Two polyketide-synthase-associated acyltransferases are required for sulfolipid biosynthesis in *Mycobacterium tuberculosis*

Kiranmai Bhatt,¹ Sudagar S. Gurcha,² Apoorva Bhatt,¹ Gurdyal S. Besra² and William R. Jacobs, Jr¹

¹Howard Hughes Medical Institute, Department of Microbiology and Immunology, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, USA
²School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

The methyl-branched fatty acyl components of sulfolipid-I (SL-I), a major glycolipid of the human pathogen *Mycobacterium tuberculosis*, are synthesized by the polyketide synthase Pks2. Rv3824c (papA1), located downstream of pks2, encodes a protein that belongs to a subfamily of acyltransferases associated with mycobacterial polyketide synthases [polyketide synthase-associated proteins (PAPs)]. The presence of a conserved acyltransferase motif (HX₃DX₄₁₄Y) suggested a role for PapA1 in acylation of sulfated trehalose to form SL-I. Targeted deletion of the H37Rv papA1 resulted in loss of SL-I, demonstrating its role in mycobacterial sulfolipid biosynthesis. Furthermore, SL-I synthesis was restored in the mutant strain following complementation with papA1, but not with mutant alleles of papA1 containing alterations of key residues in the acyltransferase motif, confirming that PapA1 was an acyltransferase. While other *M. tuberculosis* pks clusters are associated with a single PAP-encoding gene, it was demonstrated that another open reading frame, Rv3820c (papA2), located 5.8 kb downstream of papA1 is also an acyltransferase gene involved in SL-I biosynthesis: deletion of papA2 abolished SL-I production. The absence of any partially acylated intermediates in either null mutant indicated that both PapA1 and PapA2 were required for all acylation steps of SL-I assembly.

**INTRODUCTION**

*Mycobacterium tuberculosis*, the causative agent of tuberculosis, has a distinct, lipid-rich cell wall that plays an important role in virulence (Brennan & Nikaido, 1995; Puzo, 1990). Sulfolipids are one of the major *M. tuberculosis* cell wall lipids, of which sulfolipid-I (SL-I) is the most abundant (Goren, 1970a, b; Goren et al., 1976; Middlebrook et al., 1959). The presence of sulfolipids in virulent strains led to the speculation that this class of polar lipids plays an important role in pathogenesis. However, there have been mixed reports about the role of sulfolipids in virulence, with various modulatory effects being observed with purified sulfolipids (Okamoto et al., 2006; Pabst et al., 1988; Zhang et al., 1988), but no attenuation observed with mutant strains in mouse and guinea pig models of infection (Rousseau et al., 2003). SL-I consists of a tetra-acylated trehalose sulfate (Fig. 1): of the four acyl groups that are esterified to trehalose, one is palmitate while the remaining three are complex methyl-branched fatty acids, termed phthioceranic and hydroxyphthioceranic acids (Goren et al., 1971). The sulfation of trehalose is presumed to be the first step in SL-I assembly and is catalysed by the sulfotransferase Stf0 (Mougous et al., 2004). The palmitate is fatty acid synthase-I (FAS-I)-derived, the methyl-branched moieties, phthioceranic and hydroxyphthioceranic acids, are synthesized by a polyketide synthase, Pks2 (Sirakova et al., 2001). Present in the same cluster as pks2 is mmpL8, a gene encoding a transmembrane protein involved in transport (Converse et al., 2003; Domenech et al., 2004). Though the genes involved in the biosynthesis of individual components of SL-I are well studied, not much is known about the assembly of SL-I, vis-à-vis the components and sequence of events involved in acylation of trehalose-2-sulfate to SL-I. Located immediately downstream of pks2 is Rv3824c (papA1; Fig. 2a), which encodes a protein that belongs to a subfamily of acyltransferases associated with mycobacterial polyketide synthases [polyketide synthase-associated proteins (PAPs)]. The presence of a characteristic HX₃DX₄₁₄Y acyltransferase motif suggested a role for PapA1 in the acylation steps of SL-I assembly. Indeed, PapA5, another PAP, has been shown to be an acyltransferase involved in the esterification of phthiocerol with methyl-branched mycocerosic fatty acids to form the virulence lipid phthiocerol dimycocerosate (PDIM) (Onwueme et al., 2004;
Trivedi et al., 2005). In general, pks clusters in *M. tuberculosis* are associated with only one PAP-encoding gene (Cole et al., 1998). Interestingly, Rv3820c (papA2), a gene located 5.8 kb downstream of papA1 (Fig. 2a), also encodes a protein with an acyltransferase motif, indicating that SL-I biosynthesis may involve more than one PAP. A similar arrangement of homologues of pks2, papA1 and papA2 is also seen in the genome of the bovine pathogen *Mycobacterium bovis* (Garnier et al., 2003), although the *M. bovis*-derived vaccine strain BCG does not synthesize SL-I (Rivera-Marrero et al., 2002). In these studies we have used a genetic approach to investigate whether papA1 and papA2 play a role in SL-I biosynthesis. Individual null mutants of H37Rv papA1 and papA2 failed to produce SL-I. However, no intermediates were observed in either of the mutants, indicating an essential role for each acyltransferase in assembly of complete SL-I.

**METHODS**

**Plasmids, phages and culture conditions.** Plasmids and phages used in this study are listed in Table 1. *Mycobacterium tuberculosis* H37Rv was cultured in 7H9 broth (Difco) containing 10% Middlebrook OADC enrichment and 0.05% Tween 80, or on Middlebrook 7H10 agar (Difco); *Escherichia coli* strains were cultured in LB broth. The concentrations of antibiotics used were 75 μg hygromycin ml⁻¹ and 20 μg kanamycin ml⁻¹ for H37Rv, and 150 μg hygromycin ml⁻¹ and 40 μg kanamycin ml⁻¹ for *E. coli*.

**Construction of null mutants.** For generating an allelic-exchange construct designed to replace the papA1 gene with a hygromycin resistance cassette (hyg), ~1 kb sequences flanking the left and right of the *M. tuberculosis* H37Rv papA1 gene were PCR amplified from genomic DNA using the primer pairs papA1L1 (5'-CGGACTAGTGGTCTCAGCATGCAACG-3')/papA1R1 (5'-AGGATCCGAGCTCATCTGCGTGAAC-3') and papA2L1 (5'-TGGTCTCAGCATGCAACG-3')/papA2R1 (5'-AGGATCCGAGCTCATCTGCGTGAAC-3'). Following cloning into *E. coli* plasmid pCR2.1-TOPO and sequencing, the cloned PCR fragments were excised using the primer-introduced restriction sites excising the hygromycin gene from H37Rv genomic DNA. Following subcloning and sequence confirmation, the PCR product was excised and cloned as a HindIII–Clal fragment into pMV361 to yield the complementing plasmid pMV361:papA2, which was then electroporated into ΔpapA2 to yield the complemented strain ΔpapA2::pMV361:papA2.

**Complementation of mutant strains.** The primers papA1-F (5'-CATGACGGATCTGGTCAATAAGGC-3') and papA1-R (5'-GGCTGATTACCTAAGCTTCATCTC-3') were used to PCR amplify the papA1 gene from H37Rv genomic DNA. Following cloning into pCR2.1-TOPO and sequencing, the cloned PCR fragment was excised by digestion of the primer-introduced EcoRI and HpaI sites and cloned downstream of the hsp60 promoter in the integrative vector pMV361 (Stover et al., 1991). The resultant plasmid pMV361:papA1 was introduced into ΔpapA1 by electroporation to yield the complemented strain ΔpapA1::pMV361:papA1. Similarly, the primers papA2-F (5'-CTCAAGCTTTGGTATCAGCTACCA-3') and papA2-R (5'-CGATCGATTATCTGCGTGAAG-3') were used to PCR amplify the papA2 gene. Following subcloning and sequence confirmation, the PCR product was excised and cloned as a HindIII–Clal fragment into pMV361 to yield the complementing plasmid pMV361:papA2, which was then electroporated into ΔpapA2 to yield the complemented strain ΔpapA2::pMV361:papA2.
Biochemical analysis of mutant strains. Labelling of *M. tuberculosis* cultures with $^{14}$Cacetate, $^{14}$Cpropionate or Na$_2^{35}$SO$_4$ was done as described previously (Sirakova et al., 2001). Total lipids were extracted from strains using previously described protocols (Dobson et al., 1985) and separated on a TLC plate using chloroform/methanol/water (100:14:0.8, by vol.) in the first dimension and chloroform/acetone/methanol/water (50:60:2.5:3, by vol.) in the second dimension. The TLC plates were exposed to X-ray films for 24 h.

RESULTS

Generation of *papA1* and *papA2* null mutants of *M. tuberculosis*

We used specialized transduction (Bardarov et al., 2002) to replace *papA1* in the *M. tuberculosis* H37Rv genome with a hygromycin resistance marker (hyg) (Fig. 2a). Genomic DNA obtained from Hyg$^R$ *M. tuberculosis* H37Rv colonies following transduction with phAE405 was digested with KpnI and analysed by Southern blotting (Fig. 2b). The band pattern obtained confirmed that *papA1* was replaced by hyg. Similarly, the knockout phage phAE406 was used to obtain a *papA2* null mutant of H37Rv (Fig. 2c). The *papA1* and *papA2* null mutants are referred to as Δ*papA1* and Δ*papA2* respectively.

**SL-I is not synthesized in Δ*papA1***

To determine changes in the lipid profiles of the Δ*papA1* mutant, we first monitored differences in the $^{14}$Cacetate-labelled lipids from wild-type and Δ*papA1* strains by TLC. In a solvent system designed to separate *M. tuberculosis* glycolipids (Dobson et al., 1985), a $^{14}$C-labelled lipid, which was clearly present in the wild-type but missing in the mutant strain, migrated in a manner similar to an authentic SL-I standard (Fig. 3a). Three of the four acyl components of SL-I are methyl-branched fatty acids (phthioceranate and hydroxyphthioceranate) which can be labelled using $^{14}$Cpropionate. Indeed, the lipid missing in Δ*papA1*, but present in WT showed high incorporation of $^{14}$C label when propionate was used as the radiolabel (Fig. 3b). Finally, this lipid also showed incorporation of $^{35}$S from Na$_2^{35}$SO$_4$, confirming that the species was indeed SL-I (Fig. 3b). Introduction of pMV361*papA1*, a single copy integrative vector containing *papA1* cloned downstream of
the hsp60 promoter, into ΔpapA1 fully restored SL-I biosynthesis, indicating that the observed phenotype in ΔpapA1 was solely due to deletion of papA1 (Fig. 3a, b). These results demonstrated that papA1 function was necessary for SL-I biosynthesis.

**PapA2 is also required for SL-I biosynthesis**

To investigate the role of papA2 in SL-I biosynthesis, an approach similar to that used for the ΔpapA1 strain was followed. An initial analysis of total lipids extracted from cultures of H37Rv and ΔpapA2 labelled with [14C]acetate indicated that SL-I was missing in the ΔpapA2 strain (Fig. 4a). This was then confirmed using [15N]propionate or Na235SO4 as the labelling reagents (Fig. 4b). SL-I biosynthesis could be restored in ΔpapA2 following complementation with pMV361papA2, a single-copy integrative vector containing papA2 cloned downstream of the hsp60 promoter (Fig. 4a, b). Thus, like papA1, papA2 was also essential for SL-I biosynthesis.

**papA1 and papA2 encode acyltransferases**

The presence of a conserved acyltransferase motif (HX3DX14Y) had suggested a role for PapA1 and PapA2 in acylation of either trehalose or sulfated trehalose, to form SL-I. The histidine and aspartate residues of the motif are considered essential for acyltransferase activity (Bergendahl et al., 2002; Lewendon et al., 1988). To demonstrate that papA1 and papA2 do encode acyltransferases, we tested the ability of mutant alleles of the genes containing alterations in either of these key residues to restore SL-I biosynthesis in the mutant strain. First, single-copy-integrative plasmids containing two mutant alleles of papA1 containing separate point mutations in the HX3DX14Y motif were constructed using site-directed mutagenesis (Fig. 5). The mutant constructs pMV361A1-DA (Asp-175→Ala-175) and pMV361A1-HA (His-171→Ala-171) were then introduced separately into ΔpapA1 by electroporation and the resultant strains, ΔpapA1::pMV361A1-HA and ΔpapA1::pMV361A1-DA, were tested for functional complementation. Analysis of [14C]propionate- or 35S-labelled total lipids from the transformed strains revealed that neither construct reconstituted SL-I biosynthesis in ΔpapA1 (Fig. 3b). Similarly, the plasmids pMV361A2-HA (His-166→Ala-166) and pMV361A2-DA (Asp-170→Ala-170) failed to restore SL-I biosynthesis in ΔpapA2 (Fig. 4b). Thus, the key residues in the putative acyltransferase motifs of PapA1 and PapA2 were essential for functional complementation of the corresponding SL-I deficient mutants, indicating that PapA1 and PapA2 are acyltransferases involved in SL-I biosynthesis.

**DISCUSSION**

The sequencing of the *M. tuberculosis* H37Rv genome revealed five open reading frames, *papA1–papA5*, that were annotated as genes for polyketide synthase-associated proteins (PAPs) because of their linkage to genes encoding enzymes involved in polyketide biosynthesis (Cole et al., 1998). Based on the presence of an acyltransferase motif, it was speculated that PAPs may be acyltransferases involved in the assembly of complex lipid metabolites synthesized by the associated *pk*s cluster. Experimental proof for this came first from *in vitro* experiments using purified PapA5, a PAP

---

**Table 1. Plasmids and phages**

<table>
<thead>
<tr>
<th>Plasmid or phage</th>
<th>Description</th>
<th>Reference/source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCR2.1-TOPO</td>
<td>PCR product cloning vector</td>
<td>Invitrogen Life Technologies</td>
</tr>
<tr>
<td>pJSC347</td>
<td>Vector for cloning allelic-exchange substrates to be used for specialized transduction; contains λ phage cos site and HygR marker</td>
<td>Sambandamurthy et al. (2002)</td>
</tr>
<tr>
<td>pYUB2418</td>
<td>Derivative of pJSC347 designed for allelic exchange of <em>M. tuberculosis</em> papA1</td>
<td>This work</td>
</tr>
<tr>
<td>pYUB2419</td>
<td>Derivative of pJSC347 designed for allelic exchange of <em>M. tuberculosis</em> papA2</td>
<td>This work</td>
</tr>
<tr>
<td>pMV361A1</td>
<td>Single-copy mycobacterial integrative vector with hsp60 promoter</td>
<td>Stover et al. (1991)</td>
</tr>
<tr>
<td>pMV361papA1</td>
<td><em>M. tuberculosis</em> papA1 cloned in pMV361</td>
<td>This work</td>
</tr>
<tr>
<td>pMV361A1-HA</td>
<td>His-171→Ala-171 mutant of papA1 in pMV361</td>
<td>This work</td>
</tr>
<tr>
<td>pMV361A1-DA</td>
<td>Asp-175→Ala-175 mutant of papA1 in pMV361</td>
<td>This work</td>
</tr>
<tr>
<td>pMV361papA2</td>
<td><em>M. tuberculosis</em> papA2 cloned in pMV361</td>
<td>This work</td>
</tr>
<tr>
<td>pMV361A2-HA</td>
<td>His-166→Ala-166 mutant of papA2 in pMV361</td>
<td>This work</td>
</tr>
<tr>
<td>pMV361A2-DA</td>
<td>Asp-170→Ala-170 mutant of papA2 in pMV361</td>
<td>This work</td>
</tr>
<tr>
<td>Phages</td>
<td></td>
<td></td>
</tr>
<tr>
<td>phAE159</td>
<td>Conditionally replicating shuttle phasmid derived from the lytic mycobacteriophage TM4</td>
<td>Gift from J. Kriakov*</td>
</tr>
<tr>
<td>phAE405</td>
<td>Derivative of phAE159 obtained by cloning pYUB2418 in its unique PacI site</td>
<td>This work</td>
</tr>
<tr>
<td>phAE406</td>
<td>Derivative of phAE159 obtained by cloning pYUB2419 in its unique PacI site</td>
<td>This work</td>
</tr>
</tbody>
</table>

*Albert Einstein College of Medicine.*

---

The sequencing of the *M. tuberculosis* H37Rv genome revealed five open reading frames, *papA1–papA5*, that were annotated as genes for polyketide synthase-associated proteins (PAPs) because of their linkage to genes encoding enzymes involved in polyketide biosynthesis (Cole et al., 1998). Based on the presence of an acyltransferase motif, it was speculated that PAPs may be acyltransferases involved in the assembly of complex lipid metabolites synthesized by the associated *pk*s cluster. Experimental proof for this came first from *in vitro* experiments using purified PapA5, a PAP.
involved in PDIM biosynthesis (Onwueme et al., 2004; Trivedi et al., 2005). PDIM is made up of two mycocerosic acid residues esterified to the hydroxyl groups of a diol (phthiocerol). Using purified proteins and artificial substrates, Onwueme et al. (2004) first showed that PapA5 could catalyse esterification of an alcohol with an acyl group. Trivedi et al. (2005) then demonstrated that purified PapA5 catalysed trans-esterification of mycocerosic acid synthase-bound mycocerosic acids to dodecanol.

Of all the pks genes linked to PAPs, pks2 is the only one located next to two pap genes (papA3 and pks3/4 are involved in the biosynthesis of polyacylated trehalose, and the metabolite synthesized by papA4 and pks5 has not yet been identified). Deletion of pks2 resulted in loss of the principal sulfolipid SL-I in M. tuberculosis H37Rv (Sirakova et al., 2001). The absence of SL-I in ΔpapA1 and in ΔpapA2, the two mutants generated in this study, demonstrated that PapA1 and PapA2 were indeed functionally associated with Pks2. Phthioceranic acid and hydroxyphthioceranic acid, the methyl-branched products of Pks2, are found esterified to positions 3, and 6 and 6′, respectively, of sulfated trehalose in the principal sulfolipid SL-I, while position 2 is substituted with FAS-I-derived palmitate. It was thus tempting to speculate that PapA1 and PapA2 may differ in their substrate specificities, with one PAP involved in the transfer of the methyl-branched acyl chains, and the other in that of palmitate. Another possibility was that one of the pks2-associated PAPs was specific for transfer of phthioceranic acid and the other for hydroxyphthioceranic acid, with

![Fig. 3. Sulfolipid analysis of the ΔpapA1 mutant by 2D-TLC. X-ray film exposed to TLC plates following 2D-TLC of total lipid samples extracted from cultures labelled with (a) [14C]acetate or (b) [14C]propionate or Na235SO4. The spot corresponding to SL-I is indicated in the images. GMM, glucose monomycolate; DAT, diacyltrehalose; TMM, trehalose monomycolate; TDM, trehalose dimycolate.](http://mic.sgmjournals.org)

http://mic.sgmjournals.org
the transfer of palmitate being carried out by an unknown acyltransferase. In either case, if the acylation reactions were sequential, partially acylated trehalose sulfate intermediates would be expected to accumulate in either ΔpapA1 or ΔpapA2. While both ΔpapA1 and ΔpapA2 did not synthesize SL-I, neither mutant strain showed accumulation of any partially acylated intermediate. This result indicated that in one of the mutants, feedback inhibition of the initial step may have prevented the accumulation of the intermediate. Alternatively, instead of a sequential addition of different acyl groups to sulfated trehalose by PapA1 and PapA2, concurrent presence of both proteins may be required for complete assembly of a tetracylated trehalose sulfate molecule. Some indication that this might be a possibility comes from recent work on the components of the PDIM biosynthetic cluster. Jain & Cox (2005) showed that polyketide synthase PpsE directly interacted with the transmembrane protein and PDIM transporter MmpL7. This suggested that not only was MmpL7 involved in transport, but that it also probably acted as a scaffold in a tightly integrated system that couples the biosynthesis of PDIM to its secretion from the bacterial cell. In light of these findings it is plausible that MmpL8, the transporter associated with the pks2 cluster, may play a similar ‘scaffolding’ role in SL-I assembly through interactions with Pks2, PapA1 and Pap2, wherein PapA1 and PapA2 may carry out localized and simultaneous acylation reactions. In the ΔpapA1 or ΔpapA2 mutant, loss of one PAP would result in non-functional SL-I assembly machinery and complete loss of SL-I biosynthesis (rather than accumulation of a partially acylated intermediate). However, two independent reports of a M. tuberculosis mmpL8 null mutant showed that

**Fig. 4.** Sulfolipid analysis of the ΔpapA2 mutant by 2D-TLC. X-ray film exposed to TLC plates following 2D-TLC of total lipid samples extracted from cultures labelled with (a) [14C]acetate or (b) [14C]propionate or Na235SO4. The spot corresponding to SL-I is indicated in the images. GMM, glucose monomycolate; DAT, diacyltrehalose; TMM, trehalose monomycolate; TDM, trehalose dimycolate.
PapA1  \(\text{HLHADGQFVGVLMEFQSMY}\)
A1-\(\text{HA}\)  \(\text{HLHADGQFVGVLMEFQSMY}\)
A1-\(\text{DA}\)  \(\text{HLHADGQFVGVLMEFQSMY}\)

PapA2  \(\text{HLCDPDTIVGVLFIEIHHXY}\)
A2-\(\text{HA}\)  \(\text{HLCDPDTIVGVLFIEIHHXY}\)
A2-\(\text{DA}\)  \(\text{HLCDPDTIVGVLFIEIHHXY}\)

Motif consensus \(\text{HXXXDXXXXXXXXXX}\)

**Fig. 5.** Mutation of the HX_2DX_4Y motif in PapA1 and PapA2. Alignment of the amino acid sequences in the HX_2DX_4Y motif of PapA1, PapA2 and the putative proteins encoded by the mutant constructs pMV361A1-\(\text{HA}\) (A1-\(\text{HA}\)), pMV361A1-\(\text{DA}\) (A1-\(\text{DA}\)), pMV361A2-\(\text{HA}\) (A2-\(\text{HA}\)), pMV361A2-\(\text{DA}\) (A2-\(\text{DA}\)). The conserved amino acids are indicated in bold and the altered residues are underlined.

the loss of MmpL8 resulted in the accumulation of a diacylated trehalose sulfate intermediate (Converse et al., 2003; Domenech et al., 2004). Nevertheless, we did not detect accumulation of such an intermediate in either the \(\Delta\text{papA1}\) or \(\Delta\text{papA2}\) mutant. Though the absence of any partially acylated intermediates in either \(\Delta\text{papA1}\) or \(\Delta\text{papA2}\) made it difficult to assign putative substrate specificities (if any) to PapA1 and PapA2, our results have clearly demonstrated that these two distinct acyltransferases are involved in the biosynthesis of the major sulfolipid SL-1 in *M. tuberculosis*.

**ACKNOWLEDGEMENTS**

The authors would like to thank Nagatoshi Fujiwara for helpful suggestions on sulfolipid analysis and Catherine Vilchèze for comments on the manuscript. This work was supported by grants from NIH AI26170 and CFAR (AI 051519) to W.R.J.; G.S.B. acknowledges support from Mr James Bardrick in the form of a Personal Research Chair, and as a former Lister Institute Jenner Research Fellow, and of the Medical Research Council (UK).

**REFERENCES**


Edited by: S. V. Gordon