Molecular Cell, Vol. 17, 631–643, March 4, 2005, Copyright ©2005 by Elsevier Inc. DOI 10.1016/j.molcel.2005.02.009

# Dissecting the Mechanism and Assembly of a Complex Virulence Mycobacterial Lipid

**Omita A. Trivedi, Pooja Arora, Archana Vats, Mohd. Zeeshan Ansari, Rashmi Tickoo, Vijayalakshmi Sridharan, Debasisa Mohanty, and Rajesh S. Gokhale***
National Institute of Immunology
Aruna Asaf Ali Marg
New Delhi 110 067
India

## Summary

***Mycobacterium tuberculosis*** **cell envelope is a treasure house of biologically active lipids of fascinating molecular architecture. Although genetic studies have alluded to an array of genes in biosynthesis of complex lipids, their mechanistic, structural, and biochemical principles have not been investigated. Here, we have dissected the molecular logic underlying the biosynthesis of a virulence lipid phthiocerol dimycocerosate (PDIM). Cell-free reconstitution studies demonstrate that polyketide synthases, which are usually involved in the biosynthesis of secondary metabolites, are responsible for generating complex lipids in mycobacteria. We show that PapA5 protein directly transfers the protein bound mycocerosic acid analogs on phthiocerol to catalyze the final esterification step. Based on precise identification of biological functions of proteins from Pps cluster, we have rationally produced a nonmethylated variant of mycocerosate esters. Apart from elucidating mechanisms that generate chemical heterogeneity with PDIMs, this study also presents an attractive approach to explore host-pathogen interactions by altering mycobacterial surface coat.**

## Introduction

*Mycobacterium tuberculosis* is the etiologic agent of tuberculosis (TB) in humans and is responsible for more morbidity than any other bacterial disease. The pathogenic mycobacteria possess a complex cell envelope, which is considered one of the major determinants of virulence. This impermeable barrier imparts resistance against hostile environments as well as against therapeutic agents and also plays an active role in modulating the host-immune response (Glickman and Jacobs, 2001; Karakousis et al., 2004; Stewart et al., 2003). An interesting feature of the mycobacterial envelope is the extraordinary high lipid content consisting of 40% of their dry weight. The long-chain $\alpha$-alkyl, $\beta$-hydroxy mycolic acids are one of the major constituents that are associated with diverse glycolipids representing the outermost envelope of mycobacteria (Barry, 2001; Brennan and Nikaido, 1995). The chemical structures of many of these molecules have been defined in exquisite detail (Daffe, 1991; Daffe and Laneelle, 1988). However, the molecular components of biosynthetic machinery involved in the synthesis of these complex lipids have not been elucidated.

The tremendous efforts in recent years have demonstrated that a number of mycobacterial lipids are biosynthesized by the combined action of fatty acid synthases (FASs) and polyketide synthases (PKSs) (Camacho et al., 2001; Cox et al., 1999; Kolattukudy et al., 1997; Sirakova et al., 2001). Whereas FASs are known to catalyze the biosynthesis of fatty acids, PKSs usually produce secondary metabolites that are a rich source of commercially important therapeutic agents (Hopwood, 1997; Khosla, 2000). Traditionally, these multienzymatic assemblies have been studied independently, and their modes of interaction to produce hybrid molecules of fatty acids and polyketides have not been investigated (Gokhale and Tuteja, 2001; Kolattukudy et al., 1997; Minnikin et al., 2002). In mycobacteria, several PKS disruption mutants display altered lipid profiles, some of which also exhibit attenuation in their virulence properties (Sirakova et al., 2001; Dubey et al., 2002).

Genetic studies have identified an array of PKS genes that are required for biosynthesis of PDIMs and glycosylated phenolphthiocerol esters, which are surface-exposed lipids unique to the virulent strains of mycobacteria. The *pks7*, *pks10*, *pks12*, *ppsA-E*, and *pks15/1* mutant strains of *M. tuberculosis* are deficient in phthiocerol derivatives and yield strains with attenuated growth in the murine model (Kolattukudy et al., 1997; Rousseau et al., 2003; Sirakova et al., 2003a, 2003b). Most of the genes proposed to be involved in PDIM biosynthesis can be classified in two large clusters. The *pps* locus consists of three transcriptional units with 15 open reading frames and encompasses ~50 kbp of the genome. The enzymic activity for three of these proteins from *pps* cluster has been elucidated. We recently characterized FadD26 protein as fatty acyl-AMP ligase involved in the activation and transfer of long-chain fatty acids (Trivedi et al., 2004). The function of the PapA5 protein as acyltransferase, based on its ability to synthesize hydroxyl-ester products with surrogate substrates, was also very recently described (Onwueme et al., 2004). The *Mas* gene product from *M. bovis* BCG was shown to catalyze synthesis of mycocerosic acids (Azad et al., 1996). The *pks15/1* is proposed to be involved in the synthesis of phenolphthiocerol (Constant et al., 2002). The other gene cluster contains six PKS genes, *pks10*, *pks7*, *pks8*, *pks17*, *pks9*, and *pks11*, that have been newly implicated in PDIM biosynthesis (Sirakova et al., 2003a). This is a rather unusual cluster flanked on either side by type III PKS genes. Although phenotypes produced by these manipulations provide clues to the probable pathway for PDIM synthesis, the assembly of these metabolites has only been inferred from indirect observations. In order to unambiguously delineate the enzymology and assembly of these complex virulent lipids, it is imperative to isolate proteins with demonstrable functions.

Here, we report the precise roles of several proteins in the biosynthesis of PDIM by using a cell-free system.

*Correspondence: rsg@nii.res.in

Based on the retro-biosynthetic approach, we have deconvulated the PDIM synthesis in four steps: (1) priming of long-chain fatty acid and synthesis of diol-component of phthiocerol, (2) phthiocerol synthesis by PpsE protein, (3) enzymatic synthesis of mycocerosic acids, and (4) transesterification of mycocerosic acids onto the diol of phthiocerol. Our studies have identified some of the mechanisms that generate chemical heterogeneity associated both with phthiocerol and mycocerosic acids. We also demonstrate that the PapA5 protein directly transfers the protein bound mycocerosic acids onto the phthiocerol to catalyze the final esterification step. The protein interfaces that are likely to play an important role in the transfer of mycocerosic acids have been investigated by a combination of docking simulations, structural analysis, and site-directed mutagenesis. Based on our in vitro studies, mycocerosic acid synthase (Mas) protein was engineered to synthesize nonmethylated fatty acids that could be transferred to phthiocerol to produce new analogs. Because the properties of branched-chain fatty acids are different from linear-chain fatty acids, this methodology suggests an attractive approach to chemically alter the mycobacterial cell envelope. This report of the biochemical reconstitution of the essential component of the cell envelope of virulent mycobacteria characterizes several potential antimycobacterial drug targets and provides insights into the mechanisms by which biochemical functions are integrated in an organism.

## Results and Discussion

### Computational Analysis and Organization of Proteins from the *pps* Cluster

Because the organization of genes does not necessarily represent the order in which proteins would assemble the metabolite, a detailed computational analysis was carried out to provide guidelines for metabolic reconstitution experiments. The pathway along with the organization of the gene cluster for the biosynthesis of phthiocerol and phenolphthiocerol was proposed earlier (Azad et al., 1997; Kolattukudy et al., 1997; Minnikin et al., 2002), although several carrier domains were not detected in these studies. The analysis of the domain organizations of multifunctional PKS and FAS proteins investigated by using in house-developed NRPS-PKS tool (Ansari et al., 2004) suggested that all five PpsA–E proteins contained acyl carrier protein (ACP) domains characteristic of modular PKSs. Recently, the macrotetrolide PKS biosynthetic gene cluster unveiled an unprecedented organization that does not contain an ACP domain (Kwon et al., 2001; Shen, 2003). The identification and characterization of ACP and PCP domains are essential for understanding the biosynthetic pathway of complex lipids, because these carrier domains facilitate hand-to-hand transfer of intermediates across various PKS systems. The biosynthetic pathway involved in PDIM synthesis along with the proposed organization of the domains for the *pps* cluster is shown in Figure 1. The presence of N-terminal carrier domain in PpsA is characteristic of loading modules in PKSs and NRPSs and is usually involved in initiating synthesis (Admiraal et al., 2001; Cane and Walsh, 1999). The PpsE protein revealed a condensation-like domain at the carboxy-terminal end, as suggested in the earlier study (Kennedy et al., 1999). Apart from catalyzing peptide bond formation in the NRPS, condensation (C) domains when present at the carboxy-terminal end of NRPSs are involved in the release of mature peptide chains (Keating et al., 2001). Intriguingly, the PapA5 protein present adjacent to PpsE in the cluster also showed weak homology to the C domains. During the preparation of this manuscript, the PapA5 three-dimensional (3D) structure was solved (Buglino et al., 2004), and it indeed revealed structural homology to the NRPS C domains (Keating et al., 2002).

In order to directly probe the function of each protein leading to the assembly of PDIMs, we functionally expressed several proteins in a heterologous host. For technical convenience and for minimizing errors in amplifying large genes, we constructed overlapping shotgun libraries of mycobacterial BAC genomic DNA in plasmids. By suitable engineering of restriction sites at ends, these genes were cloned and expressed in *E. coli*. All proteins were expressed in the soluble form by suitably modulating the growth and expression conditions. The posttranslational modification of ACP domains of PpsA, PpsB, PpsE, and Mas proteins was achieved by coexpressing surfactin phosphopantetheinyl (P-pant) transferase protein (Sfp) in *E. coli* (Gokhale et al., 1999; Lambalot et al., 1996).

### Biochemical Investigation of *pps* Cluster: Initiation of Phthiocerol Synthesis

Because PpsA is a multifunctional protein, the catalytic integrity of individual domains was probed by using radiolabeled substrates (Figure 2A). Because the turnover rates of acyl transferase (AT) domain in general are very high, these assays were also performed at low temperatures, which showed a strict specificity for malonyl-CoA (MCoA). The posttranslational modification of the ACP domains of the PpsA protein was confirmed by hydrolyzing the transferred extender (malonyl) unit by using hydroxylamine (Figure 2B), which specifically cleaves thioester linkages. A significant loss of the radioactivity associated with the protein indicated that the protein was posttranslationally modified and was present in the holo form.

The first step in phthiocerol synthesis was investigated by using FadD26 protein to transfer lauric acid onto the purified PpsA protein (Trivedi et al., 2004). In order to extend the lauroyl chain by two carbon, the acylated PpsA protein was incubated with $^{14}$C-MCoA and NADPH to form corresponding reduced acyl-diketide product. Because PpsA protein does not contain a thioesterase domain, the enzyme bound products were released by alkali hydrolysis and analyzed by radio-TLC (Figure 2C). The product was extracted from the TLC plate and was chemically characterized to be 3-hydroxy tetradecanoic acid by tandem mass spectrometry (Figure 2D). The molecular ion peak of $[M - H]^{-1}$ was observed at 243.19, which is consistent with the molecular formula $C_{14}H_{28}O_3$ corresponding to the reduced acyl-diketide. The dehydration peak observed during MS/MS spectrum is also consistent with the presence of hydroxyl group. The product was further confirmed

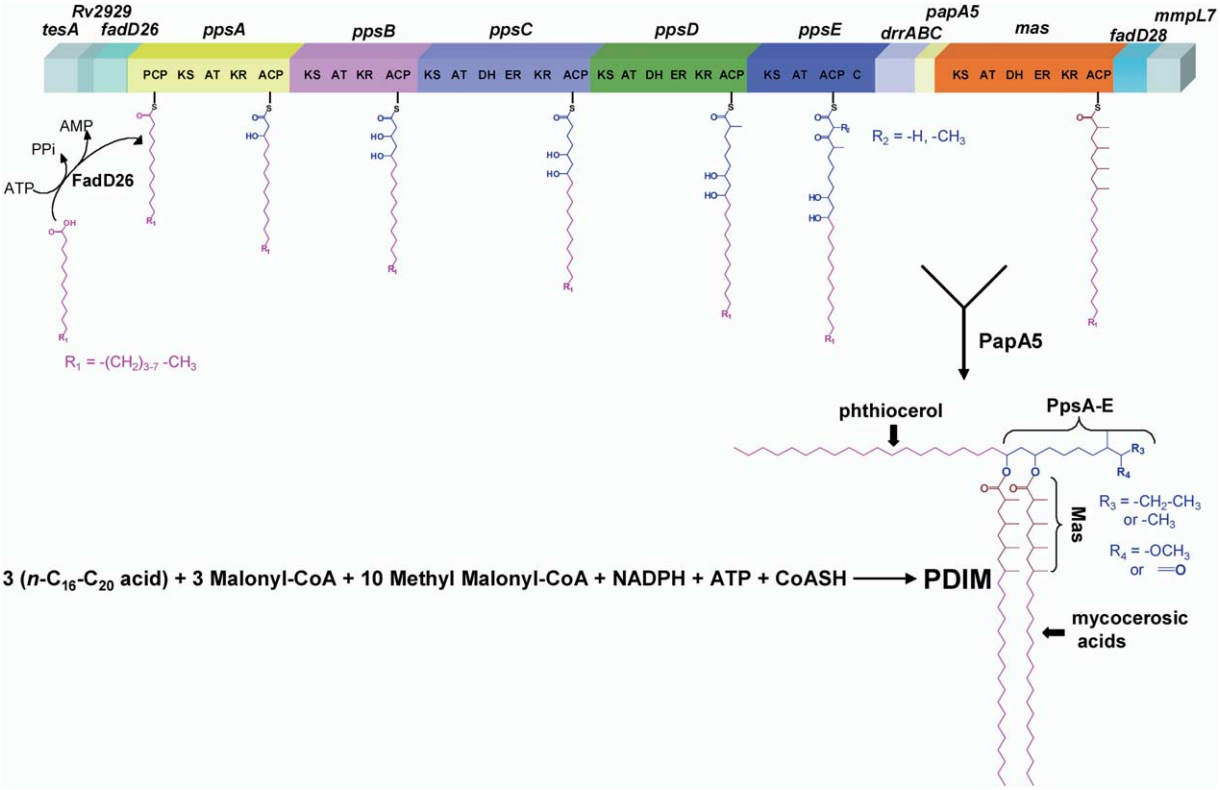


Figure 1. Biosynthesis of Diesters of Phthiocerol and Related Compounds

Schematic representation of the organization of ~50 kbp *pps* locus implicated in the virulent cell wall lipid PDIM biosynthesis. This cluster consists of 15 open reading frames constituting of at least 35 catalytic steps. The intermediates synthesized during the assembly of PDIM are covalently bound as acyl-thioesters on to the acyl carrier protein (ACP) of the respective multifunctional proteins. FadD26 activates long-chain fatty acids and transfers them to the N-terminal PCP domain of PpsA. All the six multifunctional proteins contain β-ketoacyl ACP synthase (KS), acyl transferase (AT), ketoreductase (KR) and ACP domains involved in the loading, decarboxylative condensation, and ketoreduction of the acyl intermediates. The presence of the dehydratase (DH) and enoylreductase (ER) domain converts these β-hydroxy acyl-thioester to corresponding reduced products. PapA5 mediates the direct transfer of the mycocerosic acids from the Mas protein on to the phthiocerol. The intermediates are passed on from one protein to the next, like a baton in a relay race.

by using radiolabeled extender unit, which would shift the mass by 2 Da. The kinetics of condensation of two-carbon unit and the reduction of ketone to hydroxyl proceeded with an apparent $k_{cat}$ of 0.103 pmol $mg^{-1}min^{-1}$. These studies unambiguously demonstrate that PpsA protein is involved in extension of long-chain fatty acids to initiate phthiocerol biosynthesis. The phthiocerol moiety in PDIMs is known to show heterogeneity in terms of carbon chain length. Investigation of the substrate specificity of FadD26 protein showed that this protein was able to activate fatty acids ranging from $C_{12}$ to $C_{16}$ (Figure 2E). These activated fatty acids would in turn be transferred onto the PCP domain of the PpsA protein, resulting in diverse chain lengths of phthiocerol.

### Formation of Diol Component of Phthiocerol

According to the assembly-line mechanism for phthiocerol synthesis, the PpsA protein would transfer the acyl chain from the ACP domain of the PpsA protein onto the β-ketoacyl ACP synthase (KS) domain of the PpsB protein, where the next 2-carbon extension would take place. The active site motif of the PpsB AT domain was identical to that of PpsA protein, and its specificity for incorporation of MCoA was confirmed by using radioactive substrates (Figure 3A). The PpsA bound radiolabeled acyl diketide product was incubated with purified PpsB protein along with MCoA extender units. The covalently acylated products were hydrolyzed from the protein, extracted, and analyzed on TLC (Figure 3C). In lane 2, which corresponds to reaction containing both PpsA and PpsB proteins, two closely migrating bands are observed in addition to the PpsA product. The PpsB would catalyze formation of 3,5-dihydroxy hexadecanoic acid, which on release would spontaneously cyclize to form the lauroyl-triketide lactone (Figure 3B). The formation of lauroyl triketide lactone was confirmed by ESI-MS analysis (Figure 3D). The molecular ion $[M - H]^{-1}$ peak at 269.23 corresponded to the expected mass. The MS/MS of the parent ion of 269.23 at different collision energies resulted in fragment ions of 251.23 and 233.23, corresponding to the loss of one or two water molecules, respectively. Previous mass spectrometric analysis with triketide lactones also showed similar fragmentation pattern (Kearney et al., 1999; Weissman et al., 1999). The closely related higher band on TLC corresponded to the unreduced ketolactone product, where the keto group at position three

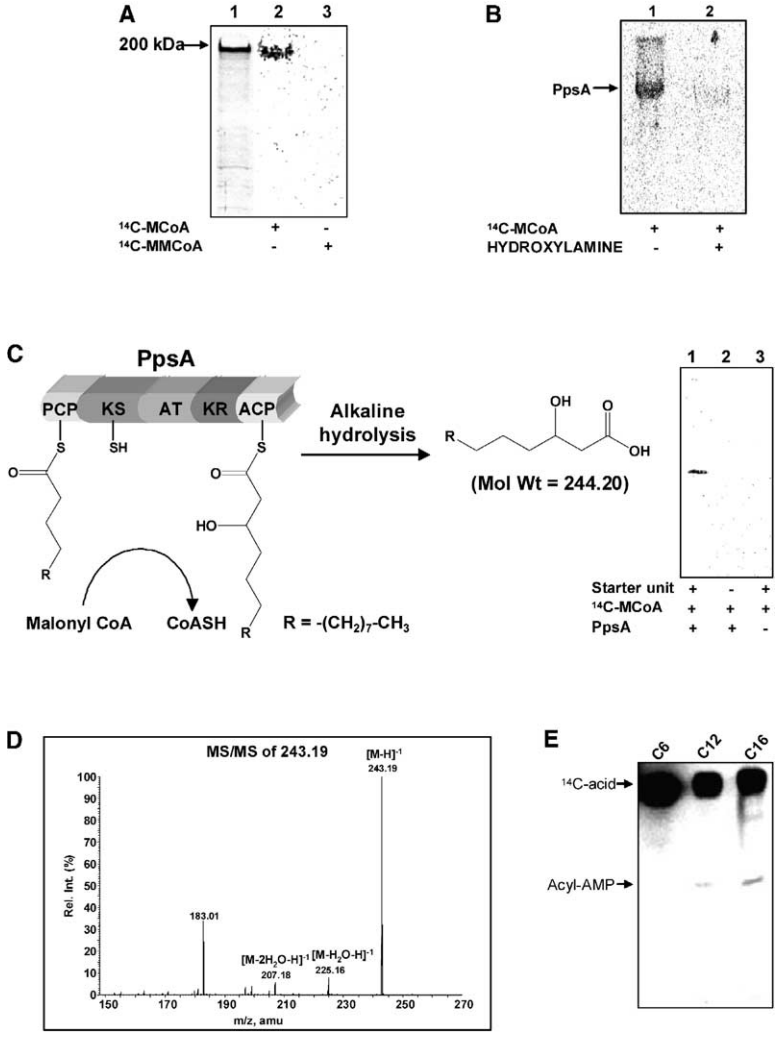


Figure 2. Initiation of Phthiocerol Synthesis by PpsA

(A) Labeling of PpsA protein with radioactive extender units. Lane 1 shows the Coomassie-stained, purified PpsA protein (200 kDa). The radiolabel is associated with the protein with $^{14}$C-MCoA (lane 2) and not with $^{14}$C-MMCoA (lane 3).

(B) Phosphopantetheinylation of the ACP domain of PpsA protein. The radiolabel associated with the protein (lane 1) is significantly decreased on treatment with hydroxylamine (lane 2).

(C) The first step of phthiocerol synthesis is catalyzed by PpsA protein. This enzymatic reaction involves extension of lauroyl chain by decarboxylative condensation of a MCoA. The acyl-diketide product formed was analyzed by radio-TLC (lane 1). No product is formed in the absence of the starter unit or the protein (lanes 2 and 3).

(D) ESI-MS/MS analysis of the PpsA product. The $[M - H]^{-1}$ molecular ion peak obtained at 243.19 corresponds to the 3-hydroxy tetradecanoic acid. Fragments corresponding to two successive dehydrations are marked.

(E) Activation of various fatty acids by FadD26. The FadD26 activates medium- to long-chain fatty acids to acyl-adenylates. The fatty acid migrates with an $R_f$ of 0.9 and the fatty acyl-AMP migrates with an $R_f$ of 0.5.

is incompletely reduced by PpsB ketoreductase (KR) domain. The absence of the dehydratase (DH) and the enoylreductase (ER) domains in PpsA and PpsB modules thus results in the formation of the diol portion of the phthiocerol moiety. The AT domains of PpsC and PpsD proteins are predicted to be specific for MCoA and methyl MCoA (MMCoA), respectively, and would be expected to convert the acyl-triketide to the pentaketide product.

## PpsE Catalyzes the Last Step of Phthiocerol Synthesis

The final step in phthiocerol biosynthesis was investigated by reconstituting PpsE protein. A surrogate synthetic substrate N-acetyl cysteamine (NAC) thioester of 9-hydroxy decanoic acid was chemically synthesized. This substrate mimics the ketide chain that would be transferred from ACP of PpsD protein onto the KS domain of PpsE. Incubation of PpsE protein with radiolabeled MCoA and MMCoA demonstrated that both these substrates could be used as extender units (Figure 4A). The enzymatic assays were performed with 9-hydroxy decanoyl-NAC as starter unit and by using both $^{14}$C-MCoA and $^{14}$C-MMCoA as extender units (Figure 4B). In the absence of any reductive domains in the PpsE protein, the keto group would not undergo any reduction and so the expected final product would have a β-keto moiety. Because the C domain present at the end of the PpsE protein did not release ketide chains efficiently under the experimental conditions, products were released by hydrolysis. The autoradiograph of the radio-TLC showed products with two distinct $R_f$ for two extender units (0.34 and 0.38 for MCoA and MMCoA, respectively) (Figure 4C). Mass spectrometric analyses in negative ion mode confirmed these products to be 11-hydroxy 3-keto dodecanoic acid (Figure 4D) and 11-hydroxy 2-methyl 3-keto dodecanoic acid (Figure 4E). Interestingly, the diesters of phthiocerol and related compounds are mixtures of long-chain β-diols, some of which also possess heterogeneity at the terminal methyl group position ($R_3$) of the phthiocerol moiety (Figure 1). The catalytic flexibility of PpsE protein to efficiently incorporate both MCoA and MMCoA readily explains the latter heterogeneity. The β-keto acid produced by PpsE protein could undergo spontaneous decarboxylation and reductive methyla-

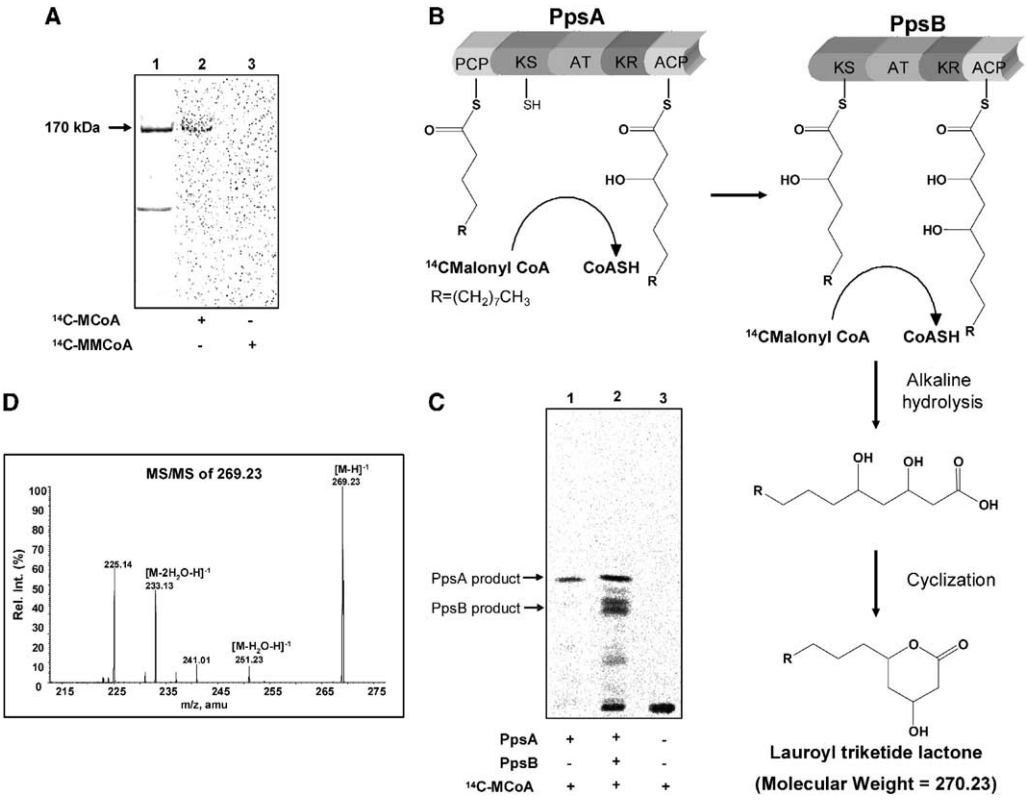


Figure 3. Formation of the Diol Moiety of Phthiocerol

(A) Specificity of AT domain. Lane 1 shows the purified PpsB protein. The PpsB protein is radiolabeled in the presence of $^{14}$C-MCoA (lane 2) and not with $^{14}$C-MMCoA (lane 3).

(B) Extension of the PpsA product by PpsB protein. The lauroyl diketide from PpsA protein is transferred to the KS domain of PpsB protein. Further two-carbon chain extension and chemical hydrolysis results in spontaneous cyclization to form triketide lactone.

(C) Radio-TLC characterization of PpsB products. Lane 1 shows the lauroyl diketide product synthesized by PpsA protein. On addition of PpsB to the reaction, a new product with lower $R_f$ value is formed (lane 2). No product is formed in the absence of both these proteins (lane 3).

(D) Characterization of lauroyl triketide product. Nanospray ESI-MS/MS spectrum of the molecular ion at 269.23 shows the collision induced fragments at 251.23 and 233.23, which correspond to successive loss of two water molecules.

tion in some phthiocerol variants. Although the precise order in which these terminal steps occur is not known, together, these studies provide molecular mechanistic insights into the programming of phthiocerol biosynthesis.

## Mas Synthesizes the Multimethylated Mycocerosic Acids

In order to reconstitute the mycocerosic acid synthesis, we cloned and overexpressed the *mas* gene from H37Rv in *E. coli*. The 232 kDa protein was obtained in the soluble form and was posttranslationally modified by coexpression with *sfp*. The Mas protein specifically incorporated MMCoA (Figure 5A). The homologous protein from *M. bovis* BCG has been earlier characterized to be an iterative fatty acid synthase that synthesizes long chain mycocerosic acids (Rainwater and Kolattukudy, 1983). The catalytic activity of Mas protein was analyzed by incubating various starter ($C_6$–$C_{20}$) acyl-CoAs, radiolabeled MMCoA, and NADPH (Figure 5B). The covalently tethered products were chemically hydrolyzed by alkali and analyzed on the TLC. Mas protein displayed promiscuous substrate specificity for the incorporation of the exogenous primers. Radio-TLC analysis showed three closely migrating radioactive bands for most of the starter units. The first conjecture was that these bands might correspond to the different cycles of iterations. However, in the same solvent system, the linear-chain fatty acids, lauric acid ($C_{12}$), myristic acid ($C_{14}$), palmitic acid ($C_{16}$), and stearic acid ($C_{18}$) could not be resolved. We therefore explored the possibility that some of these radioactive products might correspond to adducts derived from DTT. Indeed, by substituting nonthiol-based reducing agent TCEP for DTT, a single radioactive band was observed on the radio-TLC (Figure 5C). The apparent $k_{cat}$ obtained from the initial velocity measurements for $C_{12}$ starter unit was estimated to be 0.08 $min^{-1}$.

The mycocerosic acids formed by Mas protein were characterized by nanospray ESI-MS. Preliminary attempts at identification of products were complicated by the presence of background solvent contaminations interfering with the expected masses. Therefore, we decided to overcome this problem by using $^{14}$C-radiolabeled starter unit. The TOF MS spectrum of the Mas-catalyzed extension of lauroyl-CoA starter unit primarily

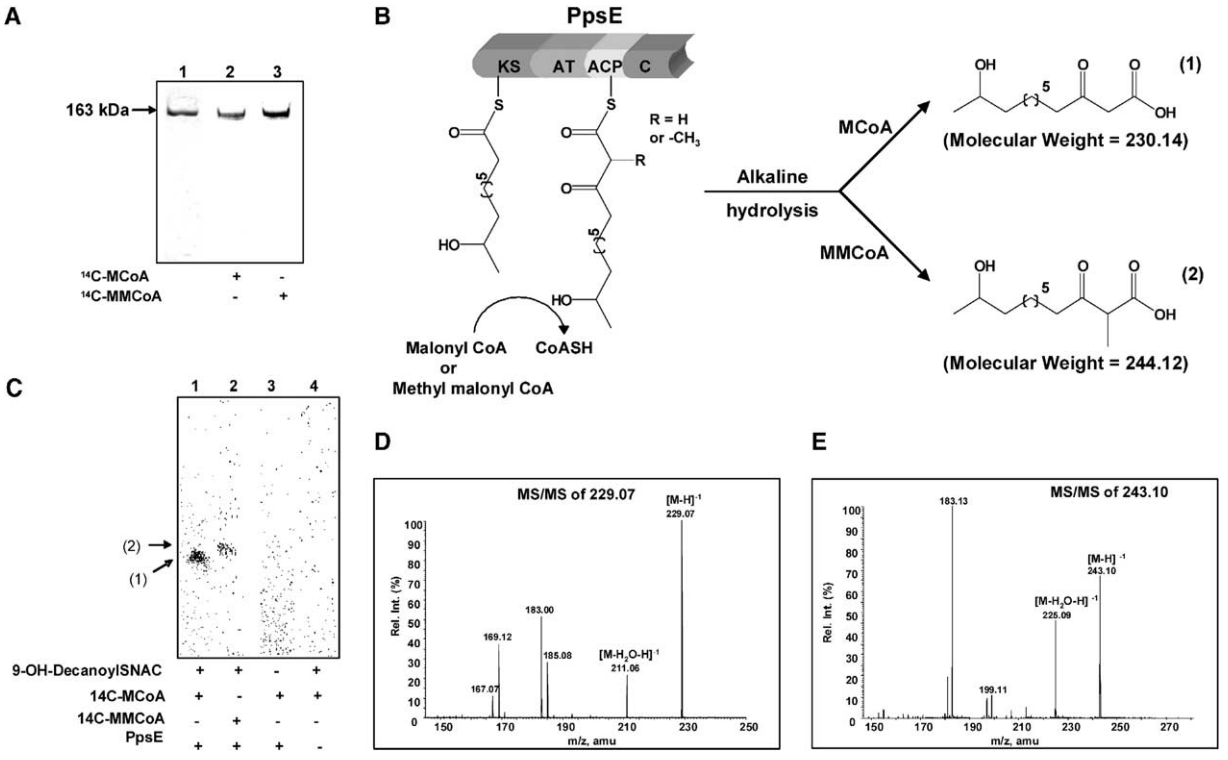


Figure 4. PpsE Catalyzes the Final Step in the Synthesis of Phthiocerol

(A) AT domain specificity of PpsE protein. Lane 1 shows the Coomassie-stained, purified PpsE protein. Incubation of the protein with radiolabeled MCoA and MMCoA (lanes 2 and 3) indicate that both the extender units can be incorporated by this protein.

(B) Enzymic activity of PpsE protein. The starter unit was exogenously incorporated in PpsE protein by using 9-hydroxy decanoyl-SNAC. The malonyl and the MMCoA extension products were analyzed.

(C) Radio-TLC characteization of PpsE products. PpsE reactions with malonate and methyl malonate extender showed distinct a $R_f$ of 0.34 and 0.38, respectively (lanes 1 and 2). The control reactions carried out in the absence of the starter unit or the protein did not show any radiolabeled bands (lanes 3 and 4).

(D) Chemical identity of the malonate-extended PpsE product. ESI-MS showed (m/z) peak at 229.07 corresponding to 11-hydroxy 3-keto dodecanoic acid. MS/MS analysis showed characteristic dehydration peak.

(E) MS analysis of methyl malonate-extended PpsE product. Molecular ion peak was obtained at 243.10 corresponding to 11-hydroxy 2-methyl 3-keto dodecanoic acid. The fragmentation pattern showed the dehydration ion at 225.09 (m/z).

showed a molecular ion peak of $[M - H]^{-1}$ 325.28 Da (Figure 5D), which corresponds to the product obtained on completion of third iteration. It can be observed that peak corresponding to the $^{14}C$-isotope (327.28 Da) was considerably more populated than the $^{13}C$-isotopic peak (326.28 Da). The MS/MS fragmentation pattern of the molecular ion peak 327.28 is also shown (Figure 5E). The weakly populated dehydration peak at 309.21 as well as other fragments with a mass difference of 14 corresponding to the $-CH_2$ units confirmed the products synthesized by Mas protein. Trimethylated mycocerosic acids esterified to the phthiocerol have been previously identified in *M. tuberculosis* cell wall (Daffe and Laneelle, 1988). Enzymatic analysis of Mas protein thus reveals some of the mechanisms by which chemical diversity is generated by using different primer units and by varying the iteration cycles. Interestingly, *M. tuberculosis* PDIMs contain D-specific mycocerosic acids, whereas closely related species *M. marinum* produce L-specific mycocerosic acids. The in vitro synthesis of the multimethylated fatty acids by *M. tuberculosis* Mas protein would provide a good opportunity to dissect such stereochemical preferences observed in related species.

### PapA5 Catalyzes the Hand-to-Hand Transfer of Intermediates Bound to the Protein to Synthesize PDIM

The final step in the assembly of PDIM would require esterification of mycocerosic acids onto the hydroxyl group of phthiocerol intermediate. Initial studies carried out with PapA5 protein by using 2-dodecanol (surrogate substrate for phthiocerol) and $^{14}C$-lauroyl-CoA showed formation of a new product with an $R_f$ of 0.8. This metabolite was analyzed on HPLC (Figure 6A) and characterized as 2-dodecanoyl dodecanoate ester by ESI-MS/MS (Figure 6B). The identity of this product was also confirmed by the synthesis of the chemical reference. Similar activity with surrogate substrates was also recently reported for PapA5 protein (Onwueme et al., 2004). Because mycocerosic acids are not actively released from the Mas protein, it is debatable whether mycocerosic acids are first released and then activated as acyl-CoA thioesters or the enzyme bound products

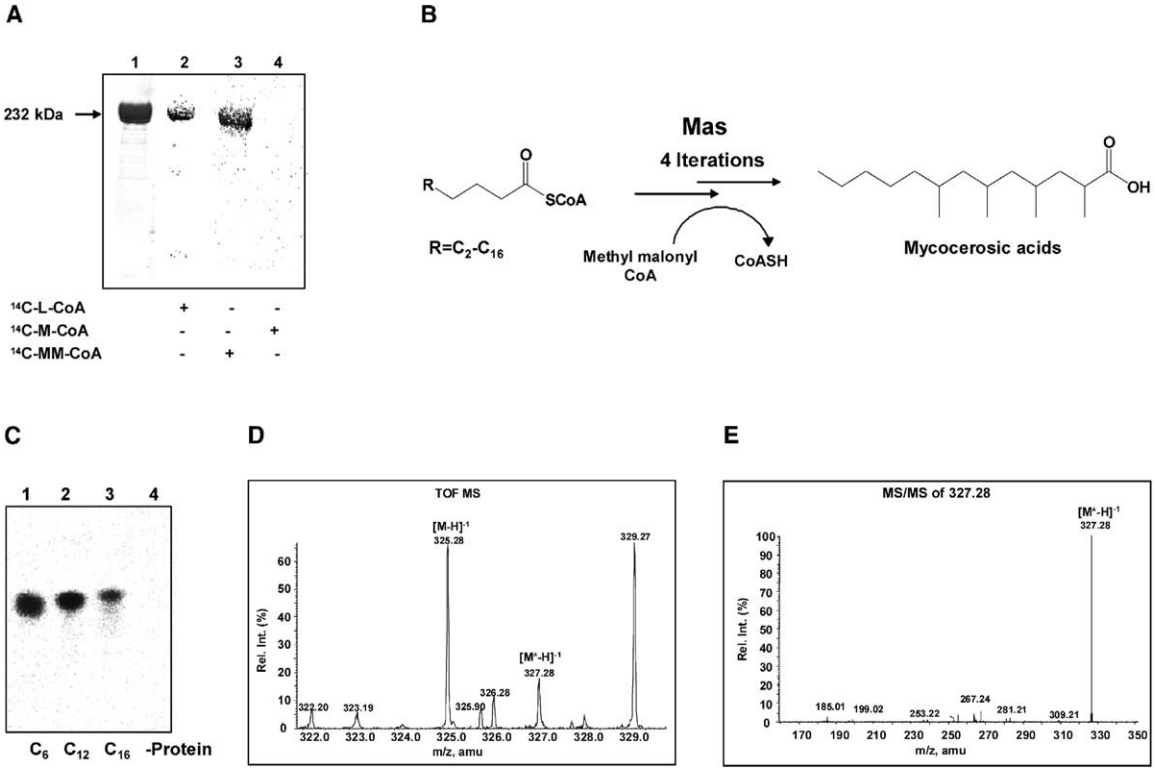


Figure 5. Synthesis of Mycocerosic Acids

(A) Labeling of the Mas protein by starter and extender units. The purified protein is shown in lane 1. The radiolabeled protein bands are observed for $^{14}$C-lauroyl-CoA (lane 2) and $^{14}$C-MMCoA (lane 3), but not with $^{14}$C-MCoA (lane 4).

(B) Synthesis of mycocerosic acids. The Mas protein elongates a range of $C_6$–$C_{20}$ starter CoAs to synthesize methylated fatty acids by iterative rounds of condensation with MMCoA units.

(C) Radio-TLC characterization of Mas products. Radiolabeled products were chemically released from the protein and analyzed by radio-TLC. The autoradiograph shows the Mas products formed by using $C_6$, $C_{12}$, and $C_{16}$ acyl-CoAs (lanes 1–3). No product is formed in the absence of the protein (lane 4).

(D) ESI-MS analysis of the Mas product formed by three iterative cycles. The left panel shows the molecular ion $[M - H]^{-1}$ at 325.27. The MS/MS fragmentation pattern of C-14 isotopic molecule corresponding to m/z 327.28 is shown in the right panel. The dehydration peak at 309.21 as well as fragments generated due to loss of $-CH_2$ unit are shown in the spectrum.

(E) Tandem MS of Mas product. Some of the ions show characteristic mass difference of 14 corresponding to the $-CH_2$ units.

can be directly transferred on to alcohols. Interestingly, PapA5 directly transferred the Mas bound, $^{14}$C-labeled mycocerosic acids analogs on to 2-dodecanol (Figure 6C). The identity of product was confirmed by mass spectrometry. The MS/MS fragmentation of 411.22 molecular ion shows the characteristic $\alpha$ cleavage resulting in the loss of alkoxy group and formation of a positively charged acid species (Figure 6D). The steady-state kinetic parameters of ester formation showed good catalytic rates and obeyed saturation kinetics with apparent kinetic parameters of $k_{cat}$ = 11.74 pmol $mg^{-1}min^{-1}$ (Figure 6E), and $K_M$ = 23.97 $\mu$M for 2-dodecanol (Figure 6F). The dependence of the rate of mycocerosate ester formation with the mycocerosic acids bound to the Mas protein showed complex kinetic profile (Figure 6G). The catalytic activity peaks at a concentration of 3.7 $\mu$M and drops thereafter. The reaction velocity at 3.7 $\mu$M corresponds to the apparent $k_{cat}$ of 13.21 pmol $mg^{-1}min^{-1}$ and is comparable to values obtained for 2-dodecanol. The inhibition effects observed beyond 4.0 $\mu$M may in fact arise from the competition with unacylated Mas protein. Indeed, 1:1 ratio of acylated to unacylated Mas protein showed ~50% decrease in ester formation (Figure 6H). This observation also provides evidence for the functional interaction between the Mas and PapA5 proteins. Similar inhibitory effects of unacylated ACP have been previously reported for actinorhodin polyketide biosynthesis (Dreier et al., 1999). Our studies thus demonstrate that PapA5 protein can directly transfer mycocerosic acid analogs from Mas protein to alcohols.

### Computational Analysis of the Interaction of Mas ACP Domain with PapA5

Biochemical studies indicated that protein-protein interaction between ACP and PapA5 is likely to play an important role in this enzymatic transfer of mycocerosic acids. Computational analysis involving docking simulations was carried out to identify putative residues mediating recognition. Although ACP domains have low-sequence homology, these proteins possess a conserved overall 3D structural fold. Based on thread-

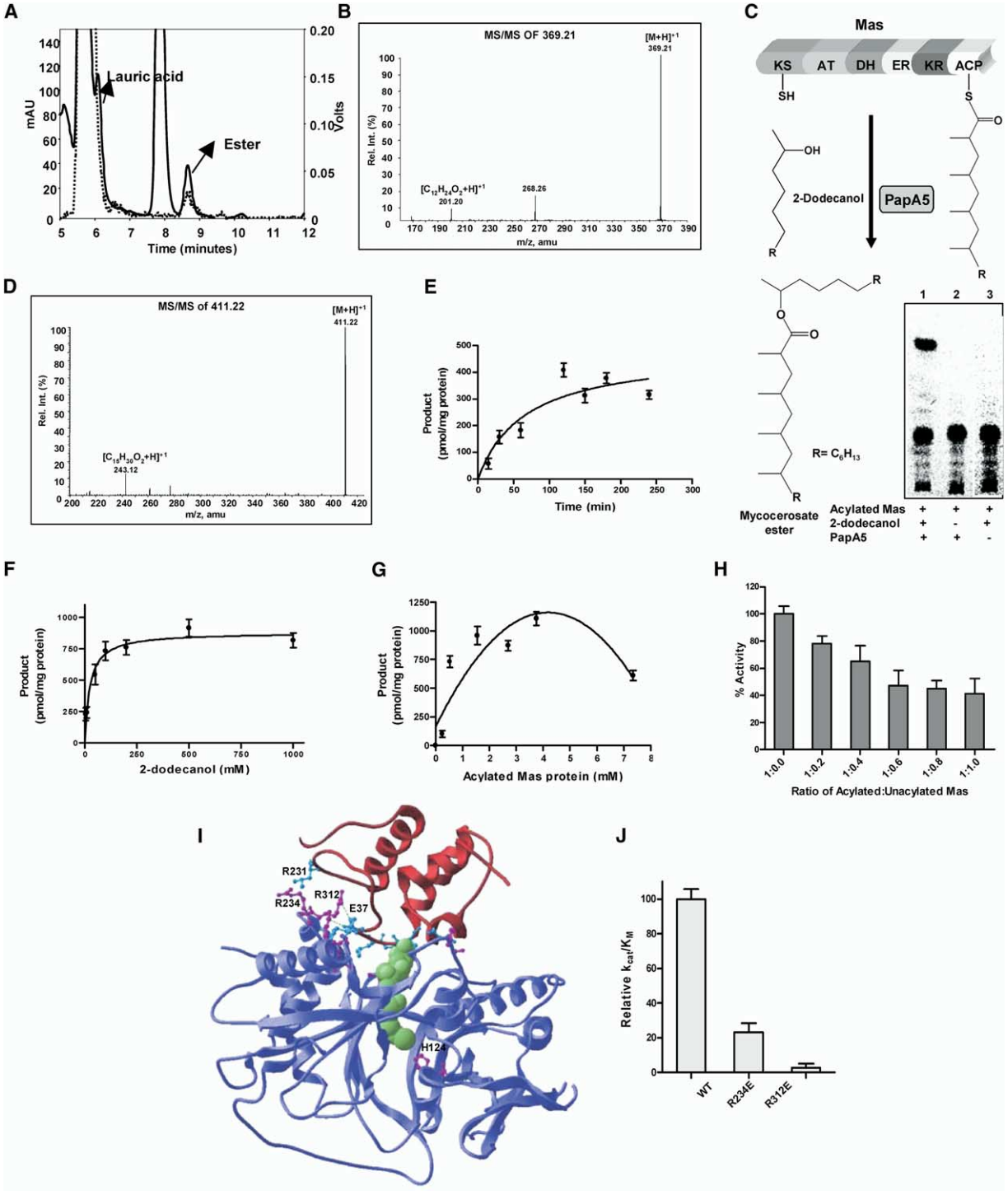


Figure 6. Esterification Reaction Catalyzed by PapA5

(A) Enzymatic assay of PapA5 protein. The radioactive product synthesized by PapA5 protein was analyzed on reverse-phase HPLC, and the chromatograms were recorded by simultaneously using photomultiplier and radioactive detectors (left). In order to enable the detection at 215 nm, the chemical standard was coinjected along with the reaction products. The 8.5 min peak recorded at 215 nm (solid line) superimposes with the radioactive measurements (dashed line).

(B) ESI-MS/MS characterization of 8.5 min peak. Positive ion MS analysis showed molecular peak at m/z 369.21 that corresponds to 2-dodecanoyl dodecanoate ester. The MS/MS-produced m/z peak at 201.2 corresponds to the characteristic $[C_{12}H_{24}O_2 + H]^{+1}$ ester fragment ion generated from the dodecanoic acid moiety.

(C) PapA5 catalyzed transesterification of mycocerosic acids. Covalently acylated products at the ACP domain of Mas protein are transferred onto the 2-dodecanol to form the mycocerosate ester. Lane 1 on the radio-TLC shows the formation of the ester with an $R_f$ value of 0.8. The ester is not formed in the absence of 2-dodecanol (lane 2).

ing analysis of Mas ACP (Weber et al., 2000), a structural model was built by using the MODELLER package. The modeled Mas ACP was docked on the crystal structure of PapA5 (1Q9J) (Buglino et al., 2004) by using the FTDOCK program (Jackson et al., 1998). The top-ranking complexes based on RPscore were analyzed such that the P-pant group attached to catalytic serine of ACP should reach the catalytic histidine of PapA5. Interestingly, one such ACP-PapA5 complex was found in the top 1% of the solutions given by FTDOCK. Most of the ACP interactions with PapA5 (Figure 6I) were restricted to residues 23–25 (after helix 1), 37–41 (the loop between helix 1 and helix 2), and 66–67 (after helix 3), analogous to previous studies (Finking et al., 2004; Keatinge-Clay et al., 2003). Although most of the contacts are hydrophobic, one of the Arg312 is involved in electrostatic interactions with a *Glu*37 of ACP. Two other Arg (231 and 234) residues have their charged ends exposed to the solvent, and their nonpolar groups make contacts with hydrophobic residues of ACP. The conformation of ACP bound P-pant group inside the PapA5 channel was modeled by superposing the crystal structure of carnitine acetyl transferase (1NDI) on PapA5 and transforming the coordinates of CoA from 1NDI (Jogl and Tong, 2003). In the energy-minimized model of ACP-PapA5 complex, the thiol group of ACP bound P-pant is within 3.7 Å of catalytic histidine amino acid.

### Direct Evidence of Physical Interaction of Mas ACP Domain with the PapA5

Because PapA5 can use both CoA- and ACP-acylated substrates, the kinetics for these two substrates were evaluated. Although the Michaelis constant ($K_M$) for lauroyl-CoA was estimated to be 118 μM, the apparent $K_M$ for the acylated Mas protein was of the order of 0.39 μM. More than 100-fold difference between the affinities for the two substrates clearly demonstrates that the Mas bound substrates are highly favored over CoA-esters. These studies thus emphasize the importance of protein-protein interactions for direct transfer of substrates. The direct physical interactions of Mas protein with PapA5 was further investigated by mutating two surface residues Arg234 and Arg312, as predicted by our model. The R312E mutant protein showed dramatic 40-fold increase in $K_M$ for Mas-ACP, whereas R234E showed only 4-fold increase, and both of these mutant proteins showed $k_{cat}$ values similar to the PapA5 protein. The relative catalytic efficiency ($k_{cat}/K_M$) for wild-type (wt) and mutant proteins are plotted in Figure 6J. The dramatic decrease in catalytic efficiency of R312E corroborates our in silico-derived model and also demonstrates that Arg residues embedded in the hydrophobic patch on the PapA5 protein are important in recognizing Mas ACP domain. A similar patch of Arg residues in the FabG protein was shown to be important to recognize highly conserved electronegative/hydrophobic residues along helix 2 of ACP protein (Zhang et al., 2003).

### Engineering Nonmethylated Mycocerosate Esters

The precise understanding of the functions of proteins from the Pps cluster provided impetus to engineer analogs by varying the substrate specificity of these proteins. Because the hallmark of mycobacterial lipids is the presence of multimethylated branched-chain fatty acids, we decided to engineer a nonmethylated analog of PDIM (Figure 7A). Our in vitro study showed that the multimethylated mycocerosic acids are biosynthesized by the iterative condensation of MMCoA with long chain starter substrates. In these megasynthases, the extender unit specificity is determined by the AT domain. Previous analysis had suggested that a Phe residue present at the substrate binding pocket of AT domains facilitates discrimination between malonate and methylmalonate extender groups (Yadav et al., 2003). Based on computational prediction, *Ser726* amino acid in Mas protein was mutated to Phe to facilitate incorporation of malony-CoA unit. The mutant protein was expressed and purified from *E. coli* analogous to wt protein. Although the wt Mas protein does not show any product with MCoA, the mutant protein readily accepted MCoA (Figure 7B). The S726F mutant protein also showed marginal activity with MMCoA. The affinity of mutant AT domains for incorporation of MCoA and MMCoA was compared. Because the rate of acylation of AT domains even at low temperatures is very fast (150 $s^{-1}$) (Rangan and Smith, 1997) in comparison to the overall rate of product formation (0.08 $min^{-1}$), it is not feasible to estimate apparent $K_M$ of AT domains by using product formation assays. Hence, we used transacylation assays to specifically determine the kinetic parameters for AT domains with N-acetylcysteamine as an acceptor. The strict specificity of Mas AT domain

(D) MS characterization of mycocerosate ester. Mas assays were carried out by using hexanoyl-CoA as the starter unit and MMCoA as the extender unit. The products formed were transesterified on 2-dodecanol by PapA5 protein. Nanospray tandem mass analysis of mycocerosate ester product in positive ion mode shows m/z fragment at 243.12 corresponding to protonated mycocerosic acids generated from three cycles of iterations.

(E) Time course of product formation catalyzed by PapA5 protein.

(F) Initial velocity measurements for the synthesis of mycocerosate ester performed at varying concentrations of 2-dodecanol.

(G) Rate of product formation by PapA5 as a function of the concentration of the mycocerosic acids acylated to Mas protein.

(H) Bar graph representing the amount of ester formed in the presence of unacylated protein as a percentage of the total product formed. The ratio of acylated to unacylated Mas protein is shown on the x axis.

(I) A computational model indicating the interaction of the PapA5 protein with the Mas ACP domain obtained by docking the modeled ACP on the crystal structure of PapA5. The PapA5 protein is shown in blue, and the docked ACP is shown in red. The P-pant arm is shown in green and positions next to the catalytic His124. The crucial interfacial interacting residues are highlighted.

(J) Relative catalytic efficiency for the wild-type and PapA5 mutant proteins are compared. Mutation of two surface arginine residues to glutamic acid results in dramatic loss of direct transfer of Mas bound acids onto phthiocerol surrogate substrate. In all the panels, results are means (n =3) with SEM values less than 15%.

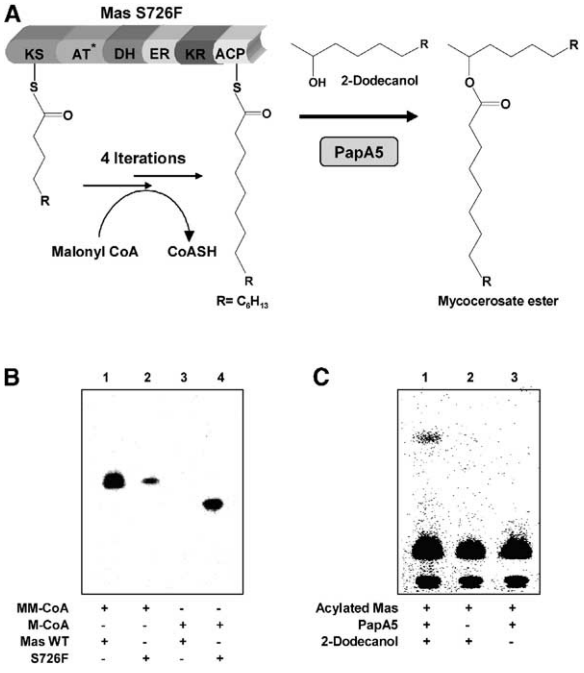


Figure 7. Rational Design of PDIM Analog

(A) The specificity of Mas AT domain was engineered (S726F) to accept MCoA extender units to biosynthesize nonmethylated mycocerosic acids. The enzyme bound products could be transferred to 2-dodecanol to generate modified nonmethylated mycocerosate esters.

(B) Characterization of S726F Mas protein products. The product observed at an $R_f$ of 0.55 corresponds to multimethylated mycocerosic acids, whereas S726F mutant protein produces linear unbranched product of lower $R_f$ 0.41 by using MCoA (lane 4). The mutant protein also makes small amounts of methylated mycocerosic acid product by using MMCoA (lane 2).

(C) Enzymatic synthesis of mycocerosate ester analog. The Mas protein bound, nonmethylated fatty acids could be transesterified onto 2-dodecanol. Lane 1 on the radio-TLC shows the formation of ester product with an $R_f$ value of 0.8. This ester product is not formed in the absence of PapA5 protein (lane2) and 2-dodecanol (lane 3).

was also reflected in the transacylation assays. The mutant protein utilized MCoA readily and obeyed saturation kinetics (Table 1). In contrast, the catalytic efficiency of S726F mutant protein for MMCoA was severely reduced to 3% of the wt protein. These studies demonstrate that alteration of single amino acid changes the AT domain specificity from MMCoA to MCoA.

In order to produce nonmethylated mycocerosate esters, we probed whether PapA5 protein would be able to transesterify modified lipids directly onto phthiocerol surrogate substrate. The extension of hexanoyl starter by three iterative condensations and associated reductions would result in the formation of lauroyl moiety. Indeed, S726F Mas bound lauroyl chain could be transesterified to 2-dodecanol. The new radioactive band at 0.8 $R_f$ observed in lane 1 (Figure 7C) comigrated with the 2-dodecanol dodecanoate synthetic standard and was confirmed by MS analysis. The studies described here demonstrate functional application of the system to generate unusual lipid analogs. Similar strategies could be extended further to generate unnatural lipids.

## Conclusions

The results presented here unravel molecular mechanisms by which mycobacteria synthesize complex lipids from ubiquitous n-long-chain fatty acids. It is interesting to note that the complete biosynthesis proceeds in the cytoplasm by using soluble enzymes, although most of the enzymatic intermediates are covalently sequestered on to the protein. Our study demonstrates that proteins present in the *pps* cluster, in principle, can biosynthesize PDIMs by condensing long-chain fatty acids with simple acyl-CoA units. The PapA5 protein can directly transfer mycocerosic acids from Mas protein to phthiocerol. This mode of hand-to-hand transfer may be the most efficient method to channelize intermediates, because free long-chain methylated mycocerosic acids or its CoA-esters can interfere with lipid metabolism. The proposed requirement of other PKS genes in PDIM biosynthesis is not clear (Rousseau et al., 2003; Sirakova et al., 2003a, 2003b) and might require systematic reinvestigations. The PKS12 gene, earlier implicated in PDIM biosynthesis, was recently shown to produce mycoketide leading to the biosynthesis of mannosyl-β-1-phosphomycoketide (Matsunaga et al., 2004). In another report, a fraction of mycobacterial mutants of the unrelated loci were shown to be deficient in PDIMs (Domenech et al., 2004). The in vitro synthesis of the components of PDIM assembly-line biosynthetic machinery described in this study would facilitate investigation of the mechanistic and structural basis of the programming of these complex lipids. The characterization of unique biosynthetic pathways of the essential component of the mycobacterial cell envelope makes them attractive targets for developing new antimycobacterial drugs.

Based on in vitro assays with Mas, we have engineered new specificity that would incorporate MCoA instead of MMCoA by changing a single amino acid. The mutant protein produces linear fatty acids, which can be transferred to phthiocerol to produce new analogs. Because the properties of branched-chain fatty acids are very different from the linear-chain fatty acids, this experiment would provide the basis to alter the mycobacterial coat to understand host-pathogen interac-

Table 1. Steady-State Kinetic Parameters for Wt and S726F Mas Proteins for Extender Units

| | Methyl Malonyl-CoA | | | Malonyl-CoA | | |
|---|---|---|---|---|---|---|
| | $K_M(\mu M)$ | $K_{cat}(min^{-1})$ | $K_{cat}/K_M(s^{-1}\ M^{-1})$ | $K_M(\mu M)$ | $K_{cat}(min^{-1})$ | $K_{cat}/K_M(s^{-1}\ M^{-1})$ |
| Wt | 70.1 | 13.21 | 3133.9 | ND | ND | ND |
| S726F | 238.6 | 1.20 | 83.8 | 117.6 | 6.11 | 965.9 |

Results are means (n = 3) with SEM values less than 15%.

tions. The implications of this experiment are immense and provide a tantalizing strategy to investigate the effect of subtle chemical changes at the mycobacterial surface coat. Cell-free reconstitution studies thus apart from demonstrating exact functions can provide new ways to explore host-pathogen interactions.

## Experimental Procedures

### Cloning, Expression, and Purification

BAC genomic DNA of *M. tuberculosis* H37Rv (Brosch et al., 1998) was digested with restriction enzymes (BglII or AscI) and the fragments were cloned in either BglII-digested Litmus28 or AscI-digested pNEB193 to generate a shotgun library. Suitable restriction enzyme sites were engineered at the 5′ and the 3′ to facilitate cloning in pET expression vectors. The *pks* (*ppsA*, *ppsB*, and *ppsE*) and the *fas* (*mas*) genes were coexpressed with *sfp* for expression in the holo P-pant form in BL21-DE3 strain of *E. coli*. All the proteins were purified by $Ni^{2+}$-NTA affinity chromatography followed by anion exchange (ResourceQ) chromatography using AKTA (Pharmacia) chromatographic system.

### Specificity of the AT Domain and Posttranslational Modification of the ACP Domain

The PKS proteins were incubated with $^{14}$C-MCoA and $^{14}$C-MMCoA at 4°C for 2 min. To confirm phosphopantetheinylation, proteins were incubated with $^{14}$C-MCoA or $^{14}$C-MMCoA and treated with 0.3 M hydroxylamine (pH 8.5) at 30°C for 30 min. In both these cases, the reactions were quenched, resolved by SDS-PAGE gel, and analyzed by phosphorimager.

### Characterization of PpsA, PpsB, and PpsE Product

The fatty acid activation by FadD26 and its transfer onto PpsA were performed as described previously.(Trivedi et al., 2004). PpsA enzymatic assays included 100 mM phosphate buffer (pH 7.0), 2 mM ATP, 8 mM $MgCl_2$, 100 μM fatty acid, 100 μM $^{14}$C-MCoA, 4 mM NADPH, 10% glycerol, 2 mM TCEP, 0.1 μM FadD26, and 5 μM PpsA in 100 μl reaction volume. After incubation at 30°C for 6 hr, the reactions were quenched and precipitated by addition of 0.9 ml of ice-cold 10% TCA (w/v). The pellet was air dried and resuspended in 100 μl of 50 mM Tris-HCl (pH 8.0), and KOH hydrolysis (350 mM) at 65°C for 20 min was carried out to release the covalently bound products. After acidification with formic acid, the product was extracted with ethyl acetate and spotted on TLC plate. The TLC plate was developed by using 60:40:5 ethyl acetate:hexanes:acetic acid, and the radioactive spots were analyzed by phosphorimager. The product was eluted from the TLC and characterized by ESI-MS (APT QSTAR Pulsar I MS/MS, Applied Biosystems). For PpsE, reactions priming was carried out by 9-hydroxy decanoyl-SNAC, and both $^{14}$C-MCoA and $^{14}$C-MMCoA were used as extender units. In order to characterize PpsB products, the PpsA bound products were separated from free substrate by $Ni^{2+}$-NTA affinity chromatography. Acylated PpsA protein was incubated along with PpsB protein and MCoA at 30°C for 8 hr. The release of PpsB bound products and its analysis was performed analogous to PpsA protein.

### Synthesis of 9-Hydroxy Decanoyl SNAC and 2-Dodecanoyl Dodecanoate Ester

Methyl ester of 9-decenoic acid was synthesized by refluxing with methanol/sulfuric acid. This was followed by mercuration-demercuration reactions with mercuric acetate and sodium borohydride to give the 9-hydroxy decanoic acid derivative. The hydroxyl group was protected by using *tert*-butyldimethyl silyl chloride (TBDMS-Cl). This protected ester was then saponified, and N-acetyl cysteamine thioester derivative was synthesized as reported earlier (Suo et al., 2000). Finally, the TBDMS was removed by hydrofluoric acid. For 2-dodecanoyl dodecanoate ester, equimolar concentrations of lauroyl chloride and 2-dodecanol were dissolved in dichloromethane (DCM), and the reaction was carried out at 0°C for 3 hr in the presence of triethylamine. All the intermediate as well as final steps were purified by using silica column chromatography, and the identity of the compound was confirmed by ESI-MS and NMR spectroscopy.

### Enzymatic Synthesis of Mycocerosic Acids and Characterization

The Mas assays contained 100 mM phosphate buffer (pH 7.0), 200 μM starter acyl-CoA ($C_6$–$C_{20}$), 200 μM $^{14}$C-MMCoA, 4 mM NADPH, 10% glycerol, 2 mM TCEP, and 2.5 μM Mas protein. Reactions were incubated at 30°C for 1 hr, and the products were released by alkaline hydrolysis, extracted, and analyzed as described previously (Rainwater and Kolattukudy, 1983) The products were extracted from the TLC and analyzed on ESI-MS.

### Characterization of PapA5 Catalyzed Synthesis of Esters

The formation of ester by PapA5 protein was monitored by radio-TLC and HPLC. The enzymatic assays contained 100 mM phosphate buffer (pH 7.0), 100 μM $^{14}$C-lauroyl-CoA, 1 mM 2-dodecanol, and PapA5 protein (at 30°C for 2 hr). The product was extracted in chloroform, concentrated, and injected on C5 reverse-phase column (solvent system isocratic 10% dichloromethane in methanol) along with 20 μg of the chemically synthesized ester for its detection at 215 nm on HPLC (Shimadzu). The radioactivity was monitored online by a radioactive detector (IN/US system β-RAM Model 3). For transesterification assays, the covalently acylated mycocerosic acids bound to the Mas protein were produced by using hexanoyl CoA as starter unit and $^{14}$C-MMCoA. To this reaction, 1 mM 2-dodecanol and PapA5 protein were added and further incubated at 30°C for various time points. The reaction products were then extracted and separated on TLC by using 70:30 hexanes:ethyl acetate. The radioactive products were analyzed by phosphorimager. Similar assays were performed with Mas S726F with $^{14}$C-MCoA for biosynthesis nonmethylated mycocerosyl-ester analogs. For inhibition assays with unacylated MAS, similar reactions were carried out, except that PapA5 protein was preincubated with unacylated Mas protein for 10 min.

### Docking of the ACP Domain of Mas onto the PapA5

Apo ACP of Mas was modeled with Modeller (Sali and Blundell, 1993) version 6v2 using PCP (PDB code 1DNY) as template. Rigid-body docking of modeled Mas ACP and PapA5 (PDB code 1Q9J) was performed by using FTDock and RPScore from 3D-Dock suite, (Jackson et al., 1998) with PapA5 as a fixed molecule. Resulting complexes were filtered with distance constraints for *Ser42* toward the mouth of the P-pant binding channel. P-pant arm of CoASH was transformed to PapA5 by superimposing it with carnitine acetyltransferase (PDB code 1NDI) using DALI (Holm and Sander, 1993).

### Site-Directed Mutagenesis

All the mutants were constructed by using QuickChange Site-Directed Mutagenesis Kit (stratagene). All the mutations were confirmed by DNA sequencing, and the mutant proteins were expressed in *E. coli* and purified in a manner analogous to wt proteins.

### Transacylation Assays to Study AT Domain Specificity

The steady-state kinetic parameters of the specificity of AT domains of Mas wt and S726F were carried out based on previously described transacylation assays (Liou et al., 2003). These assays were with apo proteins (2.5 μM) in 100 mM phosphate buffer (pH 7.0), 5 mM N-acetylcysteamine, 25–1000 μM of MCoA, or MMCoA at 30°C for 2 min. The reaction products were extracted and spotted on a TLC plate, developed with 70:30:5 of ethyl acetate: hexanes:glacial acetic acid (v/v), and analyzed by using phosphorimager.

## Acknowledgments

The authors thank Anil Bobin, Sunder Bisht, and Suresh Chand for their technical support. The authors also thank Dr. Sandip K. Basu for helpful discussions. P.A. and M.Z.A. are Senior Research Fellows of the Council of Scientific & Industrial Reaseach, India. R.S.G

is a Wellcome Trust International Senior Research Fellow in India. This work was also supported by grants to the National Institute of Immunology by Department of Biotechnology, India.

Received: August 5, 2004
Revised: December 8, 2004
Accepted: February 4, 2005
Published: March 3, 2005

## References

Admiraal, S.J., Walsh, C.T., and Khosla, C. (2001). The loading module of rifamycin synthetase is an adenylation-thiolation didomain with substrate tolerance for substituted benzoates. Biochemistry *40*, 6116–6123.

Ansari, M.Z., Yadav, G., Gokhale, R.S., and Mohanty, D. (2004). NRPS-PKS: a knowledge-based resource for analysis of NRPS/PKS megasynthases. Nucleic Acids Res. *32*, W405–W413.

Azad, A.K., Sirakova, T.D., Rogers, L.M., and Kolattukudy, P.E. (1996). Targeted replacement of the mycocerosic acid synthase gene in Mycobacterium bovis BCG produces a mutant that lacks mycosides. Proc. Natl. Acad. Sci. USA *93*, 4787–4792.

Azad, A.K., Sirakova, T.D., Fernandes, N.D., and Kolattukudy, P.E. (1997). Gene knockout reveals a novel gene cluster for the synthesis of a class of cell wall lipids unique to pathogenic mycobacteria. J. Biol. Chem. *272*, 16741–16745.

Barry, C.E., 3rd. (2001). Interpreting cell wall 'virulence factors' of Mycobacterium tuberculosis. Trends Microbiol. *9*, 237–241.

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

Brosch, R., Gordon, S.V., Billault, A., Garnier, T., Eiglmeier, K., Soravito, C., Barrell, B.G., and Cole, S.T. (1998). Use of a Mycobacterium tuberculosis H37Rv bacterial artificial chromosome library for genome mapping, sequencing, and comparative genomics. Infect. Immun. *66*, 2221–2229.

Buglino, J., Onwueme, K.C., Ferreras, J.A., Quadri, L.E., and Lima, C.D. (2004). Crystal structure of PapA5, a phthiocerol dimycocerosyl transferase from Mycobacterium tuberculosis. J. Biol. Chem. *279*, 30634–30642. Published online May 3, 2004. DOI: 10.1074/jbc.M404011200.

Camacho, L.R., Constant, P., Raynaud, C., Laneelle, M.A., Triccas, J.A., Gicquel, B., Daffe, M., and Guilhot, C. (2001). Analysis of the phthiocerol dimycocerosate locus of Mycobacterium tuberculosis. Evidence that this lipid is involved in the cell wall permeability barrier. J. Biol. Chem. *276*, 19845–19854.

Cane, D.E., and Walsh, C.T. (1999). The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases. Chem. Biol. *6*, R319–R325.

Constant, P., Perez, E., Malaga, W., Laneelle, M.A., Saurel, O., Daffe, M., and Guilhot, C. (2002). Role of the pks15/1 gene in the biosynthesis of phenolglycolipids in the Mycobacterium tuberculosis complex. Evidence that all strains synthesize glycosylated p-hydroxybenzoic methly esters and that strains devoid of phenolglycolipids harbor a frameshift mutation in the pks15/1 gene. J. Biol. Chem. *277*, 38148–38158. Published online July 22, 2002. 10.1074/jbc.M206538200

Cox, J.S., Chen, B., McNeil, M., and Jacobs, W.R., Jr. (1999). Complex lipid determines tissue-specific replication of Mycobacterium tuberculosis in mice. Nature *402*, 79–83.

Daffe, M. (1991). Further stereochemical studies of phthiocerol and phenol phthiocerol in mycobacteria. Res. Microbiol. *142*, 405–410.

Daffe, M., and Laneelle, M.A. (1988). Distribution of phthiocerol diester, phenolic mycosides and related compounds in mycobacteria. J. Gen. Microbiol. *134*, 2049–2055.

Domenech, P., Reed, M.B., Dowd, C.S., Manca, C., Kaplan, G., and Barry, C.E., 3rd. (2004). The role of MmpL8 in sulfatide biogenesis and virulence of mycobacterium tuberculosis. J. Biol. Chem. *279*, 21257–21265.

Dreier, J., Shah, A.N., and Khosla, C. (1999). Kinetic analysis of the actinorhodin aromatic polyketide synthase. J. Biol. Chem. *274*, 25108–25112.

Dubey, V.S., Sirakova, T.D., and Kolattukudy, P.E. (2002). Disruption of msl3 abolishes the synthesis of mycolipanoic and mycolipenic acids required for polyacyltrehalose synthesis in Mycobacterium tuberculosis H37Rv and causes cell aggregation. Mol. Microbiol. *45*, 1451–1459.

Finking, R., Mofid, M.R., and Marahiel, M.A. (2004). Mutational analysis of peptidyl carrier protein and acyl carrier protein synthase unveils residues involved in protein-protein recognition. Biochemistry *43*, 8946–8956.

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

Gokhale, R.S., and Tuteja, D. (2001). Biochemistry of polyketide synthases. In Biotechnology, H.-J. Rehm, G. Reed, A. Pühler, and P. Stadler, eds. (Weinheim, Germany: WILEY-VCH), pp. 341–372.

Gokhale, R.S., Tsuji, S.Y., Cane, D.E., and Khosla, C. (1999). Dissecting and exploiting intermodular communication in polyketide synthases. Science *284*, 482–485.

Holm, L., and Sander, C. (1993). Protein structure comparison by alignment of distance matrices. J. Mol. Biol. *233*, 123–138.

Hopwood, D.A. (1997). Genetic contributions to understanding polyketide synthases. Chem. Rev. *97*, 2465–2498.

Jackson, R.M., Gabb, H.A., and Sternberg, M.J. (1998). Rapid refinement of protein interfaces incorporating solvation: application to the docking problem. J. Mol. Biol. *276*, 265–285.

Jogl, G., and Tong, L. (2003). Crystal structure of carnitine acetyltransferase and implications for the catalytic mechanism and fatty acid transport. Cell *112*, 113–122.

Karakousis, P.C., Bishai, W.R., and Dorman, S.E. (2004). Mycobacterium tuberculosis cell envelope lipids and the host immune response. Cell. Microbiol. *6*, 105–116.

Kearney, G.C., Gates, P.J., Leadlay, P.F., Staunton, J., and Jones, R. (1999). Structural elucidation studies of erythromycins by electrospray tandem mass spectrometry II. Rapid Commun. Mass Spectrom. *13*, 1650–1656.

Keating, T.A., Ehmann, D.E., Kohli, R.M., Marshall, C.G., Trauger, J.W., and Walsh, C.T. (2001). Chain termination steps in nonribosomal peptide synthetase assembly lines: directed acyl-S-enzyme breakdown in antibiotic and siderophore biosynthesis. Chembiochem *2*, 99–107.

Keating, T.A., Marshall, C.G., Walsh, C.T., and Keating, A.E. (2002). The structure of VibH represents nonribosomal peptide synthetase condensation, cyclization and epimerization domains. Nat. Struct. Biol. *9*, 522–526.

Keatinge-Clay, A.T., Shelat, A.A., Savage, D.F., Tsai, S.C., Miercke, L.J., O'Connell, J.D., 3rd, Khosla, C., and Stroud, R.M. (2003). Catalysis, specificity, and ACP docking site of Streptomyces coelicolor malonyl-CoA:ACP transacylase. Structure (Camb) *11*, 147–154.

Kennedy, J., Auclair, K., Kendrew, S.G., Park, C., Vederas, J.C., and Hutchinson, C.R. (1999). Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. Science *284*, 1368–1372.

Khosla, C. (2000). Natural product biosynthesis: a new interface between enzymology and medicine. J. Org. Chem. *65*, 8127–8133.

Kolattukudy, P.E., Fernandes, N.D., Azad, A.K., Fitzmaurice, A.M., and Sirakova, T.D. (1997). Biochemistry and molecular genetics of cell-wall lipid biosynthesis in mycobacteria. Mol. Microbiol. *24*, 263–270.

Kwon, H.J., Smith, W.C., Xiang, L., and Shen, B. (2001). Cloning and heterologous expression of the macrotetrolide biosynthetic gene cluster revealed a novel polyketide synthase that lacks an acyl carrier protein. J. Am. Chem. Soc. *123*, 3385–3386.

Lambalot, R.H., Gehring, A.M., Flugel, R.S., Zuber, P., LaCelle, M., Marahiel, M.A., Reid, R., Khosla, C., and Walsh, C.T. (1996). A new enzyme superfamily - the phosphopantetheinyl transferases. Chem. Biol. *3*, 923–936.

Liou, G.F., Lau, J., Cane, D.E., and Khosla, C. (2003). Quantitative

analysis of loading and extender acyltransferases of modular polyketide synthases. Biochemistry *42*, 200–207.

Matsunaga, I., Bhatt, A., Young, D.C., Cheng, T.Y., Eyles, S.J., Besra, G.S., Briken, V., Porcelli, S.A., Costello, C.E., Jacobs, W.R., Jr., and Moody, D.B. (2004). Mycobacterium tuberculosis pks12 produces a novel polyketide presented by CD1c to T cells. J. Exp. Med. *200*, 1559–1569.

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

Onwueme, K.C., Ferreras, J.A., Buglino, J., Lima, C.D., and Quadri, L.E. (2004). Mycobacterial polyketide-associated proteins are acyltransferases: proof of principle with Mycobacterium tuberculosis PapA5. Proc. Natl. Acad. Sci. USA *101*, 4608–4613. Published online March 18, 2004. 10.1073/pnas.0306928101.

Rainwater, D.L., and Kolattukudy, P.E. (1983). Synthesis of mycocerosic acids from methylmalonyl coenzyme A by cell-free extracts of Mycobacterium tuberculosis var. bovis BCG. J. Biol. Chem. *258*, 2979–2985.

Rangan, V.S., and Smith, S. (1997). Alteration of the substrate specificity of the malonyl-CoA/acetyl-CoA:acyl carrier protein S-acyltransferase domain of the multifunctional fatty acid synthase by mutation of a single arginine residue. J. Biol. Chem. *272*, 11975–11978.

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

Sali, A., and Blundell, T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. *234*, 779–815.

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

Sirakova, T.D., Thirumala, A.K., Dubey, V.S., Sprecher, H., and Kolattukudy, P.E. (2001). The Mycobacterium tuberculosis pks2 gene encodes the synthase for the hepta- and octamethyl-branched fatty acids required for sulfolipid synthesis. J. Biol. Chem. *276*, 16833–16839.

Sirakova, T.D., Dubey, V.S., Cynamon, M.H., and Kolattukudy, P.E. (2003a). Attenuation of Mycobacterium tuberculosis by disruption of a mas-like gene or a chalcone synthase-like gene, which causes deficiency in dimycocerosyl phthiocerol synthesis. J. Bacteriol. *185*, 2999–3008.

Sirakova, T.D., Dubey, V.S., Kim, H.J., Cynamon, M.H., and Kolattukudy, P.E. (2003b). The largest open reading frame (pks12) in the Mycobacterium tuberculosis genome is involved in pathogenesis and dimycocerosyl phthiocerol synthesis. Infect. Immun. *71*, 3794–3801.

Stewart, G.R., Robertson, B.D., and Young, D.B. (2003). Tuberculosis: a problem with persistence. Nat. Rev. Microbiol. *1*, 97–105.

Suo, Z., Chen, H., and Walsh, C.T. (2000). Acyl-CoA hydrolysis by the high molecular weight protein 1 subunit of yersiniabactin synthetase: mutational evidence for a cascade of four acyl-enzyme intermediates during hydrolytic editing. Proc. Natl. Acad. Sci. USA *97*, 14188–14193.

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

Weber, T., Baumgartner, R., Renner, C., Marahiel, M.A., and Holak, T.A. (2000). Solution structure of PCP, a prototype for the peptidyl carrier domains of modular peptide synthetases. Struct. Fold. Des. *8*, 407–418.

Weissman, K.J., Kearney, G.C., Leadlay, P.F., and Staunton, J. (1999). Structural elucidation studies of polyketide tetrasubstituted delta-lactones by gas chromatography/tandem mass spectrometry and electrospray mass spectrometry. Rapid Commun. Mass Spectrom. *13*, 2103–2108.

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

Zhang, Y.M., Wu, B., Zheng, J., and Rock, C.O. (2003). Key residues responsible for acyl carrier protein and beta-ketoacyl-acyl carrier protein reductase (FabG) interaction. J. Biol. Chem. *278*, 52935–52943.