Introduction
Microorganisms produce a wide variety of biologically active metabolites representing a vast diversity of fascinating molecular architecture. The large number of metabolites found in living organisms and the calculated, unexpected small number of genes identified during genome sequencing projects confound biologists. Presently, it is not apparent how nature could generate vast metabolic repertoire by using limited number of genes. The cell envelope of *Mycobacterium tuberculosis* is a fantastic example of unusual compounds constituting 40% of the dry weight of the bacteria (Brennan and Nikaido, 1995). While structures of these molecules have been defined in exquisite chemical detail for years, the molecular mechanism of biosynthesis has been unclear (Daffe, 1991; Daffe and Laneelle, 1988). The genome sequencing projects of mycobacteria have revealed an array of genes involved in lipid metabolism, including several genes homologous to polyketide synthases (PKSs) (Cole et al., 1998). The availability of the genetic blueprint of this organism provides a tantalizing opportunity to dissect biosynthetic pathways of complex lipids. This thesis describes the molecular logic underlying the biosynthesis of a virulent lipid phthiocerol dimycocerosate (PDIM), which involves approximately 35 catalytic steps. These cell-free reconstitution studies demonstrate that PKSs, which are usually involved in the biosynthesis of secondary metabolites, are responsible for generating complex lipids in *M. tuberculosis*.

Mycobacterial genus comprises of greater than 70 species and includes the human pathogens *M. tuberculosis* and *M. leprae*. *M. tuberculosis* is a gram positive bacillus and belongs to the *Actinomycetes* family. The cell wall of mycobacteria, which contains a variety of unusual lipids, has been proposed to
play an important role in the process of survival of the bacteria in hostile environments (Glickman and Jacobs, 2001; Karakousis et al., 2004). The outermost lipid barrier contains a variety of polysaccharides and species-specific acyl trehaloses, sulpholipids, glycopeptidolipids and phenolic glycolipids, which is followed by characteristic long-chain mycolic acids anchored to the cell wall core (Brennan and Nikaido, 1995; Daffe and Draper, 1998). The genome sequencing of various mycobacterial species revealed that these bacilli have the potential to synthesize a variety of metabolites apart from those required to meet the basic cellular needs (Cole et al., 1998; Garnier et al., 2003). The bioinformatics analyses of the genomes have facilitated prediction of the biochemical functions of genes by comparison with their homologues from other organisms, apart from providing insights into the variation in genome organization across species. Sequence information and the recent development of genetic tools have provided an impetus to understand functional significance of genes in various biosynthetic systems and their involvement in bacterial pathogenesis.

The mycobacterial genome suggested the presence of a large number of fatty acid and polyketide biosynthetic systems, including enzymes usually found in higher eukaryotes (Cole et al., 1998). Fatty acids and polyketides are biosynthesized by structurally and mechanistically similar enzymes called fatty acid synthases (FASs) and PKSs respectively (O'Hagan, 1993). The PKSs have been typically characterised from *Streptomyces* and are involved in the synthesis of secondary metabolites (Hopwood, 1997; Khosla, 2000). Although *M. tuberculosis* has not been shown to produce any polyketide metabolites, *M. ulcerans* (an extracellular mycobacterial species) has been reported to synthesize polyketide
toxins called mycolactones that are required for its pathogenesis (George et al., 1999; Stinear et al., 2004). Gene inactivation studies with PKSs in *M. tuberculosis* have revealed that some of these genes might be involved in the synthesis of the complex cell wall lipids, although the precise mechanism has not been elucidated (Constant et al., 2002; Rousseau et al., 2003a; Rousseau et al., 2003b; Sirakova et al., 2003b; Sirakova et al., 2001). Several PKS and FAS genes have been implicated in the synthesis of phthiocerol dimycocerosate (PDIM), a lipid present only in the virulent strain of mycobacteria (Azad et al., 1997; Camacho et al., 2001; Cox et al., 1999; Kolattukudy et al., 1997; Rousseau et al., 2003b; Sirakova et al., 2003b).

In this thesis, we have characterised the biochemical functions of several genes present on ~70 kb region on the mycobacterial genome that are involved in the synthesis of the lipid core of PDIM. This cluster includes five PKS-like genes (*ppsA-E*), two genes homologous to acyl-CoA synthetases (*fadD26* and *fadD28*), one iterative FAS gene (*mas*), a thioesterase and a *papA5* (PKS associated protein) gene of unknown function. There are several genes (*drrABC* and *mmpL7*) homologous to genes encoding transporter proteins and glycosyl and methyl transferase proteins.

Chapter one reviews the current understanding about the three multienzymatic systems, PKSs, FASs and non-ribosomal peptide synthetases (NRPSs). The relevance of these proteins along with other lipid metabolizing enzymes in mycobacterial pathogenesis has also been reviewed. The molecules that constitute the cell wall and their probable biosynthetic mechanisms are discussed.
Based on the retro-biosynthetic approach, we have deconvulated the PDIM synthesis in four steps: a) Priming of long-chain fatty acid and synthesis of diol-component of phthiocerol b) Phthiocerol synthesis by PpsE protein c) Enzymatic synthesis of mycocerosic acids and d) Transesterification of mycocerosic acids on to the diol of phthiocerol and the results are described in the next four chapters.

The biosynthesis of phthiocerol is described in Chapter two. In this chapter, three large multifunctional PKS proteins have been reconstituted and this study provides the first direct evidence of the involvement of PKS proteins in the biosynthesis of lipids. Our studies show that the PpsA and PpsB proteins catalyse the synthesis of the diol moiety of phthiocerol and the PpsE protein catalyses the final step of phthiocerol synthesis. Metabolites biosynthesized by each protein have been characterised by mass spectrometry and this cell-free reconstitution study provides insights in the generation of the heterogeneity in the family of phthiocerol esters in *M. tuberculosis*.

The multimethylated fatty acids are a characteristic feature of the mycobacteria and the tubercle contains dedicated proteins involved in the synthesis of methyl-branched fatty acids. Chapter three describes the *in vitro* synthesis and characterisation of the mycocerosic acids synthesized by the Mas protein. The metabolites during the biosynthesis are covalently bound to the proteins as acyl-thioesters. Interestingly, the multienzymatic proteins coded by the genes of the *pps* cluster do not contain a typical thioesterase domain fused to the C-terminal end. The biochemical role of the *tesA* gene which codes for a thioesterase-like protein was investigated. Our studies suggest that this type II thioesterase may be
involved in hydrolysing the covalently bound aberrant intermediates from the proteins.

The esterification of mycocerosic acids on to the diol of phthiocerol is required for completing the synthesis of PDIM. Chapter four characterises an ester synthase which is involved in the transesterification of mycocerosic acids on to the diol moiety of the phthiocerol. Our study illustrates that PapA5 directly transfers the ACP-bound mycocerosic acids on to various long-chain alcohols.

The biochemical function of several FadD proteins is discussed in Chapter five. *M. tuberculosis* genome sequence revealed the presence of 34 copies of the *fadD* genes homologous to the acyl-CoA synthetases. Our computational, biochemical and structural studies facilitate classification of these 34 FadD proteins into two distinct families. We provide evidence of the presence of a new class of long-chain fatty acyl-AMP ligases (FAAL) which activate and transfer long-chain fatty acids onto PKSs leading to synthesis of complex hybrid metabolites.

The studies described in this thesis dissect the assembly of a complex mycobacterial lipid. Apart from providing insights into the mechanisms that mediate the crosstalk between PKS and FAS systems, this study also provides clues about how different biochemical functions are integrated within a given organism to generate metabolite diversity.