Introduction
Introduction

Genome sequencing projects that were initiated more than a decade back are now starting to provide an understanding of the biology of various organisms. Since the publication of the *Mycobacterium tuberculosis* H37Rv (Mtb) genome in 1998, several novel gene clusters, unusual protein families, and unique metabolic pathways have been identified (Cole et al., 1998). A remarkable feature of the mycobacterial genome is the presence of an unusually large number of lipid metabolizing enzymes. Notably, polyketide synthases (PKSs), which have been classically characterized to mediate the biosynthesis of secondary metabolites in various organisms (Katz and Donadio, 1993; O'Hagan, 1992; O'Hagan, 1993), are also present in large numbers. Since there are no secondary metabolites known from Mtb, the functional relevance of these PKSs was not immediately perceivable upon sequencing. Genetic studies over the last decade have revealed a role of mycobacterial PKSs in the biosynthesis of the acyl component of various mycobacterial lipids (Kolattukudy et al., 1997; Chopra and Gokhale, 2009; Gokhale et al., 2007; Jackson et al., 2007). The PKS catalyzed acyl chains are an important constituent of complex lipids like sulfolipids, polyacyltrehaloses, phthiocerol dimycocerosates, mycolic acids, and mannosyl-β-1-phosphomycoketides (MPMs), which have been implicated in mycobacterial pathogenicity. Many of these acyl chains are branched and thus provide an intercalated network of hydrophobic lipids in the cell envelope. The unusual chemical nature of these acyl chains clearly highlights that mycobacterial PKS enzymes have devised novel mechanisms to synthesize these metabolites. In this thesis, we have investigated into the biosynthetic potential of several large multifunctional PKS enzymes from Mtb, with specific emphasis on PKS12 protein,
encoded by the largest open reading frame in the Mtb genome. Our studies with the PKS12 protein have revealed a unique mechanism of polyketide biosynthesis that involves formation of a large supramolecular assembly to catalyze the biosynthesis of the acyl chain of MPMs.

In chapter one, we review the present biochemical understanding of PKS proteins. Further, we provide a perspective on their structural and functional relationship with fatty acid synthases (FASs). We have also described various modes of biosynthesis and provided insights into their multifunctional catalysis, based on the recent elucidation of their three-dimensional structures. The carrier domain, which is an integral part of the PKS enzymes, is modified post-translationally by phosphopantetheine arm and we have discussed its key role in coordinating catalysis, which is believed to involve large conformational changes. Finally, comprehensive details of each mycobacterial PKS and their role in the biosynthesis of lipids present in the cell envelope are described.

In chapter two of this thesis, we have attempted to deconvolute the MPM biosynthesis in five steps, based on a retro-biosynthetic approach: a) Biosynthesis of the mycoketide chain by PKS12 b) Release of the mycoketide chain c) Reduction of the mycoketide chain to an alcohol d) Phosphorylation of the mycoketide alcohol and e) Mannosylation of phosphomycoketide to form MPMs. Based on our analysis, Rv2047c, PKS12 and Ppm1 are the key enzymes that could be involved in this biosynthetic pathway. Since the role of Ppm1 in transfer of mannose sugar is well established, we have biochemically investigated the functional role of Rv2047c and PKS12 in MPM biosynthesis.
After establishing the *in vitro* activity of PKS12, and realizing the unusual bimodular iterative nature of polyketide catalysis, we have performed comparative analysis of PKS12 with three similar PKS proteins from Mtb. These studies are described in chapter three of this thesis. We demonstrate that PKS2 and MAS belong to the iterative class of PKSs and catalyze multiple rounds of condensation of methylmalonate extender units with the starter chain. PKS7 protein on the other hand shows a modular behaviour resulting in formation of a unique product. Interestingly, *in silico* analysis of PKS12 suggests a modular architecture, analogous to PKS7 protein. Biochemical studies, in contrast, suggest an iterative mode of catalysis.

In chapter four, the novel mechanism of bimodular iterative catalysis performed by the PKS12 protein is deciphered. We have specifically probed the role of the *N*-terminus and *C*-terminus linker residues in dictating PKS12 catalysis. Several experiments have led to a model, wherein the *C*-terminus linker residues of PKS12 can specifically interact with the *N*-terminus linker sequence of another PKS12 protein molecule to form an oligomeric organization. We refer to this mode of catalysis as the "modularly-iterative" mechanism of polyketide condensation.

In conclusion, the studies described in this thesis converge towards a novel paradigm in biosynthesis of complex metabolites by PKS proteins. Our studies expand the existing versatility of PKS catalysis and also demonstrate remarkable features of metabolite biosynthesis in mycobacteria.