The molecular mechanism of bacterial lipoprotein modification—how, when and why?
Nienke Buddelmeijer

To cite this version:

HAL Id: pasteur-01407690
https://hal-pasteur.archives-ouvertes.fr/pasteur-01407690
Submitted on 2 Dec 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives| 4.0 International License
The molecular mechanism of bacterial lipoprotein modification - how, when and why?

<table>
<thead>
<tr>
<th>Journal:</th>
<th>FEMS Microbiology Reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>FEMSRE-14-10-0036.R2</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Review Article</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>n/a</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Buddelmeijer, Nienke; Institut Pasteur, Microbiology</td>
</tr>
<tr>
<td>Keywords:</td>
<td>lipoprotein, fatty acid acylation, bacteria, acyltransferase, signal peptidase</td>
</tr>
</tbody>
</table>
The molecular mechanism of bacterial lipoprotein modification – how, when and why?

Nienke Buddelmeijer

Institut Pasteur, Biology and Genetics of the Bacterial Cell Wall Unit, Inserm Group Avenir, 28 rue du docteur Roux, Paris F-75724 cedex 15, France

Abstract

Posttranslational modification of proteins by lipidation is a common process in biological systems. Lipids provide protein stability, interaction with other membrane components, and in some cases, due to reversibility of the process, a mechanism for regulating protein localization and function. Bacterial lipoproteins possess fatty acids at their amino-termini that are derived from phospholipids and this lipid moiety anchors the proteins into the membrane. These lipids, as is the case for lipopolysaccharides and lipoteichoic acids, play an important role in signaling of the innate immune system through the interaction with Toll-like receptors. Over the past three years tremendous progress has been made in understanding the mechanism by which lipoproteins become lipated. Advanced methodology in mass spectrometry, proteomics and genome-wide analyses allowed precise characterization of lipoprotein modifications and the identification of the enzymes catalyzing the reactions in diverse bacterial species. This review will highlight new findings on bacterial lipoprotein modification with focus on the reaction mechanisms and the role of lipoproteins in cell envelope homeostasis.

Introduction
Bacterial lipoproteins are exported proteins and often exposed on the bacterial cell surface and include an important class of virulence factors expressed by many pathogenic species (Kovacs-Simon et al., 2010). Prolipoprotein is synthesized in the cytoplasm and translocated via Sec or Tat secretion machineries (Inouye et al., 1977, Hayashi & Wu, 1985, Gibbons et al., 2007, Siboo et al., 2008, Shruthi et al., 2010, Thompson et al., 2010). Lipoprotein modification occurs in the cytoplasmic (inner) membrane by the sequential action of three membrane-bound enzymes (Fig. 1) (Wu et al., 1982). Lipoproteins possess a conserved sequence, so-called lipobox, which allows specific recognition by the modification machinery. The invariant cysteine residue in the lipobox becomes lipidated and the first amino acid after complete maturation of lipoprotein. Phosphatidylglycerol:prolipoprotein diacylglyceryl transferase (Lgt) adds an sn-1,2-diacylglyceryl group from phosphatidylglycerol (PG) to the sulfhydryl group of cysteine via a thioether bond. Prolipoprotein signal peptidase (Lsp) cleaves the signal peptide thereby liberating the α-amino group of diacylglyceryl-cysteine. The third and last step is N-acylation of diacylglyceride-cysteine by Apolipoprotein N-acyltransferase (Lnt), resulting in mature N-acyl-S-diacylglyceryl-cysteine linked proteins. Lgt and Lsp are conserved in all bacterial species, while Lnt has only been identified in proteobacteria and actinomycetes. Recent findings suggest that a non-canonical lipidation pathway exist by the demonstration that lipoproteins of firmicutes and terenicutes (that include mollicutes), whose genome lack a gene encoding Lnt, exist in a triacylated form (Fig. 1) (Jan et al., 1995, Kurokawa, 2009, Asanuma et al., 2010, Serebryakova et al., 2011). Various N-terminal modifications were observed: N-acyl-S-diacylglyceryl-cysteine, but also N-acyl-S-monoacyl-cysteine, N-acetyl-S-diacylglyceryl-cysteine and N-peptidyl-S-diacylglyceryl-cysteine (Nakayama et al., 2012). Correct lipoprotein modification and consequently, the enzymes that carry them out, are involved in a process essential for bacterial viability in proteobacteria and some
actinomycetes (Table 1). Understanding the molecular details of their catalytic mechanism and function should provide detailed knowledge on bacterial envelope homeostasis and facilitate the discovery of new antibacterial agents targeting this essential pathway. Most enzymatic and functional studies have been performed with the lipoprotein modification enzymes of Escherichia coli that are discussed in this review. It is stated in the text when enzymes from other bacterial species are described.

Phospholipids are substrates for acyltransferases involved in lipoprotein modification

The lipid moiety of lipoproteins is derived from membrane phospholipids, linking the lipoprotein modification pathway to phospholipid biogenesis (Lai et al., 1980). Two different systems exist in bacteria to transfer fatty acids from acyl-acyl carrier proteins (acyl-ACP), the end products of the fatty acid synthase II pathway (FASII), to glycerol-3-phosphate. The PIsX/Y/C is a widespread pathway while the PIsB/C system is found predominantly in γ-proteobacteria, but both systems result in the synthesis of phosphatidic acid, the building block of phospholipids. Phosphatidic acid is subsequently converted into the key intermediate CDP-diacylglycerol. Two pathways then lead to the synthesis of phosphatidylethanolamine (PE) (via phosphatidylserine) and phosphatidylglycerol (PG) and cardiolipin (CL), respectively (for a recent review see (Parsons & Rock, 2013)). Movement of lipids, including PE and lipopolysaccharides (LPS), across the cytoplasmic membrane is dependent on MsbA (Zhou et al., 1998, Doerrler & Raetz, 2002). The ABC transporter MsbA functions as a dimer in which each monomer is composed of six transmembrane segments (TMS) and a nucleotide-binding domain (Ward et al., 2007). The dimer hydrolyses ATP and then undergoes large conformational changes, which is thought to lead to the trans-bilayer movement of lipids across the membrane (Doerrler et al., 2001, Doerrler et al., 2004, Doshi & van Veen, 2013). The main phospholipids of E. coli and Salmonella enterica serovar Typhimurium are
phosphatidylethanolamine (PE), phosphatidylglycerol (PG), cardiolipin (CL) and minor phospholipids phosphatidylserine (PS) and phosphatidic acid (PA) (Ames, 1968). The fatty acids of phospholipids of a typical *E. coli* strain contain about 45% palmitic acid (C16:0), 2% myristic acid (C14:0), 35% palmitoleic acid (C16:1) and 18% cis-vaccenic acid (C18:1cis-11) (Cronan & Rock, 1996). The main fatty acids of major lipoprotein Lpp (murtein-lipoprotein or *Braun’s lipoprotein*) of *E. coli* are palmitic acid (53%), cis-vaccenic acid (20.7%), 9,10-methylene-hexadecanoic acid (C17:0) (10.6%) and palmitoleic acid (9.4%) (Hantke & Braun, 1973), corresponding to the fatty acid moieties of phospholipids. Phospholipids that compose membranes of *Staphylococcus aureus* include phosphatidylglycerol, cardiolipin, lysyl-phosphatidylglycerol, and mono- and diglucosyl-diglyceride but not phosphatidylethanolamine (White & Frerman, 1968, Fischer, 1994). Mass spectrometry data has shown that lipoproteins in *S. aureus* are modified with C17:0 to C20:0 at the sn-1 position of the diacylglyceryl group and with C15:0 at the sn-2 position. The α-amino group of lipidated diacylglyceryl-cysteine have fatty acids of varying lengths, from C15:0 to C20:0 (Kurokawa *et al.*, 2012). The lipid moiety of lipoproteins thus reflects the fatty acid composition of phospholipids and varies among bacterial species.

**Phosphatidylglycerol:prolipoprotein diacylglyceryl transferase (Lgt)**

Modification of prolipoprotein by Lgt is the first step in lipoprotein modification (Tokunaga *et al.*, 1982). Upon membrane insertion of prolipoprotein by Sec or Tat, Lgt transfers the diacylglyceryl group from phosphatidylglycerol onto the sulfhydryl group of the invariant cysteine in the lipobox resulting in a thioether-linked diacylglyceryl-prolipoprotein and a sn-glycerol-1-phosphate by-product (Lai *et al.*, 1980, Sankaran & Wu, 1994). Signal sequences of lipoproteins share common features including one or several positively charged residues at the amino terminus followed by a hydrophobic stretch of about 20 amino acids that includes
the lipobox. Based on studies with Lpp, the consensus sequence of the lipobox was
determined as: (L)_{-3}-(A/S)_{-2}-(G/A)_{-1}C_{+1}, in which the lipid-modified cysteine at position +1
becomes the first amino acid in the mature protein (Wu et al., 1983, Wu, 1996, Madan Babu
& Sankaran, 2002, Babu et al., 2006). This consensus sequence has been documented in the
PROSITE database under pattern entry PS00013/PDOC0013
(http://www.expasy.ch/prosite/). Lipobox motifs are similar in firmicutes but more variable in
spirochetes (Haake, 2000, Sutcliffe & Harrington, 2002), the -3 position in particular is more
loosely defined in the latter, where a leucine residue is often replaced by another hydrophobic
amino acid. A high percentage of ORFs in spirochetes encode lipoproteins, some of which are
highly abundant proteins (Haake, 2000). For Lpp it has also been demonstrated that residue
G14 in the signal peptide sequence is important for lipidation by Lgt, and that a signal
sequence shortened by several residues still leads to modified protein (Wu et al., 1977, Yem
first biochemical characterization of Lgt was reported in the 1980-s using crude membrane
extracts. The optimal pH and reaction temperature for Lgt are pH 7.8 and 37°C, respectively
(Tokunaga et al., 1984). Lgt is stable in membrane vesicles up to 65°C, the enzyme is active
in N-octyl-β-D-glucoside but many other detergents are not compatible with enzyme activity
(Tokunaga et al., 1984, Sankaran & Wu, 1994). PG is specific as acyl donor since PE and CL
could not be used as substrate in the in vitro assay using a synthetic peptide, suggesting that
the polar headgroup of the phospholipid plays an important role in the Lgt reaction (Sankaran
& Wu, 1994). The gene encoding Lgt was identified by screening a collection of temperature
sensitive mutants of S. enterica for the accumulation of unmodified Lpp and was shown to be
located at the 5’ end of thyA (Williams et al., 1989, Bell-Pedersen et al., 1991, Gan et al.,
1993). The enzyme is localized in the cytoplasmic membrane and spans the membrane by
seven transmembrane segments (Fig. 2.) (Pailler et al., 2012). The majority of essential
residues are located in the membrane, however, it is currently unknown how Lgt interacts with its substrates PG and prolipoprotein (Qi et al., 1995, Sankaran et al., 1997, Pailler et al., 2012). Residues H103 and Y235 were identified as conserved by comparing four Lgt sequences and were shown to be essential for Lgt activity (Sankaran et al., 1997). These residues are among twelve residues highly conserved in proteobacteria, firmicutes and actinobacteria (Pailler et al., 2012). A more recent analysis on conserved residues highlighted the existence of a so-called Lgt signature motif; two amino acids in this sequence (N146 and G154) were shown to be absolutely required for function (Pailler et al., 2012). Together, these data show that regions H103-G108 and G142-G154 and residues Y26 and Y235 are critical for Lgt function. Initial kinetics studies were performed with crude membrane extracts but further biochemical characterization of the *E. coli* enzyme turned out difficult due to overproduction defects and instability of the purified enzyme (Sankaran et al., 1997, Selvan & Sankaran, 2008, Pailler et al., 2012). Lgt has been purified recently from *Lactococcus lactis* with which kinetic studies were performed (Banerjee & Sankaran, 2012), unfortunately, assay conditions were not comparable with those described previously (Sankaran et al., 1997, Sundaram et al., 2012), making it difficult to compare enzyme activities from different bacterial species.

Lgt of *S. aureus* can complement a temperature sensitive mutant of *E. coli*, suggesting that Lgt is functionally conserved among bacterial species (Qi et al., 1995). Deletion mutants of *lgt* in firmicutes often lead to shedding of lipoproteins into the culture supernatant in non-lipidated precursor form with uncleaved signal peptide (Leskelä et al., 1999) or processed by an unidentified signal peptidase (Stoll et al., 2005, Denham et al., 2009) or Lsp (Baumgärtner et al., 2007, Henneke et al., 2008). Two Lgt homologues were identified in *Streptomyces coelicolor*, Lgt1 and Lgt2, which can both complement the *lgt* mutant of *Streptomyces scabies*, suggesting that both enzymes are functional enzymes (Widdick et al., 2011).
double mutant could not be obtained suggesting that Lgt activity is essential in *S. coelicolor* (Thompson *et al.*, 2010). Lipoproteins were purified from the complementing strains and the composition of their N-terminal fatty acid was analyzed by mass spectrometry (Widdick *et al.*, 2011). Several fatty acid acylated peptides were missing in these strains compared to wild type *S. scabies*, suggesting that subtle differences in diacylglyceryl transferase activity might exist. The gene encoding Lgt in *Mycobacterium tuberculosis* and two genes in *Mycobacterium smegmatis* have recently been identified (Tschumi *et al.*, 2012). The MSMEG_3222 gene was shown to be responsible for the transfer of diacylglyceryl onto lipoproteins, and although less conserved than MSMEG_3222 the role of the second *lgt* gene (MSMEG_5408) in lipoprotein modification has not yet been addressed (Tschumi *et al.*, 2012). An *lgt* deletion mutant could not be obtained in *M. tuberculosis* suggesting that this gene is essential (Sassetti *et al.*, 2003; Tschumi *et al.*, 2012). Strikingly, MSMEG_3222 is not essential in *M. smegmatis* and deletion did not completely abolish incorporation of radioactive palmitate suggesting that either MSMEG_5408 or another, as yet unknown, enzyme possesses Lgt activity but is perhaps less efficient than MSMEG_3222 (Tschumi *et al.*, 2012). Absence of Lgt resulted in massive release of lipoproteins in the culture media but none of these proteins contained a signal peptide, suggesting that they are processed by Lsp or other signal peptidases (Tschumi *et al.*, 2012). No clear phenotype on cell wall biogenesis was observed.

Detailed biochemical analyses and comparison of kinetic parameters of various Lgt enzymes will contribute to a better understanding of the mechanism of the reaction and are predicted to highlight differences and similarities in enzyme activity. The thiol group of prolipoprotein probably performs a nucleophilic attack on the carbonyl group of PG where Lgt provides the catalytic residues to support transfer of diacylglycerol onto prolipoprotein, and positions the two substrates close together. PG is the sole acyl donor in the Lgt reaction, suggesting that the glycerol-phosphate headgroup plays an important role in substrate recognition by Lgt.
Lipoprotein signal peptidase (Lsp)

Lsp (SPase II) is localized in the cytoplasmic membrane (Tokunaga et al., 1984) and spans this membrane by four transmembrane segments with both the amino and carboxyl termini facing the cytoplasm (Munoa et al., 1991) (Fig. 2). Five conserved regions were identified, i.e. region I (amino acids residues 23-27 in \textit{E. coli} sequence), region II (residues 53-58), region III (residues 104-115) and region IV (residues 121-123) and region V (residues 137-141) (Munoa et al., 1991, Tjalsma et al., 1999). The role of these conserved regions in the peptidase reaction has not yet been addressed but probably plays a role in substrate recognition and/or binding. The active site of Lsp is composed of at least two aspartate residues (D114 and D141 in \textit{E. coli}) (Tjalsma et al., 1999). Modification of lipoproteins by addition of diacylglycerol catalyzed by Lgt is a prerequisite for processing by Lsp in proteobacteria (Tokunaga et al., 1982). The Lsp enzymes of actinomycetes have an N-terminal extension, truncations of which led to loss of Lsp function, either due to inactivation or instability of the enzyme (Thompson et al., 2010). Biochemical studies with Lsp were performed in the early 1980-s and showed that Lsp is temperature resistant in isolated membranes, but unlike Lnt still active at 80°C (Hussain et al., 1982). In detergent-solubilized membranes Lsp is active at a reaction temperature between 37°C-45°C with a pH optimum of 7.9 (Tokunaga et al., 1984). The enzyme does not require phospholipids and is active in non-ionic detergents.

Initially two enzymatic activities were shown to be involved in the proteolytic cleavage of lipoproteins, Lsp that cleaves off the signal peptide of diacylglycerol-proLpp and a signal peptide peptidase responsible for the degradation of signal peptide (SppA or Protease IV) (Suzuki et al., 1978, Novak & Dev, 1988). SppA is an intramembrane protease (Pacaud, 1982, Pacaud, 1982) utilizing a Ser/Lys dyad for catalysis (Wang et al., 2008). It adopts a
tetrameric structure with a hydrophobic interior providing access to the catalytic site (Ichihara et al., 1986, Kim et al., 2008) and cleaves its substrate in the hydrophobic region (Novak & Dev, 1988). Recently, Saito et al. (Saito et al., 2011) provided evidence that peptidase RseP (YaeL in E. coli), which belongs to the zinc-metalloprotease group (S2P) of the intramembrane-cleaving proteases (I-CLiPs) family of proteases (Urban, 2009), is responsible for cleavage of signal peptides including those obtained after cleavage by Lsp. RseP is not the exclusive signal peptide peptidase since some signal peptides were cleaved in the absence of RseP and SppA. The sole natural substrate of RseP is the anti-σE factor RseA that upon cell envelope stress is cleaved by DegS and further processed by RseP resulting in activation of σE and the expression of stress-induced genes (Alba et al., 2002). RseP has four transmembrane segments with its N- and C-termini facing the periplasm (Kanehara et al., 2001). The periplasmic PDZ domain is specific for regulated processing of RseA after truncation by DegS. Removal of the PDZ domain led to intramembrane cleavage of RseA (Hizukuri et al., 2014), suggesting that this domain is not required for processing of lipoprotein signal peptides.

Lsp proteins from firmicutes can complement an lsp temperature sensitive mutant of E. coli, suggesting that the enzymes are functionally conserved as seen for Lgt (Pragai et al., 1997), however kinetic analyses have not yet been performed to determine specificity and mechanism of the reaction. Lsp does not process prolipoprotein without S-diacylglyceryl modification in proteobacteria, indicating that the enzyme specifically recognizes the fatty acid acylated cysteine. Processing of Lpp occurs both in E. coli and Bacillus subtilis but at different rates (Hayashi et al., 1985). Apo-Lpp accumulates in B. subtilis probably due to inefficient N-acylation or N-acetylation of apoLpp (Kurokawa et al., 2012). Furthermore, the two organisms differ in their susceptibility to globomycin as shown by the processing of penicillinase (Hayashi & Wu, 1983). Myxococcus xanthus possesses four lsp genes, two of
which are part of the antibiotic TA biosynthetic cluster (Xiao & Wall, 2014). The Lsp enzymes can complement an lsp depletion strain of E. coli only in the absence of Lpp suggesting that Lsp-s of M. xanthus are less efficient than Lsp of E. coli. Lipoproteins of Streptococcus uberis are processed by Eep, a homologue of enhanced expression of pheromone from Enterococcus faecalis (An et al., 1999), in the absence of Lsp in late-stationary-phase cells (Denham et al., 2009).

Structural information is not available for any of the lipoprotein modification enzymes, including Lsp. Other bacterial signal peptidases, such as signal peptidase I (leader peptidase) and type IV pilin signal peptidase (prepilin peptidase), share their membrane-embedded characteristics and the hydrophobic nature of their substrates. The catalytic mechanism (Paetzel, 2014) and the Ser/Lys active site (Wang et al., 2008) of leader peptidase have been reported. A soluble version of the enzyme, lacking its two N-terminal transmembrane segments, is active in vitro (Kuo et al., 1993), although slightly less compared to full-length protein and it requires detergents or phospholipids for its activity (Tschantz et al., 1995). Several crystal structures of the soluble form have been solved (Kim et al., 2008) and a recent NMR structure suggests that the signal peptide remains in the bilayer in an unstructured conformation and that the active site of leader peptidase functions at the membrane surface (De Bona et al., 2012) or well within the lipid bilayer (Paetzel, 2014). The membrane is probably important in stabilizing the signal peptide while under reconstituted conditions the detergents or phospholipids would serve an analogous role. Several inhibitors have been identified including penem-type inhibitors (Allsop et al., 1996) and lipopeptides (Paetzel et al., 2002). Differences in specificity between leader peptidases from different species may be explained by structural variations in the active site (Auclair et al., 2012). Prepilin peptidase is, like Lsp, an aspartic protease that belongs to the class of aspartyl protease I-CLiPs (Erez et al., 2009) and is involved in processing of pre-pilin proteins (LaPointe & Taylor, 2000). I-
CLiPs are involved in intramembrane proteolysis liberating membrane-tethered domains of membrane proteins that can function as regulatory molecules, and are found in many different organisms (Erez et al., 2009). Prepilin peptidase is an integral membrane protein and processing of prepilin has been postulated to occur at the cytoplasmic side of the membrane, allowing the subsequent addition of methyl groups onto phenylalanine (Strom & Lory, 1987). A topology model was proposed based on data obtained with reporter fusion proteins and topology prediction programs; prepilin peptidases have eight transmembrane segments with their N- and C-termini are located in the periplasm, and the catalytic Asp dyad is facing the cytoplasm (LaPointe & Taylor, 2000). The crystal structure of aspartyl protease FlaK, has recently been solved (Hu et al., 2011). The two active site Asp residues are located on top of two α-helices, housed in a hydrophilic cavity in the membrane and facing the cytoplasm, similar to the structure of presenilin (Hu et al., 2011, Lu et al., 2014). Lsp probably shares structural features with both leader peptidase and aspartyl proteases. It interacts with diacylglyceryl-prolipoprotein as leader peptidase with signal peptides of secreted proteins in such a way that the diacylglyceryl-cysteine is accessible to the Asp dyad active site as in aspartyl proteases, with the only difference that the active site is facing the periplasm.

Globomycin – a specific inhibitor of Lsp

Globomycin was first identified in 1978, as an antibiotic produced by various Streptomyces strains with spheroplast-forming (global-shape) activity on E. coli (Inukai et al., 1978, Inukai et al., 1978). It specifically inhibits proteobacteria such as E. coli (MIC 0.4 µg/ml for E. coli B) but not Pseudomonas aeruginosa. It is also less active against firmicutes (for example, MIC >100 µg/ml for S. aureus and B. subtilis) (Inukai et al., 1978). Globomycin is a lipophilic cyclic peptide with molecular formula C_{32}H_{57}N_{5}O_{2} and a molecular weight of 655 (Nakajima et al., 1978) (Fig. 3). Indications that the target of globomycin is Lsp were
reported in the last paper in a series of articles on globomycin discovery and its characterization by Mamoru Arai’s group (Inukai et al., 1978). It was later shown that diacylglycerol-proLpp that accumulated in the cytoplasmic membrane in cells grown in the presence of globomycin (Hussain et al., 1980) could be converted into mature protein in vitro when cell envelope fractions were incubated under optimal conditions. The gene encoding Lsp was identified in E. coli by selecting clones from the Carbon-Clarke collection for globomycin resistance, or by screening for complementation of a strain that is temperature sensitive in the processing of diacylglycerol-proLpp (Tokunaga et al., 1983, Yamagata et al., 1983). Kinetics studies showed that globomycin inhibits Lsp in a non-competitive manner with a $K_i$ of 36 nM, and a $K_m$ value of 6 $\mu$M was determined for diacylglycerol-prolipoprotein substrate (Dev et al., 1985). Globomycin is mycobactericidal at elevated concentrations but killing is independent of Lsp in M. tuberculosis, suggesting an alternative pathway of inhibition in this bacterium (Banaiee et al., 2007). Globomycin analogues have been described that are more potent against proteobacteria and exhibit antibacterial activity against firmicutes (Kiho et al., 2003, Kiho et al., 2004). A long alkyl side chain, the lipophilic part of the molecule, increases activity and also shows moderate activity against S. aureus. The hydroxyl group of L-Ser is essential but not that of L-allo-Thr (Kiho et al., 2004). The lipid moiety of globomycin is probably required to anchor the antibiotic into the membrane to allow correct positioning of the peptidic part for interaction with the enzyme, but also seems to play a direct role in globomycin-enzyme interaction. Differences in membrane composition and/or membrane topology of Lsp might explain differences in antibacterial activities of globomycin and its analogues on Lsp of different bacterial species. Another antibiotic, TA or myxovirescin isolated from Myxococcus xanthus, has been shown to target Lsp (Xiao et al., 2012). Antibiotic TA is a cyclic macrolactam lactone that is active against proteobacteria and some firmicutes (Rosenberg et al., 1973).
Apolipoprotein N-acyltransferase (Lnt)

The membrane localization of Lnt of *E. coli* was demonstrated by Gupta and Wu in 1991 (Gupta & Wu, 1991), in the same study it was demonstrated that Lnt has a broad pH optimum (pH 6.5-7.5) and that it requires the presence of detergents for its activity *in vitro*. In 2005 the membrane topology of Lnt and the phenotype of an *lnt* depletion mutant was reported (Robichon *et al.*, 2005). Several essential residues in Lnt were identified that are mainly located in the catalytic periplasmic domain and include the catalytic triad E-K-C (Vidal-Ingigliardi *et al.*, 2007). It was shown that growth of the *lnt* depletion mutant could be restored by Lnt homologues of proteobacteria, but not actinobacteria (Vidal-Ingigliardi *et al.*, 2007). The reaction catalyzed by Lnt involves two steps. In the first step, the sn-1 of PE reacts with the active site cysteine of Lnt resulting in the formation of a thioester-acyl enzyme intermediate and a lysophospholipid by-product (Jackowski & Rock, 1986, Buddelmeijer & Young, 2010, Hillmann *et al.*, 2011). In the second step, the acyl group is transferred to apolipoprotein resulting in triacylated, mature lipoprotein (Buddelmeijer & Young, 2010). Kinetics studies showed that Lnt follows a ping-pong mechanism whereby lyso-phospholipid produced in the first reaction is released before the second substrate, apolipoprotein, is bound and modified. With PE and a synthetic diacylated peptide (FSL-1) a $K_m$ value of 8.0 $\mu$M and a $V_{\text{max}}$ value of 350 $\mu$mol$\text{FSL-1}$ min$^{-1}$ $\mu$mol$^{-1}$ Lnt were determined, corresponding to a specific activity of 5.6 $\mu$mol$\text{FSL-1}$ min$^{-1}$ mg$^{-1}$ Lnt. Interestingly, similar kinetics were obtained with PagP that catalyzes the transfer of an acyl group from PE to lipid A (Bishop *et al.*, 2000). Although acyltransferases Lnt and PagP are involved in completely different pathways, their substrates are highly abundant; 5 x 10$^5$ molecules of Lpp and 2 x 10$^6$ molecules of lipid A per cell (Braun & Rehn, 1969, Raetz, 1986). Furthermore, similar $K_m$ values have been reported for
Lsp and Lnt of E. coli for diacylglyceryl-proLpp and apo-lipopeptide, respectively (Dev et al., 1985, Hillmann et al., 2011).

Lnt is highly efficient in the second step of the reaction but less in the first step: $k_{cat}/K_m$ of $4.8 \times 10^3$ for PE versus $7.3 \times 10^5$ for FSL-1. Based on the kinetic data, the number of Lnt molecules per cell was estimated between 100 and 200, indicating that the $N$-acyltransferase reaction is fast and efficient since lipoproteins are very abundant. PE is the preferred substrate with an optimal fatty acid chain length of C16-18, more specifically palmitate or oleate with a saturated acyl chain in position $sn$-1 and a non-saturated acyl chain at $sn$-2 (Jackowski & Rock, 1986, Hillmann et al., 2011). Phospholipids with small polar headgroups can also serve as acyl donor (Gupta et al., 1991, Hillmann et al., 2011). PagP and LpxA, which is involved in the first step of lipopolysaccharide biosynthesis, possess a carbon-ruler that makes these enzymes highly specific for the chain length of the acyl donor substrate (Wyckoff et al., 1998, Ahn et al., 2004, Williams & Raetz, 2007). It is unlikely that Lnt has carbon ruler capacity since the acyl chain length and composition of the phospholipid substrate are variable. On the other hand, the $S$-diacylglyceryl moiety of the apolipoprotein substrate probably determines substrate specificity due to variations in lipid composition depending on the bacterial species.

The gene encoding Lnt was first identified in S. enterica by screening the same library as described for lgt for the accumulation of apol-Lpp (Gupta et al., 1993). The mutation causing the temperature sensitive defect in Lnt was a substitution of glutamate 435 to lysine; a similar mutation in E. coli did not result in a temperature sensitive phenotype (Robichon et al., 2005). Lnt activity in Actinomycetes was first reported in M. smegmatis (Tschumi et al., 2009). A strain of M. smegmatis deleted for lnt was viable and resulted in the accumulation of the apolipoprotein form of two model lipoproteins LppX and LprF, directly demonstrating its $N$-acyltransferase activity (Tschumi et al., 2009, Brülle et al., 2010). Lnt of mycobacteria transfers either palmitate or tuberculostearic acid (in M. bovis) from phospholipids onto
apolipoprotein (Brülle et al., 2013). Phospholipids of mycobacteria possess mainly octadecanoic acid and tuberculostearic acid at the sn-1 position and palmitate at the sn-2 position (Okuyama et al., 1967), suggesting that Lnt has different substrate specificity in these species. Lnt of \textit{M. smegmatis} was unable to complement a conditional \textit{lnt} null mutant of \textit{E. coli}, suggesting that indeed these enzymes have different substrate specificity. Furthermore, the enzyme was correctly inserted into the cytoplasmic membrane and shown to have a similar membrane topology as Lnt of \textit{E. coli} (Baulard et al., 2003). Among the essential residues in Lnt of \textit{E. coli}, W237 and Y388 are not conserved in Lnt of \textit{M. smegmatis} (Vidal-Ingigliardi et al., 2007, Tschumi et al., 2009). These residues were shown to be involved in the second step of the reaction, \textit{N}-acylation of apolipoprotein, suggesting they might play a role in substrate specificity (Buddelmeijer & Young, 2010). The \textit{lnt} gene in \textit{M. smegmatis} (MSMEG\_3860 or ppm2) is located adjacent to \textit{ppm1} encoding polyprenol monophosphomannose (Ppm) synthase that is involved in the transfer of mannose from GDP-mannose to polyprenol phosphate, an important step in the biosynthesis of mannose-containing lipoglycans (Gurcha et al., 2002). Ppm synthase is essential for viability in \textit{M. smegmatis} (Rana et al., 2012).

In \textit{M. tuberculosis ppm1} (Rv2051c) encodes a two-domain protein, in which the N-terminal domain is Lnt and the C-terminal domain Ppm synthase. The Lnt domain of Ppm1 enhances Ppm synthase activity (Gurcha et al., 2002). In mycobacteria where Lnt and Ppm synthase are encoded by two genes, the proteins were shown to interact (Baulard et al., 2003). In both cases Lnt is thought to stabilize the Ppm synthase (Gurcha et al., 2002, Baulard et al., 2003). A second \textit{lnt}-like gene could be identified in \textit{M. tuberculosis} (Rv2261c and Rv2262c) and \textit{M. bovis} (BCG\_2279c), but the active site cysteine residue is replaced by serine (Brülle et al., 2013). The role of these enzymes in lipoprotein modification is currently unknown.
Lipoproteins in Mycobacteria and Streptomyces are often, in addition to being acylated, glycosylated (Espitia & Mancilla, 1989, Herrmann et al., 1996, Gonzalez-Zamorano et al., 2009, Sartain & Belisle, 2009, Wehmeier et al., 2009, Brülle et al., 2010). Recent findings demonstrated a direct correlation between N-acylation and glycosylation of lipoproteins in Corynebacterium glutamicum (Mohiman et al., 2012). Lnt (Cg-Ppm2) of C. glutamicum catalyzes N-acylation of apolipoproteins by addition of a palmitoyl moiety similar to M. smegmatis (Tschumi et al., 2009, Mohiman et al., 2012). Lipid modifications in C. glutamicum included a diacylglyceryl moiety (C16:0, C18:1) and a N-palmitoyl group (C16:0) similar to lipoproteins of E. coli (Mohiman et al., 2012). N-acylation of lipoproteins was not affected in a Cg-ppm1 mutant but proteins were no longer O-glycosylated with hexose units, indicating that Ppm synthase is required for glycosylation of lipoproteins in C. glutamicum. In a Cg-ppm2 mutant lipoproteins are only diacylated and not glycosylated (Mohiman et al., 2012). A Cg-ppm1 deletion mutant does not synthesize lipoglycans, furthermore, their biosynthesis is also affected in the absence of Cg-ppm2 (Lnt), probably because Ppm synthase is less stable under these conditions (Gibson et al., 2003). Interestingly, Ppm1 of M. tuberculosis is able to complement both N-acylation and glycosylation defects in C. glutamicum (Mohiman et al., 2012) and to restore N-acylation in M. smegmatis (Tschumi et al., 2009). The Lnt domain of Ppm1 alone does not restore N-acylation but does restore glycosylation in C. glutamicum. This strengthens the interpretation that the activities of Lnt and Ppm synthase are tightly linked and illustrate that these enzymes operate in a common pathway implicated in envelope biosynthesis.

All Streptomyces species have two lnt genes (Widdick et al., 2011). In S. scabies, Lnt1 is responsible for N-acylation of lipoproteins since lipoproteins are diacylated in lnt1 and lnt1lnt2 mutants. In the absence of Lnt2, however, both di- and triacylated proteins were
detected, suggesting a role for Lnt2 in efficient N-acylation of lipoproteins (Widdick et al., 2011). Since these bacteria undergo different developmental stages, gene expression and activity of lipoprotein modifying enzymes might be tightly regulated depending on environmental conditions. In this respect it is interesting to note that the degree of fatty acid acylation varies greatly in firmicutes depending on growth conditions. For example, diacylated lipoproteins accumulated in S. aureus during late-exponential growth at low pH (Kurokawa et al., 2012). New protein synthesis is required for both accumulation of diacylated proteins and N-acylation of these proteins during pH upshift. These data suggest that regulation of lnt expression or pH-mediated enzyme activity control N-acylation of lipoproteins. Assuming that phospholipids are also substrate for N-acylation in firmicutes, Lnt-like enzymes of these species probably have different specificities for the acyl-donor substrate than those of proteobacteria.

After completion of the Lnt reaction, the lysophosphatidylethanolamine, or 2-acylglycerophosphoethanolamine (2-acyl-GPE) by-product can either be degraded by lysophospholipase PldB (Hsu et al., 1991) or recycled in the synthesis of phospholipid via the acyl-acyl carrier protein synthetase/2-acylglycerophosphoethanolamine acyltransferase pathway (Cooper et al., 1989, Jackowski et al., 1994). 2-acyl-GPE first needs to be translocated to the inner leaflet of the cytoplasmic membrane, a process catalyzed by LplT, before it can be used as a substrate for PL synthesis (Harvat et al., 2005). The Aas enzyme is a cytoplasmic membrane protein (Cooper et al., 1989) that contains both enzymatic activities with acyltransferase located at the N-terminal end and acyl ACP synthetase located at the C-terminal end (Jackowski et al., 1994). In the first step of the reaction acyl-ACP synthetase adds saturated fatty acids onto ACP and in the second step fatty acid is transferred to the sn-1 position of 2-acyl-GPE by the acyltransferase resulting in the formation of new PE.
Alternatives for removal of 2-acyl-GPE are the production of PE and GPE from two molecules of GPE (Homma & Nojima, 1982), and another phospholipase that degrades 2-acyl-GPE (Doi & Nojima, 1975).

Sorting of lipoproteins to the outer membrane

The majority of lipoproteins reside in the inner leaflet of the outer membrane (Tokuda et al., 2007), although recent examples of surface-exposed lipoproteins have been reported (see below). They reach this cellular location by a specific lipoprotein outer membrane localization pathway (Lol) that is composed of five proteins: a cytoplasmic ATP-ase LolD, two cytoplasmic membrane proteins LolC and LolE, a periplasmic chaperone LolA and an outer membrane receptor LolB (Fig. 4). Sorting and translocation of lipoproteins by Lol is dependent on the lipoprotein-sorting signal determined by specific residues located C-terminally to the diacylglyceryl-cysteine. The Lol pathway has been described in great detail in several excellent reviews (Tokuda et al., 2014, Zuckert, 2014) and will only be briefly discussed in light of functionality with enzymes involved in lipoprotein modification. N-acylation of lipoproteins by Lnt is essential for the Lol-dependent release of lipoproteins from the cytoplasmic membrane (Fukuda et al., 2002), however, overexpression of lolCDE can compensate for lack of N-acylation, in the absence of Lpp, and results in the release of diacylated lipoproteins from the membrane and correct outer membrane localization (Narita & Tokuda, 2011). However, the efficiency of outer membrane targeting is dependent on the nature of the lipoprotein (Narita & Tokuda, 2011). Model lipoproteins containing an outer membrane-sorting signal are not localized in the outer membrane, neither released from the cytoplasmic membrane in absence of Lnt. This suggests that the N-acyl moiety of lipoproteins plays an important role in lipoprotein translocation through the periplasm. Mature lipoprotein first interacts with LolE, is then transferred to LolC, and upon ATP hydrolysis by LolD, the
interaction between LolC and LolA leads to charging of LolA with lipoprotein (Okuda &
Tokuda, 2009). The interaction between LolA and lipoprotein is weaker than the interaction
between LolB and lipoprotein, driving transfer from LolA to LolB and preventing back-
transfer to LolCE (Taniguchi et al., 2005). It is unknown how lipoproteins interact with the
ABC transporter LolCDE and by which mechanism lipoproteins are released from the
cytoplasmic membrane. Crystal structures of LolA and soluble LolB, lacking its lipid-anchor,
have been solved (Takeda et al., 2003). The calculated size of the cavity is predicted to
accommodate one but is not large enough for three fatty acid chains. Structure-function
analyses of LolA of *P. aeruginosa* showed that one large hydrophobic surface patch is likely
to be involved in binding of acyl chains (Remans et al., 2010). Decreasing its hydrophobic
nature abolished LolA-dependent release of lipoproteins from the cytoplasmic membrane
(Remans et al., 2010). The fatty acid acyl moiety of lipoprotein, together with its sorting
signal, clearly plays a role in substrate recognition by the Lol machinery. LolC and LolE
share a similar membrane topology and their periplasmic domains share sequence identity
with each other as well as with LolB (Okuda & Tokuda, 2009). LolA interacts specifically
with LolC and not with LolE. Differences in binding affinity between LolE and LolC for
lipoprotein, similar as described for LolA and LolB, may explain the molecular mechanism of
transfer. Characterization of the lipid-protein interaction by structural- and biophysical studies
is likely to provide more insight into the mechanism of Lol-dependent targeting of
lipoproteins to the outer membrane.

Inhibitors of the Lol machinery were identified in a chemical genomic approach (Pathania et
al., 2009) which utilized a combination of three *in vitro* assays (Ito et al., 2007). Treatment
of cells with MAC13243 led to the accumulation of Lpp in the cytoplasmic membrane. It
directly interacts with LolA (Pathania et al., 2009). Further characterization of this compound
led to the finding that a degradation product is the active species (Barker et al., 2013). This
thiourea compound is an analog of A22, an inhibitor of bacterial actin MreB (Barker et al., 2013) and seems to inhibit bacterial growth in a non-specific fashion.

The Lol machinery is not fully conserved in proteobacteria, LolB is only found in β- and γ-
 proteobacteria and also LolCE are often missing (Okuda & Tokuda, 2011). Proteomics analyses showed that lipoproteins reside in the outer membrane of diderm actinomycetes but the sorting mechanism is completely unknown.

Lipoproteins involved in key processes of envelope biosynthesis

The most abundant lipoprotein in γ-proteobacteria is Lpp that is covalently linked to murein (peptidoglycan) and the first protein shown to be lipidated (Braun & Rehn, 1969). Lpp is also the most abundant protein in the cell and plays an important role in homeostasis of the cell envelope. A recent study identified a σE-regulated small RNA, MicL, which specifically targets lpp (Guo et al., 2014). The mechanism of regulation allows decrease of Lpp translation and acceleration of degradation of the stable lpp mRNA. σE up-regulates proteins that facilitate assembly of outer membrane proteins and transport of LPS under stress conditions. At the same time it down-regulates OMP synthesis by up-regulation of sRNAs such as MicL (Guo et al., 2014). A strain deleted for lpp is very sensitive to EDTA and cationic dyes and somewhat sensitive to detergents and it also leaks periplasmic proteins into the extracellular space (Hirota et al., 1977, Yem & Wu, 1978). One-third of Lpp is covalently linked to peptidoglycan, while two-third is in the cell envelope but not anchored to the cell wall, the so-called free form (Inouye et al., 1972). There is a dynamic equilibrium between the two forms and the protein is homogeneously distributed over the cell (Hiemstra et al., 1986, Hiemstra et al., 1987). The enzymes responsible for cross-linking Lpp to peptidoglycan were identified in 2008 (Magnet et al., 2008). YbiS was shown to catalyze the covalent anchoring of Lpp to peptidoglycan, but two other enzymes ErfK and YcfS could also perform
this reaction. These enzymes belong to the family of L,D-transpeptidases found in firmicutes (Mainardi et al., 2005, Magnet et al., 2007). YbiS is expected to cleave the peptide bond between meso-DAP and D-Ala and to link meso-DAP to the L-Lys residue in Lpp (Magnet et al., 2008). Mis-localization of Lpp in the cytoplasmic membrane is lethal when cross-linked to the peptidoglycan (Yakushi et al., 1997). The highly abundant Lpp homologue in P. aeruginosa, OprI exist in a free form and is not covalently cross-linked to peptidoglycan (Mizuno & Kageyama, 1979, Cornelis et al., 1989). When produced in E. coli it is exposed on the surface of cells (Cornelis et al., 1996). Interestingly, the free form of E. coli Lpp is also exposed on the cell surface (Cowles et al., 2011), although detection of Lpp by immunogold labeling was only successful upon permeabilization of the outer membrane with TRIS-EDTA (Hiemstra et al., 1986). Lpp is an elongated trimer composed of long α-helical domains that adopt a coiled coil, which reflects its repetitive characteristics (Braun & Bosch, 1972, Choi et al., 1986). The last amino acids of the C-terminus form a hydrophobic capping motif that is important for binding to PG and possibly involved in the recognition of lipoprotein by ligase YbiS (Zhang & Wu, 1992, Shu et al., 2000). It is thus likely that Lpp has a dual orientation, similar as described for peptidoglycan-associated Lipoprotein P6 of Haemophilus influenza (Michel et al., 2013), although the correlation between surface-exposed orientation and biological function is not yet known.

Surface exposed lipoproteins are abundant in Borrelia, which lacks LPS in the outer membrane (for a recent review see (Zuckert, 2014)). The N-terminal membrane-tethering peptide keeps lipoproteins of Borrelia burgdorferi in a translocation compatible state that allow them to be secreted to the cell surface probably via a flippase complex. Several other examples of surface-exposed lipoproteins have been reported in proteobacteria, including proteins that are translocated via type II secretion machinery (PulA) (Pugsley et al., 1986) or via an autotransporter mechanism (NalP and SphB1) (Coutte et al., 2003, van Ulsen et al.,
In case of NalP of *Neisseria meningitidis*, it was demonstrated that lipidation is required for efficient release of surface-exposed proteins by delaying its own release from the cell surface through autocatalytic processing (Roussel-Jazede *et al.*, 2013). Similarly, non-lipidated SphB1 is efficiently secreted but is unable to cleave its substrate filamentous haemagglutinin FHA (Coutte *et al.*, 2003), which is important for virulence of *Bordetella pertussis* (Coutte *et al.*, 2003).

The Lpt pathway transports LPS from the cytoplasmic membrane to the cell surface (Ruiz *et al.*, 2009). LptE is an essential OM lipoprotein that stabilizes the LptD OM channel responsible for correct translocation of LPS (Bos *et al.*, 2004, Chng *et al.*, 2010). Recent studies showed that LptE forms a plug inside the β-barrel structure of LptD (Freinkman *et al.*, 2011) and that it changes the biophysical state of LPS to allow transport and OM insertion (Malojecic *et al.*, 2014). The N-terminal lipid anchor is not required for function since a variant of LptE lacking the signal sequence including the cysteine complements a *lptE* deletion strain and is able to form a complex with LptD (Chng *et al.*, 2010). LptD is inserted into the OM by the β-barrel assembly machine (Bam complex) (Chimalakonda *et al.*, 2011). Four lipoproteins (BamB-E) are part of the Bam complex, of which BamD is essential for the assembly and/or insertion of proteins in the OM (Malinverni *et al.*, 2006). BamC is surface exposed (Webb *et al.*, 2012). How the lipid moiety is orientated in the outer membrane is still an open question.

Biosynthesis of peptidoglycan is catalyzed by several membrane-bound enzymes, among which the bifunctional penicillin-binding proteins PBP1A and PBP1B. These enzymes each require an outer membrane lipoprotein, LpoA and LpoB, respectively, for their function (Paradis-Bleau *et al.*, 2010, Typas *et al.*, 2010). The Lpo proteins bind specifically to their cognate PBP and stimulate transpeptidase activity. LpoB with an inner membrane retention signal is functional as long as the flexible N-terminal domain is present (Typas *et al.*, 2010). It is currently unknown whether non-lipidated LpoA is functional.
The role of lipoprotein modification in virulence

The role of individual lipoproteins on virulence of pathogenic bacteria has been studied (for a review see (Kovacs-Simon et al., 2010)). Signaling of the host immune system by lipoproteins and lipopeptides occurs via Toll-like receptor 2, either via TLR2-1 in case of diacylated proteins or via TLR2-6 heterodimers in case of triacylated proteins. The lipid moiety plays an important role in the interaction with these receptors (Jin et al., 2007, Kang et al., 2009). In recent years the effect of lipoprotein modification on virulence has been addressed using lgt and lsp deletion mutants of pathogenic bacteria and depending on the pathogen and the infection model used, various effects on virulence were obtained.

For example, impaired colonization of the nasopharynx and a decrease in virulence in pneumonia and septicaemia mouse models was reported for Streptococcus pneumoniae deleted for lgt (Petit et al., 2001, Chimalapati et al., 2012). TLR-2 signaling is dependent on pneumococcal lipoproteins and NF-κB and cytokine release were reduced in a Δlgt strain (Tomlinson et al., 2014). Intracellular survival of Listeria monocytogenes was affected (Baumgärtner et al., 2007) and the Δlgt strain was attenuated in a mouse model (Machata et al., 2008). Lipoproteins are released into the culture supernatant of this strain (Baumgärtner et al., 2007). Inactivation of lsp in L. monocytogenes led to a severe growth defect in macrophages and a reduced capacity of phagosomal escape leading to attenuated virulence (Reglier-Poupet et al., 2003). Expression of the lsp gene was strongly induced inside phagosomes probably caused by environmental conditions similar to pH-regulation of N-acylation in S. aureus (Reglier-Poupet et al., 2003, Kurokawa et al., 2012). Inefficient germination of spores of a lgt deletion strain of Bacillus anthracis led to attenuation of virulence (Okugawa et al., 2012). Lipoproteins of an lsp deletion mutant of M. tuberculosis are not processed and the strain was attenuated in an animal model of tuberculosis (Sander et
Reduced innate immune activation was observed with a Δlgt strain of *Streptococcus suis* (Wichgers Schreur *et al.*, 2011), but no difference in virulence was observed between a wild type strain and a *lsp* deletion mutant in an infection model in piglets (De Greeff *et al.*, 2003). Deletion of either *lgt* or *lsp* in *Streptococcus sanguinis* has only a slight effect on endocarditis virulence (Das *et al.*, 2009). No significant effect on virulence was observed in *Streptococcus equi* (Hamilton *et al.*, 2006, Das *et al.*, 2009). Escape of immune recognition caused lethal infections by a Δlgt strain of *S. aureus* but a Δlsp strain was attenuated (Bubeck-Wardenburg *et al.*, 2006). Enhanced lethality was shown in a mouse model for neonatal sepsis by Δlgt of *Streptococcus agalactiae* (Henneke *et al.*, 2008). Stimulation of the immune system via TLR2 during early stages of infection is important for elimination of pathogens. The examples described here clearly show that both the lipid moiety and the functional characteristics of lipoproteins determine virulence capacity.

**Protein lipidation in other microbes**

Lipoproteins are predicted in Archaea, but orthologous genes for lipoprotein modifying enzymes could not been identified in their genomes. Recent findings in *Haloferax volcanii* show that the Tat-dependent secretion of predicted lipoproteins depends on the conserved cysteine in the putative lipobox (Gimenez *et al.*, 2007, Storf *et al.*, 2010). Future studies will provide direct evidence for lipidation of these proteins and will demonstrate whether phospho-, glyco- or phospho-glyco-lipids that constitute archaeal membranes, and are substantially different from their bacterial counterparts, are the source of modification (Villanueva *et al.*, 2014).

In eukaryotes, including yeast, various types of protein lipidation exist including *S*-palmitoylation, *N*-myristoylation and *S*-prenylation (Aicart-Ramos *et al.*, 2011). In the process of *S*-palmitoylation, palmitate is transferred onto a cysteine residue catalyzed by
protein palmitate transferases (PAT) containing DHHC (Asp-His-His-Cys) cysteine rich
domains that results in the formation of a thioester-linked palmitoylated protein (Dietrich &
Ungermann, 2004). The PAT enzymes are polytopic integral membrane proteins located in
the ER or Golgi. The active site is facing the cytoplasm where acyl-CoA functions as the acyl
donor. Multiple DHHC enzymes have been identified in *Saccharomyces cerevisiae* but their
substrate specificity is still poorly understood. Recent work on the Erf2-Erf4 PAT complex in
*S. cerevisiae* showed that it catalyzes a two-step reaction in which Erf2 undergoes auto-
acylation (Mitchell *et al.*, 2010). Erf2 transfers the palmitate group onto the substrate and Erf4
is required for stabilization of Erf2 and for the formation of the Erf2-palmitate intermediate
(Mitchell *et al.*, 2012). S-palmitoylation is a reversible process due to the action of acyl-
protein thioesterases. Together these two processes allow for regulated control of membrane-
associated proteins, such as small GTPase Ras, the function of which depends on their lipid
anchor.

Several examples have been described of bacterial effector proteins that contain eukaryotic
lipidation motifs and that are modified by lipids upon entry into host cells. A farnesyl or
geranylgeranyl isopreny lipid moiety is added to cysteine residue of the conserved sequence
leading to enhanced membrane affinity and localization of effector proteins to membranes of
host cells (Ivanov *et al.*, 2010).

**Concluding remarks**

Lipoproteins are highly abundant in bacteria, and require therefore efficient maturation. The
acyltransferases and signal peptidase involved in this pathway are not abundant, estimated to
only a few hundred molecules per cell. It is very likely that these enzymes exist in a complex
to guarantee complete and efficient lipoprotein modification. The existence of a lipoprotein
modification complex was already suggested in 1982 (Tokunaga *et al.*, 1982). This presumed
complex might co-localize and function together with other essential protein complexes including the Sec and Tat secretion machineries. In this respect it is intriguing to understand how the lipoprotein modification machinery recognizes non-folded- and fully folded substrates that are translocated via Sec and Tat, respectively. The signal peptide of both types of lipoprotein substrates including the early mature region of the precursor protein is thought to exist in a flexible conformation inside the translocon. Lateral opening of the Sec channel can be envisaged for insertion of lipoproteins into the lipid bilayer (Chatzi et al., 2014), whether lateral gating of the Tat complex occurs is currently unknown (Patel et al., 2014).

Once inserted into the membrane, Lgt binds the signal peptide and the first few amino acids of the mature protein and catalyzes the first step in the lipidation pathway. Several examples clearly demonstrate that some essential lipoproteins are perfectly functional without their lipid anchor what raises the following questions: what is the role of the lipid moiety, and how does the degree of lipidation influence protein function? Unlike membrane proteins, lipoproteins are expected to be more mobile within lipid bilayers giving them more flexibility, although this has not been experimentally tested. For lipoproteins in monoderm bacteria it has been postulated that the N-terminal lipid anchor keeps the protein close to the membrane while the extended functional domain participates in various biological processes. Many lipoproteins in these organisms fulfill equivalent functions as periplasmic proteins in diderm bacteria. Accessory proteins such as LptE and LpoA/B interact strongly with their partner proteins, for which the lipid part is not essential, although the process in which they are involved might be more efficient when these proteins are present in lipidated form. In the case of lipoproteins that themselves fulfill specific functions and not in complex with other proteins, such as surface-exposed proteases or targeting of proteins to the OM, the lipid anchor becomes essential. The lipoprotein modification pathway seems to be tightly regulated since a lipoprotein can exist in both di- or triacylated form in one bacterial species, depending
on growth conditions. The fact that multiple copies of genes encoding modification enzymes exist in bacteria that undergo various developmental stages also suggest that regulation of their expression is environmentally controlled.

The lipoprotein modification pathway represents an attractive target for antibiotic development. It is essential for the viability of all proteobacteria examined to date, which include many important human pathogens. Furthermore, Lsp and Lnt are membrane-bound with catalytic domains facing the bacterial periplasm. This topology makes at least two steps of lipoprotein biogenesis accessible to small molecules that will be capable of inhibiting enzymes without a need to traverse the relatively impermeable cytoplasmic membrane or risk elimination by the action of efflux pumps. For future studies on lipoprotein modification and function it is important to include proteomic analysis of membrane-bound and secreted lipoproteins and to determine the lipidation state and lipid composition by mass spectrometry (Cain et al., 2014). These approaches not only allow the identification and characterization of specific lipoproteins that act as virulence factors but also demonstrate a correlation between lipidation and the stage of infection. Structural and mechanistic properties of enzymes involved in lipidation of proteins in various organisms will highlight similarities and differences and will provide ways of finding new species specific antibacterial agents.

Acknowledgements

I thank Joel Berry, Richard Wheeler and Nicolas Bayan for constructive comments and all members of the Biology and Genetics of the Bacterial Cell Wall Unit at the Institut Pasteur for their support.
Figure legends

Figure 1. Lipoprotein modification pathway in bacteria

Lipoprotein modification occurs in the cytoplasmic membrane (CM) by the sequential action of three membrane-bound enzymes Lgt, Lsp and Lnt. Prolipoprotein is inserted into the membrane via the Sec or Tat secretion machinery through recognition of its signal peptide (SP). This sequence also contains the lipobox (LB) specific for bacterial lipoproteins. Lgt catalyzes the transfer of a diacylglyceryl group from phosphatidylglycerol (PG) onto prolipoprotein resulting in diacylglyceryl-prolipoprotein. Cleavage of the signal peptide by Lsp liberates the α-amino group of diacylglyceryl-cysteine that then becomes acylated by Lnt resulting in mature triacylated lipoprotein. Lnt preferentially uses phosphatidylethanolamine (PE) as acyl donor. Alternative N-acyl modifications have been observed in lipoproteins from firmicutes and mollicutes.

Figure 2. Membrane topology of modification enzymes

Schematic representation of membrane topology and essential residues (in red) in Lgt, Lsp and Lnt. Numbering of amino acids correspond to sequences of *E. coli* proteins. C=cytoplasm, CM=cytoplasmic membrane, P=periplasm. Catalytic site residues are shown in yellow. Conserved domains are indicated as white boxes. See text for details.

Figure 3. Structure of globomycin

Globomycin is a lipophilic cyclic peptide. Derivatives have been synthesized and analyzed for their activity against a variety of bacteria, in particular the length of the lipophilic part (A), and the role of the hydroxyl groups of threonine (B) and serine (C) have been addressed.
Figure 4. Lol machinery in proteobacteria

The Lol machinery is involved in translocation of mature (triacylated) lipoproteins from the cytoplasmic membrane (CM) to the outer membrane (OM). The ABC transporter component of the machinery is composed of membrane proteins LolC and LolE and cytoplasmic ATPase LolD. The periplasmic chaperone LolA binds the lipoprotein, upon interaction with LolC and hydrolysis of ATP by LolD, and transfers it to receptor protein LolB. The lipoprotein is then inserted into the outer membrane. Several lipoproteins are located on the cell surface; however, the mechanism for their translocation and the orientation of the fatty acids in the membrane is not yet understood.
References:


Hiemstra H, de Hoop MJ, Inouye M & Witholt B (1986) Induction kinetics and cell surface
distribution of *Escherichia coli* lipoprotein under lac promoter control. *Journal of
bacteriology* **168**: 140-151.

newly synthesized lipoprotein over the outer membrane and the peptidoglycan sacculus
of an *Escherichia coli* lac-lpp strain. *Journal of bacteriology* **169**: 5434-5444.

Hillmann F, Argentini M & Buddelmeijer N (2011) Kinetics and phospholipid specificity

Hirota Y, Suzuki H, Nishimura Y & Yasuda S (1977) On the process of cellular division in
*Escherichia coli* mutants: A mutant of *E. coli* lacking a murein-lipoprotein. *Proc Natl Acad Sci USA
74*: 1417-1420.

Nogi T (2014) A structure-based model of substrate discrimination by a noncanonical

Homma H & Nojima S (1982) Synthesis of various phospholipids from 2-acyl
lysophospholipids by *Escherichia coli* extract. *Journal of biochemistry* **91**: 1103-1110.

Hsu L, Jackowski S & Rock C (1991) Isolation and characterization of *Escherichia coli* K-12
mutants lacking both 2-acylglycerophosphoethanolamine acyltransferase and acyl-acyl carrier


precursor of the outer membrane lipoprotein in the cytoplasmic membrane of

biosynthesis of the major lipoprotein of the *Escherichia coli* outer membrane. *J Biol Chem
257*: 5177-5182.

of the *sppA* gene and characterization of the encoded protease IV, a signal peptide

Inouye M, Shaw J & Shen C (1972) The assembly of a structural lipoprotein in the

residues in the hydrophobic region of the *Escherichia coli* prolipoprotein signal peptide


Zuckert WR (2014) Secretion of Bacterial Lipoproteins: Through the Cytoplasmic Membrane, the Periplasm and Beyond. *Biochim Biophys Acta.*
Table 1. Essentiality of genes encoding lipoprotein modifying enzymes in diderm bacteria.

<table>
<thead>
<tr>
<th></th>
<th>lgt</th>
<th>lsp</th>
<th>int</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E. coli</strong></td>
<td>lgt (umpA) (Gan et al., 1995)</td>
<td>lsp (Tokunaga et al., 1983, Yamagata et al., 1983)</td>
<td>int (cutE) (Gupta &amp; Wu, 1991, Rogers et al., 1991, Gupta et al., 1993, Robichon et al., 2005)</td>
</tr>
<tr>
<td><strong>S. enterica</strong></td>
<td>lgt (Gan et al., 1993)</td>
<td>lsp1, lsp2, lsp3, lsp4 (Xiao &amp; Wall, 2014)</td>
<td>int (Gupta et al., 1993)</td>
</tr>
<tr>
<td><strong>M. xanthus</strong></td>
<td>lgt (MSMEG_3222) (Tschumi et al., 2012)</td>
<td>lsp (Rv1539) (Sander et al., 2004, Rampini et al., 2008)</td>
<td>ppm1 (Rv2051c) (Tschumi et al., 2009)</td>
</tr>
<tr>
<td><strong>M. smegmatis</strong></td>
<td>lgt (Rv1614) (Sassetti et al., 2003, Tschumi et al., 2012)</td>
<td>lsp (SCO2074) (Thompson et al., 2010)</td>
<td>int (MSMEG_3860) (ppm2) (Tschumi et al., 2009)</td>
</tr>
<tr>
<td><strong>M. tuberculosis</strong></td>
<td>lgt (SCO2034), lgt2 (SCO7822) (Thompson et al., 2010)</td>
<td>lsp (SCAB68121) (Widdick et al., 2011)</td>
<td>int1 (SCO1014), int2 (SCO1336) (Widdick et al., 2011, Cordova-Davalos et al., 2014)</td>
</tr>
<tr>
<td><strong>S. coelicolor</strong></td>
<td>lgt (SCAB68531) (Widdick et al., 2011)</td>
<td>lsp (SCAB68121) (Widdick et al., 2011)</td>
<td>int1 (SCAB83111), int2 (SCAB76621) (Widdick et al., 2011)</td>
</tr>
<tr>
<td><strong>S. scabies</strong></td>
<td>lgt (SCAB68531) (Widdick et al., 2011)</td>
<td>lsp (SCAB68121) (Widdick et al., 2011)</td>
<td>int1 (SCAB83111), int2 (SCAB76621) (Widdick et al., 2011)</td>
</tr>
<tr>
<td><strong>S. lividans</strong></td>
<td>lgt (SCAB68531) (Widdick et al., 2011)</td>
<td>lsp (Gullon et al., 2013)</td>
<td>int1 (SCAB83111), int2 (SCAB76621) (Widdick et al., 2011)</td>
</tr>
</tbody>
</table>

Essential genes are shown in red, non-essential genes are shown in blue. Table includes only published findings. See text for further details.
FIGURE 2

Lgt (291 aa)  Lsp (164 aa)  Lnt (512 aa)

293x167mm (300 x 300 DPI)
FIGURE 3

128x98mm (300 x 300 DPI)
FIGURE 4

180x320mm (300 x 300 DPI)