1 **A dual motif mediates outer-membrane translocation and packing of**

2 **glycosidases into *Bacteroides* Outer Membrane Vesicles.**

3 Ezequiel Valguarnera[a], Nichollas Scott[b] and Mario F. Feldman[a]#.

4 [a]Department of Molecular Microbiology, Washington University in Saint Louis, Saint

5 Louis, Missouri, USA.

6 [b]Department of Microbiology and Immunology, The University of Melbourne, Victoria,

7 Australia.

8 Running head: Secretion of SusG into *Bacteroides* OMV

9 # Address correspondence to Mario F. Feldman, mariofeldman@wustl.edu.

10

11 ABSTRACT

12 Outer membrane vesicles (OMV) are spherical structures derived from the outer

13 membrane (OM) of Gram-negative bacteria. *Bacteroides spp. are* prominent

14 components of the human gut microbiota, and OMV produced with these species are

15 proposed to play key roles in gut homeostasis. OMV biogenesis in *Bacteroides* is a

16 poorly understood process. Here, we revisited the protein composition of *B. theta* OMVs

17 by mass spectrometry. We confirmed that OMVs produced by this organism contain

18 large quantities of glycosidases and proteases, with most of them being lipoproteins.

19 We found that most of these OMV-enriched lipoproteins are encoded by polysaccharide

20 utilization loci (PULs), such as the *sus* operon. We examined the subcellular localization

21 of the components of the Sus system, and found that the alpha-amylase SusG is highly

22   enriched in OMVs while the oligosaccharide importer SusC remains mostly in the OM.

23   We show that all OMV-enriched lipoproteins possess a lipoprotein export sequence

24   (LES) that mediates translocation of SusG from the periplasmic face of the OM towards

25   the extracellular milieu and is required for SusG to localize preferentially to OMVs. We

26   also show that surface-exposed SusG in OMVs is active and can rescue growth of

27   bacterial cells incapable of growing on starch as only carbon source. Our results

28   support the role of OMVs as "public goods" that can be utilized by other organisms with

29   different metabolic capabilities.

30   IMPORTANCE

31   Species from the *Bacteroides* genus are predominant members of the human gut

32   microbiota. OMVs in *Bacteroides* have been shown to be important for the homeostasis

33   of complex host-commensal relationships, mainly involving immune tolerance and

34   protection from disease. OMVs carry many enzymatic activities involved in the cleavage

35   of complex polysaccharides and have been proposed as public goods that can provide

36   growth to other bacterial species by release of polysaccharide breakdown products into

37   the gut lumen. Nevertheless, the mechanistic nature of OMV biogenesis is unclear for

38   *Bacteroides* spp. This works shows the presence of a negatively-charged rich amino

39   acid dual motif that is required for efficient packing of the surface-exposed alpha-

40   amylase SusG into OMVs. Discovery of this motif (LES) is the first step in the

41   generation of tailor made probiotic interventions that can exploit LES-related sequences

42   to generate *Bacteroides* strains displaying proteins of interest in OMVs.

43   **INTRODUCTION**

44   Outer membrane vesicles (OMVs) are small spherical structures derived from the outer

45   membrane (OM) of all Gram-negative bacterial species. OMVs are composed of

46   phospholipids, lipopolysaccharide (LPS) or lipooligosaccharide (LOS), and OM and

47   periplasmic proteins (1, 2). OMVs can mediate host-microbe interactions by facilitating

48   long-distance delivery of virulence factors, modulating the host-immune response, and

49   contributing to antibiotic resistance (1, 3-8).   Despite these key roles in bacterial

50   physiology, OMV biology is poorly understood.   Recent research indicates that OMVs

51   are produced by diverse mechanisms. For example, OMVs are proposed to be

52   generated by LPS remodeling in *Porphyromonas gingivalis*, *Salmonella enterica* and

53   *Pseudomonas aeuroginosa* (9-12). In contrast, OMVs from *Haemophilus influenza* and

54   *Vibrio cholerae* are thought to be the result of an accumulation of phospholipids in the

55   OM outer leaflet mediated by their specialized VacJ/Yrb transporter (13). Thus, it

56   appears that there is not a universal mechanism of OMV biogenesis. For most species,

57   including *Bacteroides* spp., OMV biogenesis remains poorly understood (5).

58   Species from the phylum Bacteroidetes compose a major part of the human gut

59   microbiota, and OMVs from these organisms are proposed to play important roles in the

60   commensal-host relationship (14, 15). These roles include the delivery of

61   immunomodulatory molecules to host immune cells, an interaction that appears to help

62   prevent colitis flares in patients with inflammatory bowel disease (IBD) (7, 8).

63   Furthermore, *Bacteroides* OMVs have been proposed to interfere with intracellular Ca^{2+}

64   signaling in host cells (16). Most studies focus primarily on two predominant species in

65   the human gut, *B. thetaiotaomicron ("B. theta") and B. fragilis.* These species produce

66   large amounts of uniformly sized OMVs that have a distinct protein composition

67 compared to the OM, indicating that these OMV particles are not by-products of

68 bacterial lysis (5). Most *Bacteroides* OMV-exclusive proteins are putative acidic

69 lipoprotein hydrolases, suggesting that proteins with similar structural and

70 physicochemical properties are selectively sorted to OMVs (5). Many *Bacteroides*

71 enzymatic lipoproteins are encoded on polysaccharide utilization loci (PULs), which

72 constitute ~20% of the *B. theta* genome and are essential for the breakdown and

73 acquisition of plant, fungi and mucosal complex polysaccharides (17-19). PULs consist

74 of at least one TonB-dependent receptor similar to a SusC-like protein; one nutrient

75 binding accessory protein or SusD; and other accessory binding proteins (20). PULs

76 can also carry two-component systems that sense nutrient variations in the media, with

77 subsequent induction of PUL genes required for the utilization of complex carbon

78 sources (21). Hence, *Bacteroides* cells can modify the enzymatic composition of OMVs

79 according to available carbon sources (5, 22). The enzymatic arsenal carried by

80 *Bacteroides* OMVs appears to carry a "social" function, as the products of OMV-

81 mediated hydrolysis can be utilized by other bacteria within the gut (22, 23).

82 Here, we further characterized the protein composition of *B. theta* OMVs. We confirmed

83 that OMVs produced by this organism mainly contain putative acidic lipoprotein

84 hydrolases or carbohydrate-binding proteins. Many of these OMV-enriched lipoproteins

85 were found to be encoded by PULs. We examined the subcellular localization of the

86 components of the archetypical PUL, the starch utilization system (Sus) (20, 24), and

87 found that the alpha-amylase SusG and other Sus lipoproteins are highly enriched in

88 OMVs. In contrast, the oligosaccharide importer SusC remains mostly in the OM. We

89 show that the presence of a lipoprotein export sequence (LES) mediates translocation

90   of SusG from the periplasmic face onto the extracellular milieu and is both required and

91   sufficient for SusG to localize preferentially to OMVs. Our results support the role of

92   OMVs as "public goods" that can be utilized by other organisms with different metabolic

93   capabilities.

94   **RESULTS**

95   **Proteomics analysis of membrane and OMV fractions from *B. theta***

96   Electron micrograph analysis confirmed that *B. theta* produces large amounts of uniform

97   OMV particles (Figure 1). We previously performed a proteomic analysis of *B. theta* OM

98   and OMV employing Triton X-100 for the purification of OM proteins (5). In this work, we

99   employed N-lauroylsarcosine (Sarkosyl), which has been widely used for the separation

100   of inner membrane (IM) and OM fractions in Gram-negative bacteria (25-27) to verify

101   that the apparent selective fractionation of proteins into OM and OMV was not due to

102   the use of a specific detergent. Using early stationary cultures from *B. theta,* we

103   prepared the different membrane fractions as indicated in Figure S1. In our previous

104   study, OMVs were not treated with the detergent employed to extract the IM proteins

105   from total membrane preparations. To rule out possible detergent effects, we added an

106   additional step consisting of incubating OMV in 1% Sarkosyl prior to the

107   ultracentrifugation of the samples to recover an OMV supernatant (OMV-S) and an

108   OMV pellet (OMV-P). Samples were lyophilized for proteomic analysis by LC-MS/MS

109   and aliquots were visualized by Coomassie Blue (Figure S1).

110   Our new dataset with the annotation and predicted function and localization for proteins

111   in all fractions is provided as Table S1. We confirmed our previous findings showing an

112   enrichment of lipoproteins in the OMV in comparison with the OM fraction (Table 1,

113   Table S2) (5). Figure 2 highlights the top enriched proteins in the OMV. We found that

114   18 out of the 23 OMV-exclusive lipoproteins from our previous study are also enriched

115   in the new OMV preparations (5). We then cloned three OMV and three OM-enriched

116   proteins identified in our studies into the expression vector pFD340 as 6xHis-tagged

117   proteins, to confirm their localization by Western Blot. Figure 3a shows that a putative

118   cell surface protein (BT_1488), as well as a putative calpain-like protease (BT_3960),

119   and a putative zinc peptidase (BT_3237), are highly enriched in OMVs. Figure 3b shows

120   that BT_0418, BT_2844 and BT_2817, identified as OM-enriched proteins by MS, are

121   retained at the OM and are not present in OMVs. BT_0418 is a Porin F ortholog (8 β-

122   strand protein with a PG-binding domain), BT_2817 is a TonB-dependent receptor (22

123   β-strand protein), and BT_2844, a lipoprotein containing tetratricopeptide repeat (TPR)

124   motifs. These set of experiments demonstrate that lipoproteins are indeed differentially

125   sorted between OM and OMV.

Common features of OMV-enriched lipoproteins

127   Consistent with our previous results, we found that lipoproteins enriched in OMVs are

128   acidic, with an average isoelectric point of 4.86 (5). We also determined that OMV-

129   enriched lipoproteins possess a negatively charged-rich amino acid motif $S(D/E)_3$

130   adjacent to the Cysteine residue required for lipoprotein attachment (Figure 4, Table

131   S2). A similar motif (K-(D/E)2 or Q-A-(D/E)2) has recently been described in the oral

132   pathogen *Capnocytophaga canimorsus*, a member of the Bacteroidetes phylum. The

133   report showed that this motif functions as a lipoprotein export signal (LES) required for

134   surface exposure of OM lipoproteins (28). We found that many OMV-enriched *B. theta*

135   lipoproteins are putative protein and sugar hydrolases, required for the breakdown of

136   complex nutritional sources. The presence of a LES on OMV proteins is consistent with

137   their annotated functions, as they are required to face the extracellular milieu to access

138   their cognate substrates. Conversely, OM-enriched lipoproteins like BT_2844, do not

139   carry a LES motif. Instead, lipoproteins preferentially retained at the OM have lower

140   sequence conservation and lower frequency of negatively-charged amino acids along

141   the residues adjacent to the lipoprotein attachment cysteine (Figure 4b, Table S2).

142   SusG and other Sus lipoproteins are enriched in OMV and SusC is retained at the

143   OM.

144   In *B. theta*, hydrolytic lipoproteins are mainly encoded in PULs. PUL and PUL-like

145   operons account for approximately 20% of the *B. theta* genome. 36 lipoproteins

146   encoded in PUL and PUL-like operons were found in our OMV-enriched proteins. One

147   of the most studied lipoproteins from *B. theta* is the α-amylase SusG, encoded by the

148   *sus* operon and essential for starch catabolism (29, 30). The *sus* operon has been

149   shown to be induced by starch and maltooligosaccharides (31). Our MS data shows that

150   even in non-inducing conditions, two lipoproteins encoded by the *sus* operon, SusD and

151   SusE, are enriched in OMVs (Table S2). Given the importance of starch in the

152   mammalian diet, we investigated the subcellular localization of the components of the

153   Sus operon. We found that all Sus lipoproteins, in particular SusG, are enriched in the

154   OMV. The only exception was the porin SusC, which is mostly retained at the OM

155   (Figure 5). This result strongly suggests that hydrolytic lipoproteins can exert their

156   biological effect not only at the level of the outer membrane but also through OMVs.

157   LES is required for SusG exposure and packing into OMVs

158   All *B. theta* endoglycanases involved in the first step of polysaccharide breakdown,

159   including SusG, are surface-exposed lipoproteins (29, 32). Because most OMV-

160   enriched lipoproteins contain a LES motif, we investigated a possible association

161   between surface exposure and packing into OMVs. SusG only contains two Asp

162   residues following the +2 Ser residue, and therefore, for the purpose of this work, we

163   provisionally define its LES motif as CSDD. We performed a mutational analysis on the

164   LES motif of SusG and analyzed the localization of the protein by fractionation and

165   Western blot (Figure 6). We cloned WT His-tagged SusG and a set of proteins carrying

166   amino acid replacements in the LES (Figure 6a) and expressed them in the Δ*susG*

167   strain (33). Replacement of the lipid attachment site (C23) by Ala resulted in abrogation

168   of SusG-recruitment into OMVs. Furthermore, substitution of a single Asp residue with

169   Ala (CSAD, CSDA, CAAD) decreased OMV-packaging (~50% relative to WT), while

170   mutation of both Asp residues by either Ala or Lys (CSAA, CAAA, CAKK) had a more

171   dramatic effect (~15% relative to WT) (Figure 6c). Conservative replacement of Asp by

172   Glu (CSEE) displayed a WT-like behavior. These results indicate that the LES motif

173   $CS(D/E)_2$ is required for SusG packaging into OMVs.

174   The LES motif has been defined as a surface exposure tag in *C. canimorsus* (28). We

175   used the WT LES SusG construct as well as the non-Asp variants (CSAA, CAAA) to

176   determine whether this motif is also required for surface exposure in *B. theta*. Whole

177   cells and OMVs from the different strains were subjected to Proteinase K (ProK)

178   sensitivity analysis. Only the WT SusG was degraded by ProK (Figure 7). As a control,

179   we employed a periplasmic soluble protein that localizes into the lumen of OMVs

180   (BT_0766) and is therefore protected from ProK degradation. Taken together, these

181 experiments confirm that LES mediates both SusG surface exposure and enrichment in

182 OMVs.

183 OMV containing SusG rescue Δ*susG* growth in starch

184 Our experiments demonstrated that SusG is packed and surface-exposed in OMVs.

185 OMVs carrying certain glycosyl hydrolases can digest complex polysaccharides

186 providing essential nutrients to bacteria unable to degrade these substrates (22, 23).

187 The Δ*susG* strain is unable to grow on minimal media with starch as the sole carbon

188 source (Figure 8). We investigated whether surface exposure of OMV-delivered SusG

189 can rescue the Δ*susG* growth phenotype on starch. For this, we employed OMVs

190 produced by the Δ*susG* strain expressing SusG with mutated LES (ASDD, CSAA, and

191 CAAA). ASDD carries soluble SusG in the periplasm, while the CSAA and CAAA SusG

192 variants face the periplasmic side of the OM (Fig. 7). Purified OMVs were added to the

193 Δ*susG* strain in minimal media containing starch. Only WT OMVs restored growth of the

194 Δ*susG* strain to wild-type levels (Figure 8). OMV from Δ*susG* carrying vector control

195 (pFD340) or expressing the non-lipoprotein mutant (ASDD), or the CSAA mutant, were

196 unable to rescue the growth of SusG, while the mutant CAAA displayed an intermediate

197 phenotype. Altogether, our results indicate that OMVs carrying surface exposed SusG

198 can mediate cross-feeding of other bacteria.

199 DISCUSSION

200 The human gut microbiome is largely composed of species from the *Bacteroides* genus.

201 *Bacteroides* spp are important for gut homeostasis (14, 15). They establish ecological

202 interactions with each other and are also involved in host-commensal relationships,

203   especially regarding the development of the immune system (7, 8, 34). Here, we show

204   that *B. theta* produces large amounts of uniformly sized OMVs. Using an optimized

205   methodology for the purification of different membrane fractions, we confirmed that

206   OMV are highly enriched with lipoproteins, particularly glycosyl hydrolases. Our MS

207   data identified the presence of a lipoprotein exposure sequence (LES) in all OMV-

208   enriched lipoproteins. Employing the α-amylase SusG as a model, we showed that the

209   LES is required for surface exposure and recruitment into OMVs. To our knowledge,

210   this is the first identification of a signal involved in protein sorting into OMVs.

211   A handful of lipoprotein surface transport mechanisms have been described in Gram-

212   negative bacteria (35-38). Shuttling of a group of *Neisseria meningitidis* lipoproteins to

213   the surface is dependent on proteins Slam1 and Slam2, although the sorting

214   mechanism has not been defined (38, 39). Moreover, the well-studied Bam system for

215   folding of beta-barrel proteins in the outer membrane has been shown to export specific

216   lipoproteins (36). We have not identified orthologs of Slam1 or Slam2 by sequence

217   similarity in any *Bacteroides* genomes. The existence of LES motifs in *B. theta*

218   lipoproteins, together with the discovery of the LES in *C. canimorsus*, suggest the

219   existence of a conserved phylum-wide mechanism that flips specific lipoproteins

220   towards the extracellular milieu among Bacteroidetes. Defining a functional LES in

221   *Bacteroides* spp. constitutes the first step towards the identification of the machinery

222   that mediates lipoprotein surface exposure in these organisms. Our experiments

223   confirmed that the two Asp residues in position +3 and +4 respecting the lipid

224   attachment site (+1) are essential components of SusG LES. The proposed *B. theta*

225   LES motif $CS(D/E)_2$ is similar but not identical to the $(K-(D/E)_2$ or $Q-A-(D/E)_2)$ LES

226   proposed for *C. canimorsus*. An exhaustive mutagenesis analysis of multiple proteins

227   from several species will be required to exactly define a consensus sequence for the

228   LES motif among Bacteroidetes.

229   Our results suggest that the LES has a dual role in surface exposure and recruitment of

230   lipoproteins into OMVs (Fig 9). It is possible that the LES is involved in protein packing

231   into OMVs through a direct interaction with an unknown component of a putative sorting

232   machinery. Another possibility is that LES only mediates surface exposure, and flipped

233   lipoproteins are subsequently directed for enrichment into OMVs. In the case of SusG,

234   we have confirmed that LES is required for packing into OMVs. Future work is required

235   to unravel the link between surface exposure and OMV packing.

236   In the current models for PUL systems, the hydrolytic enzymes and the oligosaccharide

237   importers are present at the outer membrane (20, 32). However, we determined that

238   functional glycosyl hydrolases are mainly packed into OMVs. We also found that the

239   glycosidases present in OMVs can provide substrates to support the growth of other

240   bacteria, in agreement with the previously proposed function of OMVs as public goods

241   (22, 23). We propose a model in which, in a process involving the LES, the

242   glycosidases are preferentially sorted into OMVs (Fig 9). The secreted OMVs, armed

243   with an arsenal of hydrolases, can digest diverse dietary polysaccharides and host

244   glycoconjugates, making the mono- and oligo-saccharides available to all members of

245   the microbiota. Depending on the diet composition, *Bacteroides* spp can induce and

246   package different enzyme repertoires, and therefore, members of the microbiota can act

247   as donors or acceptors in their ecological niches. In addition, pathogens such as

248 *Campylobacter*, *Salmonella*, and *Clostridium* can also benefit from the glycosidic activity

249 contained in OMVs (40, 41).

250 We have identified a set of OMV- and OM-enriched proteins that may be employed as

251 markers to investigate the process of OMV biogenesis through the visualization of OMV

252 in the mammalian gut. For example, these markers would allow the differentiation *in*

253 *vivo* between *bona-fide* OMV and cell-lysis derived material. Furthermore, the

254 identification of LES as a signal for surface exposure and transport into OMVs

255 constitutes a starting point for the design of novel probiotic interventions in animal and

256 human health. In the future, the LES of SusG or other OMV-proteins could be employed

257 to engineer *Bacteroides* strains to secrete OMVs packed with medically-relevant

258 proteins in the mammalian gut.

259 MATERIALS AND METHODS

260 **Bacterial strains and growth conditions**

261 Oligonucleotides, strains and plasmids are described in Table S3. *Bacteroides* strains

262 were grown in an anaerobic chamber (Coy Laboratories) using an atmosphere of 10%

263 H_2, 5% CO_2, 85% N_2. For liquid growth, TYG, TYM (TY media supplemented with 0.5%

264 maltose instead of 0.5% glucose), or minimal media supplemented with 0.5% glucose or

265 0.5% potato starch were prepared as previously described (17, 42). BHI-Agar with 10%

266 defibrinated horse blood was used as solid media. Antibiotics were used as indicated;

267 Amp, 100 µg/ml; Erm, 25 µg/ml; BrdU, 200 µg/ml.

268 **Construction of plasmids, mutagenesis and generation of *susC*-3xFlag strain**

269   Constructs for over-expression were built on pFD340 (43). Inserts obtained by PCR

270   were digested with restriction enzymes, purified and ligated onto digested vector

271   pFD340 as previously reported. For cloning of BT_0766-Flag into pFD340, purified PCR

272   inserts were integrated into PCR-amplified pFD340 using InFusion Cloning Kit

273   (Clontech). Mutagenesis of pFD340/*susG*-6xHis was carried out by inverse PCR with

274   Pfu Turbo (Agilent) using overlapping oligos carrying the mismatch required as

275   described elsewhere. Generation of a *susC*-3xFlag strain was carried out using the *B.*

276   *theta* Δ*tdk* strategy as previously described (44). Briefly, 1000 bp upstream and

277   downstream fragments of a C-terminal *susC*-3xFlag translational fusion were cloned

278   into pExchange-*tdk*. Constructs were conjugated into *B. theta* Δ*tdk* cells using

279   previously transformed *E. coli* S17-1 λpir as a donor, and strain plating and selection

280   was performed as previously described (44).

OMV preparations

282   Outer membrane vesicles were purified by ultracentrifugation of filtered spent media as

283   previously described by our group (5). For MS analysis, OMV preparations were

284   resuspended in 50 mM HEPES pH 7.4 and N-lauroyl sarcosine was added to 1% final

285   concentration in 1.5 ml polyallomer tubes (Beckman Coulter). Samples were incubated

286   with gentle rocking for 1 hour at room temperature (RT) and ultracentrifuged at 100000

287   x g for 2 hours at RT. Supernatants were recovered (OMV-S) and pellets (OMV-P) were

288   resuspended in 50 mM HEPES pH 7.4. Protein content was quantified using DC Protein

289   Assay Kit (Bio-Rad). Fractions were lyophilized for MS analysis.

Membrane preparations

291   Total membrane preparations were performed by cell lysis and ultracentrifugation as

292   previously described(5). For separation of IM and OM, total membranes were

293   resuspended in 50 mM HEPES pH 7.4 using a 2 ml glass tissue grinder with a PTFE

294   pestle (VWR), and N-lauroyl sarcosine was added to 1% final concentration in 1.5 ml

295   polyallomer tubes (Beckman Coulter). Samples were incubated with gentle rocking for 1

296   hour at room temperature (RT) and ultracentrifuged at 100000 x g for 2 h at RT.

297   Supernatants were recovered (IM) and pellets were resuspended in 50 mM HEPES pH

298   7.4. Protein content was quantified using DC Protein Assay Kit (Bio-Rad). Fractions

299   were lyophilized for MS analysis.

300   **Mass spectrometry analysis**

301   **Protein clean up and In-solution digestion**: Lyophilized protein preparations were

302   solubilized in lysis buffer (4% SDS, 10mM, 100mM Tris pH8.5) by boiling for 10 min and

303   the protein content assess by BCA protein assay according to the manufactures

304   instruction. 300 μg of each preparation and biological replicate were adjusted to a

305   volume of 200 μl and precipitated overnight using 800 μl of ice-cold acetone (1:4 v/v).

306   Sample were spun down at 16,000G for 10min at 0°C to the resulting protein precipitate

307   and acetone remove. Residue acetone was driven off at 70°C for 5 min. Protein

308   precipitants were resuspended in 6 M urea, 2 M thiourea, 40 mM NH_4HCO_3 and

309   reduced / alkylated prior to digestion with Lys-C (1/200 w/w) then trypsin (1/50 w/w)

310   overnight as previously described (45). Digested samples were acidified to a final

311   concentration of 0.5% formic acid and desalted with using C_{18} stage tips (46, 47)

312 **LFQ based quantitative proteome LC-MS.** Purified peptides prepared were re-

313 suspend in Buffer A* and separated using a two-column chromatography set up

314 composed of a PepMap100 C18 20 mm x 75 μm trap and a PepMap C18 500 mm x 75

315 μm analytical column (Thermo Fisher Scientific). Samples were concentrated onto the

316 trap column at 5 μL/min for 5 minutes and infused into an Orbitrap Q-exactive™ plus

317 Mass Spectrometer (Thermo Fisher Scientific) at 300 nl/minute via the analytical column

318 using a Dionex Ultimate 3000 UPLC (Thermo Fisher Scientific). 125-minute gradients

319 were run altering the buffer composition from 1% buffer B to 28% B over 95 minutes,

320 then from 28% B to 40% B over 10 minutes, then from 40% B to 100% B over 2

321 minutes, the composition was held at 100% B for 3 minutes, and then dropped to 3% B

322 over 5 minutes and held at 3% B for another 10 minutes. The Q-exactive™ Mass

323 Spectrometer was operated in a data-dependent mode automatically switching between

324 the acquisition of a single Orbitrap MS scan (60,000 resolution) and 15 MS-MS scans

325 (Orbitrap HCD, 35,000 resolution maximum fill time 110 ms and AGC $2*10^5$).

326

327 **Mass spectrometry data analysis.** Identification and LFQ analysis and was

328 accomplished using MaxQuant (v1.5.3.1) (48). Searches were performed against the *B.*

329 *thetaiotaomicron* (strain ATCC 29148 / VPI-5482) proteome (Uniprot proteome id

330 UP000001414, downloaded 20-05-2017, 4,782 entries) with carbamidomethylation of

331 cysteine set as a fixed modification and the variable modifications of oxidation of

332 methionine and acetylation of protein N-termini. Searches were performed with trypsin

333 cleavage specificity allowing 2 miscleavage events with a maximum false discovery rate

334 (FDR) of 1.0% set for protein and peptide identifications. To enhance the identification

335   of peptides between samples the Match Between Runs option was enabled with a

336   precursor match window set to 2 minutes and an alignment window of 10 minutes. For

337   label-free quantitation, the MaxLFQ option within Maxquant (49) was enabled in addition

338   to the re-quantification module. The resulting protein group output was processed within

339   the Perseus (v1.4.0.6) (50) analysis environment to remove reverse matches and

340   common protein contaminates prior. For LFQ comparisons missing values were

341   imputed using Perseus. Visualization was done using Perseus and R. Predicted

342   localization and topology analysis for proteins identified by MS was performed using

343   LipoP and TOPCONS (51, 52).

344   **SDS-PAGE and OMV/OM ratio determination**

345   Membrane and OMV fractions were analyzed by standard 10-12% Tris-Glycine SDS-

346   PAGE as described elsewhere. Briefly, 10 µg of each fraction (TM/IM/OM/OMV) were

347   loaded onto SDS-PAGE gels, transferred onto nitrocellulose membranes and Western

348   Blots were performed using the LI-COR system. Membranes were blocked using TBS-

349   based Odyssey blocking solution (LI-COR). Primary antibodies used in this study were

350   rabbit polyclonal anti-His (Thermo Fisher), mouse monoclonal anti-Flag M2 (Sigma),

351   mouse monoclonal anti-*E. coli* RNApol subunit alpha (Biolegend) and mouse polyclonal

352   anti-SusD/E/F/G (Nicole Koropatkin/Eric Martens, University of Michigan). Secondary

353   antibodies used were IR-Dye anti-rabbit 780 and IR-Dye anti-mouse 680 antibodies (LI-

354   COR). Imaging was performed using an Odyssey CLx scanner (LI-COR).

355   For validation of MS data, duplicate SDS-PAGE gels were stained by Coomassie Blue

356   as described elsewhere and gel images were acquired to determine fraction quality and

357   relative abundance (Fig. S2 and S3). For OMV/OM determination of SusG-6xHis

358   experiments, cells were grown in minimal media with glucose and fraction were

359   prepared as described. After transferring SDS-PAGE gels, nitrocellulose membranes

360   were incubated with REVERT Total Protein Stain as described by the manufacturer (LI-

361   COR) and imaged immediately at 680 nm (Fig. S4). After imaging, Western Blots using

362   SusG antibodies were carried out and membranes were scanned at 780 nm. Total

363   intensity values were calculated for each lane using the Odyssey scanner software

364   (Image Studio, LI-COR). Total intensity values for each fraction, determined by

365   scanning each lane of the REVERT stain image, were used to relativize each SusG

366   signal. Relativized OMV and OM SusG fluorescence signals were used to calculate an

367   OMV/OM SusG ratio. Such ratio was considered 100% for the WT LES SusG construct.

368   Statistical significance between different OMV/OM ratios was determined by unpaired t-

369   test for each pair of SusG LES variants.

370

371   **Proteinase K assays**

372   Strains from *B. theta* were grown in minimal media with glucose and cells were washed

373   with PBS and normalized to 9 OD_{600}/ml. 540 µl of PBS and 10 µl of a Proteinase K

374   (ProK) solution (20 mg/ml) were added to 450 µl of the cell suspension. Tubes were

375   incubated at 37°C and 200 µl aliquots were removed at different time points and

376   precipitated using trichloroacetic acid (TCA, 20% v/w final concentration). Precipitated

377   aliquots were washed twice with acetone, and pellets were resuspended with Laemmli

378   buffer for Western Blot analysis. A non-treated control of the cell suspension was

379   incubated for the longest time point of the experiment and TCA-precipitated as

380   described. A similar procedure was followed for OMV ProK treatments, 450 µl of purified

381   OMVs were treated in the same conditions as whole cells.

382

383   **Growth curves and OMV complementation**

384   For growth curves, wild-type or Δ*susG* strains were grown overnight in TYG media.

385   Cultures were washed with minimal media (MM) without any carbon source and

386   normalized by OD_{600}. Minimal media with 0.5% starch as only carbon source was

387   inoculated with WT or Δ*susG* strains to final OD_{600} = 0.05. OMVs were purified from

388   Δ*susG* strains containing different pFD340/*susG*-6xHis derivatives grown in minimal

389   media with glucose, and were added to the Δ*susG* cultures in minimal media with starch

390   (1 µg/ml final OMV concentration). Aliquots were taken at different times and OD_{600} was

391   measured to determine growth. Statistical significance when comparing two growth

392   curves was determined by performing one pair of unpaired t-test analysis per each time

393   point.

394   **Transmission electron microscopy.**

395   For negative staining and analysis by transmission electron microscopy, bacterial

396   suspensions in PBS were allowed to absorb onto freshly glow-discharged

397   Formvar/carbon-coated copper grids for 10 min. Grids were washed in dH_2O and

398   stained with 1% aqueous uranyl acetate (Ted Pella, Inc., Redding, CA) for 1 min.

399   Excess liquid was gently wicked off, and grids were allowed to air dry. Samples were

400   viewed on a JEOL 1200EX transmission electron microscope (JEOL United States,

401 Peabody, MA) equipped with an AMT 8-megapixel digital camera (Advanced

402 Microscopy Techniques, Woburn, MA).

ACKNOWLEDGMENTS

408

REFERENCES

410 1. Haurat MF, Elhenawy W, Feldman MF. 2015. Prokaryotic membrane vesicles: new insights on

411   biogenesis and biological roles. Biol Chem 396:95-109.

412 2. Schwechheimer C, Kuehn MJ. 2015. Outer-membrane vesicles from Gram-negative bacteria:

413   biogenesis and functions. Nat Rev Microbiol 13:605-19.

414 3. Alaniz RC, Deatherage BL, Lara JC, Cookson BT. 2007. Membrane vesicles are immunogenic

415   facsimiles of Salmonella typhimurium that potently activate dendritic cells, prime B and T cell

416   responses, and stimulate protective immunity in vivo. J Immunol 179:7692-701.

417 4. Bomberger JM, Maceachran DP, Coutermarsh BA, Ye S, O'Toole GA, Stanton BA. 2009. Long-

418   distance delivery of bacterial virulence factors by Pseudomonas aeruginosa outer membrane

419   vesicles. PLoS Pathog 5:e1000382.

420 5. Elhenawy W, Debelyy MO, Feldman MF. 2014. Preferential packing of acidic glycosidases and

421   proteases into Bacteroides outer membrane vesicles. MBio 5:e00909-14.

422   6.   Ellis TN, Kuehn MJ. 2010. Virulence and immunomodulatory roles of bacterial outer membrane

423   vesicles. Microbiol Mol Biol Rev 74:81-94.

424   7.   Hickey CA, Kuhn KA, Donermeyer DL, Porter NT, Jin C, Cameron EA, Jung H, Kaiko GE,

425   Wegorzewska M, Malvin NP, Glowacki RW, Hansson GC, Allen PM, Martens EC, Stappenbeck TS.

426   2015. Colitogenic Bacteroides thetaiotaomicron Antigens Access Host Immune Cells in a

427   Sulfatase-Dependent Manner via Outer Membrane Vesicles. Cell Host Microbe 17:672-80.

428   8.   Shen Y, Giardino Torchia ML, Lawson GW, Karp CL, Ashwell JD, Mazmanian SK. 2012. Outer

429   membrane vesicles of a human commensal mediate immune regulation and disease protection.

430   Cell Host Microbe 12:509-20.

431   9.   Haurat MF, Aduse-Opoku J, Rangarajan M, Dorobantu L, Gray MR, Curtis MA, Feldman MF.

432   2011. Selective sorting of cargo proteins into bacterial membrane vesicles. J Biol Chem

433   286:1269-76.

434   10.   Elhenawy W, Bording-Jorgensen M, Valguarnera E, Haurat MF, Wine E, Feldman MF. 2016. LPS

435   Remodeling Triggers Formation of Outer Membrane Vesicles in Salmonella. MBio 7.

436   11.   Bonnington KE, Kuehn MJ. 2016. Outer Membrane Vesicle Production Facilitates LPS

437   Remodeling and Outer Membrane Maintenance in Salmonella during Environmental Transitions.

438   MBio 7.

439   12.   Kadurugamuwa JL, Beveridge TJ. 1995. Virulence factors are released from Pseudomonas

440   aeruginosa in association with membrane vesicles during normal growth and exposure to

441   gentamicin: a novel mechanism of enzyme secretion. J Bacteriol 177:3998-4008.

442   13.   Roier S, Zingl FG, Cakar F, Durakovic S, Kohl P, Eichmann TO, Klug L, Gadermaier B, Weinzerl K,

443   Prassl R, Lass A, Daum G, Reidl J, Feldman MF, Schild S. 2016. A novel mechanism for the

444   biogenesis of outer membrane vesicles in Gram-negative bacteria. Nat Commun 7:10515.

445   14.   Gilbert JA, Blaser MJ, Caporaso JG, Jansson JK, Lynch SV, Knight R. 2018. Current understanding

446   of the human microbiome. Nature Medicine 24:392.

447   15.   Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. 2012. Diversity, stability and

448   resilience of the human gut microbiota. Nature 489:220-230.

449   16.   Stentz R, Osborne S, Horn N, Li AW, Hautefort I, Bongaerts R, Rouyer M, Bailey P, Shears SB,

450   Hemmings AM, Brearley CA, Carding SR. 2014. A Bacterial Homolog of a Eukaryotic Inositol

451   Phosphate Signaling Enzyme Mediates Cross-kingdom Dialog in the Mammalian Gut. Cell reports

452   6:646-56.

453   17.   Martens EC, Chiang HC, Gordon JI. 2008. Mucosal Glycan Foraging Enhances Fitness and

454   Transmission of a Saccharolytic Human Gut Bacterial Symbiont. Cell Host & Microbe 4:447-457.

455   18.   Bjursell MK, Martens EC, Gordon JI. 2006. Functional Genomic and Metabolic Studies of the

456   Adaptations of a Prominent Adult Human Gut Symbiont, Bacteroides thetaiotaomicron, to the

457   Suckling Period. Journal of Biological Chemistry 281:36269-36279.

458   19.   Martens EC, Lowe EC, Chiang H, Pudlo NA, Wu M, McNulty NP, Abbott DW, Henrissat B, Gilbert

459   HJ, Bolam DN, Gordon JI. 2011. Recognition and degradation of plant cell wall polysaccharides

460   by two human gut symbionts. PLoS Biol 9:e1001221.

461   20.   Foley MH, Cockburn DW, Koropatkin NM. 2016. The Sus operon: a model system for starch

462   uptake by the human gut Bacteroidetes. Cell Mol Life Sci 73:2603-17.

463   21.   Schwalm ND, 3rd, Townsend GE, 2nd, Groisman EA. 2016. Multiple Signals Govern Utilization of

464   a Polysaccharide in the Gut Bacterium Bacteroides thetaiotaomicron. MBio 7.

465   22.   Rakoff-Nahoum S, Coyne MJ, Comstock LE. 2014. An ecological network of polysaccharide

466   utilization among human intestinal symbionts. Curr Biol 24:40-9.

467   23.   Rakoff-Nahoum S, Foster KR, Comstock LE. 2016. The evolution of cooperation within the gut

468   microbiota. Nature 533:255-9.

469   24.   Shipman JA, Berleman JE, Salyers AA. 2000. Characterization of Four Outer Membrane Proteins

470   Involved in Binding Starch to the Cell Surface ofBacteroides thetaiotaomicron. Journal of

471   Bacteriology 182:5365-5372.

472   25.   Wilson MM, Anderson DE, Bernstein HD. 2015. Analysis of the Outer Membrane Proteome and

473   Secretome of Bacteroides fragilis Reveals a Multiplicity of Secretion Mechanisms. PLOS ONE

474   10:e0117732.

475   26.   Ottman N, Huuskonen L, Reunanen J, Boeren S, Klievink J, Smidt H, Belzer C, de Vos WM. 2016.

476   Characterization of Outer Membrane Proteome of Akkermansia muciniphila Reveals Sets of

477   Novel Proteins Exposed to the Human Intestine. Frontiers in Microbiology 7:1157.

478   27.   Renzi F, Manfredi P, Mally M, Moes S, Jenö P, Cornelis GR. 2011. The N-glycan Glycoprotein

479   Deglycosylation Complex (Gpd) from Capnocytophaga canimorsus Deglycosylates Human IgG.

480   PLOS Pathogens 7:e1002118.

481   28.   Lauber F, Cornelis GR, Renzi F. 2016. Identification of a New Lipoprotein Export Signal in Gram-

482   Negative Bacteria. MBio 7.

483   29.   Shipman JA, Cho KH, Siegel HA, Salyers AA. 1999. Physiological characterization of SusG, an

484   outer membrane protein essential for starch utilization by Bacteroides thetaiotaomicron. J

485   Bacteriol 181:7206-11.

486   30.   Koropatkin NM, Smith TJ. 2010. SusG: a unique cell-membrane-associated alpha-amylase from a

487   prominent human gut symbiont targets complex starch molecules. Structure 18:200-15.

488   31.   Spence C, Wells WG, Smith CJ. 2006. Characterization of the Primary Starch Utilization Operon in

489   the Obligate Anaerobe Bacteroides fragilis: Regulation by Carbon Source and Oxygen. Journal of

490   Bacteriology 188:4663-4672.

491   32.   Luis AS, Briggs J, Zhang X, Farnell B, Ndeh D, Labourel A, Basle A, Cartmell A, Terrapon N, Stott K,

492   Lowe EC, McLean R, Shearer K, Schuckel J, Venditto I, Ralet MC, Henrissat B, Martens EC,

493   Mosimann SC, Abbott DW, Gilbert HJ. 2018. Dietary pectic glycans are degraded by coordinated

494   enzyme pathways in human colonic Bacteroides. Nat Microbiol 3:210-219.

495 33. Cameron EA, Kwiatkowski KJ, Lee BH, Hamaker BR, Koropatkin NM, Martens EC. 2014.

496   Multifunctional nutrient-binding proteins adapt human symbiotic bacteria for glycan

497   competition in the gut by separately promoting enhanced sensing and catalysis. MBio 5:e01441-

498   14.

499 34. Stentz R, Osborne S, Horn N, Li AW, Hautefort I, Bongaerts R, Rouyer M, Bailey P, Shears SB,

500   Hemmings AM, Brearley CA, Carding SR. 2014. A bacterial homolog of a eukaryotic inositol

501   phosphate signaling enzyme mediates cross-kingdom dialog in the mammalian gut. Cell Rep

502   6:646-56.

503 35. Hooda Y, Shin HE, Bateman TJ, Moraes TF. 2017. Neisserial surface lipoproteins: structure,

504   function and biogenesis. Pathogens and Disease 75:ftx010-ftx010.

505 36. Konovalova A, Perlman DH, Cowles CE, Silhavy TJ. 2014. Transmembrane domain of surface-

506   exposed outer membrane lipoprotein RcsF is threaded through the lumen of beta-barrel

507   proteins. Proc Natl Acad Sci U S A 111:E4350-8.

508 37. Konovalova A, Silhavy TJ. 2015. Outer membrane lipoprotein biogenesis: Lol is not the end.

509   Philosophical Transactions of the Royal Society B: Biological Sciences 370.

510 38. Hooda Y, Moraes TF. 2018. Translocation of lipoproteins to the surface of gram negative

511   bacteria. Curr Opin Struct Biol 51:73-79.

512 39. Hooda Y, Lai CC, Judd A, Buckwalter CM, Shin HE, Gray-Owen SD, Moraes TF. 2016. Slam is an

513   outer membrane protein that is required for the surface display of lipidated virulence factors in

514   Neisseria. Nat Microbiol 1:16009.

515   40.   Stahl M, Friis LM, Nothaft H, Liu X, Li J, Szymanski CM, Stintzi A. 2011. L-fucose utilization

516   provides Campylobacter jejuni with a competitive advantage. Proc Natl Acad Sci U S A 108:7194-

517   9.

518   41.   Ng KM, Ferreyra JA, Higginbottom SK, Lynch JB, Kashyap PC, Gopinath S, Naidu N, Choudhury B,

519   Weimer BC, Monack DM, Sonnenburg JL. 2013. Microbiota-liberated host sugars facilitate post-

520   antibiotic expansion of enteric pathogens. Nature 502:96-9.

521   42.   Bacic MK, Smith CJ. 2008. Laboratory maintenance and cultivation of bacteroides species. Curr

522   Protoc Microbiol Chapter 13:Unit 13C 1.

523   43.   C. JEFFREY SMITH MBR, AND MARIAN L. MCKEE. 1992. Heterologous Gene Expression in

524   *Bacteroides Fragilis*. Plasmid 27:141-154.

525   44.   Koropatkin NM, Martens EC, Gordon JI, Smith TJ. 2008. Starch catabolism by a prominent

526   human gut symbiont is directed by the recognition of amylose helices. Structure (London,

527   England : 1993) 16:1105-1115.

528   45.   Scott NE, Parker BL, Connolly AM, Paulech J, Edwards AV, Crossett B, Falconer L, Kolarich D,

529   Djordjevic SP, Hojrup P, Packer NH, Larsen MR, Cordwell SJ. 2011. Simultaneous Glycan-Peptide

530   Characterization Using Hydrophilic Interaction Chromatography and Parallel Fragmentation by

531   CID, Higher Energy Collisional Dissociation, and Electron Transfer Dissociation MS Applied to the

532   N-Linked Glycoproteome of Campylobacter jejuni. Mol Cell Proteomics 10:M000031MCP201.

533   46.   Ishihama Y, Rappsilber J, Mann M. 2006. Modular stop and go extraction tips with stacked disks

534   for parallel and multidimensional Peptide fractionation in proteomics. J Proteome Res 5:988-94.

535   47.   Rappsilber J, Mann M, Ishihama Y. 2007. Protocol for micro-purification, enrichment, pre-

536   fractionation and storage of peptides for proteomics using StageTips. Nat Protoc 2:1896-906.

537   48.   Cox J, Mann M. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-

538   range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26:1367-72.

539   49.   Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. 2014. Accurate proteome-wide label-free

540   quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ.

541   Mol Cell Proteomics 13:2513-26.

542   50.   Tyanova S, Temu T, Carlson A, Sinitcyn P, Mann M, Cox J. 2015. Visualization of LC-MS/MS

543   proteomics data in MaxQuant. Proteomics 15:1453-6.

544   51.   Tsirigos KD, Peters C, Shu N, Kall L, Elofsson A. 2015. The TOPCONS web server for consensus

545   prediction of membrane protein topology and signal peptides. Nucleic Acids Res 43:W401-7.

546   52.   Juncker AS, Willenbrock H, Von Heijne G, Brunak S, Nielsen H, Krogh A. 2003. Prediction of

547   lipoprotein signal peptides in Gram-negative bacteria. Protein Sci 12:1652-62.

548   53.   Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator.

549   Genome Res 14:1188-90.

550

551   <u>Figure Legends</u>

552   **Figure 1. *B. theta* produces outer membrane vesicles (OMV).** Transmission electron

553   microscopy of a single *B. theta* cell together with OMVs in the extracellular milieu. *B.*

554   *theta* cells were swabbed from a solid media plate, suspended in PBS and processed

555   for TEM. Images were acquired at a direct magnification of 15000x. Scale bar

556   represents 500 nm. Credit to Wandy Beatty, Molecular Microbiology Imaging Facility,

557   WUSTL.

558

559   **Figure 2. *B. theta* OMV protein content is different from OM.** 300 μg of each

560   preparation and biological replicates of purified OM and OMVp proteins were digested

561   with trypsin. The resulting peptides were enriched, then analyzed via liquid

562 chromatography coupled to tandem mass spectrometry (LC-MS/MS) as explained in

563 Materials and Methods, followed by protein identification with Mascot search engine

564 using the Uniprot database. Volcano plot shows OM and OMVp protein populations.

565 Red labels indicate the proteins with the highest OMVp enrichment in comparison to

566 OM.

567

568 **Figure 3. Validation of OMV- and OM-enriched proteins.** Candidate ORFs that

569 encode proteins identified in our MS analysis as OMV- or OM-enriched were cloned into

570 pFD340 with a C-terminal 6xHis tag. Constructs were introduced into *B. theta* by

571 conjugation, generated strains were grown in TYG media and fractions were prepared.

572 Ten micrograms of each fraction were run on 12% SDS-PAGE and analyzed by Western

573 Blot using anti-His polyclonal antibodies. (a) OMV-enriched proteins; (b) OM-enriched

574 proteins. The isoelectric point as well as the residues following lipoprotein attachment

575 Cysteine (for lipoproteins) are indicated below the protein name.

576

577 **Figure 4. OMV-enriched proteins show a conserved N-terminal LES motif.** The top

578 91 OMV-enriched proteins (a) and top 24 OM-enriched proteins (b) were aligned using

579 the lipoprotein attachment cysteine as the +1 position (not shown in the logo), plus the 9

580 C-term contiguous residues. Lipoprotein export sequence (LES) consensus was

581 generated using WebLogo (https://weblogo.berkeley.edu/logo.cgi)(53).

582

583 **Figure 5. Sus lipoproteins are OMV-enriched while SusC is OM-enriched.** *B. theta*

584 cells containing a tagged genomic copy of *susC* with 3xFlag were grown in TYM media

585   (TYG recipe with 0.5% maltose instead of glucose for induction of the *sus* operon).

586   Fractions were prepared and analyzed by 12% SDS-PAGE and Western Blot using

587   specific anti-Sus, anti-Flag and anti-RnaPol antibodies.

588

589   **Figure 6. SusG LES is required for efficient packing into OMVs.** (a) SusG LES point

590   mutants were generated on pFD340/*susG*-6xHis using Quick-site directed mutagenesis.

591   (b) Constructs were introduced in the Δ*susG* background, and generated strains were

592   grown in minimal media with glucose as only carbon source. Fractions were prepared

593   and analyzed by SDS-PAGE and Western Blot using anti-SusG antibodies. (c) OMV/OM

594   ratios were calculated using the fluorescence signal values for each protein and plotted

595   as a percentage relative to the WT OMV/OM ratio (100%). Statistical significance was

596   determined by unpaired t-test of each LES variant in comparison to wild-type LES

597   strain. * p-values <0.03. Experiment is representative of three biological replicates,

598   shown are mean values with SD for two technical replicates.

599

600   **Figure 7. The LES is required for SusG surface exposure.** WT SusG and the LES

601   mutants were assessed in their surface exposure by proteinase K assays using whole

602   cells (a) and OMVs (b). Proteinase K was added and incubated at 37°C; at different

603   times, aliquots were TCA-precipitated and analyzed by SDS-PAGE and Western Blot

604   using anti-SusG antibodies. We used a 3xFlag-tagged BT_0766, a periplasmic soluble

605   protein found in OMVs, as an outer membrane and OMV integrity control.

606

607 **Figure 8. OMVs displaying WT SusG can rescue a Δ*susG* growth phenotype on**

608 **starch as carbon source.** WT and Δ*susG* strains were grown overnight in TYG media.

609 Cultures were washed with minimal media (MM) without any carbon source and

610 normalized by OD. Minimal media with 0.5% starch as only carbon source was

611 inoculated with WT or Δ*susG* strains to final OD_{600} = 0.05. OMVs purified from the

612 Δ*susG* strain containing different pFD340 derivatives, were added to the Δ*susG* cultures

613 in minimal media with starch (1 µg/ml final of OMVs). Aliquots were taken at different

614 times and OD_{600} was measured to determine growth. Statistical significance when

615 comparing two growth curves was determined by performing one pair of unpaired t-test

616 analysis per each time point. * p-values <0.01. Experiment is representative of three

617 biological replicates, shown are mean values with SD for three technical replicates.

618

619 **Figure 9. A model for SusG LES-mediated surface exposure and packing into**

620 **OMV.** The SusG lipoprotein is probably transported to the OM by a machinery

621 homologous to the LOL system. Presence of a LES sequence mediates surface

622 exposure of SusG and incorporation into OMVs. The LES could have a dual role where

623 it is required for both surface exposure and OMV packing, or it could be involved only in

624 surface exposure and additional information on each lipoprotein is recognized for

625 packing into OMVs. Surface-exposed SusG in OMVs as well as in the OM hydrolyzes

626 starch molecules into oligosaccharides that can be imported by TonB-dependent

627 receptors by the OMV-producing cell, as well as other commensal and pathogen

628 organisms.

629

630   <u>Supplementary Figure Legends</u>

631   **Figure S1. Protocol for purification and analysis of *B. theta* OMV and membrane**

632   **fractions.** We performed membrane and OMV fractionation as indicated using cultures

633   grown overnight in TYG media (a). Fractions were incubated with 1% Sarkosyl for 1 h at

634   RT and ultacentrifuged for 2 h at RT. Fractions were resuspended in 50 mM HEPES pH

635   7.4. Aliquots from each fraction were analyzed by SDS-PAGE and Coomassie Blue

636   staining (b). Fractions labelled as OM (green) and OMV-P (blue) are the most relevant

637   fractions for comparison by MS.

638

639   **Figure S2. Total protein profiles of membrane and OMV fractions from OMV-**

640   **enriched protein validations.**

641   We performed membrane and OMV fractionation as indicated using cultures grown

642   overnight in TYG media. Duplicate SDS-PAGE gels loaded with 10 μg of each protein

643   fraction were analyzed in their total protein profiles by Coomassie Blue stain. Provided

644   are representative gels used for two OMV-enriched proteins BT_3960 and BT_1488

645

646   **Figure S3. Total protein profiles of membrane and OMV fractions from OM-**

647   **enriched protein validations.**

648   We performed membrane and OMV fractionation as indicated using cultures grown

649   overnight in TYG media. Duplicate SDS-PAGE gels loaded with 10 μg of each protein

650   fraction were analyzed in their total protein profiles by Coomassie Blue stain. Provided

651   are representative gels used for two OM-enriched proteins BT_0418 and BT_2844.

652

653 **Figure S4. Total protein stain analysis of fractions for OMV/OM ratio**

654 **determination.**

655 We performed membrane and OMV fractionation as indicated using cultures grown

656 overnight in minimal media with glucose. SDS-PAGE gels loaded with 10 µg of each

657 protein fraction were transferred onto nitrocellulose membranes and subjected to

658 REVERT total protein stain. Membranes were imaged immediately and Western Blot

659 analysis was carried out using anti-SusG antibodies as described in Materials and

660 Methods. Gel shown is representative of all the variants of SusG LES that were

661 assayed for OMV/OM ratio determinations.

662

663 **Table S1. Complete MS dataset.**

664

665 **Table S2. OMV- and OM-enriched proteins.**

666

667 **Table S3. Oligonucleotides, strains and plasmids.**

668

669 **Table 1. Top 20 most OMVp-enriched proteins, including protein name and**

670 **enrichment value.** LFQ corresponds to label-free quantification values.

Gene name	Protein name	log2 LFQ (OMVp/OM)
BT_3382	Putative GH10/39 family protein	13.70
BT_1488	Putative cysteine protease	11.36
BT_3960	Putative calpain-like protease	11.07

BT_3381	Putative GH10/39 family protein	10.48
BT_3383	Putative GH10/39 family protein	10.29
BT_3148	Putative adhesin	10.12
BT_2317	Putative fimbrial protein	9.61
BT_p548229	Putative fimbrial protein	9.21
BT_3147	Putative adhesin	9.18
BT_1792	Putative cell surface protein	9.05
BT_3740	Putative cell adhesion protein	9.02
BT_3630	Putative fimbrial protein	8.78
BT_2479	Putative imelysin-like protein	8.36
BT_1826	Putative GH31-like family protein	8.29
BT_3067	Uncharacterized protein	8.09
BT_3066	Uncharacterized protein	7.84
BT_4606	Uncharacterized protein	7.83
BT_2658	Putative cell adhesion protein	7.77
BT_1063	Putative cell adhesion protein	7.72
BT_0166	Putative thioredoxin	7.69
BT_2989	Uncharacterized protein	7.63

671
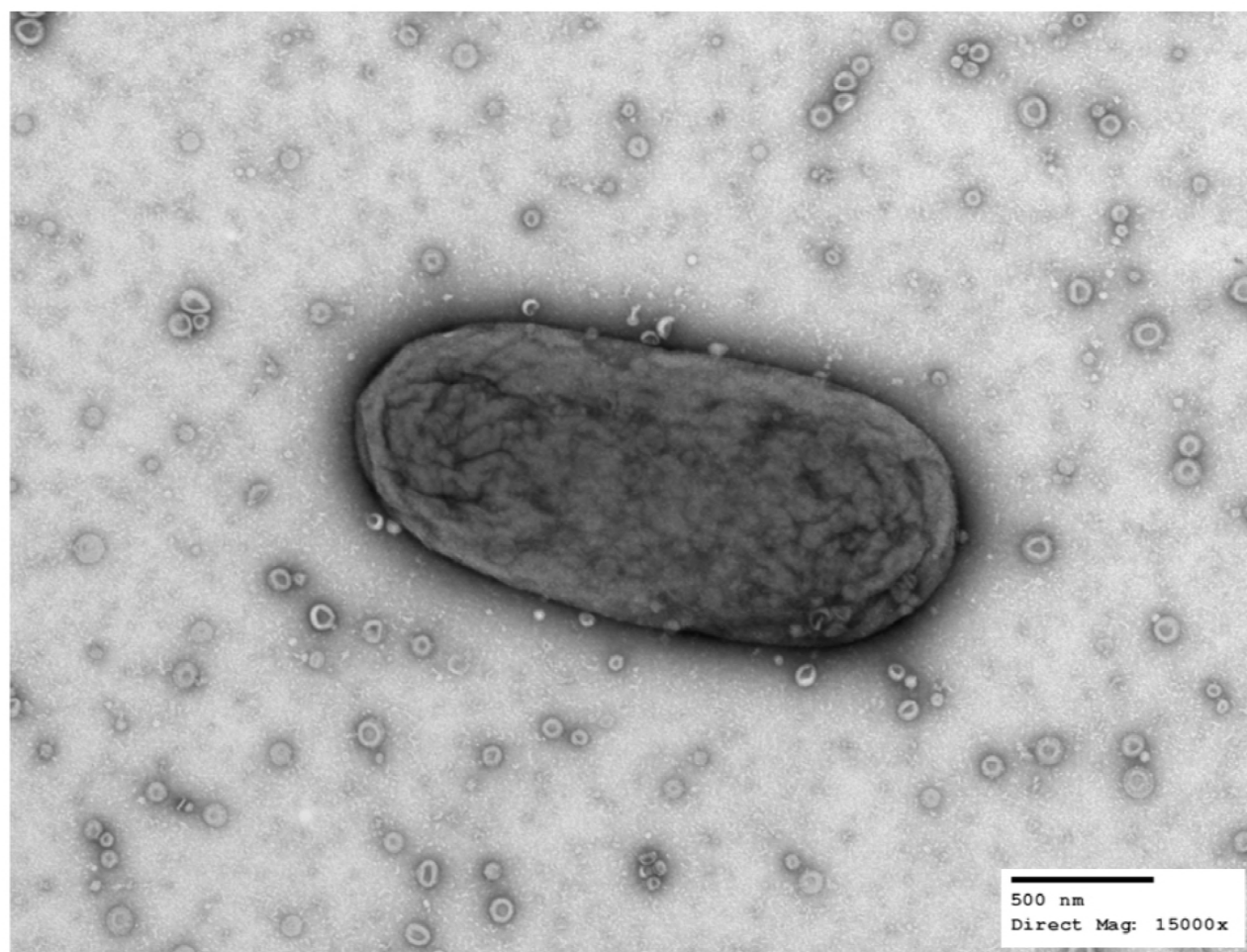
Figure 1. *B. theta* produces outer membrane vesicles (OMV). Transmission electron microscopy of a single *B. theta* cell together with OMVs in the extracellular milieu. *B. theta* cells were swabbed from a solid media plate, suspended in PBS and processed for TEM. Images were acquired at a direct magnification of 15000x. Scale bar represents 500 nm. Credit to Wandy Beatty, Molecular Microbiology Imaging Facility, WUSTL.
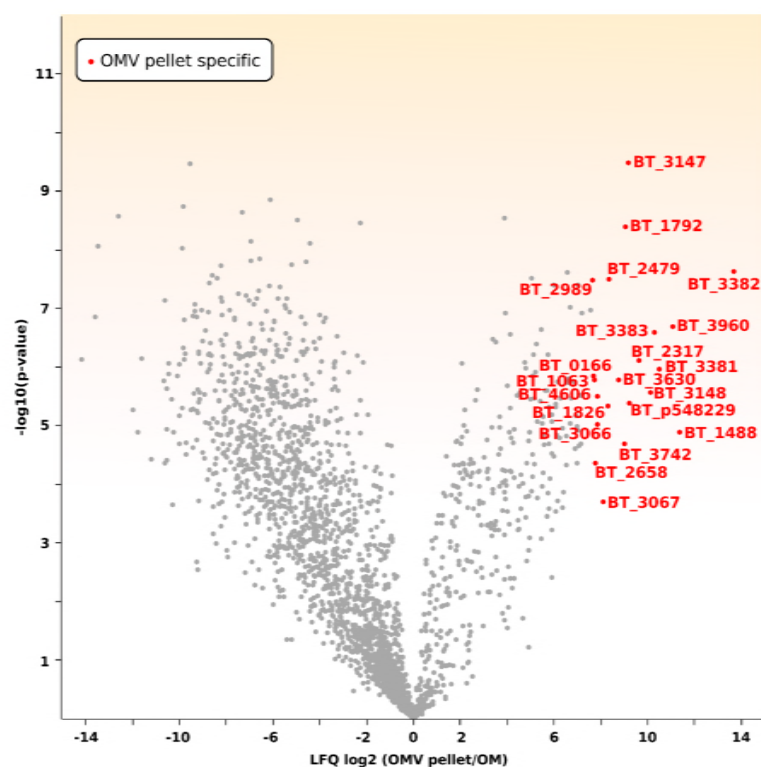
Figure 2. *B. theta* OMV protein content is different from OM. 300 μg of each preparation and biological replicates of purified OM and OMVp proteins were digested with trypsin. The resulting peptides were enriched, then analyzed via liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) as explained in Materials and Methods, followed by protein identification with Mascot search engine using the Uniprot database. Volcano plot shows OM and OMVp protein populations. Red labels indicate the proteins with the highest OMVp enrichment in comparison to OM.
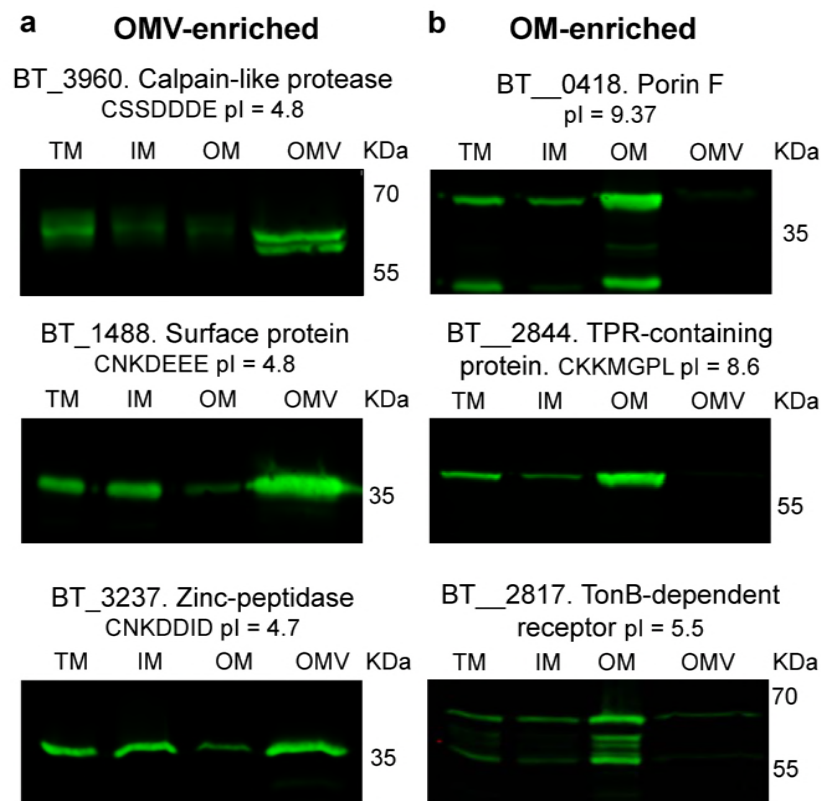
Figure 3. Validation of OMV- and OM-enriched proteins. Candidate ORFs that encode proteins identified in our MS analysis as OMV- or OM-enriched were cloned into pFD340 with a C-terminal 6xHis tag. Constructs were introduced into *B. theta* by conjugation, generated strains were grown in TYG media and fractions were prepared. Ten micrograms of each fraction were run on 12% SDS-PAGE and analyzed by Western Blot using anti-His polyclonal antibodies. (a) OMV-enriched proteins; (b) OM-enriched proteins. The isoelectric point as well as the residues following lipoprotein attachment Cysteine (for lipoproteins) are indicated below the protein name.
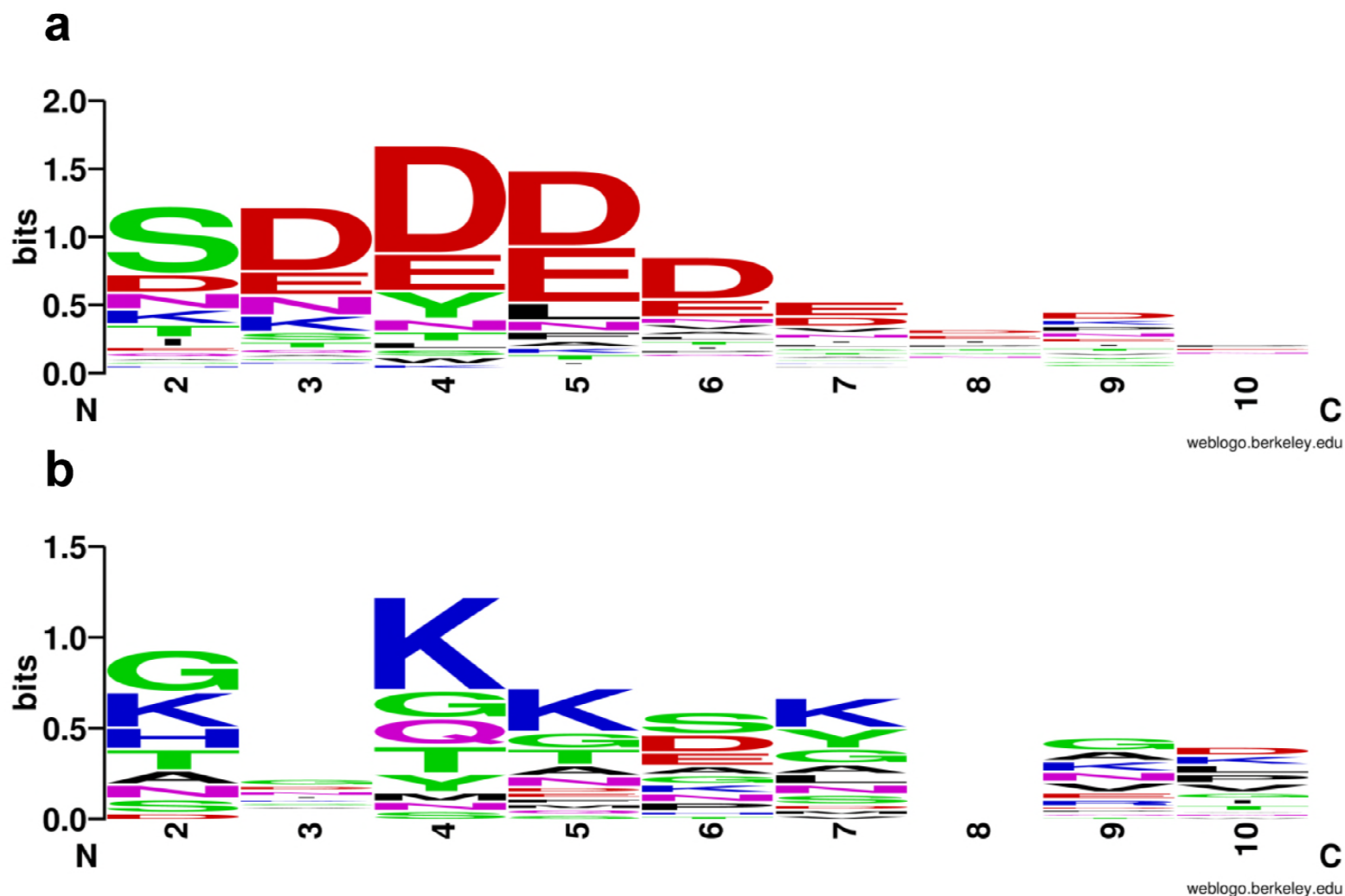
Figure 4. OMV-enriched proteins show a conserved N-terminal LES motif. The top 91 OMV-enriched proteins (a) and top 24 OM-enriched proteins (b) were aligned using the lipoprotein attachment cysteine as the +1 position (not shown in the logo), plus the 9 C-term contiguous residues. Lipoprotein export sequence (LES) consensus was generated using WebLogo (https://weblogo.berkeley.edu/logo.cgi).
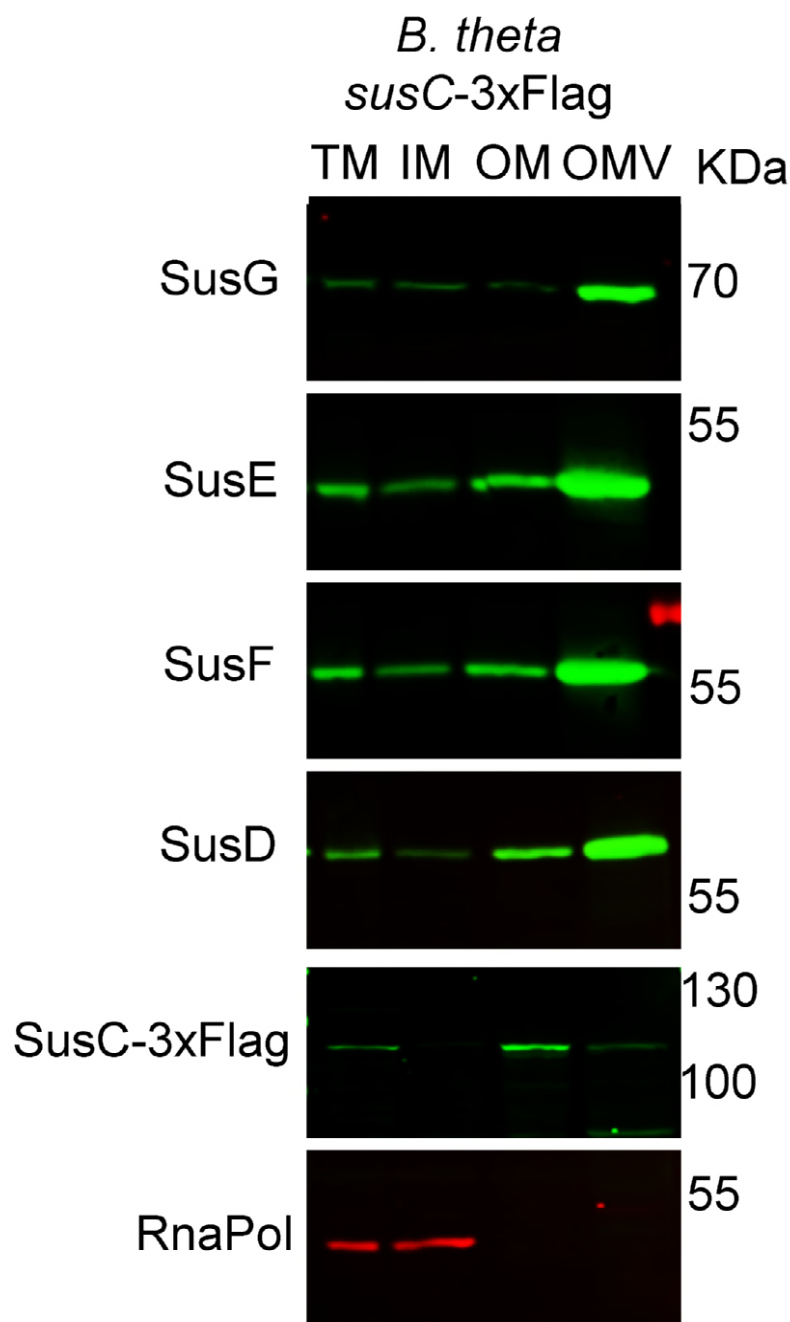
Figure 5. Sus lipoproteins are OMV-enriched while SusC is OM-enriched. *B. theta* cells containing a tagged genomic copy of *susC* with 3xFlag were grown in TYM media (TYG recipe with 0.5% maltose instead of glucose for induction of the *sus* operon). Fractions were prepared and analyzed by 12% SDS-PAGE and Western Blot using specific anti-Sus, anti-Flag and anti-RnaPol antibodies.
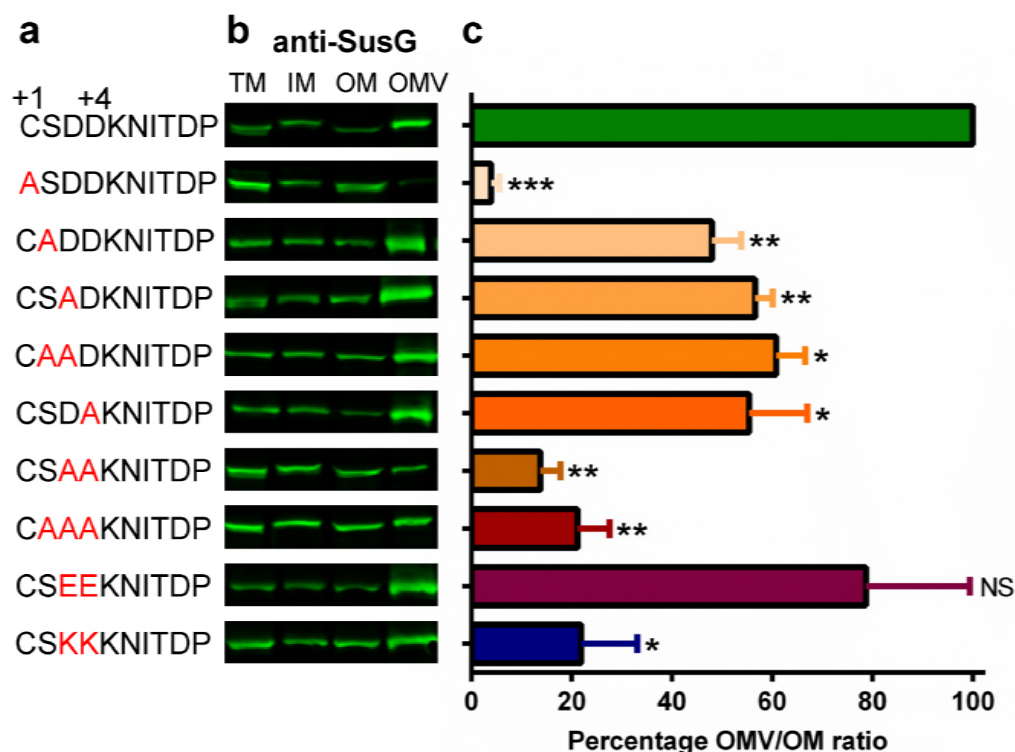
Figure 6. SusG LES is required for efficient packing into OMVs. (a) SusG LES point mutants were generated on pFD340/*susG*-6xHis using Quick-site directed mutagenesis. (b) Constructs were introduced in the Δ*susG* background, and generated strains were grown in minimal media with glucose as only carbon source. Fractions were prepared and analyzed by SDS-PAGE and Western Blot using anti-SusG antibodies. (c) OMV/OM ratios were calculated using the fluorescence signal values for each protein and plotted as a percentage relative to the WT OMV/OM ratio (100%). Statistical significance was determined by unpaired t-test of each LES variant in comparison to wild-type LES strain. * p-values <0.03 .Experiment is representative of three biological replicates, shown are mean values with SD for two technical replicates.
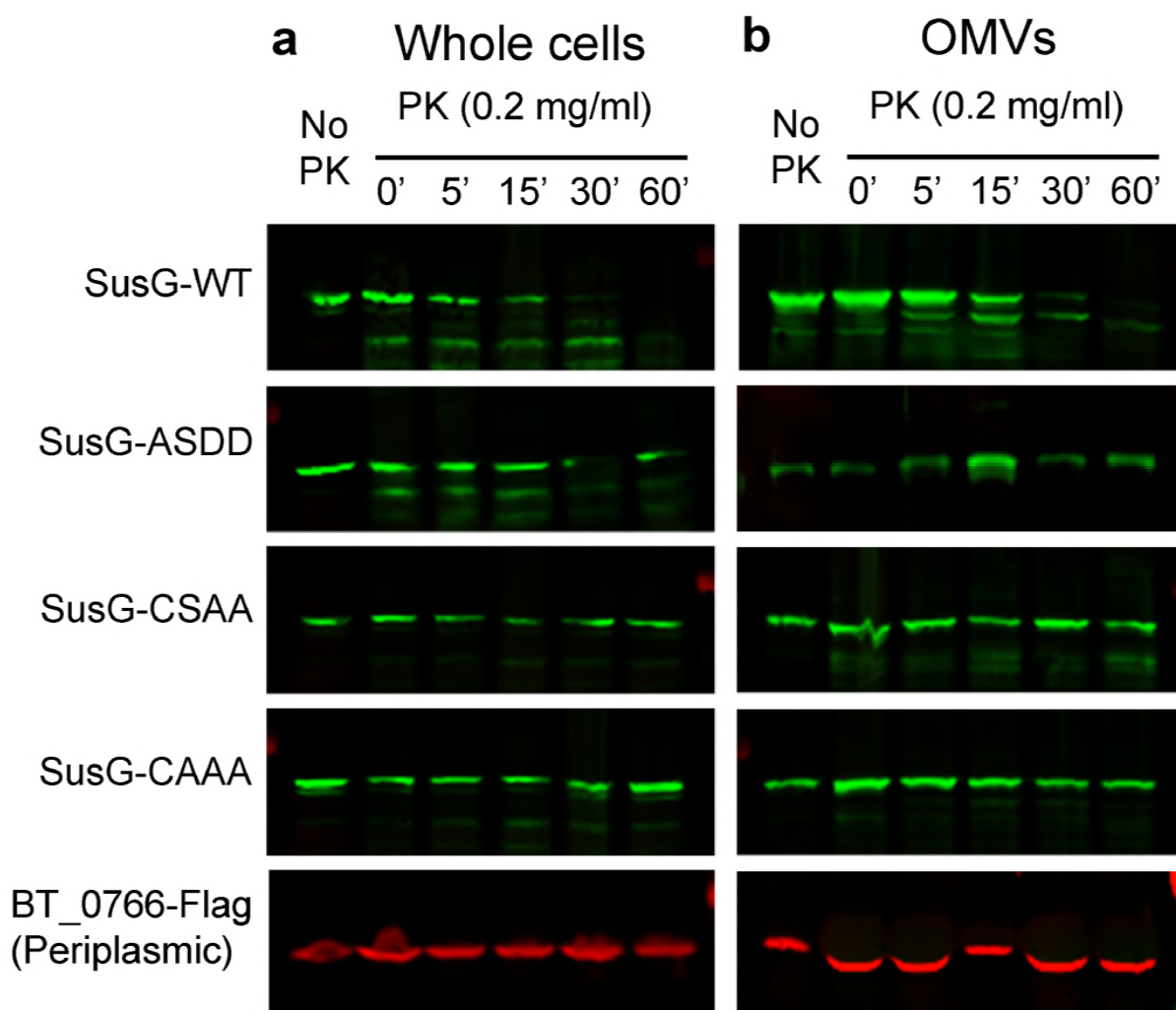
Figure 7. The LES is required for SusG surface exposure. WT SusG and the LES mutants were assessed in their surface exposure by proteinase K assays using whole cells (a) and OMVs (b). Proteinase K was added and incubated at 37°C; at different times, aliquots were TCA-precipitated and analyzed by SDS-PAGE and Western Blot using anti-SusG antibodies. We used a 3xFlag-tagged BT_0766, a periplasmic soluble protein found in OMVs, as an outer membrane integrity control.
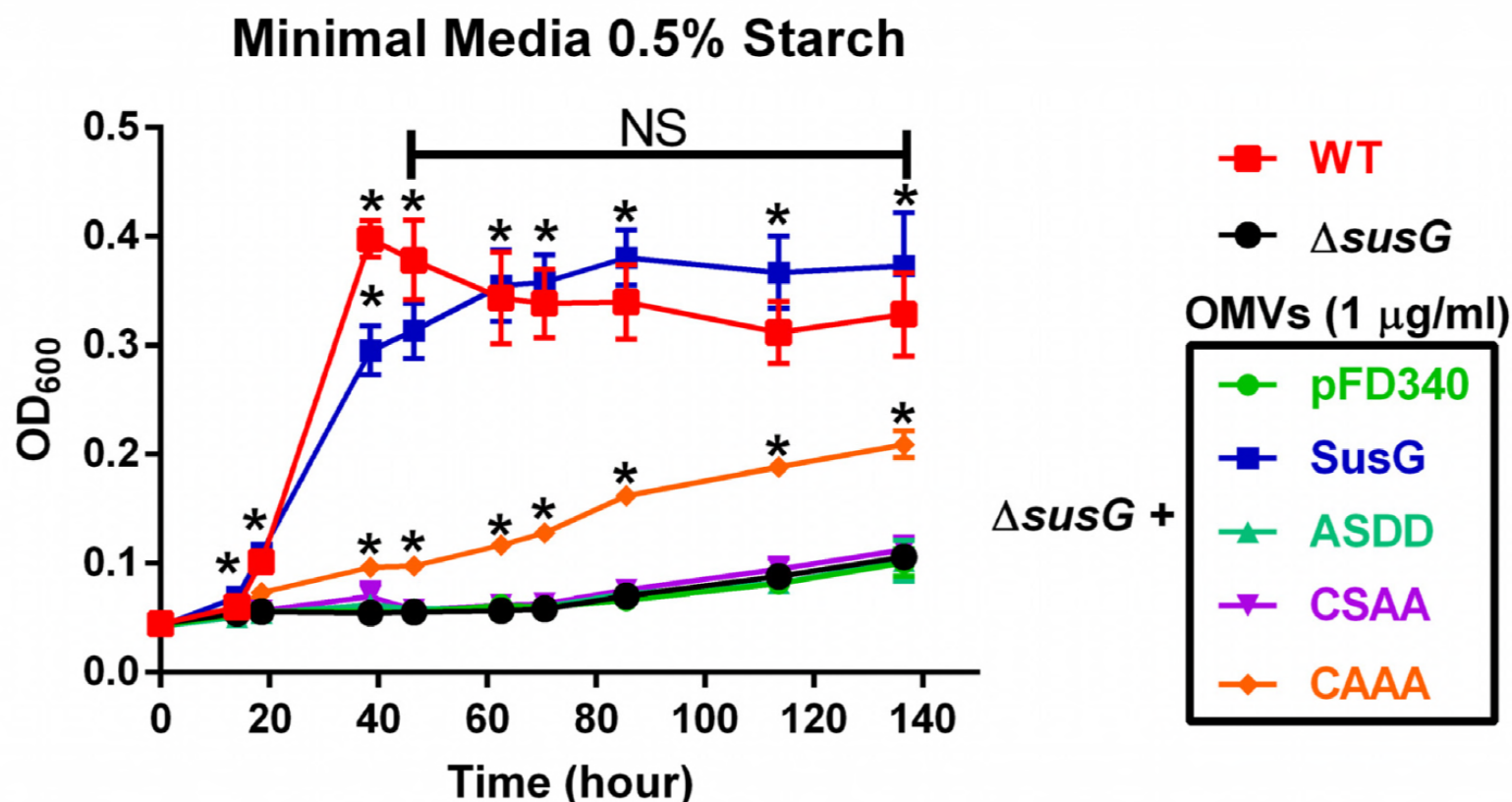
Figure 8. OMVs displaying WT SusG can rescue a ΔsusG growth phenotype on starch as carbon source. WT and ΔsusG strains were grown overnight in TYG media. Cultures were washed with minimal media (MM) without any carbon source and normalized by OD. Minimal media with 0.5% starch as only carbon source was inoculated with WT or ΔsusG strains to final OD_{600} = 0.05. OMVs purified from the ΔsusG strain containing different pFD340 derivatives, were added to the ΔsusG cultures in minimal media with starch (1 µg/ml final of OMVs). Aliquots were taken at different times and OD_{600} was measured to determine growth. Statistical significance when comparing two growth curves was determined by performing one pair of unpaired t-test analysis per each time point. * p-values <0.01. Experiment is representative of three biological replicates, shown are mean values with SD for three technical replicates.
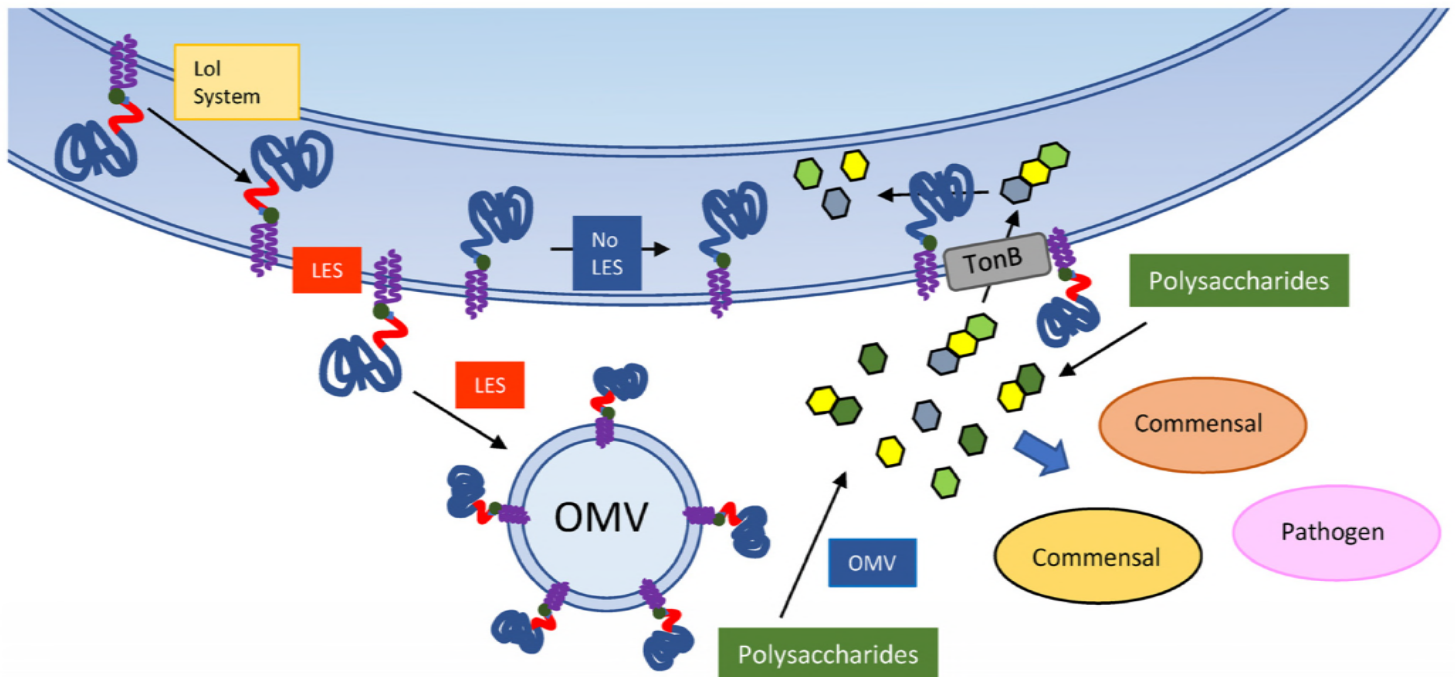
Figure 9. A model for SusG LES-mediated surface exposure and packing into OMV. The SusG lipoprotein is transported to the OM by a putative machinery orthologous to the LOL system. Presence of a LES sequence mediates surface exposure of SusG and incorporation into OMVs. The LES might have a dual role where it is required for both surface exposure and OMV packing, or it could be involved only in surface exposure and additional information on each lipoprotein could be recognized for packing into OMVs. Surface-exposed molecules of SusG in OMVs as well as at the OM hydrolyze starch molecules into oligosaccharides that can be imported by TonB-dependent receptors on the OMV-producing cell, as well as by other commensal and pathogen organisms.