hour in a shaking incubator at 225 rpm (Sambrook *et al.,* 1989). Transformants were plated out on LB+G plates containing either kanamycin (25 µg / mL) or hygromycin (200 µg / mL) depending on the requirement.

b) Electroporation

The required number of 0.2 cm Biorad micro-electroporation chambers and Falcon 2059 15 mL tubes were placed on ice prior to electroporation. *E. coli* DH10β cells (Gibco-BRL, Max Efficiency) were thawed on ice. 2 µl of ligation reaction or plasmid DNA were placed in chilled electroporation cuvettes, followed by addition of 40 µl of competent cells. DNA was electroporated into competent cells using the Gene Pulser apparatus (Bio-Rad) at 1.8 kV, 25 µF and 200 Ω. Electroporated cells were resuspended in 450 µl SOC medium and allowed to recover for 1 hour at 37°C. Transformants were plated on LB+G medium containing the appropriate antibiotic.

c) Electroduction

Electroduction was used to analyse the presence of vector in the complemented mutants. The complemented mutant was grown in 10 mL basal medium supplemented with appropriate antibiotic for 7 days. Electrocompetent cells were prepared as described above from 1 mL of culture. 20 µl of this suspension was electroporated into *E. coli* as above. Plasmid DNA was isolated and then digested with appropriate RE to confirm the size of the gene insert of the recombinant plasmid. Finally, plasmid DNA was sequenced using the primers to confirm the presence of the cloned insert.

ii) Transformation of mycobacteria

Mycobacteria were grown in 100 mL of MADC-TW medium to an OD_{600} of approximately 0.8. The harvested cells were washed twice with 50 mL of 10% glycerol, followed by two washes with 25 mL of 10% glycerol at RT and then re-suspended in a final volume of 1.5 mL in 10% glycerol. To 400 µl of this suspension

of competent cells taken in 0.2 cm electroporation cuvettes, 1 µg of plasmid DNA was added and then electroporated at 2.5 kV, 25 µF, 2000 Ω (Bio-Rad Gene Pulser). Then 10 mL of MADC-TW or MADC-TWP was added and the cells were allowed to recover at 37°C for 24 h. 100 µl of the transformation mixture was plated out on 7H10 agar, containing the appropriate antibiotic and incubated at 37°C for 6 - 8 weeks.

2.6.8. Isolation of plasmid DNA from *E. coli*

2.6.8.1. Mini prep

Plasmid DNA was purified by the QIAprep Miniprep DNA kit (Qiagen USA) according to the manufacturer's instructions. 4 mL of an overnight culture of *E. coli* was centrifuged at 12,000 g for 5 min; the pellet was re-suspended in 250µL of suspension buffer P1, followed by addition of 250 µl Buffer P2 to effect cell lysis. The bacterial lysate was neutralised by adding 350 µl of buffer N3. The neutralised lysate was then centrifuged using a bench-top centrifuge for 10 min at 13,000 rpm and the supernatant transferred to a Miniprep spin column, centrifuged for 2 min at 13,000 rpm, washed with 750 µl of purification solution followed by centrifugation at 13,000 rpm for 1 min. The spin column was transferred into fresh microcentrifuge tube and plasmid DNA was eluted by incubating with 30 µl of nuclease free water (heated to 65°C prior to use) and spinning the column / collection tube for 2 min at 12,000 rpm. The elution was repeated to increase the DNA yield.

2.6.8.2. Midi Prep

Bacterial cell pellet was re-suspended in 4 mL of buffer P1 containing RNaseA at 100 µg / mL, followed by addition of 4 mL of lysis buffer P2 with gentle mixing by inverting. The mixture was left at RT for 5 min, followed by 4 mL of chilled buffer P3. The lysate was poured into a QIA filter cartridge and incubated for 10 min. While incubating, the Hi-Speed Midi Tip was equilibrated with 4 ml of buffer QBT. The contents from the QIA filter cartridge were forced out through the resin using a plunger into the equilibrated Hi-Speed Midi tip. The lysate was allowed to clear the column by gravity flow. Plasmid DNA bound to the column was washed twice with

10 mL of QC wash buffer. The plasmid DNA was eluted with 5 mL of QF elution buffer. Plasmid DNA was precipitated with 3.5 mL of isopropanol at RT for 5 min. The contents were filtered through the QIA precipitator. Plasmid DNA bound to the QIA precipitator was washed with 2 mL of 70% (v / v) EtOH and recovered using 0.5 ml of distilled water.

2.6.9. Analysis of recombinants

a) PCR

The plasmid DNA isolated by the mini / midi prep were used as template in PCR for identification of the inserts in *E. coli*.

For mycobacterial mutants, a single colony was grown in 10 mL of MADC-TW or MADC-TWP supplemented with 200 µg / mL hygromycin or 25 µg / mL kanamycin or both for 7 days at 37°C. Two millilitres of this culture was pelleted by centrifugation at 12,000 g for 10 min at 4°C and washed three times in 2 mL of Milli-Q water. The pellet was resuspended in 100 µl of water and incubated at 80^{o}C for an hour. The solution was centrifuged briefly for 1 min and the clear supernatant was used as template in PCR.

b) DNA sequence analysis

Cycle sequencing reactions were analysed at University of Dundee, Scotland, UK using an ABI 377 DNA sequencer. For each reaction a sequence analysis document containing information on the sequence was returned. These ABI traces were analysed using the computer program Seqman (DNASTAR). Contiguous sequences were assembled using Seqman and converted to DNA files in the computer program Editseq (DNASTAR). The sequences were mapped on the *M. tuberculosis* genome using the BLASTN programs within the National Centre for Biological information (NCBI) home page (http://www.ncbi.nlm.nih.gov), Tuberculist (http:// genolist.pasteur.fr/TubercuList/), or Bovilist (htpp://genolist.Pasteur.fr/BoviList).

2.7. Generation of Knock in and Knock out mutants in mycobacteria

The mutants to be generated included the following

(a) M. tuberculosis Δ ald

(b) *M. tuberculosis Δ pykA*

(c) *M. tuberculosis Δ ald Δ pykA*

(d) *M. bovis :: ald and M. bovis* BCG *:: ald*

(e) *M. bovis :: ald :: pykA*

The overall methodology used included the following steps

1) Knock in mutants were generated using complementing plasmid vectors pSM96 (Fig. 15) and pSM 81, while knock out mutants were generated using suicide plasmid pSMT100 (Fig. 16) as vector Table 11.

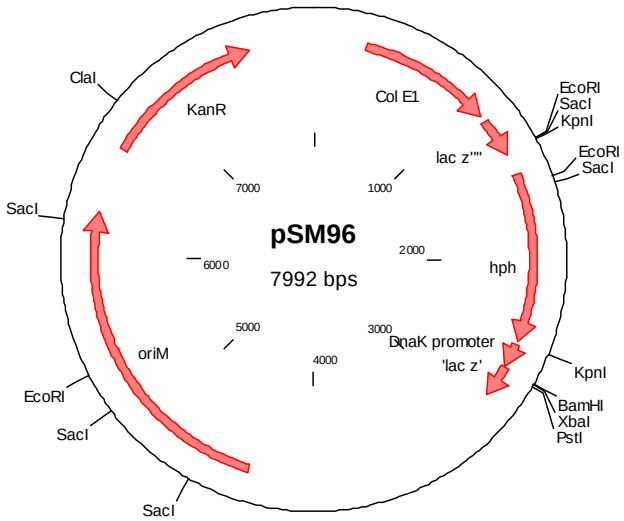


Fig. 15 - pSM96 Complementing expression vector used for creating knock in mutants

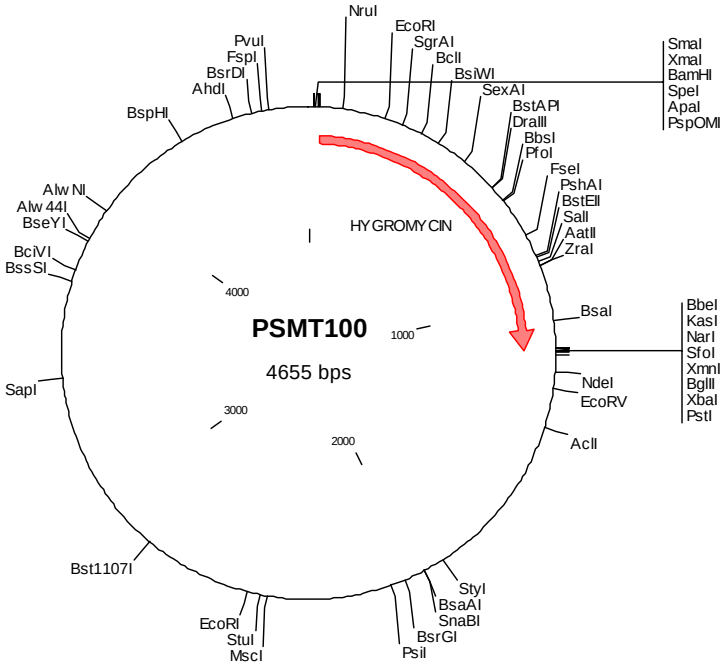


Fig. 16 - pSMT100 Suicide vector used for creating knock out mutants

Table 11 - Plasmids used in this study

Plasmid	Reference/Source	Description
PSM96	Dr Stephen Michell, VLA-Weybridge	Mycobacterial high-level expression vector with kan^R and hyg^R and contains a strong mycobacterial promoter hsp70 upstream of multiple cloning site derived from pSM81.
PSM81		Mycobacterial high-level expression vector, consist of strong mycobacterial promoter hsp 70 upstream of multiple cloning site, kan^R
pSMT100	Graham Stewart (University of Surrey, UK)	Mycobacterial suicide vector, utilised to create knock out mutants based on homologous recombination

2) Knock in mutants was created by PCR amplification of *ald* and *pykA* from *M. tuberculosis* using specific primers.

3) PCR products were cloned into *E. coli* DH5α, followed by verification of the recombinant by suitable restriction digestion for the presence and orientation of the insert. The resultant vector was verified by sequencing.

4) Knock out mutants were generated by first constructing vectors by cloning upstream and downstream sequences of the respective gene and then introducing them in the respective mycobacterial species by homologous recombination.

5) The mycobacteria were transformed by electroporation with the vectors to generate both knock in and knock out mutants.

6) All constructs were verified by PCR.

7) Characterisation of mutants were done by i) enzyme activity, ii) colony morphology and iii) lipid analysis and iv) proteome analysis.

2.7.1. Construction of *ald* knock-in *M. bovis* and *M. bovis* BCG

a) PCR amplification of *ald* from *M. tuberculosis* and *M. africanum*

The 1116 bp ald gene was amplified using the High fidelity PCR kit (Qiagen) from 2 clinical strains of *M. africanum* and *M. africanum* 2 using the primers,

Forward: ALDFC 5'- AAC TGC AGC ATG CGC GTC GGT ATT CCG ACC-3'

Reverse: ALDRC 5'- AAC TGC AGT AAC GAG CGG CCG AGA GTC AGG-3'

PCR conditions:

i) Initial denaturation at 94^{o}C for 15 s

ii) 30 cycles of 94^{o}C for 15 s, 68^{o}C for 4 min and a final elongation for 3 min at 68^{o}C.

(b) Cloning the *ald* gene into *E. coli*

The 1116 bp PCR product (*ald* gene amplified from *M. tuberculosis* H37Rv Fig. 3a) was digested with *PstI* and ligated to the *PstI* digested dephosphorylated pSM96 DNA (Fig. 3b**).** The ligation mixture was transformed into *E. coli* DH5α by heat shock method. The recombinant, henceforth called pSR1 was confirmed by

restriction digestion and sequencing. The sequence of the cloned fragment was confirmed using the following primers

ALDTBSF: 5'- CGG CGA GTC TTG TGT TCC GTT –3',

ALDTBSF2: 5'- GGC TAT CAC CGA CGC GGA TTT C -3'

ALDSR: 5'-GCT TGC ATG CCT GCA GTA AC-3'.

2.7.2. Construction of plasmid vector for *ald* knockout mutant in *M. tuberculosis*

To generate the plasmid vector for inactivation of *ald* by homologous recombination in *M. tuberculosis,* approximately 1000 bp from both upstream and downstream regions of the *ald* were PCR amplified. Due to constraint for using the restriction enzyme, the plasmid vector was designed in such a way so as to delete a 300 bp internal region from *ald.*

(a) PCR amplification and cloning of the upstream region

Forward: ALDLFF 5'- AAC TGC AGC CCA GGT TGG CGC TTA CCA C-3'

Reverse: ALDLFR 5'- AAC TGC AGC CGG TAC CAA TGC CGC ACA C-3'.

The 1030 bp PCR product, digested with *Pst*I and ligated into the *Pst*I site of *Pst*I digested, dephosphorylated pSMT100 was transformed into *E. coli* DH5α. The clones were verified for their orientation by restriction digestion.

(b) PCR amplification and cloning of the downstream region

Forward: ALDRF 5'- ATT TCC CGG GTT TCG TTG CGA CCG ACA AAC AC- 3'

Reverse: ALDLFR 5'- AAT TAC TAG TCA AGA TCT GCC CGT GTC GCAGG- 3'

The 1113 bp PCR product, digested with *XmaI* and *SpeI* and ligated to pSN1 with compatible ends was transformed into *E. coli* DH5α. The recombinant, henceforth called pSN2 were analysed by restriction digestion and sequencing. The sequence of the cloned fragment was confirmed using the following primers

ALDSRF: 5'-GGC GAT TAA GTT GGG TAA CG-3',

ALDSRF/2: 5'-CGA GAT CAT CGC CTG CAA CG-3',

ALDSRR: 5'-GTA CGG GCC GGC TTA ATT CC-3',

ALDSLF: 5'-AAC TGG CGC AGT TCC TCT GG-3',

ALDSLF/2: 5’-CGG CTG TTT CGA AGG CTC AC-3’ and
ALDSLR: 5’-AAG CAG GCA AAT GGC TAA CC-3’.

2.7.3. Construction of plasmid vector for *pykA* knockout mutant in *M. tuberculosis*

The plasmid vector for inactivation of *pykA* by homologous recombination in *M. tuberculosis* was generated by PCR amplification and cloning of both upstream and downstream regions of *pykA.*

(a) PCR amplification and cloning of the upstream region

Forward: PYKARFF 5’-ATA ATC TAG AGG TCG TGC CGG ACG GTA AAC-3’
Reverse: PYKARFR 5’-AAT ATC TAG AGG CCG ACG GCA GCT TCA CAG-3’.

The 1059 bp PCR product digested with *XbaI* and ligated into *XbaI* site vector pSMT100 was transformed into *E. coli* DH5α

(b) PCR amplification and cloning of the downstream region

Forward: PYKALFF 5’-ATT TCC CGG GTT GAG CCA ACT CAG GAA ATG- 3’
Reverse: PYKALFR 5’-ATT AAC TAG TGC CCG GGT CGG CAC ACC GGTC-3’.

The 1056 bp product, digested with *Xma I* and *Spe I* and ligated to pSC1 with compatible termini was transformed into *E. coli* DH5α. Positive clones (pSC2) were analysed by restriction digestion and sequencing. The sequence of the cloned fragment was confirmed using the following primers

PYKASRF: 5’-AAC TGG CGC AGT TCC TCT GG-3’,
PYKASRF / 2:5’-GTG AGC TGT TGC GCG GTT AC-3’,
PYKASRR: 5’-AAG CAG GCA AAT GGC TAA CC-3’,
PYKASLF: 5’- TGA ATC TCT GAG CGG AGG TG-3’,
PYKASLF / 2: 5’-CAT TGA CGC CCT GCT GAT ATT G-3’ and
PYKASLR: 5’-GTA CGG GCC GGC TTA ATT CC-3’.

2.7.4. Construction of plasmid vector for double knock out mutant (*pykA* and *ald)* in *M. tuberculosis*

The hygromycin cassette in pSC2 vector (explained above) was replaced by kanamycin resistance cassette as shown below.

(a) The kanamycin cassette present in pSM96 was PCR amplified by using primers.

Forward: KanF 5'-ATA TAC TAG TCG GGC GCG ATC CAG AAG AAC-3'

Reverse: KanR 5'-AAT AGA TCT GCT TGC CGT CCC GTC AAG TC-3'.

This 1130 bp PCR product was digested with *Spe I* and *Bgl II* and ligated into pSC2 cut with these two enzymes that removed the hygromycin cassette. The recombinant, henceforth called pCM2 was transformed into *E. coli* DH5α. Sequence confirmation was done using primers

PYKASRF: 5'-AAC TGG CGC AGT TCC TCT GG-3',

PYKASRF / 2: 5'-GTG AGC TGT TGC GCG GTT AC-3',

PYKASRR: 5'-AAG CAG GCA AAT GGC TAA CC-3',

PYKASLF: 5'- TGA ATC TCT GAG CGG AGG TG-3',

PYKASLF / 2: 5'-CAT TGA CGC CCT GCT GAT ATT G-3'

PYKASLR: 5'-GTA CGG GCC GGC TTA ATT CC-3'.

2.7.5. Construction of plasmid vector for complementing with both *pykA* and *ald* in *M. bovis* T-35

(a) Cloning *ald* into *E. coli*

The *ald* gene (amplified as shown section 2.7.1) using the primers

Forward: ALDDIF 5'- AAC TGC AGA TGC GCG TCG GTA TTC CGA CCG-3'

Reverse: ALDDIR 5'- AAC TGC AGT TAA TTA AAG TAC TCC TCC TTC AGG CCA GCA CGC TGG CGG G-3'

The PCR product was cloned into the *Pst I* site of pSM96 and transformed into *E. coli* DH5α to get the recombinant plasmid pMS1.

(b) Cloning *pykA* gene into *E. coli*

Forward:

PYKADIF 5'- CCC AGT ACT GTG ACG AGA CGC GGG AAAATC G-3'

Reverse:

PYKADIR 5'- CCG CTT AAT TAA CTA GAC GTC ATC TTC CCC GATCCG –3'.

The pykA fragment was subjected to double digestion with *Pac* I and *Sca* I and cloned into pMS1 and transformed into *E. coli* DH5α. The newly generated recombinant plasmid pMS2 was sequenced using primers

ALDDISF: 5'-AAT TGA CAG CAC ACC GCC GTC CAG G-3',

ALDDISF2: 5'-CGC GAC CGT TAC GGT TCT AGA C-3',

PYKADISF: 5'-AAT GGA CGT CGC CCG AAT GAA C-3',

PYKADISF2: 5'-AGA GGT CCC GCT GGT ACA GAA G-3' and

PYKDISR: 5'-GCT GCA AGG CGA TTA AGT TGG G-3'.

The plasmid vectors created in this study are listed in table 12.

Table 12 - Plasmids created in this study

Plasmid	Reference /Source	Description
PSR1	This study	Mycobacterial high level expression vector ($hyg^{R,}$ kan^{R}) containing *ald* gene, used for complementing *M. bovis* T-35, *M. bovis* BCG Pasteur.
PMS2	This study	Mycobacterial high level expression vector (hyg^{R}, kan^{R}) containing *ald* and *pykA* used for complementing *M. bovis* T-35.
PSN2	This study	Mycobacterial suicide vector having the upstream and downstream of the *ald* gene and hyg^{R} used for creating *M. tuberculosis* H37Rv Δ *ald*
PSC2	This study	Mycobacterial suicide vector having the upstream and downstream region of the *pyk A* gene and hyg^{R} used for creating *M. tuberculosis* H37Rv Δ *pykA*
PCM2	This study	Mycobacterial suicide vector having the upstream and downstream of the *pyk A* gene and kan^{R} used for creating *M. tuberculosis* H37Rv Δ *pyk A* Δ *ald*
PLK102	Keating *et al.*, 2005	Mycobacterial high level expression vector (kan^{R}) containing *pyk A* gene, used for complementing *M. bovis* T-35, *M. tuberculosis* H37Rv Δ *pykA*

2.8. Generation of mycobacterial mutants

2.8.1. Generation of *M. bovis*:: *ald* and *M. bovis* BCG :: *ald* by complementing with pSR1

Plasmid pSR1 was introduced into *M. bovis* and *M. bovis* BCG by electroporation. Potential recombinants were selected by growing them on Middlebrook 7H10 agar with hygromycin. Colonies that were resistant to hygromycin were grown in Middlebrook 7H9 + hygromycin for 7 days.

The recombinants were confirmed by PCR using hygromycin cassette-specific primers,

Forward: HYGF 5'- CTT CAC CGA TCC GGA GGA ACT G-3'

Reverse: HYGR 5'- TTG TGT CAC AGC GGA CCT CTA T-3'.

PCR reaction conditions:

i) Initial denaturation for 15 min at 94^{o}C,

ii) 30 cycles consisting of denaturation at 94^{o}C for 1 min, annealing at 55^{o}C for 1 min and elongation for 1 min at 72^{o}C,

iii) Final extension for 5 min at 72^{o}C.

Electroduction was also used to verify presence of plasmid and their insert in mycobacteria.

2.8.2. Generation of *M. bovis:: ald :: pykA* by complementing with pMS2

Plasmid pMS2 was introduced into *M. bovis* by electroporation. Potential recombinants were selected by growing them on Middlebrook 7H10 agar with hygromycin. Colonies that were resistant to hygromycin were grown in Middlebrook 7H9 + hygromycin for 7 days. These colonies were further screened by Hotstart PCR analysis as explained above.

2.9. Allelic replacement for creating knockout mutants in *M. tuberculosis*

2.9.1. *Mycobacterium tuberculosis Δ ald*

Plasmid pSN2 was introduced into *M. tuberculosis* H37Rv by electroporation. The recombinant knockouts were selected based on their antibiotic resistance by

growing them on Middlebrook 7H10 with hygromycin. The colonies were inoculated in Middlebrook 7H9 + hygromycin and harvested after 7 days. DNA was prepared as explained (Chapter 2. section 2.0.5.2). Colonies were further screened by Hotstart and XL PCR, followed by enzyme assays.

Hotstart PCR analysis:

(a) Forward: HygF and Reverse: HygR

(b) Forward: Hygout1 5'- AGG GTA CGG GCC GGC TTA ATT C-3' and
Reverse: ALDLFR

(c) Forward: ALDRFF and
Reverse: Hygout3 5'- CTT CAC CGA TCC GGA GAT TCC T-3'

The PCR reaction was done as per above protocol.

XLPCR analysis:

Forward: ALDFC and Reverse: ALDRC.

XLPCR reaction condition:

i) Initial denaturation for 30 s at 94^{o}C,

ii) 30 cycles of denaturation at 94^{o}C for 15 s, combined annealing and extension at 68^{o}C for 6 min and

iii) Final extension at 72^{o}C for 10 min.

2.9.2. *Mycobacterium tuberculosis* Δ *pykA*

Plasmid pSC2 was introduced into *M. tuberculosis* H37Rv by electroporation. The recombinant knockouts were selected and analysed as follows:

Hotstart PCR analysis:

(a) Forward: Hygout1 and Reverse: PYKALFF

(b) Forward: PYKARR and Reverse: Hygout3

(c) Forward: PYKABamHIF 5'-CTG AGG ATC CAT GAC GAG ACG CGG GAA AAT C-3'

Reverse: PYKABamHIR 5'-GTC CGG ATC CAC TAG TCT AGA CGT CAT CTT CCC CGA-3'.

The strain was named as *M. tuberculosis* H37Rv Δ *pykA.*

2.9.3. *Mycobacterium tuberculosis* Δ *ald* Δ *pykA*

Plasmid pMC2 was introduced into *M. tuberculosis* H37RvΔ*ald* by electroporation. The recombinant knockouts were selected as explained previously with the following changes.

Hotstart PCR analysis:

a) Forward: KANF and Reverse: KANR.
b) Forward: ALDFC and Reverse: ALDRC
c) Forward: PYKAFBamHI and Forward: PYKARBamHI

XLPCR analysis

a) Forward: KANF and Reverse: PYKARFR and
b) Reverse: KANR and Reverse: PYKALFF.

The strain was named as *M. tuberculosis* Δ *ald* Δ *pykA.*

2.9.4. Complementing *M. tuberculosis* Δ *pykA* with pLK102

Plasmid pLK102 (Keating *et al.,* 2005) was introduced by electroporation into *M. tuberculosis* Δ *pykA*, to study the effect of *pykA* by recomplementing. The recombinant knockin mutants were selected based on antibiotic resistance by growing them on Middlebrook 7H10 + Hygromycin and kanamycin. Hotstart PCR analysis using the primers

KANF6 (5’ CTC CTG ATG ATG CAT GGT TAC TCA C-3’)
KANR7 (5’-ACT CAC CGA GGC AGT TCC ATA GGA T-3’)

was done to identify the mutants containing pLK102 The strain was named as *M. tuberculosis* Δ *pykA* pLK102.

2.9.5. Complementing with expression vectors

In order to find the effect of the expression vectors, plasmids pSM96 and pSM81 without the gene of interest were electroporated into *M. bovis* and *M.*

tuberculosis Δ *pykA* respectively. Hotstart PCR analysis was used to identify the mutants.

2.10. Enzyme assays

2.10.1. Preparation of cell extracts

Cells were harvested at mid-logarithmic stage (OD_{600} = 0.5 to 0.6) by centrifuging at 4800 rpm for 10 min. The cell pellet was washed twice and resuspended in 20 mM phosphate buffer (pH 6.8) containing 5 mM $MgCl_2$. The chilled cells were disrupted using the Fast Prep® system (Hybaid) for 20 s at a speed of 6.5 m / sec four times and the supernatant containing the enzymes were harvested by centrifugation at 14, 000 x g for 5 min at 4°C (Keating *et al.,* 2005). Extracts were prepared and assayed from at least two independent cultures. When no activity was detected extracts were prepared from four independent cultures for confirmation.

The protein concentration in samples used for enzyme assays was estimated by Bradford (Bradford, 1976).

2.10.2. Pyruvate kinase assay

Coupled enzyme assay was used to measure the activity of pyruvate kinase by measuring the conversion of $NADH^+$ to NAD^+ by lactate dehydrogenase. The rate of conversion of $NADH^+$ to NAD^+ was monitored spectrophotometrically at 340 nm at 37^{o}C. The final reaction mixture in 1 mL volume consisted of 80 mM Tris-HCl pH 7.4, 2.5 mM $MgSO_4$, 10 mM KCl, 0.2 mM NADH, 4.7 mM ADP and 1.2 U / mL lactate dehydrogenase. 100 µl of crude cell extract was used as enzyme source and the reactions were initiated by adding 0.54 mM PEP (final concentration) (Keating *et al*, 2005). Two negative reactions were always performed one without the substrate and other without the enzyme. The rate of conversion was calculated using the rate constant equation calculating the amount of NAD^+ formed. Each sample was analysed at least twice and three sets of different cultures were utilised to analyse the activity of each enzyme.

2. 10.3. Alanine dehydrogenase assay

L-alanine dehydrogenase activity was measured by conversion of $NADH^+$ to NAD^+ accompanying the production of pyruvate from alanine. Final reaction mixture consisted of 200 mM ammonium sulphate, 0.2 mM NADH, 20 mM phosphate buffer (pH 6.8) and the reactions were initiated with 50 mM sodium pyruvate, as substrate. 50 µl of crude cell extract was used as the source of enzyme. The conversion of $NADH^+$ was followed spectophotometrically at 340 nm at 37^{o}C (Hutter & Singh, 1999). Two negative reactions were always performed one without the substrate and other without the enzyme. The rate of conversion was calculated using the rate constant equation calculating the amount of NAD^+ formed. Each sample was analysed at least twice and three sets of different cultures were utilised to analyse the activity of each enzyme.

2.11. Colony morphology

Cultures were grown in MADC-TW for 8 days, washed, resuspended and serially diluted in PBS / 0.05% Tween 80. 25 µl samples of each 10-fold dilution were spotted in five or six spots adjacently onto Middlebrook 7H11 agar plates using 0.5% glycerol or 40 mM pyruvate as carbon source along with OADC. Control plates were made using without any carbon source except OADC and incubated for four weeks at 37°C. Antibiotics were included as required (Keating *et al.,* 2005).

2.12. Utilisation of carbon sources

The ability of a carbon source to support growth was determined by the method of Meyers *et al.,* 1998. The effect of deletion of *pykA* and *ald* was determined by growing various mutants along with their wild type on a range of sole carbon sources in Sauton's minimal medium. Carbon sources were added to a final concentration of 40 mM for sodium pyruvate, 40 mM for glucose, 0.05% (v/v) for Tween 80 (+ 0.5% BSA) and 6% (v/v) for glycerol, or omitted as a negative control. Bovine Serum Albumin (BSA) was added to media containing Tween 80 as the sole

carbon source since this counteracts the toxic build-up of oleic acid from Tween hydrolysis; BSA does not act as a carbon source for mycobacteria.

2.13. Preparation of cells for lipid extraction

Cells were harvested at mid-logarithmic stage and washed twice with equal amounts of Milli Q water. The cells were transferred to screw capped pyrex tube and the pathogenic mycobacteria were heat killed by incubating the cells at 80^{o}C for an hour. This was then freeze dried to get dried biomass.

2.13.1 Extraction of free mycobacterial lipids: Non-polar lipids

Lipids were extracted from 10-50 mg of dried biomass by the following method. To 20 mg of dried biomass, 2 mL of aqueous MeOH (10 mL of 0.3% NaCl: 100 mL of MeOH) was added along with 1 mL of hexane or petroleum ether (60^{o}C-80^{o}C). This was mixed in blood tube rotator for 15 min and then centrifuged for 5 min at 250 g. The upper layer was collected separately, and to the lower layer, 1 mL of hexane was added and the non-polar lipid was extracted as above. Resulting upper layer was added with first and the lower layer was retained for the extraction of polar lipid. This was dried completely with a stream of nitrogen gas at 37^{o}C. The non-polar lipids were dissolved in appropriate amount of chloroform: MeOH (4:1). If it appeared turbid and hazy due to the presence of proteins, it was filtered through PTFE filters. The samples were then stored at 4^{o}C for TLC analysis using the solvent system given in Table 13.

2.13.2 Polar lipids

After the extraction of nonpolar lipids, the cell pellet was heated at 100^{o}C for 5 min to remove the hexane. Care was taken not to over boil. After cooling to RT, 2.3 mL of Chloroform: MeOH: 0.3% NaCl (90:100:30) was added, mixed for 1 h by rotation and centrifuged for 5 min at 250 g. The supernatant was aspirated carefully and stored separately. The cells were extracted by adding 0.75 mL of chloroform: MeOH: 0.3% NaCl (50:100:40), mixed for 30 min and centrifuged for 5 min. The supernatant was added with the above and the process was repeated once again. The

pellet was used for the extraction of mycolic acids and cell wall bound fatty acids. The supernatants were combined, then 1.3 mL chloroform and 1.3 mL of 0.3% NaCl was added. The tubes were mixed for 30 min. The layers were then separated by centrifuging at 250 x g for 5 min. The lower layer, which contains the polar lipids was dried completely by a stream of nitrogen gas at 37^{o}C. Polar lipids were dissolved with appropriate amount of chloroform: MeOH (2:1) and stored at 4^{o}C for further analysis by HPTLC as explained in Table 14.

Table 13 - Two dimensional HPTLC analyses of Polar and non-polar lipids

System	Dimension	Solvent composition	No. of times	Detection	Lipids Detected
A	1	Hexane: Ethyl acetate (98:2)	3	MPA	Triacylglycerols, menaquinones, PDIM
	2	Hexane: Acetone (98:2)	1		
B	1	Hexane: Acetone (92:8)	3	MPA	Trehalose mycolipenates
	2	Toluene: Acetone (95:5)	1		
C	1	Chloroform: MeOH (96:4)	1	MPA NAS	Glycosylated phenol-phthiocerols
	2	Toluene: Acetone(80:20)	1		

Table 14 - Single dimensional analysis for Polar fraction

Solvent composition	Length (cm)	No of times
Chloroform: Methanol (9:1)	20	1
Chloroform: Methanol:Water (14:6:1)	10	1

2.13.3 Extraction of cell wall bound fatty acids and mycolic acids by esterification

The cells were heated with 1 mL of 15% aqueous tertiary butyl-ammonium hydroxide (TBAH) overnight at 105^{o}C. Then they were cooled to RT and 1 mL of water (MilliQ) and 2 mL of dichloromethane was added. 100 µl of methyl iodide was

added immediately to this mixture. The upper aqueous layer was removed after 30 min incubation at RT. The lower layer was washed twice with 2 mL of 0.1M HCl and then dried completely as above (Dobson *et al,* 1985). The mycolic acids were further anlaysed by HPTLC using the solvent system Hexane: acetone (19:1).

2.14. Overall proteome analysis of mutants

2.14.1. Preparation of protein samples

Cultures were grown upto mid log phase. Bacteria were pelleted down by centrifuging at 4800 rpm for 15 minutes. The cell pellet was washed twice with 50 mL PBS. The pellet was resuspended in 3 mL of PBS and mixed well by pipetting. The samples were sonicated for 10 min with 9 s pulse using Sonics VX 500. The cell debris was removed by centrifuging at 1400 rpm for 10 min. An equal amount of denaturing buffer 1 (50 mM Tris, pH 7.8, 10% glycerol, 10 mM DTT and 0.2% SDS) was added to the supernatant. Samples were boiled at 95^{o}C for 5 min to ensure that no viable bacteria remained in the suspension. The samples were concentrated to 1 mL using 5 kDa MWCO Vivaspin concentrators. To the concentrated sample, 2 mL of freshly prepared lysis buffer (final concentration urea 5 M, thiourea 2 M, CHAPS 2%, SB3-10 2%, Pharmalytes 3-10 0.5%, DTT 100 mM, Tris base 0.01 M) was added and incubated for an hour at 37^{o}C. The samples were again concentrated to 0.5 mL and washed by dialysing with 2.5 mM ammonium bicarbonate solution. Proteins in the sample were estimated by Bradford method and stored as aliquots at -80^{o}C.

2.14.2. Trypsin digestion

100 µg of protein fraction was made up to 0.09 mL with ammonium bicarbonate (2.5 mM; pH 8.0). Trypsin (20 µg) was made up with 0.1 mL in ammonium bicarbonate (2.5 mM; pH 8.0) and 10 µl was added to the protein extracts. The solution was mixed well and incubated overnight at 37^{o}C. The hydrolysis was stopped by addition of 1 µl formic acid. Samples were centrifuged at 12,000 rpm for 1 min and supernatant (0.1 mL) was pipetted into fresh HPLC vial insert.

2.14.3 Two dimensional HPLC analysis

2.14.3.1 SCX chromatography (1st dimension)

Tryptic peptides were chromatographed using a Hewlett-Packard 1050 system on a Biobasic SCX HPLC column (100 x 2.1 mm, 5 μm). Analytes were eluted at a flow rate of 0.25ml/ min with mobile phases comprising of 2.5 mM ammonium acetate (pH 4.5): acetonitrile (75:25) (A) and 250 mM ammonium acetate (pH 4.5): acetonitrile (75:25), (B) (Table 15). HPLC effluent was recorded at 280 nm and 15 fractions of 1 min duration were collected from 8 and 22 min. The SCX fractions were taken to dryness under reduced pressure at 50^{o}C using centrifugal concentrator (Gyrovap, GT, Howe, UK).

Table 15 - HPLC gradient for SCX chromatography

Time (min)	Buffer A %	Buffer B %	Flow rate (ml/min)
0	100	0	0.25
5	100	0	0.25
18	65	35	0.25
20	0	100	0.25
22	100	0	0.25
32	100	0	0.25

2.14.3.2 Reversed phase chromatography (2nd dimension) and Mass analysis

SCX fractions were dissolved in 0.1% (v / v) formic acid and analysed by capillary-HPLC-MS using a Famos auto sampler, a surveyor HPLC pump equipped with an Accurate flow splitter (100 to 1) and LCQ ion-trap mass spectrometer (ThermoFinnigan, Hemel Hamstead, UK). SCX fractions (5μl) were chromatographed on a Picofrit column and peptides eluted at an estimated flow rate of 400 nL / min with a linear binary gradient of 0.1% formic acid (A) and acetonitrile (B) as explained in table 16. The Picofrit column (70 mm) was mounted in a steel filter (0.5 μm pore size) assembly to which the electrospray ionisation potential (2.5 kV) was applied. The Picofrit tip was positioned slightly off axis within 3 mm of the LCQ mass

spectrometer heated capillary inlet using a Protana nanospray source. Mass data for each SCX fraction was collected over a mass of 300-2000 using a Big Three acquisition method with data dependent product ion scanning of 1st, 2nd and 3rd most abundant ions above a threshold trigger of 3 x 10^5 counts / sec. The mass isolation window and collision energy were set to 4 amu and 35% respectively. The mass spectrometer was tuned to the doubly charged ion (m/z 820.5) derived from synthetic peptide FNPGELLPEAAGPTQV.

Table 16 - HPLC gradient for analysis of SCX fractions

Time (min)	Buffer A%	Buffer B%
0	100	0
40	60	40
50	40	60
52	100	0
75	100	0

2.14.3.3 Bioinformatic data analysis

The SEQUEST algorithm embedded within the bioworks software package was used for the identification of proteins from tryptic peptide mass spectra. Proteins were identified by comparison of tryptic peptide product ion mass spectra against those generated from both *M. tuberculosis* and *M. bovis* database derived from Tuberculist and Bovilist. The multi consensus report function was used to assign tryptic peptides to individual proteins and compile rank listings of the proteomes. Identified proteins were ranked in ascending order according to consensus scores and false positives identification was minimised by filtration against 4 of the 5 following criteria Xcorr >2.0. DeltCn >0.2, Sp >400, rsp <5, ions >30%. Compilations of identified proteins from each triplicate were exported to Microsoft Access for interrogation and comparison using the query function.

2.14.3.4 Assessment of assay performance

The performance of the SCX and reversed phase chromatographic separation and sensitivity of the mass spectrometer was monitored by analysis of specific standards at the beginning of each batch. The SCX procedure was assessed by analysis

of the blanks to control for carryover between samples and a standard preparation of tryptic peptides derived from BSA (25 μg / mL) to ensure reproducible chromatography and sensitivity of the mass spectrometer was assessed by monitoring the retention time and intensity of ions derived from a low level standard prepared from the synthetic peptides MRFA (50 ng / mL; $[M+H]^+$ m/z 524.2); APYELNITSATYQSAI (10g / mL; $[M+H]^{2+}$ m/z 871.5); FNPGELLPEAAGPTQV (10g / mL; $[M+H]^{2+}$ m/z 820.5) and AIQGNVVSTSIHSLLDEG (10 g / mL; $[M+H]^{2+}$ m/z 827.9) (Coldham and Woodward, 2004).

CHAPTER 3

STUDIES ON THE IRON ACQUISITION MACHINERY IN *MYCOBACTERIUM KANSASII*

3.0 Results

3.1. Establishment of growth under iron-regulated conditions

Several mycobacterial species were initially tested for growth under high (8 μg Fe / ml) and low (0.02 μg Fe / ml) iron conditions respectively. Carboxymycobactin (CMb) was used as an indicator for the establishment of low iron condition (Fig. 17).

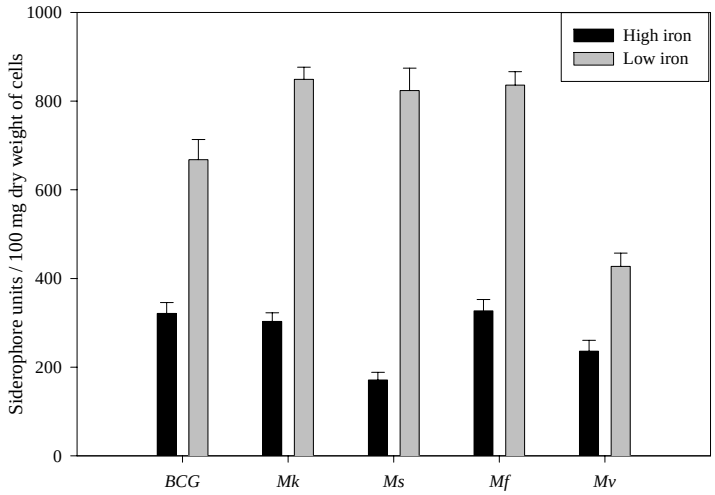


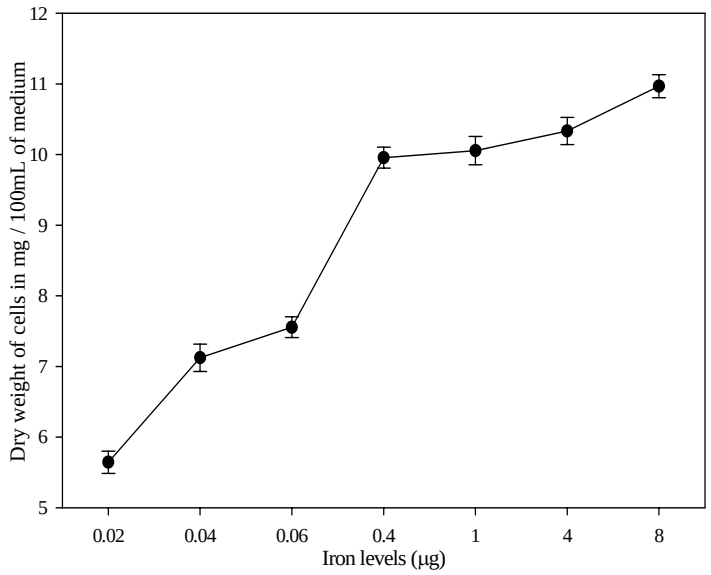
Fig. 17 - Analyses of carboxymycobactin / exochelin by CAS assay in different mycobacterial species. Cells were grown in high (8 μg Fe / ml) and low (0.02 μg Fe / ml) iron conditions and were harvested during mid log phase. The siderophores calculated from three separate experiments were used and expressed as Siderophore units / 100 mg dry weight of cells. BCG, *Mk, Ms, Mf* and *Mv* represent *M. bovis* BCG, *M. kansasii*, *M. smegmatis*, *M. fortuitum* and *M. vaccae* respectively.

3.1.1. Effect of iron levels on *M. kansasii*

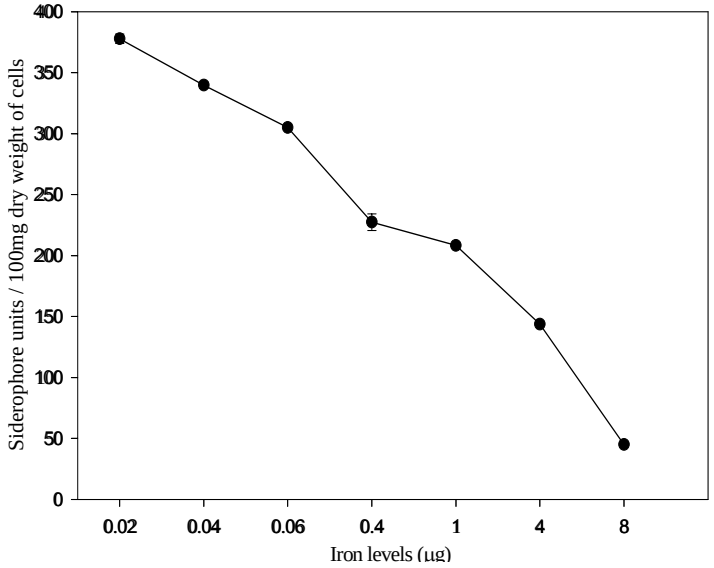
Mycobacterium kansasii was grown in medium with varying iron concentrations (ranging from 0.02 μg Fe / ml (0.36 μM) to 8 μg Fe / ml (144 μM). The growth of the organism as monitored by determination of the cell dry weight (Fig. 18a) showed increase in the cell dry weight with increase in iron levels in the medium. The levels of the two siderophores was significantly high when iron was used at 0.02 μg of Fe / mL of medium while it was negligible in organisms maintained at 8 μg of Fe / mL of medium (Fig. 18b & 18c). Subsequently, the organisms were grown under high (8 μg Fe / ml) and low iron (0.02 μg Fe / ml) conditions and time course

expression was done by harvesting over a period of 2 weeks (Fig. 19a). Both the siderophores (Fig. 19b & 19c) reached a peak on the 7th day. All further experiments were done using 8 μg Fe / ml for high and 0.02 μg Fe / ml for low iron growth and harvested after 7 days of growth.

(a)



(b)



(c)

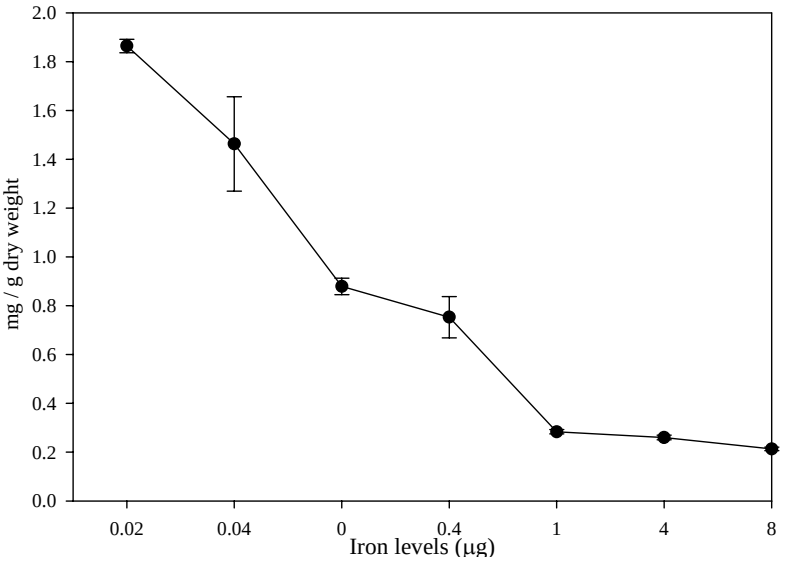
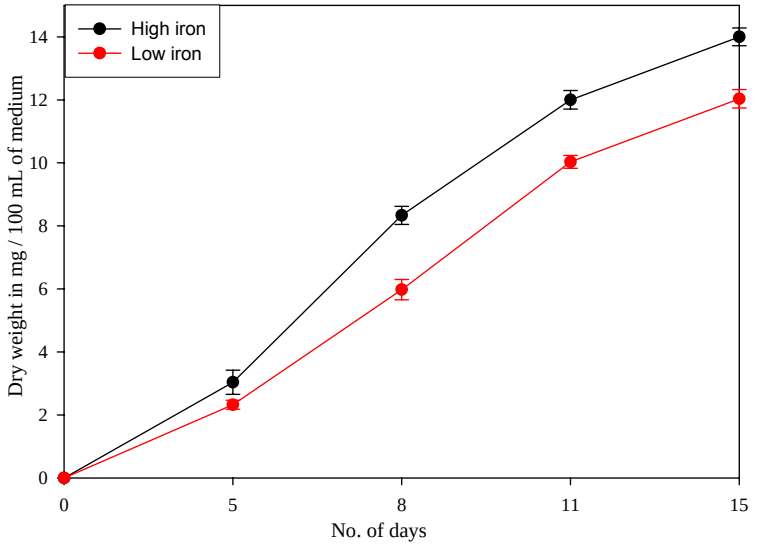
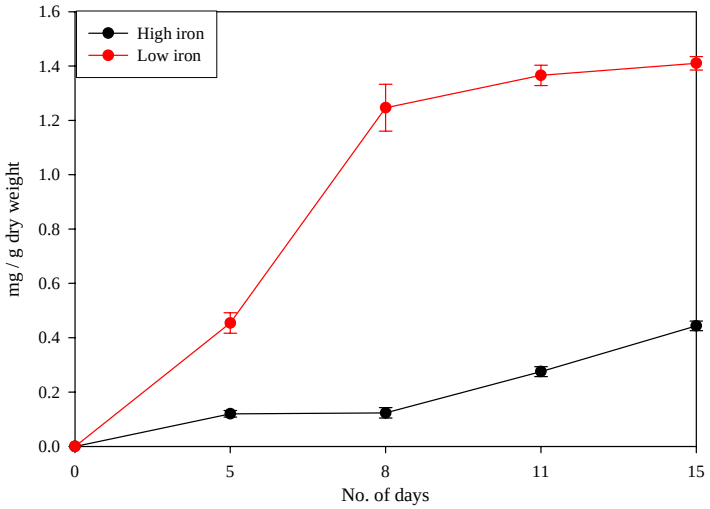


Fig. 18 - Growth and coordinated expression of CMb and Mycobactin (Mb) in *M. kansasii* with varying concentrations of iron, (a) growth curve, (b) estimation of CMb and (c) Mb. The cells were grown in varying concentrations of iron (0.02, 0.04, 0.06, 0.4, 1, 4 and 8 µg Fe / ml) and harvested during mid log phase of growth to identify the optimal concentration of iron required for growth. The values were calculated from three separate experiments.

(a)



(b)



(c)

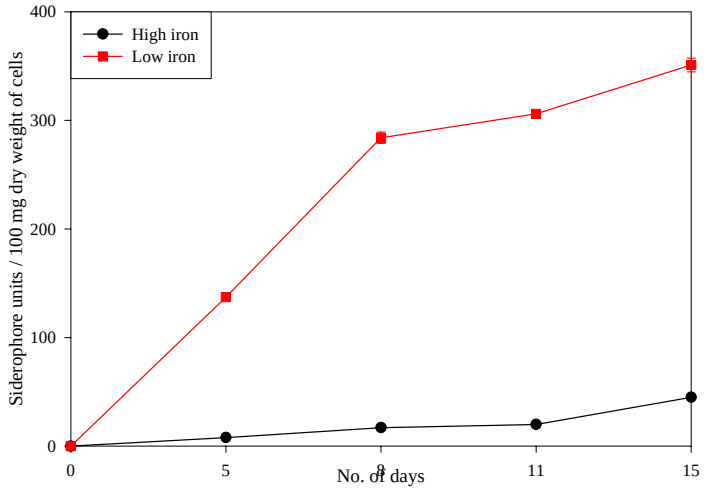


Fig. 19 - Time course expression of Mb and CMb in *M. kansasii:* (a) growth, (b) CMb and (c) Mb. The cells were grown under high (8 µg Fe / ml) and low iron (0.02 µg Fe / ml) and harvested on days 5, 8, 11 and 15 respectively. The values were calculated from three separate experiments.

3.2. Extraction, purification and structural elucidation of carboxymycobactin and mycobactin from *M. kansasii*

3.2.1. *Mycobacterium kansasii* expresses chloroform soluble carboxymycobactin

Mycobactin was easily extracted by overnight incubation with ethanol, followed by addition of ferric chloride and extraction of the wine-red ferric mycobactin into the chloroform layer. The CMb was also found to be chloroform extractable. Upon addition of chloroform, the ferric form of the CMb in the culture filtrate partitioned into the chloroform layer, as seen visually by the reddish colour of the ferric carboxymycobactins. The absorption maximum was at 450 nm.

3.2.2. IR and NMR analysis

Mycobactin and carboxymycobactin are structurally similar except for the R_1 side chain, which is a long chain fatty acid in the former and is a COOH or simple ester in CMb. Functional group analysis of CMb by IR spectroscopy (Fig. 20) confirmed that the different functional groups present in CMb were similar to the mycobactins (Table 17). Hence, NMR analysis (Fig. 21) was performed to analyse various hydrogen atoms and their positions (Table 18).

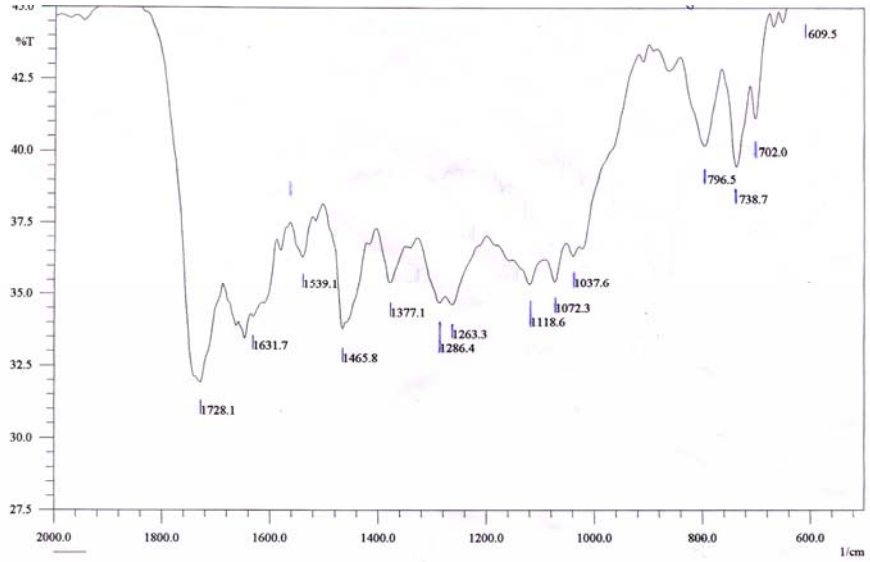


Fig. 20 - IR analysis of carboxymycobactin from *M. kansasii.* The NEAT infrared spectra for 2 mg of CMb were recorded on SHIMADZU FT-IR spectrophotometer model 8300 with polystyrene as reference.

Table 17 - IR analysis of carboxymycobactin from *M. kansasii*

Position (1 / cm)	Intensity (% T)	Groups present
702	41.167	Five adjacent hydrogen atoms present in a aromatic ring
738	39.491	
796	40.199	Three adjacent hydrogen atoms
1072	35.579	Five membered cyclic ethers
1118	35.459	
1263	34.469	CO stretching and OH deformation
1286	34.730	
1377	35.404	Phenols
1465	33.792	$\rangle CH_2$
1539	36.433	May be β– ketones
1631	34.228	Non-conjugated C=C stretching
1730	31.933	CO-NH-CO

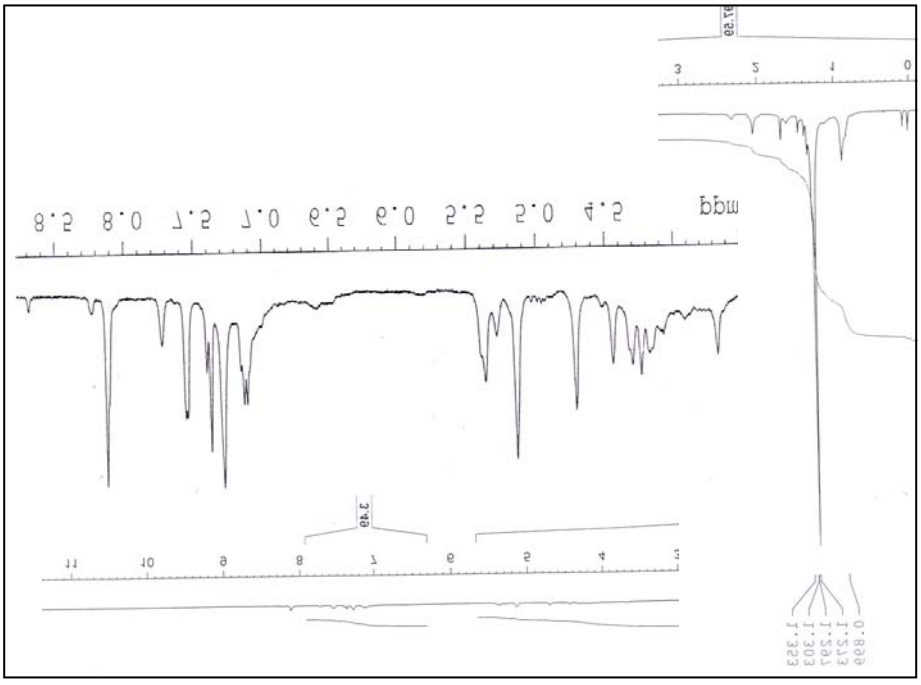


Fig. 21 - ^{1}H NMR analysis of carboxymycobactin. ^{1}H NMR 400MHz spectra were recorded on Bruker Avance-400 spectrometer with 2 mg of CMb, using chloroform-d as a solvent and TMS as reference (δ = 0 ppm) at 303 K. The chemical shifts are expressed in δ downfield from the signal of internal TMS.

Table 18 - NMR assignments

Particulars	**^{1}H (Coupling Hz)**	**Particulars**	**^{1}H (Coupling Hz)**
Cyclo lysine		**3-OH butanoic acid**	
α	4.4		
ββ’	1.69, 1.39	2’	2.5 (7.5)
δ,δ’	1.66, 1.42	3	5.10
NH	8.08	4	(d 6.5)
Lysine		R_1 group	
α	4.13	CH_2 ‘s	1.2 –1.5
ββ’	1.62, 1.71	CH_2	2. 16
NH	8.66		
Oxazoline		**Salicylic acid**	
α	5.0	5	6.92
β	4.65	4	7.45
β’	4.50	6	7.63

3.2.3. HPLC purification

Using reverse phase HPLC system, the hydrophobic Mbs were separated with a gradient of methanol, acetonitrile and water, while the relatively more hydrophilic CMb were separated with acetonitrile and water gradient. Both the siderophores separated out as group of compounds with maximal absorbance at 450 nm. Mycobactins resolved into nine separate compounds (Fig. 22) while CMb, with retention time from 19 minutes to 25 minutes separated into 15 peaks (Fig. 23) as also observed for other mycobacterial siderophores.

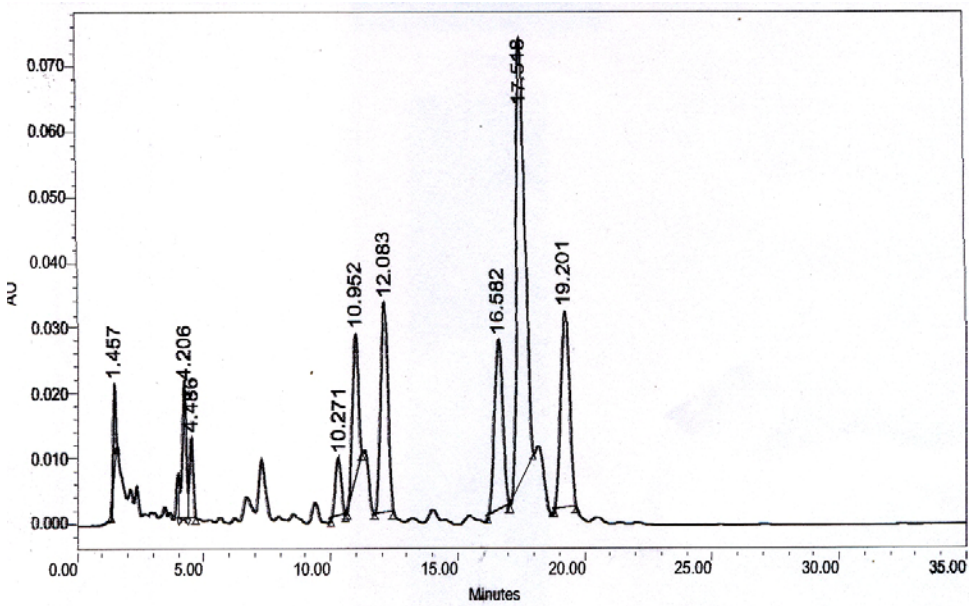


Fig. 22 - HPLC purification of mycobactins. The extract of Mb as subjected to reverse phase HPLC on a Water's C18 column (150 mm x 4.6 mm id, 5 μm) with a gradient of 0 to 100 % with buffer A (0.09 % formic acid/ 90% acetonitrile) and buffer B (0.1 % formic acid in 60% acetonitrile and 40% methanol at a flow rate of 1mL / minute for 35 minutes. Eluate was monitored continuously at 450 nm.

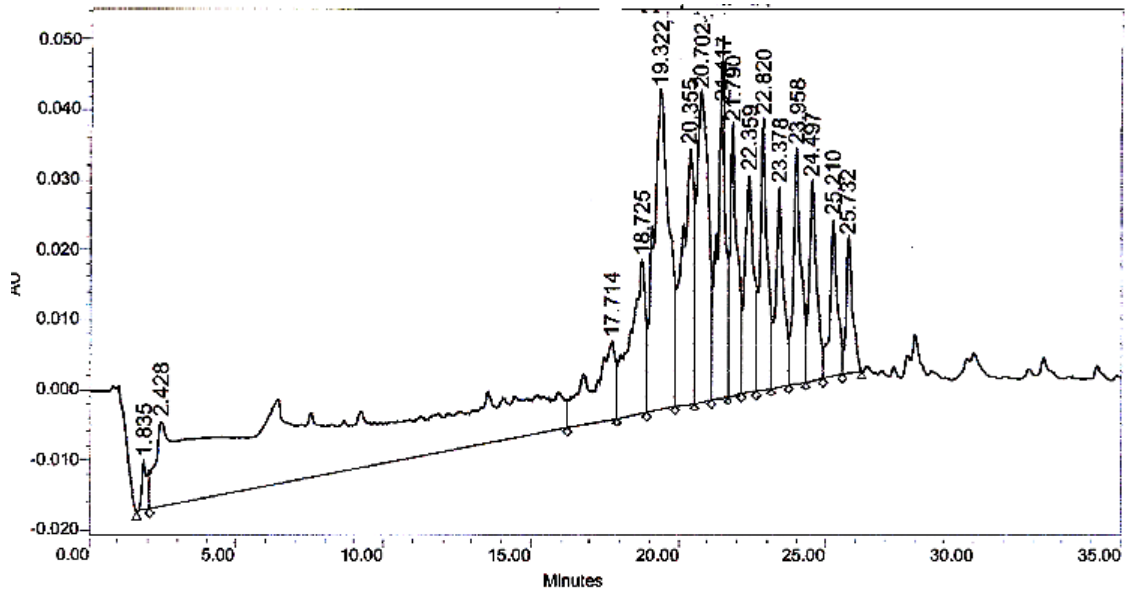


Fig. 23 - HPLC purification of carboxymycobactins. The extract of CMb as subjected to reverse phase HPLC on a Water's C18 column (150 mm x 4.6 mm id, 5 μm) with a gradient of 0 to 100 % with buffer A (0.1 % formic acid) and buffer B (0.09 % formic acid in 90 % acetonitrile) at a flow rate of 1mL / minute for 35 minutes. Eluate was monitored continuously at 450 nm.

3.2.4. LC-MS Analysis

As HPLC could not resolve the Mb and CMb into single purified fractions, LC-MS was used done as an effective and a sensitive system that helped not only in the identification of closely related compounds but which also detected the less abundant species. The HPLC / ESI-MS analysis showed that Mb and CMb comprised of a large family of molecules ranging from a mass of 810 to 910 Da (Fig. 24) and 758 to 884 Da (Fig. 25) respectively.

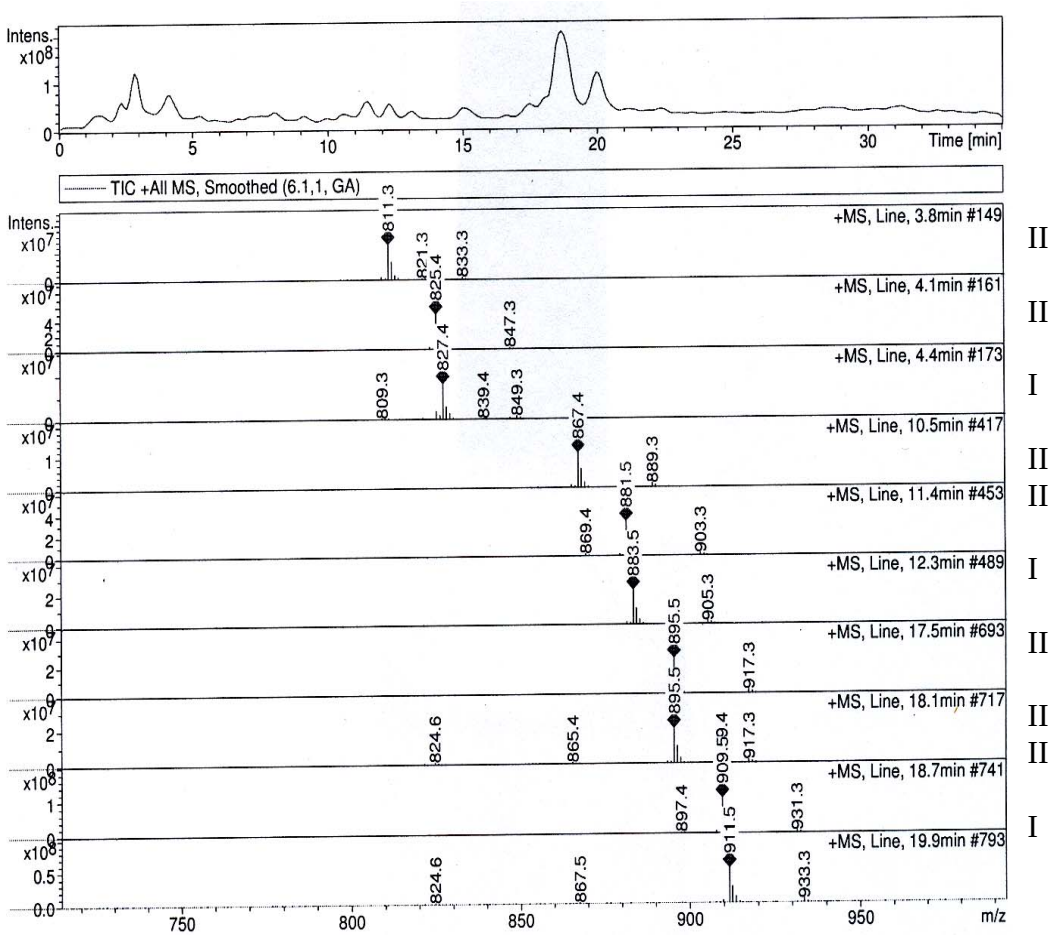


Fig. 24 - LC-ESI / MS analysis of mycobactins: The data suggests that Mb based upon its long alkyl chains exists in two series denoted as group I and group II having saturated and unsaturated alkyl chain respectively. Mycobactins (Mol. Wt. 826, 882, 910) have saturated alkyl chains, while Mb (Mol. Wt. 810, 824, 866, 880, 894, and 908) has unsaturation in their R1 side chain.

The higher molecular weights of the Mb were attributed to its large alkyl chain in the R_1 group. HPLC / ESI - MS analysis also suggested the presence of two groups of molecules, differing from each other not only by multiples of 14 Da ($\Delta M = CH_2$), but also by 2 Da (CH_2-CH_2 versus CH_2=CH_2) indicating the difference in saturation in the side chain. Thus, on the basis of the molecular mass analysis alone it can be

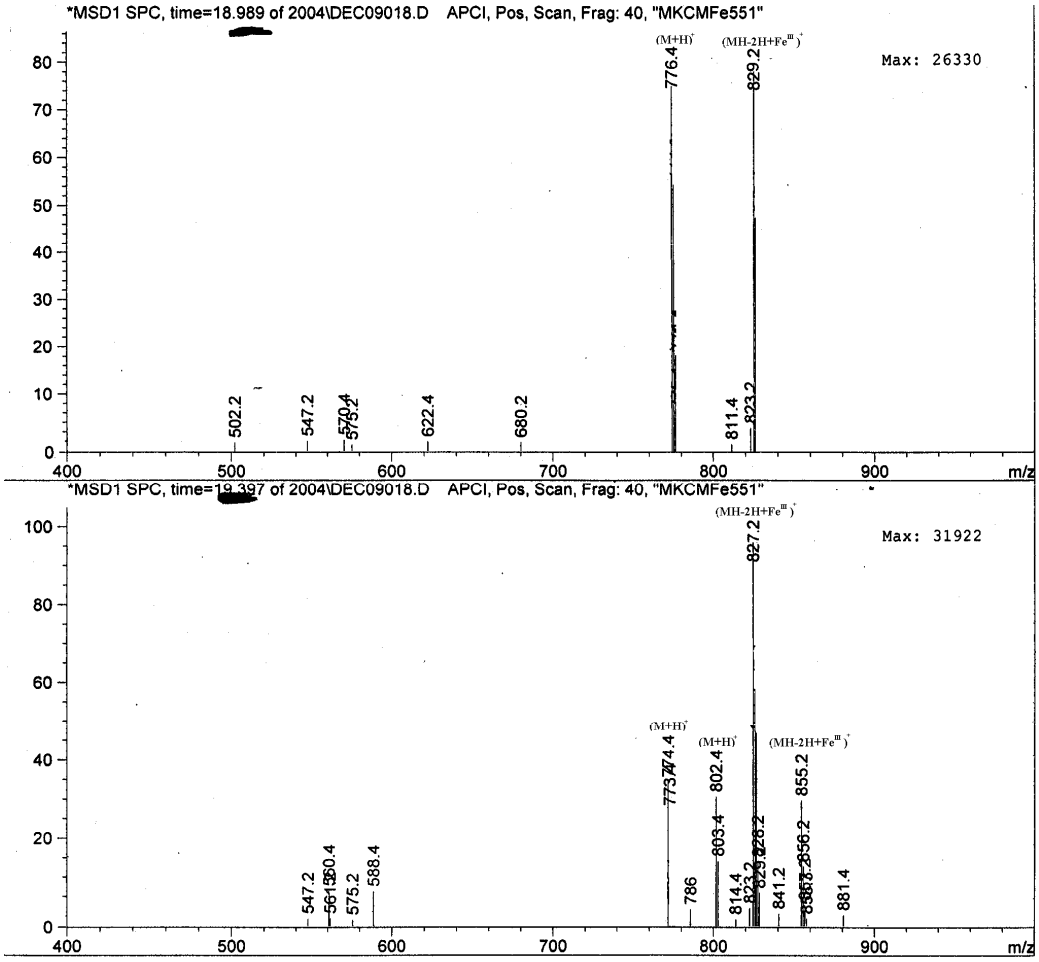


Fig. 25 - LC-ESI / MS analysis of carboxymycobactins: The analysis suggests that CMb for molecular mass (MH-2H+FeIII) 855, 827 and 829 Da. The data also suggests the CMb can be grouped based on R1 alkyl chain saturation. Carboxymycobactin with masses (M-2H+Fe3+) 758, 772, 786, 800, 814, 828, 842, 856, 870 and 884 Ad belongs to the saturated series, while the other series comprising the unsaturated CMb with masses (M-2H+Fe3+) of 798, 812, 826, 840, 854, 868 and 882 respectively.

predicted that both Mb and CMb belong to two series; one series comprising the saturated Mb with masses (M-2H + Fe^{3+}) of 826, 882 and 910 (Fig. 24) and the CMb with masses (M-2H+Fe^{3+}) of 758, 772, 786, 800, 814, 828, 842, 856, 870 and 884 Da series (Fig. 25) and the other series comprising the unsaturated Mb with masses (M-2H + Fe^{3+}) of 810, 824, 866, 880, 894 and 908 (Fig. 24) and CMb with masses (M-2H + Fe^{3+}) of 798, 812, 826, 840, 854, 868 and 882 respectively.

3.2.5. Tandem mass spectrometric analysis of carboxymycobactin

HPLC-MS / MS analysis was performed on CMb and the fragmentation pattern of CMb from *M. kansasii* was analysed. The fragment ions that resulted from the cleavage about the amide and ester bonds were assigned as A to F (Fig. 26). Both Mb and CMb had similar fragmentation patterns, differing only in the F moiety fragment ions. The fragments include a (cyclical lysine), B (β-hydrocyanic), C (lysine), D (oxazoline), E (salicylic acid) and / or F (acyl alkyl acid or methyl ester) moieties. In the fragment E, the fragmentation ions at 235 Da (DE + NH + 2H), 362 Da (CDE + H) and 592 Da (ABCDE + 2H) indicate the methyl substitution in the salicylate moiety. The presence of a cyclical threonine (R_3= CH_3) is indicated by the fragmentation ions at 235 Da (DE + NH + 2H), 362 Da (CDE + H), 592 Da (ABCDE + 2H), and 448 Da (BCDE + H). The fragmentation ions (Fig. 28) at 448 Da (BCDE + H), 557 Da (ABCF + 2H), 592 Da (ABCDE + 2H) and 602 Da (BCDEF) show that R_4 = CH_3 and R_5 = H.

From MS / MS analysis, CMb can be characterised as acids or methyl esters based on their termination in R_1 side chain. Carboxymycobactins terminating as carboxylic acids showed a characteristic loss of 44 Da $(MH\text{-}CO_2)^+$, while those terminating with methyl esters had a loss of 32 Da ($MH\text{-}CH_3OH$) and a loss of 59 Da $(MH\text{-}COOCH_3)^+$. The difference in the mass by multiples of 14 Da has shown to be the different numbers of CH_2 groups in the R_1 side chain (Table 19 & 20).

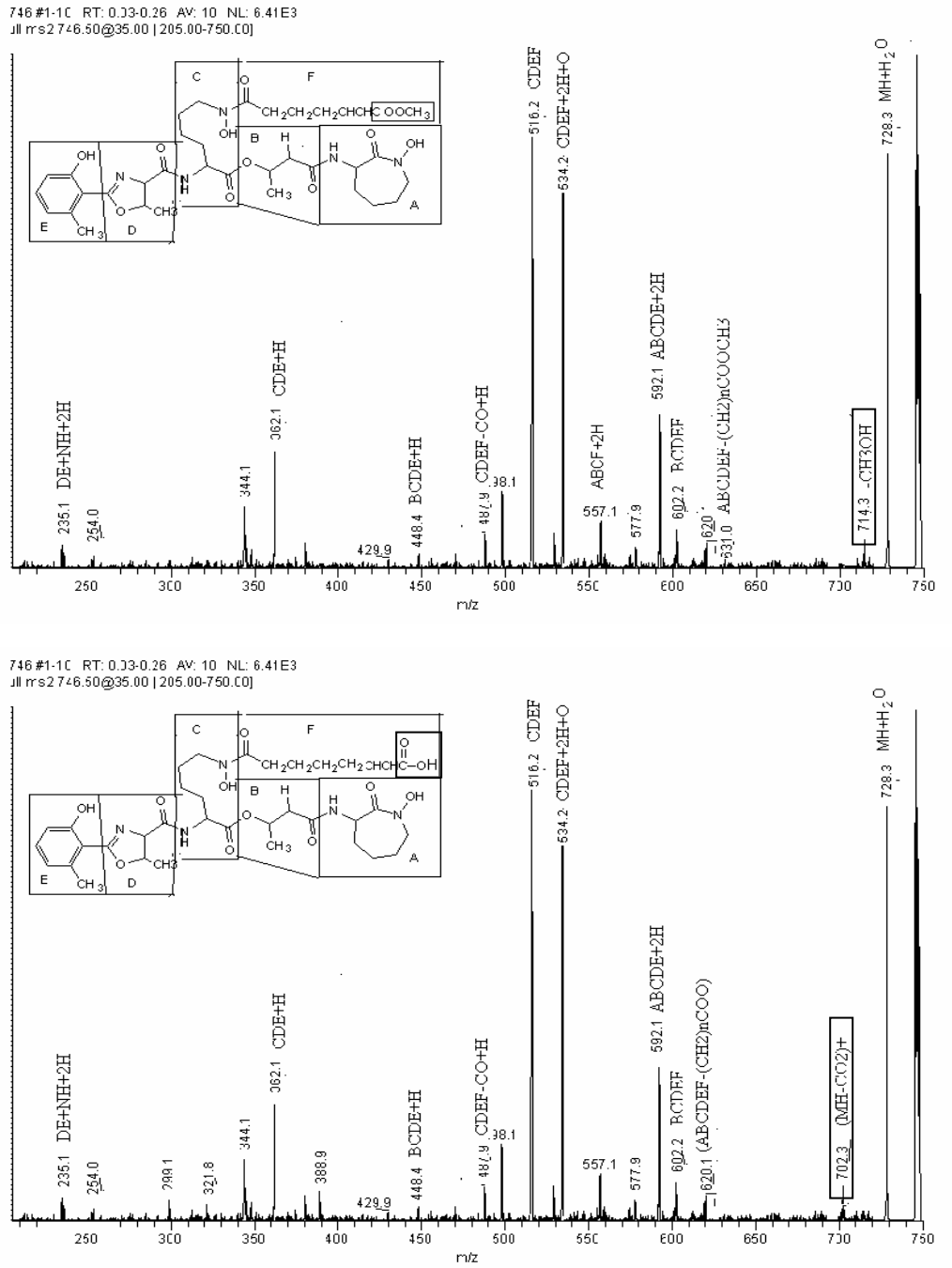


Fig. 28 - LC-MS/MS analysis of desferri CMb M+H⁺ at m/z746. The fragment ions resulted from cleavage about the amide and ester bonds were assigned as A to F where A and C (N-hydroxyl lysine), B (β-hydroxybutyrate), D (oxazoline), E (salicylic acid) and/or F (acyl acid or methyl ester) moieties.

Table 19 - Carboxymycobactins with saturated R_1 alkyl chain

Desferri CMb (MW)	Ferri CMb (MW)	R_1 $(CH_2)_n$	Acid or Methyl ester	R_2	R_3	R_4	R_5
705	**758**	3	Acid	CH_3	CH_3	CH_3	H
719	**772**	4 or 3	Both	CH_3	CH_3	CH_3	H
733	**786**	5 or 4	Both	CH_3	CH_3	CH_3	H
747	**800**	6	Acid	CH_3	CH_3	CH_3	H
761	**814**	6	Methyl ester	CH_3	CH_3	CH_3	H
775	**828**	8	Acid	CH_3	CH_3	CH_3	H

Table 20 - Carboxymycobactins with unsaturated R_1 alkyl chain

Desferri CMb (MW)	Ferri CMb (MW)	R_1 $(CH_2)_n$	Acid or Methyl ester	R_2	R_3	R_4	R_5
745	**798**	3	Methyl ester	CH_3	CH_3	CH_3	H
759	**812**	5	Acid	CH_3	CH_3	CH_3	H
773	**826**	5	Methyl ester	CH_3	CH_3	CH_3	H
787	**840**	6	Methyl ester	CH_3	CH_3	CH_3	H

3.3. Iron regulated envelope proteins as analysed by SDS-PAGE

3.3.1. Identification of IREPs in *M. kansasii*

The solubilisation of the cell wall and cell membrane by SDS did not give clear separation due to the possible interference by the cell envelope lipids. Thus, the method of Sinha and his group (2004) employing Triton X-114 was used to solubilise the wall and membrane pellet followed by separation by SDS-PAGE on a 5 - 20 % gradient gel.

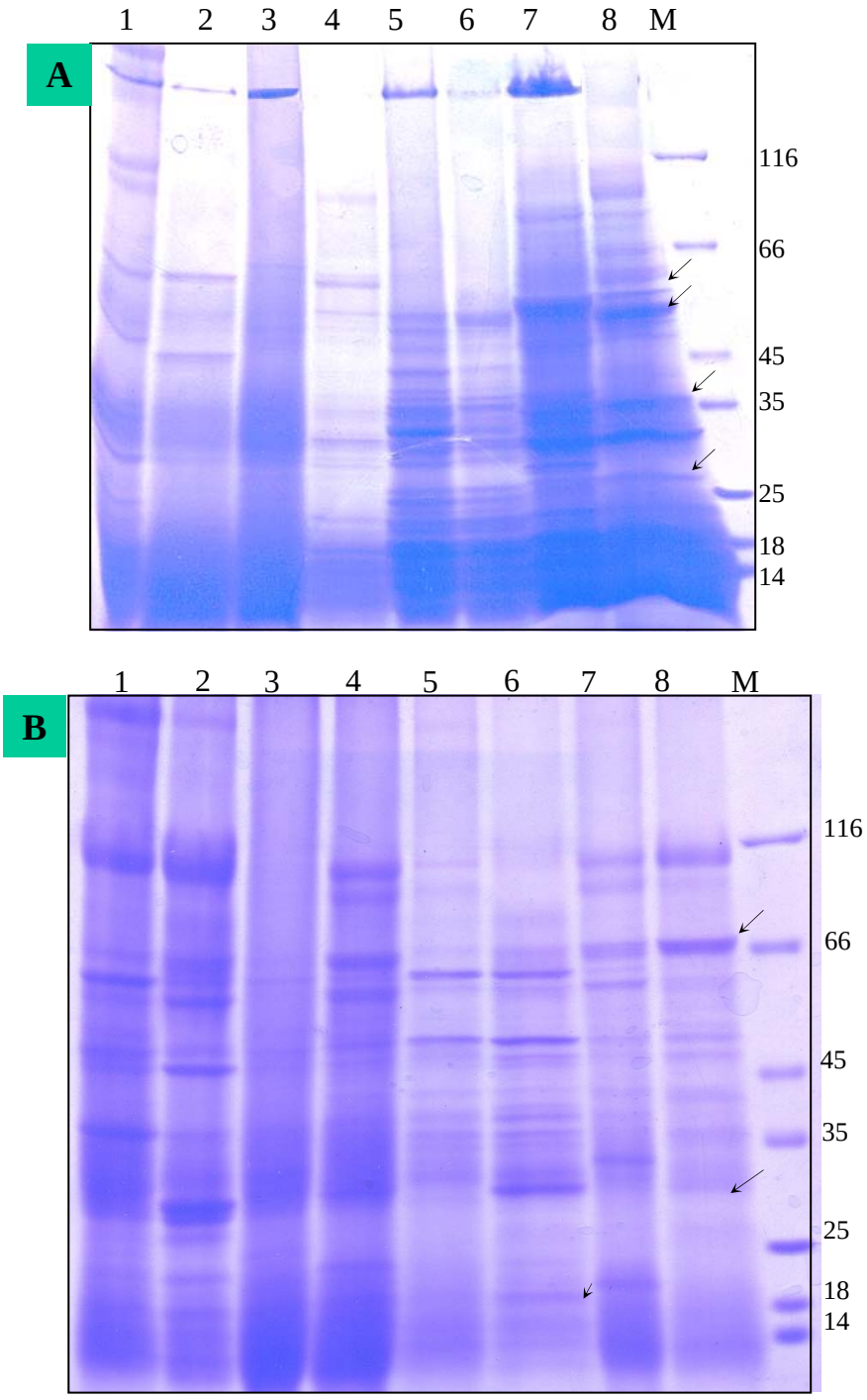


Fig. 27 - SDS-PAGE profile of cell membrane proteins - Time course study. The organisms were grown in high (8 µg Fe / mL of medium) and low (0.02 µg Fe / mL of medium) iron medium. Panel A and B represents aqueous and detergent phase proteins, where lanes 1, 3, 5, 7 represent high iron cells and 2, 4, 6, 8 represent low iron cells harvested on days 5, 8, 11 and 15 respectively.

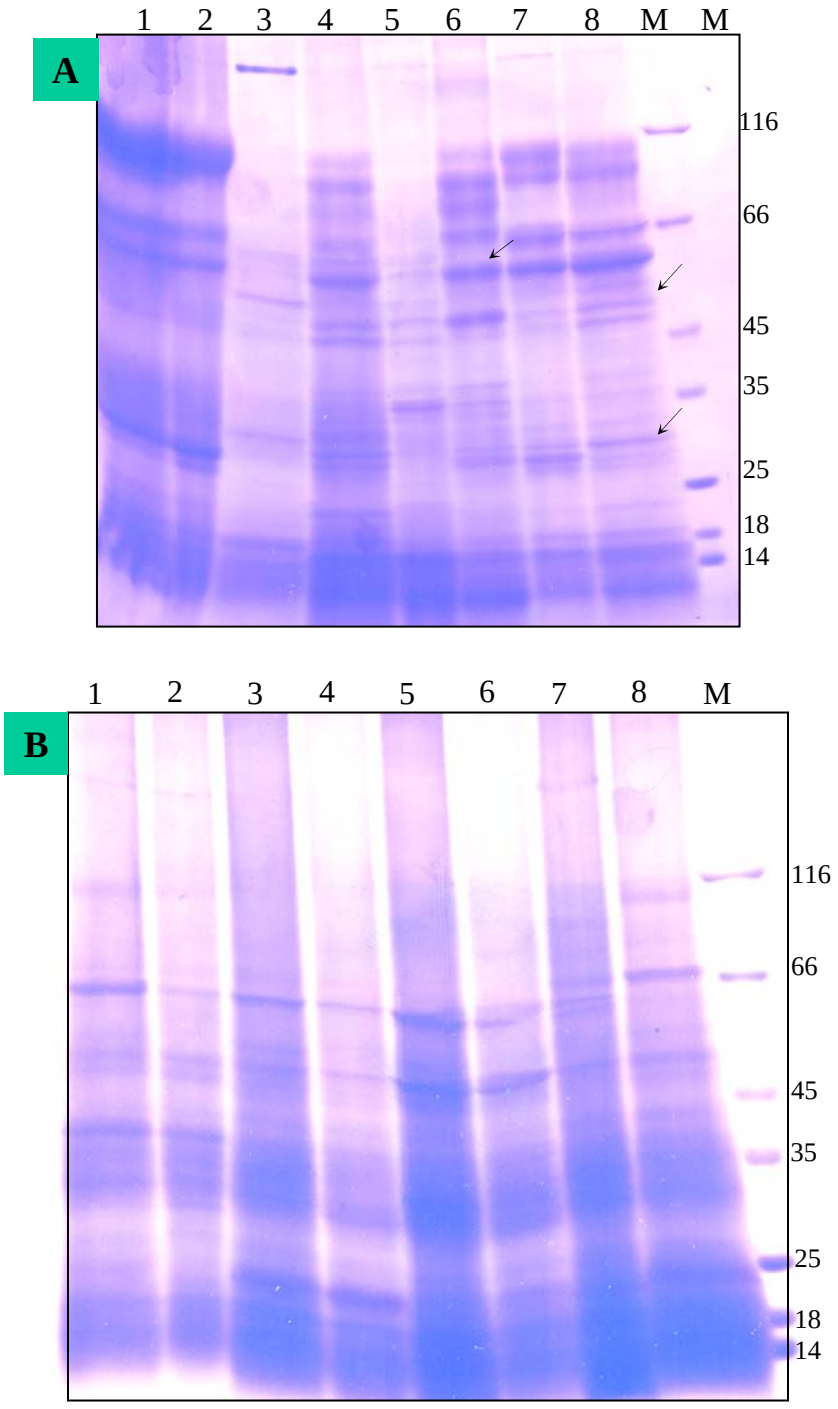


Fig. 28 - SDS-PAGE profile of cell wall proteins- Time course study. The organisms were grown in high (8 µg Fe / mL of medium) and low (0.02 µg Fe / mL of medium) iron medium. Panel A and B represents aqueous and detergent phase proteins, where lanes 1, 3, 5, 7 represent high iron cells and 2, 4, 6, 8 represent low iron cells harvested on days 5, 8, 11 and 15 respectively.

The organisms were grown in high (8 μg Fe / mL) and low (0.02 μg Fe / mL) iron medium, harvested on days 5, 8, 11 and 15 respectively (as done for estimating mycobactin and carboxymycobactin shown above) and then processed for the analysis of cell wall and membrane proteins for the expression of iron-regulated envelope proteins. SDS-PAGE analysis of aqueous cell membrane on a 5-20% gradient gel (Fig. 27a) showed significant expression of two proteins of molecular mass 52 and 35 kDa in low (0.02 μg Fe / mL of medium) iron medium. The expression of higher molecular weight protein of molecular mass 200 kDa after 5 days of growth was seen in both high and low iron cells, indicating the iron replete conditions in both the cultures. Subsequently, it was expressed only in high iron cells and represented as a marker for high iron conditions. A 31 kDa protein was expressed in high iron cells of 8 day old culture, while the 52 kDa protein was seen in low iron cells in all stages of growth except in the early phase (5 days) of growth. In the cell membrane detergent fraction (Fig. 27b), significant expression of 66, 54 and 29 kDa proteins were seen in low iron cells, of which the 66 kDa protein could not be seen in cells harvested on day 11. The expression of a 45 kDa protein was found to be up regulated only during the early phase.

Cell wall aqueous fraction (Fig. 28a) showed the expression of three proteins of approximate molecular mass 65, 52 and 29 kDa respectively in low iron organisms. These proteins were seen during late stationary phase (15 days) of high iron cells also, indicating the onset of iron deficiency during stationary phase of growth. No significant differences were observed in detergent phase cell wall proteins (Fig. 28b).

3.3.2. Expression of IREPs with varying levels of iron

To study the effect of varying concentration of iron on the expression of proteins, *M. kansasii* was grown in media with iron added at 0.02, 0.04, 0.06, 0.4, 1, 4 and 8 μg Fe / ml of medium respectively. Cell wall and cell membrane fractions were prepared as above and subjected to gradient SDS-PAGE analysis. Analysis of aqueous membrane proteins (Fig. 29a) showed increased expression of the 52 kDa protein in cells grown with iron added at 0.02, 0.04, 0.06, 0.4 μg Fe / ml of medium while the 35

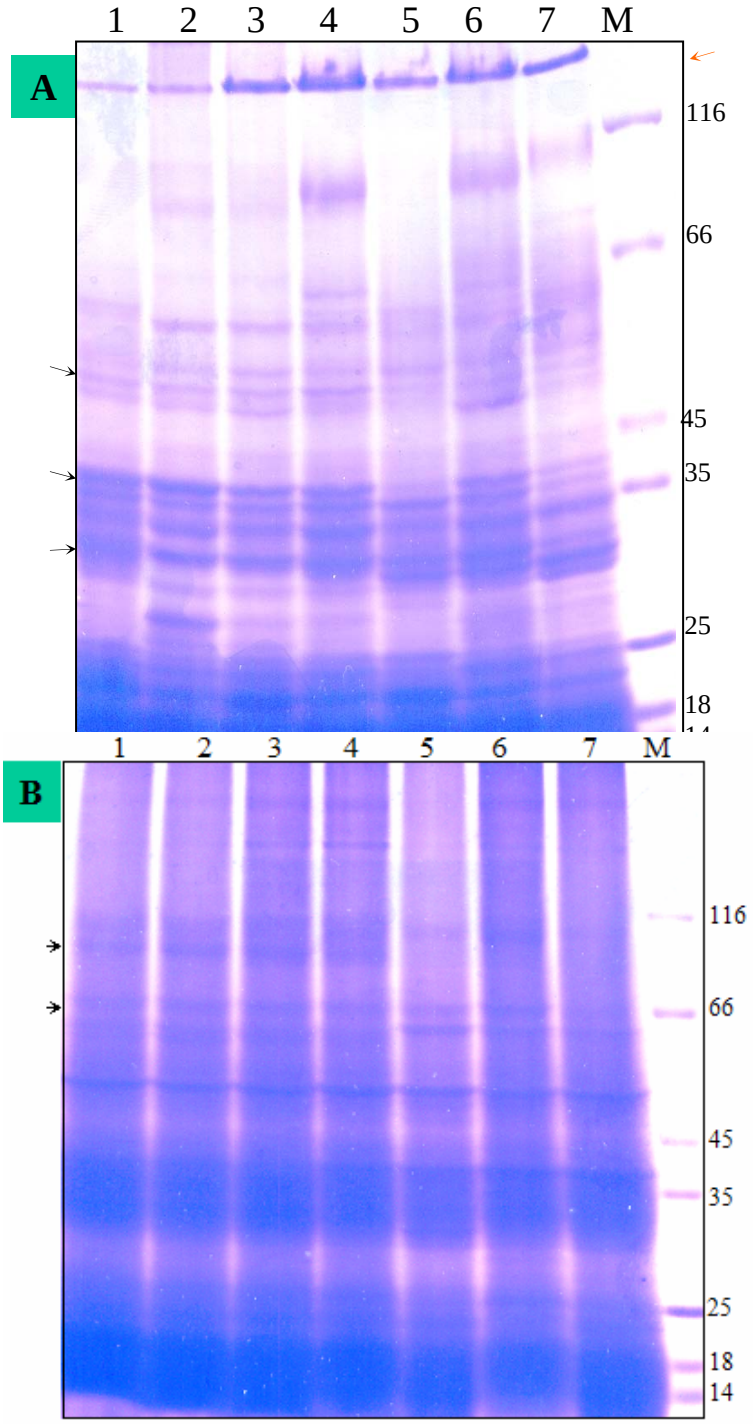


Fig. 29 - SDS-PAGE profile of cell membrane proteins of *M. kansasii* grown under different iron concentration. The organisms were grown in different iron concentrations ranging from 0.02 to 8 µg Fe / mL of medium. Panel A and B represents the aqueous and detergent phase proteins. Lane 1 to 7 represents 0.02, 0.04, 0.06, 0.4, 1, 4 and 8 µg Fe / mL of medium respectively.

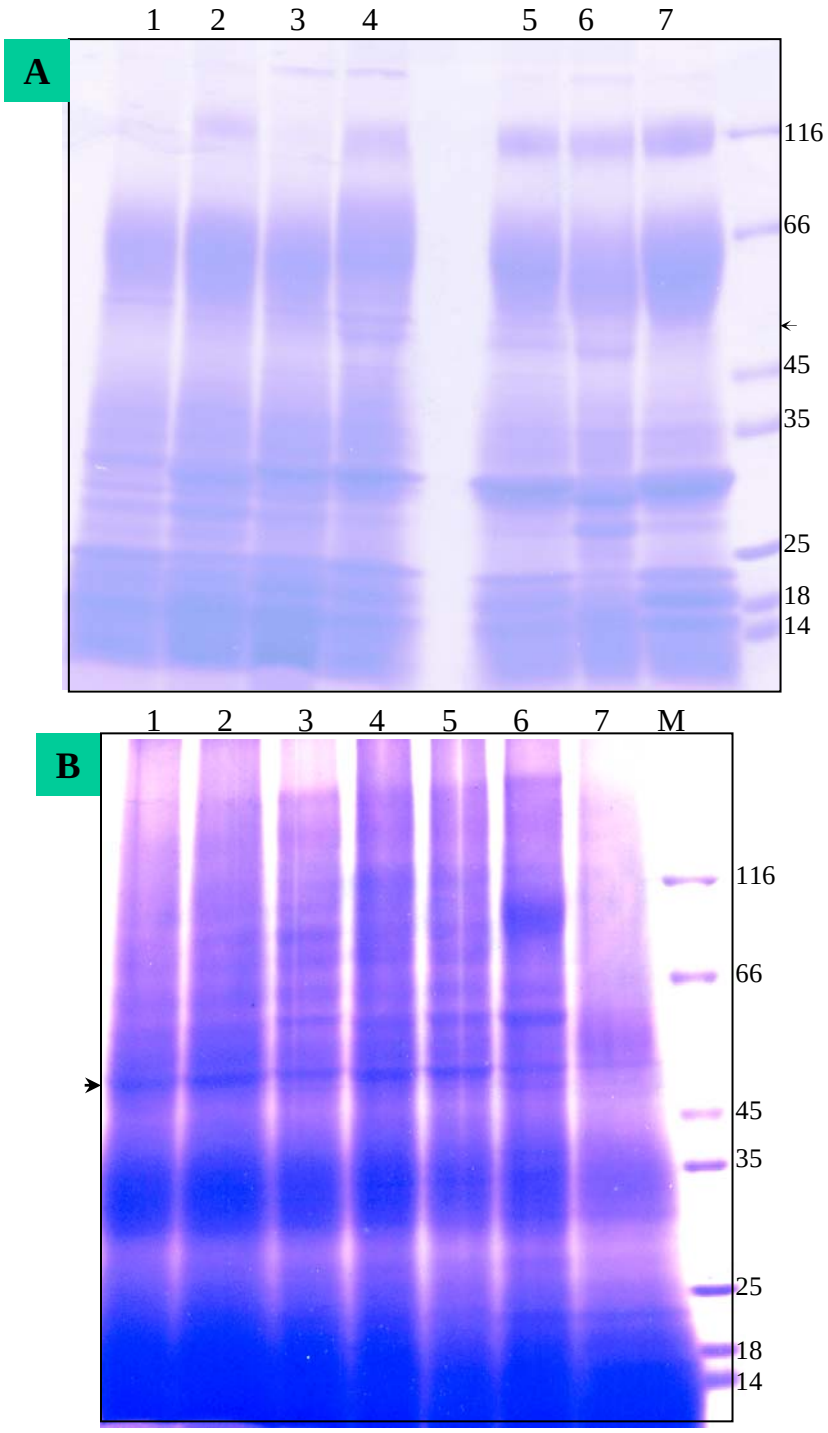


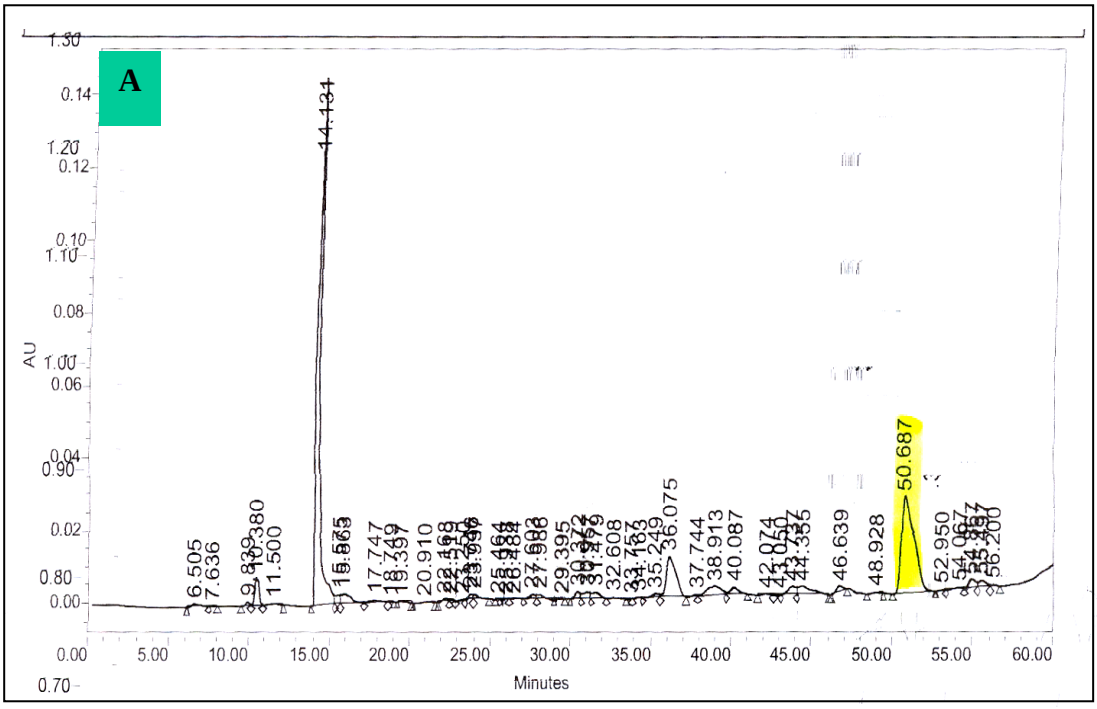
Fig. 30 - SDS-PAGE profile of cell wall proteins of *M. kansasii* grown under different iron concentrations. The organisms were grown on different iron concentrations ranging from 0.02 to 8 μg Fe / mL of medium. Panel A and B represents the aqueous and detergent phase proteins. Lane 1 to 7 represents 0.02, 0.04, 0.06, 0.4, 1, 4 and 8 μg Fe / mL of medium respectively.

kDa band was observed in all the concentration except the 8 μg Fe / ml of medium. In the cell membrane detergent pellet (Fig. 29b), there was higher expression of a protein of approximate molecular mass 66 kDa from 0.02 to 4 μg of Fe / ml of medium along with a 112 kDa protein.

There was no significant difference in cell wall aqueous proteins (Fig. 30a) while cell wall detergent fraction (Fig. 30b) showed the 52 kDa. The gel slices were collected and used for sequencing analysis by mass spectrometry.

3.3.3. Isolation of IREPs by HPLC

HPLC analysis was done only with the cell membrane proteins as the cell wall proteins were difficult to solubilise in acetonitrile. The membrane proteins from both high (8 μg Fe / mL) and low (0.02 μg Fe / mL) medium were filtered through 0.22 μm filters and separated on a Waters C_{18} column. In the sample from low iron cells, a major peak at 40 minute was seen which was absent in the corresponding sample from high iron cells (Fig. 31). This fraction, called as MK40' was collected, dried using Speed vac and used for interaction with mycobactin and carboxymycobactin as well as for identification by mass spectrometry.



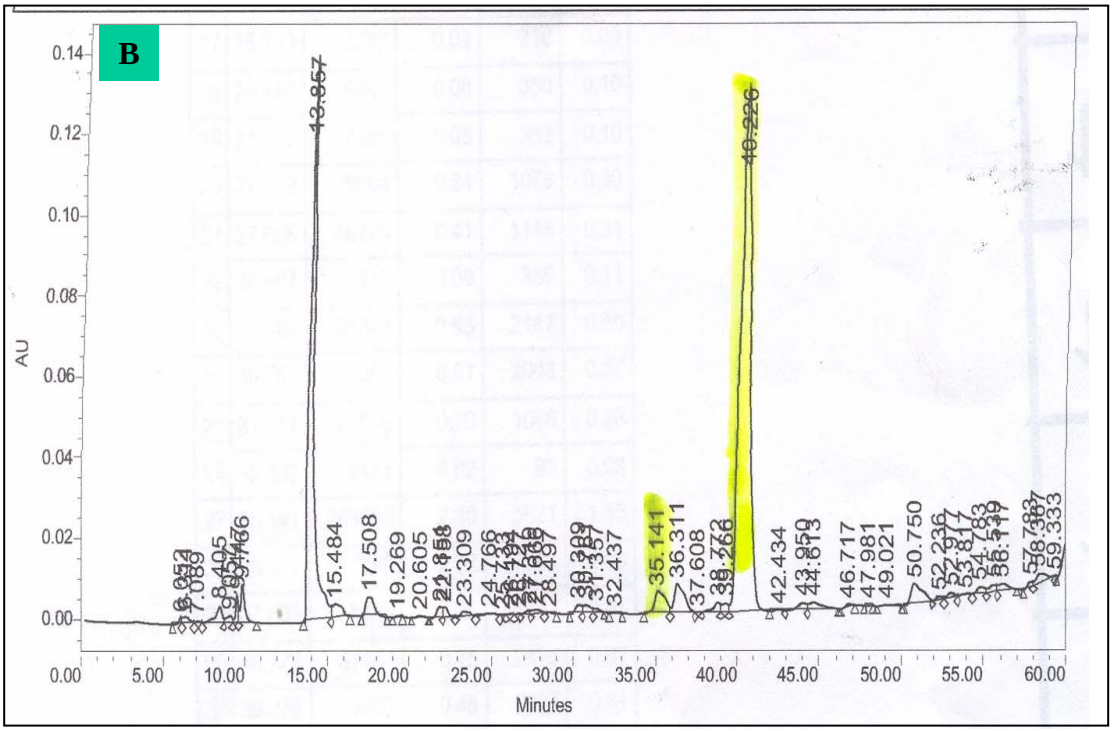


Fig. 31 - HPLC purification of membrane proteins. Panel A and B represents cell membrane high and low iron organisms. The protein was subjected to reverse phase HPLC on a Water's C18 column (150 mm x 4.6 mm id, 5 μm) with a gradient of 0 to 100 % buffer A (0.1% TFA in 90% acetonitrile) and buffer B (0.1% TFA in 10% acetonitrile) at a flow rate of 0.5 mL / min for 60 min. The eluate was monitored continuously at 280 nm.

3.3.4 Spectrofluorimetric studies with the HPLC purified fraction MK40'

Fluorescent spectroscopy was used to study the interaction of MK40' with the two siderophores (Fig. 32-33) from *M. kansasii*. The binding affinity of siderophores with this protein was calculated. K_b values for mycobactin and carboxymycobactin were calculated as 6.06×10^5 and 3.397×10^5 from the ratio of the slope to intercept (Fig. 34).

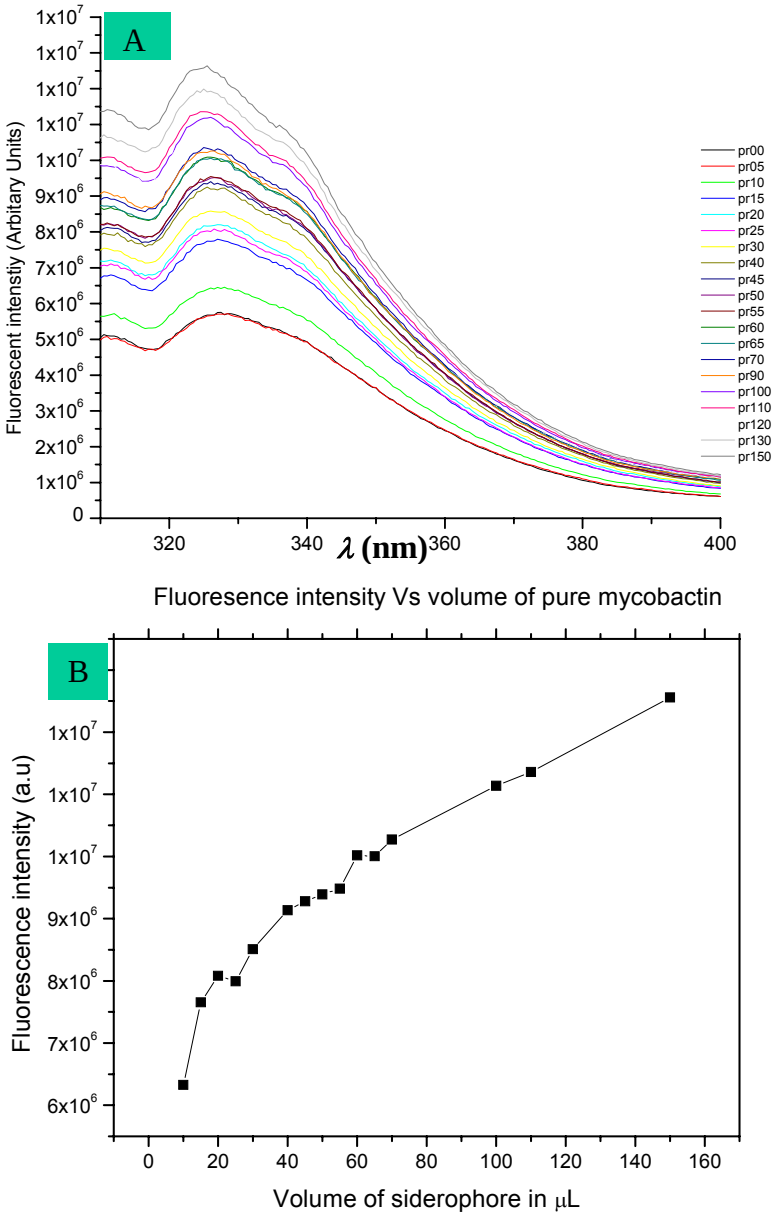


Fig. 32 - Spectrofluorimetric analysis of binding of MK40' (IREP) with mycobactin. Panel A represents the fluorescent spectrum and Panel B represents the plot of the change in fluorescent intensity with increasing addition of siderophores. The MK40' was subjected to excitation at 280 nm and the emission spectrum was recorded from 310 nm to 400 nm. Mycobactin was added at concentration ranging from 2.5 µM to 30 µM and the emission spectrum recorded after each addition. The data obtained was fitted to Wolf's equation to calculate the binding constant.

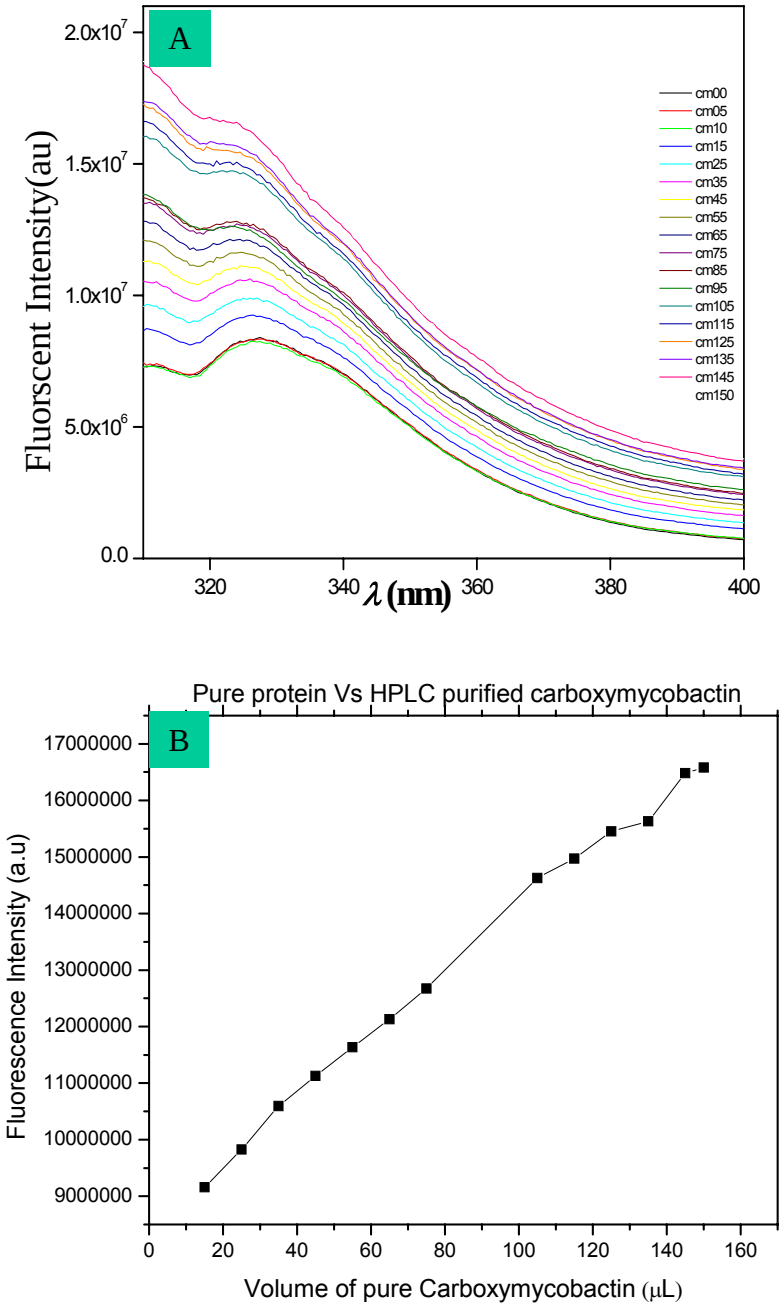


Fig. 33 - Spectrofluorimetric analysis of binding of MK40' (IREP) with carboxymycobactin. Panel A represents the fluorescent spectrum and Panel B represents the plot between the change in fluorescent intensity with addition of siderophores. The MK40' was subjected to excitation at 280 nm and the emission spectrum was recorded from 310 nm to 400 nm. Carboxymycobactin was added at concentration ranging from 2.5 µM to 30 µM and the emission spectrum recorded after each addition. The data obtained was fitted to Wolf's equation to calculate the binding constant.

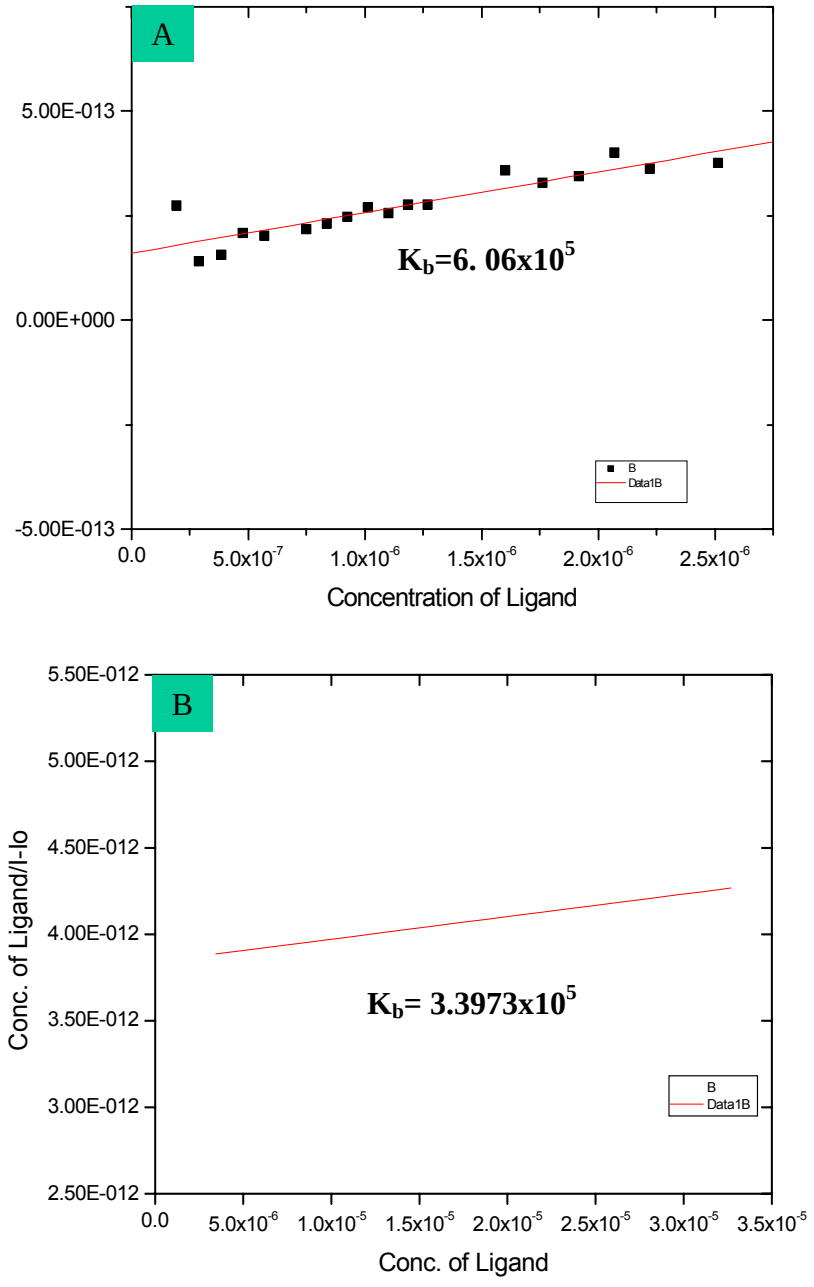


Fig. 34 - Determination of binding constants for mycobactin (Panel A) and carboxymycobactin (Panel B). The values were calculated using Wolfe's equation.

3.3.5. Sequencing of IREPs by mass spectrometry

The MK40' HPLC purified protein was identified as a transcriptional regulatory protein (Rv0078) with unknown function having molecular mass of 22 kDa.

The IREPs resolved by SDS-PAGE were analysed by mass spectrometry. The 52 kDa cell membrane protein was found to contain three proteins the DeaD (Rv1253), ponA2 (Rv3682) and conserved hypothetical protein (Rv1779c). The 52 kDa cell wall protein was identified as CysN, a bi-functional enzyme acting as the sole sulfate transporter.

3.4. Discussion

In this study, we analyzed the iron acquisition machinery of *M. kansasii* and studied the effect of iron levels on mycobactin, carboxymycobactin and the iron-regulated proteins. We show that *M. kansasii*, though a fast grower produces the chloroform soluble carboxymycobactins, similar to slow-growing pathogenic mycobacteria. The structural elucidation of these iron-chelating molecules is done and the similarities and differences with those expressed by other mycobacterial species are discussed. The rationale for choosing this organism and the potential of the components of the iron acquisition machinery as novel drug targets is discussed.

Mycobacterium kansasii is one of the important and frequently encountered non-tuberculous mycobacterial pathogens in humans, in whom it produces a chronic, progressive and caviatary lung disease similar to tuberculosis (Bloch *et al.*, 1998; Bittner *et al.*, 1996). *Mycobacterium kansasii* produces disseminated infections in patients infected with HIV virus (Canueto-Quintero *et al.*, 2003). Optimal treatment of *M. kansasii* in non-HIV-infected patients includes isoniazid, rifampicin and ethambutol. Patients with AIDS and *M. kansasii* infection are at a disadvantage with this traditional regimen, because rifampicin greatly accelerates hepatic metabolism of HIV protease inhibitors, rendering them ineffective (Graybil & Bocanegra, 2001). Due to the increase in the rate of infection, rise in multi-drug resistant cases and the

disadvantage of using rifampicin, there is an urgent need to identify novel targets and develop new drugs, not only for *M. kansasii* but also for all pathogenic mycobacteria.

One of the limiting factors in the development of new drugs against pathogenic mycobacteria is the inability of drugs to diffuse across the lipid-rich cell envelope and reach its intended target. The problem is severe in case of mycobacteria, which has its unique pattern of cell wall, coated with a polymeric complex of peptidoglycans covalently attached to a layer of arabinogalactans terminating with a densely packed array of mycolic acids, which in turn forms a second bilayer by associating itself with the glycolipids (Besra & Chatterjee, 1994). Hence, drugs that will target molecules normally used by these organisms and which are vital for their survival *in vivo* need to be developed (Vergne *et al.,* 2000). To identify such drug targets, it is necessary to understand the transport of essential nutrients, and gain a better insight into the host-pathogen interactions at various stages of infection.

Iron acquisition and transport is one of the important aspects of host-pathogen interactions in pathogenic mycobacteria. Among all the nutrients required, iron alone is virtually insoluble in water at pH values around neutrality. It has a solubility of approximately 1.4×10^{-9} M at pH 7. The solubility of Fe^{3+} increases by 10^{-3} M for every unit that the pH drops so that, for example, at pH 5 the solubility of Fe^{3+} is now 10^{-3} M, making it biologically available; however, this concentration will be further lowered within the mammalian system, where iron-withholding proteins like transferrin and lactoferrin come into play. *Mycobacterium tuberculosis* and other pathogenic mycobacteria, growing within the phagocytic vacuoles of the macrophages, where the pH is around 6.1 face a maximum concentration of free Fe^{3+} between 1 and 10 ng / ml, a value which is further lowered due to the lactoferrin in the macrophages that can hold the iron efficiently even at acid pH. Therefore, pathogenic mycobacteria, in order to grow and cause disease within a host, must compete against the host for its supply of iron. Douvas *et al*. (1993) showed that iron added to *M. avium* in macrophages enhanced bacterial growth and that serum and transferrin, but not holotransferrin (i.e. 98% replete with iron) or serum from which the transferrin had been removed, could inhibit bacterial multiplication.

The pathogenic mycobacteria succeed in acquiring the iron via the siderophores carboxymycobactin and mycobactin, their exact mechanisms of action still remaining unclear. The bacteriostatic effect of serum (via the transferrin chelating the iron) was shown by Barclay and Ratledge (1986) to be reversed by the addition of carboxymycobactin in their study with *M. avium* and *M. paratuberculosis*. Gobin and Horwitz (1996) confirmed the ability of carboxymycobactin to remove iron not only from transferrin but also, though more slowly, from ferritin to promote the growth of *M. tuberculosis*.

In this study, we show that *M. kansasii*, like other mycobacteria express mycobactin and carboxymycobactin in response to iron levels in the growth medium. When the added iron ranged between 0.02 to 4 μg Fe / ml, maximal expression of both these siderophores is seen when iron level is below 0.04 μg Fe / ml. The levels of both these siderophores decreased with increasing iron levels, the decline being sharper in mycobactin. The time-dependant expression of these molecules showed that the maximal expression was observed on the 8^{th} day of growth. Both the siderophores showed no significant increase in expression with time, though prolonged growth beyond two weeks showed that the initially iron replete cells became depleted of iron, as evident from the expression of carboxymycobactin.

A notable feature is that *M. kansasii* produces the chloroform soluble extracellular carboxymycobactins and not the peptidyl exochelins, thus resembling the pathogenic mycobacteria. Several of the pathogenic mycobacteria such as *M. tuberculosis* (Gobin *et al.,* 1995) and *M. avium* complex (Lane *et al.,* 1995) express chloroform-soluble CMb, while the non-pathogenic species generally expressed the peptidic water-soluble CMbs as in *M. smegmatis* and *M. neoaurum* (Sharman, 1995a & b). In *M. kansasii*, there are no earlier reports on the mycobactin and CMbs, though there is a mention that *M. kansasii* produces low levels of mycobactin (Snow, 1970). In this study, we show that this organism produces chorloform-soluble CMbs like the slow-growing pathogenic mycobacteria. The absorption maximum was at 450 nm unlike the peptidyl exochelins, which had absorption maximum at 430 nm (Ratledge & Ewing, 1996).

The structure of the two siderophores of *M. kansasii* was deduced by HPLC and mass spectrometry and compared with the mycobactins and carboxymycobactins reported for other mycobacterial species. All mycobactins (and the respective carboxymycobactins) from different mycobacterial species have the same core nucleus and essentially differ in the length and saturation in their R_1 side chain. Broadly, they can be classified in to four categories: 1) Saturated serine series, 2) Unsaturated serine series, 3) Saturated threonine series and 4) Unsaturated threonine series. HPLC separation of these molecules was influenced by the length of the side chain, extent of saturation (Barclay *et al.*, 1986) and on the presence of serine or the threonine series.

As was also observed by others (mycobactins from *M. tuberculosis* by Gobin *et al.*, 1995 and from *M. avium* by Lane *et al.*, 1998), several peaks, which differ slightly in molecular mass, are obtained for both siderophores. Barclay *et al.*, (1986) showed this was due to the fact that these molecules are closely related to each other and differ marginally from each other, especially in the levels of esterification of the COOH group. There were 9 peaks of mycobactin, with mass ranging from 810 to 910Da, while in carboxymycobactin there were 16 peaks, with mass ranging from 758 to 840 Da. The higher molecular weights of the mycobactins were attributed to its large alkyl chain in the R1 group. LC/ESI –MS analysis also suggested the presence of two groups of molecules, differing from each other not only by the multiples of 14Da ($\Delta M=CH_2$) and but also by 2Da (CH_2-CH_2 versus CH_2=CH_2) indicating the presence of un-saturation along the side chain. Thus, it can be predicted that both mycobactins and carboxymycobactins belong to two series, the saturated mycobactins and the unsaturated mycobactins. All these peaks were similar, as seen by the basic core nucleus in mass spectrometry. Mycobactin and carboxymycobactin had similar fragmentation pattern, differing only in the F moiety, indicating that mycobactins had long alkyl chains while the carboxymycobactin had short alkyl chains terminating either as carboxylic acid $(MH\text{-}CO_2)^+$ or as methyl ester $(MH\text{-}CH_3OH)^+$. Comparison of the fragment ions generated by *M. kansasii* showed that it was similar to that of *M. tuberculosis* and *M. avium*, the shift in the mass being primarily due to the different substituents in the B (β-hydroxybutyrate), D (oxazoline), E (salicylic acid) and/or F

(acyl / alkyl acid or methyl ester) moieties. The unique difference between the carboxymycobactins of *M. kansasii* and that of other mycobacteria is the methyl substituted salicylate (E) moiety, which was supported by the fragmentation ions at 235 and 362. In fragment D, cyclical threonine is present, thus R_3= CH_3, similar to that seen in *M. avium*; however it differs in fragment B, in that R_4 = CH_3 and R_5 = H while in *M. avium* R_4 = CH_2CH_3 and R_5 = CH_3. In the fragment B, it resembles *M. tuberculosis* in the substitutions at R_4 and R_5.

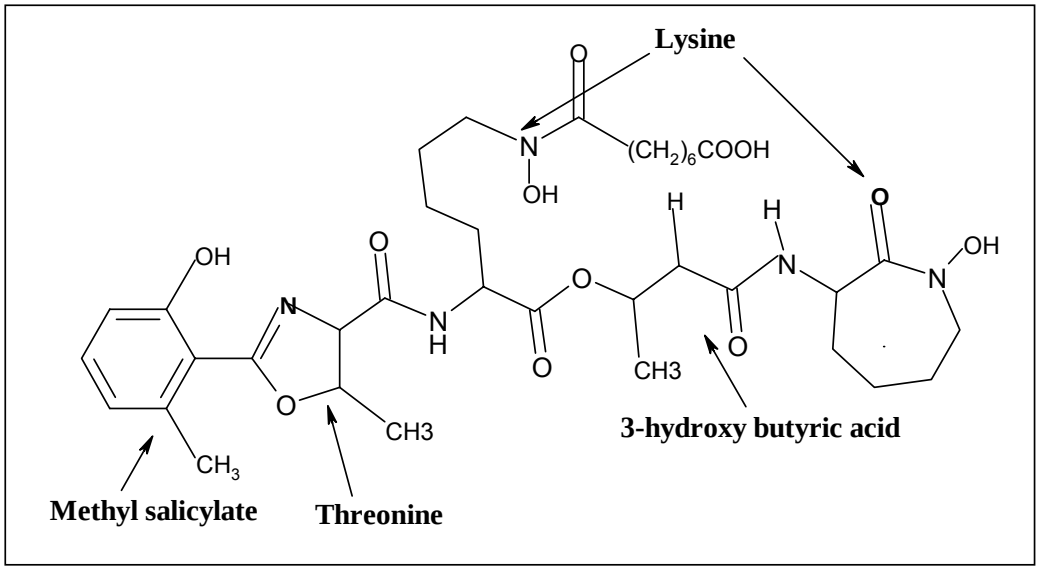


Fig. 35 - Structure of carboxymycobactin in *M. kansasii*

Why mycobacteria require two siderophores is still not clear. Mycobactin, owing to its long alkyl chain, is highly hydrophobic and is intracellular and is associated with the cell wall. The carboxymycobactin, being hydrophilic serves as the extracellular iron-solubilising agent chelating the iron from the environment. The iron, taken up via the ferri-carboxymycobactin and a possible receptor (see below for discussion on IREPs) is considered to be held by the mycobactin, that serves to store the iron temporarily, before it is transferred to the cytoplasm (Ratledge, 2004). In the cytoplasm, the excess iron is stored in the bacterioferritin (encoded by *bfrA* and *bfrB*).

The *in vivo* significance of iron and the importance of mycobactin in the growth of *M. tuberculosis* were established by knock out studies by De Voss *et al*., (2000). The study showed that a knockout mutant of *M. tuberculosis* lacking the *mbtB* gene in the mycobactin biosynthetic pathway failed to grow inside human macrophages and under *in vitro* conditions, though it was capable of growing in high iron medium, it did not grow in iron-limiting medium. The absence of mycobactin / carboxymycobactin, either singly or together could have resulted in the failure of the mutant to grown within the macrophage as the enzymatic defect was located in the beginning of the biosynthetic pathway of mycobactin / carboxymycobactin. The work of Gold *et al*., (2001) confirmed that the genes for mycobactin synthesis are active when *M. tuberculosis* is within the macrophages. They showed that a mutation in the *dtxR* allele resulted in the loss of virulence and the failure of the bacterium to acquire iron from the host. This work and other reports (Manabe *et al.,* 1999; Hobson *et al.,* 2002) gave clear evidence that *M. tuberculosis* is subjected to iron limitation within the macrophages and could acquire iron only via carboxymycobactin and mycobactin.

The iron acquisition machinery thus offers potential targets for the development of novel drugs. PAS (p-aminosalicylic acid), one of the oldest anti-TB drugs probably inhibits Mb synthesis rather than function as an inhibitor of folate biosynthesis. PAS can inhibit conversion of salicylate to Mb (Ratledge & Brown, 1972). Though PAS analogues have been made, they did not prove to be better than PAS. The enzymes of the biosynthetic pathway for Mb / CMb may not be suitable as they may be close to those involved in the synthesis of fatty acids or peptides in the mammalian system. Other efforts that have not given notable results include the use of xenosiderophores that is siderophores from other bacteria and substitution of other metal ions for iron. Gallium protoporphyrin, though it showed inhibition effect under *in vitro* conditions was not too promising. Synthetic analogs of mycobactin were tried by Hu and Miller (1997); the inhibition observed with increase in the hydrophobicity of these molecules raised the issue of toxicity to the host.

The ferric siderophores are transported into the cytoplasm via specific receptors called iron regulated envelope / membrane proteins (IREPs / IRMPs). These

are well studied in other bacterial systems. They were identified by comparing the gel electrophoresis patterns of proteins extracted from mycobacteria grown in low- and high-iron conditions.

At least five IREPs were detected in envelope extracts from *M. smegmatis* grown in iron-deficient conditions including the extensively studied 29 kDa protein (Hall *et al*, 1987). This protein can associate directly *in vitro* with ferri-exochelin, and the addition of a polyclonal antiserum generated against it to *M. smegmatis* cells significantly inhibits ferri-exochelin mediated iron uptake. Based on these observations, the 29 kDa protein has been postulated to be a ferri-exochelin receptor in *M. smegmatis*. IREPs are expressed in other mycobacteria like *M. neoaurum* (Sritharan & Ratledge, 1989), *M. avium* and ADMs (armadillo derived mycobacteria) when grown *in vitro* under iron-deficient conditions. Interestingly, IREPS were expressed in *in vivo* derived *M. avium* and *M. leprae*, isolated from infected tissue (Sritharan & Ratledge, 1990). Studies in *M. tuberculosis* (Wong *et al.*, 1999) using two-dimensional gel electrophoresis combined with mass spectrometry and sequence information have shown that the levels of several proteins change in response to iron availability. Among these, a putative cation transporting ATPase, a mycobacterial homologue of PEPCK (phosphoenolpyruvate carboxykinase) and an NADP-dependent dehydrogenase were induced in bacteria grown in a low-iron medium. On the other hand, FurA (a putative metal regulator seen in several Gram negative bacteria), a homologue of a translational factor EF-Tu and an aconitase were synthesized in higher amounts in bacteria grown in iron-rich medium.

Members of both Fur and DtxR (iron regulator identified in several Gram positive bacteria) families of metal regulators are found in mycobacteria. The annotation of the *M. tuberculosis* genome sequence revealed two Fur-like proteins, FurA and FurB, and two DtxR homologues, IdeR and SirR. In *M. leprae* genome sequence FurA and SirR are pseudogenes. In several mycobacteria including *M. tuberculosis, M. leprae, M. marinum* and *M. smegmatis, furA* is located immediately upstream of the catalase: peroxidase-encoding gene *katG* (Heym *et al.*, 1994, 1995, 1997; Pym *et al.*, 2001). In *M. tuberculosis* (Heym *et al.*, 1994) and *M. smegmatis*

(Milano *et al.,* 2001), FurA negatively regulates the expression of *katG*, thereby modulating the response to oxidative stress. In our study on the influence of iron on catalase-peroxidases, we showed a dramatic reduction of peroxidase upon iron deprivation in several mycobacterial species (Yeruva *et al.,* 2005) and we have showed the importance of this loss of peroxidase on INH activation and subsequent effect on *M. tuberculosis* (Sritharan *et al.,* 2006). FurB is currently annotated as Zur regulon as it helps in the maintenance of zinc homeostasis (Maciaq *et al.,* 2007). In *M. tuberculosis*, nothing is known about the function of SirR and currently IdeR is the metalloregulator that has been best characterized as to structure and function.

Using microarray technology and transcriptional profiling, Rodriguez and Smith (Rodriguez *et al.,* 2002) studied the response of *M. tuberculosis* to low- and high-iron conditions. The transcription of 155 ORFs was modulated by the levels of iron in the growth medium. Iron deficiency induced about two-thirds of those genes, while the remainder was up-regulated in iron-rich medium. Half the genes induced in low iron conditions were of unknown function. Among the others, two functional groups were distinguished: one part of the response is clearly focused in overcoming iron deficiency by increasing iron uptake. The siderophore biosynthesis genes (*mbt* genes) several genes encoding membrane proteins (IREPs), members of the PE-PPE protein family, putative transporters and several genes involved in basic metabolic pathways were also up-regulated during iron deficiency. The other part of the response adjusts the general metabolism of the cell to meet the challenge imposed by iron restriction. High iron levels in medium induces expression of comparatively smaller number of genes especially those involved in iron storage (*bfrA* and *bfrB*) and oxidative stress defense enzymes such as *katG* (Rodrigues *et al.,* 2002). Also it is observed that genes involved in the energy generation (*nuo*) encoding the polypeptide subunits of the iron containing respiratory enzyme complex NADH dehydrogenase and *atpE*, which encodes the ATP synthase C chain are also induced during iron rich medium (Rodrigues *et al.,* 2002). The *mtrA*, a major gene induced during iron rich conditions was found to be one of the essential genes, which encodes a two component system (Zarht & Deretic, 2000). Recently, the IrtAB operon was identified and the

expressed protein products were implicated in carboxymycobactin uptake and transport across the membrane (Rodriguez & Smith 2006). In our lab, studies on iron acquisition in *M. tuberculosis* showed that HupB (a 28 kDa protein) was coordinately expressed with Mb and CMb (Yeruva *et al*., 2006). The role of HupB remains to be deciphered, though it has been shown to be of significant clinical relevance.

In this study, we subjected *M. kansasii* to varying iron concentrations and analysed the cell wall and membrane proteins for the expression of novel IREPs. When iron was added at (0.02 μg Fe / mL medium), a concentration when maximal CMb was expressed, three cell membrane proteins of approximate molecular mass 112, 66 and 35 kDa was up-regulated, with the respective bands still seen at lower levels in cells grown with iron concentration increased to 4 μg Fe / mL medium. A 52 kDa cell membrane protein was found to be expressed only in cells grown with 0.02 to 0.4 μg Fe / ml of medium, with corresponding levels of expression of mycobactin. This 52 kDa cell membrane protein was identified by mass spectrometry as DeaD (Rv1251) with helix destabilising activity while *ponA2* encodes a membrane protein involved in the peptidoglycan synthesis.

The cell wall fraction did not show significant differences, except for a 52 kDa protein seen in low iron cells and it was identified CysN enzyme acting as the sole sulphate transporter. The gene *cysN* (Rv1286), encoding this protein was identified as an essential gene for the survival of *M. tuberculosis* by *HimarI* gene transposition (Sassetti & Rubin, 2003). The cell membrane protein Mk40', purified by HPLC analysis was found to be Rv0078, a transcriptional regulatory protein. Though we showed the interaction of this protein with mycobactin and carboxymycobactin by spectrofluorimetric analysis, further analysis is required to determine its exact function. The increased expression of transcriptional regulators has been reported in response to iron starvation (Rodrigues *et al.,* 2002).

Iron, undoubtedly plays an important role in infection. Anemia is frequently encountered in TB patients. As discussed earlier, there is evidence to show that iron limitation occurs within the macrophages. Also, we showed that the efficacy of INH was influenced by the iron status (Sritharan *et al*., 2006), the low iron cells being

unaffected by INH due to the lack of peroxidase activity under these conditions. Despite the wealth of knowledge on mycobacterial siderophores, more remains to be understood about the machinery in different mycobacterial pathogens, especially the mechanism of uptake to exploit the iron acquisition machinery as potential drug targets.

CHAPTER 4

FUNCTIONAL GENOMICS OF *PYKA* AND *ALD* MUTANTS TO STUDY GLYCEROL UTILISATION IN *M. TUBERCULOSIS* AND *M. BOVIS*

Results

4.0 *In silico* analysis of *ald* gene and identification of deletions among the members of *M. tuberculosis* complex

The *ald* gene of *M. tuberculosis* gives a functional Ald protein but *M. bovis* does not give a functional Ald protein. In the latter, due to SNP deletion at position 362 (Fig. 36), leading to the production of truncated protein, the gene was annotated as *aldA* and *aldB* in *M. bovis* genome. Similar deletion was seen in *M. bovis* BCG and in *M. microti.* This deletion as well as deletion occurring in the RD9 region diversifies the other members of the *M. tuberculosis* complex from *M. tuberculosis* and *M. canetti.* In this study, the sequence of the PCR amplified *ald* gene from the two clinical strains of *M. africanum* showed a 6 bp deletion from nucleotide 462 to 468 (Fig. 36), the significance of which remains to be ascertained.

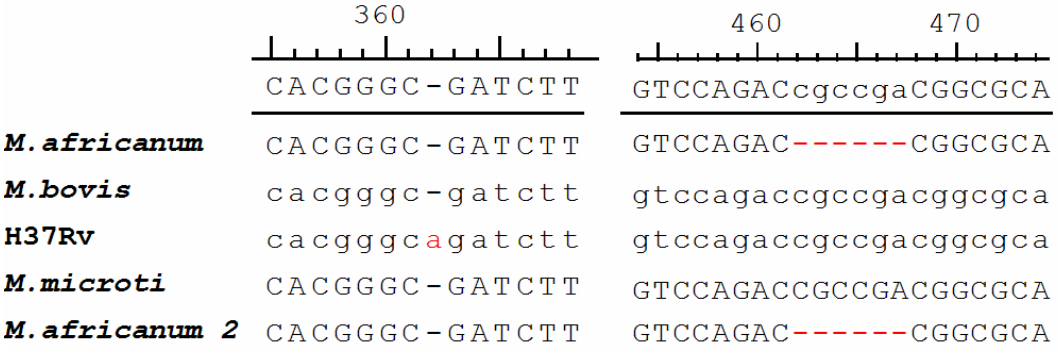


Fig. 36 - *In silico* analysis of *ald* gene among members of *M. tuberculosis* complex

4.1. Generation of recombinant plasmids

4.1.1 Recombinant pSR1 encoding *ald* gene from *M. tuberculosis*

The *ald* gene (1116 bp) from *M. tuberculosis* was amplified (Fig. 37a) and ligated with *PstI* digested pSM96 (Fig. 37b). Plasmid was transformed into *E. coli* by heat shock method. Some of the recombinant plasmids were analysed by restriction digestion (Fig. 37c) and two of the positive clones were confirmed for the orientation of *ald* gene by digestion with *EcoRI* (Fig. 37d) and named as pSR1 (Fig. 38).

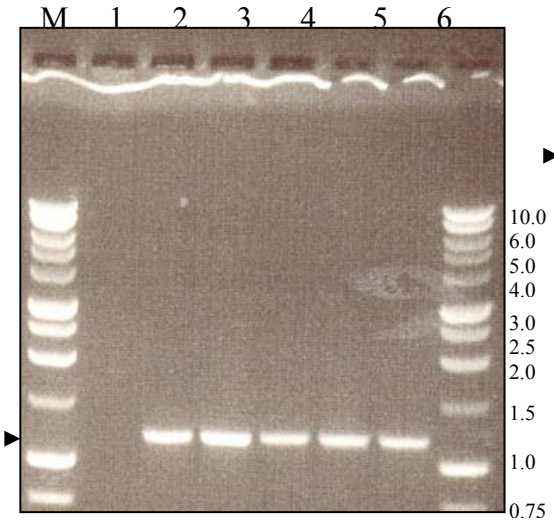


Fig. 37a - PCR amplification of *ald* (1116 bp) from *M. tuberculosis*. Positives (lanes 2-6), negative control (lane 1) & 1 kb DNA ladder (lanes M).

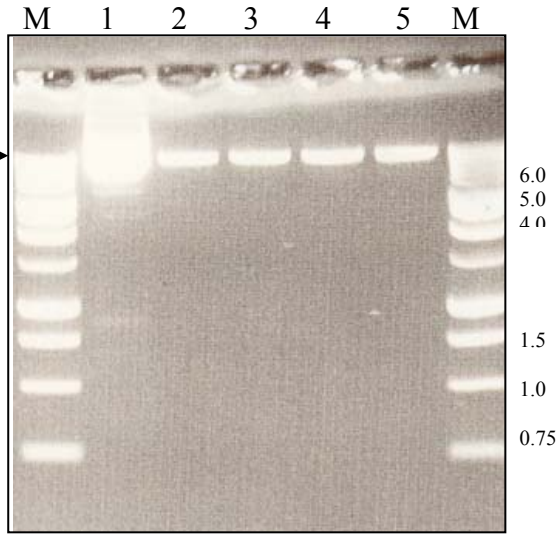


Fig. 37b - pSM96 restriction digested with *Pst I* and dephosphorylated (lanes 2-5), undigested plasmid (lane 1) &1 kb DNA ladder (lanes M).

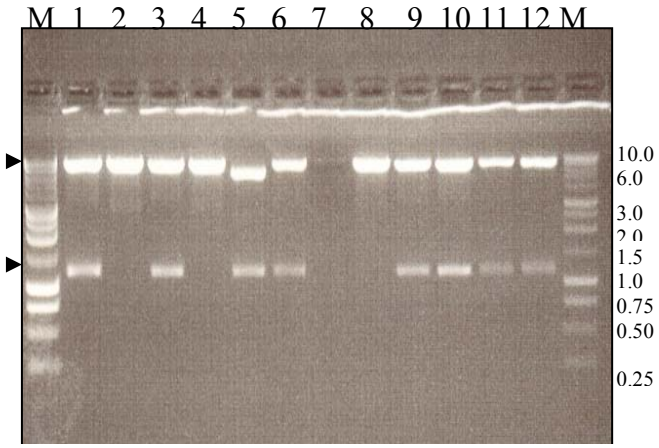


Fig. 37c - Verification of recombinant plasmids by restriction digestion with *Pst I.* Lanes 1-12 represents putative recombinant plasmids and lanes M represents 1 kb DNA ladder. Plasmids corresponding to lane 1 (pSR1) and 3 (pSR3) were selected on random basis and utilised for cloning in mycobacteria.

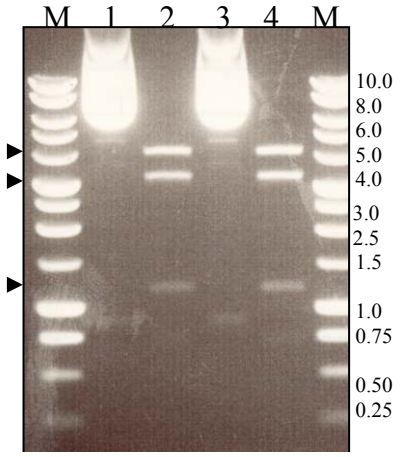


Fig. 37d - Verification of orientation in pSR1.1 and pSR1.3 by restriction digestion with *EcoRI.* Lanes 1 and 3 - undigested plasmid, lanes 2 and 4 - digested plasmids pSR1.1 and pSR1.3 respectively and lanes M 1 kb DNA ladder. Products of 3959, 3094 and 1150 bp indicate the right orientation.

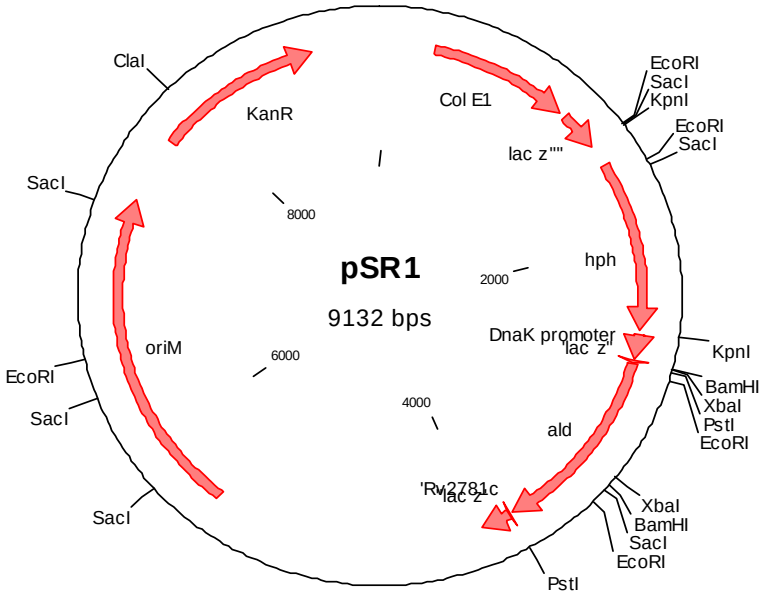


Fig. 38 - Map of pSR1. Plasmid contains *ald* gene with DnaK promoter in the upstream region.

4.1.2. Construction of the plasmid vector pSN2 for generating *ald* knock-out mutant in *M. tuberculosis*

The presence of similar restriction enzymes sites in the *ald* gene and pSMT100, causing a constraint in the use of some of the restriction enzymes led us to delete a 300 bp internal region from *ald.* Upstream (1030 bp) and down stream (1113 bp) regions of *ald* gene were PCR amplified and separately ligated with pSMT100 to generate the plasmid pSN2. The cloning of the upstream region was done by standard protocols (Fig. 39a-c). We selected the clones 4 and 6 (Fig. 39d) as they showed *EcoRI* restriction digested products of sizes 2140, 2032 and 1270 bp respectively indicating the right orientation of the *ald* gene. These plasmids pSN1.4 and pSN1.6 were used for the cloning of the 1113 bp downstream region. Fig. 39e, 39f and 39g represents the steps involved in the generation of the plasmid vector pSN2, containing the upstream 1030 bp and downstream 1113 bp region of *ald* on either side of hygromycin resistance cassette (Fig. 40).

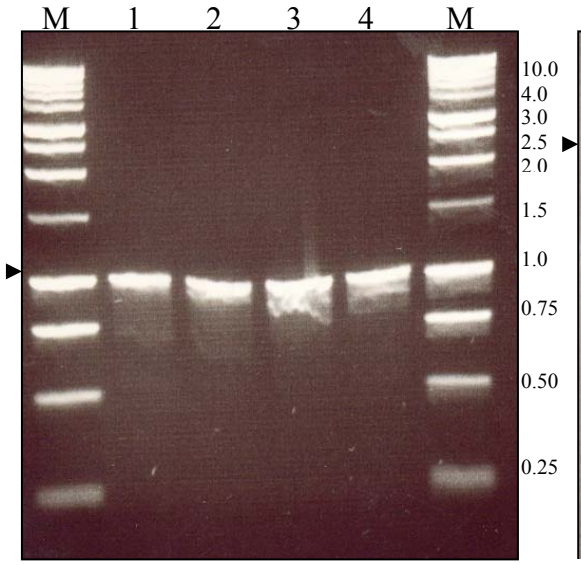


Fig. 39a - PCR amplification of upstream region (1030 bp) of *ald* from *M. tuberculosis* (lanes 1-4), 1 kb DNA ladder (lanes M).

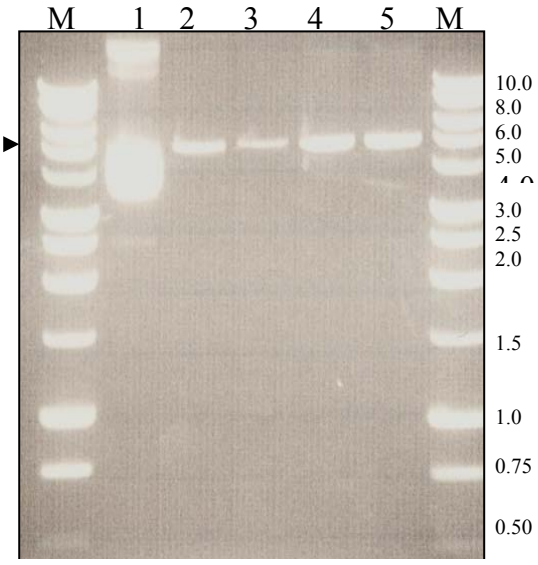


Fig. 39b - pSMT100 restriction digested with *Pst I* and dephosphorylated (lanes 2-5), undigested plasmid (lane 1), lanes M - 1 kb DNA ladder.

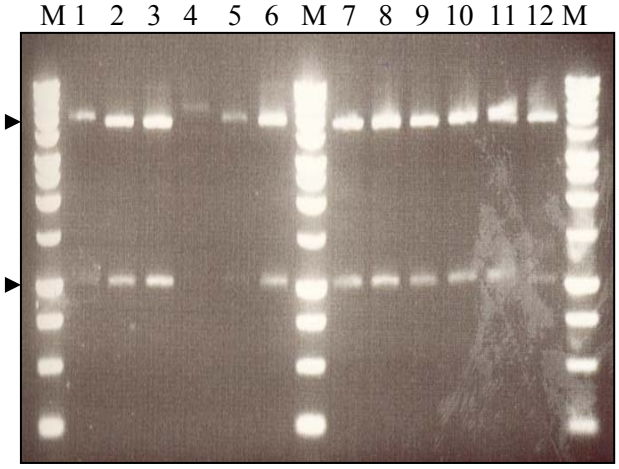


Fig. 39c - Verification of recombinant plasmids by restriction digestion with *PstI.* Lanes 1 to 12 - putative recombinant plasmids and lanes M- 1 kb DNA ladder. Plasmids corresponding to lanes 7 to 12 were checked for upstream region orientation.

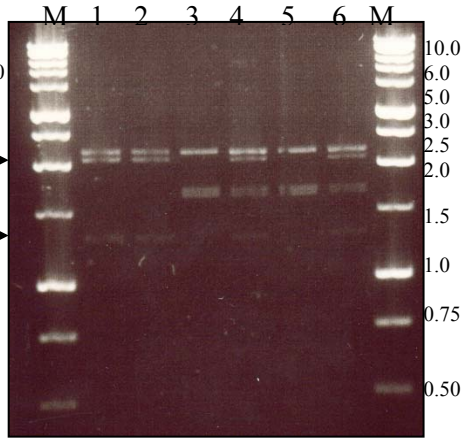


Fig. 39d - Verification of orientation of upstream region of *ald* gene by restriction digestion with *EcoRI.* Lanes 1 to 6 represents plasmids digested with *EcoRI* and lane M is 1 kb DNA ladder. Products of sizes 2140, 2032 and 1270 bp indicate right orientation

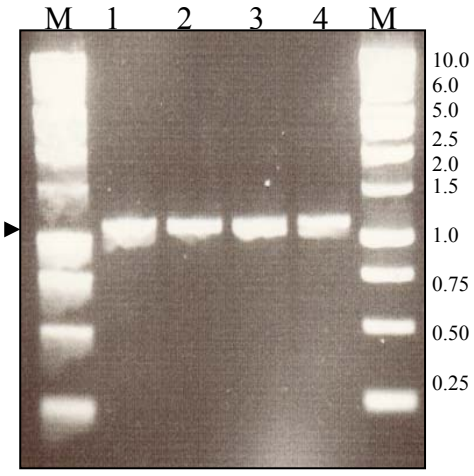


Fig. 39e - PCR amplification of downstream region (1113 bp) of *ald* from *M. tuberculosis* (lanes 1-4), 1 kb DNA ladder (lanes M).

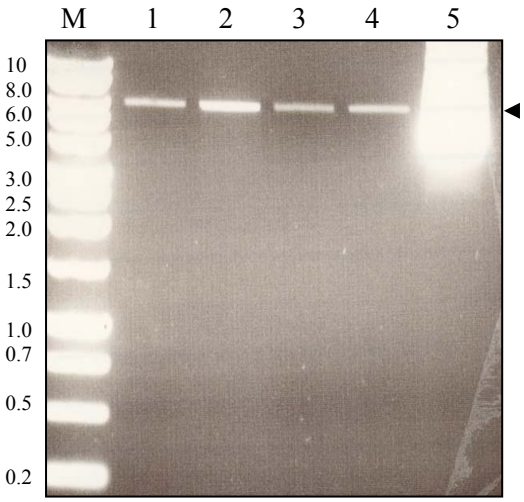


Fig. 39f - pSN1 restricted digested with *SpeI* and *XmaI* (lanes 1-4), undigested plasmid (lane 5), and 1 kb DNA ladder (lane M).

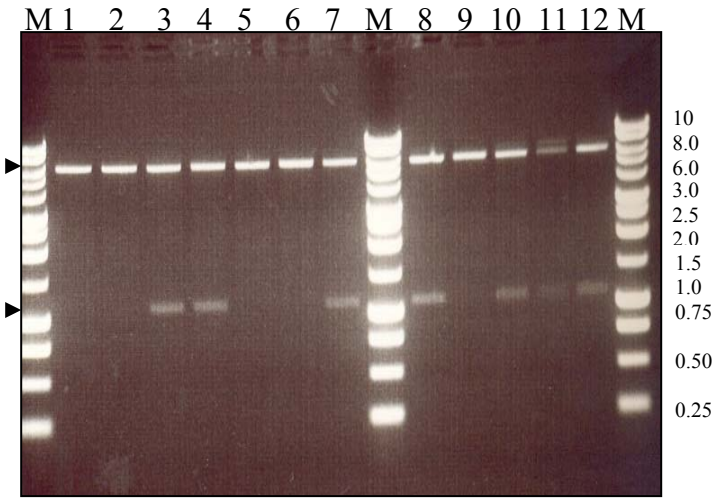


Fig. 39g - Verification of inserts by restriction digest with *XmaI* and *SpeI*. Lanes 1 to 12 represents putative recombinant plasmids and lanes M represent 1 kb DNA ladder. Plasmids in lane 7 (pSN2.7) and 8 (pSN2.8) were selected on random basis and utilised for cloning in mycobacteria.

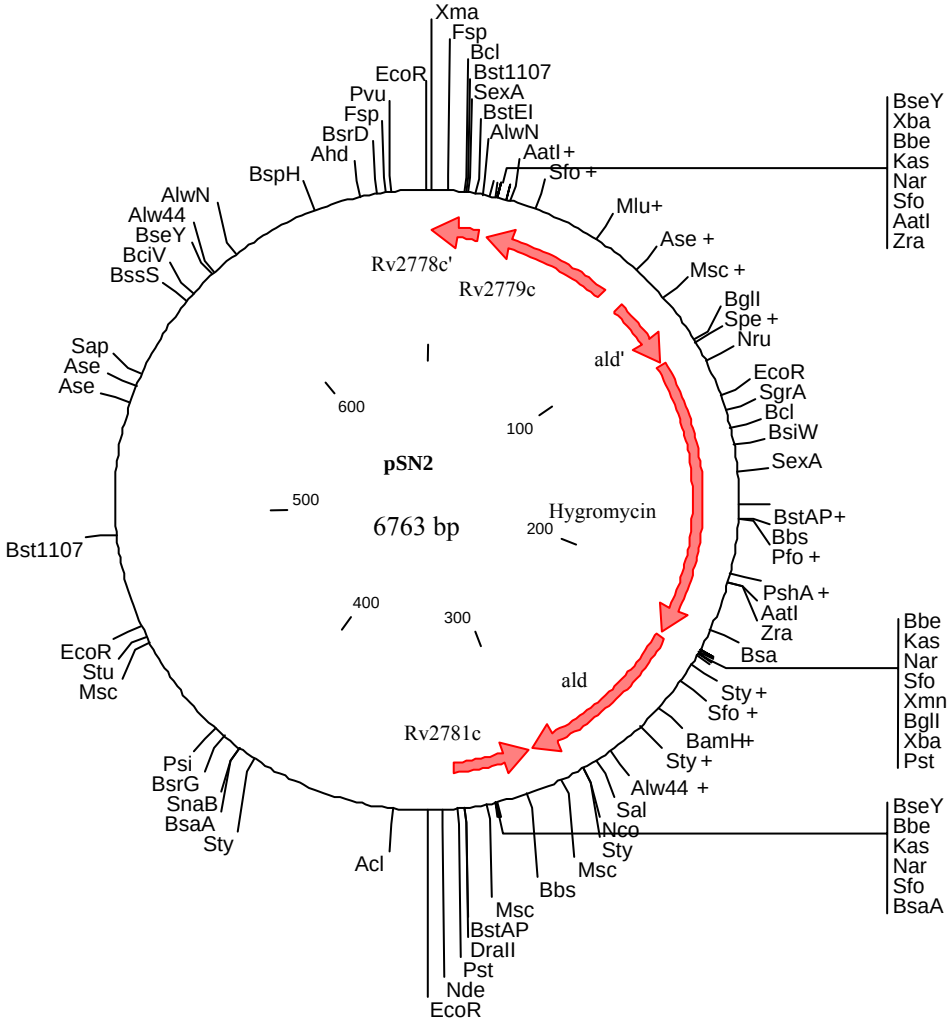


Fig. 40 - Map of pSN2. Plasmid consists of upstream and downstream regions flanking the hygromycin cassette.

4.1.3 Construction of pSC2 for generating *pykA* knock-out mutant in *M. tuberculosis*

Similar to the strategy followed for the generation of pSN2, the upstream (1059 bp) and downstream (1056 bp) regions of *pykA* from *M. tuberculosis* were PCR amplified, restriction digested and ligated into pSMT100 (Fig. 41a-g) to generate pSC2 (Fig. 42).

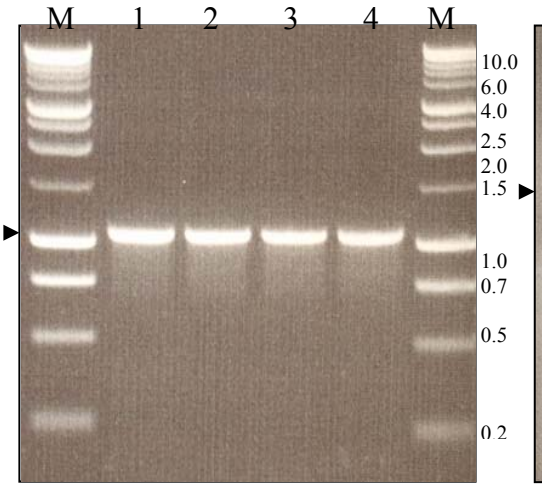


Fig. 41a - PCR amplification of upstream region (1059 bp) of *pykA* from *M. tuberculosis* (lanes1-4) & 1 kb DNA ladder (lanes M).

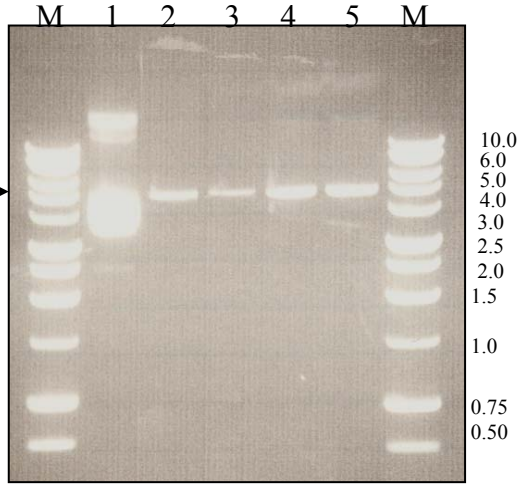


Fig. 41b - pSMT100 restriction digested with *XbaI* and dephosphorylated (lanes 2-5), undigested plasmid (lane 1) & 1 kb DNA ladder (lane M).

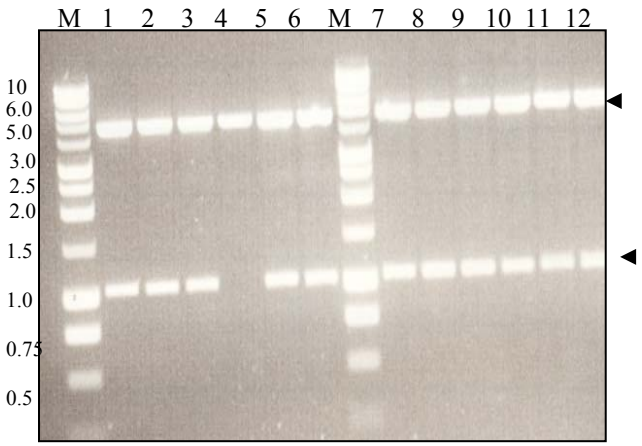


Fig. 41c - Verification of recombinant plasmids by restriction digest with *XbaI.* Lanes 1 to 12 - putative recombinant plasmids & lanes M - 1 kb DNA ladder. Plasmids corresponding to lane 2, 3, 11, 12 were checked for their orientation.

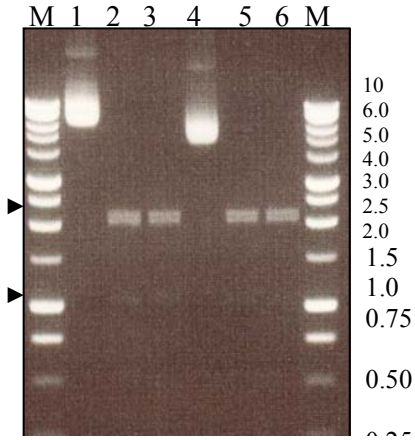


Fig. 41d - Verification of orientation in plasmids by restriction digest with *XmaI.* Lanes 1 and 4- undigested plasmids, lanes 2, 3, 5 and 6 - *Xma I* digested plasmids and lanes M - 1 kb DNA ladder. Product size of 2013 and 1024 bp indicate the right orientation. Plasmids in lane 3 (pSC1.3) and 6 (pSC1.11) were selected for construction of pSC2.

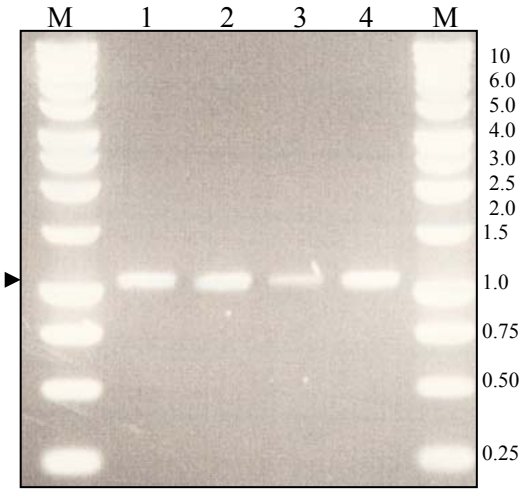


Fig. 41e - PCR amplification of downstream region of *pykA* from *M. tuberculosis* (lanes 1-4) & 1 kb DNA ladder (lane M)

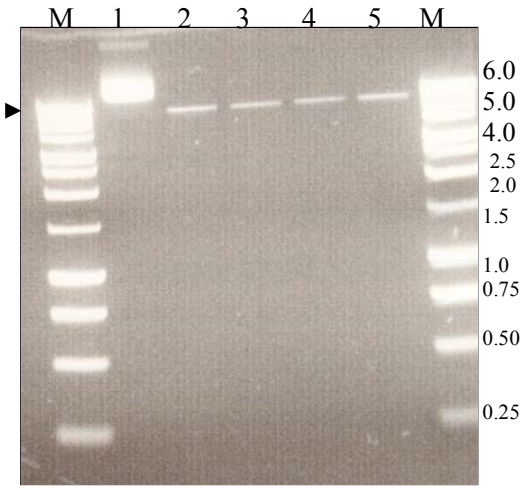


Fig. 41f - pSC1 restricted digested with *SpeI* and *XmaI*. Lanes 1 to 4 - digested plasmid, lane 5 - undigested plasmid & lanes M-1 kb DNA ladder.

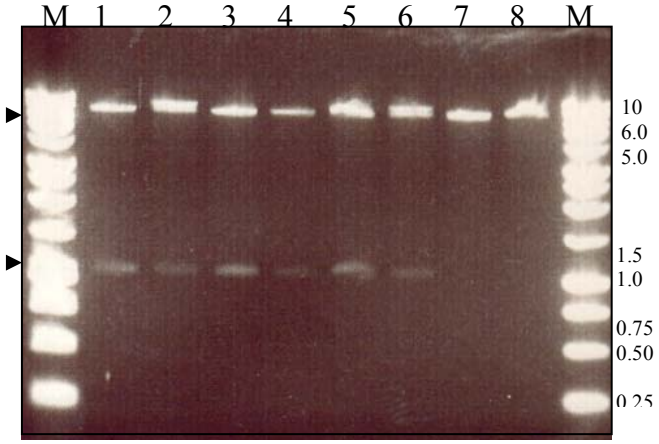


Fig. 41g - Verification of inserts by restriction digest with *XmaI* and *SpeI*. Lanes 1 to 8 are putative recombinant plasmids and lanes M -1 kb DNA ladder. Plasmids corresponding to lane 1 (pSC2.1) and 2 (pSC2.2) were selected on random basis.

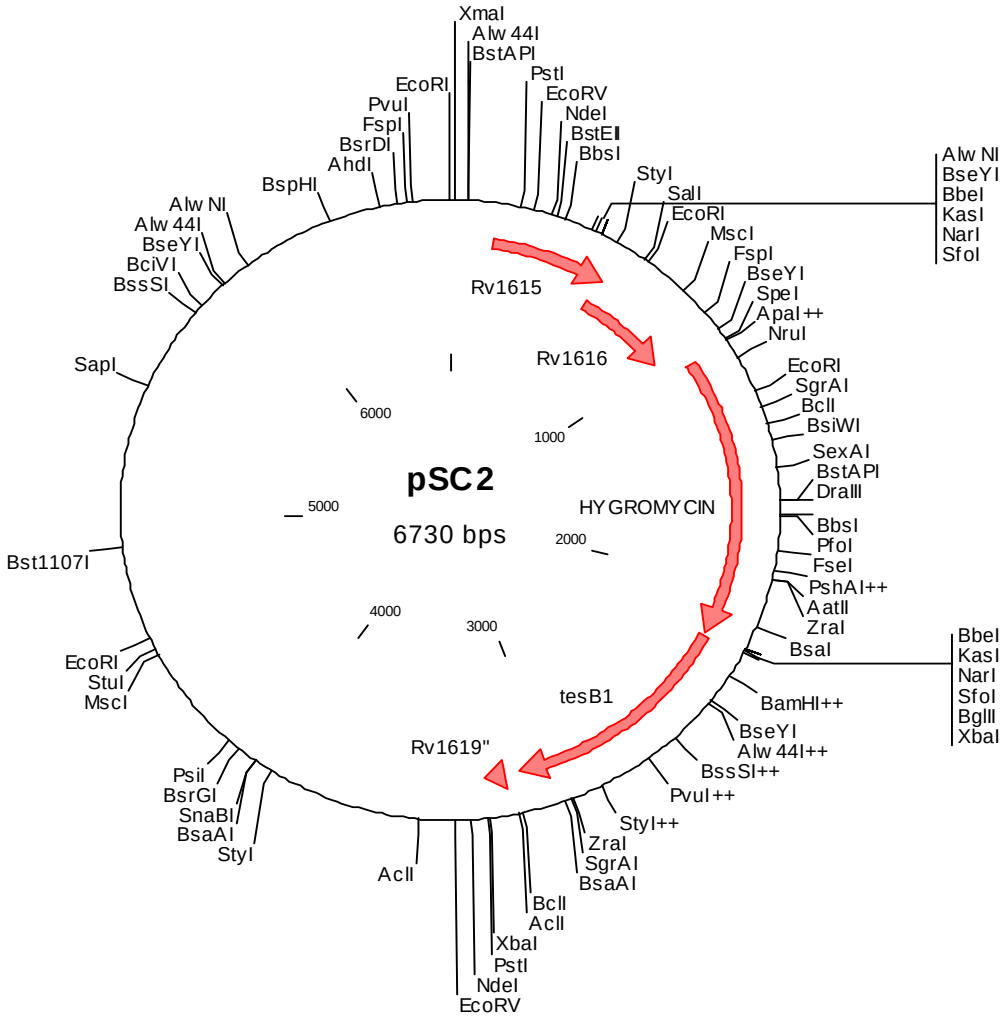


Fig. 42 - Map of pSC2. Suicide plasmid vector containing right and left flanking regions of *pykA* on either side of hygromycin resistance cassette.

4.1.4. Construction of pCM2 for generating *ald* and *pykA* double knock-out mutant in *M. tuberculosis*

The vector pSM96, bearing the kanamycin cassette was used to introduce the kanamycin resistance gene and replace the hygromycin cassette in pSC2. The 1100 bp *kanR* gene in pSM96 was PCR amplified (Fig. 43a) and ligated into pSC2, which is restriction digested with *SpeI* and *BglII* to remove the hygromycin cassette (Fig. 43b). The resulting clones were tested and plasmids pCM2.1 and pCM2.2 (Fig. 43c) were

selected. Fig. 44 represents pCM2 used for generating *M. tuberculosis Δ ald Δ pykA* double knock-out.

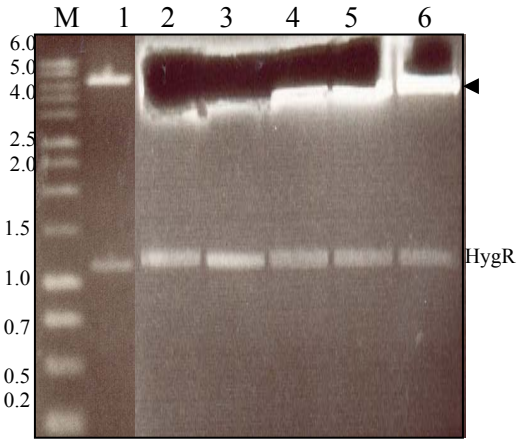


Fig. 43a - pSC2 digested with *SpeI* and *BglII* to remove hygromycin cassette (lane 1). In lanes 2-6, the arrow indicates the cut plasmid minus the Hyg^R

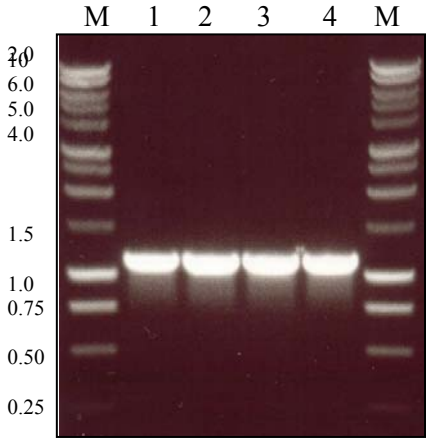


Fig. 43b - PCR amplification of Kan^R gene from pSM96. Lanes 1-4 represents PCR product and lanes M - 1 kb DNA ladder

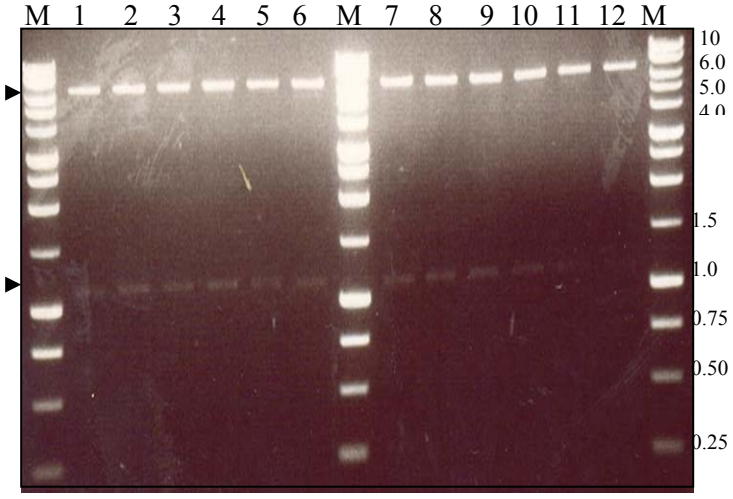


Fig. 43c - Verification of inserts by restriction digestion with *BglII* and *SpeI*. Lanes 1-12 represents putative recombinant plasmids and lane M represents 1 kb DNA ladder. Plasmids in lane 1 (pCM2.1) and 2 (pCM2.2) were selected on random basis.

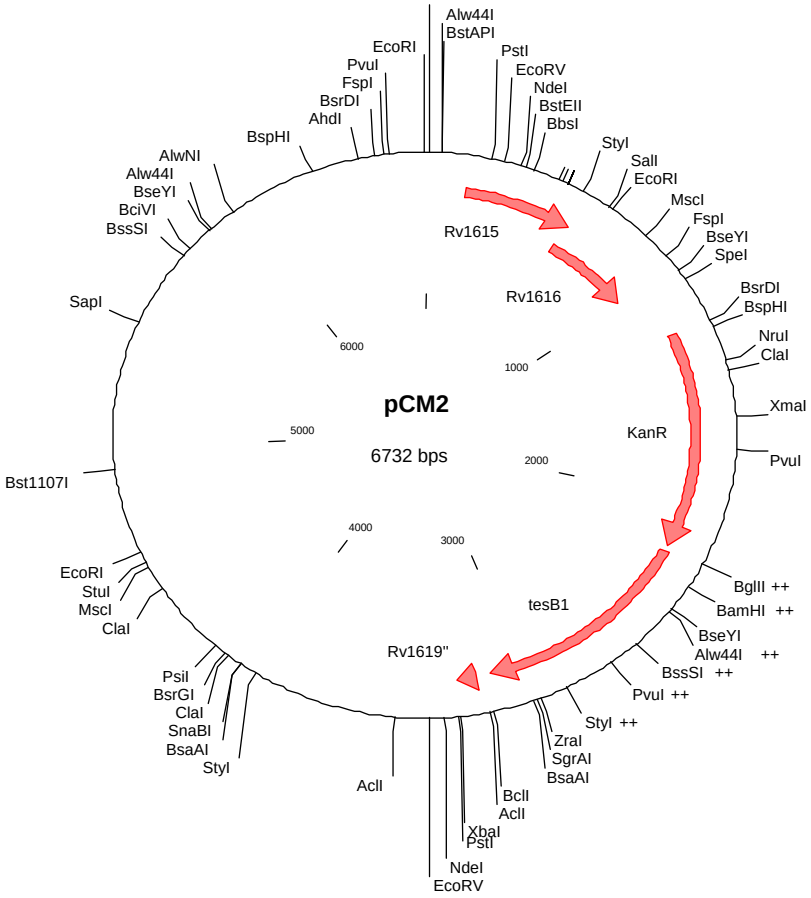


Fig. 46 - Map of pCM2 with upstream and down stream regions of *pykA* gene on either side of the kan^R cassette.

4.1.5. Construction of pMS2 for complementing *ald* and *pykA* in *M. bovis*

The *ald* and *pykA* (Fig. 45a & 45f) genes were separately amplified by PCR from *M. tuberculosis* genomic DNA. Reverse primer for the amplification of *ald* was designed with the ribosome binding site so as to allow the continuous transcription of both *ald* and *pykA* genes in the same plasmid. The *ald* PCR product was cloned into pSM96 (Fig. 45b & Fig. 45c) to generate pMS1 in which the orientation of the insert was verified (Fig. 45d & Fig. 45e). The 1448 bp *pykA* PCR product was cloned into the newly generated pMS1 (Fig. 45f, Fig. 45g & Fig. 45h). The plasmid pMS2 (Fig. 46) now has both *ald* and *pykA*.

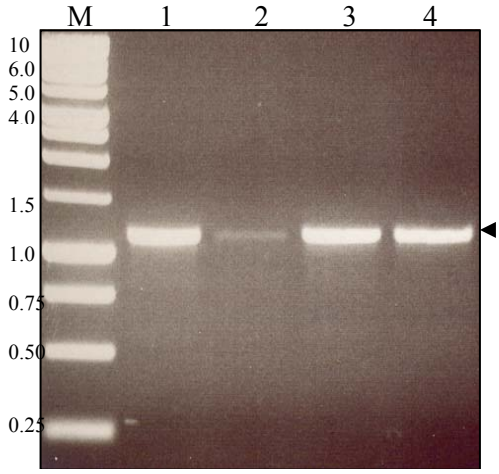


Fig. 45a - PCR amplification of *ald* from *M. tuberculosis* (lanes 1-4), lane M is 1 kb DNA ladder.

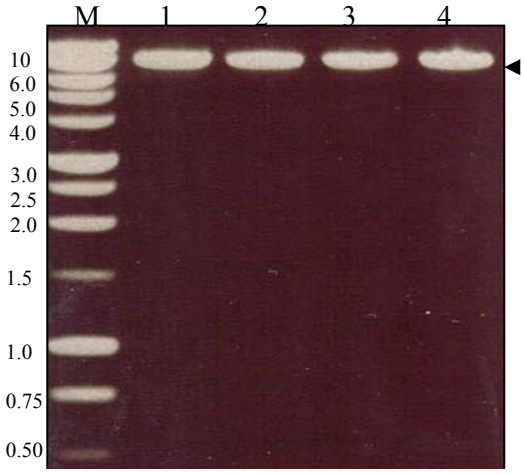


Fig. 45b - pSM96 restriction digested with *Pst I* and dephosphorylated (lanes 1-4); lane M -1 kb DNA ladder.

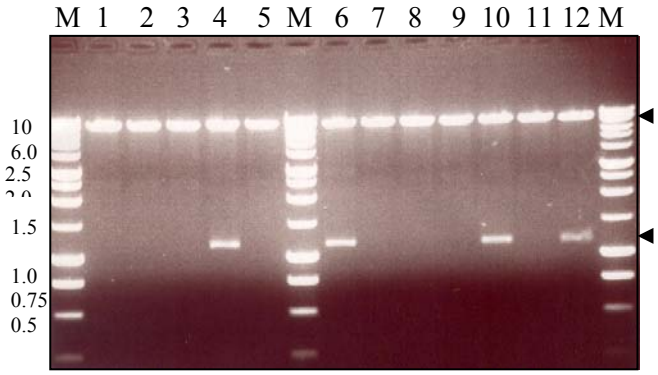
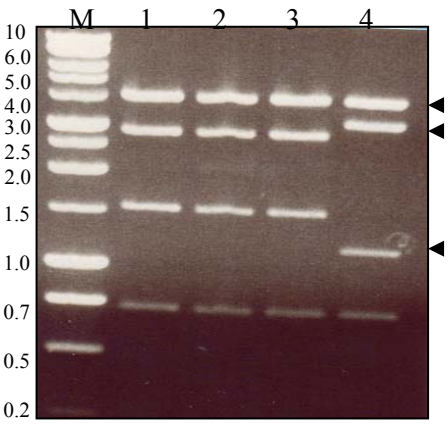


Fig. 45c - Verification of inserts by restriction digest with *Pst I.* Lanes 1-12 represents putative recombinant plasmids; lanes M -1 kb DNA ladder.



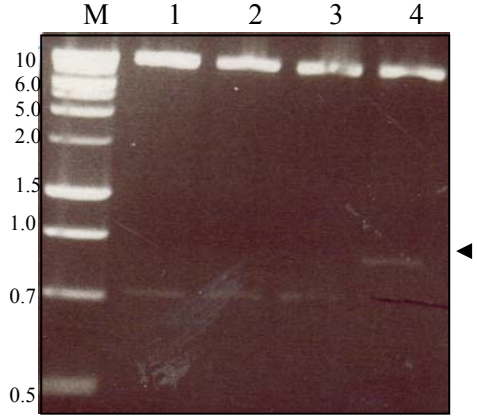


Fig. 45 d & Fig. 45e - Verification of orientation in plasmids by restriction digest with *BamHI* and *EcoRI* respectively. Lanes 1-4 represents *Bam HI* and *Eco RI* digested plasmids; lane M represents 1 kb DNA ladder. When digested with *EcoRI* product size of 3959, 3054 and 1150 bp and when digested with *BamHI* product size 8467 and 665 bp correspond to correct orientation, respectively. Plasmids corresponding to lane 4 (pMS1.4) were selected for further construction of knock-in vector.

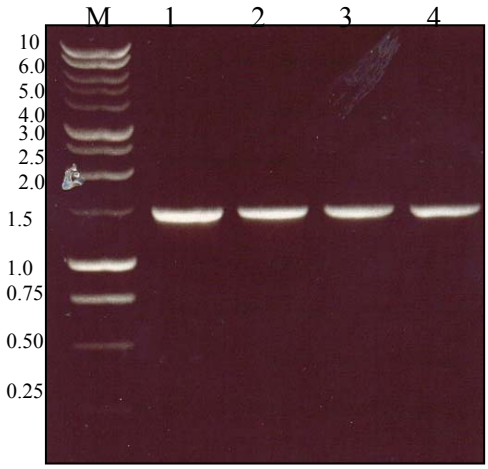


Fig. 45f - PCR amplification of *pykA* from *M. tuberculosis* (lanes 1-4) & lane M 1 kb DNA ladder.

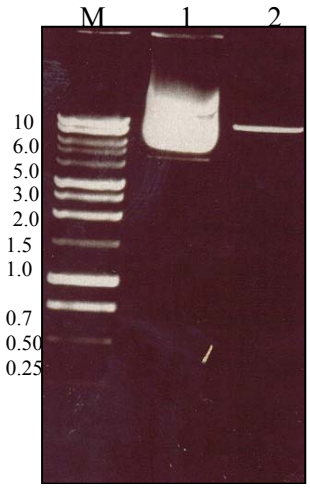


Fig. 45g - pMS1 restriction digested with *PacI* and *ScaI* (lane 2), undigested plasmid (lane 1) & 1 kb DNA ladder (lane M)

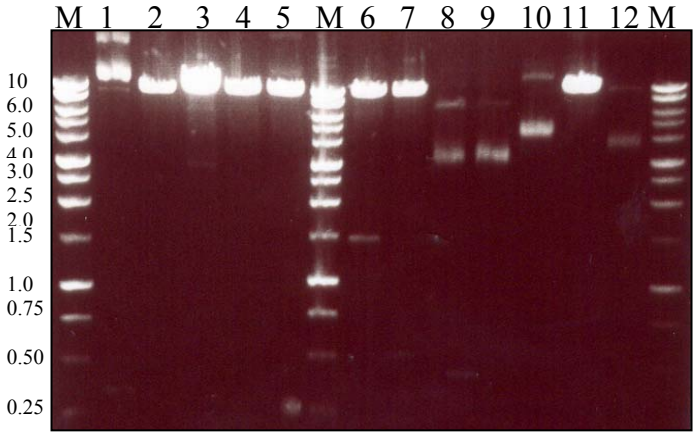


Fig. 45h - Verification of inserts by restriction digestion with *PacI* and *ScaI*. Lanes 1 to 12 - putative recombinant plasmids, lanes M - 1 kb DNA ladder. Plasmids corresponding to lane 6 (pMS2.6) was utilised for cloning in mycobacteria.

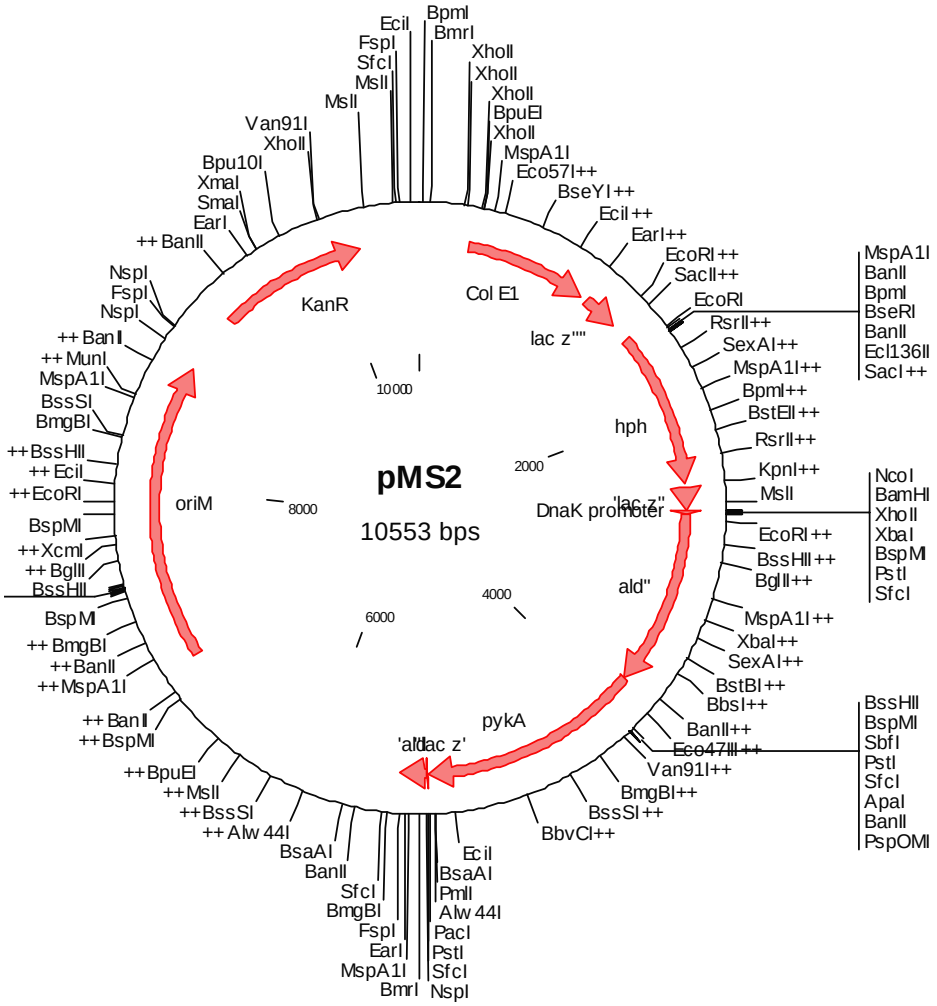


Fig. 46 - Map of pMS2 expressing *ald* and *pykA* genes with DnaK promoter region upstream.

4.2. Generation of mycobacterial mutants

4.2.1. Generation of *M. bovis* / *M. bovis* BCG mutants

4.2.1.1. Complementation of *M. bovis* / *M. bovis* BCG with *ald* gene

Both *M. bovis / M. bovis* BCG were transformed with pSR1. The knock-in mutants were initially screened by antibiotic selection. The colonies were then analysed by Hotstart PCR and those which gave the 661 bp PCR product specific for *hygR* gene (Fig. 47a & Fig. 47b), were considered as positive clones and labeled as *M. bovis* / pSR1 and *M. bovis* BCG / pSR1 respectively.

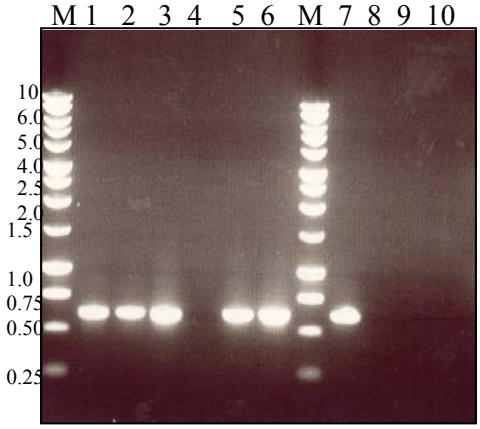


Fig. 47a - PCR verification by amplification of Hygromycin cassette. Putative *M. bovis* BCG mutants (lanes 2-10), plasmid control (lane 1), wild type (lane 11) and lane M - 1 kb DNA ladder.

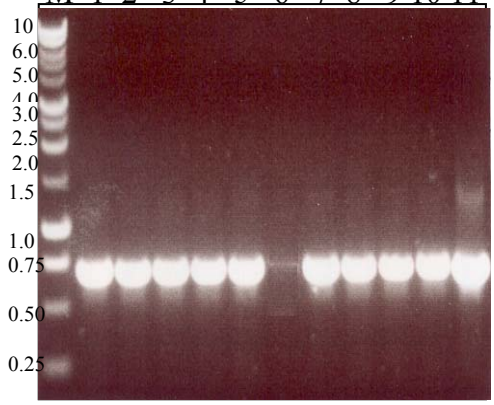


Fig. 47b - PCR verification by amplification of Hygromycin cassette. Lanes 1-5 and 7-10 putative *M. bovis* mutants, plasmid control (lane 11), wild type (lane 6) and lane M - 1 kb DNA ladder.

4.2.1.2. Complementation of *M. bovis* with *ald* and *pykA* genes

Mycobacterium bovis was transformed with pMS2. The knock-in mutants were initially screened by antibiotic selection. The colonies were then analysed by Hotstart PCR and those which gave the 442 bp PCR product specific for *kanR* gene (Fig. 48), were considered as positive clones and labeled as *M. bovis* / pMS2.

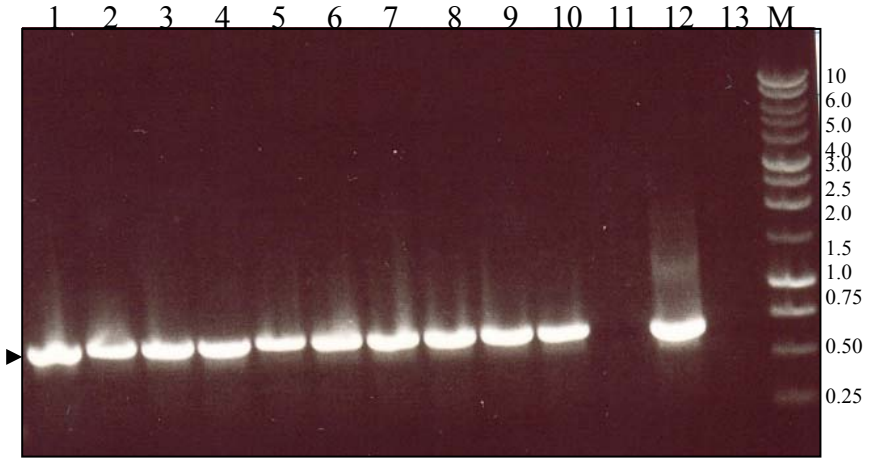


Fig. 48 - PCR verification by amplification of Hygromycin cassette. Lanes 1-10 represents putative *M. bovis* mutants, plasmid control (lane 12), wild type (lane 11), negative control (13) and lane M- 1 kb DNA ladder.

4.3. Generation of mutants of *M. tuberculosis*

4.3.1. Knock-out mutants

4.3.1.1. *M. tuberculosis Δ ald*

Knock-out mutant of *M. tuberculosis Δ ald*, generated by introducing pSN2 into *M. tuberculosis* H37Rv was due to the allelic replacement due to homologous recombination between the *ald* gene elements in pSN2 and the corresponding gene in the mycobacterial genome. Hotstart PCR confirmed the site-specific recombination (Fig. 49a & Fig. 49b) and the presence of hygromycin cassette (Fig. 49c). Since the *ald* knock-out was made by 300 bp deletion, the PCR product of 2050 bp size was considered to be positive for the selection of knock-out by XLPCR (Fig. 49d) and the 1036 bp product formed was the natural *ald* (Fig. 49d). The strain was named as *M. tuberculosis* H37Rv Δ *ald.*

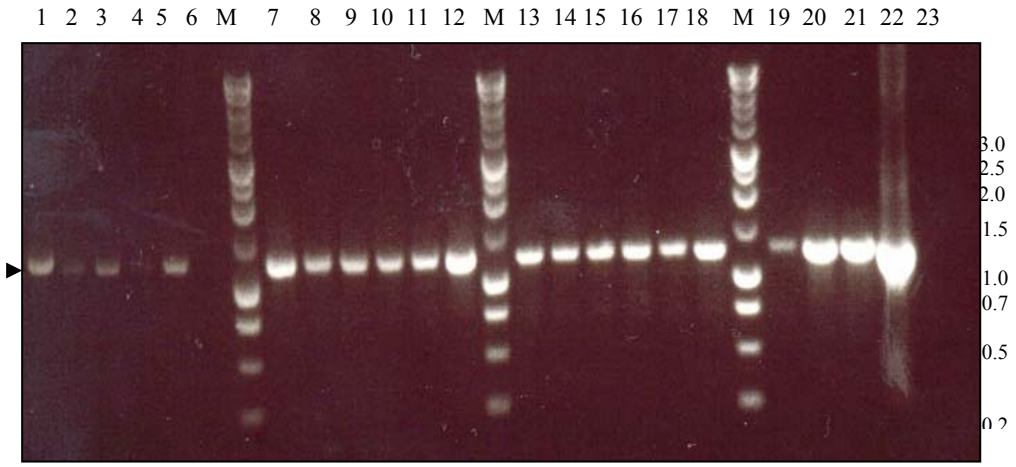


Fig. 49a - Hotstart PCR verification of upstream region with antibiotic resistance cassette from the genomic DNA of putative recombinant *M. tuberculosis Δ ald* (lanes 1-21), plasmid control (lane 22), *M. tuberculosis* H37Rv (lane 23), lanes M - 1 kb DNA ladder.

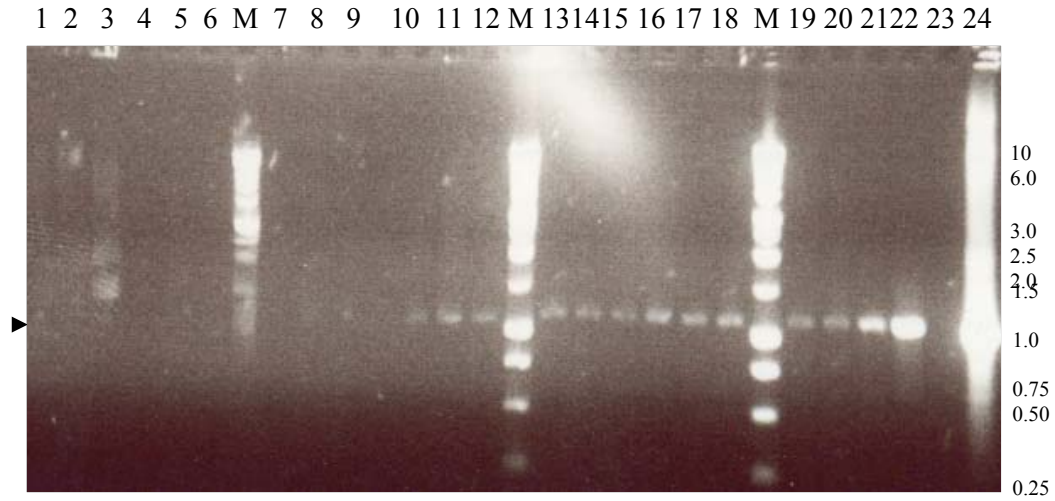


Fig. 49b - PCR verification of downstream region with antibiotic resistance cassette from the genomic DNA of putative recombinant *M. tuberculosis Δ ald* (lanes 1-22), plasmid control (lane 24), *M. tuberculosis* H37Rv (lane 23), lanes M - 1 kb DNA ladder.

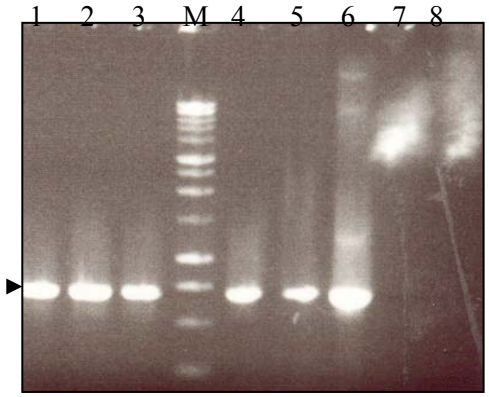


Fig. 49c - PCR verification by amplification of hygromycin cassette from mutants (lanes 1-5), wild type (lane 7), negative control (lane 8), plasmid control (lane 6) & lane M - 1 kb DNA ladder.

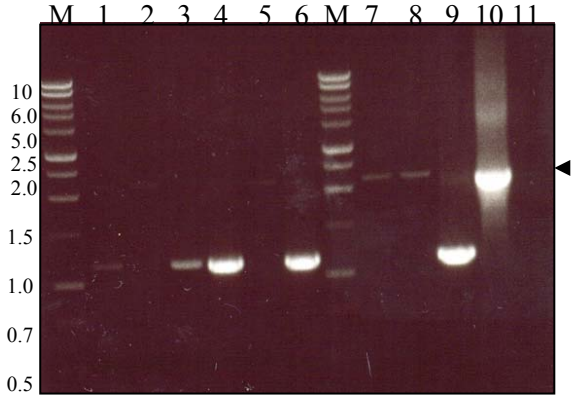


Fig. 49d - PCR verification by amplification of *ald* from mutants (lanes 1-8), *M. tuberculosis* H37Rv (lane 9), pSN2 (lane 10), negative control (lane 11) & lanes M - 1 kb DNA ladder.

4.3.1.2. *M. tuberculosis Δ pykA*

Knock-out mutant *M. tuberculosis Δ pykA* was generated by a similar strategy by introducing pSC2 into *M. tuberculosis* and the recombinant mutants were analysed by PCR (Fig. 50a-c) to identify the mutants. A negative PCR reaction with primers PykABamHIF and PykABamHIR (Fig. 50d), showed deletion of *pykA* gene in *M. tuberculosis* and labeled as *M. tuberculosis Δ pykA*.

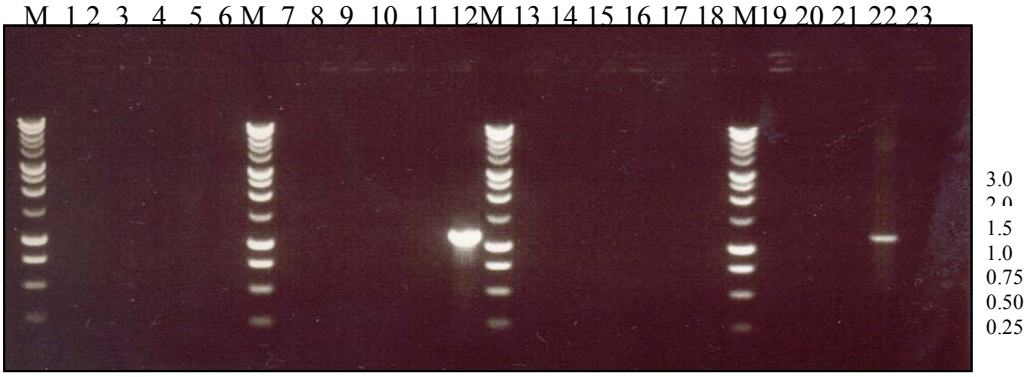


Fig. 50a - PCR amplification of upstream region with antibiotic resistance cassette from the genomic DNA of putative *M. tuberculosis Δ pykA* (lanes 1-21), plasmid control (lane 22), *M. tuberculosis* H37Rv (lane 23) & lanes M –1 kb DNA ladder.

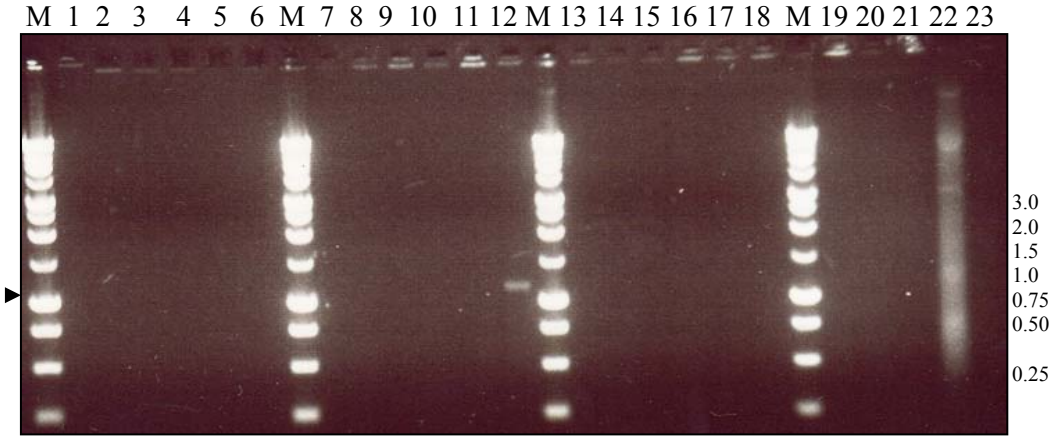


Fig. 50b - PCR verification of downstream region with antibiotic resistance cassette from the genomic DNA of putative *M. tuberculosis Δ pykA* (lanes 1-21), plasmid control (lane 22), *M. tuberculosis* (lane 23) & lanes M - 1 kb DNA ladder.

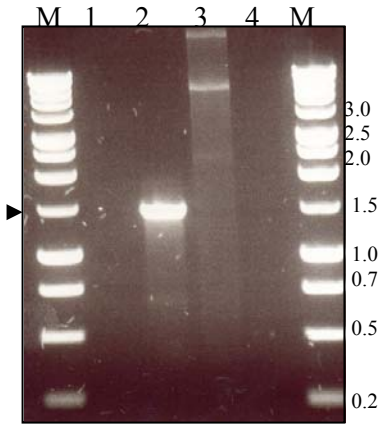


Fig. 50c - PCR verification by amplification of *pykA* from putative *M. tuberculosis* Δ *pykA* mutant (lane 1), wild type (lane 2), plasmid control (lane 3), negative control (lane 4) & lanes M - 1 kb DNA ladder.

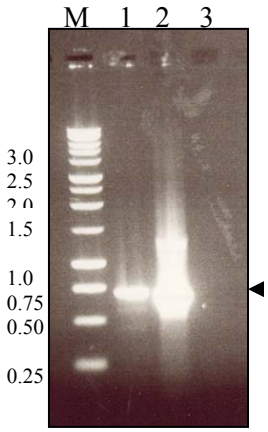


Fig. 50d - PCR verification by amplification of hygromycin cassette from putative *M. tuberculosis*Δ *pykA* mutant (lane 1) plasmid control (lane 2), wild type (lane 3) & lanes M - 1 kb DNA ladder.

4.3.1.3. *M. tuberculosis Δ ald Δ pykA*

Double knock-out mutant *M. tuberculosis Δ pykA Δ ald* was generated by introducing pCM2 into *M. tuberculosis Δ ald* and the mutants were initially screened by antibiotic selection. The colonies were further analysed by Hotstart PCR for *kanR* (Fig. 51c) and for the absence *pykA* gene in the genome as given in section 4.3.1.2 (Fig. 51e). Absence of functional *ald* gene in knock-out was confirmed as previously explained in Section 4.3.1.1 (Fig. 51d).

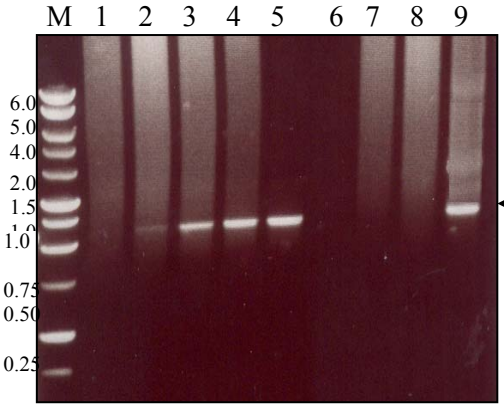


Fig. 51a - PCR verification of upstream region with antibiotic resistance cassette from the genomic DNA of putative *M. tuberculosis Δ pykA Δ ald* (lanes 1-5), *M. tuberculosis Δ pykA* (lane 6), *M. tuberculosis Δ ald* (lane 7), *M. tuberculosis* (lane 8), positive control (lane 9) & lanes M - 1 kb DNA ladder.

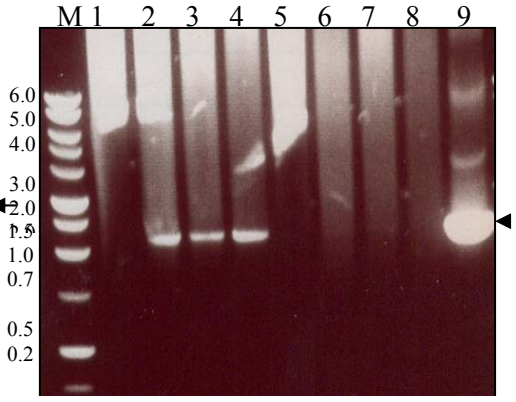


Fig. 51b - PCR verification of downstream region with antibiotic resistance cassette from the genomic DNA of putative *M. tuberculosis Δ pykA Δ ald* (lanes 1-5), *M. tuberculosis Δ pykA* (lane 6), *M. tuberculosis* Δ *ald* (lane 7), *M. tuberculosis* (lane 8), positive control (lane 9) & lanes M - 1 kb DNA ladder.

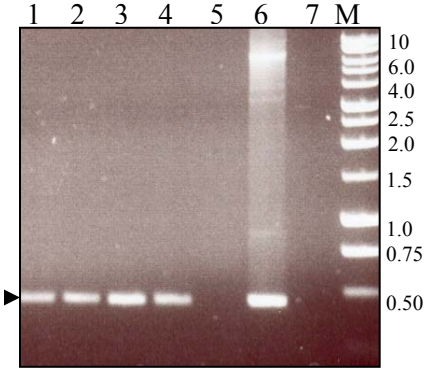


Fig. 51c - PCR verification by amplification of kanamycin cassette from putative mutants (lanes 1-4), *M. tuberculosis Δ pykA* (lane 5), *M. tuberculosis Δ ald* (lane 7), positive control (lane 6) & lanes M - 1 kb DNA ladder.

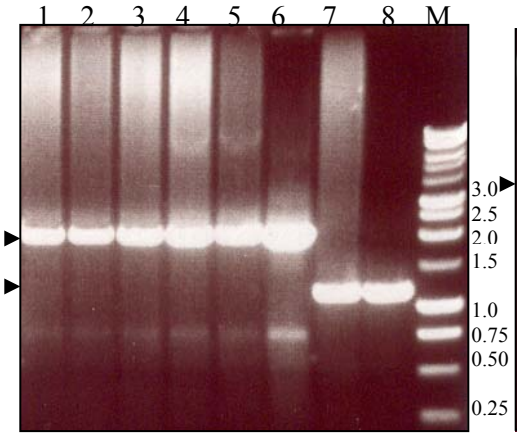


Fig. 51d - PCR verification by amplification of *ald* from putative mutants (lanes 1-4), *M. tuberculosis Δ ald* (lane 5), plasmid control (lane 6), *M. tuberculosis Δ pykA* (lane 7), wild type (lane 9) & lanes M - 1 kb DNA ladder.

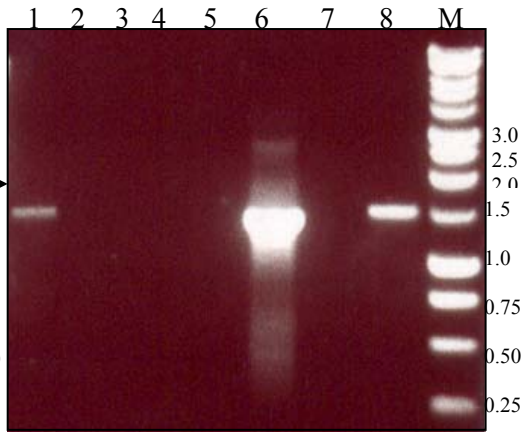


Fig. 51e - PCR verification by amplification of *pykA* from putative mutants (lanes 1-4), *M. tuberculosis Δ pykA* (lane 5), *M. tuberculosis Δ ald* (lane 6), plasmid control (lane 7), wild type (lane 8) & lanes M - 1 kb DNA ladder.

4.3.2. *M. tuberculosis Δ pykA* / pLK102

Plasmid pLK102 (a kind gift from Dr. Stephen Mitchell), containing *pykA* gene from *M. tuberculosis,* was used to complement *M. tuberculosis Δ pykA,* in order to study the effect of *pykA* gene. Colonies were analysed by restriction digestion with *BamHI* for the presence of plasmid in *M. tuberculosis Δ pykA* (Fig. 52).

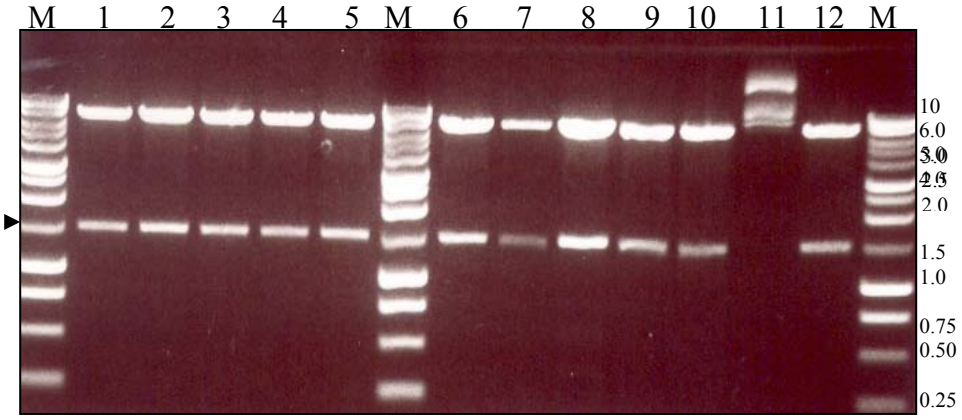


Fig. 52 - Verification of inserts by restriction digest with *BamHI.* Lanes 1-12 represents putative recombinants & lanes M -1 kb DNA ladder.

Vectors used for the construction of plasmid vector were also transformed into mycobacteria to assess the effect of expression vector on mycobacteria. *Mycobacterium bovis* was transformed with pSM96 (Fig. 53) and *M. tuberculosis Δ pykA* with pSM81 (Fig. 54).

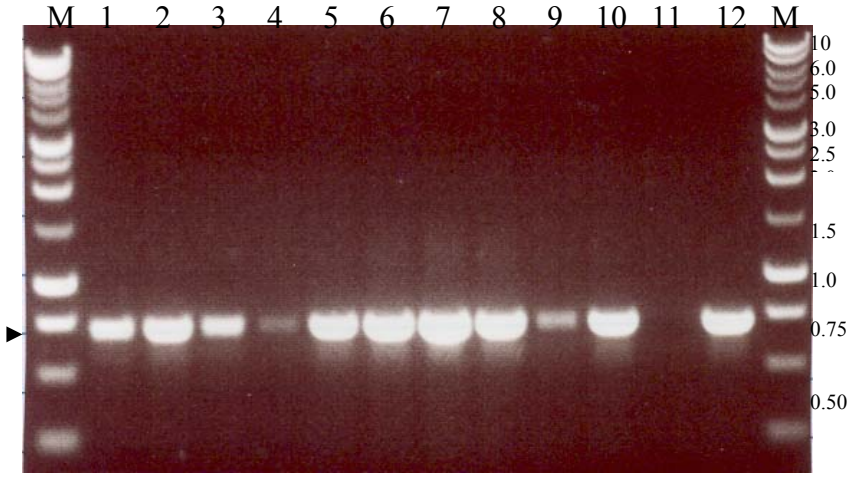


Fig. 53 - PCR verification by amplification of Hygromycin cassette from putative *M. bovis* mutants (lanes 1-10); wild type (lane 11); plasmid control (lane 12); lanes M - 1 kb DNA ladder.

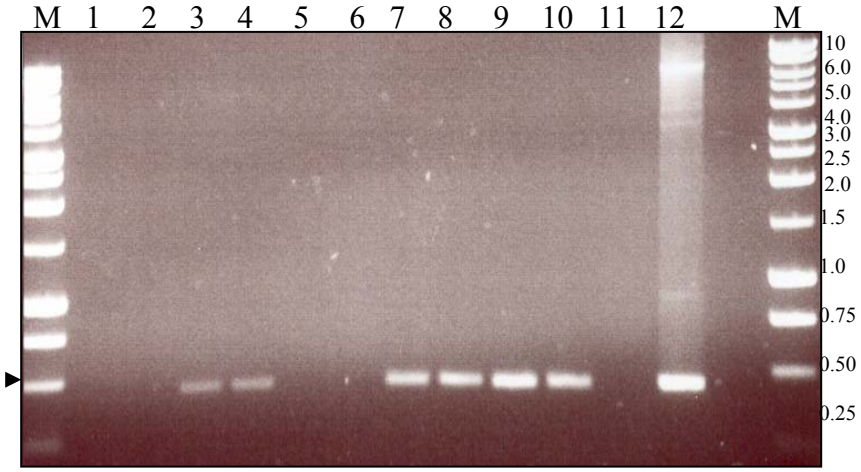


Fig. 54 - PCR verification by amplification of kanamycin cassette from putative mutants (lanes 1-10); plasmid control (lane 12); wild type (lane 11); negative control (lane 13); lanes M - 1 kb DNA ladder.

4.4. Activities of L-alanine dehydrogenase and Pyruvate kinase in the mycobacterial mutants

4.4.1. Pyruvate kinase

Notable pyruvate kinase activity was seen in wild type *M. tuberculosis* H37Rv and with no detectable activity in *M. bovis* T-35 (Fig. 55). The enzyme activity in the mutants showed a 10 to 12 fold increase in the PK activity in *M. tuberculosis Δ pykA /* pLK102 and *M. bovis /* pMS2 (mutants complemented with plasmids containing *pykA* gene) when compared to *M. tuberculosis.* No activity was detected in *M. tuberculosis Δ pykA* and *M. tuberculosis Δ pykA Δ ald* (Fig. 55).

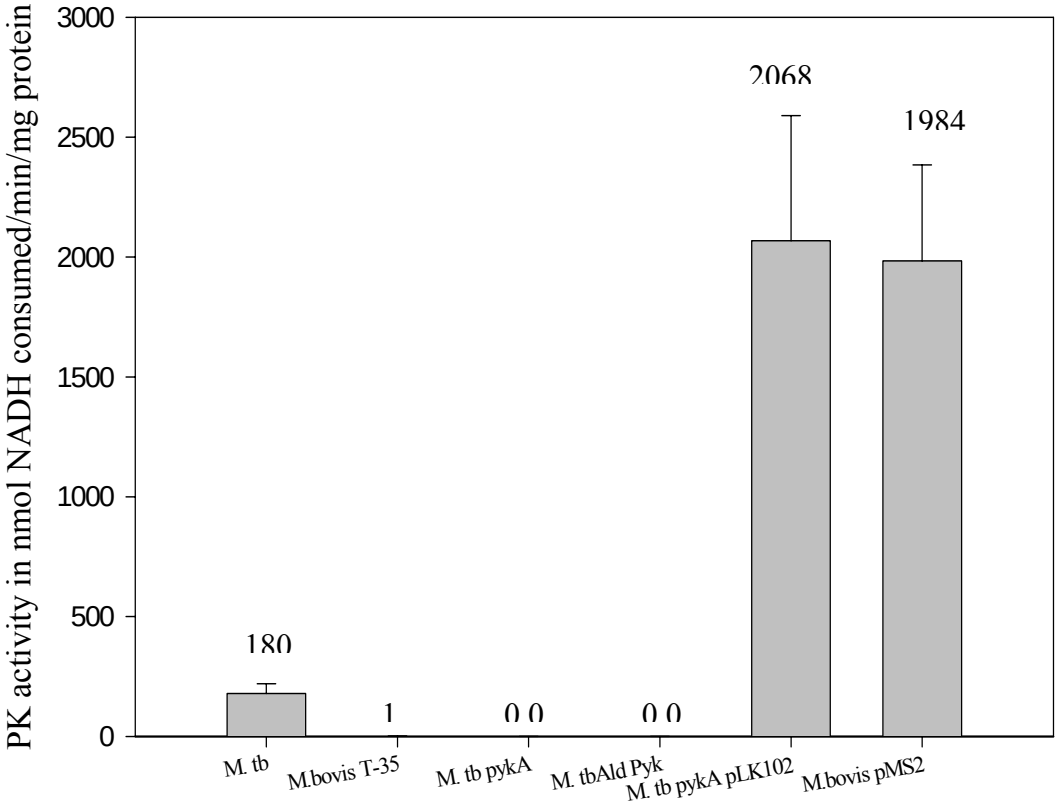


Fig. 55 - Analysis of PK activity in various mutants and wild type strains. All values are expressed in nmol NADH consumed min^{-1} mg^{-1} protein ± 95% confidence limits. Protein extracts from at least three independent mid-log cultures were assayed for pyruvate kinase activity by the method adapted by Wheeler (Wheeler, 1983). At least two determinations of activity were carried out on each protein extract. Two negative reactions were always performed one without the substrate and other without the enzyme.

4.4.2. L-alanine dehydrogenase

Ald activity was determined in *M. tuberculosis Δ ald, M. tuberculosis Δ pykA Δ ald, M. bovis* BCG / pSR1, *M. bovis /* pSR1 and *M. bovis /* pMS2, *M. tuberculosis* and *M. bovis* serving as positive and negative controls respectively. Ald was expressed in *M. tuberculosis* but not in *M. bovis.* The analysis also showed 2 fold increase in the Ald activity when complemented with plasmids containing *ald* gene (*M. bovis* BCG / pSR1 and *M. bovis /* pMS2) when compared with wild type (*M. tuberculosis*). But owing to unknown reasons, no activity was detected in *M. bovis /* pSR1. No activity was detected in *M. tuberculosis Δ ald* and *M. tuberculosis Δ pykA Δ ald* and in *M. bovis* T-35 (Fig. 56).

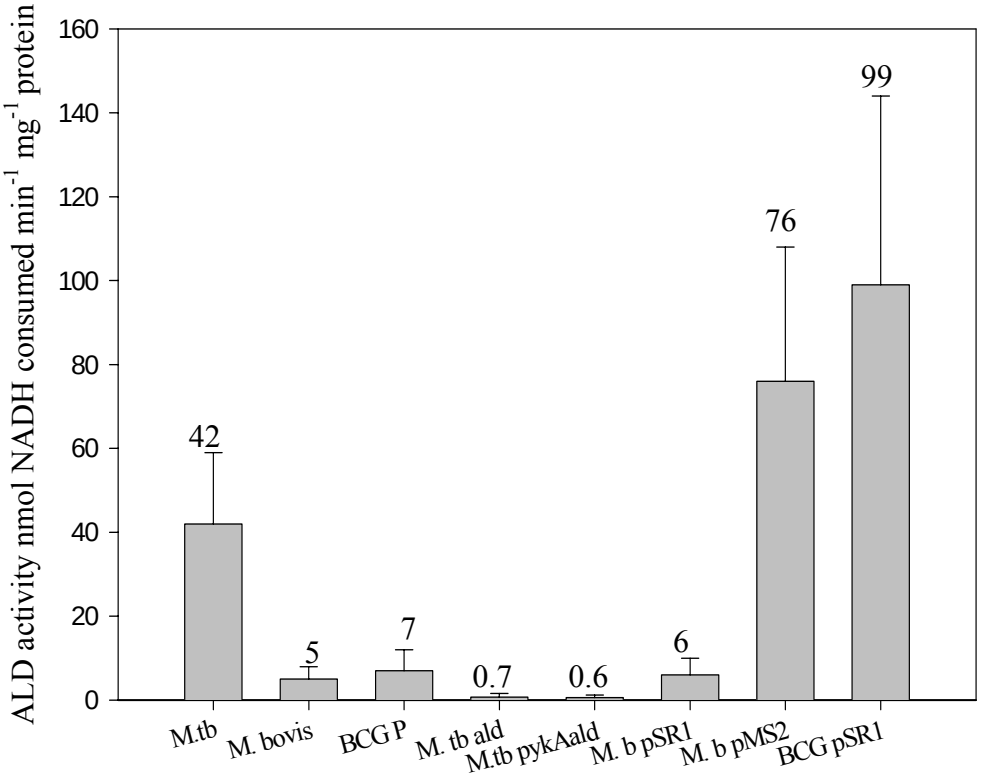


Fig. 56 - Analysis of Ald activity in various mutants and wild type strains. All values are expressed in nmol NAD^+ formed min-1 mg-1 protein ± 95% confidence limits. Protein extracts from at least three independent mid-log cultures were assayed for L-alanine dehydrogenase activity by the method adapted by Anderson. At least two determinations of activity were carried out on each protein extract. Two negative reactions were always performed one without the substrate and other without the enzyme.

4.5. Carbon source utilisation

To confirm the disruption of glycolysis by the *pykA* lesion and to identify the function of *ald* in carbon metabolism in *M. tuberculosis* complex, the various genetically manipulated strains were grown in minimal Sauton's medium with different carbon sources. As expected, when glycerol was used as the sole carbon source, *M. tuberculosis* H37Rv showed abundant growth, while *M. bovis* growth was sparse. There were no significant changes in growth between *M. tuberculosis Δ ald* and *M. tuberculosis* H37Rv, while *M. tuberculosis Δ pykA* and *M. tuberculosis Δ pykA Δ ald* exhibited considerable reduction in the growth, but not to the extent of reduction as seen in *M. bovis* T-35. Less growth was observed in *M. tuberculosis Δ pykA Δ ald* when compared to *M. tuberculosis Δ pykA* (Fig. 57).

There was little difference in growth when pyruvate (Fig. 58) or Tween 80 (Fig. 59) were used as sole carbon source. *Mycobacterium tuberculosis Δ pykA* that showed reduction in the growth in the glycerol medium was able to grow equal to the wild type when pyruvate was used in the medium. *Mycobacterium bovis* T-35 was able to grow in the pyruvate containing medium. Similar effects were observed when Tween-80 was used as the sole carbon source. There was no growth observed in any of the strains when glucose was used as the sole carbon source (Fig. 60), showing the inability of mycobacteria to utilise glucose.

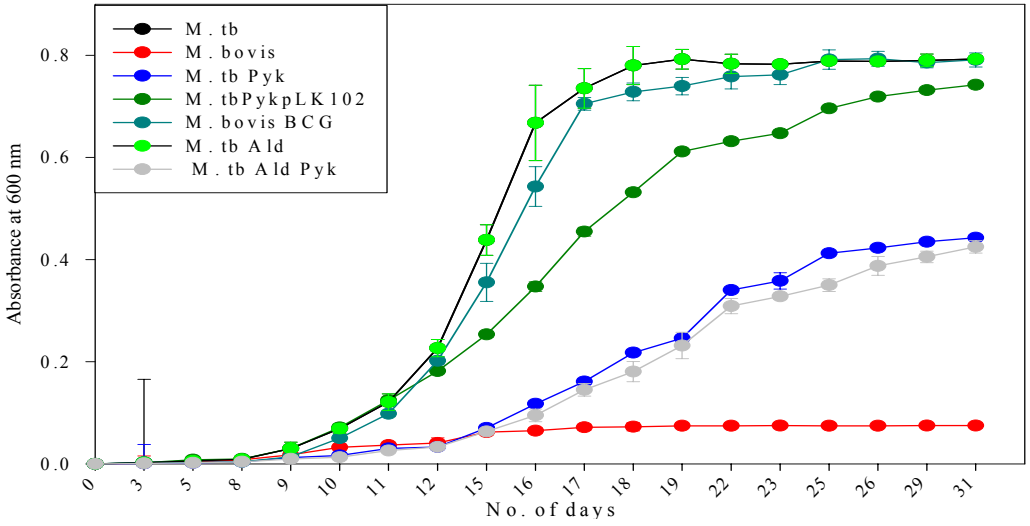


Fig. 57 - Effect of glycerol as sole carbon source on growth of various mutants grown in Sauton's medium.

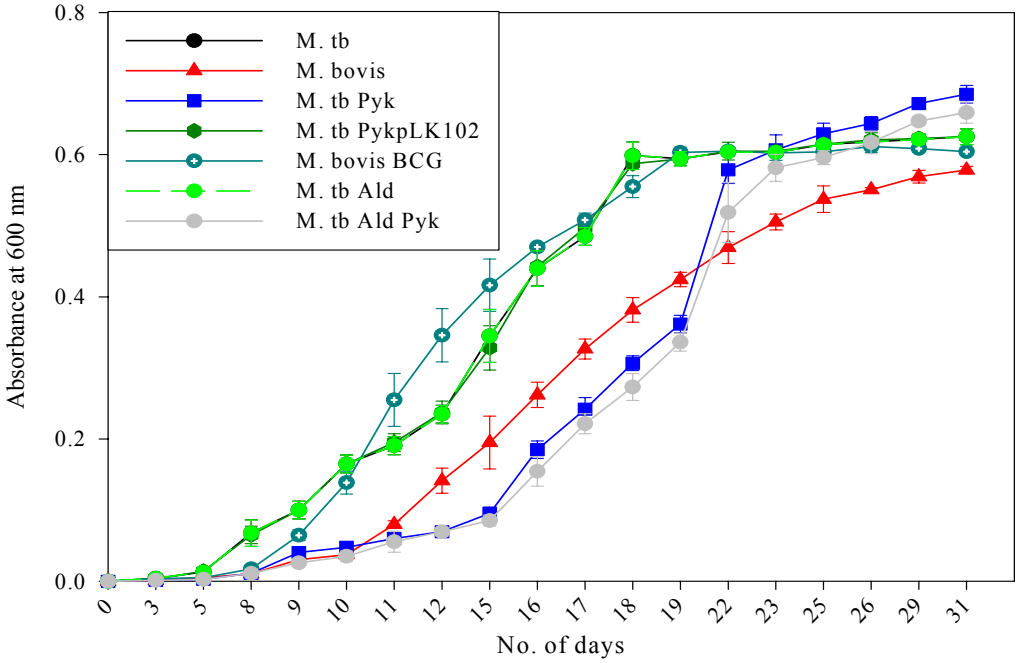


Fig. 58 - Effect of pyruvate as a sole carbon source on growth of various mutants grown in Sauton‘s medium.

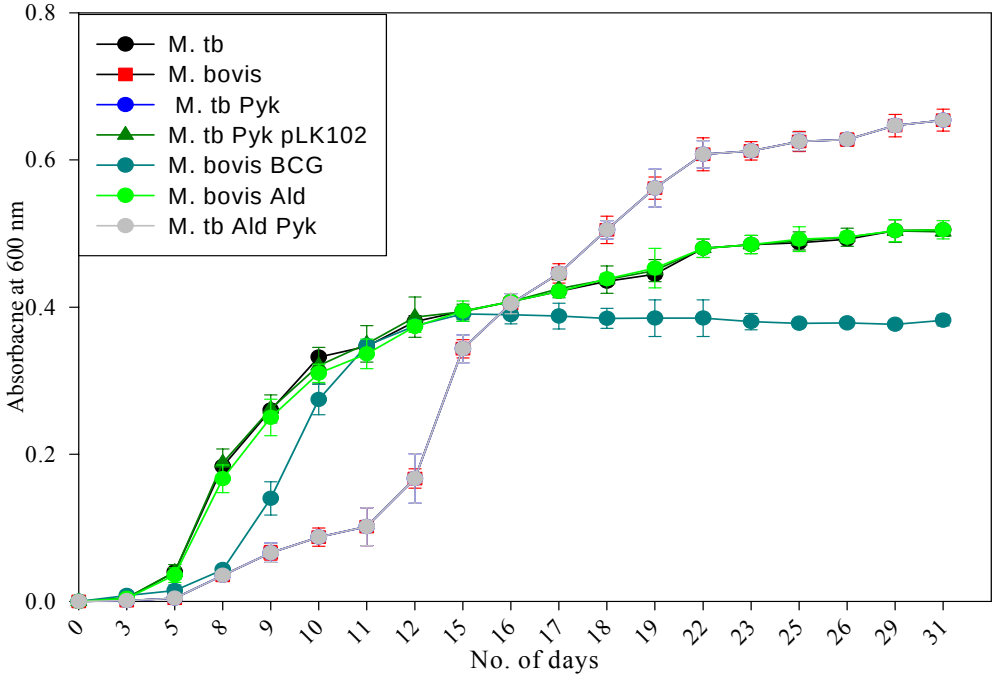


Fig. 59 - Effect of Tween-80 as a sole carbon source on growth of various mutants grown in Sauton‘s medium.

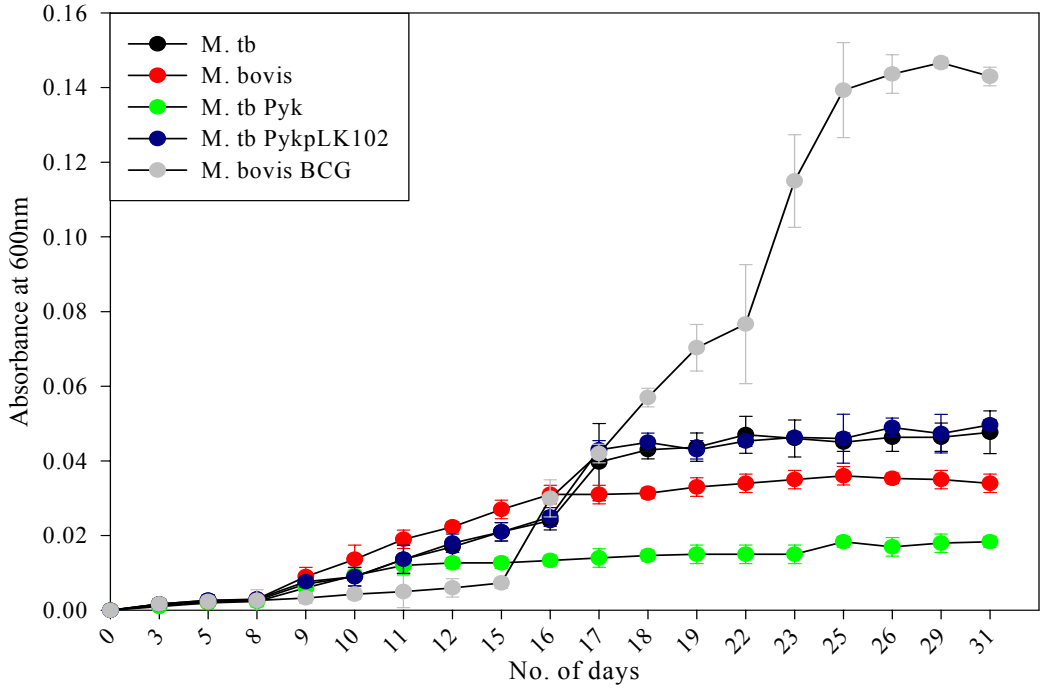


Fig. 60 - Effect of glucose as a sole carbon source on growth of various mutants grown in Sauton‘s medium.

4.6. Colony Morphology

Effect of various carbon sources on colony morphology was studied by growing the mutants on 7H11 containing either glycerol or pyruvate or without any carbon source except OADC. Considerable changes in the morphology were observed in *M. tuberculosis Δ pykA* when glycerol was added to OADC containing 7H11 medium (Fig. 61 & 62). The colony morphology was restored back to eugonic, rugose morphotype when *pykA* was restored in *M. tuberculosis Δ pykA* / pLK102. Role of various carbon sources in colony formation is explained in Table 21. All the strains when grown in pyruvate as the additional carbon source along with OADC showed eugonic growth, confirming the role of pyruvate in colony formation in the various

members of *M. tuberculosis* complex (Fig. 63 & 64). When only OADC was used as the sole carbon source all the strains showed decreased growth (Fig. 65 & 66).

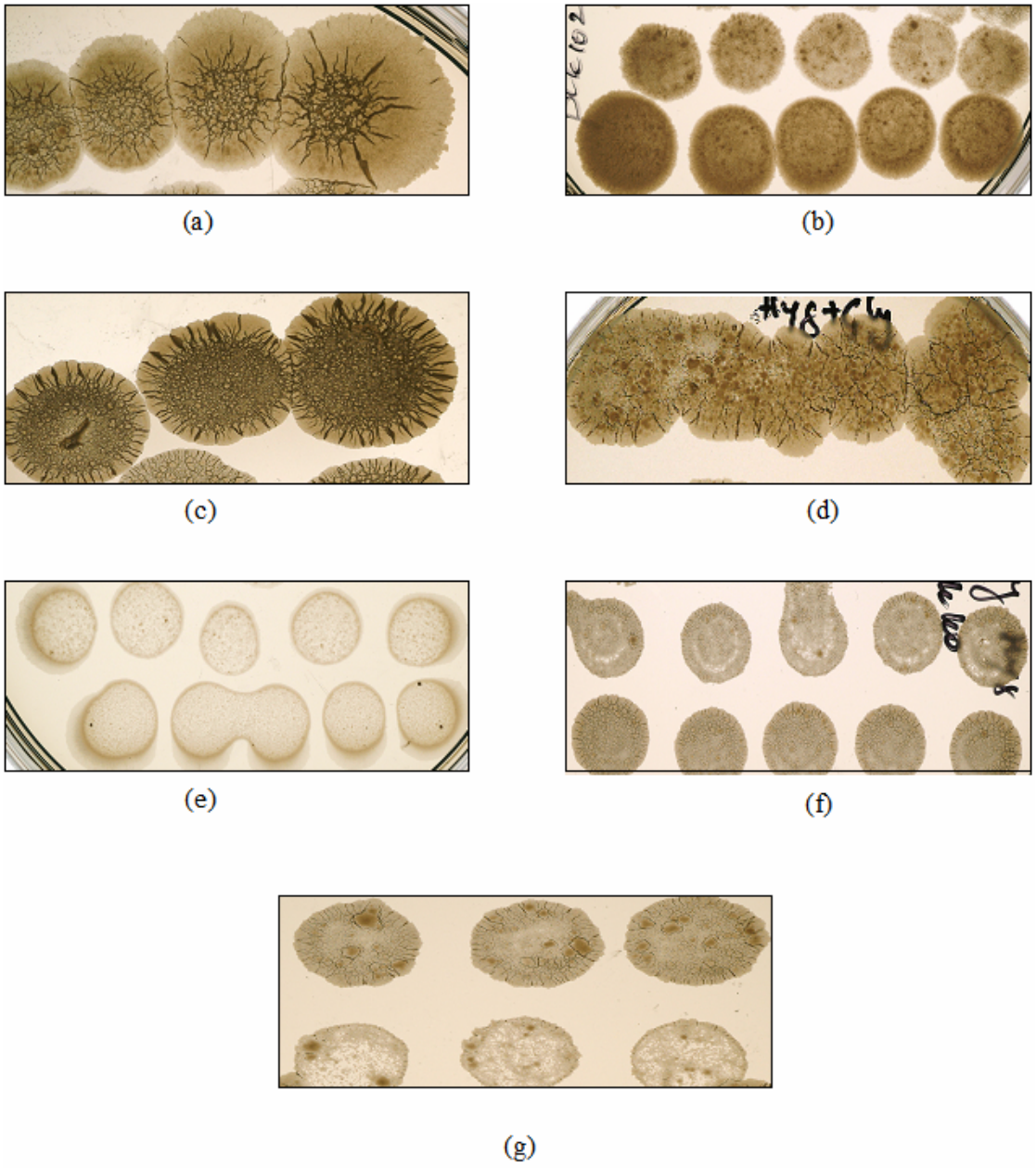


Fig. 61 - Effect of glycerol on the growth and morphology of mycobacteria. Panel a, b, c, d, e, f and g represents *M. tuberculosis, M. bovis* / pLK102*, M. tuberculosis Δ pykA* / pLK102*, M. tuberculosis Δ ald, M. bovis, M. tuberculosis Δ pykA* and *M. tuberculosis Δ pykA Δ ald* respectively.

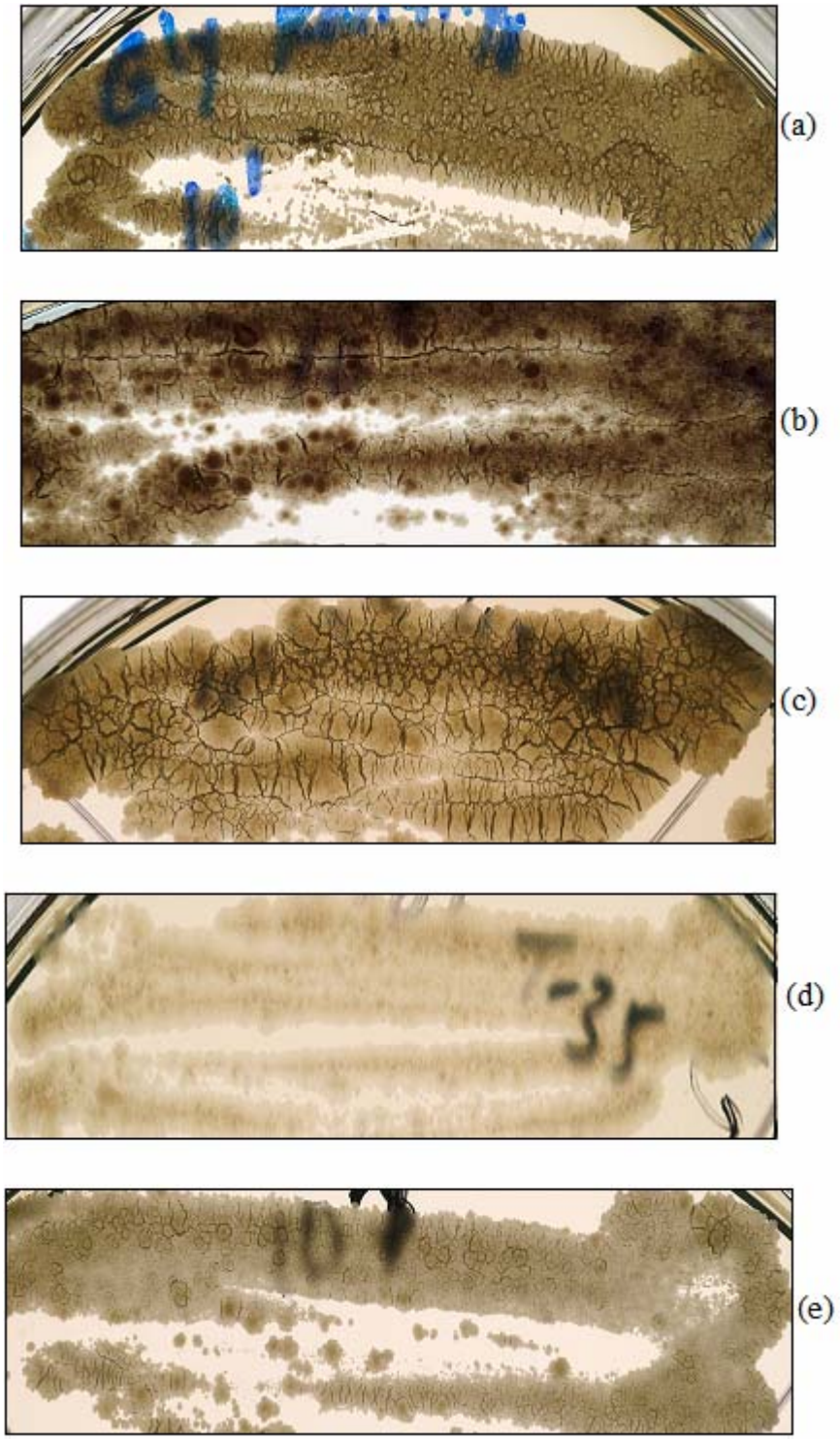


Fig. 62 - Effect of glycerol on the growth and morphology of mycobacteria. Panel a, b, c, d and e represents *M. tuberculosis, M. bovis* / pLK102*, M. tuberculosis Δ pykA* / pLK102*, M. bovis* and *M. tuberculosisΔ pykA* respectively.

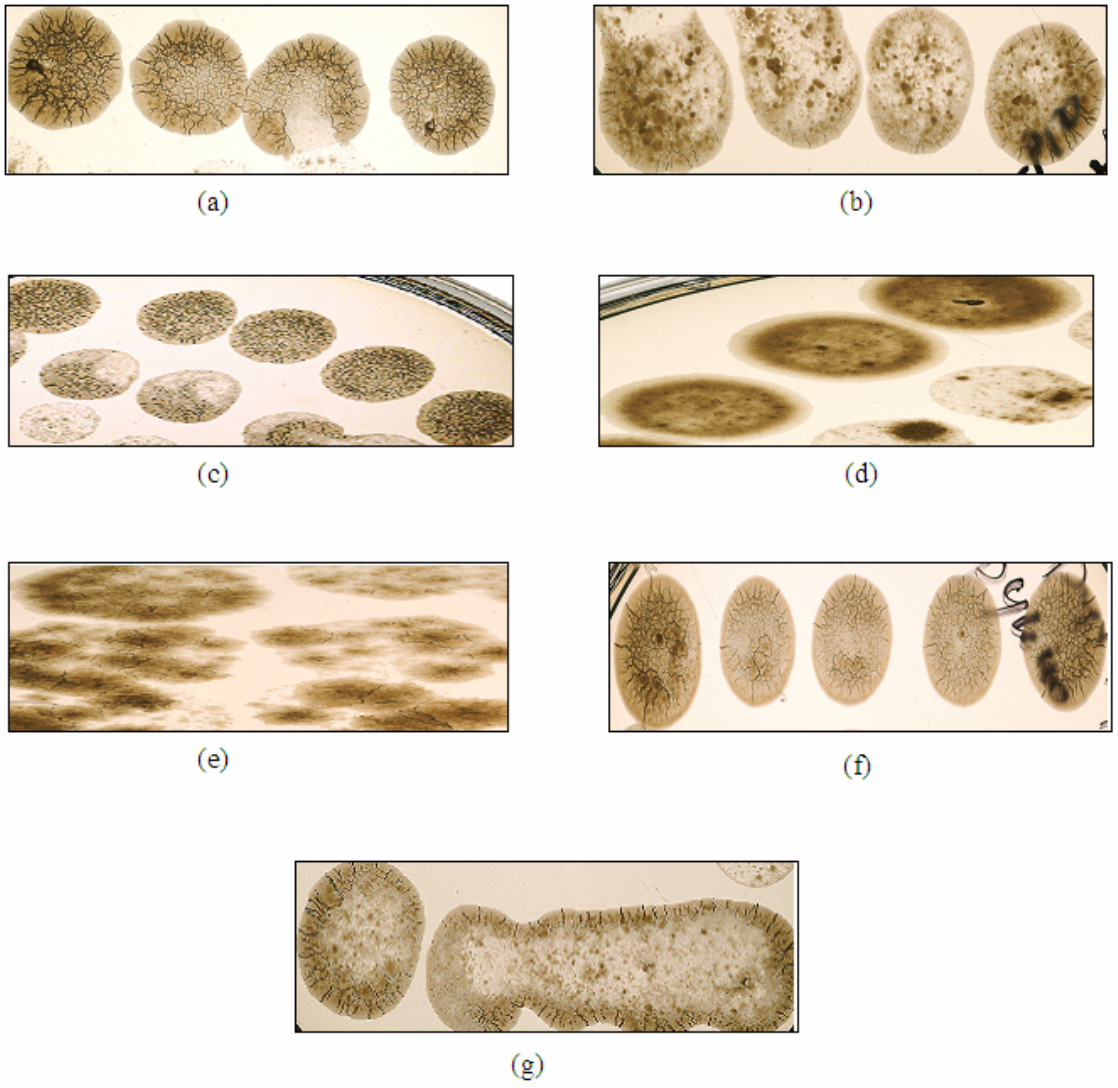


Fig. 63 - Effect of pyruvate on the growth and morphology of mycobacteria. Panel a, b, c, d, e, f and g represent *M. tuberculosis, M. tuberculosis Δ ald, M. tuberculosis Δ pykA /* pLK102*, M. bovis /* pLK102*, M. bovis, M. tuberculosis Δ pykA* and *M. tuberculosis Δ pykA Δ ald* respectively.

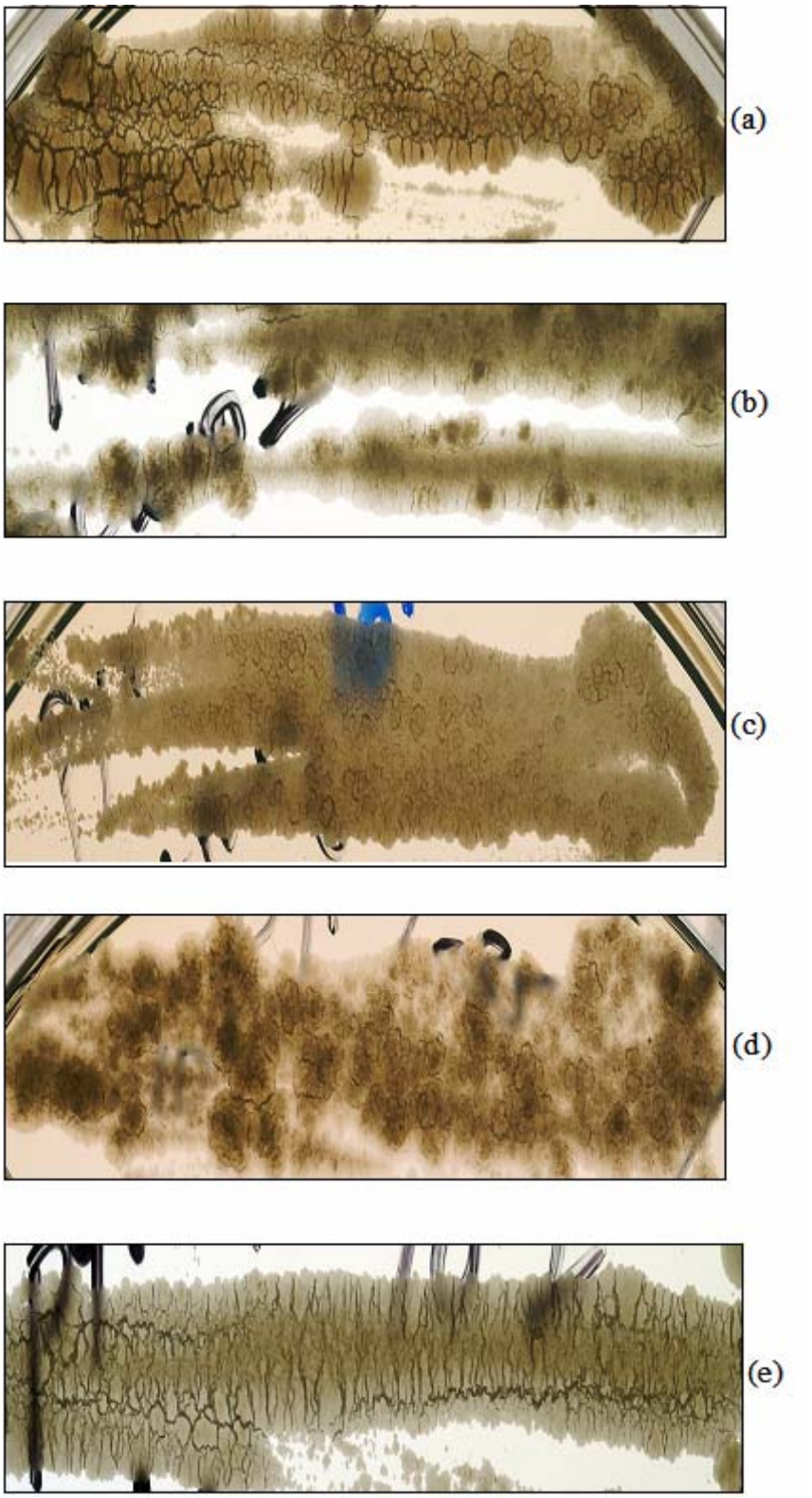


Fig. 64 - Effect of pyruvate on the growth and morphology of mycobacteria. Panel a, b, c, d and e represents *M. tuberculosis, M. bovis /* pLK102*, M. tuberculosis Δ pykA /* pLK102*, M. bovis* and *M. tuberculosis Δ pykA* respectively.

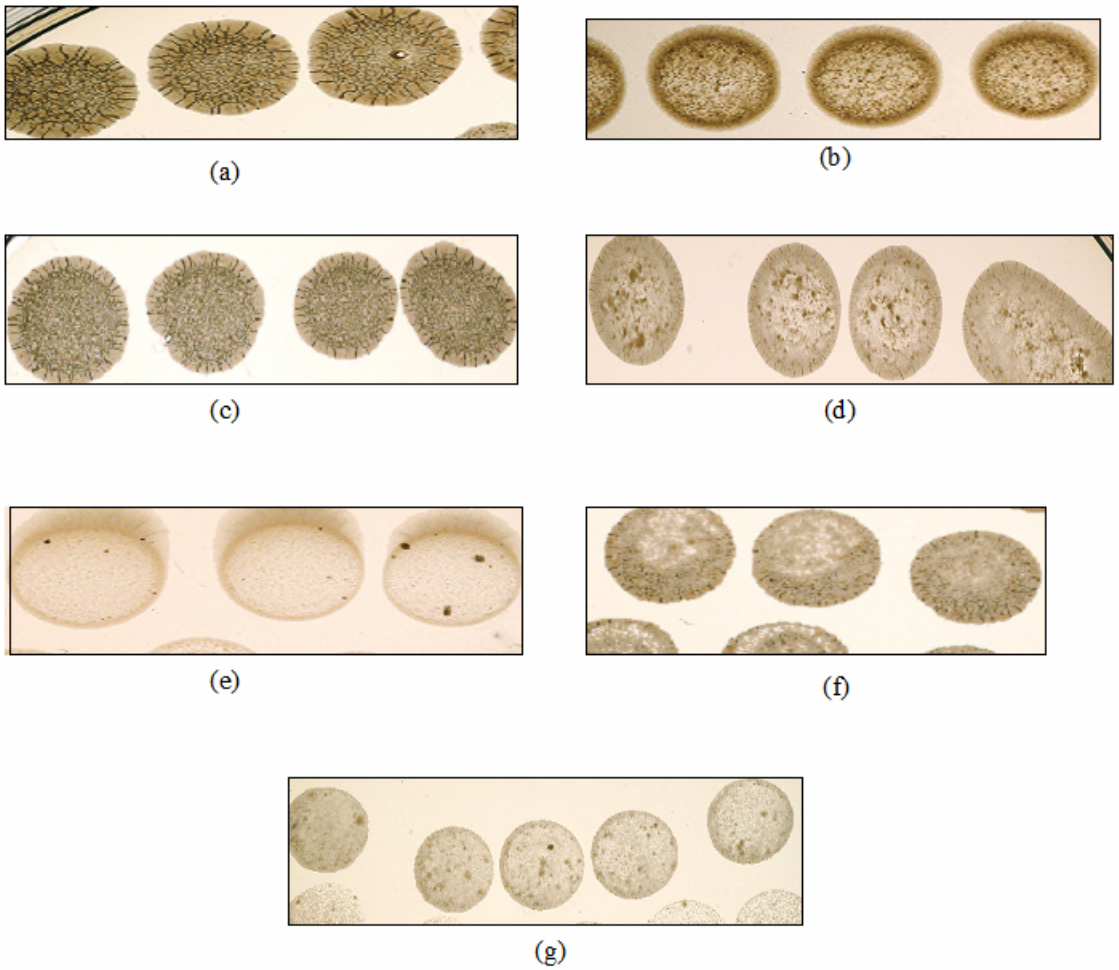


Fig. 65 - Effect of OADC on the growth and morphology of mycobacteria. Panel a, b, c, d, e, f and g represents *M. tuberculosis, M. bovis /* pLK102*, M. tuberculosis Δ pykA /* pLK102*, M. tuberculosis Δ ald, M. bovis, M. tuberculosis Δ pykA* and *M. tuberculosis Δ pykA Δ ald* respectively.

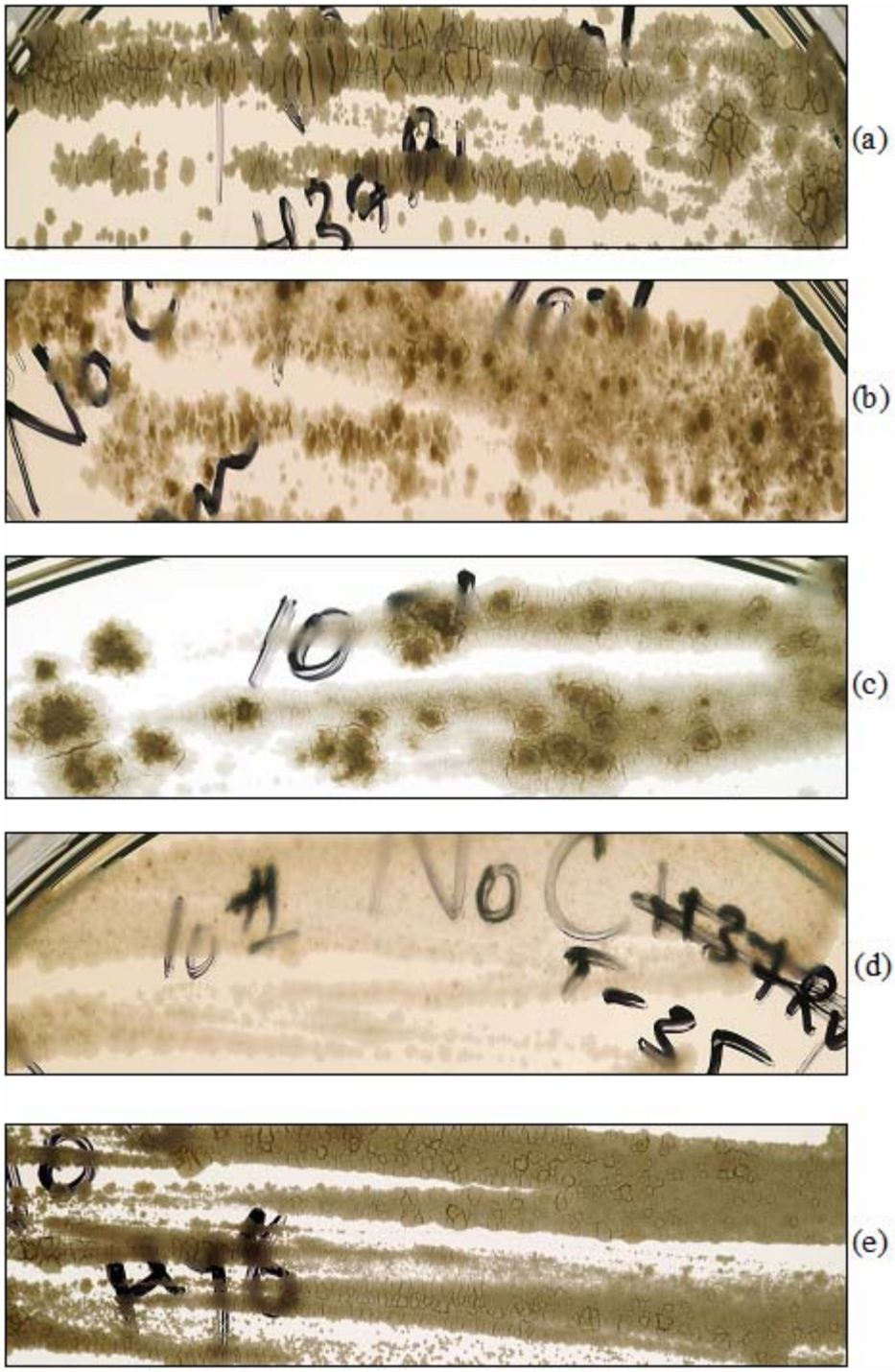


Fig. 66 - Effect of OADC on the growth and morphology of mycobacteria. Panel a, b, c, d and e represents *M. tuberculosis, M. bovis* / pLK102*, M. tuberculosis Δ pykA* / pLK102*, M. bovis* and *M. tuberculosis Δ pykA* respectively.

Table 21 - Colony morphology

Name of the strain	Glycerol + OADC		Pyruvate + OADC		Only OADC	
	Eugonic	Rugose	Eugonic	Rugose	Eugonic	Rugose
M. tuberculosis	+++	+++	++	++	+	+
M. bovis	--	--	++	++	--	--
M. tuberculosis Δ pykA	+	+	++	++	+	+
M. tuberculosis Δ ald	+++	+++	++	++	+	+
M. tuberculosis Δ pykA Δ ald	+	+	++	++	-	-
M. tuberculosis Δ pykA / *pLK102*	+++	+++	++	++	+	+
M. bovis */ pLK102*	++	++	++	++	+	+

The above table explains the effect of various carbon sources in the colony morphology, where '+++' indicates eugonic and rugose colonies while '-' indicates the dysgonic and smooth morphology.

4.7. Surface lipid analysis

4.7.1. Analysis of non-polar and polar lipids

In order to explain the morphotype change due to *pykA*, and the effect of *ald* in the production of lipids, we studied the changes in the lipid profile in various mutants. Based on the analysis of the various non-polar lipids, there is a significant level of decrease in triacylglycerol (System A, Fig. 67) and some unknown free fatty acids (System C, Fig. 68) in *M. tuberculosis Δ pykA*. The production of fatty acid was restored when *pykA* was complemented in *M. tuberculosis Δ pykA*. Analysis of polar lipids (Fig. 69) showed that *M. tuberculosis Δ pykA Δ ald* resembles *M. bovis* in phosphatidylinositol production. The overall polar lipid production was increased in *M. bovis /* pLK102 when compared to *M. bovis.* Similar observation were seen in *M. tuberculosis Δ pyk /* pLK102 when compared to *M. tuberculosis Δ pykA*. There were no significant differences in the polar lipid production between *M. tuberculosis, M. tuberculosis Δ pykA* and *M. tuberculosis Δ ald.*

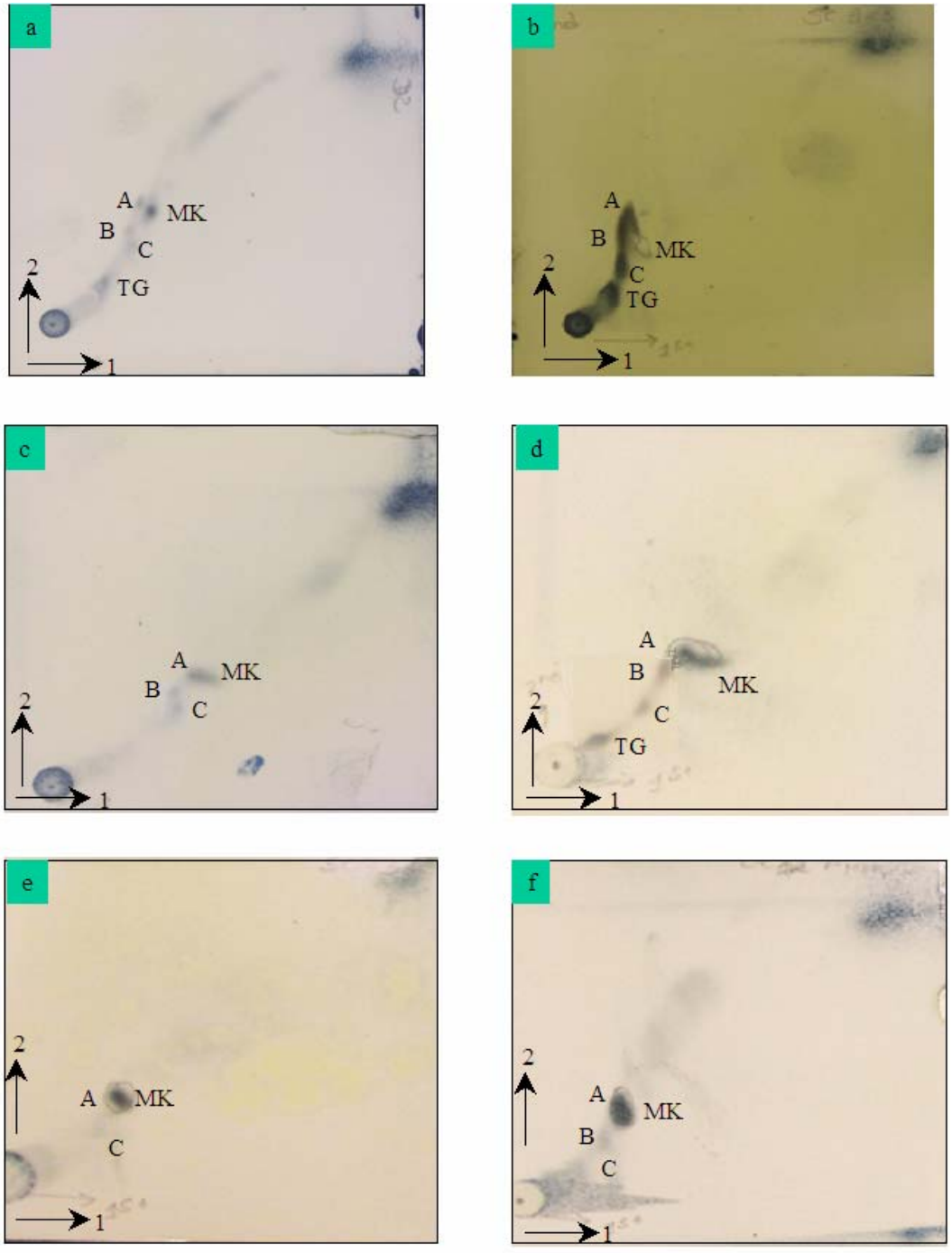


Fig. 67 - 2D-TLC of non-polar lipids with System A. First direction, Hexane: ethyl acetate (98:2) and second direction, hexane: acetone (98:2), where panel a, b, c, d, e and f represents *M. tuberculosis, M. bovis, M. tuberculosis Δ pykA, M. tuberculosis Δ ald, M. tuberculosis Δ pykA Δ ald* and *M. tuberculosis Δ pykA /* pLK102 respectively. Abbreviations A-C, dimycocerosates (III) of the phthiocerol family (IV-VI); MK, menaquinone; TG, triacylglycerol.

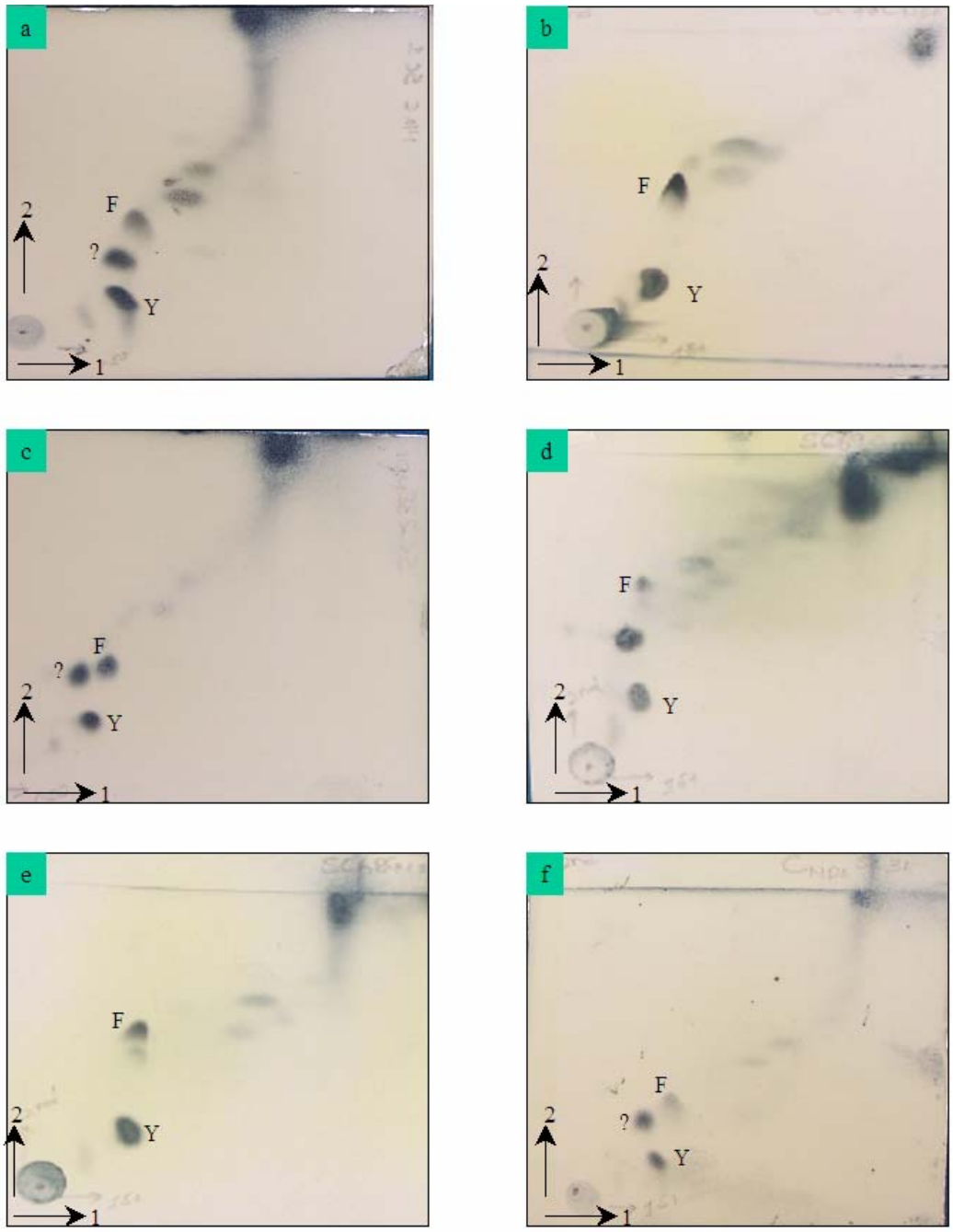


Fig. 68 - 2D-TLC of non-polar lipids with System C. First direction, Chloroform: methanol (96:4) and second direction, toluene: acetone (80:20), where panel a, b, c, d, e and f represents *M. tuberculosis, M. tuberculosis Δ ald, M. tuberculosis Δ pykA /* pLK102, *M. bovis, M. tuberculosis Δ pykA Δ ald* and *M. tuberculosis Δ pykA* respectively. Abbreviations: F, free fatty acids, Y and ?, unknowns

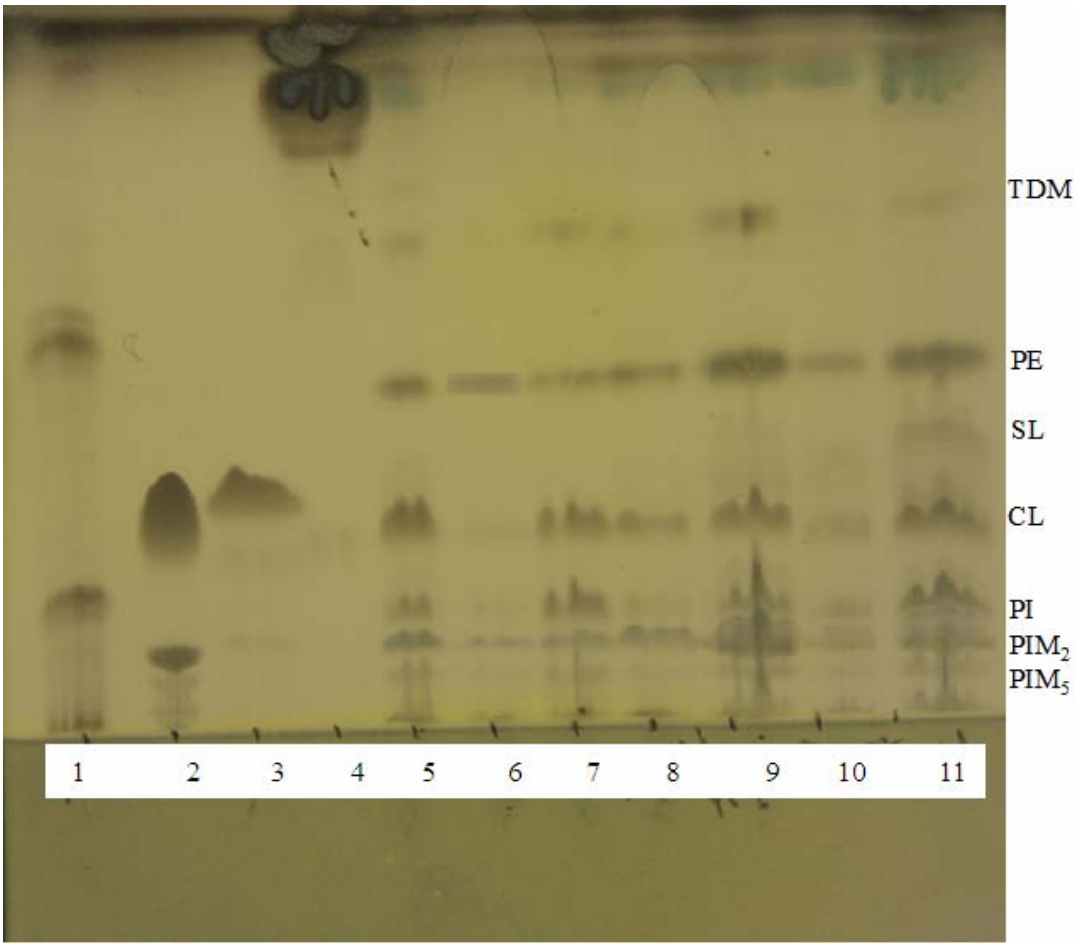


Fig. 69 - Analysis of polar lipid fraction by single dimensional HPTLC. Lanes 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12 represents PE (Phosphatidylethanolamine), PI (Phosphatidylinositol), CL (Cardiolipin), TDM (trehalose dimycolate), *M. tuberculosis, M. bovis, M. tuberculosis Δ pykA, M. tuberculosis Δ pykA Δ ald, M. tuberculosis Δ pykA /* pLK102*, M. bovis* / pLK102 and *M. tuberculosis Δ ald* respectively.

4.7.2. Analysis of mycolic acids

TLC analysis of mycolic acids showed difference in the migration between the wild type and mutants (Fig. 70). Similar results were obtained by the complementation of plasmid vector alone (Fig. 71). Hence, it was concluded that the effect is due to the plasmid and not due to the gene of interest.

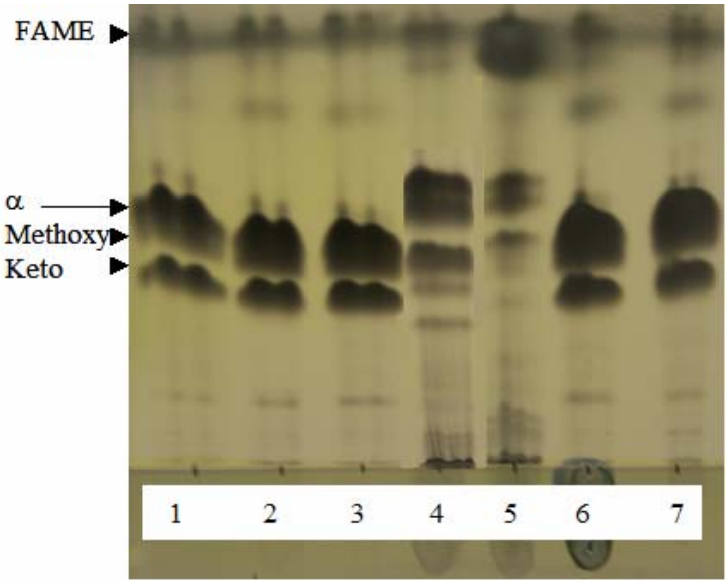


Fig. 70 - Analysis of different types of mycolic acids by HPTLC. Lanes 1, 2, 3, 4, 5, 6 and 7 represents *M. tuberculosis, M. bovis, M. tuberculosis Δ pykA, M. tuberculosis Δ pykA* / pLK102*, M. bovis* / pLK102, *M. tuberculosis Δ ald* and *M. tuberculosis Δ pykA Δ ald* respectively. α, Methoxy and ketomethoxy are various types of mycolic acid present in mycobacteria along with FAME (fatty acid methyl ester).

.

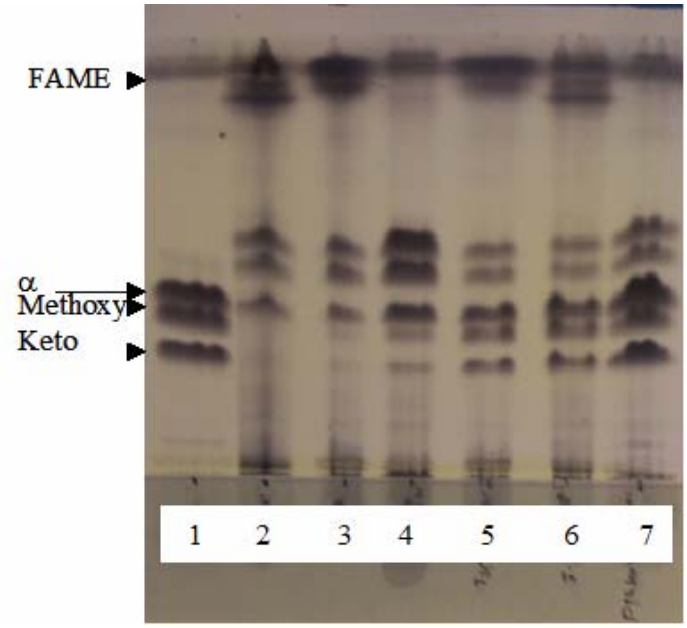


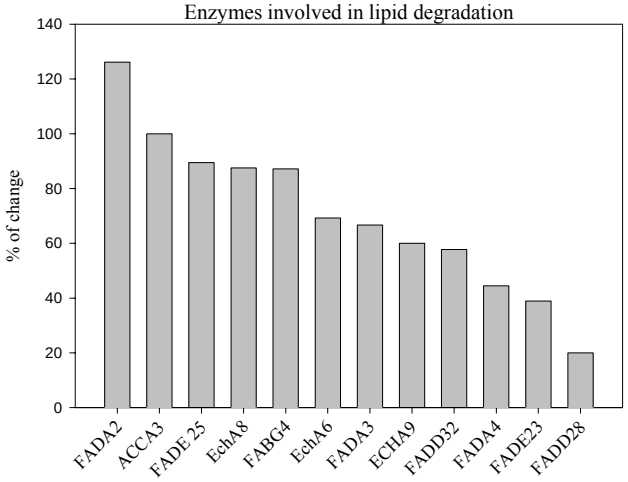
Fig. 71 - Analysis of different mycolic acids by HPTLC to ensure the difference in mycolic acids is due to complementing plasmids. Lanes 1, 2, 3, 4, 5, 6 and 7 represents *M. bovis, M. bovis* / pSM81, *M. bovis* / pLK102*, M. tuberculosis Δ pykA* / pLK102*, M. bovis*+ *M. bovis* / pSM81, *M. bovis*+ *M. bovis* pLK102 and *M. tuberculosis Δ pykA* + *M. tuberculosis Δ pykA* / pLK102. α, Methoxy and ketomethoxy are various types of mycolic acid present in mycobacteria along with FAME (fatty acid methyl ester).

4.8. Overall proteome analysis

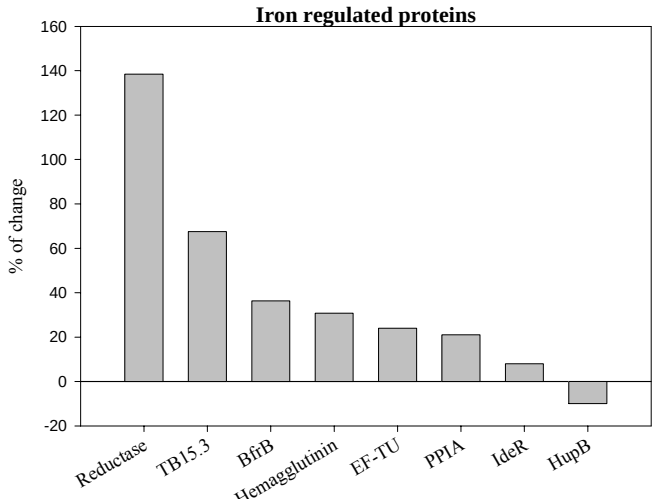
Overall proteome analysis was done to identify the effect of mutation on the protein complement of the mutant organisms. The protein expression pattern was compared between *M. tuberculosis* and *M. tuberculosis Δ pykA, M. bovis* and *M. bovis* pLK102. The analysis between *M. tuberculosis* and *M. tuberculosis Δ pykA* showed that there is increase in the level of expression of proteins involved in lipid degradation, TCA cycle and glyoxalate cycle except isocitrate lyase (Icl) along with the increase in the production of oxidative stress enzymes, iron storage proteins and proteins which are expressed during high iron conditions. There was reduction in the enzymes involved in lipid biosynthesis, phosphoenolpyruvate carboxykinase (PEPCK), Icl and enzymes involved in the biosynthesis of isoleucine and valine. Interestingly, the expression of Ald and HupB (protein which was found to be regulated during iron deficient conditions) were also reduced significantly. Apart from these enzymes several other proteins with unknown function were also found to be up-regulated (Fig. 72).

The proteome of *M. bovis* was compared with *M. bovis /* pLK102, there is a marked increase in the expression of PK, PEPCK and Icl were observed along with a significant increase in the enzymes involved in the synthesis of fatty acids, leucine biosynthesis, PDIM (Phthiocerol dimycocerosate) biosynthesis, and IdeR (Iron dependent regulator). Enzymes involved in the degradation of fatty acids and oxidative stress enzymes were reduced significantly, indicating the role of *pykA* gene in various other pathways (Fig.73).

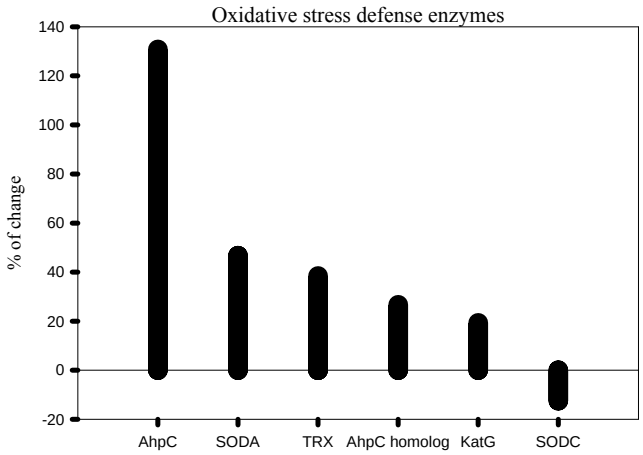
(a)



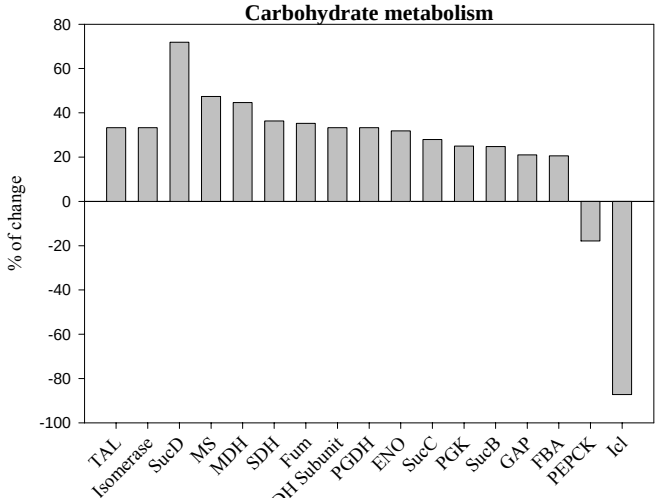
(b)


Iron regulated proteins
% of change
160
140
120
100
80
60
40
20
0
-20
Reductase
TB15.3
BfrB
Hemagglutinin
EF-TU
PPIA
IdeR
HupB

(c)


Oxidative stress defense enzymes
% of change
140
120
100
80
60
40
20
0
-20
AhpC
SODA
TRX
AhpC homolog
KatG
SODC

(d)


Carbohydrate metabolism
% of change
80
60
40
20
0
-20
-40
-60
-80
-100
TAL
Isomerase
SucD
MS
MDH
SDH
Fum
PGDH
ENO
SucC
PGK
SucB
GAP
FBA
PEPCK
Icl

(e)

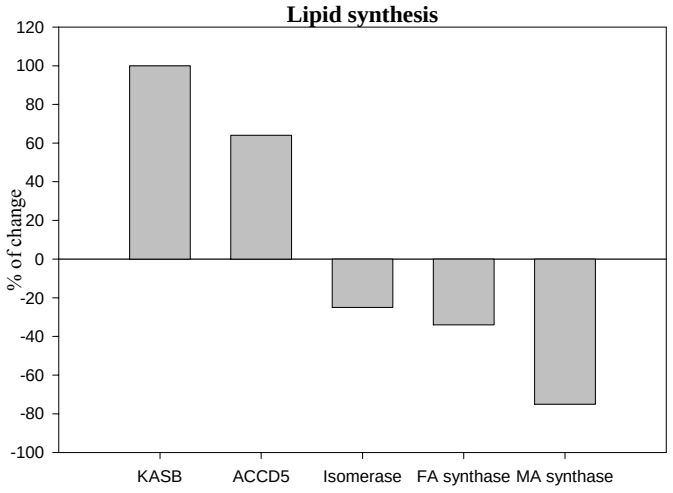
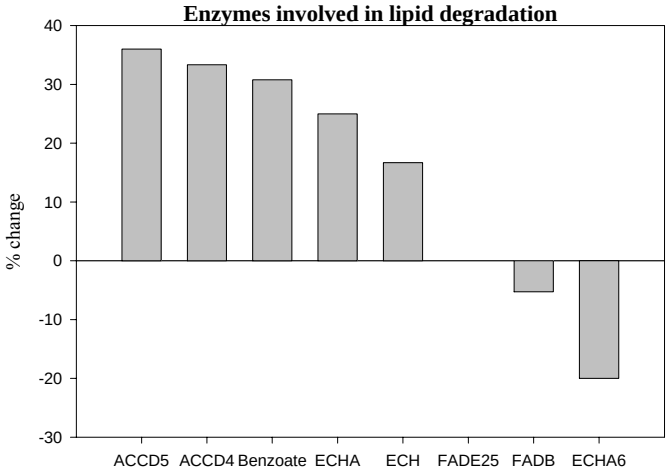
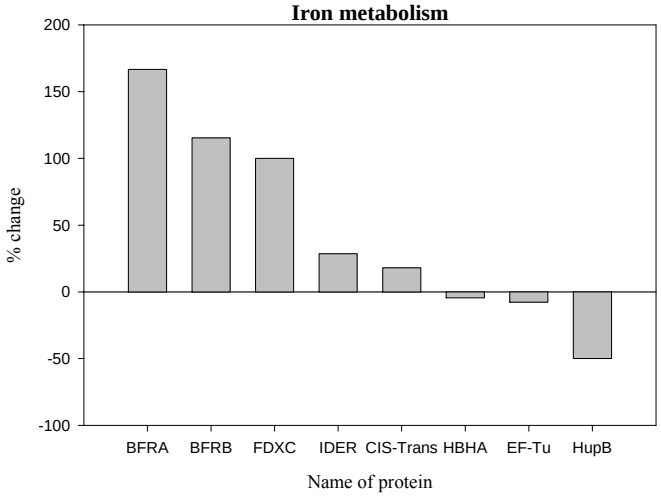


Fig. 72 - Comparison on the expression of different proteins between *M. tuberculosis* and *M. tuberculosis Δ pykA.*

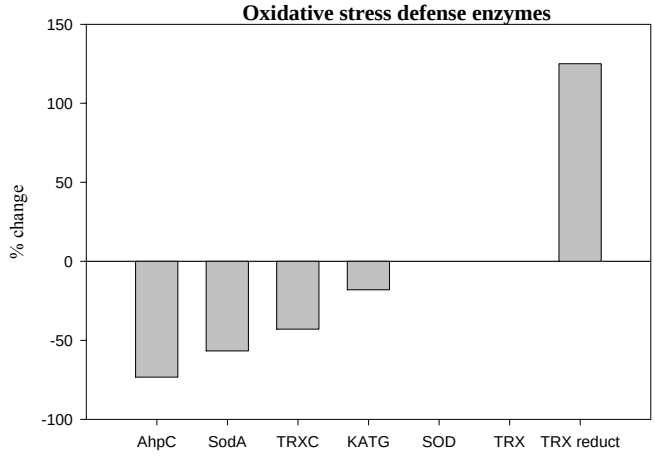
(a)



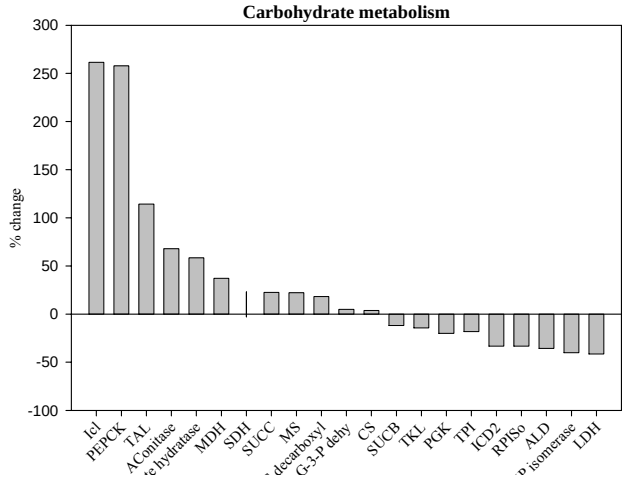
(b)



(c)



(d)



(e)

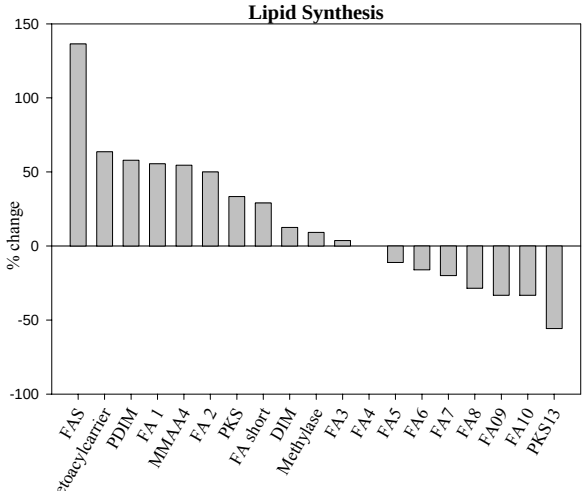


Fig. 73 - Comparison on the expression of different proteins expressed between *M. bovis* and *M. bovis* / pLK102.

4.9. Discussion

The *M. tuberculosis* complex is a group of closely genetically related yet phenotypically diverse strains, enabling their *in vitro* characteristics to differentiate the members of the complex. *Mycobacterium tuberculosis* can utilise glycerol as the sole carbon source, while *M. bovis* cannot. Glycerol utilisation in various members of *M. tuberculosis* complex has pronounced effect on the colony morphology and size of the colony. Two types of colonies are seen, namely the eugonic or abundant growth with rugose or raised crumbly pattern as in *M. tuberculosis,* and smooth, glossy colonies with less growth or dysgonic morphology as in *M. bovis.* Even though *M. bovis* field strains have dysgonic colony morphology, the laboratory-adapted strains *M. bovis* BCG and *M. bovis* AN5 are eugonic (Guerin, 1980; Paterson, 1948), an appearance normally associated with *M. tuberculosis*.

Genetic analysis showed that various genes involved in the central carbon metabolism namely *glpK*, *pykA* and *ald* showed differences across the *M. tuberculosis* complex (Garnier *et al.*, 2003) and may be the reason for their inability to grow on glycerol or glucose as the sole carbon source. Genetic and biochemical examination of glycerol kinase (GK) in a range of tubercle bacilli showed that the absence of GK activity was not the sole reason for its inability to catabolise glycerol (Keating *et al.,* 2005). Hence, emphasis was given to *pykA* that encodes pyruvate kinase (PK).

Pyruvate kinase is a highly conserved ubiquitous enzyme, which is present across prokaryotic and eukaryotic systems. Examining the PK activity across the members of *M. tuberculosis* complex showed that neither *M. bovis* nor *M. microti* has any activity while other members such as *M. tuberculosis, M. bovis* BCG and *M. bovis* AN5 have PK activity (Keating *et al.*, 2005). Several amino acid residues are essential for its catalytic activity, especially Glu_{220}, which is substituted by aspartic acid in *M. bovis* disrupting the integrity of the enzyme. Glutamic acid 220 has been shown to be conserved among 50 pyruvate kinases (Munoz & Ponce, 2003) and forms part of the active site and has been associated with the binding of ADP / ATP, phosphoenol pyruvate (Muirhead *et al*., 1986) and Mg^{2+} (Larsen *et al.,* 1997). SNP at Glu_{220} to Asp_{220} was not limited to *M. bovis* but also found in *M. microti* and *M. africanum,*

which were defined as requiring pyruvate for growth on glycerinated media (Wayne & Kubica, 1986) indicating their inability to catabolise carbohydrates to completion.

Absence of a functional PK results in defunct ATP generation through glycolysis and also via the intermediates being fed into the TCA cycle or any fermentative processes. Even the other functional pentose phosphate pathway (PPP) cannot be used as an alternative pathway as the metabolic intermediates are fed into the lower part of the glycolytic pathway and ultimately to PK. Genomic and biochemical analysis has shown Entner-Doudoroff pathway is not present in pathogenic mycobacteria (Garnier *et al.,* 2003; Ratledge, 1982). Previous study showed that *pykA* gene knock-in mutant in *M. bovis* / pLK102 and *M. bovis* / attBΔI423 utilised glycerol that conferred eugonic growth when glycerol was used as sole carbon source.

In the present study, the role of *pykA* in *M. tuberculosis* was analysed by creating knock-out and re-complementing the knock-out with *pykA* gene. Knock-out and knock-in mutants in mycobacteria are generally created by directed gene inactivation and complementation, respectively. Gene inactivation in slow growing members of mycobacteria is performed by allelic exchange. In directed gene inactivation, an antibiotic resistance cassette is inserted in the middle of the target gene and this construct is transformed into the mycobacteria either as linear or circularised plasmid, using electroporation. Most directed mutations in mycobacteria are performed with bi-functional or shuttle plasmids, which can be maintained in *E. coli.* These are called suicide plasmid vectors, as they lack OriM and cannot replicate in mycobacteria. This is a two step process, in which the plasmid containing the gene of interest disrupted with an antibiotic resistance cassette integrates into the genome by a single cross over event at the region of homology; the second cross over then takes place that excise the plasmid back bone at the other direct repeat than that initially used, when the antibiotic selection is maintained, resulting in the desired gene disruption. The second simpler and faster method involves a one step variation of single crossover. This is achieved by taking advantage of the property of an internal gene fragment carried on a circular DNA to disrupt the corresponding gene when the

DNA integrates into the chromosome by a single Campbell-type event (Fig. 74) (Morscezk, 2003).

The complementation of genes was very useful in the determination of gene function; for example, complementation of RD1 region into *M. bovis* BCG showed increased persistence of the recombinant in immunocompetent mice. Complementation of mycobacteria with heterologous or homologous DNA can be achieved either by shuttle plasmids with a mycobacterial origin of replication, or by integrating the DNA on the bacterial chromosome. The most common origin of autonomous replication used in mycobacteria was derived from pAL500, a plasmid that was first isolated from *M. fortuitum* (Morscezk, 2003).

The other methods include anti-sense method, which reduces the expression of specific genes using Anti-sense RNAs, which prevents the translation of mRNA to which they are complementary (Smith, 2003) and global gene inactivation achieved by inserting the foreign DNA, usually a transposable element, into many sites in the bacterial genome (Sassetti & Rubin, 2002).

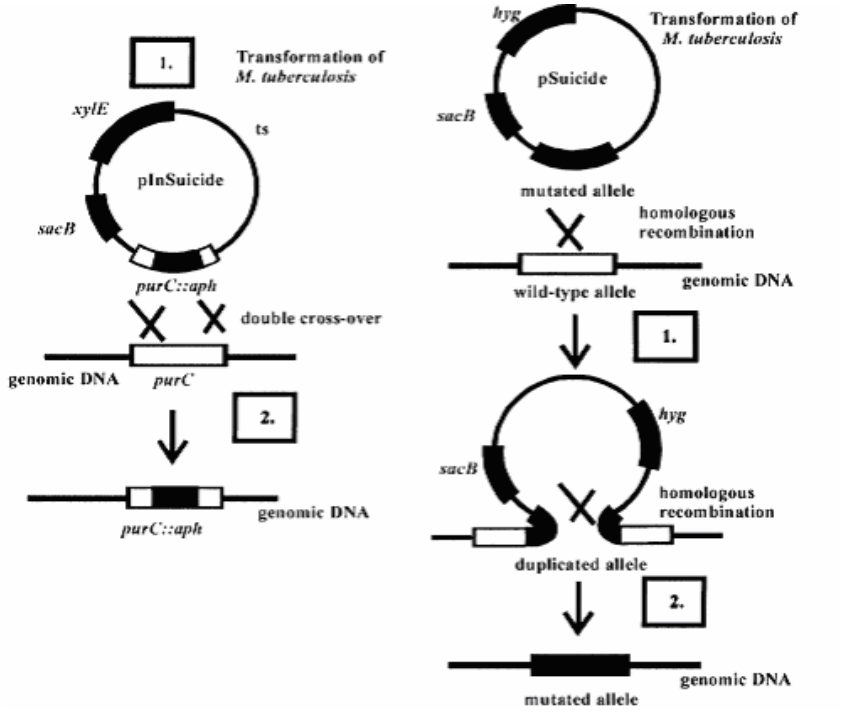


Fig. 74- Single and double selection methods for gene inactivation by allelic exchange in *M. tuberculosis* (Morscezk, 2003).

In the present study, when glycerol was used as the sole carbon source, *M. tuberculosis Δ pykA* showed significant reduction in growth when compared to *M. tuberculosis* but greater than *M. bovis.* When grown on solid media, *M. tuberculosis Δ pykA* showed remarkable change in the colony morphology to dysgonic colonies when compared to *M. tuberculosis.* The effect of complementing *pykA* in *M. bovis* restored the colony morphology from dysgonic and smooth to eugonic and rugose, when grown on glycerol medium. Strains with non-functional PK were unable to grow on glycerol and glucose, but grew readily on pyruvate (catabolised via the citric acid cycle) and Tween 80 (catabolised by beta-oxidation and the citric acid cycle). Thus, the molecular genetic basis for altered colony morphology in *M. bovis,* including *M. bovis* BCG was attributed to the utilisation of pyruvate rather than glycerol uptake by *glpK* (Keating *et al.,* 2005). Hence, it is understood that *pykA* plays a major role in the colony morphology as well as in the utilisation of carbon source.

The present study analysed the role / effect of *pykA* and *ald* in the overall composition of surface lipids and protein profile of the mutants. Surface lipid analysis suggested an overall increase in the degradation of lipids such as triacylglycerol and unknown fatty acids, with decrease in the production of lipids in the $pykA^-$ mutants resulting in the change in morphology. Decrease in the production of lipids as observed in the *M. tuberculosis Δ pykA* may be due to the increased utilisation of lipids, by increasing the expression of lipid degradative enzymes as observed by proteome analysis. Apart from the enzymes involved in lipid degradation, there is also increase in the expression of TCA cycle and glyoxalate cycle enzymes except Icl.

Lipid analysis also suggested an overall increase in the production of lipids in $pykA^+$. This was confirmed by comparison of alterations in the different proteins (proteome analysis) between *M. bovis* and *M. bovis* / pLK102 that showed an increased production of enzymes involved in the synthesis of lipids along with striking increase in levels of PEPCK, Icl and PK. In the present study, it appears that the non-availability of carbohydrate source cause increased utilisation of lipid and not the lack of pyruvate *per se* results in morphology changes. The results also demonstrate *M.*

bovis, M. africanum and *M. microti* complex utilise lipids as an essential carbon source due to the inability of these strains to catabolise carbohydrates for energy.

Our observations parallel other reports on the utilisation of fatty acids as source of energy. Transcriptome analysis of *M. tuberculosis* has demonstrated that genes involved in beta-oxidation and gluconeogenesis were up regulated relative to the *in vitro* condition (Schnappinger *et al.,* 2003). Also, genome analysis revealed the presence of multiple copies of genes for oxidation of fatty acids, suggesting that the fatty acids are the primary carbon sources rather than carbohydrates. Studies also showed that lipids are catabolised *in vivo* to provide energy and carbon in *M. tuberculosis* during infection, while the scarce carbohydrates are scavenged to provide biosynthetic precursors such as glucose-6-phosphate for the synthesis of various major macromolecules including the sugar moieties of the range of glycolipids present in the cell wall of mycobacteria (Timm *et al.,* 2003).

The sources and types of fatty acids that might be utilized by *M. tuberculosis* during infection are unknown. Genome analysis of *M. tuberculosis* revealed that apart from the prototype β-oxidation cycle required for the lipid catabolism (multifunctional FadA / FadB), it also appears to have ~100 enzymes potentially involved in alternative lipid oxidation pathways in which exogenous lipids from host cells can be degraded (Cole *et al.,* 1998). Studies suggest that in the phagosomal environment of the macrophages, fatty acids are a major carbon source for *M. tuberculosis*. Macrophage triacylglycerol stores mobilized during phagocytosis (Mason *et al*., 1972) could also be utilized by *M. tuberculosis*. Uniquely, actinomycetes (including the genus *Mycobacterium*) accumulate large quantities of triacylglycerols, which are catabolized during starvation (Alvarez & Steinbüchel, 2002). *Mycobacterium tuberculosis* may draw on its internal triacylglycerol stores (Daniel *et al*., 2004) during chronic infection, when the bacteria may experience nutritional stress and become dependent on the RelA mediated stringent response for survival (Dahl *et al*., 2003). Also, *M. tuberculosis* could hydrolyse liposome suspensions, releasing free fatty acids (Kondo *et al*., 1985) as the genome encodes as many as 20 putative lipases (Raynaud *et al.,* 2002; Cole *et al*., 1998). *Mycobacterium tuberculosis* can also metabolize dipalmitoyl

phosphatidylcholine (lung surfactant) (Grabner & Meerbach W, 1991), a long chain fatty acid in a glyoxylate cycle-dependent manner (Muñoz-Elías & McKinney, 2005). *In vivo* studies showed up-regulation of glyoxalate cycle enzymes in human macrophages (Schnappinger *et al*., 2003) and mice (Timm *et al*., 2003; Dubnau *et al*., 2005) during growth on fatty acids. Glyoxylate cycle is the only other pathway known capable of oxidation of acetyl CoA to CO_2 while simultaneously generating pyruvate and PEP for biosynthetic pathways, which may lead to the possibility that the TCA cycle may act in bifurcated mode in *M. tuberculosis* (Tian *et al.,* 2005).

The other important metabolic adaptation may be the utilisation of gluconeogenic and anaplerotic products for the survival of mycobacteria. Sassetti and Rubin (2003) have shown that inactivation of 15 genes involved in lipid metabolism and disruption of the gene *glpX*, encoding the key gluconeogenic enzyme fructose 1, 6-bisphosphatase, attenuated the mutants for *in vivo* growth. Adaptation of *M. tuberculosis* to low carbohydrate availability is suggested by preliminary evidence that a putative high-affinity ABC transporter for hexoses and a hexose kinase are required for survival in mice (Sassetti & Rubin, 2003). Evidences showed that PEPCK and Icl are found to be up-regulated in the lungs of mice and TB patients (Timm *et al.,* 2003). Collins and his group (2002) concluded from the avirulence of their *pckA* knock-out strains (lacking PEPCK) that the gluconeogenic pathway is essential for *M. bovis* virulence in guinea pigs. Similarly *pckA* deficient *M. bovis* BCG showed attenuated virulence in mice and in macrophages (Liu *et al*., 2003). The importance of lipids as sources of carbon to *M. bovis in vivo* is highlighted by the attenuation of *pckA* mutants. Disruption of *icl* in *M. tuberculosis* impaired the survival after the acute phase of infection in the lungs of mice and accounted for their inability to stay dormant as persistors in macrophages and mice (McKinney *et al.,* 2000).

In addition to acetyl-CoA, β-oxidation of odd-chain fatty acids yields propionyl-CoA, which is metabolized via several routes, including the citramalate, methylmalonyl-CoA and methylcitrate pathways. Emerging evidence suggests that the methylcitrate cycle is the dominant route for propionyl-CoA metabolism in eubacteria (Textor *et al*., 1997; Horswill & Escalante-Semerena, 1999; Claes *et al*., 2002). The

methylcitrate cycle involves the cleavage of methylisocitrate to succinate and pyruvate by methylisocitrate lyase, analogous to the reaction catalysed by Icl in the glyoxylate cycle.

The other pathway involved in the production of pyruvate is the L-alanine dehydrogenase (Ald). Ald catalyses the reversible conversion of pyruvate to alanine in the presence of ammonia and $NADH^{+}$. Based on enzyme kinetics, at physiological pH, Ald preferentially catalyses the conversion of alanine to pyruvate rather than the production of alanine. Though Ald acts as a connecting link between carbon and nitrogen metabolism, the exact role is not known. It has been postulated that the nitrogen source may act as a carbon source in the late log phase of growth in many organisms or during the non-availability of a proper carbon source. Earlier experiments on the utilisation of various amino acids as nitrogen sources by *M. tuberculosis* H37Ra showed that L-alanine augmented the growth when used along with aspartic acid or glutamic acid (Lyon *et al.,* 1970). However, the growth of *M. smegmatis* and *M. bovis* BCG were inhibited when L-alanine was used as the sole nitrogen source indicating a defect in the utilisation of L-alanine. Ald is one of the functional antigens present in the culture filtrate of *M. tuberculosis* and not in *M. bovis* BCG.

Genome analysis of *M. tuberculosis* and *M. bovis* showed a deletional SNP at position 362 of *ald* gene in *M. bovis* leading to frame shift mutation that results in the production of a truncated and a non-functional Ald (Garnier *et al.,* 2003). This was also seen in *M. microti* and *M. africanum* and *M. bovis* BCG. The functional *ald*, seen in *M. tuberculosis* and *M. cannetti* segregates these 2 members from the other members of the *M. tuberculosis* complex.

Genome analysis of *ald* suggests that the deletion occurred along with the RD9, a key deletion event, which splits *M. tuberculosis* and *M. canettii* from *M. africanum, M. microti* and *M. bovis* (Garnier *et al.,* 2003). Along with mutation in *pykA*, *ald* co-segregates with the RD9 deletion event, such that *M. tuberculosis* complex that are RD9 deleted also carry the *pykA* and *ald* mutation; hence the *pykA* and *ald* mutation are consistent with the Brosch phylogeny that suggests that *M.*

tuberculosis is closer to the common progenitor of the complex than *M. bovis* and evolution of *M. bovis* from *M. tuberculosis* is by chromosomal deletion events (Brosch *et al.,* 2002). Clearly, the loss of *pykA* and *ald* in this branch of the complex shows that these enzymes activities are not required for the virulence of *M. bovis, M. africanum,* and *M. microti*. Taken with the evidence that lipids are the key carbon source for the tubercle bacilli *in vivo*, the loss of PK and Ald activity reflects reductive evolution, with the accumulation of mutations in genes whose function is no longer required for their *in vivo* survival.

Colony morphology studies using *M. tuberculosis Δ ald* showed that *ald* may not play a direct role in the colony morphology as seen by the similar morphology of *M. bovis* BCG, with a functional PK and non-functional Ald. Ald along with PK probably play a role in lipid production, as the polar lipid fraction from *M. tuberculosis Δ ald Δ pykA* was identical to that of *M. bovis,* while *M. tuberculosis Δ ald* and *M. tuberculosis Δ pykA* have not showed significant difference when compared to *M. tuberculosis*. Attempts made to complement *ald* in *M. bovis* were futile. Complementing the gene and over-production of Ald may help in identify its role in mycobacteria.

The global effect of the alteration in the carbon metabolism was reflected by the proteome analysis. In the *M. bovis* knock-in mutant, there is a notable increase in PK due to the insertion of a functional gene. There is an increase in the lipid degradation enzymes, the increase being higher in the *M. tuberculosis* knock-out mutant. The specific observations with reference to iron-related proteins and oxidative stress enzymes is presented in Table 22. In the *M. bovis* knock-in mutant, there is an increase in IdeR, the iron regulator protein and the iron storage proteins bacterioferritin. These are much lowered in the *M. tuberculosis* knock-out mutant. Of the oxidative stress enzymes, the Fe-SOD (SodA) is notably high in the latter, with parallel increase in AhpC and KatG.

Table 22 - Comparison of the changes in the levels of the proteins involved in iron metabolism and oxidative stress in *M. tuberculosis* and *M. bovis* mutants.

Enzyme / protein	% change in *M. bovis* knock-in	% change in *M. tuberculosis* knock-out
Pyruvate kinase	2150	0
Iron storage proteins		
BfrA	166	NA
BfrB	115	36
Iron metabolism		
Aconitate hydratase	68	2
Iron regulator IdeR	28	8
50S ribosomal protein L4*	-23	100
HupB*	-50	-10
Possible iron-regulated short chain dehydrogenase / reductase	NA	138
Iron-regulated hypothetical protein	NA	67
Iron-regulated heparin binding protein	NA	30.1
Iron-regulated EF-Tu	NA	24
Oxidative stress enzymes		
Cu-Zn SOD (SodC)	0	-12.5
Fe-SOD (SodA)	-57	46
KatG	-18	17
AhpC	-73	130
* these proteins were up-regulated in *M. tuberculosis* grown under conditions of iron limitation (separate study done in the lab of Dr. Manjula Sritharan) NA is not available		

Table shows the comparison of some of the oxidative stress and iron-related proteins between the *M. bovis* knock-in mutant (change between *M. bovis* and *M. bovis /* pLK102) and *M. tuberculosis* knock-out mutant (change between *M. tuberculosis* and *M. tuberculosis Δ pykA*).

Of interest is the iron-regulated expression of HupB protein. Studies in our lab showed that HupB is an iron-regulated protein in *M. tuberculosis*, and is coordinately regulated with the expression of mycobactin and carboxymycobactin. There is up-regulation of two additional proteins of sizes 54 kDa and 26 kDa respectively, both of which, when sequenced were identified as 50S ribosomal L4 protein. In the proteome analysis of *M. tuberculosis* knock-out mutant, the level of

HupB was significantly low as also the 50S ribosomal protein L4. The levels of these two proteins were higher in the *M. bovis* knock-in mutant. The levels of free iron is probably higher in the knock-in mutant as compared to the knock-out mutant, thereby showing increase in IdeR and iron storage protein with a decrease in the HupB and 50S ribosomal L4 protein. Growth of these mutants under high and low iron conditions will probably give a better understanding of the alterations in these mutants.

CONCLUSIONS

This study involves two aspects of the metabolism of pathogenic mycobacteria, namely the iron and carbon metabolism. Both of them have significant implications on the host-pathogen interactions. There is increasing evidence to show that iron metabolism plays an important role under *in vivo* conditions. Here, though independent work on iron and carbon metabolism was done, in the latter study, we extrapolate the global effect of alteration in the carbon metabolism, especially with reference to iron and oxidative stress.

There is coordinated expression of the mycobactin, carboxymycobactin and the IREPs, though in the latter further analysis is required to understand if one or more of these proteins play the role of a receptor for the ferri-carboxymycobactin. The structural elucidation showed that the carboxymycobactin was unique in that it had the methyl-salicylate moiety unlike the carboxymycobactin of *M. tuberculosis* and *M. avium*, in which the methyl group was absent.

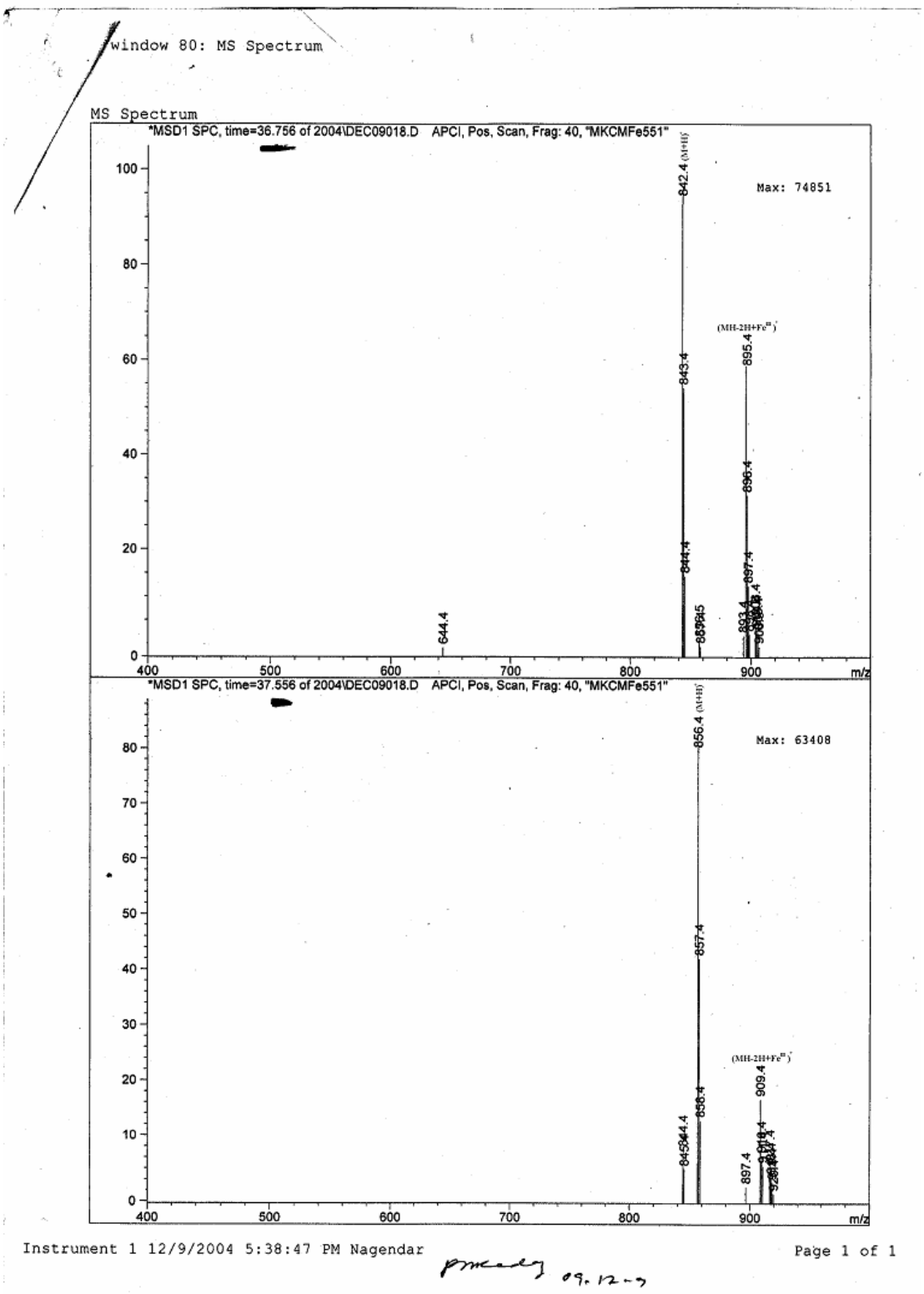
Iron acquisition is of great clinical significance and it is necessary to understand the adaptation of pathogenic mycobacteria to the environment encountered *in vivo*. Anemia is frequently seen in TB patients, indicating the host response to infection by lowering of the iron, a micronutrient that is necessary for the pathogen multiplication and survival. In addition, the characterization of the siderophores will help in the designing and proper formulation of the analogs for use as anti-tubercular molecules and may also provide more insights on the transport of molecules across the lipid cell wall. The IREPs may also act as potential drug targets and as vaccine candidates. Alternate drugs are required, especially for opportunistic pathogens like *M. kansasii*, which is being increasingly implicated as co-infecting with HIV and which cannot be treated with the regular drug regimen, as they interfere with the anti-retroviral (ARV) therapy.

The carbon metabolism is complex, with the organism adapting its metabolic pathways to optimally utilize the available carbon sources. There is increasing evidence to show that lipids are preferentially utilized under *in vivo* conditions. The global effect of a shift in the carbon metabolism was seen in the mutants used in this

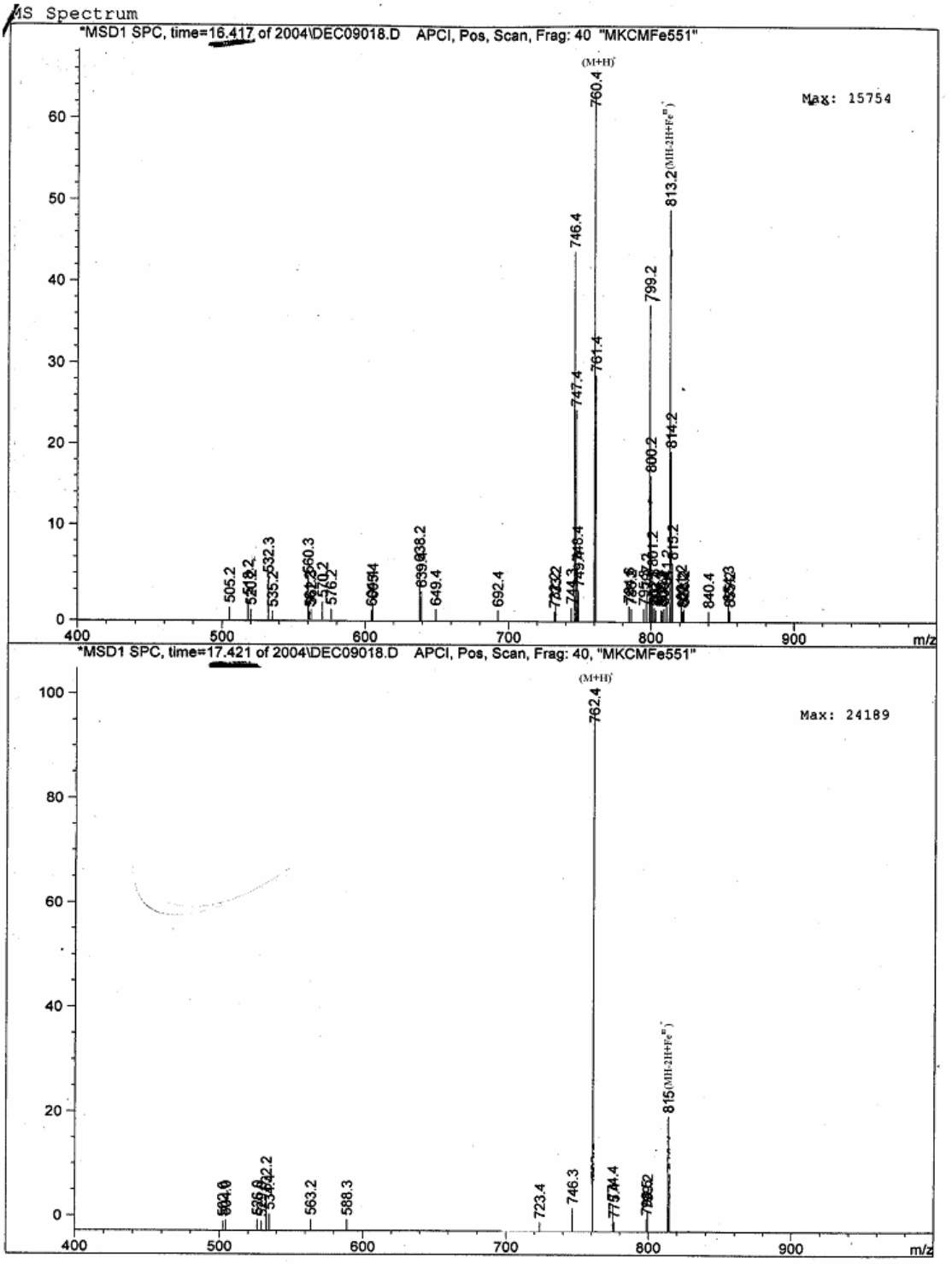
study and it had a profound influence on the iron and oxidative stress machinery. In the knock-out mutant of *M. tuberculosis* (*M. tuberculosis Δ pykA*), the up-regulation of the enzymes involved in lipid degradation show that there is increase in respiration due to increase in lipid degradation. This results in an increase in the generation of the toxic reactive oxygen species, an inference supported by an up-regulation of the oxidative stress enzymes. Of significance is the SOD, the predominant enzyme is Fe-SOD (SodA), which is notably high. The reason for the moderate rise in KatG with a steep rise in AphC is not clear. There are reports to show that when there is a mutation in *katG*, AhpC increases to tackle the oxidative stress, in other words compensate for lack of KatG. Our observation in the *M. tuberculosis Δ pykA* may have some implications in the understanding of the metabolism of *M. tuberculosis* inside a human host. With increasing evidence of utilization of lipids as carbon sources *in vivo*, the resultant effects on iron metabolism is possible and needs to be explored. A more intensive study is required to understand the association of the different metabolic pathways, as the infecting pathogen must be in dynamic equilibrium with its environment, altering its metabolism with fineness so as to establish a successful infection within the hostile environment of the macrophages.

SUPPLEMENTARY DATA

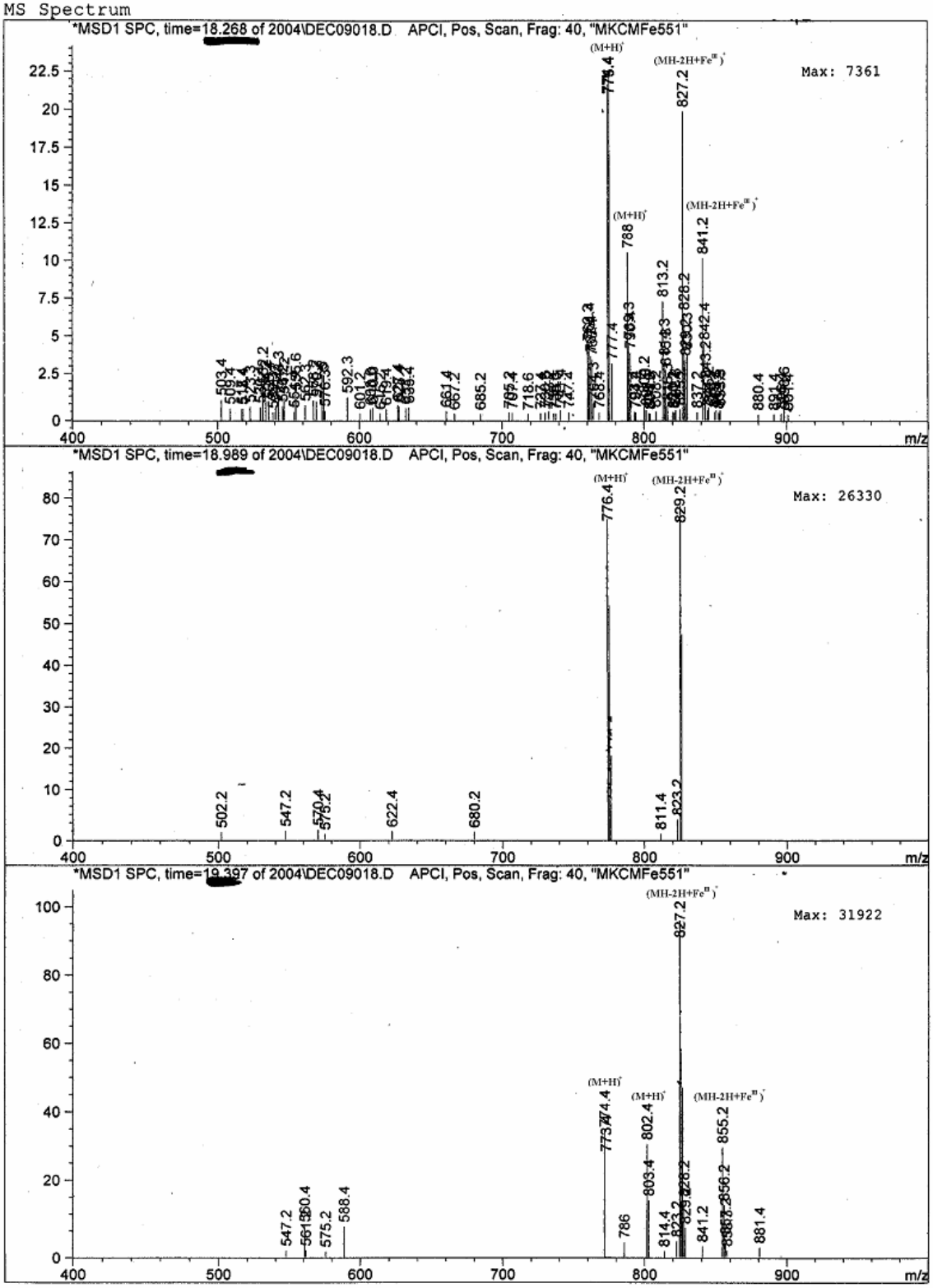
LC-MS analysis of carboxymycobactin



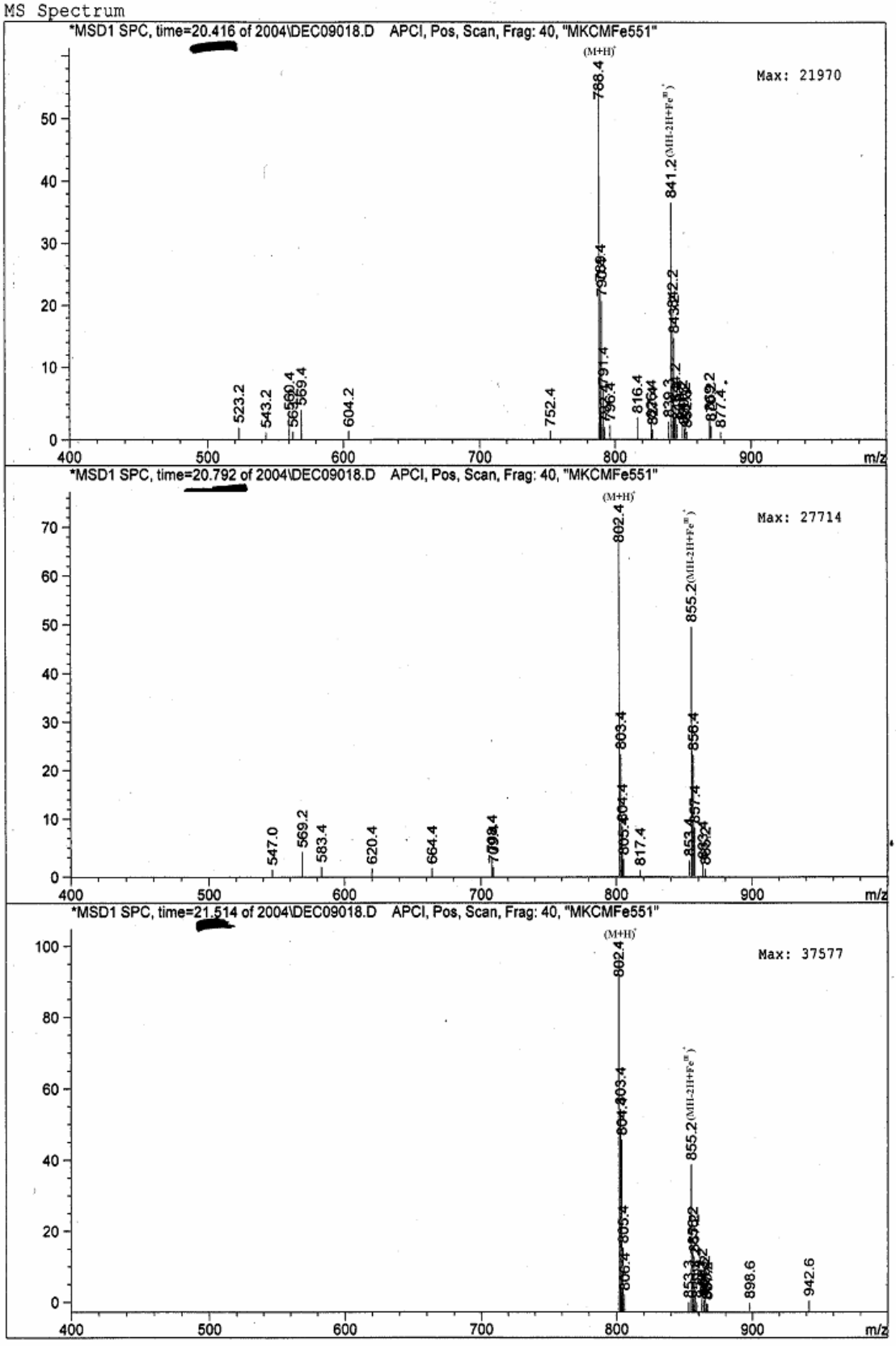
Indow 80: MS Spectrum



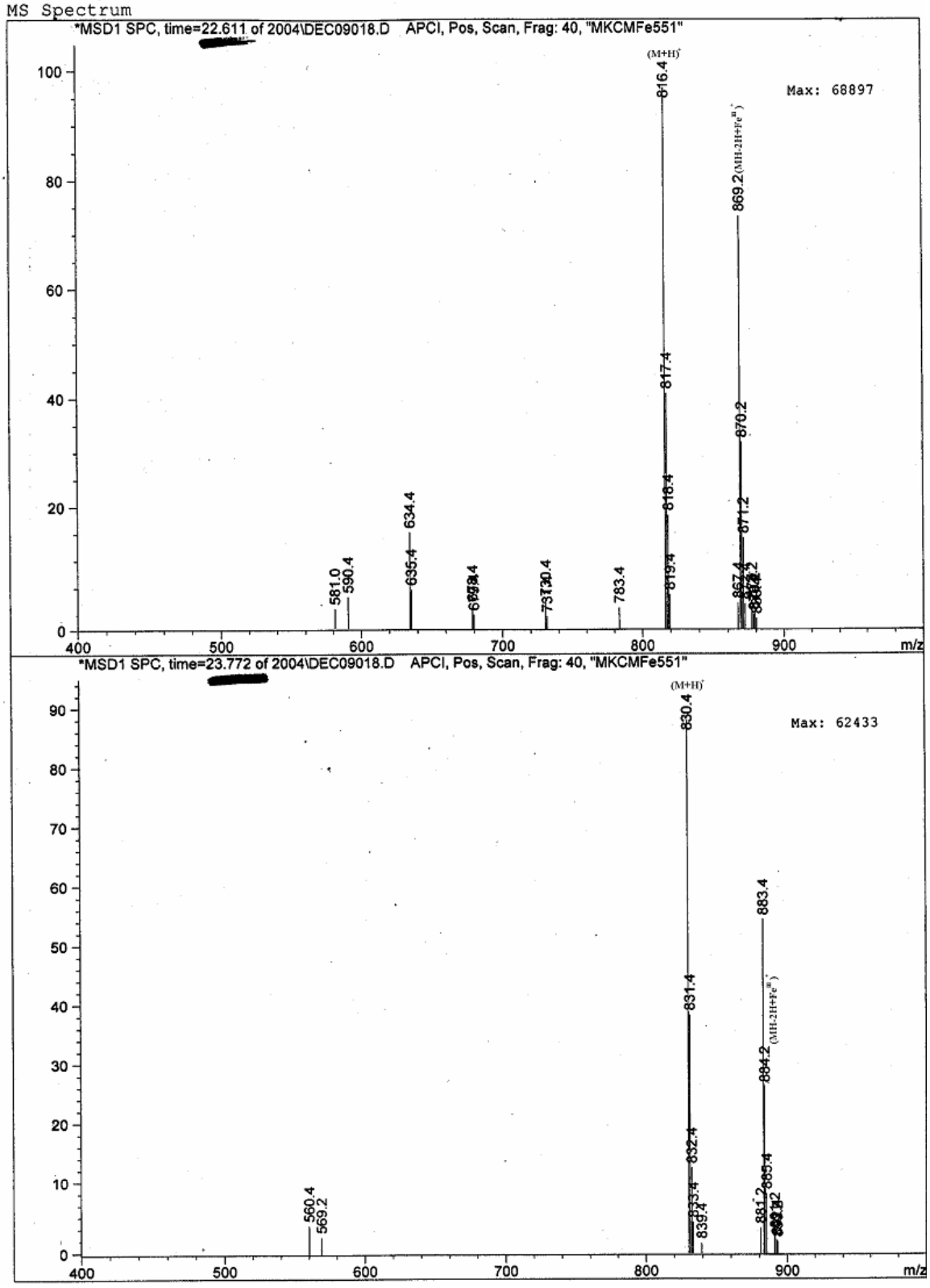
MS Spectrum

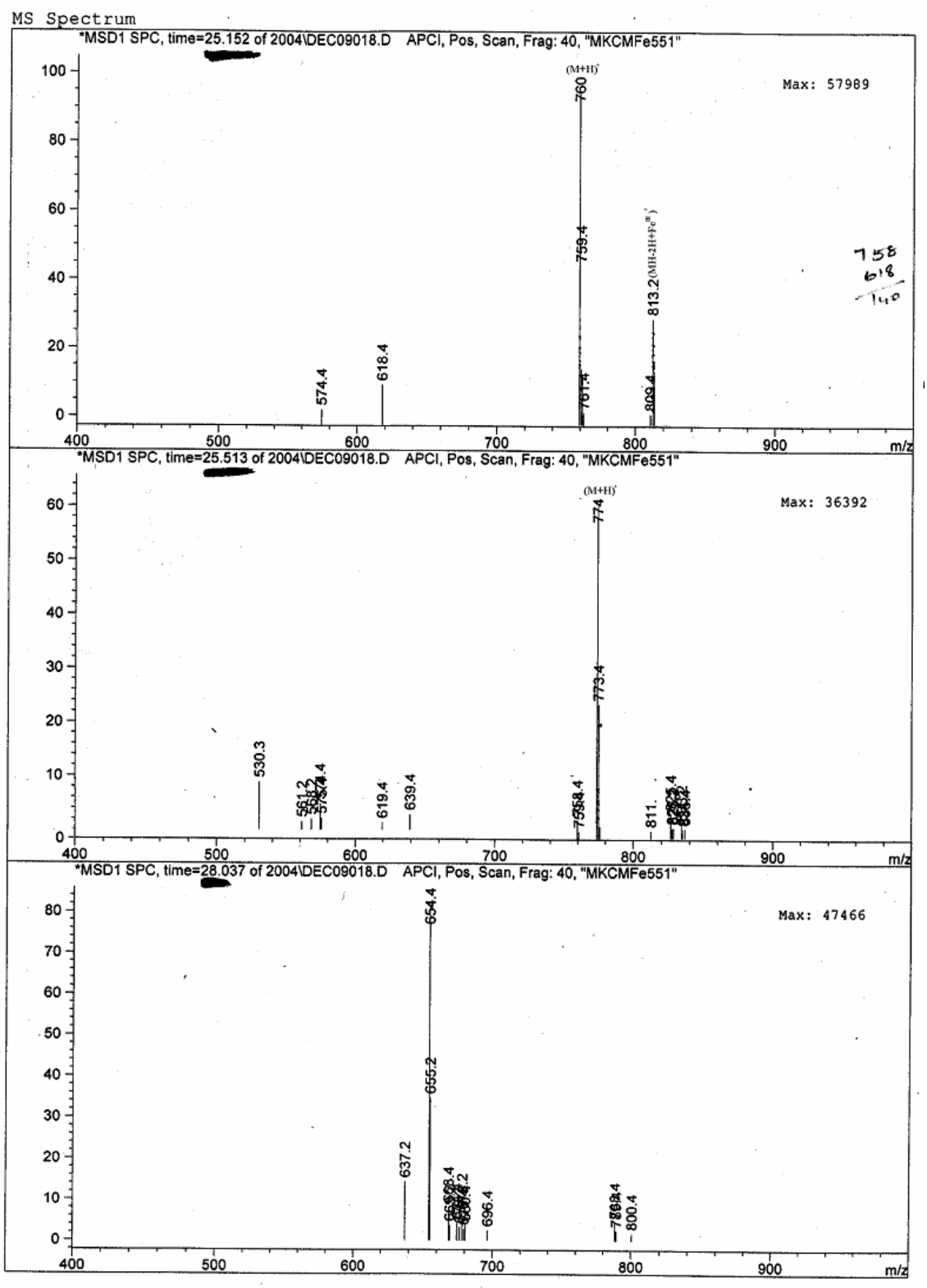

*MSD1 SPC, time=18.268 of 2004\DEC09018.D APCI, Pos, Scan, Frag: 40, "MKCMFe551"
Max: 7361
(M+H)+
(MH-2H+Fe^III)+
776.4
827.2
788
841.2
813.2
m/z
*MSD1 SPC, time=18.989 of 2004\DEC09018.D APCI, Pos, Scan, Frag: 40, "MKCMFe551"
Max: 26330
776.4
829.2
502.2
547.2
622.4
680.2
811.4
823.2
m/z
*MSD1 SPC, time=19.397 of 2004\DEC09018.D APCI, Pos, Scan, Frag: 40, "MKCMFe551"
Max: 31922
827.2
774.4
802.4
855.2
547.2
575.2
588.4
786
841.2
881.4
m/z

MS Spectrum

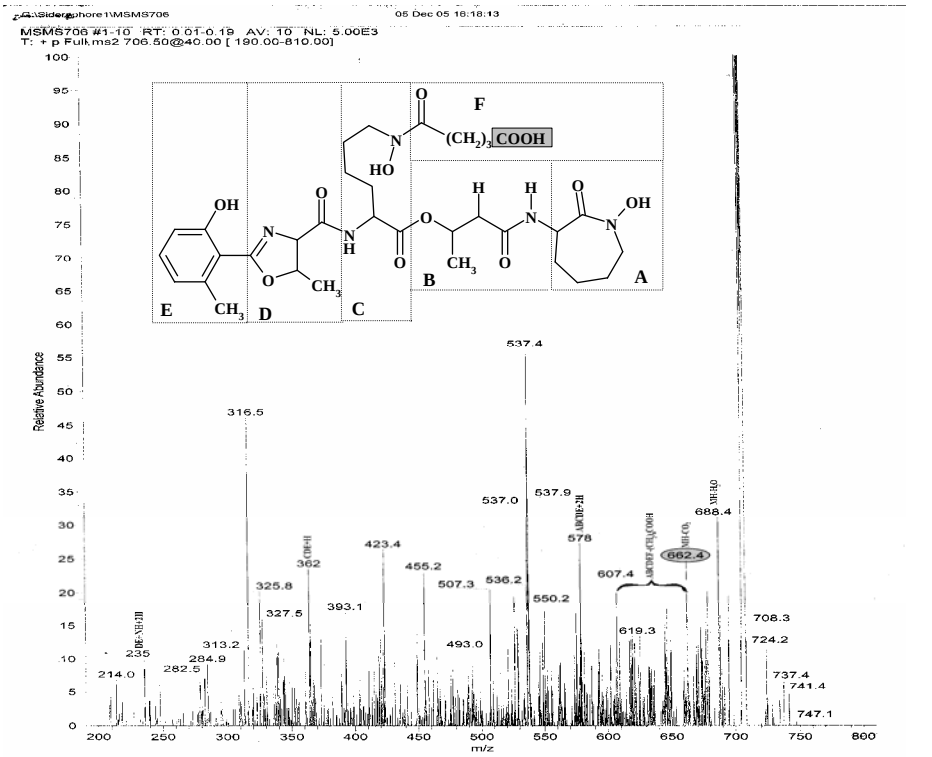

*MSD1 SPC, time=20.416 of 2004\DEC09018.D APCI, Pos, Scan, Frag: 40, "MKCMFe551"
Max: 21970
(M+H)+
786.4
841.2(MH-2H+Fe3+)+
523.2
543.2
569.4
604.2
752.4
816.4
877.4
*MSD1 SPC, time=20.792 of 2004\DEC09018.D APCI, Pos, Scan, Frag: 40, "MKCMFe551"
Max: 27714
(M+H)+
802.4
855.2(MH-2H+Fe3+)+
547.0
569.2
583.4
620.4
664.4
817.4
*MSD1 SPC, time=21.514 of 2004\DEC09018.D APCI, Pos, Scan, Frag: 40, "MKCMFe551"
Max: 37577
(M+H)+
802.4
855.2(MH-2H+Fe3+)+
898.6
942.6
m/z

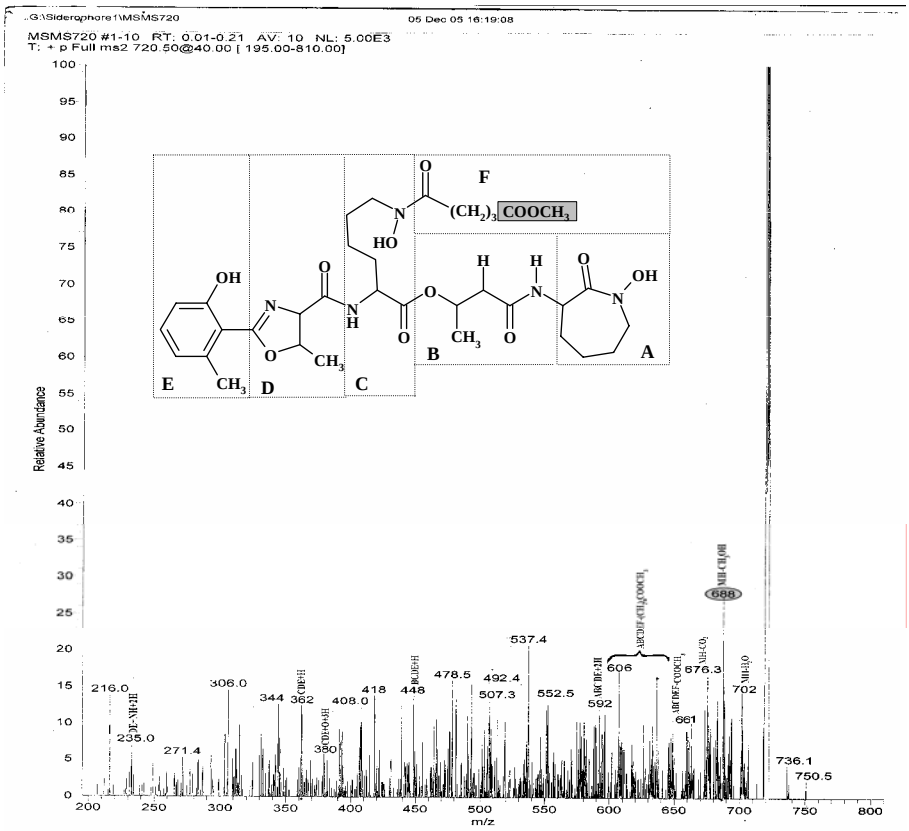
MS Spectrum

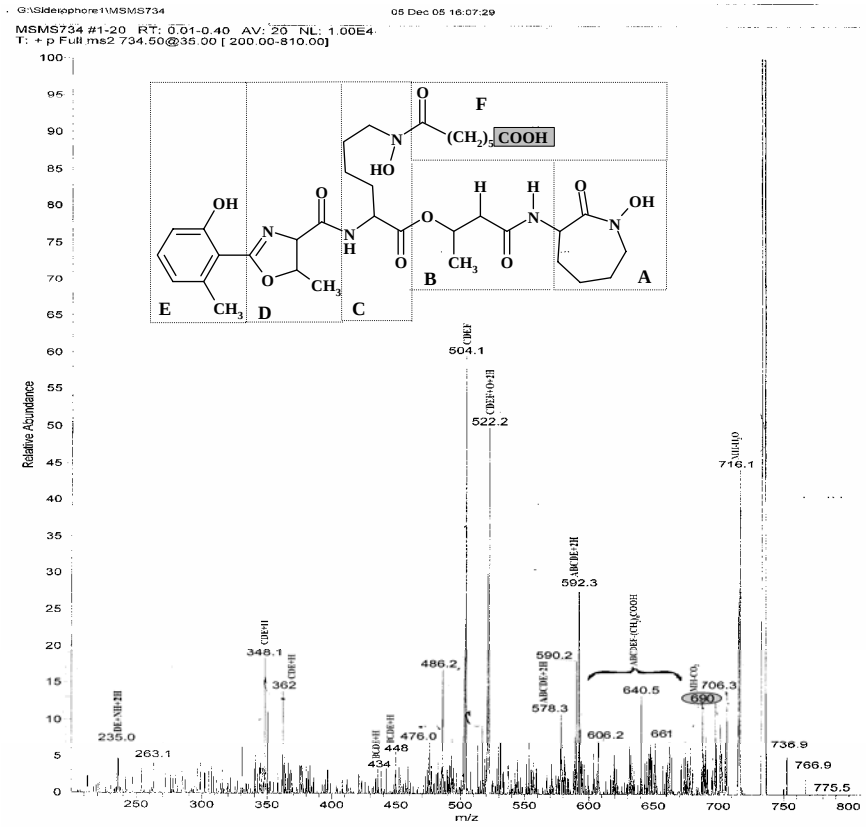

*MSD1 SPC, time=22.611 of 2004\DEC09018.D APCI, Pos, Scan, Frag: 40, "MKCMFe551"
Max: 68897
(M+H)+
816.4
869.2
(MH-2H+Fe)+
817.4
870.2
634.4
581.0
590.4
635.4
783.4
m/z
*MSD1 SPC, time=23.772 of 2004\DEC09018.D APCI, Pos, Scan, Frag: 40, "MKCMFe551"
Max: 62433
(M+H)+
830.4
883.4
831.4
884.2
(MH-2H+Fe)+
560.4
569.2
832.4
m/z

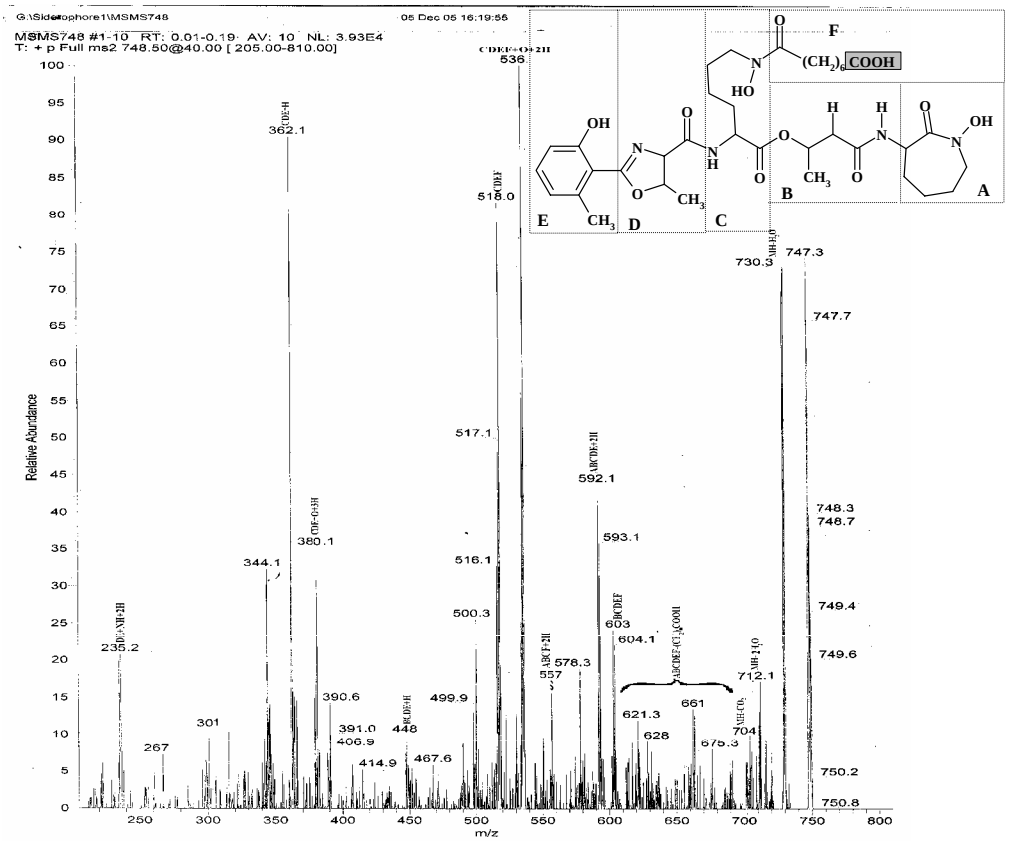

MS Spectrum
*MSD1 SPC, time=25.152 of 2004\DEC09018.D APCI, Pos, Scan, Frag: 40, "MKCMFe551"
Max: 57989
(M+H)+
760
759.4
813.2
618.4
574.4
*MSD1 SPC, time=25.513 of 2004\DEC09018.D APCI, Pos, Scan, Frag: 40, "MKCMFe551"
Max: 36392
(M+H)+
774
773.4
530.3
619.4
639.4
*MSD1 SPC, time=28.037 of 2004\DEC09018.D APCI, Pos, Scan, Frag: 40, "MKCMFe551"
Max: 47466
654.4
655.2
637.2
696.4
800.4
m/z

LC-MS / MS analysis of carboxymycobactins

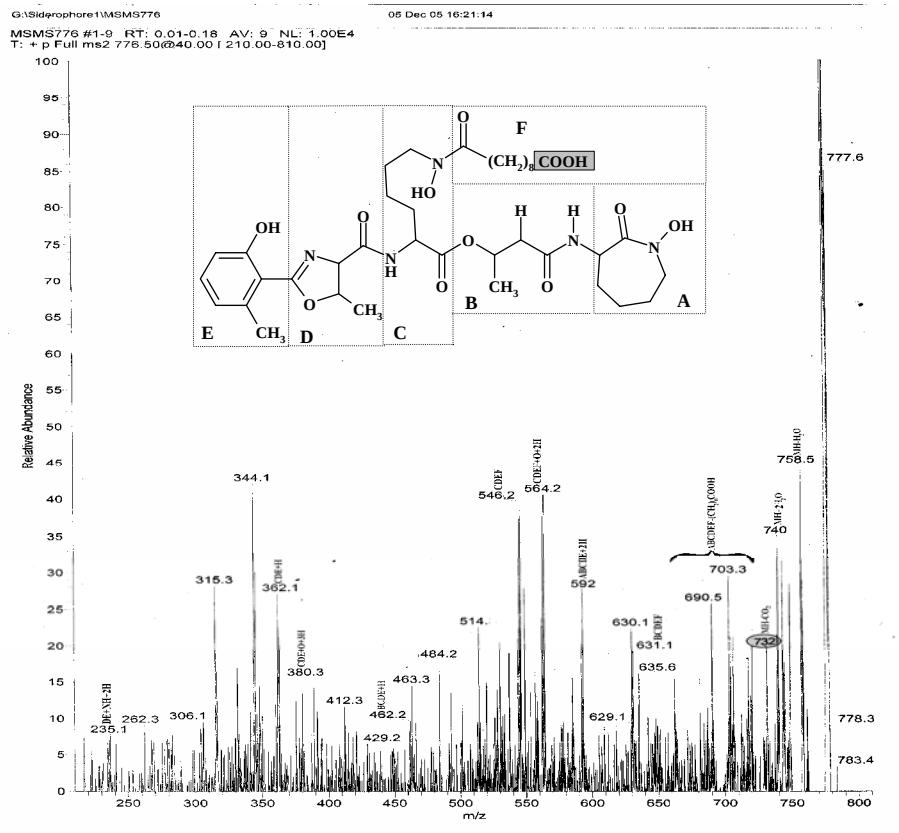


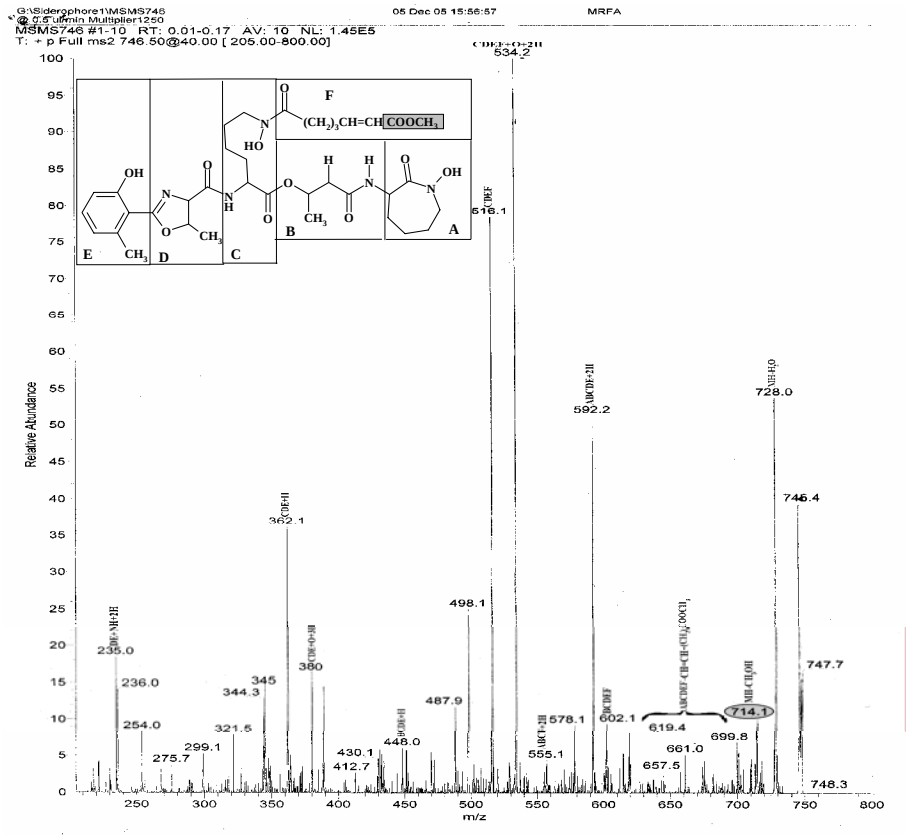


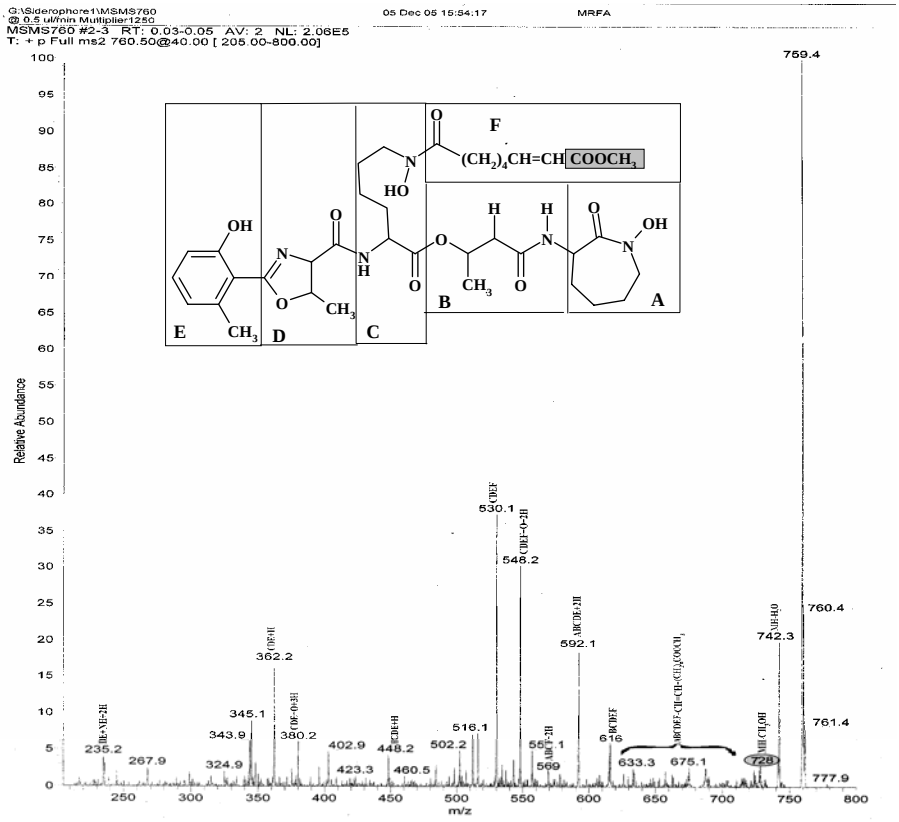

G:\Siderophore1\MSMS734
05 Dec 05 16:07:29
MSMS734 #1-20 RT: 0.01-0.40 AV: 20 NL: 1.00E4
T: + p Full ms2 734.50@35.00 [200.00-810.00]
Relative Abundance
m/z
F
(CH2)5COOH
HO
OH
N
O
H
CH3
A
B
C
D
E
504.1
522.2
716.1
592.3
348.1
362
486.2
590.2
640.5
706.3
690
578.3
606.2
661
235.0
263.1
476.0
448
434
736.9
766.9
775.5

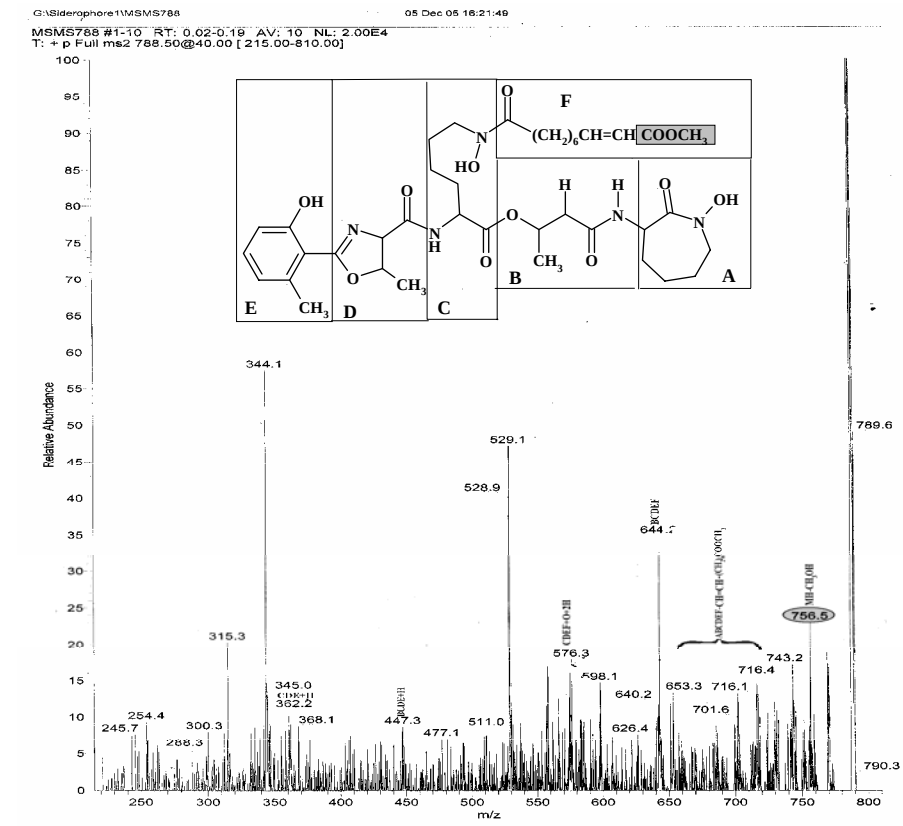

G:\Siderophore1\MSMS748
05 Dec 05 16:19:55
MSMS748 #1-10 RT: 0.01-0.19 AV: 10 NL: 3.93E4
T: + p Full ms2 748.50@40.00 [205.00-810.00]
Relative Abundance
m/z
F
(CH2)6COOH
HO
OH
N
O
H
CH3
A
B
C
D
E
536
362.1
518.0
730.3
747.3
747.7
517.1
592.1
593.1
748.3
748.7
516.1
380.1
344.1
500.3
603
604.1
749.4
749.6
578.3
557
712.1
661
621.3
628
704
675.3
390.6
499.9
391.0
406.9
448
301
267
414.9
467.6
750.2
750.8
235.2

X


G:\Siderophore1\MSMS776
05 Dec 05 16:21:14
MSMS776 #1-9 RT: 0.01-0.18 AV: 9 NL: 1.00E4
T: + p Full ms2 776.50@40.00 [210.00-810.00]
Relative Abundance
m/z
777.6
778.3
783.4
344.1
315.3
362.1
546.2
564.2
758.5
740
703.3
690.5
592
630.1
631.1
635.6
732
514.
484.2
463.3
412.3
380.3
462.2
429.2
306.1
262.3
235.1
629.1
F
(CH2)8COOH
HO
OH
CH3
B
A
E
D
C


G:\Siderophore1\MSMS746
05 Dec 05 15:56:57
MRFA
@ 6.5 Turmin Multiplier1250
MSMS746 #1-10 RT: 0.01-0.17 AV: 10 NL: 1.45E5
T: + p Full ms2 746.50@40.00 [205.00-800.00]
Relative Abundance
m/z
534.2
516.1
592.2
728.0
746.4
747.7
748.3
362.1
498.1
345
344.3
380
487.9
602.1
578.1
555.1
619.4
714.1
699.8
661.0
657.5
448.0
430.1
412.7
321.5
299.1
275.7
254.0
236.0
235.0
F
(CH2)3CH=CHCOOCH3
HO
OH
CH3
B
A
E
D
C


G:\Siderophore1\MSMS760
05 Dec 05 15:54:17
MRFA
@ 0.5 uV/min Multiplier:1250
MSMS760 #2-3 RT: 0.03-0.05 AV: 2 NL: 2.06E5
T: + p Full ms2 760.50@40.00 [205.00-800.00]
Relative Abundance
m/z
759.4
760.4
761.4
777.9
742.3
728
530.1
548.2
592.1
362.2
345.1
343.9
380.2
516.1
235.2
267.9
324.9
402.9
423.3
448.2
460.5
502.2
633.3
675.1
F
(CH2)4CH=CH COOCH3
HO
OH
N
O
NH
H
CH3
B
A
E
D
C


G:\Siderophore1\MSMS788
05 Dec 05 16:21:49
MSMS788 #1-10 RT: 0.02-0.19 AV: 10 NL: 2.00E4
T: + p Full ms2 788.50@40.00 [215.00-810.00]
Relative Abundance
m/z
789.6
790.3
756.5
743.2
716.4
344.1
529.1
528.9
644.1
315.3
576.3
598.1
640.2
653.3
716.1
701.6
626.4
345.0
362.2
368.1
447.3
477.1
511.0
254.4
245.7
300.3
288.3
F
(CH2)6CH=CH COOCH3
HO
OH
N
O
NH
H
CH3
B
A
E
D
C

Overall proteome analysis between the wild type and the mutants

Table S1 - Comparison for the expression of protein profile between *M. tuberculosis* and *M. tuberculosis Δ pykA*

Locus ID	Name of protein	% change
15609582	NDP KINASE	175
15609436	HTPG	164.7059
15607994	hypothetical protein Rv0854	150
15610360	Iron-regulated short chain dehydrogenase / reductases	138.4615
15609565	AHPC	130.7692
15607384	FADA2	126.087
15607862	50S RIBOSOMAL PROTEIN L30 RPMD	100
15607935	TRANSPOSASE FOR IS6110 (FRAGMENT)	100
57116835	hypothetical protein Rv1159A	100
15607842	50S RIBOSOMAL PROTEIN L4 RPLD	100
57116872	MOXR1	100
15610421	ACCA3	100
15609383	KASB	100
15610525	Possible dehydrogenase	92.85714
15610814	hypothetical protein Rv3678c	90
15610410	FADE25	89.47368
15608210	ECHA8 (CROTONASE)	87.5
15607383	FABG4	87.17949
15610940	Antigen 85-A FBPA	85.71429
15607208	ICD2	85.71429
15610846	LEUA	83.33333
15609277	hypothetical protein Rv2140c	77.77778
15607702	GRCC1	75
15607706	hypothetical protein Rv0566c	75
15609614	Macrolide-transport ABC transporter	75
15610165	FIXB	74.13793
15608092	SUCD	71.875
15609381	ACPM	71.42857
15608964	CFP17	71.42857
57116992	CLPP1	71.42857
15610419	SSEA	71.42857
15610812	Transcriptional regulatory protein	69.56522
15608045	ECHA6	69.23077
15608774	IRON-REGULATED TB15.3	67.56757

Locus ID	Name of protein	% change
15608214	FADA3	66.66667
15607782	MMAA4	64.28571
15607861	30s ribosomal protein S5 RPSE	64.28571
15610416	ACCD5	64
15609760	hypothetical protein Rv2623	61.90476
15610594	30 s ribosomal protein S4 RPSD	61.90476
15607954	SSEC2	61.53846
15608437	RHO	61.11111
15608764	Two-component system transcriptional regulator	60
15607849	50 s ribosomal protein L29 RPMC	60
15608211	ECHA9	60
15610026	EF-TS	58.62069
15610937	FADD32	57.69231
15609740	Highly conserved hypothetical protein	57.14286
15609117	Antigen MPT64/MPB64	55.55556
15607823	30 s ribosomal protein S7 RPSG	55
57116768	EF-G	54.83871
15607955	CYSA2	54.05405
15608527	MIHF	53.84615
15608000	FADB	53.57143
15610027	RPSB	52
57117019	Conserved 35 kda alanine rich protein	50
15609597	CLPP2	50
15607474	hypothetical protein Rv0333	50
15608592	QOR	50
15607841	50S ribosomal protein L3	50
15607290	Short-chain type dehydrogenase / reductase	50
57116724	DNAJ1	50
15607776	hypothetical protein Rv0636	50
15607198	50s ribosomal protein L9	50
15608450	ATPD	48.33333
15608974	GLCB	47.36842
15610982	SODA	46.66667
15610719	TF	45.45455
15608380	MDH	44.64286
15608463	FADA4	44.44444
15609743	Pyridoxine biosynthesis protein	43.47826
15609282	WAG31	42.85714
57117069	hypothetical protein Rv3196A	42.85714
15607858	30s ribosomal protein S8 RPSH	40
15608622	INHA	40

Locus ID	Name of protein	% change
15607792	50 s ribosomal protein L7/L12	39.47368
15607971	hypothetical protein Rv0831c	39.13043
15610276	FADE23	38.88889
15611050	TRXC	38.46154
15607492	GRPE (HSP-70 COFACTOR)	37.5
15609357	GLNA1	36.84211
15610019	RRF	36.66667
15610977	BFRB	36.36364
15607389	[Iron-Sulfur subunit] Succinic dehydrogenase	36.36364
15608238	FUM	35.29412
15608448	ATPA	34.78261
15608586	TAL	33.33333
57116727	SECE2	33.33333
57116993	ISOMERASE	33.33333
15607388	[Iron-Sulfur subunit] Succinic dehydrogenase	33.33333
57117042	SERA1	33.33333
15610236	hypothetical protein Rv3099c	33.33333
15610845	ASK	33.33333
15609228	Probable membrane protein	33.33333
15609287	FtsZ	33.33333
15610242	PRFB	33.33333
15608163	ENO	31.81818
15607791	50s ribosomal protein L10	31.57895
15608220	GREA	31.57895
15607616	HBHA	30.76923
15609247	PrcB	28.57143
57116926	Immunogenic protein MPT32	28.57143
15609023	Antigen 85-B	28.57143
15608091	SUCC	28
15609976	INFB	27.77778
15607807	RPOB	26.66667
15609168	HSPX	26.66667
15610764	PPASE	26.31579
15610382	MTRA	26.31579
15608273	METE	25.35211
15608575	PGK	25
15609352	SucB	24.7191
15607491	HSP70	24.64455
15607825	EF-TU	24
15607525	CLPB	23.72881
15607297	PNTAA	23.07692

Locus ID	Name of protein	% change
15608531	METK	22.5
57117131	CLPC1	22.44898
15609510	DNAJ2	21.42857
15608626	Exported conserved protein	21.42857
15610108	Oxidoreductase	21.42857
15607151	PPIA	21.05263
15608574	GAP	20.96774
15607504	FBA	20.58824
15607847	30 s ribosomal protein S3	20
15610078	FADD28	20
15610384	SAHH	19.56522
15609045	KATG	19.35484
15609672	PEPQ	16.66667
15607178	hypothetical protein Rv0036c	16.66667
15611005	CMP	16.66667
15610936	PKS13	15.55556
15610553	GROEL	15.21739
15608768	Ribosomal protein RPSA	14.70588
57116763	hypothetical protein Rv0634A	14.28571
15610554	GroES	14.02715
15609439	hypothetical protein Rv2302	11.11111
15607581	GroEL	10.77844
57116801	PSTS1	10.71429
15607873	ADK	10.34483
15609671	EFP	10
15608071	PKND	9.677419
15611010	CFP10	8.433735
15609848	IDER	8
15607781	50s ribosomal protein	4.761905
15608464	Thioredoxin	4.166667
15608036	GLTA2	4
15610784	CSPA	3.225806
15608613	CAN	2.040816
15607840	RPSJ	0
57117165	ESAT-6	0
15610593	RPOA	0
15608317	FDXC	0
15607780	RPLK	0
15607860	RPLR	0
15610883	hypothetical protein Rv3747	0
15610166	FIXA	0

Locus ID	Name of protein	% change
15607424	CMP	0
15609006	Probable reductase	0
15608405	hypothetical protein Rv1265	0
15609322	hypothetical protein Rv2185c	0
15607772	ECHA6	0
15608434	THRA	0
15607808	RPOC	-2.631579
15609599	TIG	-6.060606
15609840	SIGA	-7.142857
15607603	LPD	-7.692308
15607609	FADB2	-8.333333
15607162	hypothetical protein Rv0020c	-8.510638
15610082	LPPX	-9.090909
15607425	CMP	-9.302326
15607195	RPSF	-10
15610123	HUPB	-10
15610852	hypothetical protein Rv3716c	-11.11111
15607573	SODC	-12.5
15609339	CbhK	-12.5
15607330	ILVD	-12.5
15610854	hypothetical protein Rv3718c	-15.38462
15609268	CYSQ	-16.66667
15607352	PEPCK	-17.91045
15608155	RPLY	-18.75
15609836	hypothetical protein Rv2699c	-20
15608351	hypothetical protein Rv1211	-20
15610130	HHDD ISOMERASE	-25
15610053	FFH	-25
15610578	RPSI	-27.27273
15609920	GPSI	-28.57143
15607423	hypothetical protein Rv0282	-28.57143
15609886	hypothetical protein Rv2749	-28.57143
15610728	hypothetical protein Rv3592	-30
15611001	hypothetical protein Rv3865	-33.33333
15609315	AroG	-33.33333
15609661	FAS	-33.96226
15609975	RBFA	-37.5
15609231	TATA	-40
15608123	PEPD	-40
15610543	hypothetical protein Rv3407	-50
15607717	hypothetical protein Rv0577	-50

Locus ID	Name of protein	% change
15609917	TB43	-50
15611017	Ala-Gly rich protein	-56.25
15610077	MAS	-75
15607392	HSP	-78.04878
57116734	Icl	-87.27273

Table S2 - Comparison for the expression of protein profile between *M. bovis* and *M. bovis* / pLK102.

Locus ID	Name of the protein	% change
15608755	PykA	2150
31792208	RPPP	275
31791646	Icl	261.5385
31791389	PEPCK	257.8947
31792327	Homocysteine methyl transferase	235.7143
31793959	PNPASE	182.3529
31793066	BFRA	166.6667
31793705	FAS	136.3636
31792812	Two component regulatory system	133.3333
31791184	CMP	133.3333
31795093	HP	125
31795086	TRXB2	125
31794223	ADHC	118.75
31795015	BFRB	115.3846
31794065	TS	114.2857
31792642	TAL	114.2857
31794818	CSPA	110
31792816	30S ribosomal protein S1	104.7619
31793643	TF	100
31792371	FDXC	100
31791333	PNTA	100
31794760	TF	100
31791823	NUSG	100
31794570	LPQD	100
31791459	PPE family protein	100
31794148	OXIDOREDUCTASE	100
31792120	PSTS2	87.5
31793648	PEPN	87.5
31793875	Sigma Factor	81.81818
31793658	Macrolide-transport ABC transporter	80
31791758	Transcriptional regulatory protein	80
31791892	30S ribosomal protein S3	77.77778
31792490	RHO	77.77778
31794882	2-isopropylmalate synthase	72.22222
31792670	CAN	68
31794890	hypothetical protein Mb3745c	66.66667

Locus ID	Name of the protein	% change
31791819	hypothetical protein Mb0654	66.66667
31794850	ATPASE	66.66667
31795083	CMP	66.66667
31794403	short chain dehydrogenase	63.63636
31791420	3-ketoacyl-(acyl-carrier-protein) reductase	63.63636
31791641	dihydrolipoamide dehydrogenase	60
31791851	DNA-directed RNA polymerase beta subunit	60
31795026	HNS	60
31792648	QOR	60
31793026	hypothetical protein Mb1867c	60
31792291	FUM	58.33333
31791852	DNA-directed RNA polymerase beta' subunit	58.06452
31794117	MAS	57.89474
31793077	Antigen 85B complex	57.14286
31794207	FIXA	55.55556
31794975	acyl-CoA synthase	55.55556
31791826	MMAA4	54.54545
31792206	50S ribosomal protein L25	50
31794118	acyl-CoA synthase	50
31793356	AroG	50
31794428	S-adenosyl-L-homocysteine hydrolase	47.27273
31792674	MOXR1	45.16129
31795049	ESAT-6	42.85714
31791905	50S ribosomal protein L6	42.85714
31791868	EF-2	40.90909
31791419	hypothetical protein Mb0247c	40
31794635	30S ribosomal protein S11	40
31793325	WAG31	40
31792500	ATPH	38.46154
31794894	hypothetical protein Mb3749c	37.5
31793050	MPT32	37.5
31794798	inorganic pyrophosphatase	37.5
31793791	pyridoxine biosynthesis protein	37.5
31792433	malate dehydrogenase	37.2093
31791867	30S ribosomal protein S7	36.84211
31794460	ACCD5	36
31792019	hypothetical protein Mb0854c	35.71429
31794599	GroES	35.19164
31793626	NDK	33.33333
31791902	50S ribosomal protein L5	33.33333
31794620	50S ribosomal protein L13	33.33333

Locus ID	Name of the protein	% change
31792404	hypothetical protein Mb1243	33.33333
31792828	IF-3	33.33333
31794973	ACCD4	33.33333
31794112	PPSE	33.33333
31791647	3-hydroxybutyryl-CoA dehydrogenase	30.76923
31792503	ATP synthase subunit B	29.85075
31794598	GroEL	29.54545
31791421	acetyl-CoA acetyltransferase	29.03226
31794371	hypothetical protein Mb3219c	28.57143
31791885	30S ribosomal protein S10	28.57143
31793883	IDER	28.57143
31792586	SAM synthase	28.20513
31793383	hypothetical protein Mb2227c	27.27273
31794633	DNA-directed RNA polymerase alpha subunit	26.31579
31793085	hypothetical protein Mb1926	25
31792261	enoyl-CoA hydratase	25
31794636	30S ribosomal protein S13	25
31791511	hypothetical protein Mb0340	25
31791825	50S ribosomal protein L1	25
31791425	succinate dehydrogenase	23.07692
31792141	succinyl-CoA synthetase alpha subunit	22.5
31791919	adenylate kinase	22.22222
31793027	Malate synthase	22.22222
31794961	NDK regulator	20
31792845	argininosuccinate synthase	20
31791835	50S ribosomal protein L10	18.75
31792441	alpha-ketoglutarate decarboxylase	18.18182
31791186	PPIA	17.94872
31794888	hypothetical protein Mb3743c	17.64706
31792229	ESAT-6	16.66667
31791814	enoyl-CoA hydratase	16.66667
31791824	50S ribosomal protein L11	16.66667
31793059	Probable reductase	16.66667
31795055	Conserved hypothetical alanine and glycine rich protein	16.66667
31792119	PKNDa	16.66667
31794773	CLPC	15.90909
31793363	hypothetical protein Mb2207c	15.78947
31791916	hypothetical protein Mb0751	14.28571
31794685	hypothetical protein Mb3539c	14.28571
31791182	DNA topoisomerase IV subunit B	14.28571

Locus ID	Name of the protein	% change
31793017	CFP17	12.5
31791908	50S ribosomal protein L30	12.5
31791986	CFP29	12.5
31794110	PPSC	12.5
31794761	LPQE	11.11111
31792269	PRA antigien homolog	11.11111
31791213	hypothetical protein Mb0037c	11.11111
31792501	ATP synthase subunit A	10.86957
31793159	MPT64	10
31793808	Methylase	9.090909
31792657	CONSERVED ABC TRANSPORTER	9.090909
31792502	ATP synthase subunit C	8.333333
31793392	leucyl aminopeptidase	8.333333
31794172	SERA1	8.108108
31791561	CLPB	7.843137
31794634	30S ribosomal protein S4	7.692308
31794426	MTRA	7.692308
31792630	glyceraldehyde-3-phosphate dehydrogenase	5
31792084	citrate synthase	3.703704
31793424	acyl carrier protein	3.658537
31792582	MIHF	3.125
31791462	CMP	2.777778
31791618	GroEL	0
31791528	DnaK	0
31791836	50S ribosomal protein L7/L12	0
31794066	30S ribosomal protein S2	0
31791903	30S ribosomal protein S14	0
31794769	hypothetical protein Mb3623	0
31793871	hypothetical protein Mb2718c	0
31794085	30S ribosomal protein S16	0
31791617	short chain dehydrogenase	0
31792353	pterin-4-alpha-carbinolamine dehydratase	0
31794454	FADE25	0
31794923	hypothetical protein Mb3779c	0
31794122	LPPX	0
31791230	30S ribosomal protein S6	0
31793421	pyruvate dehydrogenase subunit E1	0
31792416	HTRA	0
31792368	hypothetical protein Mb1207c	0
31793315	hypothetical protein Mb2159c	0
31791610	SODC	0

Locus ID	Name of the protein	% change
31793806	CMP	0
31794470	L-lysine aminotransferase	0
31794169	HHDD ISOMERASE	0
31793402	GLNA2	0
31792520	TRx	0
31792640	OPCA	0
31793927	dihydrodipicolinate synthase [	0
31795053	Hypothetical alanine and praline rich protein	0
31794414	hypothetical protein Mb3262c	0
31794382	hypothetical protein Mb3230c	0
31792386	acyl-CoA synthase	0
31791197	hypothetical protein Mb0020c	-2.17391
31791426	succinate dehydrogenase	-2.94118
31791655	HBHA	-4.44444
31794058	RRF	-5
31792048	FADB	-5.26316
31791529	GRPE protein	-5.55556
31792499	ATP synthase subunit B	-7.14286
31793320	hypothetical protein Mb2164c	-7.69231
31791869	elongation factor Tu	-7.7381
31793394	dihydrolipoamide acyltransferase	-8.43373
31791907	30S ribosomal protein S5	-9.09091
31794206	FIXB	-9.23077
31794465	ACCA3	-11.1111
31792140	succinyl-CoA synthetase subunit beta	-11.9048
31794933	LPQH	-12.5
31791659	CMP	-12.5
31792643	transketolase	-14.2857
31792286	serine hydroxymethyltransferase	-14.2857
31794569	POSSIBLE DEHYDROGENASE	-14.2857
31794284	peptide chain release factor 2	-14.2857
31791461	hypothetical protein Mb0290	-15
31792265	acetyl-CoA acetyltransferase	-16.129
31793274	Probable membrane protein	-16.6667
31793101	KATG	-18.0645
31792632	triosephosphate isomerase	-18.1818
31792093	enoyl-CoA hydratase	-20
31793425	3-oxoacyl-(acyl carrier protein) synthase	-20
31792271	GreA	-20
31792631	phosphoglycerate kinase	-20
31794881	aspartate kinase	-20

Locus ID	Name of the protein	% change
31791886	50S ribosomal protein L3	-20
31794463	SSEA	-20
31792072	phosphoserine aminotransferase	-20
31794052	MPB70	-20.2128
31792822	hypothetical protein Mb1662	-20.9302
31793400	GLNA1	-21.8182
31791469	CMP	-22.2222
31791762	hypothetical protein Mb0595c	-23.0769
31791887	50S ribosomal protein L4	-23.0769
31792002	SSEC2	-24.3902
31791976	phosphoribosylformylglycinamidine synthase	-25
31792981	Putative ESAT-6 like protein 5	-25
31793238	30S ribosomal protein S18	-25
31795012	prephenate dehydratase	-25
31791681	pyrroline-5-carboxylate reductase	-25
31792003	CYSA2	-28.0488
31792006	Transcriptional regulatory protein	-28.5714
31792047	acetyl-CoA acetyltransferase	-28.5714
31791463	CMP	-28.75
31795048	CFP10	-29.8246
31793918	Conserved 35 kDa Alanine rich protein	-32.2581
31793426	3-oxoacyl-(acyl carrier protein) synthase	-33.3333
31793480	hypothetical protein Mb2324	-33.3333
31793647	hypothetical protein Mb2493c	-33.3333
31791243	ICD2	-33.3333
31793646	ribose-5-phosphate isomerase B	-33.3333
31792679	enoyl-(acyl carrier protein) reductases	-33.3333
31794178	ketol-acid reductoisomerase	-33.3333
31792339	Short chain type dehydrogenase / reductases	-33.3333
31794846	Transcrptional regualatory protein	-35.7143
31791540	fructose-bisphosphate aldolase	-35.7143
31791712	hypothetical protein Mb0543	-36.3636
31794723	hypothetical protein Mb3577	-37.5
31791530	DNAJ1	-37.5
31791326	Short chain type dehydrogenase / reductases	-37.931
31791471	CMP	-38.4615
31794224	hypothetical protein Mb3072c	-40
31792135	glucose-6-phosphate isomerase	-40
31791195	PPP	-40
31793062	LLDD2	-41.3793
31791231	single-strand DNA-binding protein	-41.6667

Locus ID	Name of the protein	% change
31792960	OXIDOREDUCTASE	-41.6667
31795087	TRXC	-42.8571
31795054	hypothetical protein Mb3910c	-42.8571
31793805	hypothetical protein Mb2652c	-42.8571
31793277	twin argininte translocase protein A	-42.8571
31793108	Isocitrate lyase	-44
31791556	SECE2	-44.4444
31793375	QCRB	-44.4444
31794848	hypothetical protein Mb3702c	-46.1538
31794162	HUPB	-50
31793788	hypothetical protein Mb2635c	-50
31791394	hypothetical protein Mb0222	-51.5152
31794974	PKS13	-55.814
31795020	SODA	-56.6667
31793642	CLP	-56.7568
31793893	Conserved transmembrane alanine and glycine protein	-57.1429
31792683	CP	-68
31793474	haloalkane dehalogenase	-69.2308
31792458	PKNH	-70
31794050	MPB83	-72.2543
31793608	AhpC	-73.3333
31794906	hypothetical protein Mb3761c	-80
31794346	hypothetical protein Mb3194	-84
31791429	HSP	-89.2857

REFERENCES

1. Alcaide F, Calatayud L, Santin M & Martin R. (2004). Comparative *in vitro* activities of linezolid, telithromycin, clarithomycin, levofloxacin, moxifloxacin and four conventional antimycobacterial drugs against *M.kansasii. Antimicrobial Agents and Chemotherapy*, **48,** 4562-4565.

2. Alvarez HM & Steinbüchel A. (2002). Triacylglycerols in prokaryotic microorganisms. *Applied Microbiology and Biotechnology*, **60,** 367–376.

3. Andersen AB, Andersen P & Ljungqvist L. (1992). Structure and function of a 40,000-molecular-weight protein antigen of *M. tuberculosis. Infection and Immunity*, **60,** 2317-2323.

4. Anderson AB & Wood WA. (1969). Carbohydrate metabolism in bacteria. *Annual Reviews in Microbiology*, **23,** 539–578.

5. Andrews SC (1998). Iron storage in bacteria. *Advances in Microbial Physiology*, **40,** 281-351.

6. Aoyama Y, Horiuchi T, Gotoh O, Noshiro M & Yoshida Y. (1998). CYP51-like gene of *M.tuberculosis* actually encodes a P450 similar to eukaryotic CYP51. *Journal of Biochemistry*, **124,** 694-696.

7. Arounaguiri S, Easwaramoorthy D, Ashok kumar A, Aparna Datta Gupta & Maiya BG. (2000). Cobalt(III), nickel(II), and ruthenium(II) complexes of 1,10-phenanthroline family of ligands: DNA binding photocleavagestudies. *Proceedings of Indian Academy Sciences (Chemical sciences)*, **112,** 1-17.

8. Arruda S, Bomfim G, Knights R, Huima-Byron T & Riley LW. (1993). Cloning of an *M. tuberculosis* DNA fragment associated with entry and survival inside cells. *Science*, **261,** 1454-1457.

9. Banu S, Honore N, Saint-Joanis B, Philpott D, Prevost MC & Cole ST. (2002). Are the PE-PGRS proteins of *M. tuberculosis* variable surface antigens? *Molecular Microbiology*, **44,** 9-19.

10. Barclay R &Ratledge C. (1986). Participation of iron in the growth inhibition of pathogenic strains of *M. avium* and *M. paratuberculosis* in serum. *Z Bakt Mikrobiol Hyg*, **A262,** 189-194.

11. Barclay R, Ewing DF. & Ratledge C. (1986). Isolation, identification and strucural analysis of the Mycobactins of *M. avium, M. intracellulare, M. scrofuloceum* and *M. paratuberculosis. Journal of Bacteriology*, **164,** 896-903.

12. Besra GS & Chatterjee D. (1994). Lipids and Carbohydrates of *M. tuberculosis.* In *Tuberculosis: Pathogenesis, Protection, and control*. Edited by B. R. Bloom. Washington, D.C.: American Society for Microbiology. 285-306.

13. Bittner MJ, Horowitz EA, Safranek TJ & Preheim LC. (1996). Emergence of *M. kansasii* as the leading mycobacterial pathogen isolated over 20-year period at a Midwestern veterans affairs hospital. *Clinical Infectious Diseases,* **22,** 1109-1110.

14. Bloch K. **(**1975). Fatty acid synthases from *Mycobacterium phlei*. *Methods in Enzymology,* **35,** 84-90.

15. Bloch KC, Zwerling L, Pletcher MJ, Hahn JA, Gerberding JL, Ostroff SM, Vugia DJ & Reingold AL. (1998). Incidence and clinical implications of isolation of *M. kansasii*: results of a 5-year, population based study. *Annals of Internal medicine*, **129,** 698-704.

16. Bordier C. (1981). Phase separation of integral membrane proteins in Triton X-114 solution. *The Journal of Biological Chemistry,* **256,** 1604-1607.

17. Bradford MM. (1976). A rapid and sensitive for the quantitation of microgram quantitites of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry,* **72,** 248-254.

18. Braibant M, Gilot P. & Content J. (2000). The ATP binding cassette (ABC) transport systems of *M. tuberculosis*. *FEMS Microbiology Reviews,* **24**, 449-467.

19. Brennan MJ, Delogu G, Chen Y, Bardarov S, Kriakow J, Alavi M & Jacobs WR Jr. (2001). Evidence that mycobacterial PE_PGRS proteins are cell surface constituents that influence interactions with other cells. *Infection and Immunity,* **69**, 7326-7333.

20. Brennan PJ. (2003). Structure, function and biogenesis of the cell wall of *M. tuberculosis. Tuberculosis,* **83,** 91-97.

21. Britton WJ & Lockwood DNJ. (2004). Leprosy. *The Lancet,* **363,** 1209-1219.

22. Brodie D & Schulger NW. (2005). The diagnosis of tuberculosis. *Clinical Chest Medicine,* **26,** 247-271.

23. Brosch R, Gordon SV, Buchrieser C, Pym AS, Garnier T & Cole ST. (2000a). Comparative genomics uncovers large tandem chromosomal duplications in *M. bovis* BCG Pasteur. *Yeast,* **17,** 111-123.

24. Brosch R, Gordon SV, Eiglmeier K, Garnier T & Cole ST. (2000). Comparative genomics of the leprosy and tubercle bacilli. *Research in Microbiology,* **151,** 135-142.

25. Brosch R, Gordon SV, Marmiesse M, Brodin P, Buchrieser C, Eiglmeier K, Garnier T, Gutierrez C, Hewinson G, Kremer K, Parsons LM, Pym AS,

Samper S, van Soolingen D & Cole ST. (2002)**.** A new evolutionary scenario for the *M. tuberculosis* complex. *Proceedings of National Academy of Sciences,* **99,** 3684-3689.

26. Brosch R, Gordon SV, Marmiesse M, Brodin P, Buchrieser C, Eiglmeier K, Garnier T, Gutierrez C, Hewinson G, Kremer K, Parsons LM, Pym AS, Samper S, van Soolingen D, & Cole ST. (2002). A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proceedings of the National Academy of Sciences,* **99,** 3684-3689.

27. Brosch R, Gordon SV, Pym A, Eiglmeier K, Garnier T & Cole ST. (2000). Comparative genomics of the mycobacteria. *International Journal of Medical Microbiology*. **290,** 143-152.

28. Brosch R, Pym AS, Gordon SV & Cole ST. (2001). The evolution of mycobacterial pathogenicity: clues from comparative genomics. *Trends in Microbiology,* **9**, 452-458.

29. Brown KA & Ratledge C. (1975). Iron transport in *Mycobacterium smegmatis* : Ferrimycobactin reductase , the enzyme releasing iron from its carrier . *FEBS letters,* **53,** 262-266.

30. Brown KA & Ratledge C. (1975). The effect of p-aminisalicylic acid on iron transport and assimilation in mycobacteria. *Biochimie et Biophysics acta,* **385,** 207-220.

31. Canueto-Quintero J, Caballero-Granado FJ, Herrero-Romero M, Dominguez-Castellano A, Martin–Rico A, Verdu EV, Santamaria DS, Cerquera RC & Torres-Tortosa M. (2003). Epidemiological, clinical, and prognostic differences between the diseases caused by *M. kansasii* and *M. tuberculosis* in patients infected with Human immunodeficiency virus: a multicenter study. *Clinical Infectious diseases*, **37,** 584-590.

32. Card GL, Peterson NA, Smith CA, Rupp B, Schick BM & Baker EN. (2005). The crystal structure of Rv1347c, a putative antibiotic resistance protein from *Mycobacterium tuberculosis*, reveals a gcn5-related fold and suggests an alternative function in siderophore biosynthesis. *The journal of Biological Chemistry,* **280**, 13978–13986.

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

34. Chatterjee D. (1997). The mycobacterial cell wall: structure, biosynthesis and sites of drug action. *Current Opinion Chemical Biology,* **1**, 579-588.

35. Chipperfield, JR & Ratledge C. (2000). Salicylate is not a bacterial siderophore: A theoretical study. *Biometals,* **13,** 165-168.

36. Claes WA, Puhler A, & Kalinowski J. (2002). Identification of two *prpDBC* gene clusters in *C. glutamicum* and their involvement in propionate degradation via the 2-methylcitrate cycle. *Journal of Bacteriology*, **184,** 2728–2739.

37. Coldham NG & Woodward MJ. (2004). Charactersiation of the *S. typhimurium* proteome by semi-automated two-dimensional HPLC-mass spectrometry: detection of proteins implicated in multiple antibiotic resistance. *Journal of Proteome research*, **3,** 595-603.

38. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE 3rd, Tekaia F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Barrell BG & *et al*. (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature*, **393**, 537-544.

39. Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honore N, Garnier T, Churcher C, Harris D, Mungall K, Basham D Brown, D, Chillingworth T, Connor R, Davies RM, Devlin K, Duthoy S, Feltwell T, Fraser A, Hamlin N, Holroyd S, Hornsby T, Jagels K, Lacroix C, Maclean J, Moule S, Murphy L, Oliver K, Quail MA, Rajandream MA, Rutherford KM, Rutter S, Seeger K, Simon S, Simmonds M, Skelton J, Squares R, Squares S, Stevens K, Taylor K, Whitehead S, Woodward JR. & Barrell BG. (2001). Massive gene decay in the leprosy bacillus. *Nature*, **409,** 1007-1011.

40. Cole ST. (1998). Comparative mycobacterial genomics. *Current Opinion in Microbiology*, **1,** 567-571.

41. Cole ST. (1999). Learning from the genome sequence of *Mycobacterium tuberculosis* H37Rv. *FEBS letters*, **452,** 7-10.

42. Cole ST. (2002). Comparative and functional genomics of the *M. tuberculosis* complex. *Microbiology*, **148,** 2919-2928.

43. Collins DM, Wilson T, Campbell S, Buddle BM, Wards BJ, Hotter G, & De Lisle GW. (2002). Production of avirulent mutants of *Mycobacterium bovis* with vaccine properties by the use of illegitimate recombination and screening of stationary-phase cultures. *Microbiology*, **148,** 3019-3027.

44. Corbett EL, Watt CJ, Walker N, Maher D, Williams GB, Raviglione MC & Dye C. (2003). The growing burden of tuberculosis: global trends and interaction with the HIV epidemic. *Archives of Internal Medicine*, **163,** 1009-1021.

45. Cousins DV. (2001). *Mycobacterium bovis* infection and control in domestic livestock. *Revue Scientific et technique Office International des Epizooties,* **20,** 71-85.

46. Crick DC, Mahapatra S & Brennan PJ. (2001). Biosynthesis of the arabinogalactan–peptidoglycan complex of *M. tuberculosis. Glycobiology,* **11,** 107R–108R.

47. Daffe M & Draper P. (1998). The envelope layers of mycobacteria with reference to their pathogenicity. *Advances in Microbial Physiology,* **39,** 131-203.

48. Dahl JL, Kraus CN, Boshoff HIM, Doan B, Foley K, Avarbock D. (2003). The role of RelMtb-mediated adaptation to stationary phase in long-term persistence of *M. tuberculosis* in mice. *Proceedings of National Academy of Sciences,* **100,** 10026–10031.

49. Dailey TA & Dailey HA. (2002). Identification of (2Fe-2S) clusters in microbial ferrochelatase. *Journal Of Bacteriology,* **184,** 2460-2464.

50. Daniel J, Deb C, Dubey VS, Sirakova TD, Abomoelak B, Morbidoni HR & Kolattukudy PE. (2004). Induction of a novel class of diacylglycerol acyltransferases and triacylglycerol accumulation in *M. tuberculosis* and tuberculosis patients. *Proceedings of National Academy of Sciences,* **100,** 14321–14326.

51. De Voss JJ, Rutter K, Schroeder BG, & Barry CE III. (1999). Iron acquisition and metabolism by mycobacteria. *Journal of Bacteriology,* **181,** 4443-4451.

52. De Voss JJ, Rutter K, Schroeder BG, Su H, Zhu Y & Barry CE III. (2000). The salicylate derived mycobactin siderophores of *M. tuberculosis* are essential for growth in macrophages. *Proceedings of National Academy of Sciences,* **97,** 1252-1257.

53. Delogu G & Brennan PJ. (2001). Comparative immune response to PE and PE_PGRS antigens of *M. tuberculosis. Infection and Immunity,* **69,** 5606-5611.

54. Dobson G, Minnikin DE, Minnikin SM, Parlett JH, Goodfellow M, Ridell M & Magnusson M. (1985). Systemic analysis of complex mycobacterial lipids. *Chemical Methods in Bacterial Systematics,* p. 237-265.

55. Douvas GS, May MH & Crowle AJ. (1993). Transferrin, iron and serum lipids enhance or inhibit *M. avium* replication in human macrophages. *Journal of Infectious diseaseas,* **167,** 857-864.

56. Dover LG & Ratledge C. (1996). Identification of 29 kDa protein in the envelope of *M. smegmatis* as a putative ferriexochelin receptor. *Microbiology*, **142,** 1521-1530.

57. Dubnau E, Chan J, Mohan VP & Smith I. (2005). Responses of *M. tuberculosis* to growth in the mouse lung. *Infection and Immunity*, **73,** 3754–3757.

58. Dusserget O, Rodriguez M & Smith I. (1996). An *ideR* mutant of *M. smegmatis* has derepressed siderophore production and an altered oxidative stress response. *Molecular microbiology*, **22,** 535-544.

59. Fetherson JD, Bertolino VJ & Perry RD. (1999). YbtP and YbtQ two ABC transporter required for iron uptake in *Yersinia pestis. Molecular Microbiology*, **32,** 289-299.

60. Filliol I, Motiwala AS, Cavatore M, Qi W, Hazbon MH, Bobadilla del Valle M, Fyfe J, Garcia-Garcia L, Rastogi N, Sola C, Zozio T, Guerrero MI, Leon CI, Crabtree J, Angiuoli S, Eisenach KD, Durmaz R, Joloba ML, Rendon A, Sifuentes-Osornio J, Ponce de Leon A, Cave MD, Fleischmann R, Whittam TS & Alland D. (2006). Global phylogeny of Mycobacterium tuberculosis based on single nucleotide polymorphism (SNP) analysis: insights into tuberculosis evolution, phylogenetic accuracy of other DNA fingerprinting systems, and recommendations for a minimal standard SNP set. *Journal of Bacteriology*, **188,** 759-772.

61. Fiss EH, Yu S & Jacobs WR. (1994). Identification of genes involved in the sequestration of iron in mycobacteria: the ferric exochelin biosynthetic and uptake pathways. *Molecular microbiology*, **14,** 557-569.

62. Garmony HS & Titball RW. (2004). ATP- binding cassette transporters are targets for the development of antibacterial vaccines and therapies. *Infection and Immunity*, **72,** 66757-66763.

63. Garnier T, Eiglmeier K, Camus JC, Medina N, Mansoor H, Pryor M, Duthoy S, Grondin S, Lacroix C, Monsempe C, Simon S, Harris B, Atkin R, Doggett J, Mayes R, Keating L, Wheeler PR, Parkhill J, Barrell BG, Cole ST, Gordon SV & Hewinson RG (2003). The complete genome sequence of *M. bovis. Proceedings of the National Academy of Sciences,* **100,** 7877-7882.

64. Glassroth J. (2005). Tuberculosis 2004: Challenges and opportunities. *Transactions of the American clinical and climatological association,* **116,** 293-310.

65. Gobin J & Horwitz MA. (1996). Exochelins of *M. tuberculosis* remove iron from human iron binding proteins and donate iron to mycobactins in the *M. tuberculosis* cell wall. *Journal of Experimental Medicine*, **183,** 1527-1532.

66. Gobin J, Moore CH, Reeve JR Jr, Wong DK, Gibson BW & Horwitz MA. (1995). Iron acquisition by *M. tuberculosis*: isolation and characterisation of a family of iron binding exochelins. *Proceedings of National Academy of Sciences,* **92,** 5189-5193.

67. Gobin J, Wong DK, Gibson BW & Horwitz MA. (1999). Characterisation of exochelins of the *M. bovis* strain and BCG substrains. *Infection and Immunity*, **67,** 2035-2039.

68. Gold B, Rodriguez GM, Marras SAM, Pentecost M & Smith I. (2001). The *M. tuberculosis* IdeR is a dual functional regulator that controls transcription of genes involved in iron acquisition, iron storage and survival in macrophages. *Molecular Microbiology*, **42,** 851-865.

69. Gordon SV, Brosch R, Billault A, Garnier T, Eiglmeier K & Cole ST. (1999). Identification of variable regions in the genomes of tubercle bacilli using bacterial artificial chromosome arrays. *Molecular Microbiology,* **32**, 643-655.

70. Gordon SV, Brosch R, Eiglmeier K, Garnier T, Hewinson RG & Cole ST. (2002). Royal Society of Tropical Medicine and Hygiene Meeting at Manson House, London, 18th January 2001. Pathogen genomes and human health. Mycobacterial genomics. *Transactions of Royal society for tropical medicine and Hygiene.* **96,** 1-6.

71. Gordon SV, Eiglmeier K, Garnier T, Brosch R, Parknhill J, Barrell B, Cole ST & Hewinson RG. (2001). Genomics of *Mycobacterium bovis*. *Tuberculosis,* **81,** 157-163.

72. Gordon SV, Heym B, Parkhill J, Barrell B & Cole ST. (1999). New insertion sequences and a novel repeated sequence in the genome of *M. tuberculosis* H37Rv. *Microbiology*, **145,** 881-892.

73. Goren MB & Brennan PJ. (1980). Mycobacterial lipids: chemistry and biological activities. In: *Youmans GP, editor. Tuberculosis. Philadelphia:* W. B. Saunders, p. 63–193.

74. Grabner R & Meerbach W. (1991). Phagocytosis of surfactant by alveolar macrophages *in vitro*. *American Journal of Physiology Lung Cellular Molecular Physiology*, **261,** L472–L477.

75. Graybil JR & Bocanegra R. (2001). Treatment alternatives for *M. kansasii*. *Journal of Antimicrobials and Chemotherapy*, **47,** 417-420.

76. Guerin C. (1980). The history of BCG. In *BCG Vaccine: Tuberculosis-Cancer*. Rosenthal, S. (ed). Littleton, Mass: PSG Publishing, *pp*. 35-43.

77. Hall R, Sritharan M & Ratledge C. (1987). Iron transport in *M. smegmatis*. Occurance of iron regulated envelop proteins as potential receptors for iron uptake. *Journal of General Microbiology,* **133,** 2107-2114.

78. Heym B, Alzari PM, Honore N & Cole ST. (1995). Missense mutations in the catalase-peroxidase gene, *katG,* are associated with isoniazid resistance in *M. tuberculosis. Molecular Microbiology,* **15,** 235-245.

79. Heym B, Honore N, Truffot-Pernot C, Banarjee A, Schurra C, Jacobs Jr. WR, van Embden JDA, Grosset JH & Cole ST. (1994). The implications of the mutidrug resistance for the future of short course chemotherapyof tuberculosis: a molecular study. *Lancet,* **344,** 293-298.

80. Heym B, Stavropoulos E, Honore N, Domenech P, Saint-Joanis B, Wilson TM, Collins DM, Colston MJ & Cole ST. (1997). Effects of over expression of the alkyl hydroperoxide reductases AhpC on the virulence and isoniazid resistance of *M. tuberculosis. Infection and Immunity,* **65,** 1395-1401.

81. Hobson RJ, McBride AJA, Kempsell KE & Dale JW. (2002). Use of an assayed promoter probe library for the identification of macrophage regulated genes in *M. tuberculosis. Microbiology,* **148,** 1571-1579.

82. Horswill AR & Escalante-Semerena JC. (1999). *Salmonella typhimurium* LT2 catabolizes propionate via the 2-methylcitric acid cycle. *Journal of Bacteriology,* **181,** 5615–5623.

83. Hu J & Miller MJ. (1997). Total synthesis of a mycobactin S, a siderophore and growth promoter of *M. smegmatis,* and determination of its growth inhibitory activity against *M. tuberculosis. Journal of American Chemical Society,* **119,** 3462-3468.

84. Hutter B & Dick T. (1998). Increased alanine dehydrogenase activity during dormancy in *M. smegmatis*. *FEMS Microbiololgy Letters,* **167,** 7–11.

85. Hutter B & Singh M. (1999). Properties of the 40-kDa antigen of *M. tuberculosis*, a functional L-alanine dehydrogenase. *Biochemical Journal*, **343,** 669-672.

86. Jackson M, Crick DC & Brennan PJ. (2000). Phosphatidylinositol is an essential phospholipid of mycobacteria. *Journal of Biological chemistry*, **275,** 30092–30099.

87. Keating LA, Wheeler PR, Mansoor H, Inwald JK, Dale J, Hewinson RG & Gordon SV. (2005). The pyruvate requirement of some *M. tuberculosis* complex members is due to an inactive pyruvate kinase: implications for in vivo growth. *Molecular Microbiology,* **56,** 163-174.

88. Kellam P. (2001). Post-genomic virology: the impact of bioinformatics, microarrays and proteomics on investigating host and pathogen interactions. *Reviews in medical virology,* **11**, 313-329.

89. Kent PT & Kubica GP. (1985). Quality control in the mycobacteriology laboratory. In *Public Health mycobacteriology, A guide for the level III laboratory*. pp. 189-202.

90. Khasnobis S, Escuyer VE & Chatterjee D. (2002). Emerging therapeutic targets in tuberculosis: post-genomic era. *Expert Opinions on Therapeutic targets,* **6,** 21-40.

91. Kinter M & Sherman NE. Protein sequencing and identification using tandem mass spectrometry. Chapter 6.

92. Kochan I. (1976). Role of iron in regulation of nutritional immunity. *Bioorganic chemistry*, **2,** 55-57.

93. Kondo E, Suzuki K, Kanai K & Yasuda T. (1985). Liposomes-mycobacteria incubation systems as a partial model of host–parasite interaction at cell membrane level. *Japan Journal of Medicinal Science Biology,* **38,** 169–180.

94. Kornberg HL & Krebs HA. (1957). Synthesis of cell constituents from C2-units by a modified tricarboxylic acid cycle. *Nature,* **179,** 988–991.

95. Laemmli UK. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature,* **227,** 680-684.

96. Lane SJ, Marshall PS, Upton RJ & Rateldge C. (1998). Isolation and characterisation of carboxymycobactins as the second extracelllular siderophores in *M. smegmatis. Biometals,* **11,** 13-20.

97. Lane SJ, Marshall PS, Upton RJ, Ratledge C & Ewing M. (1995). Novel extracellular mycobactins, the carboxymycobactins from *M. avium. Tetrahedron letters* 36, 4129-4132.

98. Larsen TM, Benning MM, Wesenberg GE, Rayment I & Reed GH. (1997). Ligand-induced domain movement in pyruvate kinase: structure of the enzyme from rabbit muscle with Mg^{2+}, K^{+}, and L-phospholactate at 2.7 Å resolutions. *Archives of Biochemistry and Biophysics,* **345**, 199-206.

99. Lehmann KB & Neumann R. (1896). *Atlas undGrundriss der Bakteriologie und Lehrbuch der speciellen bakteriologischen*. Munich.

100. Liu K, Yu J, & Russell DG. (2003). pckA-deficient *Mycobacterium bovis* BCG shows attenuated virulence in mice and in macrophages. *Microbiology,* **149,** 1829-1835.

101. Maciaq A, Dainese E, Rodriguez GM, Milano A, Provvedi R, Pasca MR, Smith I, Palu G, Riccardi G & Manganelli R. (2007). Global analysis of the *M. tuberculosis* Zur (FurB) regulon. *Journal of Bacteriology*, **189,** 730-740.

102. Mahairas GG, Sabo PJ, Hickey MJ, Singh DC & Stover CK. (1996). Molecular analysis of genetic differences between *M. bovis* BCG and virulent *M. bovis*. *Journal of Bacteriology*, **178**, 1274-1282.

103. Manabe YC, Hatem CL, Kesavan AK, Durack J & Murphy JR. (2005). Both *C. diptheriae* DtxR (E175K) and *M. tuberculosis* IdeR (D177K) are dominant repressors of IdeR-regulated genes in *M. tuberculosis. Infection and Immunity*, **73,** 5988-5994.

104. Manabe YC, Saviola BJ, Sun L, Murphy JR, Bishai WR. (1999). Attenuation of virulence in *M. tuberculosis* expressing a constitutively active iron repressor. *Proceedings of National Academy of Sciences,* **96,** 12844-12848.

105. Markwell MK, Hass SM, Bieber LL & Tolbert NE. (1978). A modification of the Lowry procedure to simplify protein determination in membrane and lipoprotein samples. *Analytical Biochemistry,* **87,** 206-210.

106. Mason RJ, Stossel TP & Vaughan M. (1972). Lipids of alveolar macrophages, polymorphonuclear leukocytes, and their phagocytic vesicles. *Journal of Clinical Investigation,* **51,** 2399–2407.

107. McKinney JD, Honer zu Bentrup K, Munoz-Elias EJ, Miczak A, Chen B, Chan WT, Swenson D, Sacchettini JC, Jacobs WR, Jr, & Russell DG. (2000). Persistence of *Mycobacterium tuberculosis* in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase. *Nature*, **406**, 735-738.

108. Mdluli K, Slayden RA, Zhu Y, Ramaswamy S, Pan X, Mead D, Crane DD, Musser JM & Barry CE III. (1998). Inhibition of *M. tuberculosis* b-ketoacyl ACP synthase by isoniazid. *Science*, **280,** 1607-1610.

109. Meyers PR, Bourn WR, Steyn LM, van Helden PD, Beyers AD & Brown GD. (1998). Novel method for rapid measurement of growth of mycobacteria in detergent-free media. *Journal of Clinical Microbiology*, **36**, 2752-2754.

110. Milano A, Forti F, Sala C, Riccardi G & Ghisotti D. (2001). Transcriptional regulation of *furA* and *katG* upon oxidative stress in *M. smegmatis. Journal of Bacteriology,* **183,** 6801-6806.

111. Mostowy S, Inwald J, Gordon S, Martin C, Warren R, Kremer K, Cousins D. & Behr MA. (2005). Revisiting the evolution of *M. bovis. Journal of Bacteriology*, **187,** 6386-6395.

112. Muirhead H, Clayden DA, Barford D, Lorimer CG, Fothergill-Gilmore LA, Schiltz E, & Schmitt, W. (1986). The structure of cat muscle pyruvate kinase. *EMBO Journal,* **5,** 475-481.

113. Munoz ME & Ponce E. (2003). Pyruvate kinase: current status of regulatory and functional properties. *Comparative Biochemistry, Physiology, B Biochemistry and Molecular Biology*, **135,** 197-218.

114. Muñoz-Elías EJ & McKinney JD. (2005). *Mycobacterium tuberculosis* isocitrate lyases 1 and 2 are jointly required for *in vivo* growth and virulence. *Nature Medicine*, **11,** 638–644.

115. Ochs DC, Mc Conkey EH & Sammons DW. (1981). Silver stains for proteins in polyacrylamide gels: A comparison of six methods. *Electrophoresis,* **2,** 304-310.

116. Oliver S. (2000). Guilt-by-association goes global. *Nature,* **403,** 601-603.

117. Olsen I, Tryland M, Wiker HG & Reitan LJ. (2001). AhpC, AhpD, and a secreted 14-kilodalton antigen from *M.avium* Subsp. p*aratuberculosis* distinguish between paratuberculosis and bovine tuberculosis in an Enzyme-Linked Immunosorbent Assay. *Clinical and diagnostic Laboratory Immunology,* **8,** 797-801.

118. Parish T, Lewis J & Stoker NG. (2001). Use of the mycobacteriophage L5 excisionase in *M. tuberculosis* to demonstrate gene essentiality. *Tuberculosis,* **81**, 359-364.

119. Paterson AB. (1948). The production of bovine tuberculoprotein. *Journal of Comparative Pathology,* **58,** 302-313.

120. Pessolani MCV, Smith DR, Rivoire B McCormick J, Hefta SA, Cole ST & Brennan PJ. (1994). Purification, charecterisation, gene sequence and significanc of a bacterioferritin from *M. leprae. Journal of Experimental Medicine,* **180,** 319-327.

121. Pohl E, Holmes RK & Hol WG. (1999). Crystal structure of the Iron dependent regulator (Ide R) from *M. tuberculosis* shows both metal binding sites fully occupied. *Journal of molecular biology*, **285,** 1145-1156.

122. Pym AS, Domenech P, Honore N, Song J, Deretic V & Cole ST. (2001). Regulation of catalase-peroxidase (KatG) expression, isoniazid activity and virulence by *furA* of *M. tuberculosis. Molecular Microbiology,* **40,** 879-889.

123. Quadri LE, Sello J, Keating TA, Weinreb PH & Walsh CT. (1998). Identification of a *M. tuberculosis* gene cluster encoding the biosynthetic enzymes for the assembly of the virulence conferring siderophores mycobactin. *Chemical Biology,* **5,** 631-645.

124. Ramakrishnan L, Federspiel NA & Falkow S. (2000). Granuloma-specific expression of *Mycobacterium* virulence proteins from the gycine-rich PE-PGRS family. *Science,* **547**, 1436-1439.

125. Ratledge C & Brown KA. (1972). Inhibition of mycobactin formation in *M. smegmatis* by p-aminosalicylic acid. A new proposal for the mode of action of PAS. *American reviews in Respiratory Diseases,* **106,** 774-776.

126. Ratledge C & Dover LG. (2000). Iron metabolism in pathogenic mycobacteria. *Annual reviews in microbiology,* **54,** 881-941.

127. Ratledge C & Ewing M. (1996). The occurrence of carboxymycobactin, the siderophore of pathogenic mycobacteria, as a second extracellular siderophore in *M. smegmatis. Microbiology,* **142,** 2207-2212.

128. Ratledge C, Patel PV & Mundy J. (1982). Iron transport in *M. smegmatis:* the location of mycobactin by electron microscopy. *Journal of general Microbiology,* **128,** 1559-1565.

129. Ratledge C. & Ewing M. (1996). The occurrence of carboxymycobactin, the siderophore of pathogenic mycobacteria, as a second extracellular siderophore in *M. smegmatis. Microbiology,* **142,** 2207-2212.

130. Ratledge C. (1982). Nutrition, growth and metabolism. In *Biology of the Mycobacteria.* Vol. 1. Ratledge C and Stanford JL. (eds): Academic Press, Inc. (London) Ltd., London, pp. 186-212.

131. Ratledge C. (2004). Iron, mycobacteria and tuberculosis. *Tuberculosis,* **84,** 110-130.

132. Raynaud C, Guilhot C, Rauzier J, Bordat Y, Pelicic V & Manganelli R. (2002) Phospholipases C are involved in the virulence of *M. tuberculosis. Molecular Microbiology,* **45,** 203–217.

133. Riley LW. (2006). Of mice, men and elephants: *M. tuberculosis* cell envelope lipids and pathogeneisis. *Journal of Clinical Investigation,* **116,** 1475-1478.

134. Rodriguez GM & Smith I. (2006). Identification of an ABC transporter required for iron acquisition and virulence in *M. tuberculosis. Journal of Bacteriology,* **188,** 424-430.

135. Rodriguez GM, Voskuil MI, Gold B, Schoolnik GK & Smith I. (2002). *ideR,* an essential gene in *M. tuberculosis:* role of IdeR in iron-dependent gene expression, iron metabolism and oxidative stress response. *Infection and Immunity,* **70,** 3371-3381.

136. Rosenkrands I, Slayden RA, Crawford J, Aagaard C, Barry III, CE & Andersen P. (2002). Hypoxic response of *M. tuberculosis* studied by metabolic labeling and proteome analysis of cellular and extracellular proteins. *Journal of Bacteriology,* **184,** 3485-3491.

137. Sambrook J, Fritsch EF & Maniatis T. (1989). In Molecular cloning, A laboratory manual.

138. Sassetti C & Rubin EJ. (2002). Genomic analysis of microbial virulence. *Current Opinion in Microbiology*, **5,** 27-32.

139. Sassetti CM & Rubin EJ. (2003). Genetic requirements for mycobacterial survival during infection. *Proceedings of National Academy of Sciences*, **100,** 12989–12994.

140. Sassetti CM & Rubin EJ. (2003). Genetic requirements for mycobacterial survival during infection. *Proceedings of the National Academy of Sciences*, **100,** 12989-12994.

141. Schnappinger D, Ehrt S, Voskuil MI, Liu Y, Mangan JA, Monahan IM, Dolganov G, Efron B, Butcher PD, Nathan C & Schoolnik GK. (2003). Transcriptional adaptation of *M. tuberculosis* within Macrophages: Insights into the Phagosomal Environment. *Journal of Experimental Medcine*, **198,** 693-704.

142. Schroeder EK, de Souza ON, Santos DS, Blanchard JS & Basso LA. **(**2002). Drugs that inhibit mycolic acid biosynthesis in *M. tuberculosis*. *Current Pharmaceutical Biotechnology,* **3,** 197- 225.

143. Schwyn B & Neilands JB. (1987). Universal chemical assay for the detection and determination of siderophores. *Analytical Biochemistry*, **160,** 47-56.

144. Scorpio A & Zhang Y. (1996). Mutations in *pncA*, a gene encoding pyrazinamidase/nicotinamidase, casuse resistance to the antituberculous drug pyrazinamide in tubercle bacillus. *Nature Medicine*, **2,** 662-667.

145. Sharman DJ, Williams DH, Ewing DF & Ratledge C. (1995a). Determination of structure of exochelin MN, the extracellular siderophores from *M. neoaurum. Chemical Biology,* **2,** 553-561.

146. Sharman DJ, Williams DH, Ewing DF & Ratledge C. (1995b). Isolation purification and structure of exochelin MS, the extracellular siderophores from *M. smegmatis*. *Biochemistry Journal*, **305,** 187-196.

147. Sherman DR, Voskuil M, Schnappinger D, Liao R, Harrell MI, Singh KK, Zhang X, Patibandla AS, Chien P Jr. & Laal S. (2001). Antigens of *M. tuberculosis* expressed during preclinical tuberculosis: Serologicsl

immunodominance of proteins with repetetive amino acid sequences. *Infection and Immunity,* **69**, 4185-4191.

148. Sinha S, Arora S, Kosalai K, Namane A, Pym AS, & Cole ST. (2002). Proteome analysis of the plasma membrane of *M. tuberculosis. Comparative and Functional Genomics,* **3,** 470-483.

149. Slayden RA & Barry CE III. (2002). The role of KasA and KasB in the biosynthesis of meromycolic acids and isoniazid resistance in *M. tuberculosis*. *Tuberculosis,* **82,** 149-160.

150. Smith I. (2003). *Mycobacterium tuberculosis* pathogenesis and molecular determinants of virulence. *Clinical Microbiology Reviews,* **16,** 463-496.

151. Snow GA. (1970). Mycobactins: iron chelating growth factors from mycobacteria. *Bacteriological Reviews,* **34,** 99-125.

152. Spigelman M & Gillipsie S. (2006). Tuberculosis drug development pipeline: progress and hope. *Lancet,* **367,** 945-947.

153. Sreevatsan S, Pan X, Stockbauer KE, Connel ND, Kreiswirth BN, Whittham TS & Musser JM. (1997). Restricted structural gene polymorphism in the *M.tuberculosis* complex indicates evolutionarily recent global dissemination. *Proceedings of National Academy of Sciences,* **94,** 9869-9874.

154. Sritharan M & Ratledge C. (1989). Co-ordinated expression of the components of iron transport (mycobactin, expchelin and envelope proteins) in *M. neoaurum. FEMS Microbiology Letters,* **51,** 183-185.

155. Sritharan M & Ratledge C. (1990). Iron regulated envelope proteins of mycobacteria grown *in vitro* and their occurrence in *M. avium* and *M. leprae* grown *in vivo. Biometals,* **2,** 203-208.

156. Sritharan M, Yeruva VC, Sivasailappan SC & Duggirala S. (2006). Iron enhances the susceptibility of pathogenic mycobacteria to isoniazid, an antitubercular drug. *World Journal of Microbiology and Biotechnlogy,* **22,** 1357-1364.

157. Sritharan M. (2000). Iron as a candidate in virulence and pathogenesis in mycobacteria and other microorganisms. *World Journal of Microbiology and Biotechnology,* **16,** 769-780.

158. Stephenson MC & Ratledge C. (1980). Specificity of exochelins for iron transport in three species of mycobacteria. *Journal of General Microbiology,* **116,** 521-523.

159. Tateson JE. (1970). Early steps in the biosynthesis of mycobactins P and S. *Biochemistry Journal,* **118,** 785-792.

160. Textor S, Wendisch VF, De Graaf AA, Muller U, Linder MI, Linder D & Buckel W. (1997). Propionate oxidation in *E. coli*: evidence for operation of a methylcitrate cycle in bacteria. *Archives in Microbiology*, **168,** 428–436.

161. Thomas GH. (2001). Metabolomics breaks the silence. *Trends in Microbiology*, **9,** 158-163.

162. Tian J, Bryk R, Shi S, Erdjument-Bromage H, Tempst P & Nathan C. (2005). *Mycobacterium tuberculosis* appears to lack α–ketoglutarate dehydrogenase and encodes pyruvate dehydrogenase in widely separated genes. *Molecular Microbiology*, **57,** 859-868.

163. Timm J, Post FA, Bekker LG, Walther GB, Wainwright HC, Manganelli R, Chan WT, Tsenova L, Gold B, Smith I, Kaplan G & McKinney JD. (2003). Differential expression of iron-, carbon-, and oxygen-responsive mycobacterial genes in the lungs of chronically infected mice and tuberculosis patients. *Proceedings of the National Academy of Sciences*, **100,** 14321-14326.

164. Tortoli E, Simonetti MT, Lacchini C, Penati V & Urbano P. (1994). Tentative evidence of AIDS-associated biotype of *M.kansasii*. *Journal of Clinical Microbiology*, **32,** 1779-1782.

165. Trias J & Benz R. (1992). Porins in the cell wall of Mycobacteria. *Science*, **258**, 1479-1481.

166. Ulrichs T & Kaufmann SHE. (2006). New insights into the function of granulomas in human tuberculosis. *Journal of Pathology*, **208,** 261-269.

167. van Soolingen D, deHass PEW, Hermans PWM & van Embden JDA. (1994). DNA fingerprinting of *M. tuberculosis. Methods in Enzymology*, **235,** 197-205.

168. Vergne AF, Walz AJ &Miller MJ. (2000). Iron chelators from mycobacteria (1954-1999) and potential therapeutic applications. *Nature Products Reviews*, **17,** 99-116.

169. Vordermeier HM, Chambers MA, Buddle BM, Pollock JM & Hewinson RG. (2006). Progress in the development of vaccines and diagnostic reagents to control tuberculosis in cattle. *The Veterinary Journal*, **171,** 229-244.

170. Vordermeier HM, Cockle PC, Whelan A, Rhodes S, Palmer N, Bakker D & Hewinson RG. (1999). Development of diagnostic reagents to differentiate between *M. bovis* BCG vaccination and *M. bovis* infection in cattle. *Clinical Diagnostics and Laboratory Immunology*, **6,** 675-682.

171. Wagner D, Maser J, Lai B, Cai Z, Barry CE III, Honer zu Bentrup K, Russell DG & Bermudez L. (2005). Elemental analysis of *M. avium, M. tuberculosis* and *M. smegmatis* containing phagosomes indicates pathogen induced microenvironments within the host cell's endosomal system. *Journal of Immunology,* **174,** 1491-1500.

172. Wagner D, Maser J, Moric I, Boechat N, Vogt S, Gicquel B, Lai B, Reyrat JM & Bermudez L. (2005). Changes of the phagosomal elemental concentrations by *M. tuberculosis* Mramp. *Microbiology,* **151,** 323-332.

173. Wasinger VC, Cordwell SJ, Cerpa-Poljak A, Yan JX, Gooley AA, Wilkins MR, Duncan MW, Harris R, Williams KL & Humphery-Smith I. (1995). Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium. Electrophoresis,* **16,** 1090-1094.

174. Wayne LG & Kubica GP. (1986). Mycobacteria. *in* Bergey's Manual of Systematic Bacteriology. Vol. 2. Sneath, P.A.H., Mair, N.S., Sharpe, M.E. and Holt, J.G. (eds): The Williams & Wilkins Co., Baltimore, Md, pp. 1436-1457.

175. Wheeler P & Ratledge C. (1994). Metabolism of *M. tuberculosis*. In *Tuberculosis: Pathogeneisis, Protection and Control,* pp. 353-385. Edited By Bloom, B.R. Wahsington, DC: American Society of Microbiology.

176. WHO. (2005). Global tuberculosis control: Surveillance, planning, financing. (World Health Organisation, Geneva).

177. WHO**.** (2006). Global tuberculosis control: Surveillance, planning, financing. Geneva: World Health Organisation, **242.**

178. WHO. The world health report 2004: changing history. Geneva: World Health Organisation, 2004.

179. Wiechelman KJ, Braun RD & Fitzpatrick JD. (1988)**.** Investigation of the bicinchoninic acid protein assay: Identification of the groups responsible for color formation. *Analytical Biochemistry,* **175,** 231-237.

180. Wong DK, Gobin J, Horwitz MA & Gibson BW. (1996)**.** Characterisation of exochelins of *M. avium* : evidence for saturated and unsaturated and for acid and ester forms. *Journal of Bacteriology,* **178,** 6394-6398.

181. Wong DK, Lee BY, Horwitz MA & Gibson BW. (1999). Identification of Fur, aconitase and other proteins expressed by *M. tuberculosis* under conditions of low and high concentrations of iron by combined two dimensional gel electrophoresis and mass spectrometry. *Infection and Immunity,* **67,** 327-336.

182. World Health Organisation, WHO. (1993). Tuberculosis on the increase. *World Health Forum,* **14,** 1-94.

183. Yeruva VC, Duggirala S, Lakshmi V, Kolarich D, Altmann F & Sritharan M. (2006). Identification and characterisation of a major cell wall associated iron regulated envelope protein (Irep-28) in *M. tuberculosis. Clinical and Vaccine Immunology,* **13,** 1137-1142.

184. Yeruva VC, Sivagami sundaram CAS & Sritharan M. (2005). Influence of iron concentration on the expression and activity of catalases-peroxidases in mycobacteria. *Indian Journal of Biochemistry and Biophysics,* **42,** 28-33.

185. Yu S, Fiss E & Jacobs WR Jr. (1998). Analysis of exochelin locus in *M. smegmatis:* biosynthesis genes have homology with genes of the peptide synthetic family. *Journal of Bacteriology,* **180,** 4676-4685.

186. Zahrt TC & Deretic V. (2000). An essential two-component signal transduction system in *M. tuberculosis. Journal of Bacteriology,* **182,** 3832-3838.

187. Zhu W, Arceneaux JEL, Beggs ML, Byers BR, Eisenach KD & Lundrigan MD. **(**1998). Exochelin genes from *M. smegmatis*: identification of an ABC transporter and two nonribosomal peptide synthetase genes. *Molecular microbiology,* **29,** 629-639.