

1 **Disorder is a critical component of lipoprotein sorting in Gram-negative bacteria**

2

3 Jessica El Rayes^{1,2§}, Joanna Szewczyk^{1,2§}, Michael Deghelt^{1,2}, André Matagne³, Bogdan I.

4 Iorga⁴, Seung-Hyun Cho^{1,2}, and Jean-François Collet^{1,2*}

5

6

7 ¹WELBIO, Avenue Hippocrate 75, 1200 Brussels, Belgium.

8 ²de Duve Institute, Université catholique de Louvain, Avenue Hippocrate 75, 1200 Brussels,
9 Belgium.

10 ³Centre d'ingénierie des Protéines, Institut de Chimie B6, Université de Liège, Allée de la
11 Chimie 3, 4000 Liège, Sart Tilman, Belgium.

12 ⁴Université Paris-Saclay, CNRS UPR 2301, Institut de Chimie des Substances Naturelles,
13 91198 Gif-sur-Yvette, France.

14

15 [§]Both authors contributed equally to the work

16

17 *Correspondence: jfcollet@uclouvain.be

18 **Abstract (150 max)**

19

20 **Gram-negative bacteria express structurally diverse lipoproteins in their envelope. Here**
21 **we found that approximately half of lipoproteins destined to the *Escherichia coli* outer**
22 **membrane display an intrinsically disordered linker at their N-terminus. Intrinsically**
23 **disordered regions are common in proteins, but establishing their importance *in vivo* has**
24 **remained challenging. Here, as we sought to unravel how lipoproteins mature, we**
25 **discovered that unstructured linkers are required for optimal trafficking by the Lol**
26 **lipoprotein sorting system: linker deletion re-routes three unrelated lipoproteins to the**
27 **inner membrane. Focusing on the stress sensor RcsF, we found that replacing the linker**
28 **with an artificial peptide restored normal outer membrane targeting only when the**
29 **peptide was of similar length and disordered. Overall, this study reveals the role played**
30 **by intrinsic disorder in lipoprotein sorting, providing mechanistic insight into the**
31 **biogenesis of these proteins and suggesting that evolution can select for intrinsic disorder**
32 **that supports protein function.**

33 **Introduction**

34 The cell envelope is the morphological hallmark of *Escherichia coli* and other Gram-negative
35 bacteria. It is composed of the inner membrane, a classical phospholipid bilayer, as well as the
36 outer membrane, an asymmetric bilayer with phospholipids in the inner leaflet and
37 lipopolysaccharides in the outer leaflet¹. This lipid asymmetry enables the outer membrane to
38 function as a barrier that effectively prevents the diffusion of toxic compounds in the
39 environment into the cell. The inner and outer membranes are separated by the periplasm, a
40 viscous compartment that contains a thin layer of peptidoglycan also known as the cell wall¹.
41 The cell envelope is essential for growth and survival, as illustrated by the fact that several
42 antibiotics such as the β -lactams target mechanisms of envelope assembly. Mechanisms
43 involved in envelope biogenesis and maintenance are therefore attractive targets for novel
44 antibacterial strategies.

45
46 Approximately one-third of *E. coli* proteins are targeted to the envelope, either as soluble
47 proteins present in the periplasm or as proteins inserted in one of the two membranes². While
48 inner membrane proteins cross the lipid bilayer via one or more hydrophobic α -helices, proteins
49 inserted in the outer membrane generally adopt a β -barrel conformation³. Another important
50 group of envelope proteins is the lipoproteins, which are globular proteins anchored to one of
51 the two membranes by a lipid moiety. Lipoproteins carry out a variety of important functions
52 in the cell envelope: they participate in the biogenesis of the outer membrane by inserting
53 lipopolysaccharide molecules^{4,5} and β -barrel proteins⁶, they function as stress sensors triggering
54 signal transduction cascades when envelope integrity is altered⁷, and they control processes that
55 are important for virulence⁸. The diverse roles played by lipoproteins in the cell envelope has
56 drawn a lot of attention lately, revealing how crucial these proteins are in a wide range of vital
57 processes and identifying them as attractive targets for antibiotic development. Yet, a detailed

58 understanding of the mechanisms involved in lipoprotein maturation and trafficking is still
59 missing.

60

61 Lipoproteins are synthesized in the cytoplasm as precursors with an N-terminal signal peptide⁹.
62 The last four C-terminal residues of this signal peptide, known as the lipobox, function as a
63 molecular determinant of lipid modification unique to bacteria; only the cysteine at the last
64 position of the lipobox is strictly conserved¹⁰. After secretion of the lipoprotein into the
65 periplasm, the thiol side-chain of the cysteine is first modified with a diacylglyceryl moiety by
66 prolipoprotein diacylglyceryl transferase (Lgt)⁹ (**Extended Data Fig. 1a**, step 1). Then, signal
67 peptidase II (LspA) catalyzes cleavage of the signal peptide N-terminally of the lipidated
68 cysteine before apolipoprotein N-acyltransferase (Lnt) adds a third acyl group to the N-terminal
69 amino group of the cysteine (**Extended Data Fig. 1a**, steps 2-3). Most mature lipoproteins are
70 then transported to the outer membrane by the Lol system. Lol consists of LolCDE, an ABC
71 transporter that extracts lipoproteins from the inner membrane and transfers them to the soluble
72 periplasmic chaperone LolA (**Extended Data Fig. 1a**, steps 4-5)¹¹. LolA escorts lipoproteins
73 across the periplasm, binding their hydrophobic lipid tail, and delivers them to the outer
74 membrane lipoprotein LolB (**Extended Data Fig. 1a**, step 6). LolB finally anchors lipoproteins
75 to the inner leaflet of the outer membrane using a mechanism that remains poorly characterized
76 (**Extended Data Fig. 1a**, step 7).

77

78 In most Gram-negative bacteria, a few lipoproteins remain in the inner membrane^{12,13}. The
79 current view is that inner membrane retention depends on the identity of the two residues
80 located immediately downstream of the N-terminal cysteine on which the lipid moiety is
81 attached¹⁴; this sequence, two amino acids in length, is known as the Lol sorting signal. When
82 lipoproteins have an aspartate at position +2 and an aspartate, glutamate, or glutamine at

83 position +3, they remain in the inner membrane^{15,16}, possibly because strong electrostatic
84 interactions between the +2 aspartate and membrane phospholipids prevent their interaction
85 with LolCDE¹⁷. However, this model is largely based on data obtained in *E. coli* and variations
86 have been described in other bacteria. For instance, in the pathogen *Pseudomonas aeruginosa*,
87 an aspartate is rarely found at position +2 and inner membrane retention appears to be
88 determined by residues +3 and +4^{18,19}. Surprisingly, lipoproteins are well sorted in *P.*
89 *aeruginosa* cells expressing the *E. coli* LolCDE complex²⁰, despite their different Lol sorting
90 signal. This result cannot be explained by the current model of lipoprotein sorting, underscoring
91 that our comprehension of the precise mechanism that governs the triage of lipoproteins remains
92 incomplete.

93

94 Excitingly, more unresolved questions regarding lipoprotein biogenesis have recently been
95 raised. First, it was reported that a LolA-LolB-independent trafficking route to the outer
96 membrane exists in *E. coli*²¹, but the factors involved have remained unknown. Second,
97 although lipoproteins have traditionally been considered to be exposed to the periplasm in *E.*
98 *coli* and many other bacterial models⁹, a series of investigations have started to challenge this
99 view by identifying lipoproteins on the surface of *E. coli*, *Vibrio cholerae*, and *Salmonella*
100 Typhimurium²²⁻²⁶. Overall, the field is beginning to explore a lipoprotein topological landscape
101 that is more complex than previously assumed and raising intriguing questions about the signals
102 that control surface targeting and exposure.

103

104 Here, stimulated by the hypothesis that crucial details of the mechanisms underlying lipoprotein
105 maturation remained to be elucidated, we sought to identify novel molecular determinants
106 controlling lipoprotein biogenesis. First, we systematically analyzed the sequence of the 66
107 lipoproteins with validated localization²⁷ encoded by the *E. coli* K12 genome²⁷ and found that

108 half of the outer membrane lipoproteins display a long and intrinsically disordered linker at
109 their N-terminus. Intrigued by these unstructured segments, we then probed their importance
110 for the biogenesis of RcsF, NlpD, and Pal, three structurally and functionally unrelated outer
111 membrane lipoproteins. Unexpectedly, we found that deleting the linker—while keeping the
112 Lol sorting signal intact—altered the targeting of all three lipoproteins to the outer membrane,
113 with physiological consequences. Focusing on RcsF, we determined that both the length and
114 disordered character of the linker were important. Remarkably, lowering the load of the Lol
115 system by deleting *lpp*, which encodes the most abundant lipoprotein (~1 million copies per
116 cell²⁸), restored normal outer membrane targeting of linker-less RcsF, indicating that the N-
117 terminal linker is required for optimal lipoprotein processing by Lol. Taken together, these
118 observations reveal the unsuspected role played by protein intrinsic disorder in lipoprotein
119 biogenesis.

120 **Results**

121

122 **Half of *E. coli* lipoproteins present long disordered segments at their N-termini**

123 In an attempt to discover novel molecular determinants controlling the biogenesis of
124 lipoproteins, we decided to systematically analyze the sequence of the lipoproteins encoded by
125 the *E. coli* genome (strain MG1655) in search of unidentified structural features. *E. coli* encodes
126 ~80 validated lipoproteins²⁹, of which 58 have been experimentally shown to localize in the
127 outer membrane²⁷. Comparative modeling of existing X-ray, cryogenic electron microscopy
128 (cryo-EM), and nuclear magnetic resonance (NMR) structures revealed that approximately half
129 of these outer membrane lipoproteins display a long segment (>22 residues) that is predicted to
130 be disordered at the N-terminus (**Fig. 1, Extended Data Fig. 2, Extended Data Table 1**). In
131 contrast, only one of the 8 lipoproteins that remain in the inner membrane (DcrB; **Extended**
132 **Data Fig. 2, Extended Data Table 1**) had a long, disordered linker, suggesting that disordered
133 peptides may be important for lipoprotein sorting.

134

135 **Deleting the N-terminal linker of RcsF, NlpD, and Pal perturbs their targeting to the outer** 136 **membrane**

137 Intrigued by the presence of these N-terminal disordered segments in so many outer membrane
138 lipoproteins, we decided to investigate their functional importance. We selected three
139 structurally unrelated lipoproteins whose function could easily be assessed: the stress sensor
140 RcsF (which triggers the Rcs signaling cascade when damage occurs in the envelope³⁰), NlpD
141 (which activates the periplasmic N-acetylmuramyl-L-alanine amidase AmiC, which is involved
142 in peptidoglycan cleavage during cell division^{31,32}), and the peptidoglycan-binding lipoprotein
143 Pal (which is important for outer membrane constriction during cell division³³).

144

145 We began by preparing truncated versions of RcsF, NlpD, and Pal devoid of their N-terminal
146 unstructured linkers (**Extended Data Fig. 1b, Extended Data Fig. 2**; RcsF $_{\Delta 19-47}$, Pal $_{\Delta 26-56}$, and
147 NlpD $_{\Delta 29-64}$). Note that the lipidated cysteine residue (+1) and the Lol sorting signal (the amino
148 acids at positions +2 and +3) were not altered in RcsF $_{\Delta 19-47}$, Pal $_{\Delta 26-56}$, and NlpD $_{\Delta 29-64}$, nor in any
149 of the constructs discussed below (**Extended Data Table 2**). For Pal, although the unstructured
150 linker spans residues 25-68 (**Fig. 1**), we used Pal $_{\Delta 26-56}$ because Pal $_{\Delta 25-68}$ was either degraded or
151 not detected by the antibody (data not shown). We first tested whether the truncated lipoproteins
152 were still correctly extracted from the inner membrane and transported to the outer membrane.
153 The membrane fraction was prepared from cells expressing the three variants independently,
154 and the outer and inner membranes were separated using sucrose density gradients (Methods).
155 Whereas wild-type RcsF, NlpD, and Pal were mostly detected (>90%) in the outer membrane
156 fraction, as expected, ~50% of RcsF $_{\Delta 19-47}$ and ~60% of NlpD $_{\Delta 29-64}$ were retained in the inner
157 membrane (**Fig. 2a, 2b**). The sorting of Pal was also affected, although to a lesser extent: 15%
158 of Pal $_{\Delta 26-56}$ was retained in the inner membrane (**Fig. 2c**). Notably, the expression levels of the
159 three linker-less variants were similar (NlpD $_{\Delta 29-64}$) or lower (RcsF $_{\Delta 19-47}$; Pal $_{\Delta 26-56}$) than those of
160 the wild-type proteins (**Extended Data Fig. 3**), indicating that accumulation in the inner
161 membrane did not result from increased protein abundance.

162
163 We then tested the impact of linker deletion on the function of these three proteins. In cells
164 expressing RcsF $_{\Delta 19-47}$, the Rcs system was constitutively turned on (**Fig. 2d**); when RcsF
165 accumulates in the inner membrane, it becomes available for interaction with IgaA, its
166 downstream Rcs partner in the inner membrane^{30,34}. Likewise, expression of NlpD $_{\Delta 29-64}$ did not
167 rescue the chaining phenotype (**Fig. 2e**)³⁵ exhibited by cells lacking both *nlpD* and *envC*, an
168 activator of the amidases AmiA and AmiB³². Finally, Pal $_{\Delta 26-56}$ partially rescued the sensitivity
169 of the *pal* mutant to SDS-EDTA that results from increased membrane permeability³⁶ (**Fig. 2f**).

170 However, this observation needs to be considered with caution given that Pal_{Δ26-56} seemed to
171 be expressed at lower levels than wild-type Pal (**Extended Data Fig. 3**). Thus, preventing
172 normal targeting of RcsF, NlpD and Pal to the outer membrane had functional consequences.

173

174 **RcsF variants with unstructured artificial linkers of similar lengths are normally targeted**
175 **to the outer membrane**

176 The results above were surprising because they revealed that the normal targeting of RcsF,
177 NlpD, and Pal to the outer membrane does not only require an appropriate Lol sorting signal,
178 as proposed by the current model for lipoprotein sorting⁹, but also the presence of an N-terminal
179 linker. We selected RcsF, whose accumulation in the inner membrane can be easily tracked by
180 monitoring Rcs activity^{30,37}, to investigate the structural features of the linker controlling
181 lipoprotein maturation; keeping as little as 10% of the total pool of RcsF molecules in the inner
182 membrane is sufficient to fully activate Rcs³⁰.

183

184 We first tested whether changing the sequence of the N-terminal segment while preserving its
185 disordered character still yielded normal targeting of the protein to the outer membrane. To that
186 end, we prepared an RcsF variant in which the N-terminal linker was replaced by an artificial,
187 unstructured sequence (**Extended Data Table 2, Extended Data Fig. 2, Extended Data Fig.**
188 **4**) of similar length and consisting mostly of GS repeats (RcsF_{GS}). Substituting the wild-type
189 linker with this artificial sequence was remarkably well tolerated by RcsF: RcsF_{GS} was targeted
190 normally to the outer membrane (**Fig. 3a**) and did not constitutively activate the stress system
191 (**Fig. 3b**). Thus, although RcsF_{GS} has an N-terminus with a completely different primary
192 structure, it behaved like the wild-type protein.

193

194 We then investigated whether the N-terminal linker required a minimal length for proper
195 targeting and function. We therefore constructed two RcsF variants with shorter, unstructured,
196 artificial linkers (RcsF_{GS2} and RcsF_{GS3}, with linkers of 18 and 10 residues, respectively;
197 **Extended Data Table 2, Extended Data Fig. 2, Extended Data Fig. 4**). Importantly, RcsF_{GS2}
198 and, to a greater extent, RcsF_{GS3} did not properly localize to the outer membrane: the shorter
199 the linker, the more RcsF remained in the inner membrane (**Fig. 3a**). Consistent with the amount
200 of RcsF_{GS2} and RcsF_{GS3} retained in the inner membrane, Rcs activation levels were inversely
201 related to linker length (**Fig. 3b**).

202

203 **The disordered character of the linker is required for normal targeting**

204 Taken together, the results above demonstrated that the RcsF linker can be replaced with an
205 artificial sequence lacking secondary structure, provided that it is of appropriate length. Next,
206 we sought to directly probe the importance of having a disordered linker by replacing the RcsF
207 linker with an alpha-helical segment 35 amino acids long from the periplasmic chaperone FkpA
208 (RcsF_{FkpA}; **Extended Data Table 2, Extended Data Fig. 2, Extended Data Fig. 4**).
209 Introducing order at the N-terminus of RcsF dramatically impacted the protein distribution
210 between the two membranes: RcsF_{FkpA} was substantially retained in the inner membrane (**Fig.**
211 **3c**) and constitutively activated Rcs (**Fig. 3d**). As alpha-helical segments are considerably
212 shorter than unstructured sequences containing a similar number of amino acids, we also
213 prepared an RcsF variant (RcsF_{col}) with a longer alpha helix from the helical segment of colicin
214 Ia, which is 73 amino acids in length and also predicted to remain folded in the RcsF_{col} construct
215 (**Extended Data Table 2, Extended Data Fig. 2, Extended Data Fig. 4**). However, doubling
216 the size of the helix had no impact, with RcsF_{col} behaving similarly to RcsF_{FkpA} (**Fig. 3c, 3d**).
217 Together, these data demonstrate that having an N-terminal disordered linker downstream of
218 the Lol sorting signal is required to correctly target RcsF to the outer membrane. The length of

219 the linker is important, but the sequence is not, on the condition that the linker does not fold
220 into a defined secondary structure.

221

222 **The disordered linker is required for optimal processing by Lol**

223 Our finding that N-terminal disordered linkers function as molecular determinants of the
224 targeting of lipoproteins to the outer membrane raised the question of whether these linkers
225 work in a Lol-dependent or Lol-independent manner. To address this mechanistic question, we
226 tested the impact of deleting *lpp* on the targeting of RcsF $_{\Delta 19-47}$. The lipoprotein Lpp, also known
227 as the Braun lipoprotein, covalently tethers the outer membrane to the peptidoglycan and
228 controls the size of the periplasm^{38,39}. Being expressed at ~1 million copies per cell²⁸, Lpp is
229 numerically the most abundant protein in *E. coli*. Thus, by deleting *lpp*, we considerably
230 decreased the load on the Lol system by removing its most abundant substrate. Remarkably,
231 *lpp* deletion fully rescued the targeting of RcsF $_{\Delta 19-47}$ to the outer membrane (**Fig. 4a**), indicating
232 that the linker functions in a Lol-dependent manner and suggesting that accumulation of
233 RcsF $_{\Delta 19-47}$ in the inner membrane results from a decreased ability of the Lol system to process
234 the linker-less RcsF variant. Importantly, similar results were obtained with NlpD $_{\Delta 29-64}$, which
235 was also correctly targeted to the outer membrane in cells lacking Lpp (**Fig. 4a**). Pal $_{\Delta 26-56}$ could
236 not be tested because membrane fractionation failed with *lpp pal* double mutant cells whether
237 or not they expressed Pal $_{\Delta 26-56}$ (data not shown).

238

239 To obtain further insights into the mechanism at play here, we next monitored whether linker
240 deletion impacted the transfer of RcsF from LolA to LolB *in vitro*. LolA with a C-terminal His-
241 tag was expressed in the periplasm of cells expressing wild-type RcsF or RcsF $_{\Delta 19-47}$ and purified
242 to near homogeneity via affinity chromatography (Methods; **Extended Data Fig. 5**). Both RcsF
243 and RcsF $_{\Delta 19-47}$ were detected in immunoblots of the fractions containing purified LolA

244 **(Extended Data Fig. 5)**, indicating that both proteins form a soluble complex with LolA and
245 confirming that they use this chaperone for transport across the periplasm. LolB was expressed
246 as a soluble protein in the cytoplasm and purified by taking advantage of a C-terminal Strep-
247 tag; LolB was then incubated with LolA-RcsF or LolA-RcsF $_{\Delta 19-47}$ and pulled-down using
248 Streptactin beads (Methods). As both RcsF and RcsF $_{\Delta 19-47}$ were detected in the LolB-containing
249 pulled-down fractions (**Fig. 4b**), we conclude that both proteins were transferred from LolA to
250 LolB. Thus, the linker is not required for the transfer of RcsF from LolA to LolB.

251

252 Finally, we focused on the LolCDE ABC transporter in charge of extracting outer membrane
253 lipoproteins and transferring them to LolA. Over-expression (**Extended Data Fig. 6a**) of all
254 components of this complex failed to rescue normal targeting of RcsF $_{\Delta 19-47}$ to the outer
255 membrane (**Extended Data Fig. 6b**). Likewise, over-expressing the enzymes involved in
256 lipoprotein maturation (Lgt, LspA, and Lnt; **Fig. 1**) had no impact on membrane targeting
257 (**Extended Data Fig. 7a, 7b**). Thus, taken together, our results suggest that retention of RcsF $_{\Delta 19-}$
258 $_{47}$ in the inner membrane does not result from the impairment of a specific step, but rather from
259 less efficient processing of the truncated lipoprotein by the entire lipoprotein maturation
260 pathway (see Discussion).

261 **Discussion**

262

263 Lipoproteins are crucial for essential cellular processes such as envelope assembly and
264 virulence. However, despite their functional importance and their potential as targets for new
265 antibacterial therapies, we only have a vague understanding of the molecular factors that control
266 their biogenesis. By discovering the role played by N-terminal disordered linkers in lipoprotein
267 sorting, this study adds an important new layer to our comprehension of lipoprotein biogenesis
268 in Gram-negative bacteria. Critically, it also indicates that the current model of lipoprotein
269 sorting—that sorting between the two membranes is controlled by the 2 or 3 residues that are
270 adjacent to the lipidated cysteine⁴⁰—needs to be revised. Lipoproteins with unstructured linkers
271 at their N-terminus are commonly found in Gram-negative bacteria including many pathogens
272 (see below); further work will be required to determine whether these linkers control lipoprotein
273 targeting in organisms other than *E. coli*, laying the foundation for designing new antibiotics.

274

275 It was previously shown that both *lolA* and *lolB* (but not *lolCDE*) can be deleted under specific
276 conditions²¹, suggesting at least one alternate route for the transport of lipoproteins across the
277 periplasm and their delivery to the outer membrane. During this investigation, we envisaged
278 the possibility that the linker could be required to transport lipoproteins via a yet-to-be-
279 identified pathway independent of LolA/LolB. However, our observations that both RcsF and
280 RcsF_{Δ19-47} were found in complex with LolA (**Extended Data Fig. 5**) and were transferred by
281 LolA to LolB (**Fig. 4b**) does not support this hypothesis. Instead, our data clearly indicate that
282 lipoproteins with N-terminal linkers still depend on the Lol system for extraction from the inner
283 membrane and transport to the outer membrane (**Extended Data Fig. 1a**); they also suggest
284 that N-terminal linkers improve lipoprotein processing by Lol (see below).

285

286 We note that two of the lipoproteins under investigation here, Pal and RcsF, have been reported
287 to be surface-exposed^{30,41,42}. A topology model has been proposed to explain how RcsF reaches
288 the surface: the lipid moiety of RcsF is anchored in the outer leaflet of the outer membrane
289 while the N-terminal linker is exposed on the cell surface before being threaded through the
290 lumen of β -barrel proteins⁴². Thus, in this topology, the linker allows RcsF to cross the outer
291 membrane. It is therefore tempting to speculate that N-terminal disordered linkers may be used
292 by lipoproteins as a structural device to cross the outer membrane and reach the cell surface. It
293 is worth noting that N-terminal linkers are commonly found in lipoproteins expressed by the
294 pathogens *Borrelia burgdorferi* and *Neisseria meningitides*^{24,43,44}; lipoprotein surface exposure
295 is common in these pathogens. In addition, the accumulation of RcsF $_{\Delta 19-47}$ in the inner
296 membrane (**Fig. 2a**) also suggests that Lol may be using N-terminal linkers to recognize
297 lipoproteins destined to the cell surface before their extraction from the inner membrane in
298 order to optimize their targeting to the machinery exporting them to their final destination
299 (BAM in the case of RcsF^{30,42,45}). Investigating whether a dedicated Lol-dependent route exists
300 for surface-exposed lipoproteins will be the subject of future research.

301

302 Our work also delivers crucial insights into the functional importance of disordered segments
303 in proteins in general. Most proteins are thought to present portions that are intrinsically
304 disordered. For instance, it is estimated that 30-50% of eukaryotic proteins contain regions that
305 do not adopt a defined secondary structure *in vitro*⁴⁶. However, demonstrating that these
306 unstructured regions are functionally important *in vivo* is challenging. By showing that an N-
307 terminal disordered segment downstream of the Lol signal is required for the correct sorting of
308 lipoproteins, our work provides direct evidence that evolution has selected intrinsic disorder by
309 function.

310

311 In conclusion, the data reported here establish that the triage of lipoproteins between the inner
312 and outer membranes is not solely controlled by the Lol sorting signal; additional molecular
313 determinants, such as protein intrinsic disorder, are also involved. Our data further highlight
314 the previously unrecognized heterogeneity of the important lipoprotein family and call for a
315 careful evaluation of the maturation pathways of these lipoproteins.

316

317 DATA AVAILABILITY

318 All data generated or analysed during this study are included in this published article and its
319 supplementary information file.

320

321 REFERENCES

- 322 1. Silhavy, T.J., Kahne, D. & Walker, S. The bacterial cell envelope. *Cold Spring Harb*
323 *Perspect Biol* **2**, a000414 (2010).
- 324 2. Weiner, J.H. & Li, L. Proteome of the Escherichia coli envelope and technological
325 challenges in membrane proteome analysis. *Biochim Biophys Acta* **1778**, 1698-713
326 (2008).
- 327 3. Ricci, D.P. & Silhavy, T.J. Outer Membrane Protein Insertion by the β -barrel Assembly
328 Machine. *EcoSal Plus* **8**(2019).
- 329 4. Chimalakonda, G. et al. Lipoprotein LptE is required for the assembly of LptD by the
330 beta-barrel assembly machine in the outer membrane of Escherichia coli. *Proc Natl*
331 *Acad Sci U S A* **108**, 2492-7 (2011).
- 332 5. Sherman, D.J. et al. Lipopolysaccharide is transported to the cell surface by a
333 membrane-to-membrane protein bridge. *Science* **359**, 798-801 (2018).
- 334 6. Malinverni, J.C. et al. YfiO stabilizes the YaeT complex and is essential for outer
335 membrane protein assembly in Escherichia coli. *Mol Microbiol* **61**, 151-64 (2006).
- 336 7. Laloux, G. & Collet, J.F. "Major Tom to ground control: how lipoproteins
337 communicate extra-cytoplasmic stress to the decision center of the cell". *J Bacteriol*
338 (2017).
- 339 8. Kovacs-Simon, A., Titball, R.W. & Michell, S.L. Lipoproteins of bacterial pathogens.
340 *Infect Immun* **79**, 548-61 (2011).
- 341 9. Szewczyk, J. & Collet, J.F. The Journey of Lipoproteins Through the Cell: One
342 Birthplace, Multiple Destinations. *Adv Microb Physiol* **69**, 1-50 (2016).
- 343 10. Babu, M.M. et al. A database of bacterial lipoproteins (DOLOP) with functional
344 assignments to predicted lipoproteins. *J Bacteriol* **188**, 2761-73 (2006).
- 345 11. Narita, S.I. & Tokuda, H. Bacterial lipoproteins; biogenesis, sorting and quality
346 control. *Biochim Biophys Acta Mol Cell Biol Lipids* **1862**, 1414-1423 (2017).
- 347 12. Horler, R.S., Butcher, A., Papangelopoulos, N., Ashton, P.D. & Thomas, G.H.
348 EchoLOCATION: an in silico analysis of the subcellular locations of Escherichia coli

- 349 proteins and comparison with experimentally derived locations. *Bioinformatics* **25**,
350 163-6 (2009).
- 351 13. Tokuda, H. Biogenesis of outer membranes in Gram-negative bacteria. *Biosci*
352 *Biotechnol Biochem* **73**, 465-73 (2009).
- 353 14. Tokuda, H. & Matsuyama, S. Sorting of lipoproteins to the outer membrane in *E. coli*.
354 *Biochim Biophys Acta* **1694**, IN1-9 (2004).
- 355 15. Gennity, J.M. & Inouye, M. The protein sequence responsible for lipoprotein
356 membrane localization in *Escherichia coli* exhibits remarkable specificity. *J Biol Chem*
357 **266**, 16458-64 (1991).
- 358 16. Terada, M., Kuroda, T., Matsuyama, S.I. & Tokuda, H. Lipoprotein sorting signals
359 evaluated as the LolA-dependent release of lipoproteins from the cytoplasmic
360 membrane of *Escherichia coli*. *J Biol Chem* **276**, 47690-4 (2001).
- 361 17. Hara, T., Matsuyama, S. & Tokuda, H. Mechanism underlying the inner membrane
362 retention of *Escherichia coli* lipoproteins caused by Lol avoidance signals. *J Biol Chem*
363 **278**, 40408-14 (2003).
- 364 18. Narita, S. & Tokuda, H. Amino acids at positions 3 and 4 determine the membrane
365 specificity of *Pseudomonas aeruginosa* lipoproteins. *J Biol Chem* **282**, 13372-8 (2007).
- 366 19. Lewenza, S., Mhlanga, M.M. & Pugsley, A.P. Novel inner membrane retention signals
367 in *Pseudomonas aeruginosa* lipoproteins. *J Bacteriol* **190**, 6119-25 (2008).
- 368 20. Lorenz, C., Dougherty, T.J. & Lory, S. Correct Sorting of Lipoproteins into the Inner
369 and Outer Membranes of *Pseudomonas aeruginosa* by the *Escherichia coli* LolCDE
370 Transport System. *mBio* **10**(2019).
- 371 21. Grabowicz, M. & Silhavy, T.J. Redefining the essential trafficking pathway for outer
372 membrane lipoproteins. *Proc Natl Acad Sci U S A* **114**, 4769-4774 (2017).
- 373 22. Konovalova, A. & Silhavy, T.J. Outer membrane lipoprotein biogenesis: Lol is not the
374 end. *Philos Trans R Soc Lond B Biol Sci* **370**(2015).
- 375 23. Wilson, M.M. & Bernstein, H.D. Surface-Exposed Lipoproteins: An Emerging Secretion
376 Phenomenon in Gram-Negative Bacteria. *Trends Microbiol* **24**, 198-208 (2016).
- 377 24. Zuckert, W.R. Secretion of bacterial lipoproteins: through the cytoplasmic
378 membrane, the periplasm and beyond. *Biochim Biophys Acta* **1843**, 1509-16 (2014).
- 379 25. Pride, A.C., Herrera, C.M., Guan, Z., Giles, D.K. & Trent, M.S. The outer surface
380 lipoprotein VolA mediates utilization of exogenous lipids by *Vibrio cholerae*. *MBio* **4**,
381 e00305-13 (2013).
- 382 26. Valguarnera, E., Scott, N.E., Azimzadeh, P. & Feldman, M.F. Surface Exposure and
383 Packing of Lipoproteins into Outer Membrane Vesicles Are Coupled Processes in
384 *Bacteroides*. *mSphere* **3**(2018).
- 385 27. Sueki, A., Stein, F., Savitski, M.M., Selkrig, J. & Typas, A. Systematic Localization of
386 *Escherichia coli* Membrane Proteins. *mSystems* **5**(2020).
- 387 28. Li, G.W., Burkhardt, D., Gross, C. & Weissman, J.S. Quantifying absolute protein
388 synthesis rates reveals principles underlying allocation of cellular resources. *Cell* **157**,
389 624-35 (2014).
- 390 29. Gonnet, P., Rudd, K.E. & Lisacek, F. Fine-tuning the prediction of sequences cleaved
391 by signal peptidase II: a curated set of proven and predicted lipoproteins of
392 *Escherichia coli* K-12. *Proteomics* **4**, 1597-613 (2004).
- 393 30. Cho, S.H. et al. Detecting Envelope Stress by Monitoring beta-Barrel Assembly. *Cell*
394 **159**, 1652-64 (2014).

- 395 31. Heidrich, C. et al. Involvement of N-acetylmuramyl-L-alanine amidases in cell
396 separation and antibiotic-induced autolysis of *Escherichia coli*. *Mol Microbiol* **41**, 167-
397 78 (2001).
- 398 32. Uehara, T., Parzych, K.R., Dinh, T. & Bernhardt, T.G. Daughter cell separation is
399 controlled by cytokinetic ring-activated cell wall hydrolysis. *EMBO J* **29**, 1412-22
400 (2010).
- 401 33. Gerding, M.A., Ogata, Y., Pecora, N.D., Niki, H. & de Boer, P.A. The trans-envelope
402 Tol-Pal complex is part of the cell division machinery and required for proper outer-
403 membrane invagination during cell constriction in *E. coli*. *Mol Microbiol* **63**, 1008-25
404 (2007).
- 405 34. Hussein, N.A., Cho, S.H., Laloux, G., Siam, R. & Collet, J.F. Distinct domains of
406 *Escherichia coli* IgaA connect envelope stress sensing and down-regulation of the Rcs
407 phosphorelay across subcellular compartments. *PLoS Genet* **14**, e1007398 (2018).
- 408 35. Tsang, M.J., Yakhnina, A.A. & Bernhardt, T.G. NlpD links cell wall remodeling and
409 outer membrane invagination during cytokinesis in *Escherichia coli*. *PLoS Genet* **13**,
410 e1006888 (2017).
- 411 36. Shrivastava, R., Jiang, X. & Chng, S.S. Outer membrane lipid homeostasis via
412 retrograde phospholipid transport in *Escherichia coli*. *Mol Microbiol* **106**, 395-408
413 (2017).
- 414 37. Farris, C., Sanowar, S., Bader, M.W., Pfuetzner, R. & Miller, S.I. Antimicrobial peptides
415 activate the Rcs regulon through the outer membrane lipoprotein RcsF. *J Bacteriol*
416 **192**, 4894-903 (2010).
- 417 38. Cohen, E.J., Ferreira, J.L., Ladinsky, M.S., Beeby, M. & Hughes, K.T. Nanoscale-length
418 control of the flagellar driveshaft requires hitting the tethered outer membrane.
419 *Science* **356**, 197-200 (2017).
- 420 39. Asmar, A.T. et al. Communication across the bacterial cell envelope depends on the
421 size of the periplasm. *PLoS Biol* **15**, e2004303 (2017).
- 422 40. Grabowicz, M. Lipoprotein Transport: Greasing the Machines of Outer Membrane
423 Biogenesis: Re-Examining Lipoprotein Transport Mechanisms Among Diverse Gram-
424 Negative Bacteria While Exploring New Discoveries and Questions. *Bioessays* **40**,
425 e1700187 (2018).
- 426 41. Michel, L.V. et al. Dual orientation of the outer membrane lipoprotein Pal in
427 *Escherichia coli*. *Microbiology* **161**, 1251-9 (2015).
- 428 42. Konovalova, A., Perlman, D.H., Cowles, C.E. & Silhavy, T.J. Transmembrane domain of
429 surface-exposed outer membrane lipoprotein RcsF is threaded through the lumen of
430 beta-barrel proteins. *Proc Natl Acad Sci U S A* **111**, E4350-8 (2014).
- 431 43. Brooks, C.L., Arutyunova, E. & Lemieux, M.J. The structure of lactoferrin-binding
432 protein B from *Neisseria meningitidis* suggests roles in iron acquisition and
433 neutralization of host defences. *Acta Crystallogr F Struct Biol Commun* **70**, 1312-7
434 (2014).
- 435 44. Noinaj, N. et al. Structural basis for iron piracy by pathogenic *Neisseria*. *Nature* **483**,
436 53-8 (2012).
- 437 45. Rodriguez-Alonso, R. et al. Structural insight into the formation of lipoprotein-beta-
438 barrel complexes. *Nat Chem Biol* **16**, 1019-1025 (2020).
- 439 46. Bardwell, J.C. & Jakob, U. Conditional disorder in chaperone action. *Trends Biochem*
440 *Sci* **37**, 517-25 (2012).

- 441 47. Majdalani, N., Hernandez, D. & Gottesman, S. Regulation and mode of action of the
442 second small RNA activator of RpoS translation, RprA. *Mol Microbiol* **46**, 813-26
443 (2002).
- 444 48. Baba, T. et al. Construction of Escherichia coli K-12 in-frame, single-gene knockout
445 mutants: the Keio collection. *Mol Syst Biol* **2**, 2006 0008 (2006).
- 446 49. Cherepanov, P.P. & Wackernagel, W. Gene disruption in Escherichia coli: TcR and
447 KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance
448 determinant. *Gene* **158**, 9-14 (1995).
- 449 50. Gil, D. & Bouche, J.P. ColE1-type vectors with fully repressible replication. *Gene* **105**,
450 17-22 (1991).
- 451 51. Yu, D. et al. An efficient recombination system for chromosome engineering in
452 Escherichia coli. *Proc Natl Acad Sci U S A* **97**, 5978-83 (2000).
- 453 52. Sklar, J.G. et al. Lipoprotein SmpA is a component of the YaeT complex that
454 assembles outer membrane proteins in Escherichia coli. *Proc Natl Acad Sci U S A* **104**,
455 6400-5 (2007).
- 456 53. Miller, J.C. *Experiments in Molecular Genetics*, (Cold Spring Harbor Laboratory Press,
457 New York, 1972).
- 458 54. Šali, A. & Blundell, T.L. Comparative Protein Modelling by Satisfaction of Spatial
459 Restraints. *Journal of Molecular Biology* **234**, 779-815 (1993).
- 460 55. Pettersen, E.F. et al. UCSF Chimera - A visualization system for exploratory research
461 and analysis. *Journal of Computational Chemistry* **25**, 1605-1612 (2004).
- 462 56. Guzman, L.M., Belin, D., Carson, M.J. & Beckwith, J. Tight regulation, modulation, and
463 high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol*
464 **177**, 4121-30 (1995).
- 465

466

467 **ACKNOWLEDGMENTS**

468 We thank Asma Boujtat for technical help. We are indebted to the members of the Collet
469 laboratory and to Nassos Typas (EMBL, Heidelberg) for helpful suggestions and discussions
470 and to Tom Silhavy (Princeton) for providing bacterial strains. J.S. was a research fellow of the
471 FRIA and J.F.C. is an Investigator of the FRFS-WELBIO. This work was funded by the
472 WELBIO, by grants from the F.R.S.-FNRS, from the Fédération Wallonie-Bruxelles (ARC
473 17/22-087), from the European Commission via the International Training Network
474 Train2Target (721484), and from the EOS Excellence in Research Program of the FWO and
475 FRS-FNRS (G0G0818N).

476

477 **AUTHOR CONTRIBUTIONS**

478 J.-F.C., J.E.R., J.S., and S.H.C. designed and performed the experiments. J.E.R., J.S., and
479 S.H.C. constructed the strains and cloned the constructs. J.-F.C., J.E.R., J.S., S.H.C., and A.M.
480 analyzed and interpreted the data. B.I.I. performed the structural analysis. J.-F.C., J.E.R., and
481 J.S. wrote the manuscript. All authors discussed the results and commented on the manuscript.

482 **FIGURE LEGENDS**

483

484 **Figure 1. Structural analysis of lipoproteins reveals that half of outer membrane**
485 **lipoproteins display an intrinsically disordered linker at the N-terminus.**

486 Structures were generated via comparative modeling (Methods). X-ray and cryo-EM structures
487 are green, NMR structures are cyan, and structures built via comparative modeling from the
488 closest analog in the same PFAM group are orange. In all cases, the N-terminal linker is
489 magenta. Lipoproteins targeting the outer membrane: Pal, OsmE, NlpE, NlpC, MltB, NlpI,
490 MltC, RcsF, YajI, YcfL, YbaY, RlpA, NlpD, YcaL. The 29 remaining lipoproteins are shown
491 in Extended Data Figure 2.

492

493 **Figure 2. The N-terminal linker displayed by lipoproteins is important for outer**
494 **membrane targeting.**

495 **a, b, c.** The outer membrane (OM) and inner membrane (IM) were separated via centrifugation
496 in a three-step sucrose density gradient (Methods). While (c) RcsF_{WT}, (d) NlpD_{WT}, and (e)
497 Pal_{WT} were found predominantly in the OM, RcsF_{Δ19-47}, NlpD_{Δ29-64}, and Pal_{Δ26-56} were
498 substantially retained in the IM. Data are presented as the ratio of signal intensity in a single
499 fraction to the total intensity in all fractions. All variants were expressed from plasmids
500 (**Extended Data Table 4**). DsbD and Lpp were used as controls for the OM and IM,
501 respectively. **d.** The Rcs system is constitutively active when RcsF's linker is missing. Rcs
502 activity was measured with a beta-galactosidase assay in a strain harboring a transcriptional
503 *rprA::lacZ* fusion (Methods). Results were normalized to expression levels of RcsF variants
504 (mean ± standard deviation; n = 6 biologically independent experiments) **e.** Phase-contrast
505 images of the *envC::kan ΔnlpD* mutant complemented with NlpD_{WT} or NlpD_{Δ29-64}. NlpD_{Δ29-64}
506 only partially rescues the chaining phenotype of the *envC::kan ΔnlpD* double mutant. Scale

507 bar, 5 μm . **f.** Expression of Pal $_{\Delta 26-56}$ does not rescue the sensitivity of the *pal::kan* mutant to
508 SDS-EDTA. Cells were grown in LB medium at 37 °C until OD $_{600}$ = 0.5. Tenfold serial
509 dilutions were made in LB, plated onto LB agar or LB agar supplemented with 0.01% SDS and
510 0.5 mM EDTA, and incubated at 37 °C. Images in **a**, **b**, **c**, **e**, and **f** are representative of
511 biological triplicates. Graphs in **a**, **b**, and **c** were created by spline analysis of curves
512 representing a mean of three independent experiments.

513

514 **Figure 3. The length and the disordered character of the RcsF linker play key roles in**
515 **RcsF targeting to the outer membrane.**

516 **a.** The outer membrane (OM) and inner membrane (IM) were separated via centrifugation in a
517 three-step sucrose density gradient (Methods). DsbD and Lpp were used as controls for the OM
518 and IM, respectively. The longer the linker, the more protein was correctly translocated to the
519 IM. Bar graphs denote mean \pm standard deviation of $n = 3$ biologically independent
520 experiments. Images are representative of experiments and immunoblots performed in
521 biological triplicate. **b.** Rcs activity was measured with a beta-galactosidase assay in a strain
522 harboring a transcriptional *rprA::lacZ* fusion (Methods). Results were normalized to expression
523 levels of RcsF variants (mean \pm standard deviation of $n = 6$ biologically independent
524 experiments). Rcs activity relates to the quantity of RcsF retained in the inner membrane. **c.**
525 RcsF mutants harboring alpha helical linkers (RcsF $_{\text{FkpA}}$ and RcsF $_{\text{col}}$) were subjected to two
526 consecutive centrifugations in sucrose density gradients (Methods). Both mutants were
527 inefficiently translocated from the IM to the OM (mean \pm standard deviation of $n = 3$
528 biologically independent experiments). Images are representative of experiments and
529 immunoblots performed in biological triplicate. **d.** The Rcs system was constitutively active in
530 RcsF $_{\text{FkpA}}$ and RcsF $_{\text{col}}$ strains; activation levels were comparable to those of RcsF $_{\Delta 19-47}$. Rcs
531 activity was measured as in **b**. Results were normalized as in **b**.

532

533 **Figure 4. N-terminal disordered linkers interact with the Lol system to target lipoproteins**
534 **to the outer membrane.**

535 **a.** Deleting Lpp rescues normal targeting of RcsF $_{\Delta 19-47}$ and NlpD $_{\Delta 29-64}$ to the outer membrane.

536 The outer and inner membranes were separated via centrifugation in a sucrose density gradient

537 (Methods). Whereas RcsF $_{\Delta 19-47}$ and NlpD $_{\Delta 29-64}$ accumulate in the inner membrane of cells

538 expressing Lpp, the most abundant Lol substrate, they are normally targeted to the outer

539 membrane in cells lacking Lpp (mean \pm standard deviation of n = 3 biologically independent

540 experiments). **b.** *In vitro* pull-down experiments show that RcsF_{WT} and RcsF $_{\Delta 19-47}$ are

541 transferred from LolA to LolB. LolA-RcsF_{WT} and LolA- RcsF $_{\Delta 19-47}$ complexes were obtained

542 by LolA-His affinity chromatography followed by size exclusion chromatography (Methods).

543 Each complex was incubated with LolB-Strep that was previously purified via Strep-Tactin

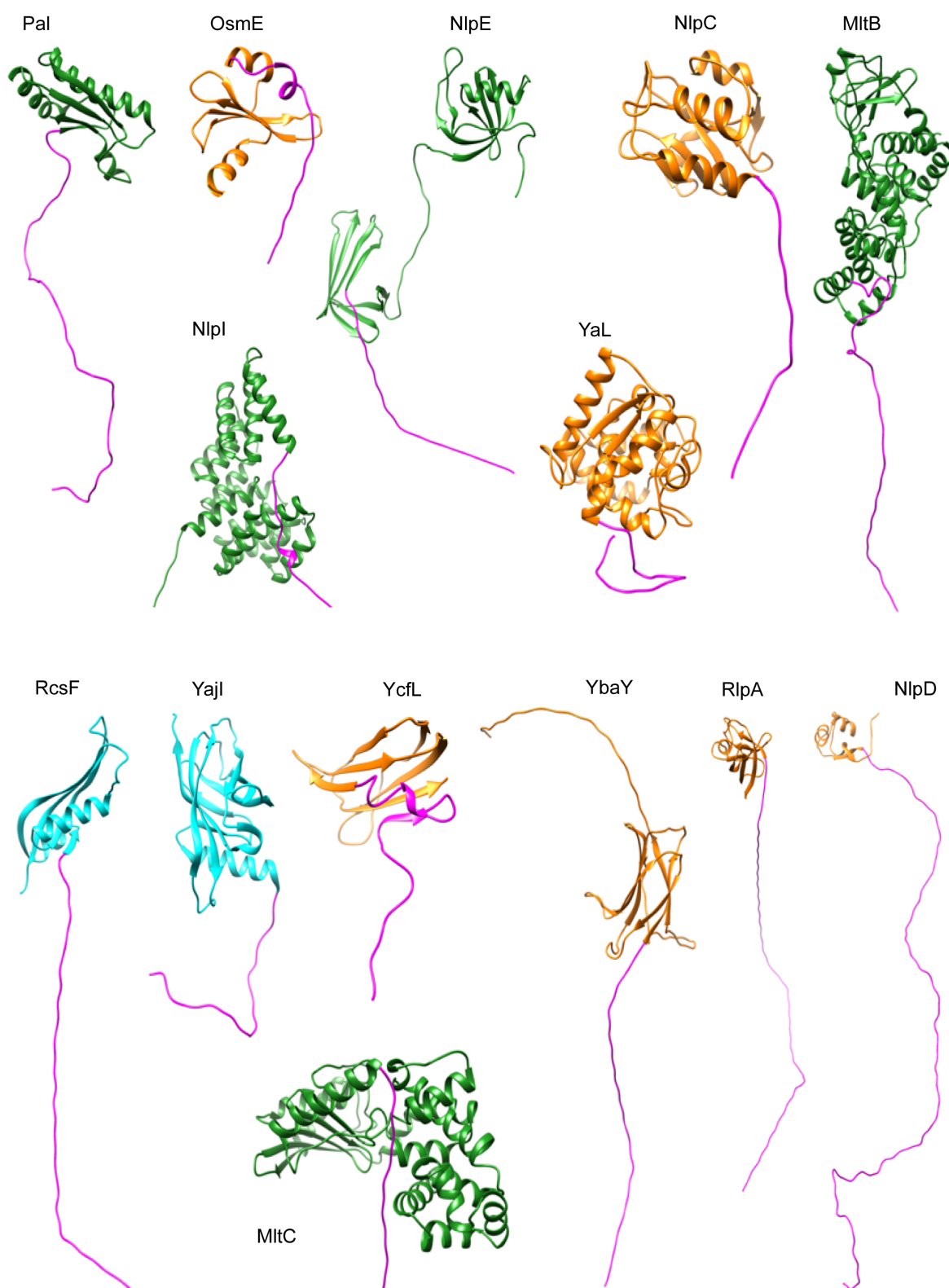
544 affinity chromatography (Methods). Both RcsF variants were eluted in complex with LolB-

545 strep, while LolA was only present in the flow through. I, input; FT, flow through; E, eluate.

546

547 **FIGURES**

548 **Figure 1**

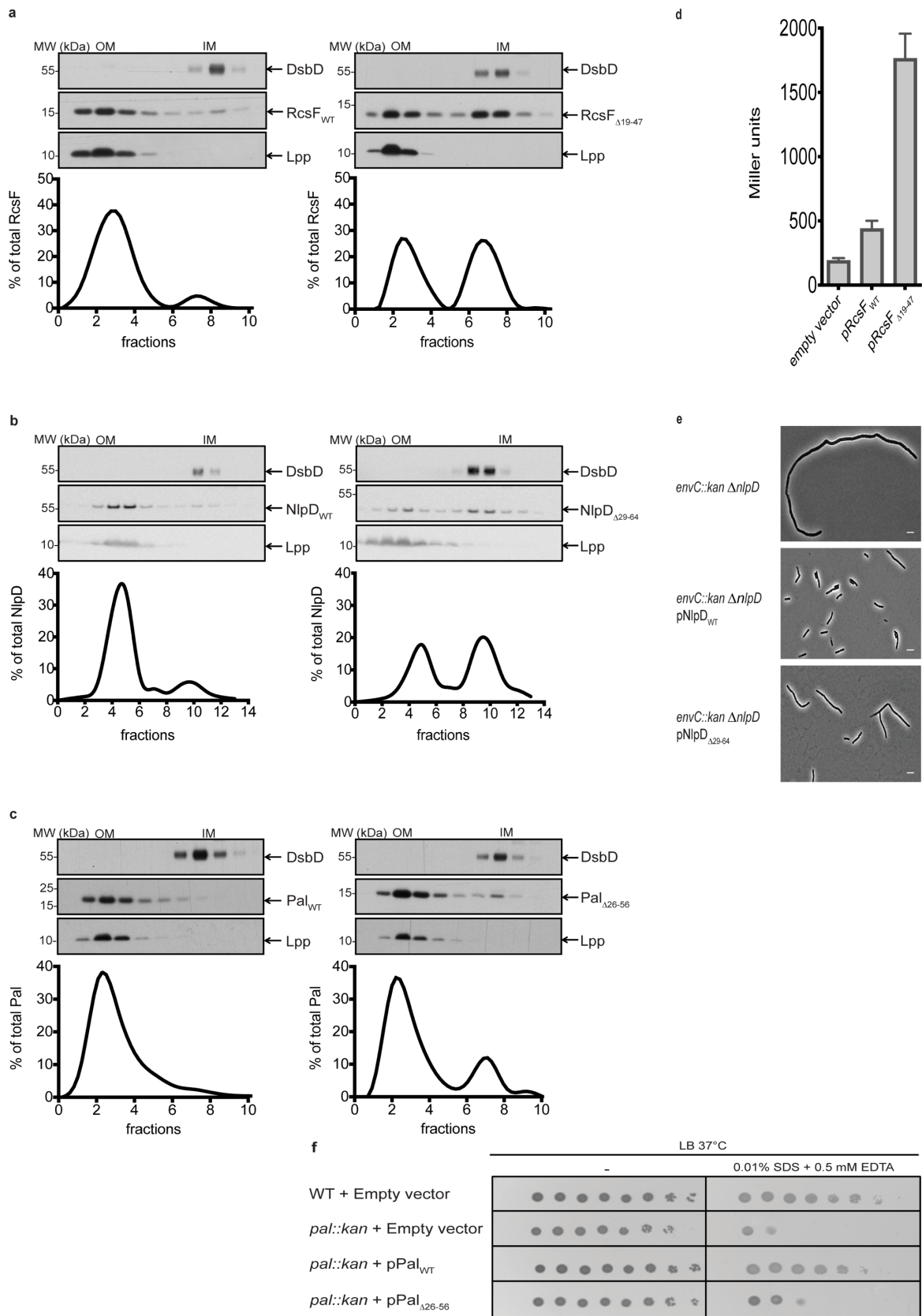


549

550

551

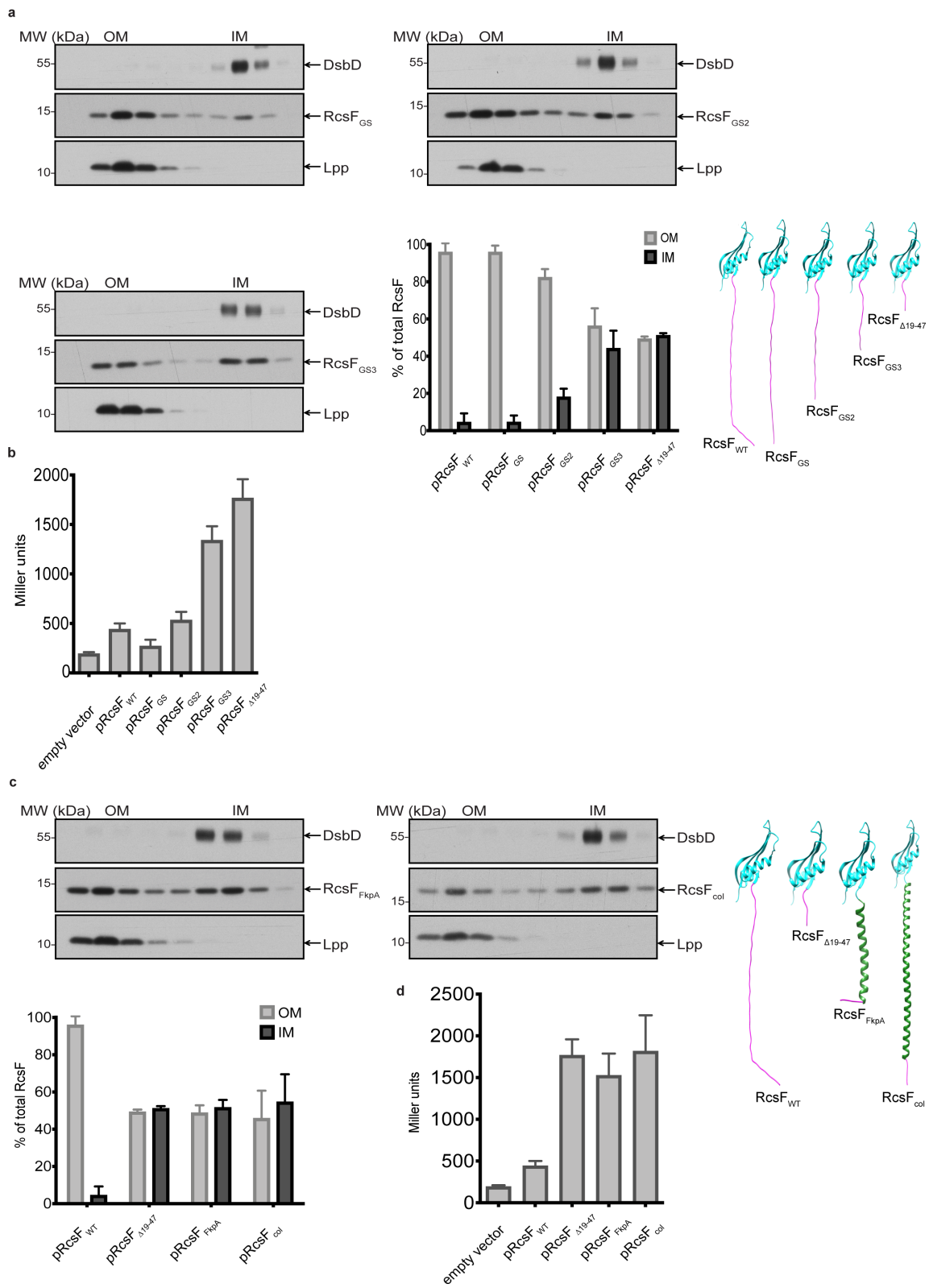
552 **Figure 2**



553

554

555 **Figure 3**

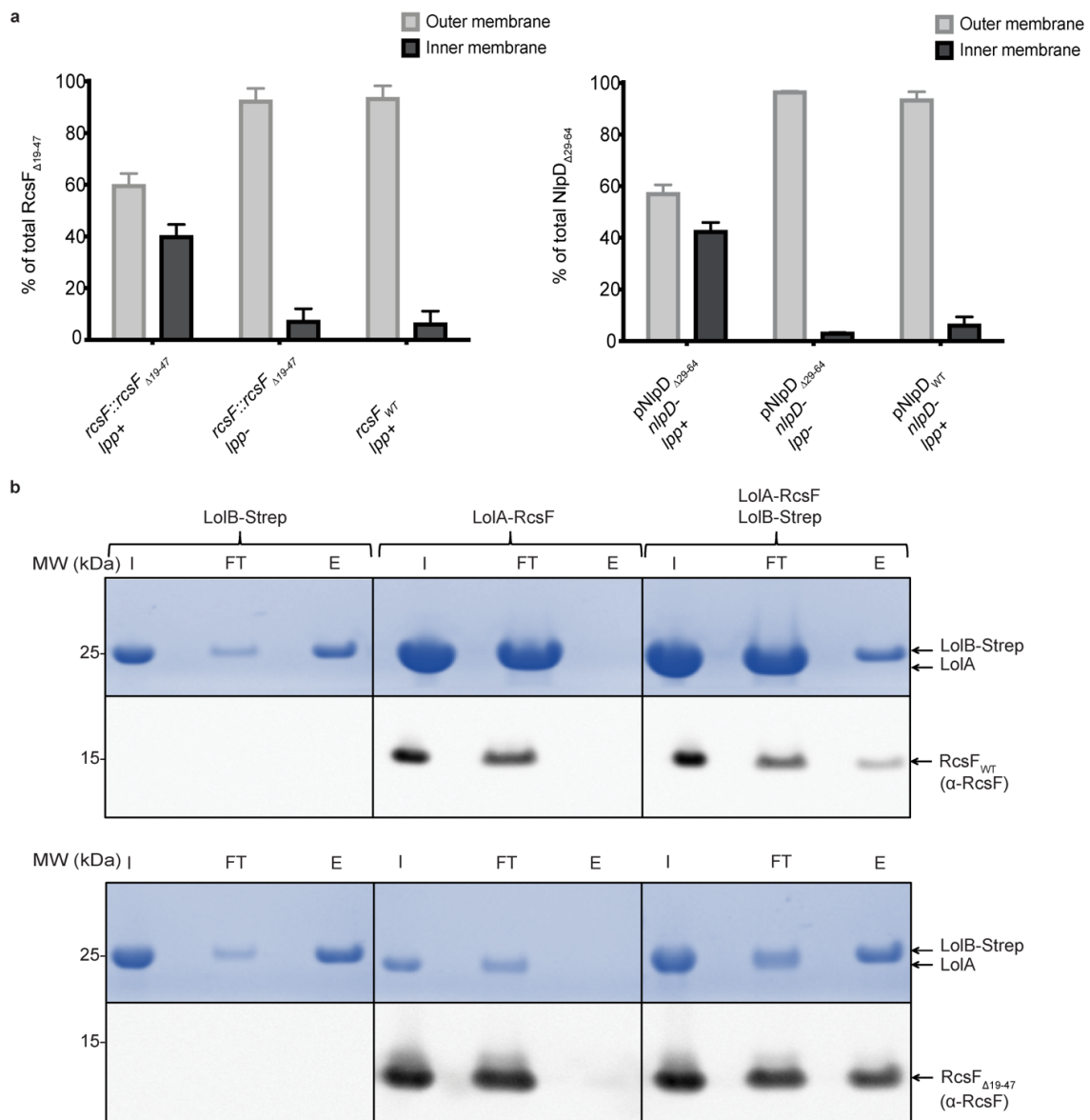


556

557

558

559 **Figure 4**



560

561

562

563

564

565

566

567

568 **METHODS**

569

570 **Bacterial growth conditions**

571 Bacterial strains used in this study are listed in **Extended Data Table 3**. Bacterial cells were
572 cultured in Luria broth (LB) at 37 °C unless stated otherwise. The following antibiotics were
573 added when appropriate: spectinomycin (100 µg/mL), ampicillin (200 µg/mL),
574 chloramphenicol (25 µg/mL), and kanamycin (50 µg/mL). L-arabinose (0.2%) and isopropyl-
575 β-D-thiogalactoside (IPTG) were used for induction when appropriate.

576

577 **Bacterial strains and plasmids**

578 DH300 (a derivative of *Escherichia coli* MG1655 carrying a chromosomal *rprA::lacZ* fusion at
579 the λ attachment site⁴⁷) was used as wildtype throughout the study. All deletion mutants were
580 obtained by transferring the corresponding alleles from the Keio collection⁴⁸ (kan^R) into
581 DH300⁴⁷ via P1 phage transduction. Deletions were verified by PCR and the absence of the
582 protein was verified via immunoblotting (when possible). If necessary, the kanamycin cassette
583 was removed via site-specific recombination mediated by the yeast Flp recombinase with
584 pCP20 vector⁴⁹. All strains expressing the RcsF mutants used for subcellular fractionation
585 lacked *rscB* in order to prevent induction of Rcs.

586

587 The plasmids used in this study are listed in **Extended Data Table 4** and the primers appear in
588 **Extended Data Table 5**. RcsF, Pal, and NlpD were expressed from the low-copy vector
589 pAM238⁵⁰ containing the SC101 origin of replication and the *lac* promoter. To produce pSC202
590 for RcsF expression, *rscF* (including approximately 30 base pairs upstream of the coding
591 sequence) was amplified by PCR from the chromosome of DH300 (primer pair SH_RcsF(PstI)-

592 R and SH_RcsFU-R (kpnI)-F). The amplification product was digested with KpnI and PstI and
593 inserted into pAM238, resulting in pSC202. *nlpD* was amplified using primers JR1 and JR2
594 and *pal* was amplified with primers JS145 and JS146. Amplification products were digested
595 with PstI-XbaI and KpnI-XbaI, respectively, generating pJR8 (for NlpD expression) and pJS20
596 (for Pal expression). To clone *rcsF*_{Δ19-47}, the nucleotides encoding the RcsF signal sequence
597 were amplified using primers SH_RcsFUR(kpnI)_F and SH_RcsFss-Fsg (NcoI)_R, and those
598 encoding the RcsF signaling domain were amplified using primers SH_RcsFss-Fsg (NcoI)_R
599 and SH_RcsF(PstI)_R. In both cases, pSC202 was used as template. Then, overlapping PCR
600 was performed using SH_RcsFUR(kpnI)_F and SH_RcsF(PstI)_R from the two PCR products
601 previously obtained. The final product was digested with KpnI and PstI, and ligated with
602 pAM238 pre-digested with the same enzymes, yielding pSC201. To add a GS linker (Ser-Gly-
603 Ser-Gly-Ser-Gly-Ala-Met) into pSC201, the primers SH_GS linker_F and SH_GS linker_R
604 were mixed, boiled, annealed at room temperature, and ligated with pSC201 pre-digested with
605 NcoI, generating pSC198. pSC199 was generated similarly, but using primers SH_SG linker_F
606 and SH_SG linker_R and plasmid pSC198. pSC200 was generated using primers SH_Da
607 linker_F and SH_SG linker_R and plasmid pSC199. The *pal* allele lacking the linker region
608 (*pal*_{Δ26-56}) was created via overlapping PCR. The pJS20 plasmid served as template for PCR
609 with the M13R/M13F external primers and JS152/JS153 internal primers. The truncated allele
610 was cloned into pAM238 at the same restriction sites as the full-length allele, producing pJS24.
611 The *nlpD* allele lacking the linker regions (*nlpD*_{Δ29-64}) was created via overlapping PCR. *E. coli*
612 chromosomal *nlpD* served as template for the PCR, with JR1/JR2 as external primers and
613 JR7/JR8 as internal primers. The truncated allele was then cloned into pAM238 at the same
614 restriction sites as the full-length allele, producing pJR10.
615

616 *rcsF_{FkpA}* and *rcsF_{col}* were obtained by inserting DNA sequences corresponding to helical linker
617 fragments (FkpA Ser94-Glu125 and colicin IA Ile213-Lys282) into *rcsF_{Δ19-47}* at NcoI and RsrII
618 restriction sites. The *fkpA* gene fragment was amplified from the *E. coli* MC4100 chromosome
619 (JS50/JS51 primers) and the *cia* gene fragment was chemically synthesized as a gene block by
620 Integrated DNA Technologies (IDT). The resulting plasmids were pJS18 and pJS27,
621 respectively. pAM238 does not contain the *lacIq* repressor. Therefore, to enable expression-
622 level regulation by IPTG, strains containing the pAM238 plasmids expressing RcsF variants
623 were co-transformed with pET22b, a high-copy plasmid from a different incompatibility group
624 (pBR223 origin of replication; Novagen) containing the *lacIq* repressor. Chromosomal
625 insertion of RcsF_{Δ19-47} was performed via λ-Red recombineering⁵¹ with pSIM5-Tet plasmid (a
626 gift of D. Hughes). In the first step, the cat-sacB cassette was introduced and later replaced by
627 mutant *rcsF*.

628

629 The chromosomal *lolCDE* operon was amplified via PCR using primers JS277 and JS278
630 (adding a C-terminal His-tag to LolE) and then inserted into pBAD33 using the restriction sites
631 PstI and XbaI, resulting in pJR203. The expression level of LolE-His was verified via
632 immunoblotting.

633

634 The sequence encoding *lolB* without its N-terminal cysteine was first amplified from the
635 chromosome via PCR using primers JR50/PL387 (adding a C-terminal Strep-tag). It was then
636 cloned into pET28a using the restriction sites XbaI and PstI. *lolA* was amplified using
637 chromosomal *lolA* as PCR template for primers JR30/JR31 (JR31 contains the sequence of a
638 His-tag) and then cloned into pBAD18 using KpnI and XbaI, resulting in pJR48.

639

640 The genes encoding Lgt and Lnt were amplified from the chromosome with PCR primers
641 AG389/AG403 and AG393/JR74, respectively. AG403 and JR74 also encode a Myc-tag. PCR
642 products were cloned into pAM238 using KpnI and PstI. Expression levels were verified via
643 immunoblotting (data not shown). *lspA* was amplified with PCR primers JR77/JR78. The PCR
644 product was cloned into pSC213, a modified pAM238 with a ribosome binding site and a C-
645 terminal Flag tag, using NcoI and BamHI. Expression of LspA-Flag was induced by adding 25
646 μ M IPTG. Expression levels were verified with immunoblots (data not shown).

647

648 **Cell fractionation and sucrose density gradients**

649 Cell fractionation was performed as described previously⁵² with some modifications. Four
650 hundred milliliters of cells were grown until the optical density at 600 nm (OD_{600}) of the culture
651 reached 0.7. Cells were harvested via centrifugation at 6,000 x g at 4 °C for 15 min, washed
652 with TE buffer (50 mM Tris-HCl pH 8, 1 mM EDTA), and resuspended in 20 mL of the same
653 buffer. The washing step was skipped with the *Alpp* strains to prevent the loss of outer
654 membrane vesicles. DNase I (1 mg; Roche), 1 mg RNase A (Thermo Scientific), and a half
655 tablet of a protease inhibitor cocktail (cOmplete EDTA-free Protease Inhibitor Cocktail tablets;
656 Roche) were added to cell suspensions, and cells were passed through a French pressure cell at
657 1,500 psi. After adding $MgCl_2$ to a final concentration of 2 mM, the lysate was centrifuged at
658 5,000 x g at 4 °C for 15 min in order to remove cell debris. Then, 16 mL of supernatant were
659 placed on top of a two-step sucrose gradient (2.3 mL of 2.02 M sucrose in 10 mM HEPES pH
660 7.5 and 6.6 mL of 0.77 M sucrose in 10 mM HEPES pH 7.5). The samples were centrifuged at
661 180,000 x g for 3 h at 4 °C in a 55.2 Ti Beckman rotor. After centrifugation, the soluble fraction
662 and the membrane fraction were collected. The membrane fraction was diluted four times with
663 10 mM HEPES pH 7.5. To separate the inner and the outer membranes, 7 mL of the diluted
664 membrane fraction were loaded on top of a second sucrose gradient (10.5 mL of 2.02 M sucrose,

665 12.5 mL of 1.44 M sucrose, 7 mL of 0.77 M sucrose, always in 10 mM HEPES pH 7.5). The
666 samples were then centrifuged at 112,000 x g for 16 h at 10 °C in a SW 28 Beckman rotor.
667 Approximately 30 fractions of 1.5 mL were collected and odd-numbered fractions were
668 subjected to SDS-PAGE, transferred onto a nitrocellulose membrane, and probed with specific
669 antibodies. Graphs were created in GraphPad Prism 9 via spline analysis of the curves
670 representing a mean of three independent experiments.

671

672 **Immunoblotting**

673 Protein samples were separated via 10% or 4-12% SDS-PAGE (Life Technologies) and
674 transferred onto nitrocellulose membranes (GE Healthcare Life Sciences). The membranes
675 were blocked with 5% skim milk in 50 mM Tris-HCl pH 7.6, 0.15 M NaCl, and 0.1% Tween
676 20 (TBS-T). TBS-T was used in all subsequent immunoblotting steps. The primary antibodies
677 were diluted 5,000 to 20,000 times in 1% skim milk in TBS-T and incubated with the membrane
678 for 1 h at room temperature. The anti-RcsF, anti-DsbD, anti-Lpp, anti-NlpD, anti-LolA, and
679 anti-LolB antisera were generated by our lab. Anti-Pal was a gift from R. Lloubès, and anti-His
680 is a peroxidase-conjugated antibody (Qiagen). The membranes were incubated for 1 h at room
681 temperature with horseradish peroxidase-conjugated goat anti-rabbit IgG (Sigma) at a 1:10,000
682 dilution. Labelled proteins were detected via enhanced chemiluminescence (Pierce ECL
683 Western Blotting Substrate, Thermo Scientific) and visualized using X-ray film (Fuji) or a
684 camera (Image Quant LAS 4000 and Vilber Fusion solo S). In order to quantify proteins levels,
685 band intensities were measured using ImageJ version 1.46r (National Institutes of Health).

686

687 **β-galactosidase assay**

688 β-galactosidase activity was measured as described previously⁵³. Graphs representing a mean
689 of six experiments with standard deviation were prepared in GraphPad Prism. Expression-level

690 estimations were performed as follows. Cultures used for β -galactosidase activity (0.5 mL per
691 culture) were precipitated with 10% trichloroacetic acid, washed with ice-cold acetone, and
692 resuspended in 0.2 mL Laemmli SDS sample buffer. Samples (5 μ L) were subjected to SDS-
693 PAGE and immunoblotted with anti-RcsF antibody.

694

695 **SDS-EDTA sensitivity assay**

696 Cells were grown in LB at 37 °C until they reached an OD₆₀₀ of 0.7. Tenfold serial dilutions
697 were made in LB and plated on LB agar supplemented with spectinomycin (100 μ g/mL) when
698 necessary. Plates were incubated at 37 °C. To evaluate the sensitivity of the *pal* mutant, plates
699 were supplemented with 0.01% SDS and 0.5 mM EDTA.

700

701 **Microscopy image acquisition**

702 Cells were grown in LB at 37 °C until OD₆₀₀ = 0.5. Cells growing in exponential phase were
703 spotted onto a 1% agarose phosphate-buffered saline pad for imaging. Cells were imaged on a
704 Nikon Eclipse Ti2-E inverted fluorescence microscope with a CFI Plan Apochromat DM
705 Lambda 100X Oil, N.A. 1.45, W.D. 0.13 mm objective. Images were collected on a Prime 95B
706 25 mm camera (Photometrics). We used a Cy5-4050C (32 mm) filter cube (Nikon). Image
707 acquisition was performed with NIS-Element Advance Research version 4.5.

708

709 **Protein purification**

710 JR90 cells were grown in LB supplemented with kanamycin (50 μ g/mL) at 37 °C. When the
711 culture OD₆₀₀ = 0.5, the expression of cytoplasmic LolB-Strep was induced with 1 mM IPTG.
712 Cells (1 L) were pelleted when they reached OD₆₀₀ = 3 and resuspended in 25 mL of buffer A
713 (200 mM NaCl and 50 mM NaPi, pH 8) containing one tablet of cComplete EDTA-free Protease
714 Inhibitor Cocktail (Roche). Cells were lysed via two passages through a French pressure cell at

715 1,500 psi. The lysate was centrifuged at 30,000 x *g* for 40 min at 4 °C in a JA 20 rotor and the
716 supernatant was mixed with Strep-Tactin resin (IBA Lifesciences) previously equilibrated with
717 buffer A. After washing the resin with 10 column volumes of buffer A, LolB-Strep was eluted
718 with 5 column volumes of buffer A supplemented with 5 mM desthiobiotin. LolB-Strep was
719 finally desalted using a PD10 column (GE Healthcare).

720

721 Soluble LolA-RcsF_{WT} and LolA-RcsF_{Δ19-47} complexes were purified via affinity
722 chromatography as follows. Cells co-expressing LolA either with wild-type RcsF (JR47) or
723 RcsF_{Δ19-47} (JR44) were grown in LB at 37 °C supplemented with 200 μg/mL ampicillin until
724 OD₆₀₀ = 0.5. Protein expression was then induced with 0.2% arabinose. Cells (1 L) were
725 pelleted at OD₆₀₀ = 3 and resuspended in 25 mL of buffer A containing one tablet of protease
726 inhibitor cocktail. Cells were lysed via two passages through a French pressure cell at 1,500
727 psi. The lysate was centrifuged at 45,000 x *g* for 30 min at 4 °C using a 55.2 Ti Beckman rotor.
728 To obtain the soluble fraction, the supernatant was centrifuged at 180,000 x *g* for 1 h at 4 °C
729 using the same rotor. The supernatant was added to a His Trap HP column (Merck) previously
730 equilibrated with buffer A. The column was washed with 10 column volumes of buffer A
731 supplemented with 20 mM imidazole and LolA-His was eluted using a gradient of imidazole
732 (from 20 mM to 300 mM). The fractions obtained were analyzed via SDS-PAGE; LolA was
733 detected around 25 kDa (data not shown). RcsF variants were detected via immunoblotting with
734 an anti-RcsF antibody. Fractions containing LolA-RcsF variants were pooled, concentrated to
735 1 mL using a Vivaspin 4 Turbo concentrator (Cut-off 5 kDa; Sartorius), and purified via size-
736 exclusion chromatography with a Superdex S75-10/300 column (GE Healthcare).

737

738 **Pull down and transfer of RcsF variants from LolA to LolB**

739 LolB-Strep was incubated at 30 °C for 20 min under agitation with LolA-RcsF_{WT} or with LolA-
740 RcsF_{Δ19-47} (LolA-RcsF_{WT} and LolA-RcsF_{Δ19-47} complexes were purified as described above).
741 The mixture was added to magnetic Strep beads (MagStrep type 3 beads, IBA Life science)
742 previously equilibrated with buffer A and incubated for 30 min at 4 °C on a roller. After washing
743 the beads with the same buffer, LolB-Strep was eluted with buffer A supplemented with 50 mM
744 biotin. Samples were analyzed via SDS-PAGE and LolA and LolB were detected with
745 Coomassie Brilliant Blue (Bio-Rad). RcsF was detected via immunoblotting with an anti-RcsF
746 antibody.

747

748 **Structural analysis of lipoproteins**

749 When X-ray, cryo-EM, or NMR structures were available, the missing residues were completed
750 through comparative modeling using MODELLER version 9.22⁵⁴. If no structure of the
751 lipoprotein was available, then the most pertinent analogous structure from proteins belonging
752 to the same PFAM group was used as template for comparative modeling. The linker was
753 defined as the unstructured fragment from the N-terminal Cys of the mature form until the first
754 residue with well-defined secondary structure (α -helix or β -strand) belonging to a globular
755 domain. Short, intermediate, and long linkers had lengths of <12, 12-22, and >22 residues,
756 respectively. Images were generated using UCSF Chimera version 1.13.1⁵⁵.

757

758 **LEGENDS FOR FIGURES IN THE EXTENDED DATA**

759

760 **Extended Data Figure 1. Lipoprotein maturation and sorting in the *E. coli* cell envelope.**

761 **a.** After processing by Lgt (step 1), LspA (step 2), and Lnt (step 3), a new lipoprotein either
762 remains in the inner membrane or is extracted by the LolCDE complex (step 4), depending on
763 the residues at position +2 and +3. LolCDE transfers the lipoprotein to the periplasmic
764 chaperone LolA (step 5), which delivers the lipoprotein to LolB (step 6). LolB, a lipoprotein
765 itself, inserts the lipoprotein in the outer membrane using a poorly understood mechanism (step
766 7). **b.** Schematic of lipoprotein structural domains. The N-terminal signal sequence targets the
767 lipoprotein to the cell envelope; the last four amino acid residues of the signal sequence form
768 the lipobox. The last residue of the lipobox is the invariant cysteine that undergoes lipidation.
769 This cysteine, which is the first residue of the mature lipoprotein, is directly followed by the
770 sorting signal, a sequence of 2 or 3 amino acids that controls the sorting of mature lipoproteins
771 between the inner and outer membranes. The C-terminal portion of a mature lipoprotein is a
772 globular domain. An intrinsically disordered linker separates the sorting signal from the
773 globular domain in about half of *E. coli* lipoproteins (**Fig. 1; Extended Data Fig. 2; Extended**
774 **Data Table 1**). The lengths of the deleted disordered linkers of the unrelated lipoproteins RcsF,
775 Pal, and NlpD are indicated. LP, lipoprotein.

776

777 **Extended Data Figure 2. Structural analysis of lipoproteins reveals that half of outer**
778 **membrane lipoproteins display an intrinsically disordered linker at the N-terminus.**

779 Structures were generated via comparative modeling. X-ray and cryo-EM structures are green,
780 NMR structures are cyan, and structures built via comparative modeling from the closest analog
781 in the same PFAM group are orange. In all cases, the N-terminal linker is magenta. Lipoproteins
782 targeting the outer membrane: AmiD, BamB, BamC, HslJ, MltA, LoiP, LpoB, Blc, BamE,
783 CsgG, EmtA, GfcE, BamD, LpoA, LolB, LptE, MlaA, MliC, YddW, YedD, YghG, YfeY,

784 YbjP, YiaD, YbhC, PqiC, YgeR, YfiB, YraP. Lipoproteins targeting the IM: DcrB, MetQ,
785 NlpA, YcjN, YehR, ApbE. Synthetic constructs: RcsF_{GS}, RcsF_{GS2}, RcsF_{GS3}, RcsF_{Δ19-47},
786 RcsF_{FkpA}, RcsF_{col}, NlpD_{Δ29-64}, Pal_{Δ26-56}.

787

788 **Extended Data Figure 3. Expression levels of RcsF_{Δ19-47}, Pal_{Δ26-56}, and NlpD_{Δ29-64}.**

789 Cells were grown at 37 °C in LB until OD₆₀₀ = 0.5 and precipitated with trichloroacetic acid
790 (Methods). Immunoblots were performed with α-RcsF, α-NlpD, and α-Pal antibodies
791 (Methods). All images are representative of three independent experiments.

792

793 **Extended Data Figure 4. Schematic of RcsF variants used in this study and their**
794 **distributions in the outer membrane (OM) and inner membrane (IM).**

795 RcsF_{GS}, RcsF_{GS2}, and RcsF_{GS3} have linkers that are disordered and mostly consist of GS repeats.
796 The linker of RcsF_{GS} is the same length as the linker of RcsF_{WT}. RcsF_{GS2} and RcsF_{GS3} are shorter
797 than RcsF_{WT}. Regions of RcsF_{FkpA} and RcsF_{col} fold into alpha helices borrowed from the
798 sequences of FkpA and colicin Ia, respectively.

799

800 **Extended Data Figure 5. Complexes between LolA and RcsF_{WT} or RcsF_{Δ19-47} can be**
801 **purified.**

802 Both RcsF_{WT} (a) and RcsF_{Δ19-47} (b) were eluted in complex with LolA-His via affinity
803 chromatography followed by size exclusion chromatography. Gel filtration was performed with
804 a Superdex S75-10/300 column. Samples were analyzed via SDS-PAGE and proteins,
805 including LolA-His, were stained with Coomassie Brilliant Blue (Methods). RcsF variants were
806 detected by immunoblotting fractions with α-RcsF antibodies. Images are representative of
807 three independent experiments.

808

809 **Extended Data Figure 6. Overexpression of Lol CDE does not restore targeting of RcsF_{Δ19-}**

810 47.

811 **a.** Expression level of LolCDE-His. Cells were grown in LB plus 0.2% arabinose at 37 °C until
812 OD₆₀₀ = 0.7 (Methods). Membrane and soluble fractions were separated with a sucrose density
813 gradient (Methods). LolE-His was detected in the membrane fraction by immunoblotting with
814 α-His (Methods). Images are representative of three independent experiments. **b.** The outer
815 membrane (OM) and inner membrane (IM) were separated with a sucrose density gradient.
816 Expression of LolCDE did not rescue OM targeting of RcsF_{Δ19-47}. Images are representative of
817 experiments performed in biological triplicate.

818

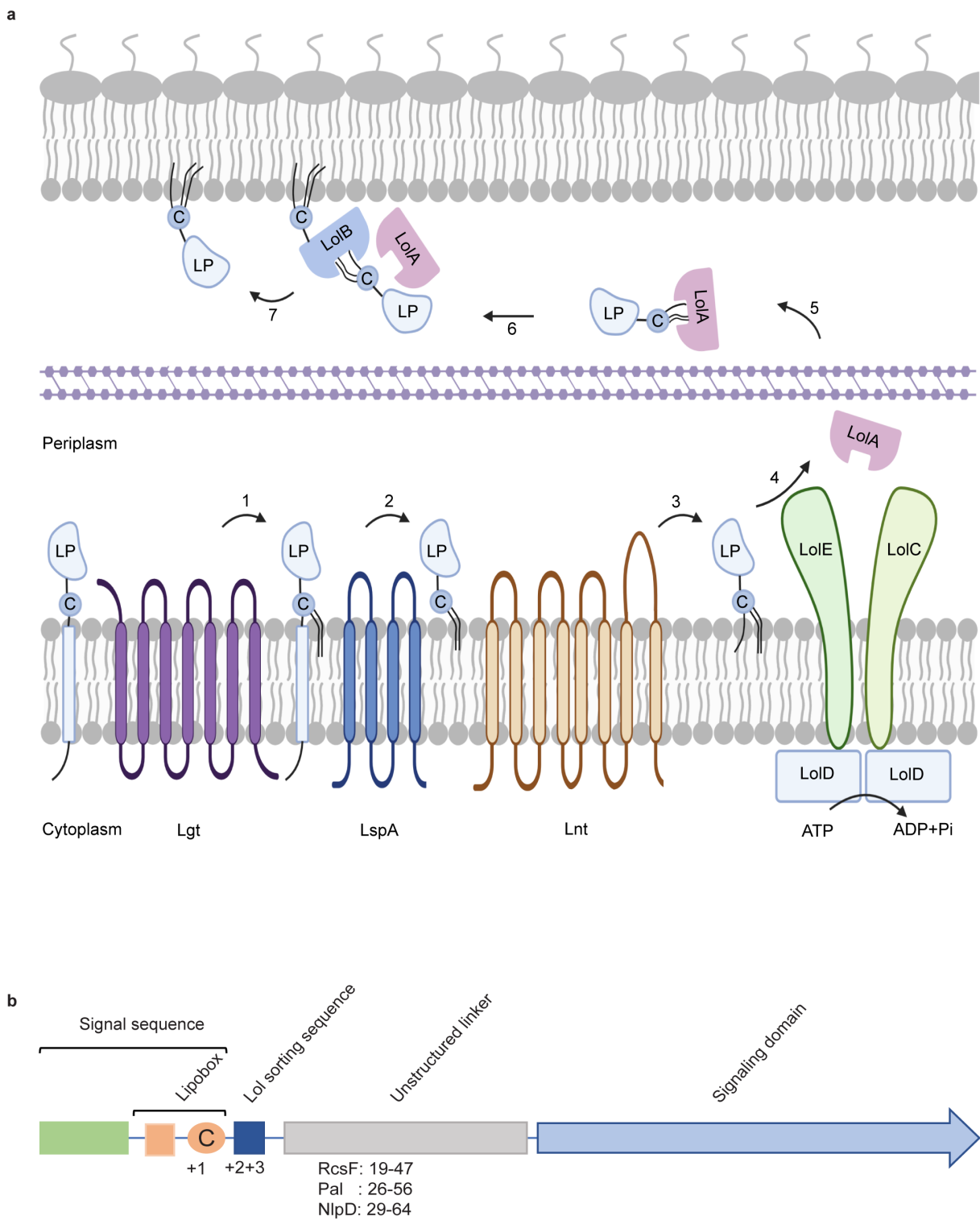
819 **Extended Data Figure 7. Overexpressing Lgt, LspA, and Lnt does not rescue the targeting**
820 **of RcsF_{Δ19-47} to the outer membrane.**

821 **a.** Expression levels of Lgt, LspA, and Lnt. Cells were grown in LB (plus 25 μM IPTG for cells
822 expressing LspA) at 37 °C until OD₆₀₀ = 0.7 (Methods). Outer membrane (OM) and inner
823 membrane (IM) were separated with a sucrose density gradient (Methods). Lgt-Myc and Lnt-
824 Myc were detected in the IM via immunoblotting with α-Myc. LspA-Flag was detected in the
825 IM with α-Flag. **b.** Cells overexpressing Lgt, LspA, or Lnt were exposed to a sucrose density
826 gradient (Methods). RcsF_{Δ19-47} was retained in the IM in all conditions. Images are
827 representative of three independent experiments.

828

829 **EXTENDED DATA FIGURES**

830 **Extended Data Figure 1**



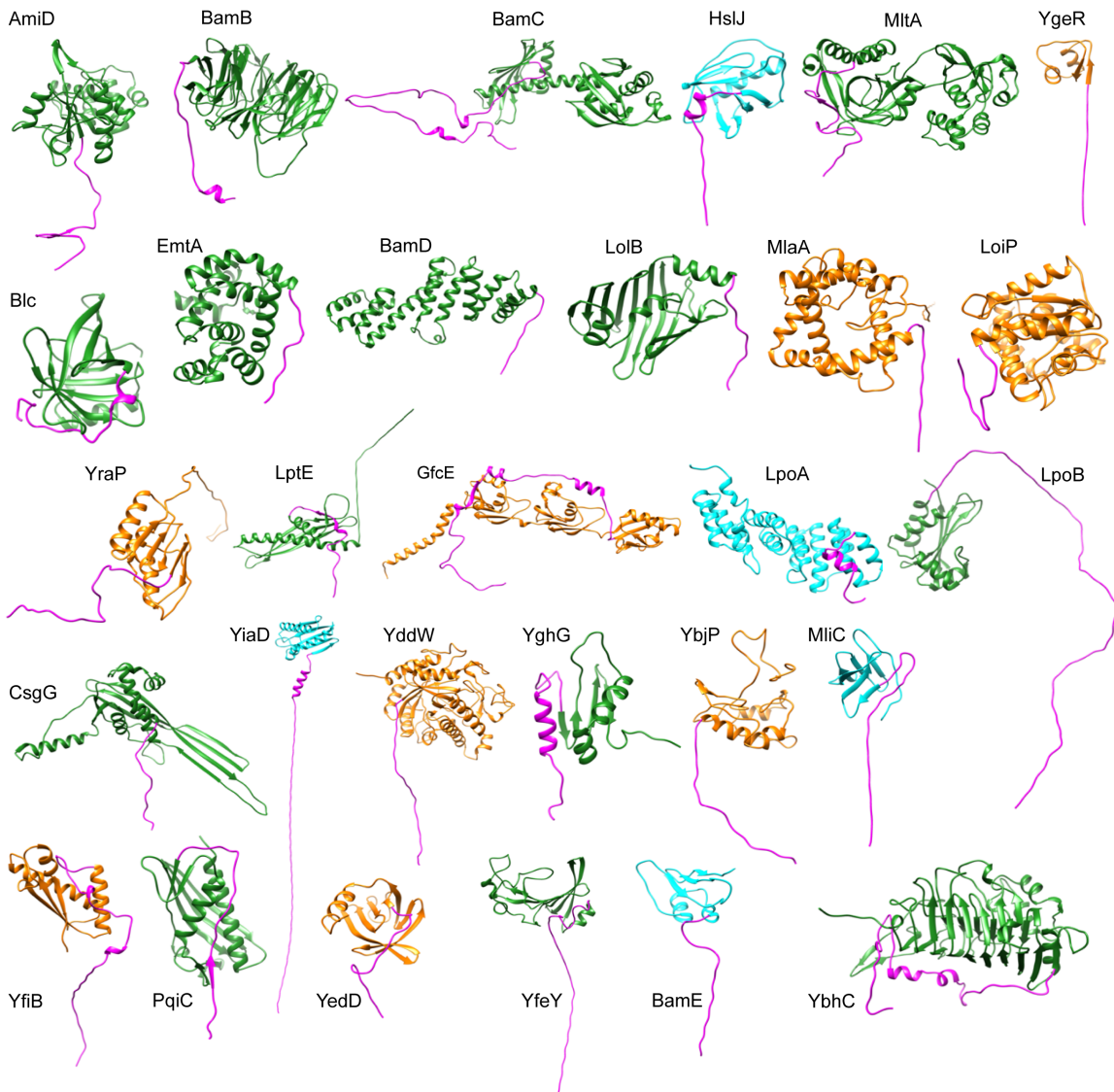
831

832

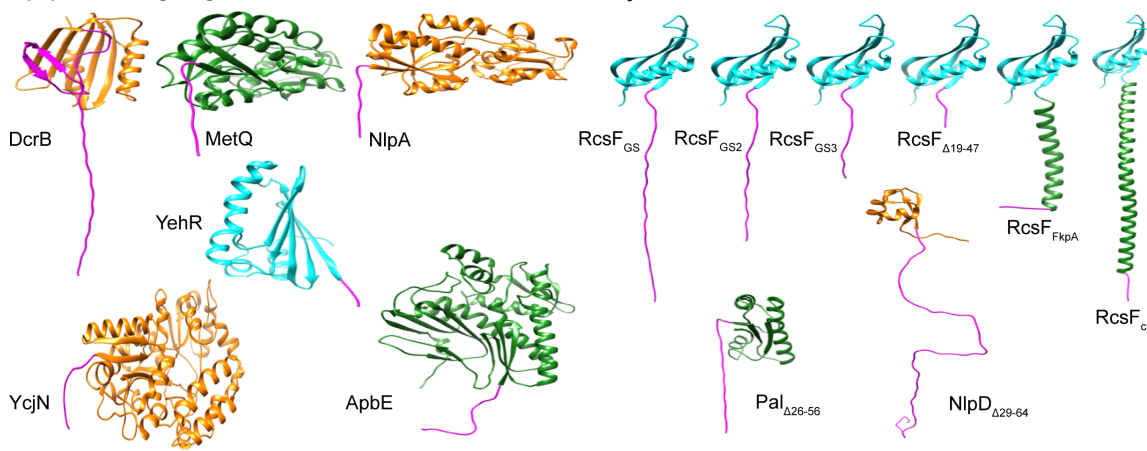
833

834 **Extended Data Figure 2**

Lipoproteins targeting the outer membrane



Lipoproteins targeting the inner membrane

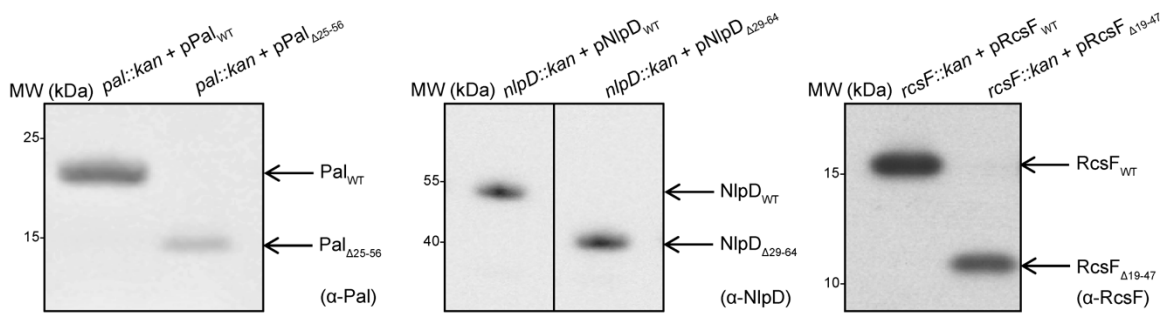


835

836

837 **Extended Data Figure 3**

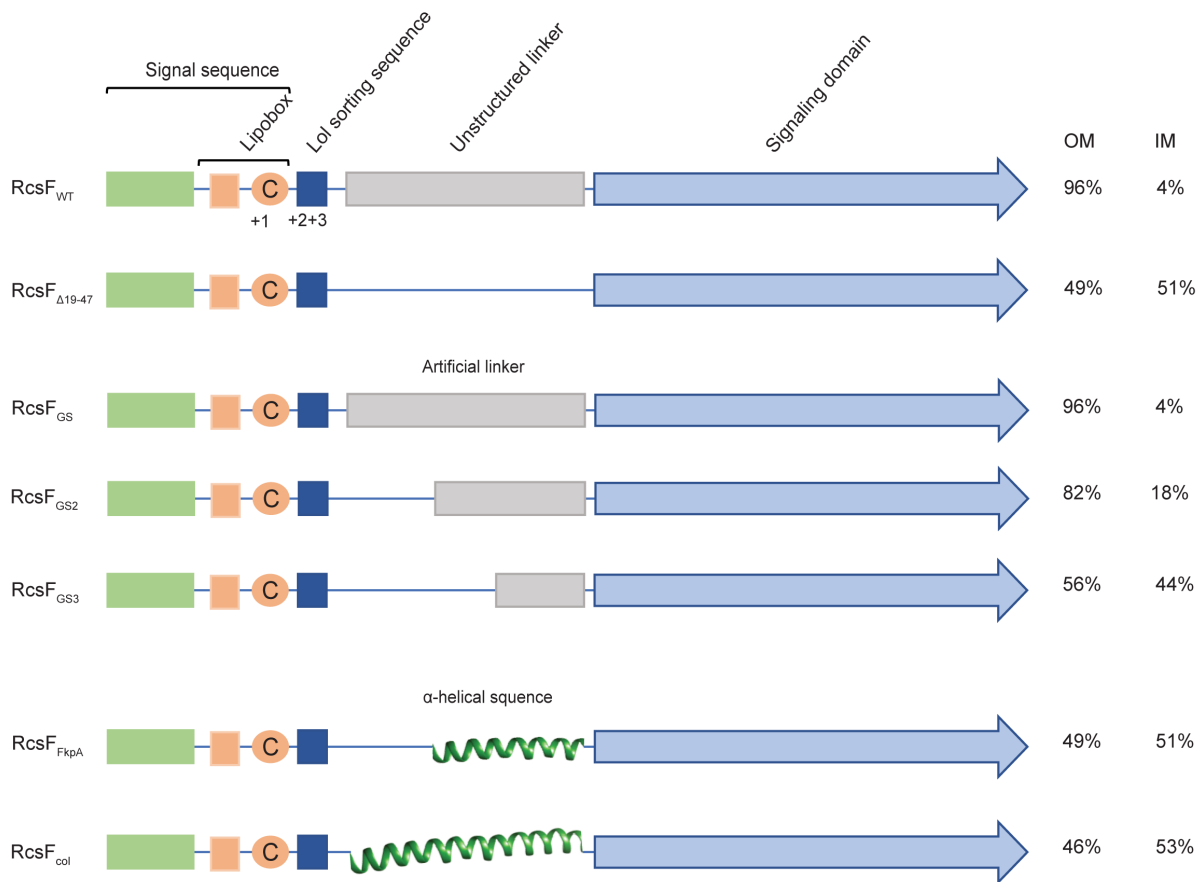
838



839

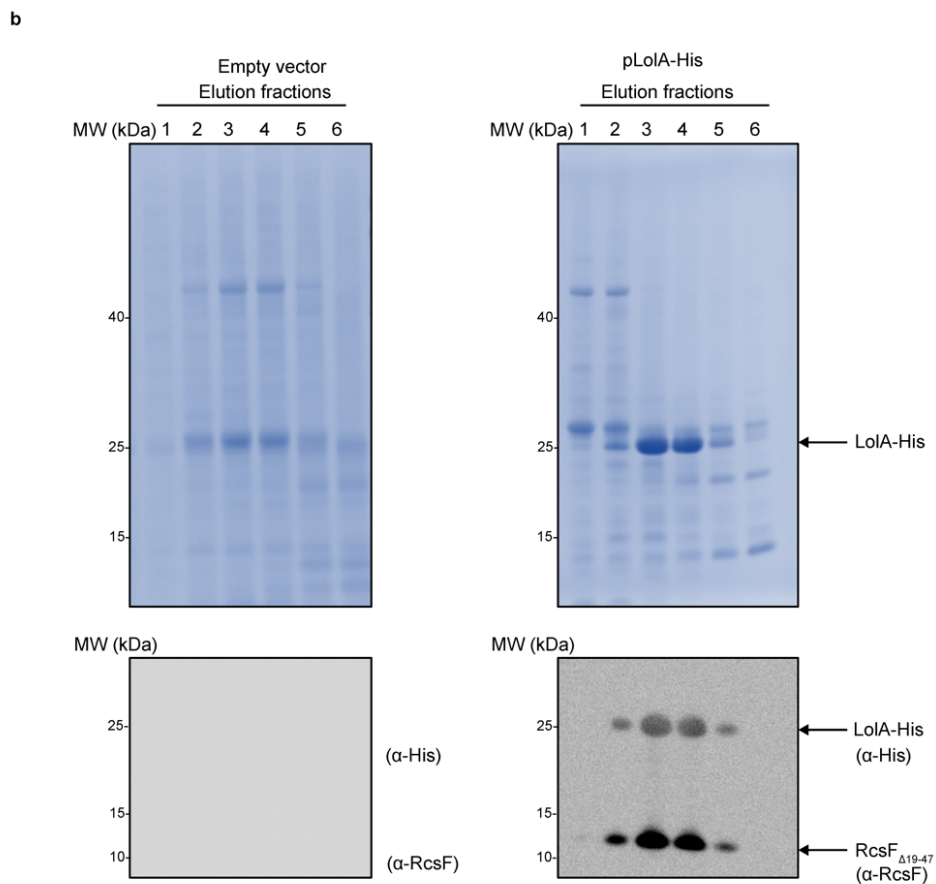
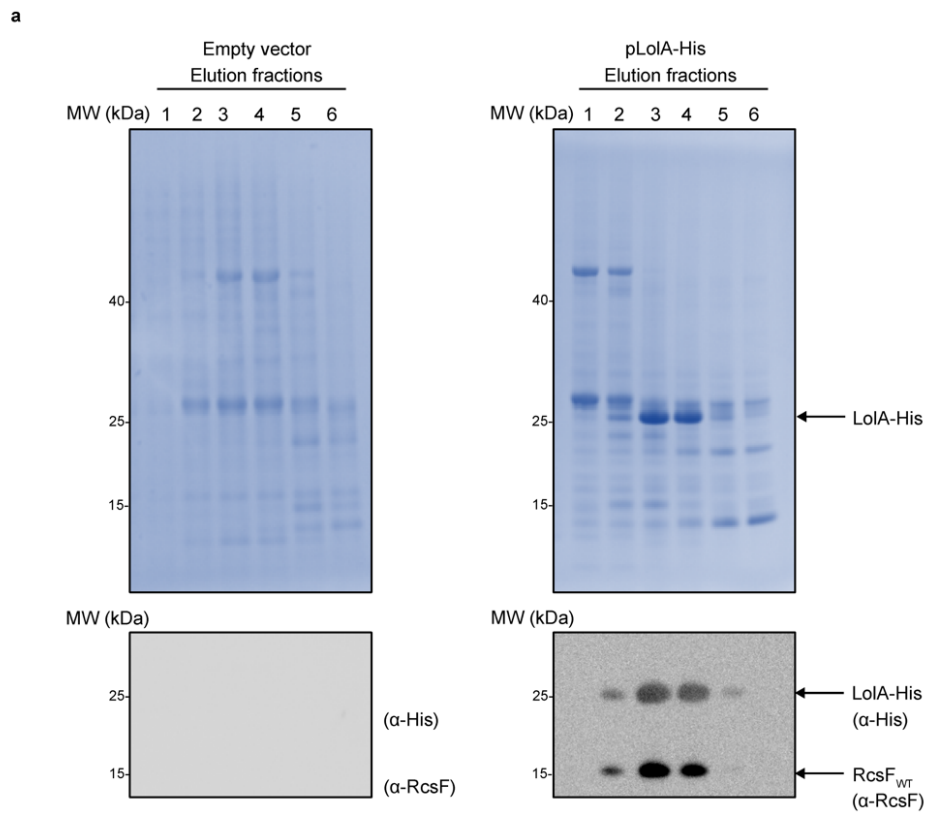
840 **Extended Data Figure 4**

841



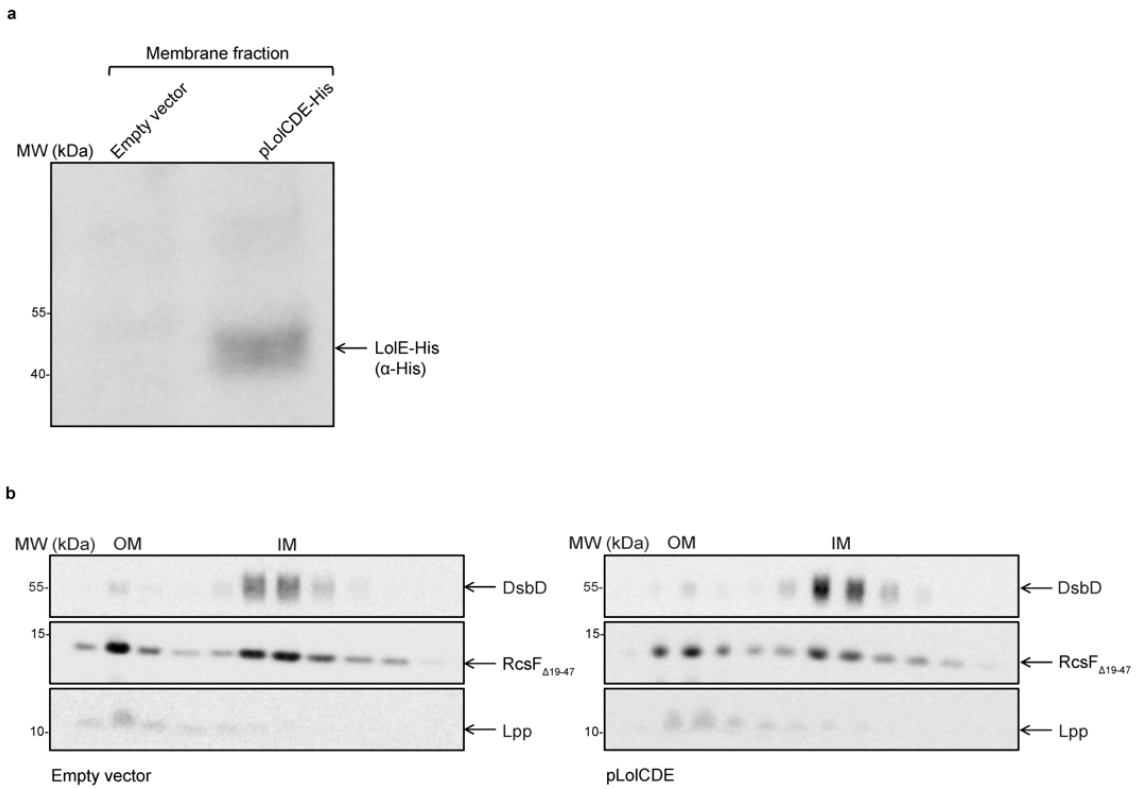
842

843 **Extended Data Figure 5**



844

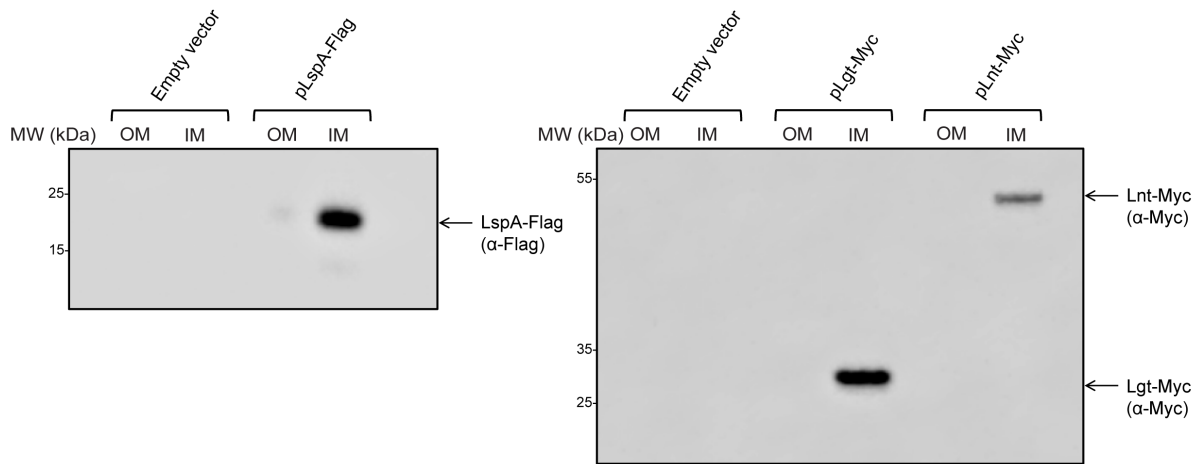
845 **Extended Data Figure 6**



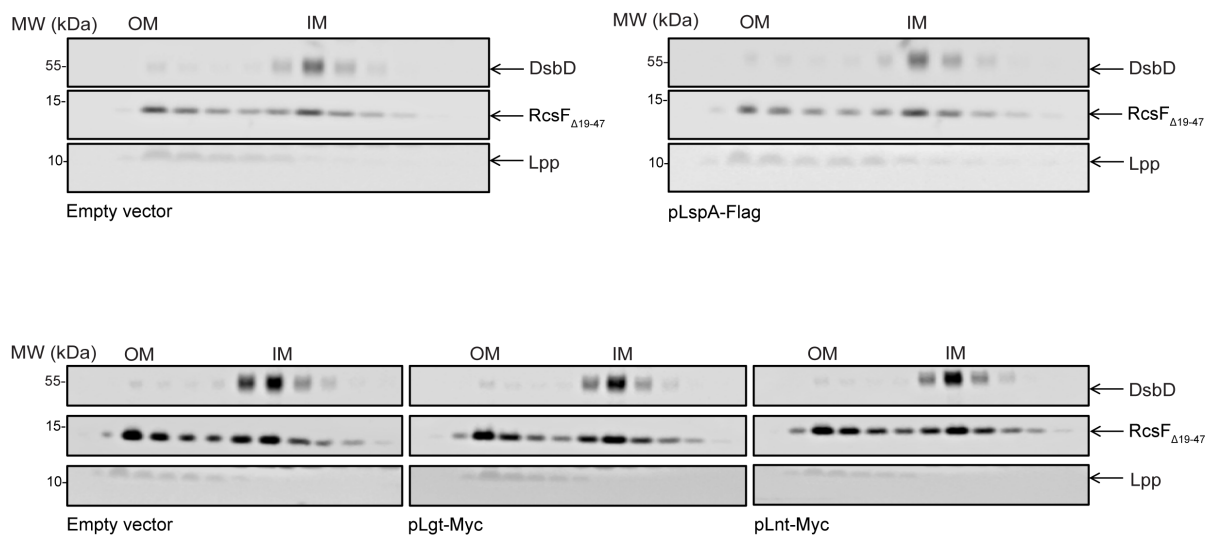
846

847 **Extended Data Figure 7**

a



b



848

849 **EXTENDED DATA TABLES**

850

851 **Extended Data Table 1: List of the verified lipoproteins of *E. coli* used for the structural**
852 **analysis in this study.**

853 **Attached Excel sheet**

854

855 **Extended Data Table 2: RcsF mutants used in this study and the amino acid sequences of**
856 **their corresponding N-terminal linkers. The acylated cysteine is the first residue listed.**

RcsF linkers	Amino acid sequence
RcsF _{WT}	CSMLSRSPVEPVQSTAPQPKAEPKPKAPRATPV
RcsF _{Δ19-47}	CSMGPV
RcsF _{GS}	CSMSLFDAPAMSGSGSGAMSGSGSGAMPV
RcsF _{GS2}	CSMSGSGSGAMSGSGSGAMPV
RcsF _{GS3}	CSMSGSGSGAMPV
RcsF _{FkpA}	CSMGSDQEIEQTLQAFEARVKSSAQAKMEKDAADNEPV
RcsF _{col}	CSMGILDTRLSELEKNGGAALAVLDAQQARLLGQQTRNDRAISEARNKL SSVTESLNTARNALTRAEQQLTQQKPV

857

858

859

860 **Extended Data Table 3: *E. coli* strains used in this study.**

Strains	Genotype and description	Source
DH300	<i>rprA-lacZ</i> MG1655 (<i>argF-lac</i>) U169	⁴⁷
Keio collection single mutants	<i>rcsF::kan, rcsB::kan, pal::kan, nlpD::kan, envC::kan</i>	⁴⁸
XL1-Blue	<i>endA1 gyrA96 (nal^R) thi-1 recA1 relA1 lac glnV44F' [::Tn10 proAB⁺ lacI^q Δ(lacZ)M15] hsdR17 (r_K⁻ m_K⁺)</i>	Stratagene
BL21	F- <i>ompT hsdSB</i> (rB- mB-) <i>gal dcm</i> (DE3)	Novagen
JS41	DH300 Δ <i>rcsF</i> pAM238	This study
JS265	DH300 Δ <i>rcsF</i> pJS18	This study
JS346	DH300 Δ <i>rcsF rcsB::kan</i> pET22b	This study
JS267	JS346 pJS18	This study
JS325	DH300 <i>pal::kan</i>	This study
JS331	JS325 pJS20	This study
JS345	JS325 pJS24	This study
JS360	DH300 Δ <i>rcsF</i> pJS27	This study
JS363	JS346 pJS27	This study
JS364	DH300 Δ <i>rcsF</i> pSC202	This study
JS372	DH300 Δ <i>rcsF</i> pSC201	This study
JS395	JS346 pSC198	This study
JS396	JS346 pSC199	This study
JS397	JS346 pSC200	This study
JS398	JS346 pSC201	This study
JS573	JS346 pSC202	This study
JS574	DH300 Δ <i>rcsF</i> pSC198	This study
JS575	DH300 Δ <i>rcsF</i> pSC199	This study

JS576	DH300 $\Delta rcsF$ pSC200	This study
JS639	$\Delta rcsB$ <i>lpp::kan rcsF::rcsF_{Δ19-47}</i>	This study
JR30	<i>nlpD::kan</i>	This study
JR31	JR30 pJR8	This study
JR32	JR30 pJR10	This study
JR2	DH300 pAM238	This study
JR88	BL21 <i>rcsF::kan</i>	This study
JR90	JR88 pET28-cytoplasmic LolB-Strep	This study
JR187	<i>rcsB::kan rcsF::rcsF_{Δ19-47}</i>	This study
JR149	$\Delta nlpD$	This study
JR121	$\Delta nlpD$ <i>envC::kan</i>	This study
JR122	JR121 pJR8	This study
JR123	JR121 pJR10	This study
JR188	JR187 pAM238	This study
JR191	JR187 pAG833	This study
JR204	JR187 pJR203	This study
JR194	JR187 pBAD33	This study
JR211	JR187 pJR209	This study
JR257	JR187 pJR239	This study
JR274	JR149 <i>lpp::kan</i>	This study
JR279	JR274 pJR10	This study
JR292	JS325 pAM238	This study
JR293	JR187 pSC213	This study
JR44	<i>rcsB::kan rcsF::rcsF_{Δ19-47}</i> pJR48	This study

JR47	<i>rcsB::kan</i> pJR48	This study
JR77	<i>rcsB::kan rcsF::rcsF_{Δ19-47}</i> pBAD18	This study
JR78	<i>rcsB::kan</i> pBAD18	This study

861

862

863 **Extended Data Table 4: Plasmids used in this study.**

Plasmids	Features	Source
pAM238	IPTG-regulated P_{lac} , pSC101-based, spectinomycin (no <i>lacIQ</i>)	⁵⁰
pBAD18	Arabinose inducible P_{BAD} , ampicillin	⁵⁶
pBAD33	Arabinose inducible P_{BAD} , chloramphenicol	⁵⁶
pET28a	IPTG regulated T7 promoter, kanamycin	Novagen
pET22b	IPTG regulated T7 promoter, ampicillin	Novagen
pCP20	FLP ⁺ , λ cI857 ⁺ , λ_{PR} Rep ^{ts} , ampicillin, chloramphenicol	⁴⁹
pSIM5-Tet	pSC101 plasmid, <i>repAt</i> ^s , tetRA, λ -Red (Gram-Beta-Exo), cI857, tetracycline	Gift from D. Hughes
pJS18	pAM238 RcsF _{FKpA} FkpA linker (S94-E125)	This study
pJS20	pAM238 Pal _{WT}	This study
pJS24	pAM238 Pal _{Δ26-56}	This study
pJS27	pAM238 RcsF _{col} Colicin Ia linker (I213-K282)	This study
pSC198	pAM238 RcsF _{GS3} (C ₁₆ S ₁₇ M ₁₈ S ₁₉ GSGSGAMG)	This study
pSC199	pAM238 RcsF _{GS2} (C ₁₆ S ₁₇ M ₁₈ S ₁₉ GSGSGAMSGSGSGAMG)	This study
pSC200	pAM238 RcsF _{GS} (C ₁₆ S ₁₇ M ₁₈ S ₁₉ LFDAPAMSGSGSGAMSGSGSGAMG)	This study
pSC201	pAM238 RcsF _{Δ19-47} (C ₁₆ S ₁₇ M ₁₈ G ₁₉ P ₂₀)	This study
pSC202	pAM238 RcsF _{WT}	This study
pJR8	pAM238 NlpD _{WT}	This study
pJR10	pAM238 NlpD _{Δ29-64} (C ₂₆ S ₂₇ D ₂₈ A ₂₉)	This study
pJR48	pBAD18 LolA-6xHis	This study
pJR90	pET28 Cytoplasmic LolB-Strep	This study

pJR203	pBAD33 LolCDE-6xHis	This study
pJR209	pAM238 Lnt-Myc	This study
pJR239	pSC213 LspA-Flag	This study
pSC213	pAM238, IPTG-regulated P_{lac} , <i>lacIQ</i> , triple Flag tag	This study
pAG833	pAM238 Lgt-Myc	This study

864

865

866

867

868 **Extended Data Table 5: Primers used in this study.**

Primer	Sequence 5' to 3'
JS50_FkpAlinker_fw	acatccatgggggtccgaccaagagatcgaac
JS51_FkpAlinker_rv	atgtcggaccgggttcgttatcagccgcgtc
JS143_Pal_-100b	cgtcttccggcaactgatgg
JS144_Pal_+100b	ttggtgcctgagcaaaagcg
JS145_Pal_fw	ACATggtaccTTAATTGAATAGTAAAGGAATC
JS146_Pal_rv	ATGTtctagaTTAgtaaaccagtaccgcac
JS152_PalNoLinker_overlapPCR_fw	tgttcttccaacCAGGCTCGTCTGCAAATG
JS153_PalNoLinker_overlapPCR_rv	CAGACGAGCCTGggtggaagaacatgccgc
JS277_LolCDEHis_fw	ACATtctagaTCTTTGCTACAGCAACCAGAC
JS278_LolCDE_His_rv	ATGTctgcagTTAGTGATGGTGATGGTGATGACCctggccgctaaggactcg
JS289_lred_catSaccBin_RcsF_fw	tcctgattcaatattgacgtttgatcatacattaggaaactAAAATGAGACGTTGATCGG CACG

JS290_lred_catSa cBin_RcsF_rev	tatagggcgagcgaataacgcctatttgctcgaactggaaactgcATCAAAGGGAAAACCTGT CCA
JS291_lred_RcsF _catSacBout_fw	tcctgattcaatattgacgtttgatcatacattgaggaaactATGCGTGCTTTACCGATCTG TT
JS292_lred_RcsF _catSacBout_rv	tatagggcgagcgaataacgcctatttgctcgaactggaaactgcTCATTTGCGCCGTAATGTT AAGC
JS293_junction1lr ed_RcsFup_fw	gcgagctgttaaaggctg
JS294_junction2lr ed_RcsFdown_rv	gagcaatgagatgcagtgc
JS295_junction1lr ed_cat-out_rv	CGGGCAAGAATGTGAATAAAGG
JS296_junction2lr ed_sacB-out_fw	GCTGTACCTCAAGCGAAAGG
M13R	CAGGAAACAGCTATGACCATG
M13F	TGTAAAACGACGGCCAGT
PL145_rcsF_- 100b	cgcttttaccagacctggc
PL146_rcsF_+10 0	atatcattcaggacggcgcttgccc
PL153_rcsB_- 100b	acatctgattcgtgagaagg
PL154_rcsB+100 b	taatgggaatcgtaggccgg
PL168_Fw_lpp_- 100	CAATTTTTTTATCTAAAACCCAGCG
PL169_Rv_lpp_+ 100	CCAGAGCAAGGGAATATGTTACGCG
SH_Da linker_F	CATGaGcTTATTCGACGCGCCGGc
SH_Da linker_R	catggCCGGCGCGTTCGAATAAgCt
SH_RcsF(PstI)_R	gagaCTGCAGtcaTTTCGCCGTAATGTTAAG
SH_RcsFUR(kpn I)_F	GAGGGTACCcgtttgatcatacattg
RcsFss-Fsg (NcoI)_F	GCGGCTGTTCCATGGggccggtccgaattatac
RcsFss-Fsg (NcoI)_R	ggaccggccCCATGGAACAGCCGCTTAGCATGAG
SH_GS linker_F	CATGagtggctctggatctggtgc

SH_GS linker_R	catggcaccagatccagagccact
JR1_NlpD_fw	GAGATCTAGATTATTAACCAATTTTCTGGGGGATAA
JR2_NlpD_rv	AGAGCTGCAGTTATCGCTGCGGCAAATAACGCA
JR7_NlpDoverlap_fw	GGCTGGCAGGCTGTTCTGACGCGCAGCAACCGCAAATTCA
JR8_NlpDoverlap_rv	TGAATTTGCGGTTGCTGCGCGTCAGAACAGCCTGCCAGCC
JR23_Fw_NlpD-98	CAGGTCAGCGTATCGTGAACATC
JR24_Rv_NlpD+100	TCATTTAAATCATGAACTTTCAGCG
JR30_Fw_LolA_-28_pBAD18	ACATGGTACCCGGGAGTGACGTAATTTGAGGAAT
JR31_Rev_LolA_His_pBAD18	ATGTTCTAGAttaatgatgatgatgatgctcgaGCTTACGTTGATCATCTACC GTGAC
JR50_Rev_cytoplasmic_LolB_nostop_StrepTag_stop	CCAACCTCGAGTCACTTTTCGAACTGCGGGTGGCTCCAGCTTGCTTT CACTATCCAGTTATCCAT
JR56-Fw--100-envC	GTTGTCGCTG ATGGGTA
JR57-Rev-+100envC	AATCATCAATGACGATGGCA
JR74-Rev-Lnt-myctag-PstI	AAAAACTGCAGctacaggtcttctcgctaatcagttctgttcgcttgcTTTACGTCGCTG ACGCAGAC
JR77-Fw-NcoI-LspA	gagaCCATGGgtAGTCAATCGATCTGTTCAAC
JR78-Rev-LspA-no stop-BamHI	gagaGGATCCTTGTTTTTTCGCTCTAG
AG389_lgt_-49_Fw_KpnI	AAAAAggtaccTTCAATCGCTGTTCTCTTTC
AG393_lnt_-49_Fw_KpnI	AAAAAggtaccACCCAGCCGAAGCTGGATG
AG403_lgt_mycCT_PstI	AAAAACTGCAGctacaggtcttctcgctaatcagttctgttcgcttgcGGAAACGTGTT GCTGTGGGC
PL387-LolBwoss-Fw-NcoI	acacCCATGGccgttaccagcccaagg
ColicinIalinker_geneBLOCK	acatccatggggATTCTGGACACGCGGTTGTCAGAGCTGGAAAAAATG GCGGGGCAGCCCTTGCCGTTCTTGATGCACAACAGGCCCGTCTGC TCGGGCAGCAGACACGGAATGACAGGGCCATTCAGAGGCACGG AATAAACTCAGTTCAGTGACGGAATCGCTTAACACGGCCCGTAAT

	GCATTAACCAGAGCTGAACAACAGCTGACGCAACAGAAAgcggtccg acat
--	---

869

