Identification of apolipoprotein n-acyltransferase (LNT) in mycobacteria

Tschumi, A; Nai, C; Auchli, Y; Hunziker, P; Gehrig, P; Keller, P; Grau, T; Sander, P
Identification of apolipoprotein n-acyltransferase (LNT) in mycobacteria

Abstract

Lipoproteins of Gram-negative and Gram-positive bacteria carry a thioether bound diacylglycerol but differ by a fatty acid amide-bound to the alpha-amino group of the universally conserved cysteine. In Escherichia coli the N-terminal acylation is catalyzed by the N-acyltransferase Lnt. Using E. coli Lnt as a query in a BLASTp search, we identified putative lnt genes also in Gram-positive mycobacteria. The Mycobacterium tuberculosis lipoprotein LppX, heterologously expressed in Mycobacterium smegmatis, was N-acylated at the N-terminal cysteine, whereas LppX expressed in a M. smegmatis lnt::aph knock-out mutant was accessible for N-terminal sequencing. Western blot analyses of a truncated and tagged form of LppX indicated a smaller size of about 0.3 kDa in the lnt::aph mutant compared to the parental strain. MALDI-TOF/TOF analyses of a trypsin digest of LppX proved the presence of the diacylglycerol modification in both strains, parental strain and lnt::aph mutant. N-acylation was found exclusively in the M. smegmatis parental strain. Complementation of the lnt::aph mutant with M. tuberculosis ppm1 restored N-acylation. The substrate for N-acylation is a C16 fatty acid while the two fatty acids of the diacylglycerol residue were identified as C16 and C19:0 fatty acid, the latter most likely tuberculostearic acid. We demonstrate that mycobacterial lipoproteins are triacylated. For the first time to our knowledge, we identify Lnt activity in Gram-positive bacteria and assigned the responsible genes. In M. smegmatis and M. tuberculosis the open reading frames are annotated as MSMEG_3860 and M. tuberculosis ppm1, respectively.
IDENTIFICATION OF APOLIPOPROTEIN N-ACYLTRANSFERASE (LNT) IN MYCOBACTERIA

Andreas Tschumi¹, Corrado Nai¹, Yolanda Auchli², Peter Hunziker², Peter Gehrig², Peter Keller¹, Thomas Grau¹, Peter Sander¹,³

¹Institut für Medizinische Mikrobiologie, Universität Zürich, Gloriastrasse 30/32, CH-8006 Zurich, Switzerland, ²Functional Genomics Center Zurich, Zurich, Switzerland, ³Nationales Zentrum für Mykobakterien, Gloriastrasse 30, CH-8006 Zurich, Switzerland

Address correspondence to: P. Sander, Gloriastr. 30/32, CH-8006 Zurich, Switzerland, Phone: +41 44 634 2684, Fax: +41 44 634 4906, E-mail: psander@imm.uzh.ch

Lipoproteins of Gram-negative and Gram-positive bacteria carry a thioether bound diacylglyceryl but differ by a fatty acid amide-bound to the α-amino group of the universally conserved cysteine. In Escherichia coli the N-terminal acylation is catalyzed by the N-acyltransferase Lnt. Using E. coli Lnt as a query in a BLASTp search, we identified putative lnt genes also in Gram-positive mycobacteria. The Mycobacterium tuberculosis lipoprotein LppX, heterologously expressed in Mycobacterium smegmatis, was N-acylated at the N-terminal cysteine, whereas LppX expressed in a M. smegmatis lnt::aph knock-out mutant was accessible for N-terminal sequencing. Western blot analyses of a truncated and tagged form of LppX indicated a smaller size of about 0.3 kDa in the lnt::aph mutant compared to the parental strain. MALDI-TOF/TOF analyses of a trypsin digest of LppX proved the presence of the diacylglyceryl modification in both strains, parental strain and lnt::aph mutant. N-acylation was found exclusively in the M. smegmatis parental strain. Complementation of the lnt::aph mutant with M. tuberculosis ppm1 restored N-acylation. The substrate for N-acylation is a C16 fatty acid while the two fatty acids of the diacylglyceryl residue were identified as C16 and C19:0 fatty acid, the latter most likely tuberculostearic acid. We demonstrate that mycobacterial lipoproteins are triacylated. For the first time to our knowledge, we identify Lnt activity in Gram-positive bacteria and assigned the responsible genes. In M. smegmatis and M. tuberculosis the open reading frames are annotated as MSMEG_3860 and M. tuberculosis ppm1, respectively.

INTRODUCTION

Proteins of various organisms are modified in numerous ways, one of them is lipidation. Lipid modification of proteins is common in eucaryal and bacterial organisms and can involve myristoyl, palmitoyl, isoprenyl polymers of various lengths or aminoglycan-linked phospholipids (1,2). Lipoprotein modifications investigated here are restricted to bacteria. The lipoprotein biosynthesis pathway is a major virulence factor in Mycobacterium tuberculosis, the causative agent of human tuberculosis. Every year 1.6 million people fall prey to tuberculosis and one third of the world’s population is infected (http://www.who.int/mediacentre/factsheets/fs104/en/index.html). Thus, tuberculosis is responsible for 2.5 % of deaths in the world, which is the highest rate claimed by a single infectious agent. An M. tuberculosis knock-out mutant deficient in lipoprotein signal peptidase lspa showed reduced multiplication in bone-marrow derived macrophages, complete absence of lung pathology and a 1000 fold reduced number of colony forming units in a mouse model of infection (3,4). Likewise, lipoprotein synthesis contributes to virulence of other Gram-positive pathogens, Listeria, Staphylococci and Streptococci (5).

Bacterial lipoproteins are a functionally diverse class of lipiddated proteins involved in cell wall synthesis, nutrient uptake, adhesion and transmembrane signalling (6) and about 2 % of open reading frames encode this kind of proteins.
Lipidation allows anchoring of these proteins to the cell surface. Lipoproteins are characterized by the presence of a consensus sequence, the “lipobox”, located in the C-terminal part of the leader sequence and consisting of four amino acids [LVI/ASTVI/GAS/C] (7). Precursor lipoproteins are mainly translocated in a Sec-dependent manner across the plasma membrane and modified subsequently on the universally conserved, essential cysteine residue located in the lipobox motif. The modifications taking place after translocation are consecutively mediated by three enzymes: 1. formation of a thioether linkage between the conserved cysteine residue and a diacylglycerol catalyzed by phosphatidylglycerol: pre-prolipoprotein diacylglyceryl transferase (Lgt), 2. cleavage of the N-terminal signal peptide by prolipoprotein signal peptidase/signal peptidase II (LspA) and 3. in case of Gram-negative bacteria, aminoacylation of the N-terminal cysteine residue by phospholipid:apolipoprotein N-acyltransferase (Lnt) (6-8). In E. coli, most of the mature triacylated lipoproteins are finally transported across the periplasm by the LolABCDE transport system (9). Homologues of the Lol-transport system are absent in Mycobacteria. Although lipoprotein modifying enzymes act sequentially, Lgt-independent LspA-mediated signal sequence cleavage has recently been demonstrated in Listeria monocytogenes (10). While Lgt and LspA are universally present in both, Gram-positive and Gram-negative bacteria, Lnt has been reported to be restricted to Gram-negatives (11), although some indications for N-acylation in Bacillus subtilis and Staphylococcus aureus were reported (12-15).

Mycobacterial lipoproteins are immunodominant antigens (16) and several manipulate innate immune mechanisms and antigen presenting cells (17). It is known that mycobacterial lipoproteins, e.g. the 19 kDa lipoprotein, activate toll-like receptor 2 (TLR2) and co–receptors TLR1, which recognize triacylated peptides, but also TLR6, which recognize diacylated peptides (18,19). However, the lipid linkage of mycobacterial lipoproteins has not been determined.

In this study, we show that Lnt activity is more widely distributed than previously assumed. We demonstrate apolipoprotein N-acyltransferase activity in a Gram-positive Mycobacterium and give complete structural information about the lipid modification of mycobacterial lipoproteins. Hereby, the functionality of Lnt homologues in Actinomycetes is revealed (5). We show that mycobacterial lipoproteins are triacylated and carry mycobacteria specific fatty acids.

**EXPERIMENTAL PROCEDURES**

**Bacterial strains and growth conditions**

*Mycobacterium smegmatis* was grown on Middlebrook 7H10 agar supplemented with oleic acid albumin dextrose (OADC, Difco) or LB agar. Tween 80 (0.05%) was added to liquid broth to avoid clumping; when appropriate, antibiotics were added at the following concentrations: kanamycin 50 µg ml$^{-1}$; streptomycin 100 µg ml$^{-1}$; hygromycin 25 µg ml$^{-1}$. Strain designations were as follows: *M. smegmatis* SmR5 (20), a derivative of *M. smegmatis mc²* (21) carrying a non-restrictive rpsL mutation conferring streptomycin resistance (= parental strain); int::aph = *Int* knock-out mutant; int::aph-lntMs = *M. smegmatis* int::aph transformed with complementing vector pMV361-hyg-lntMs; int::aph-ppm1Tb = *M. smegmatis* int::aph transformed with complementing vector pMV361-hyg-ppm1Tb.

**Complementation of conditional E. coli Int mutant PAP8508**

LntMs was amplified by PCR and cloned into the EcoRI / BamHI sites of pUC18 resulting in pUC18-lntMs. Plasmids pUC18-lntMs323W, pUC18-lntMs477Y and pUC18-lntMs323W/477Y were generated by standard mutagenesis-PCR techniques.

The *E. coli* conditional *Int* mutant PAP8508 and its parental strain PAP105 (a generous gift of N. Buddelmeijer) were used for complementation analysis (22). Strains were plated on LB agar supplemented with 1 mM IPTG, 100 µg / ml ampicillin and either 0.4 % (w/v) glucose or 0.2 % (w/v) arabinose.

**Disruption of Int in M. smegmatis**

A 3.8 kbp genomic fragment of *M. smegmatis* from position 3’929’396 to 3’933’223 (http://cmr.jcvi.org/cgi-bin/CMR/GenomePage.cgi?org=gms) spanning the entire *IntMs* gene was PCR amplified and cloned into pGem-T Easy (Promega) to result in pGem-T Easy-IntMs. For functional inactivation of *IntMs* a 1.04 kbp SfiI /
EcoRV fragment was replaced with a 1.4 kbp SmaI / HpaI kanamycin resistance cassette from pUC4K (Pharmacia) subcloned in pMCS5-Kan (unpublished). Subsequently a 4.6 kbp PvuII fragment containing the inactivated IntMs allele (IntMs::aph) was inserted into the EcoRV site of ptrpA1-rpsL (20) to result in ptrpA1-rpsL-IntMs::aph. The IntMs::aph allele was substituted for IntMs in the M. smegmatis chromosome as described previously (23) and confirmed by Southern blot analyses with a 0.2 kbp SmaI / NcoI IntMs upstream probe.

For complementation with M. smegmatis Int, a 4.3 kbp PvuII fragment from pGem-T Easy-IntMs comprising the entire IntMs gene was cloned into the HpaI site of plasmid pMV361-hyg (24) to result in pMV361-hyg-IntMs. For complementation with M. tuberculosis ppm1 a 6.3 kbp fragment from M. tuberculosis’ genomic position 2’306’187 to 2’312’526 spanning the entire ppm1 gene was cloned into pGem-T Easy to result in pGem-T Easy-ppm1Tb and subsequently subcloned as a 6.3 kbp EcoRI fragment into the HpaI site of plasmid pMV361-hyg (24) to result in pMV361-hyg-ppm1Tb. Complementation was confirmed by Southern blot analyses with a 0.2 kbp SmaI / NcoI IntMs upstream probe and a 0.2 kbp KpnI / HindIII ppm1Tb upstream probe.

Construction of expression vector pMV261-Gm-FusLppX
Plasmid pMV261-Gm a derivative of pMV261 is a shuttle vector replicating in E. coli as well as in mycobacteria (25). M. tuberculosis LppX was amplified by PCR from genomic DNA and fused to the M. tuberculosis 19 kDa promoter. Two sequences encoding a hemagglutinin and a hexa-His epitope were fused to the 3’ part of the gene to facilitate subsequent purification and detection on Western blot and the insert was cloned into the EcoRI site to result in pMV261-Gm-FusLppX.

Preparation of cell extracts and Western blot analysis
Bacteria from 2 L cultures were harvested, resuspended in PBS containing Complete EDTA free tablets (Roche) to inhibit protein degradation and subjected to two French press cycles (American Instrument Company) at 2 × 10⁸ Pa. Extracts were treated with 2 % Sodium N-lauroylsarcosine (SLS) for 1 h at room temperature and subsequently incubated at 4 °C over night. Soluble and insoluble fractions were separated by centrifugation at 30’000 g for 1 h at 4 °C. Extracts corresponding to 1–5 µg of total protein were separated by SDS-PAGE (12%) and analyzed by Western blot. Antiserum against HA epitope (Roche) was diluted 1:300.

FPLC protein purification
The soluble fraction of cell extracts was diluted with buffer containing 20 mM NaH₂PO₄, 0.5 M NaCl to 1 % SLS and loaded on HisTrap™ HP column (GE Healthcare) equilibrated with buffer containing 20 mM NaH₂PO₄, 0.5 M NaCl, 0.2 % SLS and 20 mM imidazole. Proteins were eluted with 0.125-0.5 M imidazole.

Thrombin cleavage of RecLppX
Purified RecLppX was dialyzed against PBS pH 7.4 at 4 °C because imidazole can lower thrombin activity (26). About 0.1-1 µg RecLppX was digested with 4-30 NIH units of thrombin from bovine plasma (Sigma) for 16 h at 37 °C with continuous shaking (25 rpm). Reaction was stopped through incubation at 95 °C for 5 min.

MALDI-TOF/TOF
Purified lipoprotein (100-200 pmol) were prepared and analysed according to Ujihara et al. (27). After tryptic digestion samples were resuspended in 5 µl 0.1 % trifluoroacetic acid, 66 % acetonitrile. 1.2 µl were loaded onto the target and covered with 1 µl matrix (α-cyano-4-hydroxy-cinnamic acid (Bruker Daltonics), 5 mg/ml in 0.1 % TFA, 50 % acetonitrile). Mass spectra were recorded on an Ultraflex II MALDI-TOF/TOF instrument with smartbeam laser upgrade (Bruker, Germany). The frequency-tripled Nd:YAG laser using a structured-focus profile (smartbeam, Bruker-Daltonics) was set to a repetition rate of 100 Hz and the ion acceleration voltage was 29.5 kV. The mass measurements were performed in the positive ion reflector mode.

RESULTS
Mycobacterial lipoproteins are modified at the N-terminus
We chose the well characterized M. tuberculosis lipoprotein LppX (28) as a model substrate for
mycobacterial lipoprotein synthesis and generated the expression vector pMV261-Gm-FusLppX. Plasmid pMV261-Gm-FusLppX was transformed into \textit{M. smegmatis} SmR5 \cite{20}, \textit{M. smegmatis} \textit{ΔlspA} and complemented \textit{M. smegmatis} \textit{ΔlspA-lspA} (unpublished). Whole cell extracts of these strains were subjected to Western blot analysis with anti-HA antibody. We observed bands with an apparent size of 23 kDa in parental and complemented strain and 26 kDa in the \textit{ΔlspA} mutant and to a little amount also in the parental strain. After purification of LppX, the 26 kDa band was also detected to a higher amount in the parental strain. The 23 kDa band corresponds to the predicted mass of mature recombinant LppX-HA-His and 26 kDa band to the predicted mass of the prolipoprotein form of LppX-HA-His. These results indicate LspA-dependent signal peptide cleavage of recombinant LppX-HA-His in \textit{M. smegmatis} and verify its post-translational modification within the lipoprotein synthesis pathway (Figure 1a).

Purified LppX-HA-His from \textit{M. smegmatis} parental strain was subjected to protein sequence analysis. Edman degradation of the prolipoprotein revealed a sequence starting at the initial methionine of the signal peptide of LppX (Figure 1b). In contrast, no sequence was obtained from the mature LppX indicating a modification of the N-terminal amino group.

\textbf{Identification of putative N-acyltransferases in bacterial genomes}

In \textit{E. coli}, N-acylation of lipoproteins is conferred by Lnt \cite{29}. We performed a BLAST search analysis (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi) with \textit{E. coli} Lnt as a query to investigate the distribution of Lnt homologues in the bacterial kingdom and to identify putative homologues in mycobacteria. Lnt homologues are widely distributed in Gram-negative bacteria (\(α,β,γ,δ,ε\) Proteobacteria, Spirochetes, Aquifex, Cytophaga, Thermotoga), but absent from all classes (Clostridia, Mollicutes, Bacilli) of low GC Gram-positive bacteria (Firmicutes), although some indications for N-acylation in low GC Gram positives have been reported \cite{12-15}. In contrast, Lnt homologues were identified in all classes of high GC Gram-positive bacteria (Actinobacteria, e.g. Streptomyces, Nocardia, Corynebacteria and Mycobacteria) (Figure 2a), but Lnt activity of those homologues could not be demonstrated \cite{22}. The cell envelope of the phylum Actinobacteria is more complex than the cell envelope of Firmicutes. In \textit{M. tuberculosis} and \textit{M. smegmatis}, \textit{Rv2051c} (Ppm1) and MSMEG_3860 (Ppm2) have the highest similarity to \textit{E. coli} Lnt. \textit{M. tuberculosis} \textit{Rv2051c} encodes a two-domain protein, of which the N-terminal part shows similarity to \textit{E. coli} Lnt. The C-terminal part of the protein encodes a polyprenol-phosphomannose (Ppm) synthase, an enzyme involved in lipomannan and lipoarabinomannan synthesis \cite{30}. MSMEG_3860 has been shown to stabilize \textit{M. smegmatis} Ppm1 in the bacterial membrane and therefore has been annotated as Ppm2 \cite{31}. MSMEG_3860 will be referred to as LntMs. Lnt homologues are also present in \textit{Mycobacterium avium} and \textit{Mycobacterium leprae} and are encoded by a separate open reading frame as in \textit{M. smegmatis}. The genomic region surrounding Lnt homologues is conserved in mycobacteria (Figure 2b).

\textit{M. smegmatis} \textit{Lnt} does not restore growth of a conditional \textit{E. coli lnt} mutant

All enzymes of the lipoprotein synthesis pathway are essential in Gram-negatives. We intended to demonstrate mycobacterial Lnt activity by complementation of an \textit{E. coli} conditional \textit{lnt} mutant. Since LntMs is encoded by a separate open reading frame and is not fused to a second domain as in \textit{M. tuberculosis}, we chose \textit{M. smegmatis} \textit{Lnt} (MSMEG_3860) instead of \textit{M. tuberculosis} \textit{Ppm1} (\textit{Rv2051c}) for complementation. LntMs encodes a protein of 654 amino acids with a 25 % identity and 40 % similarity to \textit{E. coli} and a 63 % identity and 73 % similarity to the N-terminal part of \textit{M. tuberculosis} \textit{Rv2051c}. LntMs was cloned into vector pUC18 (Fermentas) to result in pUC18-LntMs and transformed into the conditional \textit{E. coli lnt} mutant PAP8508 \cite{22}. However, we could not restore growth of the PAP8508 mutant under restrictive conditions (data not shown). Seven amino acids (W237, E267, K335, E343, C387, Y388, E389) are reported to be essential for \textit{E. coli} Lnt function \cite{22}. Five of these seven residues are conserved in LntMs, while two are altered (LntEc W237 corresponds to LntMs E323, LntEc Y388 corresponds to LntMs W477). We exploited site directed mutagenesis to introduce these \textit{E. coli}
codons into the *M. smegmatis* sequence of pUC18-lntMs to result in pUC18-lntMs323W, pUC18-lntMs477Y and pUC18-lntMs323W/477Y. However transformation of none of these vectors complemented the conditional *E. coli* *int* mutant (data not shown).

**Generation and characterization of *M. smegmatis int::aph* mutant**

Since we were unable to complement an *E. coli* *int* mutant, we decided to investigate Lnt activity directly in mycobacteria by generating a *M. smegmatis int* deletion mutant. The deletion mutant was constructed by transformation of *M. smegmatis* SmR5 with the suicide plasmid ptrpA1-rpsL-lntMs::aph using the *rpsL* counterselection strategy (20). The mutant strain resulting from allelic replacement is here referred to as *M. smegmatis int::aph*. Deletion of *lntMs* was verified by Southern blot analysis using a 5' *lntMs* DNA probe (Figure S1). The probe hybridized to a 1.4 kbp fragment of the parental strain and a 6.4 kbp fragment of the *int::aph* mutant. Transformation of these plasmids into *M. smegmatis int::aph* mutant resulted in strains *M. smegmatis int::aph-lntMs* and *M. smegmatis int::aph-ppm1Tb*.

Western blot analysis of extracts from *M. smegmatis int::aph* expressing LppX-HA-His revealed a molecular mass of the detected protein, which can not be distinguished from that of LppX-HA-His expressed in *M. smegmatis* parental strain. However, N-terminal sequencing revealed that LppX-HA-His purified from *M. smegmatis int::aph* is accessible to Edman degradation (sequence CSSP) indicating that the N-terminal amino group is not blocked anymore.

**LntMs and Ppm1Tb are apolipoprotein N-acyltransferases**

Since fatty acids of membrane phospholipids are the substrates for N-acylation of lipoproteins in *E. coli* (32-34), its lipoproteins are modified with myristic, palmitic, palmitoleic, oleic or vaccinic acid (35). Phospholipids in mycobacteria mainly consist of palmitic, palmitoleic, oleic and tuberculostearic acid (10-methyloctadecanoic acid) (36). Therefore we hypothesized that N-acylation of lipoproteins in mycobacteria increases the molecular mass by approximately 0.3 kDa. To differentiate between lipoproteins with a free or acylated N-terminus, we cloned an additional expression vector, RecLppX. It differs from LppX-HA-His by a hemagglutinin epitope followed by a thrombin cleavage site inserted after amino acid Ala (+19) of the mature LppX (Figure 1b). The thrombin cleavage site LVPRGS was inserted to produce a small N-terminal fragment of 33 residues (about 3.5 kDa) after thrombin cleavage. To ensure that the insertion of a HA-epitope and a thrombin site does not abolish recognition of RecLppX as a lipoprotein, we analyzed total lysates of *M. smegmatis* parental strain, *M. smegmatis ΔlspA* and *M. smegmatis ΔlspA-lspA* by Western blot (Figure 3a). Temperature-sensitive *lspA* mutants of *E. coli* and *lspA* knock-out mutants of Gram-positive bacteria accumulate prolipoproteins (37,38). Immunoblotting of total lysates of *M. smegmatis* parental strain, ΔlspA and ΔlspA-lspA with antiserum against the HA-epitope (Roche) revealed the presence of a 25 kDa band in parental and complemented strains. In contrast, a band with a slightly larger size (appr. 27 kDa, the increase corresponds to the mass of the signal sequence) was observed in the ΔlspA strain. This result shows that the insertion of a HA-epitope and a thrombin cleavage site did not impair the recognition of RecLppX as a lipoprotein. We then investigated thrombin digested RecLppX for Lnt-dependent modification by Western blot analyses (Figure 3b). In the *int::aph* mutant we observed a slightly smaller size of the N-terminal part of RecLppX suggesting, that there are fewer modifications on the protein compared to parental strain and both complemented strains *int::aph-lntMs* and *int::aph-ppm1Tb*. In all strains, we also found a double band of the N-terminal part of RecLppX, indicating partial modification of RecLppX by other enzymes than Lnt or LspA. We also observed a deviation from the calculated size of the N-terminal part of mature RecLppX. The molecular mass was calculated to be 3.5 kDa but in the parental strain we found a band corresponding to a size of about 6 kDa. This
difference in size is probably due to an altered migration behaviour because of lipid modifications. It can be excluded that these bands at 6 kDa are prolipoprotein forms, still containing the signal peptide, because the N-terminal part of pro-RecLppX from ΔspA mutant is running at about 8.5 kDa. These results show that RecLppX is modified by LspA as well as by LntMs. Ppm1Tb is sufficient to replace LntMs implicating that similar lipoprotein modifications take place in *M. tuberculosis*. Recombinant *M. tuberculosis* LppX (FusLppX) was heterologously expressed and purified. Tryptic fragments of FusLppX were analyzed by MALDI-TOF/TOF mass spectrometry to characterize modifications taking place on lipoproteins in *M. smegmatis* at the molecular level. Purified mature LppX from parental strain, int::aph mutant and int::aph-ppm1Tb was prepared for analysis according to Ujihara et al. (27). For identification of the modifications of the universally conserved cysteine, the structure of the N-terminal tryptic peptide was determined. Experimentally found m/z values are summarized and compared to calculated m/z values in Table 1. Trypsin cleavage sites of LppX are given in Figure 1b. The expected monoisotopic molecular mass of the unmodified N-terminal tryptic peptide of LppX is 2963.46 Da. Instead, we found a [M+H]⁺ signal at m/z = 3795.42 for the N-terminal tryptic peptide of LppX from parental strain and a signal at m/z = 3557.01 from the int::aph mutant. The N-terminal tryptic peptide from int::aph-ppm1Tb is a mixture and showed signals of m/z = 3795.32 and 3557.07 (Figure 4). This indicates that the N-terminal peptide of LppX is modified in an LntMs dependent manner. To identify the found modification, we calculated diacylglycerol modifications with all theoretical combinations of the four fatty acids found in mycobacterial phospholipids: palmitic acid (C16:0), palmitoleic acid (C16:1), oleic acid (C18:1) and tuberculostearic acid (C19:0). The difference in molecular mass between the peptide of the int::aph mutant and the unmodified peptide is 592.55 Da indicating a diacylglycerol modification with ester-linked C19:0 and C16:0 fatty acid. These fatty acids are most likely tuberculostearic acid and palmitic acid. The difference in molecular mass of 238.41 Da between int::aph mutant (m/z = 3557.07) and parental strain (m/z = 3795.42) indicates an additional modification with a C16:0 fatty acid in the parental strain. By thorough analyzing the signal at m/z = 3795.42 in high resolution, we also observed a minor signal (approx. 10 %) at m/z = 3793.35. This indicates the presence of a C16:1 fatty acid in place of the C16:0 fatty acid, but only to a little amount (data not shown). In the complemented mutant int::aph-ppm1Tb both forms of the N-terminal tryptic peptide of LppX were found. This finding indicates partial complementation of LntMs by *M. tuberculosis* Ppm1.

In the int::aph and the int::aph-ppm1Tb mutant we additionally found an N-terminal tryptic peptide at a signal of m/z = of 3739.18 and 3739.32, respectively. But no evidence for an amino-acyl modification was found in the MS/MS fragmentation pattern. In contrast, the release of 183 Da most likely corresponds to a covalent modification of the free N-terminal amino group of LppX in these two strains with 4-(2-Aminoethyl)-benzenesulfonyl fluoride (AEBSF), a component of the protease inhibitor mix which has been used. AEBSF is known to modify hydroxylated amino acids and to a lesser extent also free amino groups (http://www.abrf.org/index.cfm/dm.details?DMID=235&AvgMass=183&MARGIN=0).

To obtain information about linkage of the modifications, we elucidated the structure of the triacylated N-terminal tryptic peptide of LppX (m/z = 3795.42) was investigated by MS/MS (Figure 5, Figure S2). A summary of all found eliminations in the three strains are given in Table 2. The ions at m/z = 3539.59 and at m/z = 3496.52 correspond to the neutral loss of a C16 (Δ = 255.82 Da) and a C19:0 fatty acid (Δ = 298.89 Da), respectively. The most intense fragment ion at m/z = 3169.21 corresponds to the elimination of a diacylthioglycerol carrying both O-linked C19:0 and C16 fatty acids (Δ = 626.22 Da). In addition the release of 370.38 Da from the ion at m/z = 3539.59 corresponds to the elimination of a C19:0 fatty acid α-thioglyceryl ester and the release of 327.31 Da from the ion at m/z = 3496.52 corresponds to the elimination of a C16 fatty acid α-thioglyceryl ester. This fragmentation pattern shows, that the +1 cysteine is modified at the sulfhydryl group by a diacylglycerol-residue carrying ester bound C16...
fatty acid and C19:0 fatty acid. Whether the C19:0 fatty acid is in the SN1 or SN2 position can not be determined. The release of 255.08 Da from the ion at m/z = 3169.21 indicates the release of a palmitamid acid derived from N-terminal C16 fatty acid. The MS/MS fragmentation pattern of LppX N-terminal tryptic peptide of the int::aph mutant (m/z = 3557.07) showed only the elimination of the diacylthioglycerol, but no release of palmitamid acid, as expected (Figure 5). In order to verify the diacylglycerol modification, we also analyzed tryptic peptides of LppX from the ΔlspA mutant by MALDI-TOF/TOF mass spectrometry (Figure 6). Experimentally found m/z values are summarized and compared to calculated m/z values in Table 3. Trypsin cleavage sites of pro-LppX are given in Figure 1b. We found three m/z signals corresponding to the tryptic peptide containing the 1 cysteine. The signal at m/z = 4887.30 corresponds to the peptide with a disulfide bridge between the two present cysteines (position -8 and +1). The signal at m/z = 5041.23 corresponds to the peptide with both cysteines being modified by β-mercaptoethanol, a buffer component used for SDS-PAGE. The signal at m/z = 5558.12 corresponds to the peptide with one cysteine being modified with a β-mercaptoethanol but the other being modified with the diacylglycerol carrying O-linked C16 and C19:0 fatty acids also found in the previously analyzed strains (Figure 6). This result shows that LppX purified from the ΔlspA mutant is a mixture of pre-pro-LppX and pro-LppX. Taken together the results show that the universally conserved cysteine of M. tuberculosis LppX is modified with a thioether linked diacylglycerol residue carrying an ester-bound C19:0 and an ester-bound C16 fatty acid. In addition, it is modified with an amide-linked third C16 fatty acid. The C19:0 fatty acid corresponds most likely to the mycobacterial specific tuberculostearic acid. It is also proved, that LntMs is an N-acyltransferase and M. tuberculosis Ppm1 is able to complement M. smegmatis int::aph mutant and therefore Ppm1 seems to be a bifunctional protein. Within this protein the N-terminal domain presumably exhibits N-acyltransferase activity (our data) and the C-terminal domain exhibits mannosyl transferase activity (30).

**DISCUSSION**

The lipoprotein biosynthesis pathway consisting of the three enzymes Lgt, LspA and Lnt has been intensively studied in *E. coli* and has been shown to be essential and necessary for transport of lipoproteins to the outer membrane of Gram-negative bacteria (11,39,40). In mycobacteria, little is known about synthesis and localization of lipoproteins, only a few lipoproteins are functionally characterized and annotation is mainly based on theoretical considerations instead of experimental evidence. However, consistent with the biosynthetic pathway in *E. coli*, putative lgt (Rv1614) and lsp (Rv1539) genes have been identified in the *M. tuberculosis* genome (41). In previous studies we could show that in mycobacteria the lipoprotein pathway is a major virulence factor (3,4). For fundamental knowledge and further investigations, we were interested in how lipoproteins are modified in mycobacteria. In the present study, we investigated the lipid moieties of a representative mycobacterial lipoprotein.

We identified Lnt homologues in mycobacteria, corynebacteria and streptomyces species. In low GC Gram positive bacteria Lnt homologues are completely absent (Figure 2), but in 1985 the first indirect detection of N-acylation in the Gram-positive Bacillus subtilis was published and in S. aureus triacylation of lipoprotein SitC was recently reported (14,15), while another lipoprotein (SAOUHSC_02699) was only found to be diacylated (13). The protein responsible for attaching the third fatty acid to lipoproteins in *S. aureus* has not been identified. It may be differentially expressed depending on culture conditions or may have a narrow substrate specificity. In *M. tuberculosis* the Lnt homologue found, is annotated as Rv2051c. This ORF was originally annotated as a two-domain enzyme with a putative N-terminal Lnt domain and a C-terminal polyprenol monophosphomannose synthase (Ppm1) domain and was characterized as the latter one (30). Although the putative Lnt domain is not needed for Ppm1 activity, on overexpression in *M. smegmatis* it appeared to enhance the mannosyltransferase activity. Interestingly, the two domains of *M. tuberculosis* Ppm1 are encoded by separate, adjacent open reading frames in the genomes of other mycobacteria (Figure 2b).
Previous attempts to complement a conditional *E. coli* *lnt* mutant with Lnt homologues from other bacterial species corresponding to the order Actinomycetales (Streptomyces, Corynebacterium) failed (22). Likewise we were unable to complement this *E. coli* strain with a mycobacterial Lnt homologue. Even after exchange of the two essential amino acids differing between *M. smegmatis* and *E. coli*, complementation of *E. coli* *lnt* mutant failed. LntMs as *E. coli* Lnt attaches a C16 fatty acid to the free amino group of the universally conserved cysteine. Therefore the failure of complementation is not due to absence of fatty acid substrates. Rather mycobacterial lipoproteins are modified with a diacylglycerol carrying mycobacterial specific fatty acids. Failure of complementation of PAP8508 therefore is probably due to the fact that LntMs recognizes only lipoproteins modified with a diacylglycerol residue carrying at least one ester bound mycobacterial specific fatty acid. This implies that LntMs does not recognize lipoproteins modified with diacylglycerol residues carrying only small fatty acids like palmitic or palmitoleic acid. Specificity could be tested in an *in vitro* assay system. Alternatively, the expression level or enzymatic activity of mycobacterial Lnt homologues may not sustain growth of fast growing *E. coli*.

We then investigated LntMs and *M. tuberculosis* Ppm1 activity in a mycobacterial background. As *lntMs* is not an essential gene in mycobacteria, we generated an isogenic *M. smegmatis* *lnt::aph* mutant. After thrombin cleavage, the recombinant lipoprotein (RecLppX) extracted from *M. smegmatis* *lnt::aph* mutant showed a faster running behaviour on SDS-PAGE than RecLppX extracted from the parental strain. The size was about 0.3 kDa smaller corresponding to fewer modifications of RecLppX in the *lnt::aph* mutant. We also recognized a double band of digested RecLppX in all strains used as well as a discrepancy between the calculated and the apparent molecular mass. The altered running behaviour is probably due to the modifications on the small N-terminal fragment and the observed double band indicates partial processing of RecLppX by enzymes other than Lgt, LspA or Lnt. Glycosylation of RecLppX is one possibility, but information about the structure, function, and biosynthetic pathways of prokaryotic glycoproteins is scarce. Glycosylation of *M. tuberculosis* lipoproteins has been confirmed for the 45/47 kDa protein, SodC and for the *Mycobacterium bovis* MPB83 protein (Rv2873) (42-44). Glycosylation of SodC influences its ultimate subcellular localization and also its proteolytic processing.

By performing MALDI-TOF/TOF analyses of a trypsin digest of purified LppX we unambiguously identified modifications at the universally conserved cysteine. After trypsin cleavage of LppX, the N-terminal peptide from the parental strain has a mass of 3794.41 Da instead of 2963.46 Da predicted for the unmodified peptide, whereas the N-terminal peptide from the isogenic *lnt::aph* mutant showed a mass of 3556.01 Da. This strongly indicates that LntMs covalently links a C16 fatty acid (C16:0 or C16:1) to the N-terminus of the peptide. The additional increase by 592 Da corresponds to a diacylglycerol residue with C16 fatty acid and a C19:0 fatty acid, which corresponds most likely to tuberculostearic acid (10-methyloctadecanoic acid), forming a thioether-linkage to the sulfhydryl group of the cysteine. Whether the C19:0 fatty acid is in the SN1 or SN2 position can not be determined. The same diacylglycerol modification was also found in the Δ*lspA* mutant. The loss of the N-acyl modification in the *lnt::aph* mutant is complemented by the *M. tuberculosis* homologue Ppm1 suggesting that mature *M. tuberculosis* lipoproteins are N-acylated. N-acylation affects the interaction of lipoproteins with innate immune receptors (45). Ppm1 was shown to exhibit polyrenol-monophosphomannose synthetase activity (30) although the major part of the protein has homology to *E. coli* Lnt. Both masses of the N-terminal tryptic peptide (3556.07 Da and 3794.32 Da) were found in the complemented mutant indicating that not all apolipoprotein was converted to mature lipoprotein by Ppm1. As *M. tuberculosis* has a generation time of about 24 hours and *M. smegmatis* only of about 3 hours, it is possible that *M. tuberculosis* Ppm1 has a lower enzymatic activity than LntMs. Alternatively, the expression level of *M. tuberculosis* Ppm1 is lower than that of *M. smegmatis* Lnt.

The identification of O-linked tuberculostearic acid shows that mycobacterial lipoproteins are modified with mycobacterial specific fatty acids and differ from lipoproteins modified in *E. coli*. 
In this study we directly show, that Gram-positive mycobacteria synthesize triacylated lipoproteins. This is the first time to our knowledge that responsible genes for Lnt activity are assigned in Gram-positive bacteria. LntMs and M. tuberculosis Ppm1 are functional homologues of E. coli Lnt as they catalyze the transfer of the third acyl moiety to the free α-amino group of the N-terminal amino acid of lipoproteins. Most likely mycobacterial Lnt homologues differ in substrate specificity from E. coli Lnt. N-acylation is a prerequisite for transport of E. coli lipoproteins to the outer membrane (46). Likewise, N-acylation of mycobacterial lipoproteins may be required for transport to the outer most lipid layer of mycobacteria which according to recent investigations resembles the outer membrane of Gram-negatives (47,48).

ACKNOWLEDGEMENTS

This work was supported in part by the Swiss National Science Foundation (3100A0-120326) and the European Union (TB-Drug: LSHP-CT-2006-037217).

REFERENCES

1. Eichler, J., and Adams, M. W. (2005) Microbiol Mol Biol Rev 69, 393-425
2. Nadolski, M. J., and Linder, M. E. (2007) FEBS J 274, 5202-5210
3. Rampini, S. K., Selchow, P., Keller, C., Ehlers, S., Bottger, E. C., and Sander, P. (2008) Microbiology 154, 2991-3001
4. Sander, P., Rezwan, M., Walker, B., Rampini, S. K., Kroppenstedt, R. M., Ehlers, S., Keller, C., Keeble, J. R., Hagemeier, M., Colston, M. J., Springer, B., and Bottger, E. C. (2004) Mol Microbiol 52, 1543-1552
5. Hutchings, M. I., Palmer, T., Harrington, D. J., and Sutcliffe, I. C. (2009) Trends Microbiol 17, 13-21
6. Sutcliffe, I. C., and Harrington, D. J. (2004) FEMS Microbiol Rev 28, 645-659
7. Babu, M. M., Priya, M. L., Selvan, A. T., Madera, M., Gough, J., Aravind, L., and Sankaran, K. (2006) J Bacteriol 188, 2761-2773
8. Rezwan, M., Grau, T., Tschumi, A., and Sander, P. (2007) Microbiology 153, 652-658
9. Yakushi, T., Masuda, K., Narita, S., Matsuyama, S., and Tokuda, H. (2000) Nat Cell Biol 2, 212-218
10. Baumgartner, M., Karst, U., Gerstel, B., Loessner, M., Wehland, J., and Jansch, L. (2007) J Bacteriol 189, 313-324
11. Wu, H. C. (1996) in E.coli and Salmonella typhimurium, Neidhardt F.C. and others Ed., pp. 1005-1014, American Society for Microbiology., Washington, DC
12. Navarre, W. W., Daefler, S., and Schneewind, O. (1996) J Bacteriol 178, 441-446
13. Tawaratsumida, K., Furuyashiki, M., Katsumoto, M., Fujimoto, Y., Fukase, K., Suda, Y., and Hashimoto, M. (2009) J Biol Chem 284, 9147-9152
14. Hayashi, S., Chang, S. Y., Chang, S., Giam, C. Z., and Wu, H. C. (1985) J Biol Chem 260, 5753-5759
15. Kurokawa, K., Lee, H., Roh, K. B., Asanuma, M., Kim, Y. S., Nakayama, H., Shiratsuchi, A., Choi, Y., Takeuchi, O., Kang, H. J., Dohmae, N., Nakanishi, Y., Akira, S., Sekimizu, K., and Lee, B. L. (2009) J Biol Chem 284, 8406-8411
16. Young, D. B., and Garbe, T. R. (1991) Res Microbiol 142, 55-65
17. Pecora, N. D., Gehring, A. J., Canaday, D. H., Boom, W. H., and Harding, C. V. (2006) *J Immunol* **177**, 422-429
18. Drage, M. G., Pecora, N. D., Hise, A. G., Febbraio, M., Silverstein, R. L., Golenbock, D. T., Boom, W. H., and Harding, C. V. (2009) *Cell Immunol* **258**, 29-37
19. Buwitt-Beckmann, U., Heine, H., Wiesmuller, K. H., Jung, G., Brock, R., Akira, S., and Ulmer, A. J. (2006) *J Biol Chem* **281**, 9049-9057
20. Sander, P., Meier, A., and Bottger, E. C. (1995) *Mol Microbiol* **16**, 991-1000
21. Snapper, S. B., Melton, R. E., Mustafa, S., Kieser, T., and Jacobs, W. R., Jr. (1990) *Mol Microbiol* **4**, 1911-1919
22. Vidal-Ingigliardi, D., Lewenza, S., and Buddelmeijer, N. (2007) *J Bacteriol* **189**, 4456-4464
23. Sander, P., Springer B., Böttger E. C. (2001) in *Mycobacterium tuberculosis Protocols* (Parish, T., Stoker, N. G., ed), pp. 93-104, Humana Press
24. Sander, P., Prammananan, T., Meier, A., Frischkorn, K., and Bottger, E. C. (1997) *Mol Microbiol* **26**, 469-480
25. Stover, C. K., de la Cruz, V. F., Fuerst, T. R., Burlein, J. E., Benson, L. A., Bennett, L. T., Bansal, G. P., Young, J. F., Lee, M. H., Hatfull, G. F., et al. (1991) *Nature* **351**, 456-460
26. Isaacs, R. C., Solinsky, M. G., Cutrona, K. J., Newton, C. L., Naylor-Olsen, A. M., Krueger, J. A., Lewis, S. D., and Lucas, B. J. (2006) *Bioorg Med Chem Lett* **16**, 338-342
27. Ujihara, T., Sakurai, I., Mizusawa, N., and Wada, H. (2008) *Anal Biochem* **374**, 429-431
28. Sulzenbacher, G., Canaan, S., Bordat, Y., Neyrolles, O., Stadthagen, G., Roig-Zamboni, V., Rauzier, J., Maurin, D., Laval, F., Daffe, M., Cambillau, C., Gicquel, B., Bourne, Y., and Jackson, M. (2006) *EMBO J* **25**, 1436-1444
29. Gupta, S. D., Gan, K., Schmid, M. B., and Wu, H. C. (1993) *J Biol Chem* **268**, 16551-16556
30. Gurcha, S. S., Baulard, A. R., Kremer, L., Locht, C., Moody, D. B., Muhlecker, W., Costello, C. E., Crick, D. C., Brennan, P. J., and Besra, G. S. (2002) *Biochem J* **365**, 441-450
31. Baulard, A. R., Gurcha, S. S., Engohang-Ndong, J., Gouffi, K., Locht, C., and Besra, G. S. (2003) *J Biol Chem* **278**, 2242-2248
32. Lai, J. S., and Wu, H. C. (1980) *J Bacteriol* **144**, 451-453
33. Gupta, S. D., Dowhan, W., and Wu, H. C. (1991) *J Biol Chem* **266**, 9983-9986
34. Jackowski, S., and Rock, C. O. (1984) *J Biol Chem* **261**, 11328-11333
35. Lin, J. J., Kanazawa, H., and Wu, H. C. (1980) *J Biol Chem* **255**, 1160-1163
36. Goren, M. B., and P. J. Brennan. (1979) in *Tuberculosis* ((ed.), G. P. Y., ed), pp. 63-193, The W. B. Saunders Co., Philadelphia, Pa.
37. Venema, R., Tjalsma, H., van Dijl, J. M., de Jong, A., Leenhouts, K., Buist, G., and Venema, G. (2003) *J Biol Chem* **278**, 14739-14746
38. Sankaran, K., and Wu, H. C. (1995) *Methods Enzymol* **248**, 169-180
39. Robichon, C., Bonhivers, M., and Pugsley, A. P. (2003) *Mol Microbiol* **49**, 1145-1154
40. Robichon, C., Vidal-Ingigliardi, D., and Pugsley, A. P. (2005) *J Biol Chem* **280**, 974-983
41. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S.,
Murphy, L., Oliver, K., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998) Nature 393, 537-544

42. Dobos, K. M., Khoo, K. H., Swiderek, K. M., Brennan, P. J., and Belisle, J. T. (1996) J Bacteriol 178, 2498-2506

43. Michell, S. L., Whelan, A. O., Wheeler, P. R., Panico, M., Easton, R. L., Etienne, A. T., Haslam, S. M., Dell, A., Morris, H. R., Reason, A. J., Herrmann, J. L., Young, D. B., and Hewinson, R. G. (2003) J Biol Chem 278, 16423-16432

44. Sartain, M. J., and Belisle, J. T. (2009) Glycobiology 19, 38-51

45. Akira, S. (2003) Curr Opin Immunol 15, 5-11

46. Fukuda, A., Matsuyama, S., Hara, T., Nakayama, J., Nagasawa, H., and Tokuda, H. (2002) J Biol Chem 277, 43512-43518

47. Hoffmann, C., Leis, A., Niederweis, M., Plitzko, J. M., and Engelhardt, H. (2008) Proc Natl Acad Sci U S A 105, 3963-3967

48. Zuber, B., Chami, M., Houssin, C., Dubochet, J., Griffiths, G., and Daffe, M. (2008) J Bacteriol 190, 5672-5680
**FIGURE LEGENDS**

**Figure 1:** Western blot analysis of LppX
a) total lysates of *M. smegmatis* parental strain, Δ*lspA, ΔlspA-lspA* expressing LppX with hemagglutinin (HA) and hexa-His epitope analysed with mouse anti-hemagglutinin monoclonal antibody; b) amino acid sequence of LppX-HA-His. Inverted commas show trypsin cleavage sites. The arrow shows the +1 cysteine modified by Lgt and Lnt. Bold letters indicate the signal sequence cleaved off by LspA reducing the molecular weight by 2.6 kDa. Italic letters indicate the modified tryptic peptide in the Δ*lspA* mutant, underlined is the N-terminal tryptic peptide found in *M. smegmatis* parental strain, *lnt::aph* mutant and *lnt::aph-ppm1Tb*.

**Figure 2:** Lnt-BLASTp.
a) *E. coli* Lnt was used as a query to identify homologues on the National Center for Biotechnology Information BLASTp server (http://www.ncbi.nlm.gov/sutils/genom_table.cgi). The sequence filtering option was switched off and the expect value was set at 10, the cut-off value set at 10^-4. b) a comparison of the genomic region of Lnt (black) and Ppm1 (grey) homologues in mycobacteria.

**Figure 3:** Western blot analysis of RecLppX
a) total lysates of *M. smegmatis* wildtype, Δ*lspA, ΔlspA-lspA* expressing RecLppX with hemagglutinin and hexa-His epitope analyzed with mouse anti-hemagglutinin monoclonal antibody. b) RecLppX isolated from indicated strains was digested with thrombin and analyzed with mouse anti-hemagglutinin monoclonal antibody.

**Figure 4:** MALDI-TOF analysis of a trypsin digest of purified LppX
MS analysis of LppX tryptic peptides purified from *M. smegmatis* parental strain (1); *M. smegmatis lnt::aph* (2) and *M. smegmatis lnt::aph-ppm1Tb* (3). Filled triangle = diacylglycerol plus N-acyl modified N-terminal peptide; open triangle = diacylglycerol modified N-terminal peptide.

**Figure 5:** MALDI-TOF-TOF analysis of the N-terminal peptides of LppX
MS/MS analysis of N-terminal peptides of LppX purified from *M. smegmatis* parental strain (1), *M. smegmatis lnt::aph* (2) and *M. smegmatis lnt::aph-ppm1Tb* (3). A schematic drawing of the modified +1 cysteine with the cleavage sites of each identified m/z signal is depicted in the upper part of each spectrum. The eliminated fragment of LppX modifications are shown on the right side of spectrum (1): palmitic acid; tuberculostearic acid; diacylthioglycerol; tuberculostearic acid α-thioglyceryl ester; palmitic acid α-thioglyceryl ester; palmitamide. Note that the ester-linked palmitic acid and tuberculostearic acid may be coupled to either position SN1 or SN2. Only one conformation is depicted.

**Figure 6:** MALDI-TOF analysis of peptides resulting from trypsin digestion of LppX from Δ*lspA* mutant
MS analysis of LppX tryptic peptides purified from *M. smegmatis ΔlspA* mutant. Arrows indicate the N-terminal peptides. Asterisk = N-terminal peptide without fatty acid modifications; open triangle = diacylglycerol modified N-terminal peptide.
Table 1: Comparison of m/z values of LppX N-terminal tryptic peptides found in the different mutants

| peptide                  | calculated m/z | parental strain m/z | int::aph m/z | int::aph-ppm1Tb m/z |
|--------------------------|----------------|---------------------|--------------|---------------------|
| CSS...EIR                | 2964.46        | -                   | -            | -                   |
| CSS...EIR + diacylglycerl (C19,C16) | 3557.01 (+ 592.54) | -                   | 3557.01 (+ 592.55) | 3557.07 (+ 592.61) |
| CSS...EIR + diacylglycerl (C19,C16) + N-acyl (C16) | 3795.24 (+ 830.77) | 3795.42 (+ 830.95) | -            | 3795.32 (+ 830.86) |

CSS...EIR corresponds to the N-terminal tryptic peptide of LppX upon cleavage of the signal peptide by LspA (Figure 1b). Mass differences to the corresponding unmodified peptide (upper row) due to modifications are given in brackets. Observed modifications are: diacylglycerl with a C16 fatty acid and tuberculostearic acid (C19:0) (+592.54 Da), plus eventually N-acyl with C16 fatty acid (+238.23 Da, Σ = 830.77 Da).

Table 2: Comparison of experimentally determined eliminations from N-terminal tryptic peptides of LppX in the MALDI-TOF/TOF spectra of the different mutants with theoretically calculated eliminations

| modification              | eliminated fragment | calculated mass of eliminated fragment [Da]* | experimentally determined mass of eliminated fragment in parental strain [Da] | int::aph [Da] | int::aph-ppm1Tb [Da] |
|---------------------------|---------------------|----------------------------------------------|--------------------------------------------------------------------------------|---------------|-----------------------|
| O-linked palmitoyl        | palmitic acid       | 256.24                                       | 255.84                                                                          | 256.36        | 255.68                |
| N-linked palmitoyl        | palmitamide         | 255.26                                       | 255.08                                                                          | -             | 255.80                |
| O-linked tuberculostearyl | tuberculostearic acid | 298.29                                      | 298.92                                                                          | -             | -                     |
| diacylglycerl (C16,C19)   | diacylthioglycerl (C16,C19) | 626.53                                      | 626.22                                                                          | 626.78        | 626.16                |
|                           | C16 fatty acid α-thioglycerl ester | 328.24                                      | 327.30                                                                          | -             | -                     |
|                           | tuberculostearic α-thioglycerl ester | 370.29                                      | 370.38                                                                          | 370.43        | 370.48                |

*: calculated masses are monoisotopic masses
Table 3: Comparison of m/z values of tryptic peptides of pro-LppX containing the +1 cysteine found in the ΔlspA mutant

AVT...EIR corresponds to the tryptic peptide of pro-LppX containing the +1 cysteine (Figure 1b). Mass differences to the corresponding unmodified peptide due to modifications are given in brackets. Observed modifications are: diacylglycerol with C16 and C19 fatty acid (+ 592.54 Da), β-mercaptoethanyl (+ 76.00 Da).

| Peptide                                      | Calculated m/z | ΔlspA m/z     |
|----------------------------------------------|----------------|---------------|
| AVT...EIR with unmodified cysteines          | 4888.55        | -             |
| AVT...EIR with S-S                           | 4886.53        | 4887.30       |
| AVT...EIR                                   | 5040.54(+ 152.00) | 5041.23(+ 152.68) |
| AVT...EIR                                   | 5557.09(+ 668.54) | 5558.12(+ 669.57) |
Figure 1

(a) 

[Image: A gel showing bands labeled Pro-LppX-HA-His and LppX-HA-His.]

(b) 

\[ \text{MNDGK'RAVTSAVVLVGACALWLGCSNPK} \]
\[ \text{PDAAEQYQYPTASPDALLAEIR} \]
\[ \text{QSLDATK'GL} \]
\[ \text{TSVHAVR'TTGK'VDSLPGITSADVVR'ANPL} \]
\[ \text{AAK'GVCYNEQVFPR'VQGDNISVK'LFDD} \]
\[ \text{WSNLGSISELSTR'VLDPAGTQLSNGTNL} \]
\[ \text{QAGTEVIDGISTTK'TGTIPASSVK'MLDGPA} \]
\[ \text{K'SARPATWIAQDGSHHLVR'ASIDLGSQILT} \]
\[ \text{QSK'WNEPVNVDYDVPDYATWHHHHHH} \]

HA-tag \quad \text{His-tag}
Figure 3

(a) 

Figure showing molecular weight markers (kDa) with lanes labeled as wt, ΔiSpA clone 1, ΔiSpA clone 2, ΔiSpA ΔiSpA, and bands indicated as Pro-RecLppX and RecLppX.

(b) 

Figure showing molecular weight markers (kDa) with lanes labeled as wt, int::aph, int::aph::intMs, int::aph::pp1Tb, int::aph::pp1Tb, wt, ΔiSpA, ΔiSpA ΔiSpA.
