



**HAL**  
open science

## Residues located on membrane-embedded flexible loops are essential for the second step of the apolipoprotein N-acyltransferase reaction

Sébastien Gélis-Jeanvoine, Stephen Lory, Jacques Oberto, Nienke Buddelmeijer

### ► To cite this version:

Sébastien Gélis-Jeanvoine, Stephen Lory, Jacques Oberto, Nienke Buddelmeijer. Residues located on membrane-embedded flexible loops are essential for the second step of the apolipoprotein N-acyltransferase reaction. *Molecular Microbiology*, 2015, 95 (4), pp.692 - 705. 10.1111/mmi.12897 . pasteur-01407474

HAL Id: pasteur-01407474

<https://pasteur.hal.science/pasteur-01407474>

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

1 Residues located on **membrane-associated flexible** loops are essential for the second  
2 step of the apolipoprotein N-acyltransferase (Lnt) reaction

3

4 Sébastien Gélis-Jeanvoine<sup>1§</sup>, Stephen Lory<sup>1#</sup>, Jacques Oberto<sup>2</sup> and Nienke  
5 Buddelmeijer<sup>1\*</sup>

6

7 <sup>1</sup>Institut Pasteur, Biology and Genetics of the Bacterial Cell Wall Unit, Inserm Group  
8 Avenir, 28 rue du docteur Roux, Paris F-75724 cedex 15, France.

9 <sup>2</sup>Institut de Génétique et Microbiologie, CNRS UMR 8621, Orsay F-91405 cedex,  
10 **France**

11 <sup>§</sup>Current address: Micalis INRA, Domaine de Vilvert, Jouy-en-Josas F-78352 cedex,  
12 France.

13 <sup>#</sup>Permanent address: Harvard Medical School, Department of Microbiology and  
14 Immunobiology, 77 Ave. Louis Pasteur, Boston, MA 02115, USA

15

16 \*Address for correspondence: Institut Pasteur, Biology and Genetics of the Bacterial  
17 Cell Wall Unit, Inserm Group Avenir, 28 rue du docteur Roux, Paris F-75015, France.

18 Phone +33(0)145688015, Fax +33(0)140613756 E-mail: [niebud@pasteur.fr](mailto:niebud@pasteur.fr)

19

20 Running title: Functional analysis of Apolipoprotein N-acyltransferase (Lnt) of *E. coli*

21 Key words: lipoprotein modification, Apolipoprotein N-acyltransferase, membrane  
22 topology, essential residues, error-prone PCR mutagenesis.

23

24 Abbreviations:

25 Isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG), hydroxylamine (HA), (2-  
26 sulfonatoethyl)-methane thiosulfonate (MTSES), maleimide polyethylene glycol 5000  
27 (malPEG).  
28

29 **Summary**

30

31 Apolipoprotein N-acyltransferase (Lnt) is an essential membrane-bound enzyme that  
32 catalyzes the third and last step in the post-translational modification of bacterial  
33 lipoproteins. In order to identify essential residues implicated in substrate recognition  
34 and/or binding we screened for non-functional variants of Lnt obtained by error-prone  
35 PCR in a complementation assay using a *lnt* depletion strain. Mutations included  
36 amino acid substitutions in the active site and of residues located on flexible loops in  
37 the catalytic periplasmic domain. All, but one mutation, led to the formation of the  
38 thioester acyl-enzyme intermediate and to the accumulation of apo-Lpp, suggesting  
39 that these residues are involved in the second step of the reaction. A large cytoplasmic  
40 loop contains a highly conserved region and two hydrophobic segments. Accessibility  
41 analysis to alkylating reagents of substituted cysteine residues introduced in this  
42 region demonstrated that the hydrophobic segments do not completely span the  
43 membrane. Two residues in the highly conserved cytoplasmic region were shown to  
44 be essential for Lnt function. Together, our data suggest that amino acids located on  
45 cytoplasmic and periplasmic membrane-associated loops are required for efficient *N*-  
46 acylation of lipoproteins.

47

## 48 **Introduction**

49

50 Bacterial lipoproteins fulfill a variety of functions **that include** nutrient uptake,  
51 maintenance of cell wall integrity, cell wall growth, adhesion to surfaces and host  
52 cells, modulation of inflammatory processes and translocation of virulence factors  
53 (Kovacs-Simon *et al.*, 2010, Sutcliffe & Russell, 1995). Although their primary  
54 sequences and three-dimensional structures differ to a large extent, they all possess  
55 fatty acids at their amino terminus the source of which are membrane phospholipids  
56 (Lai *et al.*, 1980). Modification of lipoproteins occurs in a post-translational fashion  
57 by the sequential action of three membrane-bound enzymes: two acyltransferases and  
58 one signal peptidase. Lipoproteins are synthesized in the cytoplasm and inserted into  
59 the cytoplasmic membrane via the Sec or Tat secretion machinery (Zuckert, 2014,  
60 Sugai & Wu, 1992, Thompson *et al.*, 2010). A specific signature sequence, composed  
61 of a signal peptide sequence containing a so-called lipobox (consensus sequence  
62 **Leu**(Ala/Val)-**Leu**-Ala(Ser)-Gly(Ala)-Cys for lipoproteins in *Proteobacteria*), allows  
63 membrane insertion and recognition by the modification machinery (**Wu *et al.*, 1983**).  
64 The invariant cysteine in the lipobox becomes fatty-acid acylated. After complete  
65 maturation this cysteine is the first amino acid of mature lipoprotein. The first step in  
66 the modification process is catalyzed by prolipoprotein-  
67 diacylglyceryl:phosphatidylglycerol transferase (Lgt), which attaches a diacylglyceryl  
68 from phosphatidylglycerol to the thiol group of cysteine resulting in a thioether linked  
69 *S*-diacylglyceryl group of prolipoprotein (Sankaran & Wu, 1994). The second step in  
70 lipoprotein maturation is the cleavage of the signal peptide from *S*-diacylglyceryl  
71 prolipoprotein by signal peptidase II (Lsp) resulting in a free  $\alpha$ -amino group of *S*-  
72 diacylglyceryl-cysteine (Hussain *et al.*, 1982). The third and last step is catalyzed by

73 apolipoprotein N-acyltransferase (Lnt) that attaches a fatty acid from phospholipids to  
74 this  $\alpha$ -amino group resulting in mature triacylated lipoprotein (Hillmann *et al.*, 2011,  
75 Jackowski & Rock, 1986). The mechanism of the Lnt-catalyzed reaction involves a  
76 two-step ping-pong mechanism in which, upon hydrolysis of phospholipids, the  
77 enzyme is first auto-acylated on the active site cysteine residue resulting in the  
78 formation of a thioester acyl-enzyme intermediate and lyso-phospholipid  
79 (Buddelmeijer & Young, 2010, Hillmann *et al.*, 2011). In the second step, the  $\alpha$ -  
80 amino group of diacylglyceryl-cysteine of apolipoprotein is acylated. Phospholipids  
81 with small polar head groups, preferentially phosphatidylethanolamine (PE), and  
82 saturated- and unsaturated fatty acid on *sn*-1 and *sn*-2 position, respectively, are the  
83 preferred substrates for Lnt of *E. coli* (Hillmann *et al.*, 2011, Jackowski & Rock,  
84 1986). Incomplete modification of lipoproteins in  *$\gamma$ -proteobacteria*, for example  
85 through inhibition of Lsp activity by globomycin (Inukai *et al.*, 1978) or myxovirescin  
86 (Xiao *et al.*, 2012), or depletion of genes encoding the modification enzymes (Pailler  
87 *et al.*, 2012, Robichon *et al.*, 2005, Xiao & Wall, 2014), leads to arrest of cell growth  
88 and cell lysis. Accumulation of the abundant outer membrane lipoprotein Lpp  
89 (Braun's lipoprotein) in an unmodified form and covalently linked to the cell wall in  
90 the cytoplasmic membrane, is the main cause of cell death (Hussain *et al.*, 1980).  
91 However, modification and correct sorting of other essential outer membrane  
92 lipoproteins, such as BamD (Malinverni *et al.*, 2006) LpoA and LpoB (Paradis-Bleau  
93 *et al.*, 2010, Typas *et al.*, 2010), is important since deletion of *lpp* still leads to cell  
94 lysis in the absence of Lgt (Pailler *et al.*, 2012). Retention of these outer membrane  
95 lipoproteins in the cytoplasmic membrane likely leads to cell death. The lipoprotein  
96 modification machinery is conserved in all bacterial species, and although *N*-acylation  
97 occurs in low GC Gram-positive bacteria such as firmicutes and mollicutes, the gene

98 encoding Lnt has not yet been identified in these species (Nakayama *et al.*, 2012). In  
99 these organisms not only does the composition of the N-acyl group vary, but also the  
100 length and degree of saturation of the *S*-diacylglyceryl entity of lipoprotein is variable  
101 (Asanuma *et al.*, 2010, Kurokawa, 2009, Kurokawa *et al.*, 2012b, Serebryakova *et al.*,  
102 2011). Recent findings suggest that the extent of N-acylation varies depending on  
103 growth conditions probably by regulation of *lnt* expression (Kurokawa *et al.*, 2012a).  
104 The role of Lnt in lipoprotein modification has also been studied in high GC Gram-  
105 positive bacteria (*Actinomycetales*). In *Mycobacterium smegmatis* (Tschumi *et al.*,  
106 2009) and *Corynebacterium glutamicum* (Mohiman *et al.*, 2012) a palmitoyl group is  
107 transferred by Lnt onto apolipoproteins. It was shown recently that lipoproteins in *M.*  
108 *bovis* also contain *N*-acyl tuberculostearic acid (C19) (Brülle *et al.*, 2013). Thus, both  
109 the *S*-diacylglyceryl and the *N*-acyl moieties of lipoproteins in *Mycobacteria* contain  
110 tuberculostearic acid (Tschumi *et al.*, 2009). Furthermore, several lipoproteins in  
111 actinomycetes are modified by sugars, mainly hexose, via O-glycosylation (Brülle *et*  
112 *al.*, 2010, Mohiman *et al.*, 2012, Sartain & Belisle, 2009, Garbe *et al.*, 1993). In *C.*  
113 *glutamicum* apolipoprotein *N*-acyltransferase (Lnt or Cg-ppm2) and *O*-  
114 glycosyltransferase (Cg-ppm1) operate in a common biosynthetic pathway resulting  
115 in glycosylated lipoproteins in this organism (Mohiman *et al.*, 2012), but N-acylation  
116 is not a prerequisite for glycosylation in *M. bovis* (Brülle *et al.*, 2013). Two *lnt* genes  
117 were identified in *Streptomyces scabies*. Lnt1 is the enzyme responsible for *N*-  
118 acylation of lipoproteins since lipoproteins are diacylated in *lnt1* and *lnt1lnt2* mutants.  
119 Lnt2 is required for efficient fatty acid acylation of lipoproteins since in its absence  
120 lipoproteins are both di- and triacylated (Widdick *et al.*, 2011).  
121 Although the mechanism of apolipoprotein *N*-acyltransferase is likely to be conserved  
122 among bacteria, differences in substrate specificity result in different types of

123 lipidation. This led us to ask the question of how Lnt recognizes and interacts with its  
124 substrates, phospholipid and apolipoprotein, and what residues are involved in the two  
125 steps of the reaction. Here we identified various essential residues in Lnt of *E. coli*  
126 that are located on flexible loops and are predicted to be membrane-associated. Our  
127 results suggest that amino acids located on the cytoplasmic- and on the periplasmic  
128 side of the cytoplasmic membrane are involved in the transfer of the acyl group onto  
129 the apolipoprotein substrate.

130

131

132



133 **Results**

134

135 *Identification of essential residues in full-length Lnt*

136

137 Topology studies with Lnt from *E. coli* using reporter fusions with cytoplasmic  $\beta$ -  
138 galactosidase and periplasmic alkaline phosphatase showed that this enzyme spans the  
139 cytoplasmic membrane via six transmembrane segments (TMS) with both its amino-  
140 terminus and carboxy-terminus facing the cytoplasm (Robichon *et al.*, 2005). Lnt  
141 contains a large cytoplasmic loop (R<sub>114</sub>-R<sub>193</sub>) (CL-2) between TMS-IV and TMS-V,  
142 with two hydrophobic sequences (TMS-IVa and TMS-IVb) and a highly conserved  
143 sequence (T<sub>144</sub>GFPWL<sub>149</sub>) (Fig. 1).

144 In order to identify residues in Lnt involved in substrate recognition and/or binding  
145 we performed random PCR mutagenesis on full-length *lnt* using a two-step PCR  
146 protocol where in the first step long complementary oligos were synthesized using  
147 error-prone DNA polymerase and in the second step the entire plasmid was  
148 synthesized using the mutagenized oligos. The PCR conditions were chosen such that  
149 on average one to three mutations would occur per kilobase (kb) of target DNA. **The**  
150 ***lnt* gene is 1.5 kb in length and between 1 and 5 mutations were expected.**

151 **Functionality of the mutated Lnt-s was addressed in a complementation assay using**  
152 **the *lnt* depletion strain (Robichon *et al.*, 2005). In this strain endogenous *lnt* is under**  
153 **control of an arabinose-inducible promoter. In the presence of L-arabinose *lnt* is**  
154 **expressed, in the absence of L-arabinose *lnt* is not expressed and cells are depleted for**  
155 **Lnt.** The newly created plasmids were, either transformed directly into the *lnt*  
156 depletion strain and selected for growth under permissive conditions (with L-  
157 arabinose), or *lnt* was first re-cloned into a new vector backbone and then transformed

158 into the *lnt* depletion strain. Gene expression was not induced by IPTG and  
159 complementation was observed due to basal expression levels of *lnt* mutants. Colonies  
160 that did not grow in the absence of L-arabinose, and thus express non-functional  
161 variants of *lnt*, were analyzed further. We screened 450 colonies from two  
162 independent rounds of mutagenesis and 34 of those did not complement the *lnt*  
163 depletion strain (Table S1). The *lnt* gene from mutants obtained after direct screening  
164 for non-complementation of the *lnt* depletion strain was also re-cloned into a new  
165 vector backbone to exclude plasmid-linked mutations. Several mutants contained  
166 substitutions on C387 of the Lnt active site (E267-K335-C387), and some had stop  
167 codons. More than two substitutions were frequently observed and in several cases  
168 multiple mutations resulted in non-functional variants (Table S1). Among those,  
169 several mutants contained substitutions located in the first transmembrane segment of  
170 Lnt (Fig. 1) (Robichon *et al.*, 2005) and five mutations in NB24 map in TMS-VI. This  
171 suggests that membrane insertion of Lnt was probably affected in these cases. Most  
172 mutants possess substitutions of amino acids located in both CL-2 and the periplasmic  
173 catalytic domain. We have shown that Lnt is a low abundant enzyme, about 100-200  
174 molecules exist per cell (Hillmann *et al.*, 2011). Single mutations that only slightly  
175 affect Lnt function may not be found because higher protein levels due to expression  
176 from a high-copy plasmid may overcome a functional defect. This might explain why  
177 several mutations are required to observe a strong phenotype in our screen.

178 Mutants that contained one or two substitutions and produced Lnt protein were  
179 analyzed in more detail. Re-cloning of *lnt* from NB8 (L183R, T404S), NB10 (R123C,  
180 I126N), NB14 (W192R, I211T) led to restoration of growth of the *lnt* depletion strain  
181 in the absence of L-arabinose suggesting that additional mutations had occurred either  
182 on the plasmid or on the chromosome. Reconstruction of Q155R from NB34 (Q155R,

183 T315N) was not successful, perhaps due to a dominant negative phenotype of this  
184 mutation, but T315N alone did not affect Lnt function.

185 The substitutions in mutants NB12 (F358V, T481R), NB22 (N244I, Y406C) and  
186 NB31 (P353S, L392H) were constructed individually by site-directed mutagenesis  
187 and analyzed for functionality in the *lnt* depletion strain. A spot dilution assay was  
188 used as a complementation assay that allowed comparison of the various mutants. A  
189 single colony from a fresh plate containing L-arabinose was inoculated in liquid  
190 medium without L-arabinose and cells were grown for two hours to deplete Lnt. Cells  
191 were serial diluted and spotted on plates containing or not L-arabinose and/or IPTG.  
192 The *lnt* depletion strain shows a slight growth defect on L-arabinose when a non-  
193 functional copy of Lnt is present (Fig. 2). Depletion of Lnt in these strains probably  
194 resulted in loss of viability prior to plating. The results show that both F358V and  
195 T481R result in non-functional Lnt. L392H is functional. When combining P353S and  
196 L392H into a single protein, a non-complementing phenotype is again observed.  
197 Complementation was also observed with G447S but the strain has a mucoid  
198 morphology. Both N244I and Y406C (from NB22) resulted in functional enzyme  
199 (data not shown).

200 Previous studies identified W237 and E343 as essential for Lnt function and were  
201 shown to be located on flexible loops in the catalytic domain (Vidal-Ingigliardi *et al.*,  
202 2007). Two substitutions, P361A (NB30) and M362I (NB28), are found close to F358  
203 and P353. Construction of M362I resulted in non-functional Lnt (Fig. 2). NB30  
204 contains several substitutions including P147L that results in non-functional Lnt (see  
205 below). We have previously shown that when highly conserved residues P340 or  
206 P346, located closely to E343, resulted in functional Lnt (Vidal-Ingigliardi *et al.*,  
207 2007). P361A was therefore not analyzed separately.

208 Addition of IPTG resulted in restoration of growth in the absence of arabinose in non-  
209 functional Lnt variants T481R and P353S/L392H and to a lesser extent in F358V and  
210 M362I, suggesting that the enzymes are less efficient and that increase in enzyme  
211 concentration compensates for their reduced activity. All variants produced protein,  
212 although less in the case of T481R, as judged by Western blotting (Fig. S1).

213 In conclusion, the essential residues identified here are located in the catalytic  
214 periplasmic domain of Lnt.

215

216 *Conserved residues R139 and P147 in cytoplasmic loop 2 are required for Lnt*  
217 *function*

218

219 In the screen for essential residues using full-length *lnt* described above, we did not  
220 identify amino acids located in the cytoplasmic loop 2 (CL-2) region of Lnt. Due to  
221 their hydrophobic character, TMS-IVa and TMS-IVb might position the CL-2 with  
222 the conserved T<sub>144</sub>GFPWL<sub>149</sub> sequence in close proximity to or associated with the  
223 membrane (Fig. 1). Upstream of this sequence a charge distribution is conserved in  
224 TMS-IVa (E136 and R139 in Lnt of *E. coli*) (Fig. 1). To gain additional insights into  
225 the role of this loop in Lnt function we performed random PCR mutagenesis targeting  
226 the part of *lnt* encoding CL-2. Restriction sites not affecting the amino acids were  
227 created in codons that allowed for exchange of CL-2 encoding fragments between  
228 wild type *lnt* and mutated fragments. Induction of the plasmids was not required to  
229 observe full complementation with plasmids expressing wild type Lnt. Deletion of  
230 CL-2 resulted in non-functional Lnt, demonstrating that this region is required for Lnt  
231 function (data not shown). A total of 200 mutated clones were screened for loss of Lnt  
232 activity. All clones restored growth of the *lnt* depletion strain under restrictive

233 conditions. Efficient mutagenesis was confirmed by sequencing of several clones.  
234 Several mutants contained multiple substitutions as described above (Table S2).  
235 Conserved residues E136 and R139 were replaced separately by alanine using site-  
236 directed mutagenesis. Lnt with R139A partially complements the *lnt* depletion  
237 mutant, but Lnt containing E136A was functional (Fig. 2). We previously showed that  
238 Lnt in which conserved residue P147 was substituted by alanine was still functional  
239 (Vidal-Ingigliardi *et al.*, 2007). NB30 contained P147L among others substitutions  
240 (Table S1) and was reconstructed and analyzed separately. Lnt with P147L was not  
241 functional, but addition of IPTG restored growth in the absence of L-arabinose (Fig.  
242 2). Both R139A and P147L result in a mixed-colony phenotype. Substitution of P147  
243 by leucine increases the overall hydrophobicity of the region containing TMS-IVa and  
244 T<sub>144</sub>GFPWL<sub>149</sub> and probably results in an orientation of CL-2 relative to the  
245 membrane that leads to inactive Lnt. Thus, R139 is essential and P147 is important for  
246 Lnt function.

247

248 *TMS-IVa and TMS-IVb do not completely span the cytoplasmic membrane*

249

250 CL-2, with 79 amino acids in length, is a relative long cytoplasmic loop for a bacterial  
251 membrane protein. Furthermore, although not always predicted, the two hydrophobic  
252 regions in CL-2 are putative transmembrane segments, where TMS-IVa is less  
253 hydrophobic than TMS-IVb. We used the substituted cysteine accessibility method  
254 (SCAM) to determine whether TMS-IVa and TMS-IVb are real transmembrane  
255 segments such that CL-2 is exposed to the periplasm (Bogdanov *et al.*, 2005, Pailier *et*  
256 *al.*, 2012). Single cysteine substitutions were constructed in previously identified  
257 cytoplasmic and periplasmic loops and between the putative membrane-spanning

258 segments TMS-IVa and TMS-IVb (Fig. 1). The substitutions were introduced by site-  
259 directed PCR mutagenesis in *lnt* encoding both an active variant and an inactive  
260 variant that has active site cysteine 387 replaced by serine (C387S). To facilitate  
261 determination of membrane topology by SCAM, all constructs have cysteine 23 and  
262 cysteine 62 replaced by alanine. These substitutions do not affect Lnt function  
263 (Buddelmeijer & Young, 2010). The cysteine substitutions in active Lnt were used to  
264 test the biological activity in a complementation assay with the *lnt* depletion mutant  
265 (Robichon *et al.*, 2005). All variants restored growth of the *lnt* depletion mutant under  
266 restrictive growth conditions, in the absence of L-arabinose, indicating that all Lnt  
267 cysteine mutants are functional enzymes (data not shown).

268 The second series of cysteine substitutions in inactive Lnt was used to determine  
269 membrane topology. Whole cells were first treated with EDTA to permeabilize the  
270 outer membrane. Cells were then treated with or without membrane-impermeable  
271 alkylating reagent MTSES and subsequently labeled with a second alkylating reagent  
272 maleimide polyethylene glycol (malPEG). The PEG buffer contains urea, SDS and  
273 EDTA that denatures all proteins and renders all unprotected cysteine residues  
274 accessible to malPEG. When cysteine residues are accessible to MTSES, and thus  
275 exposed to the periplasm, their free thiol groups covalently react with MTSES and as  
276 a consequence are blocked for alkylation with malPEG. In the case of cytoplasmic or  
277 membrane-embedded cysteine residues, MTSES does not react with thiol groups and  
278 proteins are alkylated by malPEG resulting in an increase in molecular weight of 5  
279 kDa. Active Lnt and inactive Lnt were not alkylated by malPEG independent of  
280 treatment with MTSES since C387 is acylated in both proteins, as a thioester in Lnt  
281 and as an oxygen-ester in Lnt<sup>C387S</sup> (Fig. 3). Substitutions Y31C, G87C and T221C  
282 were protected by MTSES and not alkylated by malPEG. S5C, S56C, T120C, S154C,

283 N191C and S513C were not protected by MTSES and therefore alkylated with  
284 malPEG (Fig. 3). Alkylation by malPEG of N191C and S513C was less efficient  
285 compared to the other cysteines. These results suggest that residues Y31, G87 and  
286 T221 are accessible to MTSES and are facing the periplasm and that S5, S56, T120,  
287 S154, N191 and S513 are not accessible to MTSES and are located either in the  
288 cytoplasm or embedded in the membrane. Similar results were obtained with an  
289 inactive variant of Lnt in which C387 was replaced by alanine and thus unable to  
290 form an acyl-enzyme intermediate (data not shown) (Buddelmeijer & Young, 2010).  
291 This indicates that the overall membrane topology of Lnt does not change upon auto-  
292 acylation.

293 Our membrane topology analyses show that S154C is not accessible to MTSES in  
294 whole cells and therefore not exposed to the periplasm. Thus the hydrophobic  
295 segments TMS-IVa and TMS-IVb do not completely span the membrane (Fig. 1).  
296 This indicates that the conserved sequence T<sub>144</sub>GFPWL<sub>149</sub> is located in the cytoplasm  
297 and may be in close proximity to the membrane due to the hydrophobic character of  
298 TMS-IVa and TMS-IVb.

299

### 300 *Conservation of essential residues in Lnt of other bacterial species*

301

302 The Lnt sequence of *E. coli* was used to retrieve all Lnt orthologous sequences from  
303 over 2700 complete bacterial genomes. The workflow included comparison of the Lnt  
304 *E. coli* sequence to itself using BLASTP to determine the maximal bit score then  
305 matching the Lnt sequence against sequenced genomes using TBLASTN in a single  
306 representative for each genus. Lnt orthologs with a score of 15% or higher were  
307 analyzed further using Clustal Omega alignment and the conserved residues were

308 identified (Fig. S2). Known Lnt sequences from high GC gram-positive bacteria fell  
309 below this cut-off and were analyzed separately (Fig. S3).

310 Residue G447 is found in Lnt of all bacterial species and P147 and T481 are  
311 conserved in  $\gamma$ -proteobacteria (Fig. 4, Fig. S2). R139 is conserved in  $\gamma$ -proteobacteria  
312 and is substituted by K in some species. One species (*Tolumonas*) has M at this  
313 position. Position 136 in *E. coli* Lnt and the corresponding location in the orthologues  
314 contains a conserved acidic residue, usually an E and less frequently a D. F358 and  
315 M362 are less conserved but are nevertheless shown to be essential in Lnt of *E. coli*.  
316 Interestingly, several Lnt sequences have F358 replaced by valine and M362 replaced  
317 by isoleucine, substitutions that lead to inactivation of Lnt of *E. coli*. Other poorly  
318 conserved residues are probably required to act together with F358 and M362 for Lnt  
319 to function correctly. In the structural model the carboxyl group of L338 interacts  
320 with M362 (Vidal-Ingigliardi *et al.*, 2007). Substitution of L338 by alanine did not  
321 affect Lnt function, since the amino group of L338 does not play a role in the M362-  
322 L338 interaction (Vidal-Ingigliardi *et al.*, 2007). P353 is also poorly conserved and is  
323 only essential in combination with an additional mutation (L392H). This supports the  
324 idea that conformational changes take place in the catalytic domain that involves  
325 movement of flexible loops to facilitate binding of substrate.

326

327 *Essential residues are located on flexible loops in the catalytic periplasmic domain*

328

329 We have previously reported a structural model for the catalytic periplasmic domain  
330 of Lnt (Vidal-Ingigliardi *et al.*, 2007). The  $\alpha$ - $\beta$ - $\beta$ - $\alpha$  fold is characteristic for members  
331 of the nitrilase, or carbon-nitrogen hydrolase, family (Pace & Brenner, 2001). The  
332 localization of the essential residues on a revised version of the structural model



333 obtained with PHYRE2 (Kelley & Sternberg, 2009) is shown in Fig. 5A. The less  
334 conserved essential residues are located on flexible loops. The structural modeling  
335 program I-TASSER (Zhang, 2008) did not lead to a more precise or different  
336 secondary structure prediction than that obtained with PHYRE2 (data not shown).  
337 Essential residue T481 is part of a  $\alpha$ -helix on top of the  $\alpha$ - $\beta$ - $\beta$ - $\alpha$  folded structure of  
338 Lnt. Replacement of T481 by arginine probably results in a conformational change in  
339 this  $\alpha$ -helix affecting the overall structure of Lnt. Levels of mutant Lnt (T481R)  
340 detected by immunoblotting are lower compared to wild type, suggesting that the  
341 mutated protein is less stable (Fig. S1, Fig. 6). G447 and L392 are found close to the  
342 active site triad. P353, F358 and M362 are located on a flexible loop  $\beta$ 5/ $\beta$ 6 (residues  
343 N336-L382) facing away from the active site in the model. E343 was previously  
344 identified as essential and required for the first step in the reaction. It is located on  
345 loop  $\beta$ 5/ $\beta$ 6 in close proximity to P353, F358 and M362 (Fig. 5A). Essential residue  
346 W237 is located on loop  $\beta$ 1/ $\alpha$ 1 (Vidal-Ingigliardi *et al.*, 2007). Similar structural  
347 models were predicted for the catalytic domain of Lnt from *C. glutamicum*  
348 (NCgl\_1424), *M. smegmatis* (MSMEG\_3860), *M. bovis* (BCG\_2070c), *S. coelicolor*  
349 (SCO\_1014) and *S. scabies* (SCAB\_83111) (Fig. 5B). Whereas P340-E343 is highly  
350 conserved, other essential residues located on this loop, including P353, F358 and  
351 M362, are not. These residues probably have a role in correctly positioning the  
352 conserved residues and/or substrate(s) to efficiently N-acylate lipoproteins.

353

354 *Essential residues located on the cytoplasmic loop CL-2 and the  $\beta$ 5/ $\beta$ 6 periplasmic*  
355 *loop in the catalytic domain are involved in the second step of the Lnt reaction*

356

357 A stable thioester acyl-enzyme intermediate is formed as part of the two-step reaction  
358 catalyzed by Lnt (Buddelmeijer & Young, 2010). The effect of the various amino acid  
359 substitutions resulting in non-functional Lnt on the formation of the thioester acyl-  
360 enzyme intermediate was assessed by alkylation of C387 by malPEG. The acyl group  
361 bound to C387 can be specifically cleaved by neutral hydroxylamine (HA), after  
362 which the free thiol group is accessible for alkylation by malPEG. Whole cells from  
363 strains expressing *lnt* variants from an arabinose-controlled promoter were treated  
364 with or without hydroxylamine and after protein precipitation samples were treated  
365 with malPEG. Wild type Lnt was alkylated by malPEG after HA treatment,  
366 corresponding to the presence of a thioester acyl-enzyme intermediate. Lnt (T481R)  
367 was alkylated both in the absence and presence of HA, suggesting that a thioester  
368 acyl-enzyme intermediate was not formed (Fig. 6). The amount of Lnt protein in this  
369 mutant is lower compared to Lnt. All other mutants tested were only alkylated in the  
370 presence of HA and thus all Lnt variants form a stable thioester acyl-enzyme  
371 intermediate (Fig. 6). This suggests that the essential residues identified in this study,  
372 including those located in CL-2, do not affect the first step of the N-acyltransferase  
373 reaction. We analyzed endogenous Lpp in the presence of non-functional Lnt variants  
374 in the *lnt* depletion strain by Western blotting to confirm that accumulation of apo-  
375 Lpp occurred under restrictive growth conditions. Since basal level expression of *lnt-  
376 myc<sub>2</sub>* from a strong pTrc promoter was sufficient to observe complementation in the  
377 case of wild type Lnt we did not add IPTG to avoid overexpression effects. The  
378 results shown in Fig. 7 demonstrate that apo-Lpp accumulates in strains where Lnt  
379 with R139A, P147L, F358V, M362I, T481R or P353S/L392H substitutions is the  
380 only form of enzyme present. Mature Lpp was detected in the presence of Lnt with  
381 G447S substitution. This residue probably plays a role in the efficiency of the

382 **reaction.** This demonstrates that the non-functional variants are affected in the second  
383 step of the apolipoprotein N-acyltransferase reaction.  
384

385 **Discussion**

386

387 Two acyltransferases and one signal peptidase are involved in the post-translational  
388 modification of lipoproteins but little is known about the molecular mechanism of the  
389 catalytic reactions. In this study we searched for essential residues that are either  
390 involved in the first step (auto-acylation of Lnt) or in the second step (N-acylation of  
391 apolipoprotein) of the N-acyltransferase reaction with the hope to identify residues  
392 involved in recognition or binding of substrate.

393 We identified essential residues and regions in Apolipoprotein *N*-acyltransferase (Lnt)  
394 of *E. coli* using random and site-directed PCR mutagenesis. A loop (CL-2) located  
395 between TMS-IV and TMS-V containing two hydrophobic segments and a conserved  
396 region, is located in the cytoplasm. Secondary structure predictions of CL-2 of Lnt  
397 demonstrated that the hydrophobic segments are  $\alpha$ -helices with high hydrophobicity  
398 and with a length of 18-20 amino acids that resemble transmembrane segments,  
399 however, our results demonstrate that these segments do not completely span the  
400 cytoplasmic membrane but may partially enter into the lipid bilayer. Reentrant loops  
401 are structural motifs that go halfway through the membrane and then turn back to the  
402 side of entry (Yan & Luo, 2010). These loops are found in transporters such as ion-  
403 channels and undergo conformational changes following substrate binding  
404 (Grunewald *et al.*, 2002, Yagur-Kroll *et al.*, 2009). In general, reentrant loops are  
405 shorter and less hydrophobic than transmembrane segments, they have very low  
406 hydrophobicity around the turning point and the two sides usually have different  
407 secondary structure (Yan & Luo, 2010). Since CL-2 does not have these  
408 characteristics, we hypothesize that the  $\alpha$ -helices form a tilted fold in the membrane

409 as described for intra-membrane proteases like GlpG in *E. coli* (Ben-Shem *et al.*,  
410 2007).

411 The first hydrophobic segment (TMS-IVa) has a pair of amino acids with opposite  
412 charges (E136 and R139 in *E. coli*), an arrangement found in all Lnt enzymes.  
413 Immediately downstream of this hydrophobic segment a conserved sequence  
414 (T<sub>144</sub>GFPWL<sub>149</sub>) was identified. In a previous study we have replaced several of these  
415 conserved amino acids by alanine and found that all Lnt variants were functional  
416 (Vidal-Ingigliardi *et al.*, 2007). Here we identified P147 as an important residue for  
417 Lnt function since replacement by leucine resulted in an inactive variant. By  
418 increasing the hydrophobicity of this region or by eliminating the proline-imposed  
419 turn, it is likely that the orientation of CL-2 changed resulting in an inactive  
420 conformation. Substitution of R139 by alanine also resulted in non-functional Lnt but  
421 no effect was observed when E136 was replaced by alanine. This result was surprising  
422 since E136 is as conserved as R139. Association of CL-2 with the membrane  
423 positions residues R139 and P147 in close proximity to the lipid bilayer.

424 Other essential residues were located in the catalytic periplasmic domain of Lnt and  
425 several on flexible loops facing away from the active site triad in the structural model  
426 predicted using PHYRE2. The degree of amino acid conservation was high for G447  
427 and T481 but low for P353, F358, M362 and L392 in sequences more than 15%  
428 identical to Lnt from *E. coli*. Residues P353, F358 and M362 are located on the  
429 flexible loop  $\beta 5/\beta 6$  between N336 and L382. Two other highly conserved residues,  
430 W237 and E343, were previously identified as being essential for Lnt activity. E343 is  
431 also located on the  $\beta 5/\beta 6$  loop and W237 is found on loop  $\beta 1/\alpha 1$  between G229 and  
432 G240 (Vidal-Ingigliardi *et al.*, 2007). All residues, except E343 (Buddelmeijer &  
433 Young, 2010), are involved in the second step of the reaction since the thioester acyl-

434 enzyme intermediate was formed but apolipoprotein accumulated and mature  
435 lipoprotein was not formed. The loops likely adopt a secondary structure upon  
436 substrate binding that brings the  $\alpha$ -amino group of *S*-diacylglyceryl-cysteine of the  
437 substrate in close proximity to the active site cysteine. We hypothesize that the CL-2  
438 region and the catalytic flexible loops are in close proximity and together contribute  
439 to efficient transfer of fatty acids onto apolipoprotein. One likely possibility is that the  
440 six transmembrane segments surround the partially membrane-embedded periplasmic  
441 and cytoplasmic loops adopting a barrel shape where the active site cavity becomes  
442 into close contact with the carbonyl group of PE. The bottom of this barrel shape is  
443 closed by CL-2 that sticks into the membrane towards the active site. This model  
444 suggests that access of the apolipoprotein substrate to the thioester-acyl group of Lnt  
445 occurs via lateral opening of the barrel.

446 The CL-2 region was targeted separately in an error-prone mutagenesis approach but  
447 essential residues were not identified. Lnt is a low abundant enzyme that we estimated  
448 to exist in about 100-200 molecules per cell (Hillmann *et al.*, 2011). This might  
449 explain why mutations resulting in slightly less efficient enzyme were not identified  
450 and that only mutations leading to a severe phenotype were obtained in our screen.  
451 The fact that an increase in protein level of non-functional Lnt enzymes restores  
452 growth of the *lnt* depletion strain supports this hypothesis.

453 Together with the catalytic triad, residues E389 and E343 are conserved even in the  
454 more distantly related actinobacterial Lnt-s (Widdick *et al.*, 2011). Lipoproteins from  
455 *C. glutamicum* are modified with a *N*-acyl (C16:0) and *S*-diacylglyceryl (C16:0,  
456 C18:1) moiety like in *E. coli* (Mohiman *et al.*, 2012), whereas lipoproteins in  
457 *Mycobacteria* contain tuberculostearic acid (C19:0) as part of their *S*-diacylglyceryl  
458 moiety and palmitic or tuberculostearic acid as *N*-acyl group (Brülle *et al.*, 2010,

459 Tschumi *et al.*, 2009, Brülle *et al.*, 2013). The fatty acid composition of the *N*-acyl  
460 group of lipoproteins in *S. scabies* varies from C15 to C19, but details about the *N*-  
461 acyl and *S*-diacyl composition of lipoproteins are not known. It appears likely that Lnt  
462 attaches palmitic acid and tuberculostearic acid onto apolipoprotein. E343 is  
463 involved in the first step of the Lnt reaction, formation of the thioester acyl-enzyme  
464 intermediate between fatty acid and Lnt (Buddelmeijer & Young, 2010), suggesting  
465 that it is involved in the interaction and/or hydrolysis of phospholipid. Lnt of *M.*  
466 *smegmatis* could not complement a *lnt* depletion strain of *E. coli* (Tschumi *et al.*,  
467 2009). It is conceivable that the enzyme is not inserted into the *E. coli* membrane in  
468 an active conformation and therefore, unable to correctly interact with the  
469 apolipoprotein substrate. Alternatively, the lipid moiety of the substrate might not be  
470 correctly fitted into the enzyme for *N*-acylation to occur. Interestingly, the *sn*-1  
471 position of mycobacterial phospholipids is occupied by C18 related fatty acids and *sn*-  
472 2 by palmitic acid (Okuyama *et al.*, 1967). This implies that mycobacterial Lnt uses  
473 fatty acids only from *sn*-2 and not from *sn*-1, as is the case for the *E. coli* enzyme  
474 (Jackowski & Rock, 1986, Hillmann *et al.*, 2011, Nakayama *et al.*, 2012). The  
475 substrate specificity for both the phospholipid and the apolipoprotein substrates is  
476 different for Lnt of different species. The 3D structural models for the catalytic  
477 domain of all Lnt enzymes identified to date are highly similar and illustrate that the  
478 mechanism of *N*-acyltransferase is identical. However, several residues located on the  
479 disordered  $\beta 5/\beta 6$  loop are essential but poorly conserved and may play a role in  
480 apolipoprotein substrate specificity. Sortases are membrane-associated enzymes that  
481 covalently link proteins to peptidoglycan or polymerize proteins to construct pili in  
482 Gram-positive bacteria (Spirig *et al.*, 2011). Structural analyses have demonstrated  
483 that the sortase catalytic mechanism involves opening and closing of loops upon

484 binding of substrate (Suree *et al.*, 2009) and that the loops are involved in substrate  
485 recognition (Bentley *et al.*, 2007). We hypothesize that the  $\beta 5/\beta 6$  flexible loop in Lnt  
486 is involved in recognition of the *S*-diacylglyceryl-cysteine of apolipoprotein and that  
487 binding of substrate leads to conformational changes and a substrate-bound closed  
488 form of Lnt similar to sortases.

489 Our data suggest that cytoplasmic and periplasmic membrane-associated loops are  
490 required for *N*-acylation of lipoproteins and that specific amino acids in these loops  
491 may be involved in substrate recognition and/or binding. Ongoing three-dimensional  
492 structural analyses of full-length Lnt enzymes are expected to highlight differences  
493 between various species and to establish how the enzyme interacts with its substrates.

494



## 495 | Experimental Procedures

496

### 497 *Bacterial strains and growth conditions*

498

499 Strains and plasmids are listed in Table S3. Cells were grown in LB medium  
500 supplemented with ampicillin (100 µg/ml) when required. Induction of gene  
501 expression from a pBAD promoter was obtained by addition of L-arabinose (0.2%)  
502 and from pTrc promoter with IPTG (1 mM).

503

### 504 *Standard DNA manipulation*

505

506 Silent sites were introduced in *Int* flanking the cytoplasmic loop 2 (CL-2) coding  
507 sequence using the QuickChange II site-directed mutagenesis protocol (Stratagene).  
508 The procedure included a two-step PCR reaction where large oligomers were  
509 synthesized in the first PCR reaction that contained only one of the two mutant  
510 primers and synthesis of the entire mutated plasmid in the second reaction that  
511 contained the large oligomers and DNA template. A PmlI site was introduced at  
512 **codon 120 (T120)** with primers T120-PmlI\_for and T120-PmlI\_rev, an NheI site at  
513 L185 with primers L185-NheI\_for and L185-NheI\_rev, and a StuI site at A186 with  
514 primers Del\_StuI-5' and Del\_StuI-3' (primers are listed in Table S3). **The PmlI and**  
515 **NheI sites were combined in pCHAP9610 and the CL-2 segment was removed by**  
516 **deleting the PmlI-StuI fragment, resulting in pCHAP9609.**

517 The gene encoding *Int* was amplified by PCR using primers 5'-cutE and IntXbaIrev  
518 and was inserted between the EcoRI and XbaI sites of pUC18, the 3' end of *Int* is in  
519 frame with a double *c-myc* tag, resulting in pCHAP7547. Site-directed mutagenesis

520 was used to create alanine substitutions for cysteines (C23 and C62) on pCHAP7547  
521 resulting in plasmid pCHAP7556 (encoding Lnt<sup>C23AC62A</sup>) (Buddelmeijer & Young,  
522 2010). C387 was replaced by serine or alanine as described previously (Buddelmeijer  
523 & Young, 2010, Vidal-Ingigliardi *et al.*, 2007). An EcoRI-HindIII fragment from  
524 pCHAP7556 was inserted into pDSW204 and pBAD18, digested with the same  
525 enzymes, resulting in pCHAP7721 and pCHAP7626, respectively.

526

#### 527 *Screening for non-functional Lnt variants*

528

529 Random PCR mutagenesis was performed on *lnt* using GeneMorph II EZ Clone  
530 (Agilent) on plasmid pCHAP7547 or pCHAP7721 as template for full-length *lnt* and  
531 on plasmid pCHAP9515 (Hillmann *et al.*, 2011) as template for the CL-2 region of  
532 *lnt*. Plasmid DNA from mutagenesis on pCHAP7547 or pCHAP7721 was either  
533 directly transformed into the *lnt* depletion strain (PAP9502) and cells were selected  
534 for growth on LB agar plates containing ampicillin and L-arabinose, or DNA was  
535 transformed into DH5 $\alpha$  and *lnt* fragments re-cloned into a new vector backbone  
536 before transformation into PAP9502. Colonies were re-streaked on LB agar plates  
537 containing ampicillin with or without L-arabinose. In the case of mutagenesis of the  
538 CL-2 region, fragments were first re-cloned in template pCHAP9610 using PmlI and  
539 NheI sites, and then screened for non-complementing phenotype in PAP9502 on  
540 plates lacking L-arabinose.

541 For detailed analysis of the mutants, strains were grown in medium without 0.2% L-  
542 arabinose starting from one colony grown on a fresh plate with L-arabinose. Cells  
543 were grown for two hours at 37°C and serially diluted from 10<sup>-1</sup> to 10<sup>-6</sup>. Cells were

544 spotted on plates with or without L-arabinose in the presence or absence of 1 mM

545 IPTG.

546

547 *Determination of thioester-acyl enzyme intermediate*

548

549 The procedure to determine the acylation state of Lnt was described previously  
550 (Buddelmeijer & Young, 2010). Briefly, whole cells were treated with 1 M neutral  
551 hydroxylamine (HA) that cleaves specifically thioester bonds resulting in the case of  
552 Lnt in a free thiol group of active site residue C387. Control samples were incubated  
553 with 1 M Tris-HCl (pH 7.0). Subsequent labeling of free SH groups with maleimide-  
554 polyethyleneglycol (malPEG) resulted in alkylation of SH by maleimide and in an  
555 increase of molecular weight by 5 kDa.

556

557 *Cysteine protection assay*

558

559 Accessibility of cysteine residues in whole cells was determined using the substituted  
560 cysteine accessibility method (SCAM) as described previously (Bogdanov *et al.*,  
561 2005, Pailler *et al.*, 2012). In summary, whole cells were resuspended in phosphate  
562 buffer (PB) containing EDTA and treated with the membrane-impermeable reagent  
563 (2-sulfonatoethyl)-methane thiosulfonate) (30 mM MTSES). The reaction was  
564 quenched with 100 mM L-Cysteine and cells were washed twice with PB and  
565 resuspended in PB. Proteins were precipitated with either trichloroacetic acid (10%  
566 TCA) or methanol/chloroform. Pellets were resuspended in PEG buffer (1 M Tris-  
567 HCl pH 7.0, 10 M urea, 1% SDS, 1 mM EDTA). 0.2 mM MalPEG was added to all

568 samples and incubated for 60 minutes at room temperature. Proteins were precipitated  
569 with TCA or MeOH/CHCl<sub>3</sub> and analyzed by SDS-PAGE and Western blotting.

570

#### 571 *Gel electrophoresis and Western blotting*

572

573 Protein samples were solubilized in SDS sample buffer containing 4 mM DTT or 5%  
574  $\beta$ -mercaptoethanol ( $\beta$ -ME) and were heated at 100°C for 5 minutes. Samples were  
575 separated on polyacrylamide gels (Biorad) and transferred onto nitrocellulose  
576 membranes. Lnt was detected by chemiluminescence (Pierce) using anti-c-myc  
577 antibodies (Sigma) and secondary anti-rabbit antibodies conjugated with horseradish  
578 peroxidase (HRP). Lpp was detected with antibodies against Lpp (a gift from H.  
579 Tokuda).

580

#### 581 *Bioinformatics analyses*

582

583 The presence of Lnt orthologs was investigated in all bacteria and separately in the  $\gamma$ -  
584 *proteobacteria* taxonomic class using *E. coli* Lnt (NP\_415190) as bait as follows. For  
585 each genus, a single representative and sequenced genome originating from the NCBI  
586 repository (<ftp://ftp.ncbi.nih.gov/genomes/Bacteria/>) was selected alphabetically and  
587 submitted to TBLASTN analysis with Lnt using an in-house C# script. Taxonomic  
588 relationships between individual species were determined with the SyntTax web  
589 service (Oberto, 2013). Normalized genomic BLAST scores were used to measure the  
590 degree of orthology as described (Lerat *et al.*, 2003). The threshold was set  
591 empirically at the value of 15% in agreement with the annotations of the selected  
592 genomes. Lnt orthologs with a score equal or greater than 15% were collected and

593 submitted to multiple protein sequence alignment using the Clustal Omega web  
594 service (<https://www.ebi.ac.uk/Tools/msa/clustalo/>).  
595

596 **Acknowledgements**

597

598 We thank all members of the Molecular Genetics Unit and Biology and Genetics of  
599 the Cell Wall Unit of the Institut Pasteur for their support. We thank Nicolas Bayan  
600 and colleagues for helpful discussions and constructive comments on the work.

601 **References**

602

- 603 Asanuma, M., K. Kurokawa, R. Ichikawa, K. Ryu, J. Chae, N. Dohmae, B.L. Lee & H.  
604 Nakayama, (2010) Structural evidence of  $\alpha$ -aminoacylated lipoproteins of  
605 *Staphylococcus aureus*. *FEBS J.* **278**: 716-728.
- 606 Ben-Shem, A., D. Fass & E. Bibi, (2007) Structural basis for intramembrane  
607 proteolysis by rhomboid serine proteases. *Proc. Natl. Acad. Sci. U S A* **104**:  
608 462-466.
- 609 Bentley, M.L., H. Gaweska, J.M. Kielec & D.G. McCafferty, (2007) Engineering the  
610 substrate specificity of *Staphylococcus aureus* Sortase A. The beta6/beta7  
611 loop from SrtB confers NPQTN recognition to SrtA. *J. Biol. Chem.* **282**:  
612 6571-6581.
- 613 Bogdanov, M., W. Zhang, J. Xie & W. Dowhan, (2005) Transmembrane protein  
614 topology mapping by the substituted cysteine accessibility method  
615 (SCAM): applications to lipid-specific membrane protein topogenesis.  
616 *Methods (San Diego, Calif)* **36**: 148-171.
- 617 Brülle, J.K., T. Grau, A. Tschumi, Y. Auchli, R. Burri, S. Polsfuss, P.M. Keller, P.  
618 Hunziker & P. Sander, (2010) Cloning, expression and characterization of  
619 *Mycobacterium tuberculosis* lipoprotein LprF. *Biochem. Biophys. Res.*  
620 *Comm.* **391**: 679-684.
- 621 Brülle, J.K., A. Tschumi & P. Sander, (2013) Lipoproteins of slow-growing  
622 *Mycobacteria* carry three fatty acids and are N-acylated by apolipoprotein  
623 N-acyltransferase BCG\_2070c. *BMC Microbiol.* **13**: 223.
- 624 Buddelmeijer, N. & R. Young, (2010) The essential *Escherichia coli*  
625 Apolipoprotein N-acyltransferase (Lnt) exists as an extracytoplasmic  
626 thioester acyl-enzyme intermediate. *Biochemistry* **49**: 341-346.
- 627 Garbe, T., D. Harris, M. Vordermeier, R. Lathigra, J. Ivanyi & D. Young, (1993)  
628 Expression of the *Mycobacterium tuberculosis* 19-kilodalton antigen in  
629 *Mycobacterium smegmatis*: immunological analysis and evidence of  
630 glycosylation. *Infect. Immun.* **61**: 260-267.
- 631 Grunewald, M., D. Menaker & B.I. Kanner, (2002) Cysteine-scanning mutagenesis  
632 reveals a conformationally sensitive reentrant pore-loop in the glutamate  
633 transporter GLT-1. *J. Biol. Chem.* **277**: 26074-26080.
- 634 Hillmann, F., M. Argentini & N. Buddelmeijer, (2011) Kinetics and phospholipid  
635 specificity of apolipoprotein N-acyltransferase. *J. Biol. Chem.* **286**: 27936-  
636 27946.
- 637 Hussain, M., S. Ichihara & S. Mizushima, (1980) Accumulation of glyceride-  
638 containing precursor of the outer membrane lipoprotein in the  
639 cytoplasmic membrane of *Escherichia coli* treated with globomycin. *J. Biol.*  
640 *Chem.* **255**: 3707-3712.
- 641 Hussain, M., S. Ichihara & S. Mizushima, (1982) Mechanism of signal peptide  
642 cleavage in the biosynthesis of the major lipoprotein of the *Escherichia*  
643 *coli* outer membrane. *J. Biol. Chem.* **257**: 5177-5182.
- 644 Inukai, M., K. Takeuchi, K. Shimizu & M. Arai, (1978) Mechanism of action of  
645 globomycin. *J. Antibiot.* **31**: 1203-1205.

646 Jackowski, S. & C.O. Rock, (1986) Transfer of fatty acids from the 1-position of  
647 phosphatidylethanolamine to the major outer membrane lipoprotein of  
648 *Escherichia coli*. *J. Biol. Chem.* **261**: 11328-11333.

649 Kelley, L.A. & M.J. Sternberg, (2009) Protein structure prediction on the Web: a  
650 case study using the Phyre server. *Nature protocols* **4**: 363-371.

651 Kovacs-Simon, A., R.W. Titball & S.L. Michell, (2010) Lipoproteins of bacterial  
652 pathogens. *Infect. Immun.* **79**: 548-561.

653 Kurokawa, K., M.S. Kim, R. Ichikawa, K.H. Ryu, N. Dohmae, H. Nakayama & B.L.  
654 Lee, (2012a) Environment-mediated accumulation of diacyl lipoproteins  
655 over their triacyl counterpart in *Staphylococcus aureus*. *J. Bacteriol.* **194**:  
656 3299-3306.

657 Kurokawa, K., Lee H., Roh K.-B., Asanuma M., Kim Y S, Nakayama H, Shiratsuchi  
658 A., choi Y, Takeuchi O, Kang H J, Dohmae N, Nakanishi Y, Akira S, Sekimizu  
659 K, Lee B L, (2009) The triacylated ATP binding cluster transporter  
660 substrate-binding lipoprotein of *Staphylococcus aureus* functions as a  
661 native ligand for Toll-like receptor 2. *J. Biol. Chem.* **284**: 8406-8411.

662 Kurokawa, K., K.H. Ryu, R. Ichikawa, A. Masuda, M.S. Kim, H. Lee, J.H. Chae, T.  
663 Shimizu, T. Saitoh, K. Kuwano, S. Akira, N. Dohmae, H. Nakayama & B.L.  
664 Lee, (2012b) Novel bacterial lipoprotein structures conserved in low-GC  
665 content Gram-positive bacteria are recognized by Toll-like receptor 2. *J.*  
666 *Biol. Chem.* **287**: 13170-13181.

667 Lai, J.-S., W.M. Philbrick & H.C. Wu, (1980) Acyl moieties in phospholipids are the  
668 precursors for the fatty acids in murein lipoprotein in *Escherichia coli*. *J.*  
669 *Biol. Chem.* **255**: 5384-5387.

670 Lerat, E., V. Daubin & N.A. Moran, (2003) From gene trees to organismal  
671 phylogeny in prokaryotes: the case of the gamma-Proteobacteria. *PLoS*  
672 *Biol.* **1**: E19.

673 Malinverni, J.C., J. Werner, S. Kim, J.G. Sklar, D. Kahne, R. Misra & T.J. Silhavy,  
674 (2006) YfiO stabilizes the YaeT complex and is essential for outer  
675 membrane protein assembly in *Escherichia coli*. *Mol. Microbiol.* **61**: 151-  
676 164.

677 Mohiman, N., M. Argentini, S.M. Batt, D. Cornu, M. Masi, L. Eggeling, G. Besra & N.  
678 Bayan, (2012) The ppm operon is essential for acylation and glycosylation  
679 of lipoproteins in *Corynebacterium glutamicum*. *PloS One* **7**: e46225.

680 Nakayama, H., K. Kurokawa & B.L. Lee, (2012) Lipoproteins in bacteria:  
681 structures and biosynthetic pathways. *FEBS J.* **279**: 4247-4268.

682 Oberto, J., (2013) SyntTax: a web server linking synteny to prokaryotic  
683 taxonomy. *BMC bioinformatics* **14**: 4.

684 Okuyama, H., T. Kankura & S. Nojima, (1967) Positional distribution of fatty acids  
685 in phospholipids from Mycobacteria. *J. Biochem.* **61**: 732-737.

686 Pace, H.C. & C. Brenner, (2001) The nitrilase superfamily: classification, structure  
687 and function. *Genome Biol.* **2**: 1-9.

688 Pailler, J., W. Aucher, M. Pires & N. Buddelmeijer, (2012)  
689 Phosphatidylglycerol::prolipoprotein diacylglyceryltransferase (Lgt) of *E.*  
690 *coli* has seven transmembrane segments and its essential residues are  
691 embedded in the membrane. *J. Bacteriol.* **194**: 2142-2151.

692 Paradis-Bleau, C., M. Markovski, T. Uehara, T.J. Lupoli, S. Walker, D.E. Kahne &  
693 T.G. Bernhardt, (2010) Lipoprotein cofactors located in the outer  
694 membrane activate bacterial cell wall polymerases. *Cell* **143**: 1110-1120.



695 Robichon, C., D. Vidal-Ingigliardi & A.P. Pugsley, (2005) Depletion of  
696 apolipoprotein N-acyltransferase causes mislocalization of outer  
697 membrane lipoproteins in *Escherichia coli*. *J. Biol. Chem.* **280**: 974-983.

698 Sankaran, K. & H.C. Wu, (1994) Lipid modification of bacterial prolipoprotein.  
699 Transfer of diacylglyceryl moiety from phosphatidylglycerol. *J. Biol. Chem.*  
700 **269**: 19701-19706.

701 Sartain, M.J. & J.T. Belisle, (2009) N-Terminal clustering of the O-glycosylation  
702 sites in the *Mycobacterium tuberculosis* lipoprotein SodC. *Glycobiology* **19**:  
703 38-51.

704 Serebryakova, M.V., I.A. Demina, M.A. Galyamina, I.G. Kondratov, V.G. Ladygina &  
705 V.M. Govorun, (2011) The acylation state of surface lipoproteins of  
706 mollicute *Acholeplasma laidlawii*. *J. Biol. Chem.* **286**: 22769-22776.

707 Spirig, T., E.M. Weiner & R.T. Clubb, (2011) Sortase enzymes in Gram-positive  
708 bacteria. *Mol. Microbiol.* **82**: 1044-1059.

709 Sugai, M. & H.C. Wu, (1992) Export of the outer membrane lipoprotein is  
710 defective in *secD*, *secE*, and *secF* mutants of *Escherichia coli*. *J. Bacteriol.*  
711 **174**: 2511-2516.

712 Suree, N., C.K. Liew, V.A. Villareal, W. Thieu, E.A. Fadeev, J.J. Clemens, M.E. Jung &  
713 R.T. Clubb, (2009) The structure of the *Staphylococcus aureus* sortase-  
714 substrate complex reveals how the universally conserved LPXTG sorting  
715 signal is recognized. *J. Biol. Chem.* **284**: 24465-24477.

716 Sutcliffe, I.C. & R.R.B. Russell, (1995) Lipoproteins of gram-positive bacteria. *J.*  
717 *Bacteriol.* **177**: 1123-1128.

718 Thompson, B.J., D.A. Widdick, M.G. Hicks, G. Chandra, I.C. Sutcliffe, T. Palmer &  
719 M.I. Hutchings, (2010) Investigating lipoprotein biogenesis and function  
720 in the model Gram-positive bacterium *Streptomyces coelicolor*. *Mol.*  
721 *Microbiol.* **77**: 943-957.

722 Tschumi, A., C. Nai, Y. Auchli, P. Hunziker, P. Gehrig, P. Keller, T. Grau & P. Sander,  
723 (2009) Identification of apolipoprotein N-acyltransferase (Lnt) in  
724 Mycobacteria. *J. Biol. Chem.* **284**: 27146-27156.

725 Typas, A., M. Banzhaf, B. van den Berg van Saparoea, J. Verheul, J. Biboy, R.  
726 Nichols, M. Zietek, K. Beilharz, K. Kannenberg, M. von Rechenberg, E.  
727 Breukink, T. den Blaauwen, C.A. Gross & W. Vollmer, (2010) Regulation of  
728 peptidoglycan synthesis by outer-membrane proteins. *Cell* **143**: 1097-  
729 1109.

730 Vidal-Ingigliardi, D., S. Lewenza & N. Buddelmeijer, (2007) Identification of  
731 essential residues in apolipoprotein N-acyltransferase, a member of the  
732 CN hydrolase family. *J. Bacteriol.* **189**: 4456-4464.

733 Widdick, D.A., M.G. Hicks, B.J. Thompson, A. Tschumi, G. Chandra, I.C. Sutcliffe, J.K.  
734 Brulle, P. Sander, T. Palmer & M.I. Hutchings, (2011) Dissecting the  
735 complete lipoprotein biogenesis pathway in *Streptomyces scabies*. *Mol.*  
736 *Microbiol.* **80**: 1395-1412.

737 Wu, H.C., M. Tokunaga, H. Tokunaga, S. Hayashi & C.Z. Giam, (1983)  
738 Posttranslational modification and processing of membrane lipoproteins  
739 in bacteria. *Journal of cellular biochemistry* **22**: 161-171.

740 Xiao, Y., K. Gerth, R. Müller & D. Wall, (2012) Myxobacterium-produced  
741 antiobiotic TA (Myxovirescin) inhibits type II signal peptidase.  
742 *Antimicrob. Agents and Chemother.* **56**: 2014-2021.

743 Xiao, Y. & D. Wall, (2014) Genetic redundancy, proximity, and functionality of  
744 lspA, the target of antibiotic TA, in the Myxococcus xanthus producer  
745 strain. *Journal of bacteriology* **196**: 1174-1183.  
746 Yagur-Kroll, S., A. Ido & O. Amster-Choder, (2009) Spatial arrangement of the  
747 beta-glucoside transporter from *Escherichia coli*. *J. Bacteriol.* **191**: 3086-  
748 3094.  
749 Yan, C. & J. Luo, (2010) An analysis of reentrant loops. *Protein J.* **29**: 350-354.  
750 Zhang, Y., (2008) I-TASSER server for protein 3D structure prediction. *BMC*  
751 *bioinformatics* **9**: 40.  
752 Zuckert, W.R., (2014) Secretion of Bacterial Lipoproteins: Through the  
753 Cytoplasmic Membrane, the Periplasm and Beyond. *Biochim. Biophys.*  
754 *Acta*.  
755  
756

757 **Figure legends**

758

759 Fig. 1. Membrane topology, essential residues and protein sequence of cytoplasmic  
760 loop 2 (CL-2) of Lnt from *E. coli*. Lnt has six transmembrane segments (TMS)  
761 indicated as I-VI and is located in the cytoplasmic membrane (CM). Cytoplasmic loop  
762 2 (CL-2) corresponds to residues R<sub>114</sub>-R<sub>193</sub>. Positively (+) and negatively (-) charged  
763 residues are indicated. Residues substituted by cysteine are indicated as white circles.  
764 Amino acid sequence of CL-2 is shown including TMS-IVa and TMS-IVb  
765 (undelined). Essential residues (red), important residues (bleu) and the active site triad  
766 (green) are indicated.

767

768 Fig. 2. Complementation spot-dilution assay of Lnt mutants in the *lnt* depletion strain.  
769 Strains were depleted from Lnt by growing the cells in the absence of L-arabinose in  
770 liquid medium. Serial dilutions were spotted on plates with or without L-arabinose  
771 (0.2%) (+ARA) in the presence or absence of IPTG (1 mM). Lnt(C387S) is an active  
772 site mutant of Lnt.

773

774 Fig. 3. Accessibility to MTSES of substituted cysteine residues in whole cells.  
775 Western blot analysis of Lnt-myc<sub>2</sub> from cell lysates treated with (+) or without (-)  
776 MTSES and subsequent alkylation by malPEG. Lnt enzymes containing active site  
777 cysteine 387 (+) or C387S substitution are included as controls (-). Individual cysteine  
778 substitutions are introduced in C387S variant of Lnt. All Lnt proteins contain C23A  
779 and C62A substitutions. Cysteine substitutions were introduced in known cytoplasmic  
780 and periplasmic loops of Lnt. Protein bands corresponding to Lnt-c-myc<sub>2</sub> (Lnt, 57

781 kDa) and the alkylated form of Lnt-c-myc<sub>2</sub> (Lnt-malPEG) are indicated. Proteins were  
782 detected with antibodies against c-myc.

783

784 Fig. 4. LOGO representation of conserved residues in Lnt of  $\gamma$ -proteobacteria. Lnt of  
785  $\gamma$ -proteobacteria were aligned by Clustal Omega and conserved residues were  
786 identified (see also Fig. S2). Active site residues are shown in red. Essential and  
787 important residues identified in this study are indicated in the LOGO coordinates in  
788 red and blue, respectively. Amino acid substitutions of all mutants identified here are  
789 shown in bold and in blue. The transmembrane segments are shown in bold as TMS-I  
790 to VI, including membrane associated sequences TMS-IVa and TMS-IVb. Asterisks  
791 indicate gaps in the alignment, some of which result in conserved residues excluding  
792 the *E. coli* sequence. The predicted secondary structure of the periplasmic catalytic  
793 domain is indicated (see also Fig. 5).

794

795 Fig. 5. Localization of essential residues on a structural model of Lnt. A structural  
796 model for the catalytic periplasmic domain of Lnt (residues T221-N488) was obtained  
797 using PHYRE2 program (A). The catalytic triad (E267-K335-C378) is indicated in  
798 red, essential residues are shown in green. W237, E343 were identified previously  
799 (Vidal-Ingigliardi *et al.*, 2007). Loop  $\beta$ 5/ $\beta$ 6 is colored blue. Structural models of the  
800 catalytic domain of Lnt orthologues are shown (B). Color indications are the same as  
801 for *E. coli*.

802

803 Fig. 6. Acylation state of C387 in non-functional Lnt variants. Western blot analysis  
804 of cell lysates from cells treated with (+) or without (-) hydroxylamine (HA) and  
805 subsequent alkylation with malPEG. Detection of Lnt-c-myc<sub>2</sub> (Lnt) and alkylated Lnt-

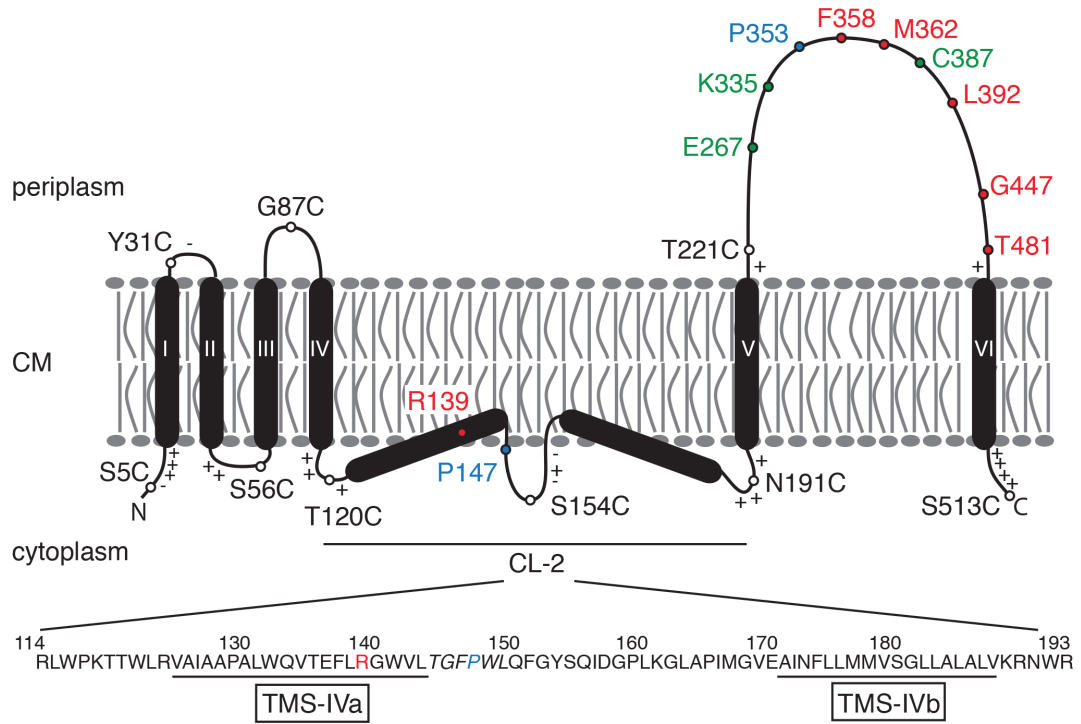
806 c-myc<sub>2</sub> (Lnt-malPEG) was obtained with anti-myc antibodies. Lnt<sup>C23AC62A</sup>  
807 corresponds to Lnt with C23A and C62A substitutions but containing active site  
808 C387.

809

810 Fig. 7. Accumulation of apo-Lpp in presence of non-functional Lnt. Western blot  
811 analysis of Lpp in whole cell lysates from the *lnt* depletion strain expressing non-  
812 functional *lnt* variants as described in Fig. 6 under depletion conditions (in the  
813 absence of L-arabinose). Lysates from a wild type strain and a  $\Delta lpp$  mutant were  
814 loaded as controls. Detection of Lpp was obtained with anti-Lpp antibodies, mature  
815 Lpp (Lpp) and the apo form of Lpp (apo-Lpp) are indicated.

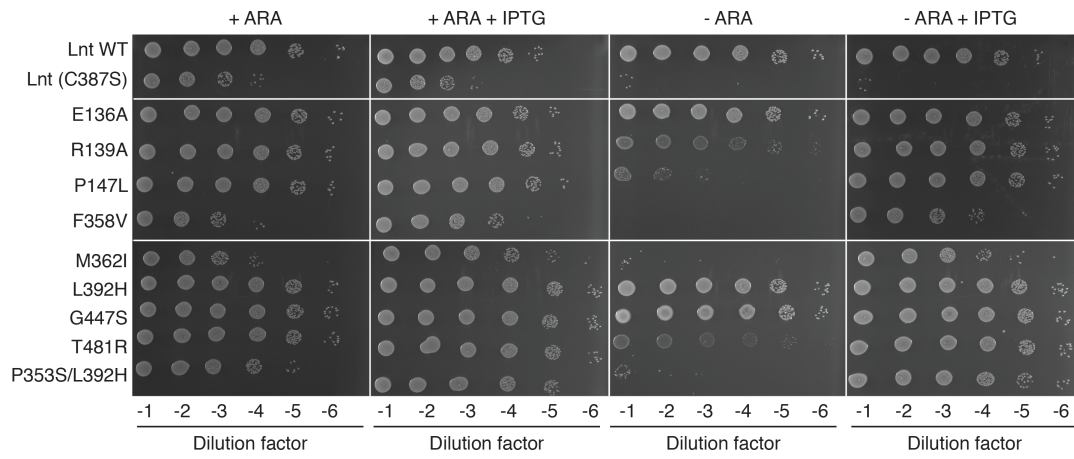
816

FIGURE 1



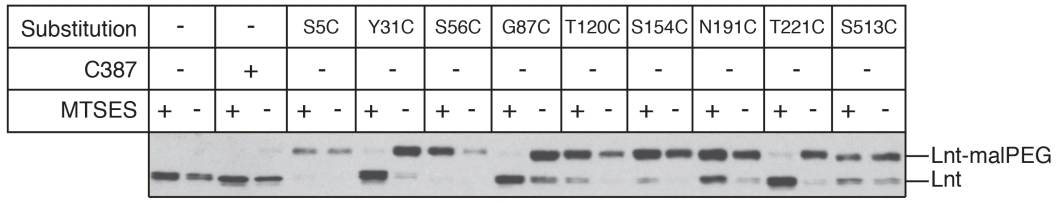
817  
818  
819

FIGURE 2



820  
821  
822

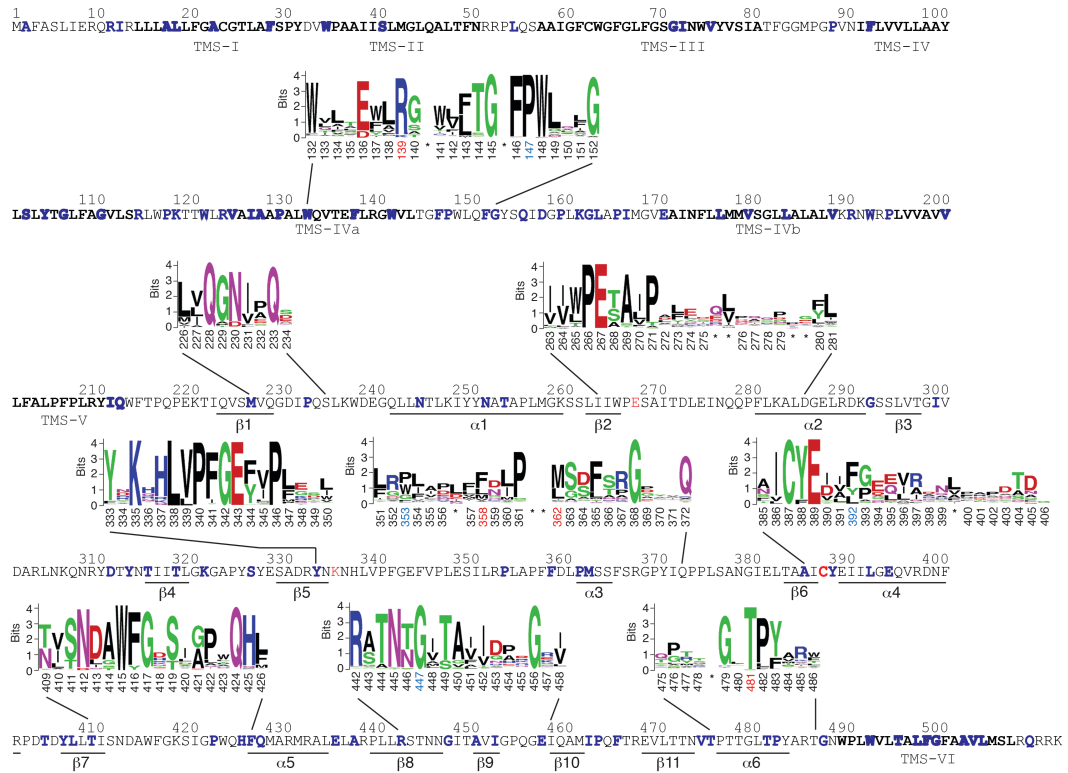
FIGURE 3



823  
824  
825



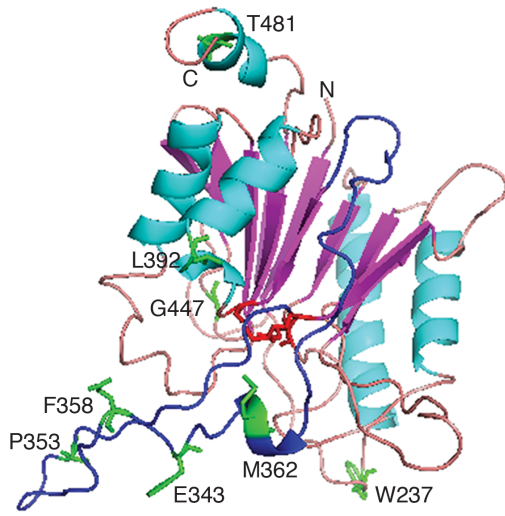
FIGURE 4



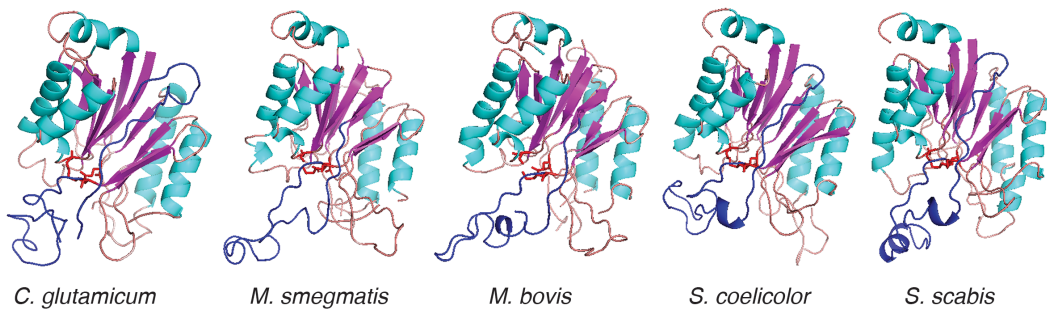
826  
 827  
 828

FIGURE 5

A

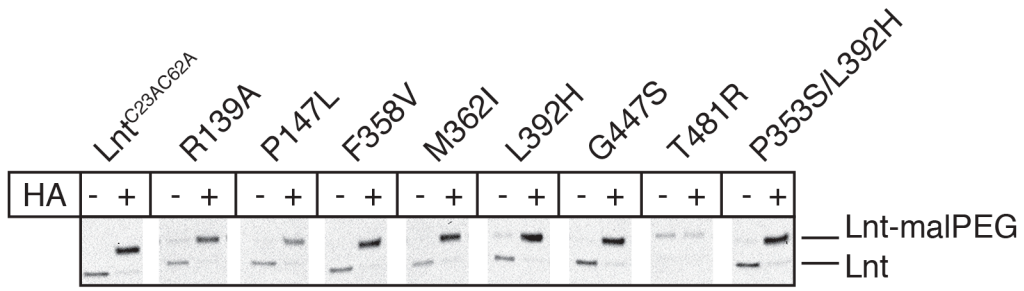


B



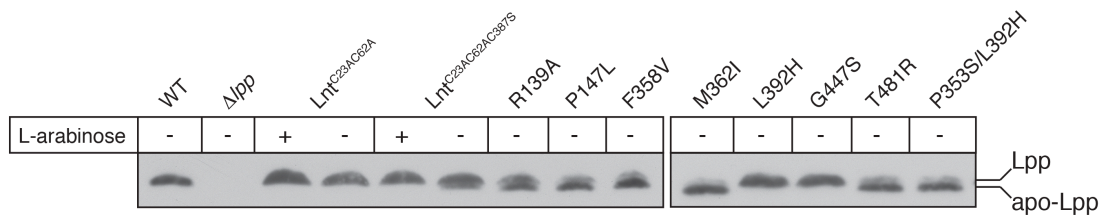
829  
830  
831

FIGURE 6



832  
833  
834

FIGURE 7



835  
836  
837