Chapter 4

Transesterification of mycocerosic acids onto alcohol
4.1 INTRODUCTION

Phthiocerol esters of mycocerosic acids are waxy esters present in the outermost coat of *M. tuberculosis* (Brennan and Nikaido, 1995). Based on the retro-biosynthetic analysis of PDIM, it can be proposed that the final step would involve esterification of mycocerosic acids on to the phthiocerol. The wax ester synthases have been characterised in plants and *Acinetobacter* species and these proteins typically catalyse the esterification of alcohols using long chain acyl-CoAs or acyl-ACPs (Kalscheuer and Steinbuchel, 2003; Routaboul et al., 1999). Similar proteins in *M. tuberculosis* were recently identified and are proposed to catalyse the synthesis of triacylglycerol (Daniel et al., 2004). The PDIM gene cluster contains two domains that show homology to the C-domains of NRPS (PapA5 and the C-domain of PpsE).

During the course of our study, Onwueme et al. showed that the PapA5 mutants of *M. tuberculosis* lack the PDIM in the cell wall (Onwueme et al., 2004). Although the biochemical activity of PapA5 protein using simple substrate mimics was demonstrated, the mechanism of esterification of mycocerosic acids on to the phthiocerol still needs to be elucidated. The crystallographic studies of the PapA5 protein from the same group (Buglino et al., 2004) revealed a two-domain structure homologous to chloramphenicol acetyl transferase (CAT) and carnitine acetyl transferase (Jogl and Tong, 2003; Leslie, 1990). The CAT and the carnitine acetyl transferase belong to a large family of CoA-dependent acyl transferases and recently a C-domain from vibriobactin synthetase (VibH) was also shown to have a similar protein architecture (Keating et al., 2002). PapA5 protein structure
revealed two hydrophobic channels that extend from the active site to the surface of the protein and based on molecular modelling studies, it was proposed that these channels might be important for binding and orientation of the long chain acyl substrates (Buglino et al., 2004).

In this chapter, we have demonstrated that the PapA5 protein can directly transfer mycocerosic acids bound to the Mas protein on to the pthiocerol. Based on molecular docking studies, a putative model for interaction of PapA5 with ACP domain of Mas has been proposed.

4.2 METHODS

4.2.1 Expression and purification of PapA5

The *papA5* gene was cloned in our laboratory by Vijayalakshmi. The BL21-DE3 strain of *E. coli* carrying the PapA5 expression plasmid (pVM2) was induced at 30°C for 6-8 hrs with 0.5 mM IPTG to obtain the protein in the soluble form. The protein was purified by Ni$^{2+}$-NTA affinity chromatography followed by anion exchange chromatography. The cells were thawed in 50 mM Tris-HCl (pH: 8.0) buffer and disrupted using French Press at 1100 psi pressure. The cell debris was removed by centrifugation at 50,000 g for 30 min at 4°C. The proteins in the supernatant were loaded on to a Ni$^{2+}$-NTA agarose column and washed with buffer to remove the contaminating proteins. The PapA5 protein was eluted from the beads using increasing concentration of imidazole in the buffer. The fractions containing the proteins of interest were further purified by using a ResourceQ column. The protein elution profile was monitored at 280 nm and fractions were analysed on SDS-PAGE gel. The protein was stored at -80°C till further use.
4.2.2 *In vitro* reconstitution of PapA5 activity

The ability of the PapA5 protein to form esters using various alcohols was monitored by incubating the reactions at 30°C for 4 hrs. The reaction mixtures included 100 mM phosphate buffer (pH: 7.0), 100 µM $^{14}$C-lauroyl-CoA, 1 mM medium chain alcohol (1-octanol, 9-decenol or 1-dodecanol) and PapA5 protein. The ester products were extracted using chloroform, concentrated and spotted on the silica gel TLC plate. The products were separated by developing the TLC plate using hexanes/ethyl acetate (70:30) as solvent system. Control reactions in the absence of substrate and protein were also simultaneously carried out. The radioactive products were analysed using phosphorimager.

4.2.3 Synthesis of standard esters

1-dodecanoyl dodecanoate ester was synthesized by using lauroyl chloride and 1-dodecanol with the help of Archana in our laboratory. Equimolar concentrations of lauroyl chloride and 1-dodecanol were dissolved in dichloromethane and the reaction was carried out at 0°C for 3 hrs in the presence of triethylamine. 5 ml of water was added to the reaction and the product was then extracted thrice using chloroform. The solvent was dried by using anhydrous sodium sulphate and the product was concentrated using a rotary evaporator. The ester thus obtained was purified using column chromatography and the identity of the compound was confirmed by ESI-MS/MS. Similarly mono- and di-dodecanolesters of 1, 2-dodecanediol were synthesized using two equivalents of lauroyl-chloride in the reaction.
4.2.4 HPLC analysis of ester formation

The PapA5 reaction was scaled up (four times) for HPLC analysis. The product was concentrated after extraction and dissolved in methanol. 20 μg of the chemically synthesized ester was added to enable detection at 215 nm on HPLC. The ester was resolved using an isocratic gradient of methanol/dichloromethane (90:10) on a reverse phase (C5) column. The radioactivity elution profile was monitored by using an online radioactive detector (IN/US system β-RAM Model 3).

4.2.5 Mass spectrometric analysis

The HPLC peak corresponding to the ester was collected and analysed on nanospray ESI-MS/MS in the positive ion mode.

4.2.6 In vitro synthesis of the mycocerosate ester

Using the standard Mas reaction, the mycocerosic acids bound to the Mas protein were synthesized with hexanoyl-CoA as starter unit and 14C- methyl malonyl-CoA as extender unit. After incubation of 8-10 hrs, 1 mM 2-dodecanol and 7 μM PapA5 protein were added and the reactions were further incubated at 30°C. Products were then extracted twice using chloroform and separated on silica gel TLC plate using hexanes/ethyl acetate (70:30) as solvent system. The radioactive products were analysed by phosphorimager.
4.3 RESULTS

4.3.1 Comparison of active site residues of C-domains of NRPS

C-domains of NRPS possess a conserved HHXXXDG active site motif (Finking and Marahiel, 2004; Walsh, 2004). Sequence analysis revealed that this motif was completely conserved in PapA5 as shown in Figure 4.1. In the C-domain of PpsE, the second His and the terminal Gly residues were replaced by Ala and Asp respectively. In the case of vibriobactin synthetase, active site residues of VibH have been investigated using site-directed mutagenesis. While the mutation of the second His residue in the VibH did not drastically affect the catalysis, the H124A mutation in PapA5 led to a substantial decrease in the activity. This indicated that the C-domain of PpsE might not be competent to catalyse the esterification.

CONSERVED MOTIF
PpsE C-DOMAIN
PapA5

Figure 4.1: Comparison of active site residues. Comparison of the active site residues of PapA5 and the C-domain of PpsE with the conserved active site motif of C-domains of NRPS (as show with bold letters)

4.3.2 Expression of PapA5 protein

The PapA5 protein was over expressed as a C-terminal hexa-histidine tagged protein in E. coli. The supernatant containing the protein was loaded onto a Ni\(^{2+}\)-NTA agarose column. The bound protein was eluted at 2, 20, and 50 mM imidazole concentrations (Figure 4.2a). The protein was then further purified to homogeneity using anion exchange chromatography with increasing NaCl
gradient. The protein eluted from the ResourceQ column at 35% B (350 mM NaCl) concentration. Small aliquots of the protein were analysed on 10% SDS-PAGE gel (Figure 4.2b). Fractions 10-13 had maximum amount of protein (lanes 7-11). 2 mM TCEP was added to the protein containing fractions before freezing the protein at -80°C in small aliquots.

**Figure 4.2: Purification of PapA5 protein**
(a) SDS-PAGE analysis of affinity purification of PapA5
1. Supernatant
2. Flow through 8-15. 20 mM imidazole fractions
3. Wash 16. Protein Marker
4-6. 2 mM imidazole wash 17-20. 50 mM imidazole fractions

**Figure 4.2: Purification of PapA5 protein**
(b) Analysis of FPLC purification of PapA5
1. Flow through 8. 35% B Fr 11 15. 45% B Fr 17
2. 15% B Fr 4 9. 38% B Fr 12 16. Protein Marker
3. 20% B Fr 5 10. Protein Marker 17. 47% B Fr 18
4. 30% B Fr 7 11. 38% B Fr 13 18. 50% B Fr 19
5. 30% B Fr 8 12. 40% B Fr 14 19. 50% B Fr 21
6. 30% B Fr 9 13. 40% B Fr 15 20. 90% B Fr 24
7. 35% B Fr 10 14. 45% B Fr 16
4.3.3 Ester bond formation by PapA5

The final step in the biosynthesis of PDIM would require the esterification of mycocerosic acids on to the hydroxyl group of phthiocerol intermediate. Since variant forms of C-domain in NRPS clusters can catalyse the formation of intermolecular amide bond (Keating et al., 2001), the function of two C-like domains (C domain of PpsE protein and PapA5) in catalysing the esterification step was explored. Initial studies were carried out by using substrate mimics. The PapA5 protein was incubated with $^{14}$C-lauroyl-CoA and various alcohols (1-octanol, 9-decenol and 1-dodecanol) and the formation of the ester was monitored by silica gel TLC using 70:30 hexanes/ethyl acetate as solvent system. Analysis of the reaction products using radio-TLC showed formation of new products with higher $R_f$ (0.7-0.75) as compared to $R_f$ of lauric acid (0.26) (Figure 4.3). The product formed with shorter chain alcohol (1-octanol) migrated with lower $R_f$ when compared to the product with longer chain alcohol (1-dodecanol) as observed in lanes 2-4. These radioactive bands were not observed either in the absence of substrate or PapA5 protein (lanes 1 and 5). Similar activity for surrogate substrates was also recently reported for PapA5 protein and based on this ester formation, PapA5 protein was proposed to be involved in the diesterification of phthiocerol (Onwueme et al., 2004). We decided to characterise the product formed by lauroyl-CoA and 1-dodecanol.
Figure 4.3: Reaction catalysed by PapA5. The TLC shows the formation of the ester product using radio-labelled lauroyl-CoA and various alcohols as nucleophiles. No product is observed in the controls without alcohol (lane 1) and PapA5 (lane 5). The $R_f$ of the ester products formed increases with the increasing chain length of various alcohols used (lane 2-4).

Standard 1-dodecanoyl dodecanoate ester was chemically synthesized using dodecanoyl chloride and 1-dodecanol. The PapA5 reaction product was co-spotted with the standard and analysed on the TLC. The TLC analysis showed that the chemical standard comigrated with the radioactive PapA5 product with an $R_f$ of 0.75. This was again confirmed by using reverse phase HPLC analysis by simultaneously detecting the radioactivity as well as UV absorbance at 215 nm. The radioactive detector showed two major peaks corresponding to lauric acid and ester. The huge radioactive peak at 5.8 min corresponded to the $^{14}$C-lauric acid.
and the radioactive ester peak coeluted with the chemical reference peak obtained at 8.5 min (Figure 4.4). The peak obtained at 8.5 min was collected and analysed on ESI-MS/MS in the positive ion mode. The ESI-MS spectra of the ester in the positive ion mode showed a very small m/z peak at 369.37 which corresponded to the [M+H]+ molecular ion (Figure 4.5a). A significant sodium adduct peak was observed at 391.37. Further MS/MS analysis of the 369.37 peak gave a fragment ion at 201.2, which corresponds to the [C_{12}H_{24}O_2+H]^+ ion of the dodecanoic acid (Figure 4.5b). Such a fragmentation pattern for long chain esters have been described previously.

![Figure 4.4: HPLC analysis of PapA5 product.](image)

The product formation with $^{14}$C-lauroyl-CoA and 1-octanol was carried out for 2, 4, 6, 10 hrs at two different protein concentrations. There was an increase in the amount of product formed with increasing time and protein concentration which indicated that the observed reaction was indeed enzyme catalysed (Figure 4.6). The amount of product formed was quantified from the radio-TLC and based on these values, the $k_{cat}$ of 1.12±0.05 pmol mg$^{-1}$ min$^{-1}$ was obtained.
Figure 4.5: ESI-MS analysis of the 1-dodecanoyl dodecanoate ester (a) TOF MS spectrum of the ester in the positive ion mode. The molecular ion peak is observed at 369.37 whereas a prominent peak of the Na⁺ adduct is seen at 391.37. (b) MS/MS spectrum of the molecular ion at 369.37. The fragment ion at 201.20 corresponds to the lauric acid.

Figure 4.6: Kinetics analysis of the ester formation. Time-dependent formation of the 1-octyl dodecanoate ester catalysed by PapA5 at two different concentrations.
4.3.4 Hand-to-hand transfer of protein-bound intermediates to synthesize PDIM

Since our analysis suggested that the mycocerosic acids were not actively released from the Mas protein, it was debatable whether the mycocerosic acids were first released and activated as acyl-CoA thioesters or the enzyme bound products were directly transferred on to the alcohol. We developed a TLC-based assay to probe whether the mycocerosic acids covalently bound to the protein could be directly transesterified onto the alcohols. After carrying out the Mas reactions, the acylated Mas protein was incubated with PapA5/C-domain of PpsE and various alcohols (1-dodecanol, 2-dodecanol and 1, 2-dodecanediol) to monitor the formation of the ester. Once again, the C-domain of PpsE failed to show any activity under the experimental conditions used. Analysis of the PapA5 products on the radio-TLC indicated formation of products with higher R_f (Figure 4.7). 2-dodecanol was used as a nucleophile since it was a closer mimic of the actual substrate phthiocerol and 1, 2-dodecanediol to monitor the formation of diester. In the reactions with 1-dodecanol and 2-dodecanol, a radioactive product with an R_f of 0.8 was obtained (Figure 4.7, lanes 1 and 2). The Mas products migrated on the TLC with an R_f of 0.34. The control reaction carried out in the absence of the acceptor alcohol did not show any product formation (lane 4). In the case of the 1,2-dodecanediol, two new products could be observed on the TLC plate. The major product migrated with an R_f of 0.5 and a very faint band with R_f of 0.85 was observed. In order to confirm these products, we synthesized the mono- and the di-esters of 1,2-dodecanediol using lauroyl chloride. The mono- and the di-esters were purified by silica gel chromatography and confirmed by ESI-MS/MS.
Figure 4.7: PapA5 catalyzed transesterification of mycocerosic acids. The PapA5 catalyses direct transfer of covalently acylated mycocerosic acids on to the alcohol to form the mycocerosate ester. Lane 1-3 on the radio-TLC shows the formation of the ester with different alcohols. The ester is not formed in the absence of alcohol (lane 4).

The TOF MS spectrum obtained for the mono-ester is shown in Figure 4.8a. The molecular ion [M+H]$^{+1}$ at 385.38 was obtained in the positive ion mode. A major peak was obtained at 367.38 amu in the MS/MS spectra of 385.38, which corresponds to the loss of water molecule characteristic of the presence of the hydroxyl moiety (Figure 4.8b). The m/z peaks at 201.22 and 183.21 would correspond to the molecular ions formed from the lauric acid moiety and its dehydration product. The [M+H]$^{+1}$ peak at 567.51 was obtained from TOF-MS
spectrum of the diester (Figure 4.9a) and the MS/MS spectrum showed the presence of the 367.37, 201.23 and 183.23 peaks as observed earlier for the monoester (Figure 4.9b). The mono- and the di-ester were then spotted along with the radioactive products and it was confirmed that the major product of $R_f$ -0.5 corresponded with the lauroyl mono-ester of 1,2-dodecanediol.

![Graph](image)

**Figure 4.8:** ESI-MS analysis of the mono acyl ester of 1,2-dodecanediol  (a) TOF MS spectrum of monoester of the diol. The molecular ion peak is observed at 385.38. (b) MS/MS spectrum of the molecular ion of 385.38. The fragment ion at 367.38 is formed due to the loss of water molecule while the 201.22 fragment ion corresponds to the lauric acid ion.
Figure 4.9: ESI-MS analysis of the diacyl ester of 1,2-dodecanediol (a) TOF MS spectrum of diester of the diol. The molecular ion peak is observed at 567.51. The signal at 585.51 corresponds to the Na$^{+1}$-adduct of the diester (b) MS/MS spectrum of the molecular ion of 567.51. The fragment ion at 367.37 is formed due to the loss of an acyl group while the 201.25 fragment ion corresponds to the lauric acid ion.

4.3.5 Determination of kinetic parameters for PapA5 catalysed ester formation

The formation of ester by PapA5 using acylated Mas protein and 2-dodecanol was monitored over a period of 4 hrs. The Mas reactions were carried out by using hexanoyl-CoA and $^{14}$C-methyl malonyl-CoA at 30°C. PapA5 protein and 2-dodecanol were then added to each reaction and further incubated for various time intervals at 30°C. The product was extracted with chloroform and analysed on radio-TLC. The amount of product formed was quantified by using
phosphorimager. The $v$ vs. $t$ curve is shown in Figure 4.10a. The saturation was obtained at 3 hrs. The steady state kinetic parameters showed good catalytic rates and obeyed saturation kinetics with apparent kinetic parameters of $k_{cat} = 11.74 \text{ pmol.mg}^{-1}\text{min}^{-1}$ and $K_m$ of 23.97 $\mu$M for 2-dodecanol (Figure 4.10b). Similar experiment for the dependence of the rate of the ester formation with the mycocerosic acids bound to the Mas protein showed complex kinetic profile (Figure 4.10c). The maximum catalytic activity was observed at a concentration of 3.7 $\mu$M of acylated Mas protein and dropped thereafter. The reaction velocity at 3.7 $\mu$M corresponds to the apparent $k_{cat}$ of 13.21 pmol.mg$^{-1}\text{min}^{-1}$ and an apparent $K_m$ of 0.39 $\mu$M. The inhibition effects observed beyond 4.0 $\mu$M might have arisen
the presence of unacylated Mas proteins, which would compete with its acylated counterpart for binding to the PapA5. In order to demonstrate this, PapA5 protein was first incubated with unacylated Mas protein and then incubated with the mycocerosic acids bound to the Mas protein. Figure 4.11a shows the radio-TLC analysis of the formation of the mycocerosate ester product with increasing concentration of the unacylated Mas protein. The percentage change in activity is plotted in Figure 4.11b and it can be observed that presence of 1:1 ratio of the unacylated to acylated Mas protein resulted in loss of more than 50% of the ester formation. Similar inhibition with apo-ACP has been previously demonstrated with actinorhodin polyketide biosynthesis (ref).

![Figure 4.11: Inhibition of mycocerosyl ester formation by unacylated Mas protein](image)

(a) TLC indicating decrease in ester formation with addition of unacylated Mas protein (b) Amount of ester formed in presence of unacylated protein as a percentage of the total product formed

### 4.3.6 Interaction of the Mas ACP domain with PapA5

Biochemical studies indicated that PapA5 is involved in the transfer of ACP bound mycocerosic acids on to the phthiocerol moiety by catalysing formation of an ester bond. Protein-protein interactions between ACP and PapA5 are likely to play an important role in this enzymatic transfer of mycocerosic acids.
Computational analysis involving docking simulations was carried out by Zeeshan in Dr. Mohanty’s group to identify putative residues mediating recognition between Mas ACP and PapA5. Although the Mas ACP did not have significant sequence similarity with the available 3D structures of ACP/PCP domains, the PCP domain of tyrocidine synthetase (1DNY) (Weber et al., 2000) showed maximum similarity as evaluated by threading analysis. Based on threading alignment, a structural model was built for Mas ACP using the MODELLER package. The modelled Mas ACP was docked on the crystal structure of PapA5 (1Q9J) (Buglino et al., 2004) using FTDOCK program (Jackson et al., 1998). All 10,000 ACP-PapA5 complexes were ranked using a residue based statistical potential given by the program RPscore. A functional constraint was set up such that the P-pant group attached to the catalytic Ser of ACP should reach the catalytic His of PapA5 through the hydrophilic channel of PapA5. The top ranking complexes were analysed to identify docked complexes consistent with the functional constraint. One ACP-PapA5 complex was consistent with the functional constraint and was found in the top 1% solutions given by the FTDOCK. The three regions of the ACP that interact with PapA5 in this complex (Figure 4.12) were residues 23-25 (immediately after helix 1), residues 37-41 (in the loop region between helix 1 and 2) and residues 66-67 (after helix 3). Similar ACP interacting residues have been previously documented in other studies (Finking et al., 2004; Keatinge-Clay et al., 2003). The contacting residues from PapA5 involved three Arg (231, 234 and 312) and six other hydrophobic residues (Figure 4.12). While most of the contacts between ACP and PapA5 were hydrophobic, one of the Arg312 was observed to be involved in electrostatic
interactions with a Glu37 of ACP. The two other Arg (231, 234) residues had their charged ends exposed to the solvent and their non-polar groups made contacts with hydrophobic residues of ACP. The conformation of the ACP bound P-pant arm inside the PapA5 channel was modelled by superposing the crystal structure of carnitine acetyl transferase (1NDI) on PapA5 and the co-ordinates of CoASH were transformed from 1NDI (Jogl and Tong, 2003). In the energy minimized model of ACP-PapA5 complex, the thiol group of ACP bound P-pant arm was within 3.7 Å of the catalytic His residue. Mutational analysis of these residues can be carried out to validate this model.

Figure 4.12: Interaction of the PapA5 protein with the ACP domain of Mas. A computational model indicating the interaction of the PapA5 protein with the Mas ACP domain was obtained by docking the 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.
4.4 DISCUSSION

The wax ester synthases catalyse the ester formation by condensation of long chain acyl-CoAs with long chain alcohols (Kalscheuer and Steinbuchel, 2003; Routaboul et al., 1999). Similar mechanism of condensation of the mycocerosic acids to the phthiocerol is necessary for the complete synthesis of PDIM. Our studies demonstrated that the PapA5 could transfer the mycocerosic acids from the Mas directly to the alcohol to form the ester. Similar O-acyl transferase activity has been reported in the lipid A biosynthesis in *E. coli*. The LpxA protein catalyses the esterification of the R-3-hydroxy myristoyl chain from the R-3-hydroxy myristoyl-ACP directly on to the UDP-N-Acetyl glucosamine (UDP-GlcNAc) (Anderson et al., 1985; Anderson and Raetz, 1987). The LpxA protein is highly specific for R-3-hydroxy myristoyl-ACP and does not accept R-3-hydroxy myristoyl-CoA or myristoyl-ACP (Andersen et al., 1991). Unlike the LpxA, the PapA5 however uses a wide range of acyl-CoAs for transfer on to the various alcohols. The ability of the PapA5 to transfer the mycocerosic acids directly from the Mas protein also explains the absence of any release mechanism of the final product from the Mas protein. This mode of hand-to-hand transfer may be the most efficient method to channelize intermediates, since long-chain methylated mycocerosic acids or their CoA-esters can interfere with lipid metabolism. The other Pap homologues in *M. tuberculosis* are also present adjacent to genes involved in lipid metabolism. It is plausible that these Pap proteins would probably condense individual components through a similar mechanism.