Classification: Biological Sciences, Microbiology

Title: "Cell surface associated LapA of *Pseudomonas fluorescens* is anchored inside its Type-1 Secretion TolC-like Pore"

Authors: T. Jarrod Smith[1], Holger Sondermann[2] and George A. O'Toole[1]

Author Affiliation: [1]Department of Microbiology and Immunology, Geisel School of Medicine at Dartmouth, Hanover, NH 03755, USA

[2]Department of Molecular Medicine, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853 USA

Corresponding Author: name: George A. O'Toole; address: Geisel School of Medicine at Dartmouth, Remsen 202, Hanover, NH 03755, USA; phone: 603 650 1247; email: George.A.O'Toole.Jr@dartmouth.edu.

Keywords: biofilms, bacterial secretion, adhesins, *Pseudomonas fluorescens*

Abstract.

The type-1 secretion system (T1SS) of gram-negative bacteria enables a one-step translocation strategy known to move functionally diverse proteins from the cytoplasm into the extracellular environment without a periplasmic intermediate. LapA of *Pseudomonas fluorescens* Pf0-1 is a giant type-1 secreted (T1S) adhesin that facilitates biofilm formation only when displayed at the cell surface. A LapA-targeting periplasmic protease, LapG, connects intracellular cyclic diguanylate (c-di-GMP) levels with cell surface-associated LapA by cleaving and absolving LapA from the cell surface under conditions unsuitable for biofilm formation. Here, we demonstrate that LapA contains a novel N-terminal element, called the retention module (RM), which prohibits classical one-step T1S of LapA. We provide evidence that the RM of LapA tethers LapA at the cell surface through its outer membrane TolC-like pore, LapE, where LapA is accessible to the periplasmic protease LapG. We also demonstrate that this unusual retention strategy is likely conserved among LapA-like proteins and represents a new subclass of T1SS ABC transporters exclusively involved in transporting LapA-like adhesins.

Significance Statement.

Bacteria have evolved multiple secretion strategies to interact with their environment. For many bacteria, the secretion of cell surface associated adhesins is often key for initiating contact with a preferred substrate to facilitate biofilm formation. Our work demonstrates that *P. fluorescens* uses a previously unrecognized secretion strategy to retain the giant adhesin LapA at its cell surface. Further, we identify likely LapA-like adhesins in various pathogenic and commensal Proteobacteria and provide phylogenetic evidence these adhesins are secreted by a new subclass of T1SS ABC transporters.

47 **Text.**

48 The biofilm lifestyle is profoundly consequential to human health and industry, for
49 better and worse (1, 2). Although bacteria initiate surface contact and biofilm formation
50 through a variety of strategies, many microbes need cell surface-associated protein
51 adhesins to bind a surface. The family of giant (>200 kDa) T1S repeats-in-toxin (RTX)-
52 containing adhesins is known to be critical for biofilm formation or surface binding by a
53 variety of organisms, including *Pseudomonas, Bordetella*, *Legionella*, *Vibrio*, *Shewanella* and
54 *Marinomonas* (3–8).

55 Mechanistic studies from our lab on the giant RTX adhesin, LapA, of *Pseudomonas*
56 *fluorescens* Pf0-1 identified a novel regulatory node, LapDG, that links cell surface
57 associated LapA levels to intracellular c-di-GMP levels. Here, the activity of the LapA-
58 targeting periplasmic protease, LapG, is inhibited when its effector, LapD, a transmembrane
59 protein, is bound to cellular c-di-GMP. Conversely, when c-di-GMP levels decrease in the
60 cell, LapD releases LapG. Free LapG in turn cleaves LapA from the cell surface, releasing the
61 adhesin into the supernatant where it is inoperative for biofilm formation (9, 10).
62 Homologs of the *lapDG* genes are found through the Proteobacteria, suggesting that this
63 poorly understood surface display strategy may be quite common (11, 12).

64 Like other T1S proteins, such as HlyB of *E. coli* and CyaA of *B. bronchiseptica*, LapA's C-
65 terminal secretion signal and cognate T1S machinery (LapEBC) are required for LapA
66 secretion and thus biofilm formation (6, 13). T1SS is considered a one-step secretion
67 strategy that lacks any periplasmic intermediate (14), and it is currently unclear how LapA
68 localizes to the cell surface with an N-terminal dialanine cleavage motif that is accessible to
69 the periplasmic protease, LapG. Here, we demonstrate that LapA is not secreted in a

canonical one-step T1SS fashion, but rather tethers to the cell surface through its T1SS apparatus. A cleavable retention module at the N-terminus of LapA prohibits complete secretion until cleaved by the LapG protease. We also provide evidence that this previously unappreciated retention strategy is broadly conserved in Proteobacteria and represents a third, distinct subclass of T1S systems.

Results

Bioinformatic Identification of LapG Substrates. Despite the importance of LapA and related proteins as key biofilm adhesins (15), it is difficult to identify LapG substrates due to their relatively low sequence similarly. Additionally, ORF analysis programs often overlook or misannotate these large and complex adhesins (5, 16, 17). To overcome the first limitation, we took advantage of the observation that the two proteins that control LapA localization, the c-di-GMP-receptor LapD and the LapD-regulated protease LapG, show high sequence similarity and functional conservation between microbes (3, 12, 18) and can be identified by their respective domain architectures (LapG, pfam06035; LapD, pfam16448). We utilized the NCBI conserved domain database (CDD) and genome database to identify bacterial species encoding *lapDG* homologs; ~1300 such *lapGD*-encoding species spanning 120 genera were identified. Each annotated genome was investigated for proteins containing hallmarks of LapA: an N-terminal LapG cleavage site and C-terminal RTX motifs. To accomplish this task, we developed an algorithm to recognize large proteins (>1000 aa) with RTX-motifs (Dx[L/I]x(4)GxDx[L/I]xGGx(3)D) and a canonical N-terminal dialanine LapG cleavage motif ([T/A/P]AA[G/V]).

Although our approach is constrained to properly annotated LapA-like ORFs, we still identified over 500 putative LapG substrates in ~50 genera throughout Proteobacteria,

including *Legionella* and *Vibrio* species (Table S1). Importantly, characterized LapG substrates LapA and BrtA (3, 19) were identified. Interestingly, some species likely encode multiple LapG substrates, including *P. fluorescens* Pf0-1 (LapA and Pfl01_1463) and *V. cholerae* O395 (FrhA and VC0395_0388).

LapG Substrates Predicted *in silico* Are Processed *in vitro*. LapG homologs cleave a variety of LapA-like N-termini from unrelated species *in vitro* and *in vivo* (3, 18). To help validate the utility of our algorithm and confirm predicted LapG substrates, we cloned and expressed the N-terminal elements (~250-350 aa) of the putative LapG-proteolyzed adhesins Pfl01_1463 from *P. fluorescens* and VC0395_0388 from *Vibrio cholerae* (Fig 1A). Cell lysates of *E. coli* expressing C-terminally 6HIS-tagged, N-terminal fragments of Pfl01_1463 (M1-240S, 117TAAG120) or VC0395_0388 (M1-363G, 127AAAG130) were mixed with a lysate made from *E. coli* expressing *P. fluorescens* Pf0-1 LapG from a plasmid or the empty vector control. The LapG-dependent cleavage product was tracked via Western blot, with the N-terminus of LapA (M1-235V, 107TAAG110) and an uncleavable variant (LapA^{TRRG}, AA108-109RR) serving as positive and negative controls, respectively (9). *P. fluorescens* Pf0-1 LapG cleaves both predicted substrates (Fig 1 B & C). Unlike FrhA (4), a role for VC0395_0388 in infection and/or biofilm formation is currently unknown; however, these data implicate VC0395_0388 as a cell-surface associated, c-di-GMP-regulated biofilm-promoting adhesin.

Conversely, *P. fluorescens* Pf0-1 LapG is unable to cleave the N-terminus of PA1874(M1-251T, 137AAA*I*G141) (Fig 1D), a large 238 kDa putative outer membrane adhesin encoded by *Pseudomonas aeruginosa* PAO1. Although PA1874 resembles LapA, it was not detected by our algorithm because it contains a degenerate LapG cleavage motif (137AAA*I*G141).

116 Instead, we manually identified PA1874 by its proximity to *lapEBC* homologs (20).
117 Together, these data support the predictive power of our algorithm for detecting LapA-like
118 LapG substrates *in silico*, and demonstrate the breadth of this c-di-GMP-regulated biofilm
119 strategy in Proteobacteria.
120 ***P. fluorescens* LapA Homologs Contain a Conserved N-terminal Domain.** Previous work
121 from our lab implicated the N-terminus of LapA in retaining the giant adhesin at the cell
122 surface (13, 19). To gain insight into LapA's retention mechanism, we aligned putative
123 LapA-like proteins from closely related *P. fluorescens* strains detected *in silico* and color-
124 filled residues according to the default Clustal X color parameters in Jalview to highlight
125 highly similar regions (Fig S1A, Table S2). Our analysis revealed these adhesins contain
126 two highly similar regions: a C-terminal region that corresponds to LapA's T1S signal
127 required for secretion (Fig S1A, blue box) and a N-terminal region that extends ~20
128 residues beyond the LapG dialanine cleavage motif (Fig S1A, red box and Fig 2A). Given the
129 C- to N-terminal secretion directionality of T1S substrates (21), we speculated that the N-
130 terminus may be involved in retaining these giant adhesins at the cell surface. Consistent
131 with this idea, analysis of BrtA and RtxA homologues from *B. bronchiseptica* and *L.*
132 *pneumophila* strains, respectively, revealed similar patterns of high identity N- and C-
133 termini (Fig S1B,C).
134 **The N-terminus of LapA is Required for Surface Retention.** Closer examination of the N-
135 terminus of *P. fluorescens* LapA homologs indicates their sequence similarity breaks down
136 shortly after a poly-glycine region (Figure 2A, Poly-G). This poly-glycine region is present
137 in many of the LapA-like proteins predicted by our algorithm (Fig S2A). Because poly-
138 glycine regions often serve as unstructured domain linkers, we hypothesized the N-

terminal region of LapA encompassing up to this linker may function as a retention module (M1-125S).

To test this idea, we made targeted deletions in this putative N-terminal domain. Analysis of the primary sequences and predicted secondary structures suggests the N-termini of the LapA-like proteins identified by our algorithm share little sequence identity, but may adopt similar secondary structures (Fig S2A-B). Given that glycine residues are known to punctuate secondary structures, we used a "gly-to-gly" targeted truncation strategy to disrupt secondary structures within the N-terminus of LapA, using the alignment from Fig S2A as guide.

LapA is typically undetectable in the supernatant and enriched at the cell surface of the *lapG* mutant (Fig 2B and 2C, WT vs *lapG*). Therefore, RM mutants were engineered into the hyper-biofilm forming *lapG* gene deletion strain using unmarked allelic replacement, allowing us to decouple retention defects from LapG-mediated proteolysis. Biofilm formation and LapA localization were assayed by comparing each retention mutant to the parental *lapG* mutant, and Western blotting for LapA in the whole cell (WC), supernatant (S) and cell surface (CS) fractions.

Consistent with our hypothesis, a LapA truncation mutant lacking residues D31-95A (ΔD31-95A) is unable to form a biofilm (Fig 2D), and this LapA variant is found in the supernatant (Fig 2B) but not retained at the cell surface (Fig 2C). While the ΔV23-95A and ΔD31-95A phenotypes were indistinguishable, most mutants engineered for this study showed only a slightly reduced biofilm phenotype (Fig S2B-C). For those mutants with a biofilm defect (i.e., ΔS02-107A), Western analysis showed that the mutant protein was unstable. Together, these data indicate the N-terminus of LapA functions as a retention

module, contributing to the localization of LapA to the cell surface, and also suggests that this region plays a critical, but unclear, role in LapA stability.

Cell Surface Associated LapA Engages the LapEBC T1SS Machinery Through its N-Terminal Domain. Next, we investigated how the RM tethers LapA to the cell surface. LapA is found in the outer membrane fraction (24); however, LapA's RM lacks elements previously shown to be involved in forming outer membrane pores and translocation structures in target cell membranes (15, 22, 23). Thus, we considered the possibility that LapA is anchored to the cell surface by remaining in the T1SS apparatus as a secretion intermediate, using its N-terminal RM to prevent complete secretion of the adhesion. Although unheard of for T1S proteins, we previously demonstrated the LapG substrate from *Pseudomonas aeruginosa*, CdrA, a two partner secreted protein, uses a "cysteine-hook" formed by an intramolecular disulfide bond to anchor itself to the cell surface through its outer membrane pore, CdrB (12, 25). Additionally, recent structural analysis and modeling performed on MpIBP, a large LapA-like adhesion of *M. primoryensis*, suggests the RM of MpIBP, which shares secondary structure features with LapA, may form a plug that prohibits complete secretion of the adhesin through its outer membrane T1S-associated TolC-like protein (26). However, to date, there has been no experimental data supporting the retention model we explore here.

To determine if LapA is secreted via the classical one-step T1SS model or retained within its translocation machinery, we conducted a secretion competition experiment to compare secretion of the C-terminal secretion domain of LapA tagged with a HA-epitope tag for Western blotting (HA-C235; Fig 3A, far right). This tool allows us to discern if cell surface-associated LapA impacts the availability of LapEBC T1SS to secrete the HA-C235

peptide. We examined secretion of the HA-C235 protein in strains where LapA is locked at the cell surface (a Δ*lapG* mutant), or alternatively, in strains where LapA is continuously secreted into the supernatant (Δ*lapD* or *lapG lapAΔD31-95A* mutants). In the Δ*lapD* mutant, constitutive LapG activity removes LapA from the cell surface, while the *lapG lapAΔD31-95A* strain expresses a variant of LapA that lacks the complete retention signal (see Fig 2).

Western blot analysis of the supernatant fraction indicates like LapA, HA-C235 peptide secretion is LapEBC–dependent. Deletion of *lapE* results in loss of biofilm formation, as reported previously (6) and eliminates HA-C235 secretion (Fig 3B). HA-C235 secretion is restored when HA-C235 and *lapE* are introduced into the Δ*lapE* mutant and expressed as a transcriptional fusion from a plasmid (Fig 3B, far right).

In the Δ*lapA*Δ*lapG* mutant, HA-C235 is constitutively secreted because it lacks the LapA RM (Fig 3C). HA-C235 supernatant levels in the Δ*lapA*Δ*lapG* and Δ*lapD* mutants are identical, indicating that lack of LapA or constitutive LapA secretion, respectively, does not impinge on HA-C235 secretion (Fig 3C, left). However, when LapA is locked at the cell surface in the Δ*lapG* mutant strain, HA-C235 secretion is nearly abolished (Fig 3C, center). HA-C235 secretion is restored in a Δ*lapG* mutant when LapA does not associate with the cell surface (compare Δ*lapG* mutant to Δ*lapG lapAΔD31-95A*; Fig 3C). Together, these data indicate retention of LapA, but not its secretion, limits LapEBC availability for secreting other peptides, and is consistent with the model that the RM of LapA is retained in the LapEBC secretion apparatus.

Chemical Inhibition of LapG Locks LapA at Cell Surface, Occluding the LapEBC T1SS Machinery. LapG is a calcium-dependent cysteine protease that can be chemically inhibited *in vivo* and *in vitro* with micromolar levels of EGTA (19). Addition of EGTA to the

208 Δ*lapD* mutant, which usually exhibits constitutive LapG activity, restores LapA cell-surface
209 localization by inhibiting LapG (19). To determine if chemical inhibition of LapG in the
210 *ΔlapD* background can block HA-C235 secretion by enhancing surface associated LapA, we
211 compared biofilm formation and HA-C235 secretion in the Δ*lapD* mutant grown in the
212 absence or presence of 500 μM EGTA (-EGTA or +EGTA). Each strain was grown for 5 hr,
213 then washed and resuspended for 20 min in the indicated medium (Fig 3D, growth/wash).
214 The supernatants were then collected and HA-C235 protein was probed via Western blot
215 analysis.
216 EGTA treatment restored biofilm formation and inhibited HA-C235 secretion in the
217 Δ*lapD* mutant (Fig 3D). Additionally, washing EGTA grown cells in medium lacking EGTA
218 restored C235 secretion (Fig 3D, right), indicating that removal of EGTA reactivated LapG
219 activity, allowing this protease to cleave LapA from the cell surface and thus allow
220 secretion of HA-C235. These data support our model that LapA localizes to the cell surface
221 via anchoring in its T1S machinery.
222 **LapE Overexpression Rescues HA-C235 Secretion in a *lapG* Mutant Background.**The
223 LapEBC proteins likely form a tripartite complex (LapE-LapBC) to secrete LapA. Given that
224 secretion of HA-C235 is blocked when LapA is locked at the cell surface (i.e., a *lapG* mutant
225 or addition of EGTA), we reasoned that a T1SS component(s) participating in cell surface
226 localization of LapA is likely limiting. Therefore, overexpression of the limiting secretion
227 component should allow additional secretion, and thus should rescue HA-C235 secretion in
228 a Δ*lapG* mutant. Previous work from our lab comparing *lapA*, *lapE*, *lapB*, and *lapC* gene
229 expression in the presence and absence of the biofilm promoting nutrient phosphate
230 indicates only *lapE* is down regulated under phosphate-limiting conditions that inhibit

biofilm formation (27). Thus, we suspected that LapE might be the limiting component of the secretion machinery.

To determine if *lapE* overexpression can rescue HA-C235 secretion in a *lapG* mutant, we assayed for HA-C235 in the supernatant of the Δ*lapG* mutant carrying a pC235-*lapE* transcriptional fusion construct. This construct allows simultaneous expression of HA-C235 and LapE protein from the same promoter. Western analysis of the supernatant fraction indicates co-expression of HA-C235 and LapE, but not HA-C235 alone, can rescue HA-C235 secretion in a *lapG* mutant background (Fig 3E, right). HA-C235 secretion in the Δ*lapA*Δ*lapG* mutant was used as a baseline secretion control. Together, these data are consistent with the model that LapA is retained on the cell surface by LapE, LapE levels are limiting, and retained LapA occupies the secretion pathway thus blocking secretion of any other substrate.

Evidence of a Conserved Retention Strategy. PHYRE analysis (28) predicts the N-termini of putative LapG substrates may adopt similar secondary structures, suggesting these adhesins could tether through the outer membrane like LapA. To test this idea, we replaced the RM of LapA (M1-125S) with low-identity putative RMs from *V. cholerae* and *P. aeruginosa* proteins described earlier (Fig 3F,G, also Fig 1C,D). Figure 3A shows the LapG cleavage site in LapA and VC0395_0338 (boxed in green), the degenerate LapG cleavage site in PA1874 (boxed in yellow), and the putative polyglycine linker common to all three N-termini (boxed in red). Based on our *in vitro* LapG cleavage experiments, the putative RMs may represent separate classes of LapA-like adhesins. VC0395_0338 has a conserved LapG cleavage motif (AAAG) and is processed by LapG *in vitro* (Fig 1C). In contrast, while the N-terminus of the *P. aeruginosa* protein PA1874 has features similar to LapA, it appears

to contain a degenerate LapG processing motif (AAA*I*G) and is not processed in vitro (Fig. 1D). Thus, we predicted PA1874 anchors to the cell surface but is not subject to LapG-mediated proteolysis and release. Using allelic exchange, we replaced the DNA encoding the N-terminus of LapA with DNA corresponding to the amino acids detailed in the alignment in Fig 3A in both the wild type *P. fluorescens* and the Δ*lapG* mutant background. These chimeras are called VC-Swap and PA-Swap. Biofilm formation and chimera localization were compared for the wild type and the Δ*lapG* mutant strains expressing the chimeras to determine if these proteins were retained on the surface, could support biofilm formation in *P. fluorescens* and/or were subject to LapG proteolysis to release the chimeric adhesins from the cell surface.

Western blot analysis of the whole-cell fraction indicates both VC-Swap and PA-Swap chimeras are stable in the wild type (light grey) and *lapG* (dark grey) mutant backgrounds (Fig 3F, WC). The biofilm phenotype and chimera localization of the VC-Swap and PA-Swap retention chimeras are consistent with their differential susceptibility to LapG cleavage *in vitro* (Fig 1 C and D). Western blot analysis of the VC-Swap chimera in the supernatant (Fig 3F, S) and cell surface fractions (Fig 3G, CS) indicates deletion of the *lapG* gene decreases levels of the chimera in the supernatant fraction (Fig 3F), and more importantly, enhances chimera levels at the cell surface (Fig 3F, G; VC-Swap, light grey vs. dark grey) and increased biofilm formation (Fig 3H, VC-Swap; light grey vs. red bar). This trend mimics wild type LapA in the parental wild type and *lapG* backgrounds (Fig 3H WT; light grey vs. red bar). These data indicate the RM from VC0395_0338 (Fig 3A, bottom) can complement LapA localization and LapG-dependent release from the cell surface, suggesting LapA and VC0395_0338 localize to the cell surface by a similar mechanism. Conversely, PA-Swap

277 chimera localization and corresponding biofilm formation were not impacted by LapG
278 activity. Consistent with our *in vitro* LapG cleavage analysis, these data indicate that the
279 PA1874 RM can complement LapA's RM for cell surface retention and biofilm formation,
280 but not for cleavage by LapG due the presence of a degenerate cleavage motif (Fig 3F-H).
281 **ABC Transporters of LapA-like Adhesins Form a Distinct T1SS Subgroup.** ABC
282 transporters are found throughout all three domains of life, but can often be functionally
283 grouped based on common residues that are critical for secreting their substrate(s) (29–
284 31). Studies of the T1SS ATPase HlyB of *E. coli*, which transports the RTX toxin HlyA,
285 demonstrated HlyB contains a N-terminal domain that is critical for binding HlyA's C-
286 terminal RTX motifs (31). The N-terminal domain of HlyB resembles the C39 peptidase
287 domain (C39) typically found at the N-terminus of the ATPase component of bacteriocin
288 ABC transporters, but HlyB lacks the catalytic cysteine residue required to cleave and
289 activate immature bacteriocins during secretion (32). Instead, the C39-like domain (CLD)
290 of HlyB contains a tryptophan involved in binding HlyA that is conserved among many ABC
291 transporters involved in secreting RTX toxins, including the ABC transporter secreting
292 CyaB of *B. bronchiseptica*. Thus, the C39 and CLD of bacteriocin and RTX toxin transporters
293 can be distinguished by their amino acid sequence (31).
294 Interestingly, the T1SS ATPase for LapA, called LapB, contains a N-terminal domain that
295 lacks both the conserved cysteine and tryptophan residues of C39 and CLD involved in
296 bacteriocin-processing and RTX-binding, respectively. Given LapA's unusual retention
297 strategy and the observation that LapEBC is often encoded nearby LapDG homologs (11,
298 33), we were curious if LapB-like transporters may represent a distinct subgroup of ABC
299 transporters involved in secreting adhesins with the newly defined N-terminal RM. To test

this idea, we assessed the phylogenetic relationship between C39 and CLD sequences from the ATPases of bacteriocin and RTX toxin transporters, and the N-terminal domain of the ATPase from transporters involved in adhesin retention. The phylogenetic relationships of N-terminal domains from the ATPase component several characterized RTX toxin and bacteriocin transporters were identified using Interpro (http://www.ebi.ac.uk/interpro/), and compared with N-terminal domain of the putative LapB-like ATPase component. LapB-like ABC transporters were defined as those encoded near LapDG homologs. We also included LssB of *L. pneumophila* and PA1876 of *P. aeruginosa* in the analysis. Recent studies indicate LssB is an ABC transporter involved with secreting an RTX adhesin, RtxA, which was predicted to be a LapG substrate by our algorithm (Fig S1B) (34). PA1876 is likely involved with transporting PA1874, the LapA-like protein with a degenerate cleavage motif described here (Fig 1D and Fig 3F-H) (20).

Our phylogenic analysis suggests the ATPase component of the ABC transporters of LapA-like adhesins form a distinct group that lack the functional residues critical for RTX motif-binding and bacteriocin-processing (Fig 4A, see WebLogo). These differences at the amino acid level likely reflect functional differences, rather than phylogenetic diversity of the organisms analyzed, because characterized RTX toxin and LapA-like adhesin ATPases encoded within the same genome cluster with their predicted substrate type. For example, the RTX toxin transporter CyaB (Fig 4A; orange circle) and LapA-like adhesin transporter BB1189 (red circle) encoded by *B. bronchiseptica* map to different clusters in this analysis. These data support the idea that adhesins with an N-terminal RM are secreted by a distinct group of T1SS ABC transporters.

Discussion. Here, we describe a novel N-terminal element in LapA - called the retention module - that is responsible for localizing this giant adhesin to the cell surface. We propose structural features within LapA's RM play a role in tethering LapA to the cell surface and show that disrupting these predicted secondary structures can lead to uncontrolled release of the adhesin from the cell surface, as was observed for the LapAΔD31-95A mutant. Consistent with these observations, low sequence identity RMs from putative LapA-like proteins share similar predicted secondary structures and can complement LapA function and localization, suggesting LapA's unusual retention strategy is broadly conserved among this group of adhesins. The bioinformatic analysis presented here and previously (33, 35) suggest LapA-like proteins are secreted by an unappreciated subgroup of T1SS ABC transporters that engage TolC-like pores dedicated to the secretion and retention of these large adhesins.

How does the RM contribute to LapA's cell surface localization? We propose the RM stalls the final steps of LapA translocation, leaving LapA threaded through the TolC-like outer membrane pore, LapE, with its RM localized in the periplasm, accessible to LapG, and C-terminal adhesive repeats displayed at the cell surface (Fig 4B). Although this model opposes the classical one-step paradigm detailed in over ~25 years, the model is consistent with our artificial HA-C235 substrate competition experiments, EGTA-mediated inhibition of LapG activity, and expression studies with the TolC-like LapE outer membrane protein presented here. Given the abundance of LapA-like adhesins in pathogens and environmental microbes, and the conserved role of c-di-GMP in regulating their cell surface localization, this means of secreting and anchoring an adhesion appears to be a general strategy used for cell-surface and perhaps cell-cell adherence.

Based on the recently reported *Marinomonas* MpIBP structural modeling (26), and the data presented here and reported previously (13), we built a model of the LapA retention module (Fig 4B). Our genetic analysis of the N-terminal RM is in good agreement with the proposed N-terminal domain of MpIBP being localized to the periplasm. Interestingly, the proposed model for the MpIBP adhesin of *Marinomonas* (26) indicates the LapG proteolysis site may be obscured by the TolC pore, and our mutational analysis suggests regions not included in the MpIBP NMR structure are also important for retention and LapG processing (13). In contrast, we propose that the LapG processing site is accessible in the periplasm, a conclusion consistent with our modeling (Fig. 4B) and our previous biochemical and genetic studies (9, 12, 24).

Materials and Methods.

Strains and media. *P. fluorescens* and *E. coli* strains listed in Table S3 were grown on lysogeny broth at 30˚C and 37˚C, respectively. Gentamycin was used when appropriate (10 µg/mL for *E. coli*, 30µg/mL for *P. fluorescens*). For biofilm and LapA localization analysis, *P. fluorescens* strains subcultured in K10T-1 for 6 hr statically and with rotation, respectively (27).

Static Biofilm Assay. The static biofilm assay was performed and quantified as described previously (13).

LapA Localization. Whole cell, supernatant, and cell surface localization analysis was preformed as previously using a HA-tagged variant of LapA (13).

HA-C235 Localization. Each strain was subcultured for 5.5 hr, washed, and then resuspended for 20 min in the indicated medium under non-inducing conditions. For chemical inhibition assays, K10T-1 medium was supplemented with 500µM EGTA as noted.

Whole cell and supernatant fractions were prepared as for LapA localization. Western blot analysis against the HA epitope was used to detect HA-C235.

***In Vitro* LapG Cleavage Analysis.** *In vitro* cleavage analysis was performed as described previously (13).

***In Silico* Prediction of LapG Substrates.** LapG and LapD homologs were defined as ORFs encoding proteins with the pfam06035 and pfam16448 domains, respectively. The NCBI CDD was utilized to generate a list of LapG- and LapD-encoding bacteria, and the programming language R was used to determine the intersecting LapD/LapG-encoding bacteria. The protein annotation of these genomes were downloaded from the NCBI genome database and each annotated locus was interrogated for the presence of a LapG cleavage site within amino acids 80-150 ([T/A/P]AA[G/V]) and at least one RTX motif (Dx[L/I]x(4)GxDx[L/I]xGGx(3)D).

Acknowledgment. We thank C. Boyd for building one of the control strains used in this study, T. Silhavy for his insight interpreting some of the data presented here.

Literature Cited.

1. Hall CW, Mah T-F (2017) Molecular mechanisms of biofilm-based antibiotic resistance and tolerance in pathogenic bacteria. *FEMS Microbiol Rev* 41(3):276–301.
2. Liu T, et al. (2015) Enhanced *Shewanella* biofilm promotes bioelectricity generation. *Biotechnol Bioeng* 112(10):2051–2059.
3. Ambrosis N, Boyd CD, O'Toole GA, Fernández J, Sisti F (2016) Homologs of the LapD-LapG c-di-GMP effector system control biofilm formation by *Bordetella bronchiseptica*. *PLoS One* 11(7):1–16.
4. Syed KA, et al. (2009) The Vibrio cholerae flagellar regulatory hierarchy controls

expression of virulence factors. *J Bacteriol* 191(21):6555–6570.

5. Guo S, Garnham CP, Whitney JC, Graham LA, Davies PL (2012) Re-evaluation of a bacterial antifreeze protein as an adhesin with ice-binding activity. 7(11). doi:10.1371/journal.pone.0048805.

6. Hinsa SM, Espinosa-Urgel M, Ramos JL, O'Toole GA (2003) Transition from reversible to irreversible attachment during biofilm formation by *Pseudomonas fluorescens* WCS365 requires an ABC transporter and a large secreted protein. *Mol Microbiol* 49(4):905–918.

7. Cirillo SLG, Bermudez LE, El-Etr SH, Duhamel GE, Cirillo JD (2001) *Legionella pneumophila* entry gene rtxA is involved in virulence. *Infect Immun* 69(1):508–517.

8. Wu C, et al. (2013) Oxygen promotes biofilm formation of *Shewanella putrefaciens* CN32 through a diguanylate cyclase and an adhesin. *Sci Rep* 3(1):1945.

9. Newell PD, Boyd CD, Sondermann H, O'Toole GA (2011) A c-di-GMP effector system controls cell adhesion by inside-out signaling and surface protein cleavage. *PLoS Biol* 9(2). doi:10.1371/journal.pbio.1000587.

10. Newell PD, Monds RD, O'Toole GA (2009) LapD is a bis-(3',5')-cyclic dimeric GMP-binding protein that regulates surface attachment by *Pseudomonas fluorescens* Pf0-1. *Proc Natl Acad Sci U S A* 106(9):3461–6.

11. Navarro MVAS, et al. (2011) Structural basis for c-di-GMP-mediated inside-out signaling controlling periplasmic proteolysis. *PLoS Biol* 9(2). doi:10.1371/journal.pbio.1000588.

12. Cooley RB, et al. (2016) Cyclic di-GMP-regulated periplasmic proteolysis of a *Pseudomonas aeruginosa* type Vb secretion system substrate. *J Bacteriol* 198(1):66–

76.

13. Boyd CD, et al. (2014) Structural features of the *Pseudomonas fluorescens* biofilm adhesin LapA required for LapG-dependent cleavage, biofilm formation, and cell surface localization. *J Bacteriol* 196(15):2775–2788.

14. Koronakis V, Hughes C, Koronakis E (1991) Energetically distinct early and late stages of HlyB/HlyD-dependent secretion across both *Escherichia coli* membranes. *EMBO J* 10(11):3263–72.

15. Satchell KJF (2011) Structure and function of MARTX toxins and other large repetitive RTX Proteins. *Annu Rev Microbiol* 65(1):71–90.

16. Chatterjee R, Nag S, Chaudhuri K (2008) Identification of a new RTX-like gene cluster in *Vibrio cholerae*. *FEMS Microbiol Lett* 284(2):165–171.

17. Cirillo SLG, Lum J, Cirillo JD (2000) Identification of novel loci involved in entry by *Legionella pneumophila*. *Direct* (2000):1345–1359.

18. Chatterjee D, Boyd CD, O'Toole GA, Sondermann H (2012) Structural characterization of a conserved, calcium-dependent periplasmic protease from *Legionella pneumophila*. *J Bacteriol* 194(16):4415–4425.

19. Boyd CD, Chatterjee D, Sondermann H, O'Toole GA (2012) LapG, required for modulating biofilm formation by *Pseudomonas fluorescens* Pf0-1, is a calcium-dependent protease. *J Bacteriol* 194(16):4406–4414.

20. Zhang L, Mah TF (2008) Involvement of a novel efflux system in biofilm-specific resistance to antibiotics. *J Bacteriol* 190(13):4447–4452.

21. Lenders MHH, et al. (2015) Directionality of substrate translocation of the hemolysin A Type I secretion system. *Sci Rep* 5(1):12470.

22. Gadeberg O V, Orskov I (1984) In vitro cytotoxic effect of alpha-hemolytic *Escherichia coli* on human blood granulocytes. *Infect Immun* 45(1):255–60.

23. Fullner Satchell KJ (2007) MARTX, multifunctional autoprocessing repeats-in-toxin toxins. *Infect Immun* 75(11):5079–5084.

24. Boyd CD, et al. (2014) Structural features of the *Pseudomonas fluorescens* biofilm adhesin LapA required for LapG-dependent cleavage, biofilm formation, and cell surface localization. *J Bacteriol* 196(15):2775–2788.

25. Borlee BR, et al. (2010) *Pseudomonas aeruginosa* uses a cyclic-di-GMP-regulated adhesin to reinforce the biofilm extracellular matrix. *Mol Microbiol* 75(4):827–842.

26. Guo S, et al. (2017) Structure of a 1.5-MDa adhesin that binds its Antarctic bacterium to diatoms and ice. (August):1–10.

27. Monds RD, Newell PD, Gross RH, O'Toole GA (2007) Phosphate-dependent modulation of c-di-GMP levels regulates *Pseudomonas fluorescens* Pf0-1 biofilm formation by controlling secretion of the adhesin LapA. *Mol Microbiol* 63(3):656–679.

28. Kelly LA, Mezulis S, Yates C, Wass M, Sternberg M (2015) The Phyre2 web portal for protein modelling, prediction, and analysis. *Nat Protoc* 10(6):845–858.

29. Kanonenberg K, Schwarz CKW, Schmitt L (2013) Type I secretion systems - a story of appendices. *Res Microbiol* 164(6):596–604.

30. ter Beek J, Guskov A, Slotboom DJ (2014) Structural diversity of ABC transporters. *J Gen Physiol* 143(4):419–435.

31. Lecher J, et al. (2012) An RTX transporter tethers its unfolded substrate during secretion via a unique N-terminal domain. *Structure* 20(10):1778–1787.

460 32. Havarstein LS, Diep DB, Nes IF (1995) A family of bacteriocin ABC transporters carry
461 out proteolytic processing of their substrates concomitant with export. *Mol Microbiol*
462 16(2):229–240.
463 33. Ginalski K, Kinch L, Rychlewski L, Grishin N (2004) BTLCP proteins: a novel family of
464 bacterial transglutaminase-like cysteine proteinases. *Trends Biochem Sci* 29(8):389–
465 392.
466 34. Fuche F, Vianney A, Andrea C, Doublet P, Gilbert C (2015) Functional type 1 secretion
467 system involved in *Legionella pneumophila* virulence. *J Bacteriol* 197(3):563–571.
468 35. Theunissen S, et al. (2009) The agglutination protein AggA from *Shewanella*
469 *oneidensis* MR-1 is a TolC-like protein and forms active channels in vitro. *Biochem*
470 *Biophys Res Commun* 386(2):380–385.
471

472 **Figures**

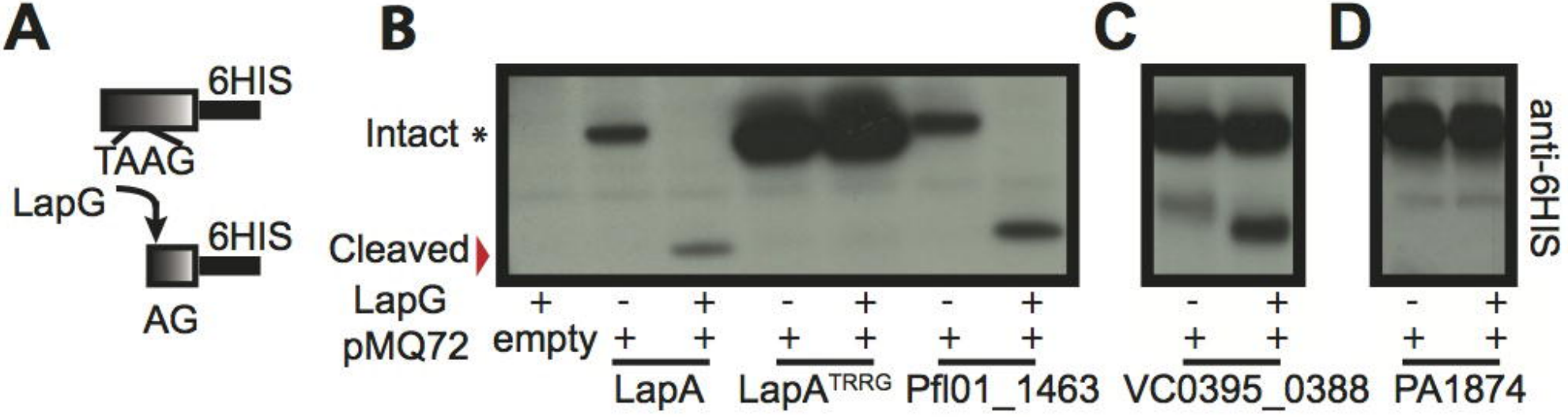
473 **Figure 1. *P. fluorescens* Pf0-1 LapG *in vitro* cleavage analysis.** (A) Overview of *in vitro*
474 LapG cleavage assay. (B-D). Intact (*) and cleaved (red arrow) N-termini were visualized
475 by Western blot. Equal protein concentrations of the substrate and extracts from LapG-
476 expressing strains were used.

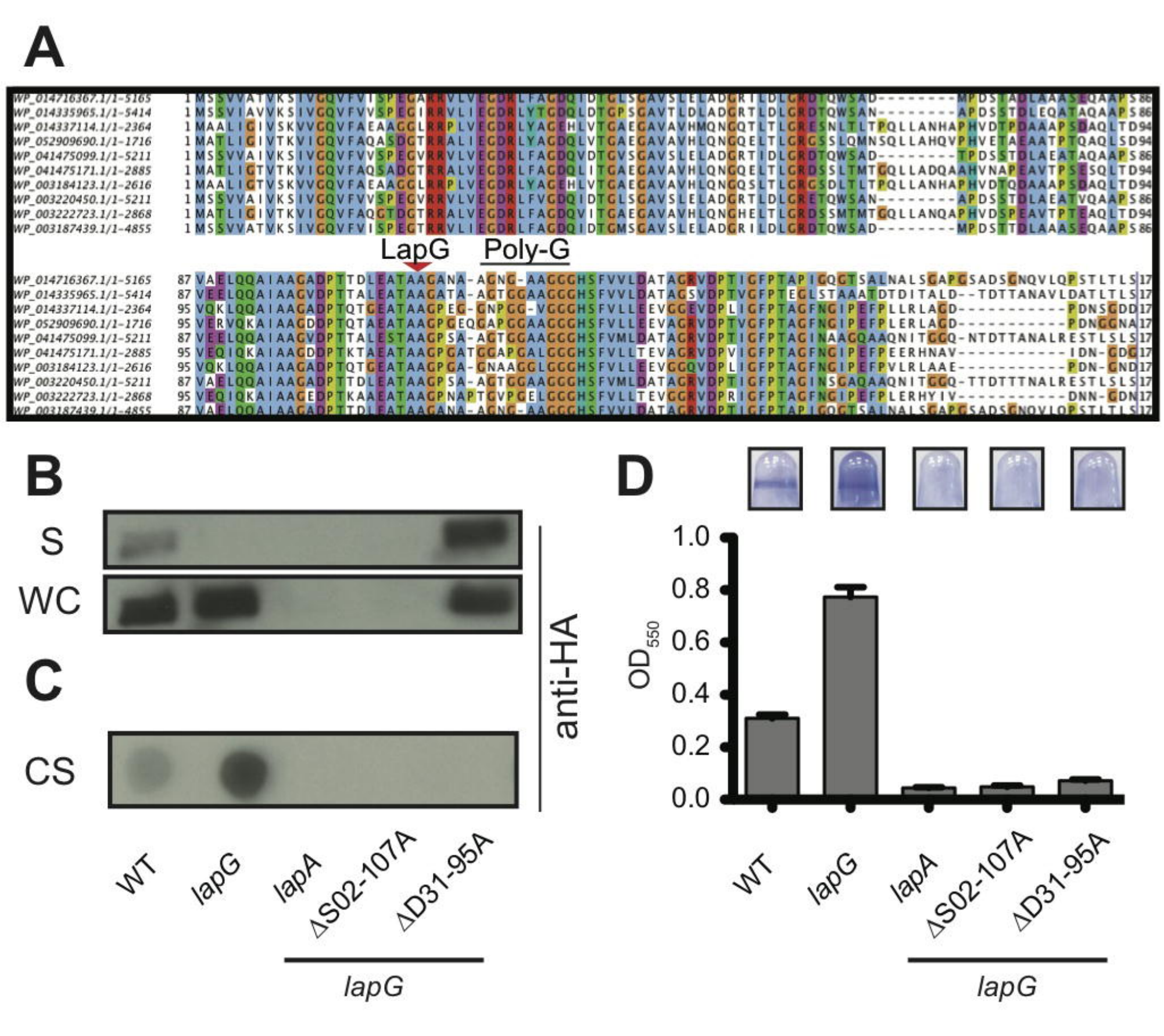
477 **Figure 2. LapA's N-terminus Serves as a Retention Module.** (A) The first ~175 aa of the
478 *P. fluorescens* LapA-like N-termini (Fig. S1A, red box) with LapG cleavage motif (red arrow)
479 and putative polyglycine linker (Poly-G, black bar). (B) Western blot analysis of
480 supernatant and whole cell fractions for LapA. (C) Dot blot analysis of surface-associated
481 LapA. (D) Biofilm analysis of LapA N-terminal mutants (n=8, +/-SEM).

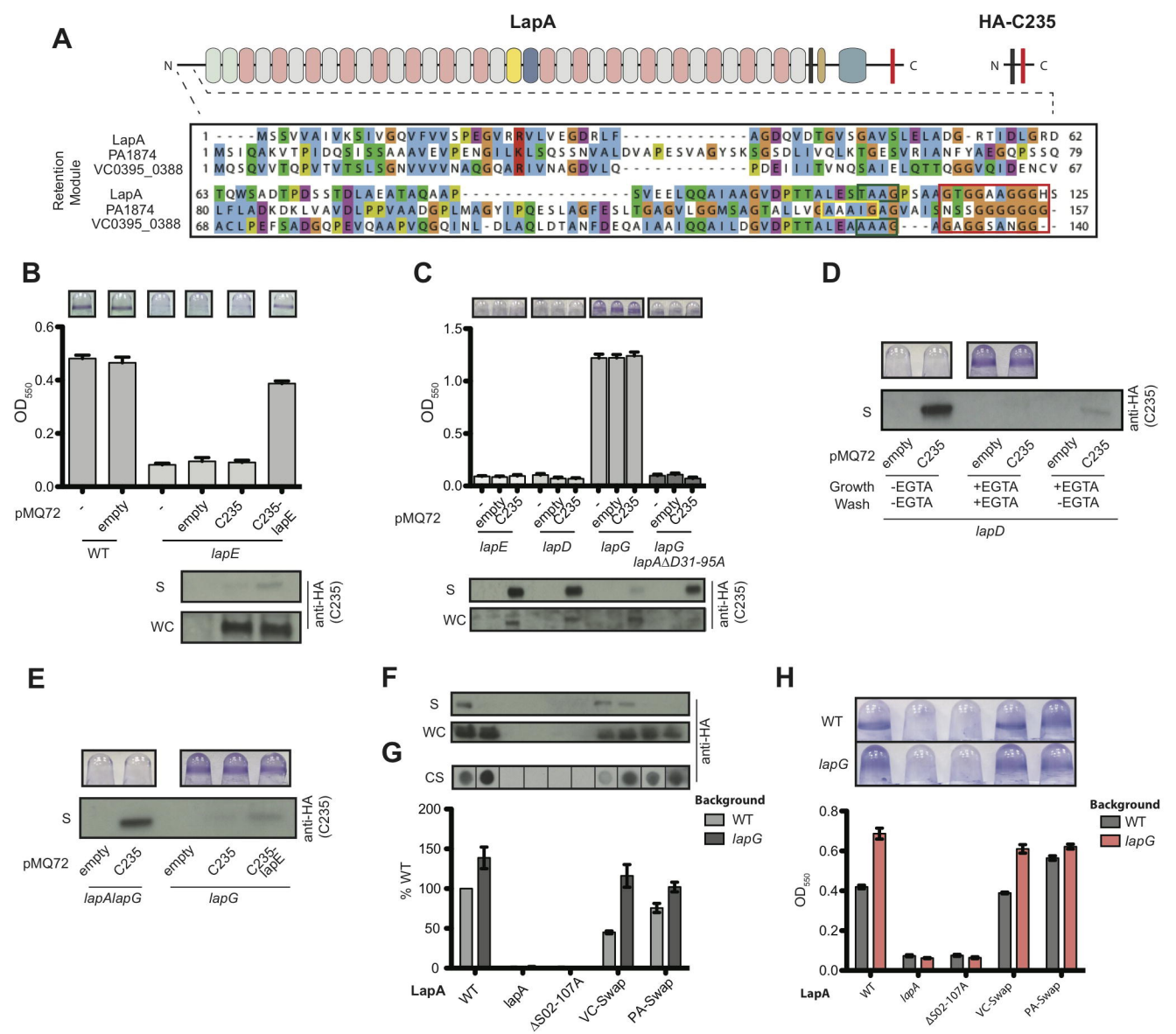
Figure 3. LapA Cell Surface Localization Impacts LapEBC Activity. (A) Top: Scaled representation of LapA and HA-C235. Black bar indicates 3XHA tag in LapA and HA-C235. Bottom: MUSCLE alignment of putative RMs from *P. fluorescens*, *P. aeruginosa*, and *V. cholerae* adhesins with residues colored according to the default Clustal X coloring scheme in Jalview. The putative polyglycine linker is boxed in red while the canonical and degenerate LapG cleavage sites are boxed in green and yellow, respectively. (B-C) Biofilm formation (top) and Western blot analysis of HA-C235 in the supernatant (S) and whole cell (WC) fractions (bottom). In these experiments, LapE and HA-C235 are co-expressed as a transcriptional fusion from a plasmid. (D) Biofilm analysis (top) and HA-C235 secretion (bottom) in Δ*lapD* mutants with and without EGTA treatment. (E) Biofilm analysis (top) and HA-C235 secretion (bottom) in Δ*lapG* mutants. (F) Western blot analysis of LapA RM chimeras VC0395_0388 (VC-Swap; also panel A, bottom.) and PA1874 (PA-Swap; also panel A, bottom) in the supernatant (S) and whole cell (WC) fractions in the indicated genetic backgrounds. (G) Western blot analysis of cell surface associated LapA RM chimeras (top) and quantification (bottom). Strains in panel G correspond to strains in panel F. (H) Biofilm formation of LapA RM chimeras from strains in panel G. For biofilm and dot blot quantification (n=8, +/-SEM).

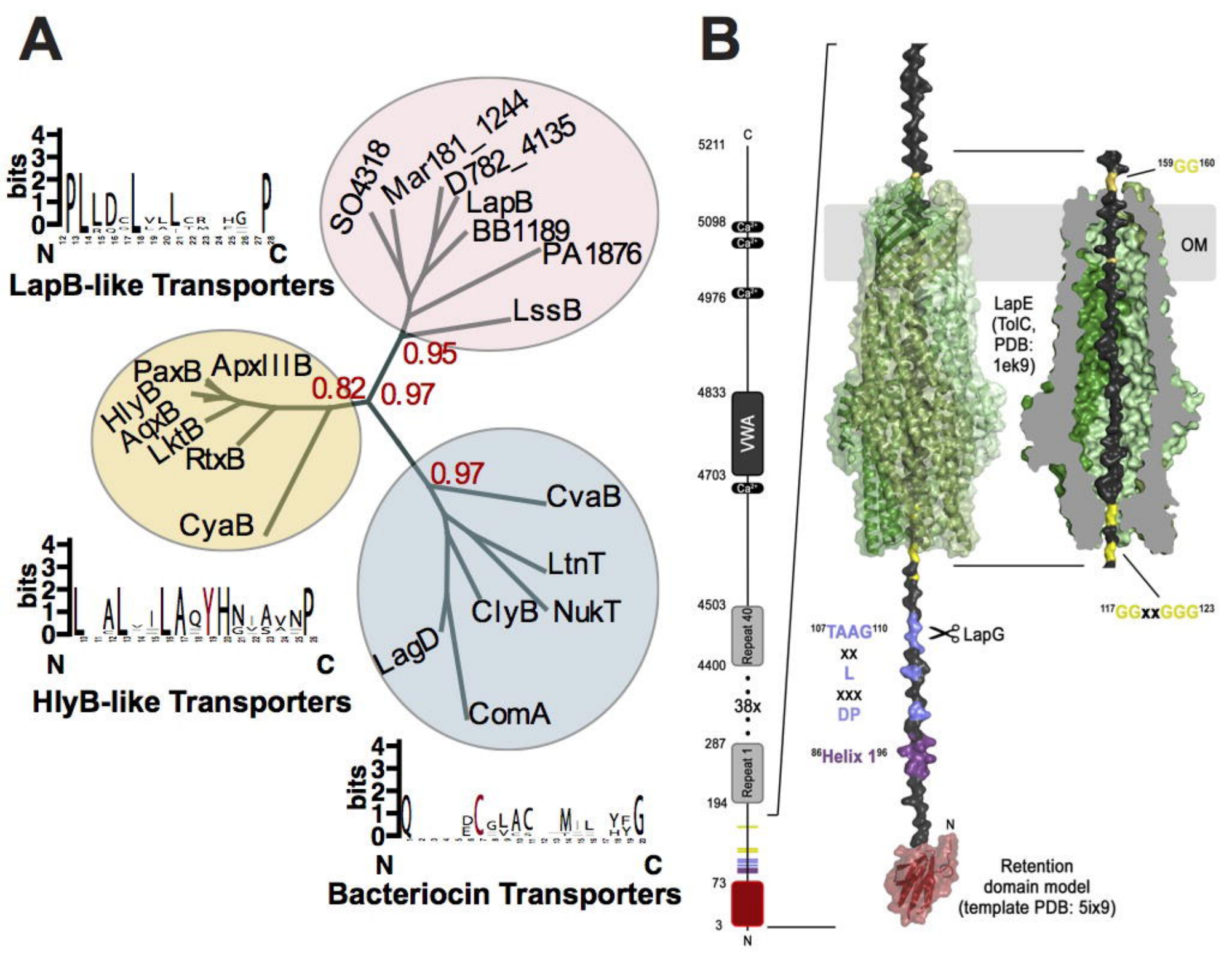
Figure 4. Phylogenetic Analysis of ABC Transporter Subfamilies. (A) Phylogenetic analysis of C39 peptidase domain, C39-like peptidase domain (CLD), and N-terminal ~125 aa of LapB-like ATPases (Table S4) using the online phylogenetic analysis program phylogeny.fr under default settings (http://www. phylogeny.fr). WebLogos were generated by MUSCLE aligning members from each branch and truncated to highlight functional residues when applicable (red). Bootstrap values are indicated in red. (B) Left: Domain

505 architecture of LapA. Right: Surface retention model for LapA. The structural model of the
506 LapA retention module was generated using the Phyre2 protein fold recognition server and
507 PDB 5ix9 as the template (red domain). The sequence following the retention domain was
508 modeled as an extended peptide to reflect secondary structure predictions indicating this
509 region (residues 97-173) to be flexible. The unstructured sequence was threaded through
510 LapE, represented by the crystal structure of TolC (PDB: 1ek9). The LapG-targeting
511 dialanine motif (TAAG) and regions critical for LapG binding (Helix 1 and DPxxxLxx) are
512 shown, and proposed to be exposed in the periplasm. Two glycine-rich motifs in LapA,
513 which flank the sequences within the TolC pore, are indicated in yellow.


A
6HIS
TAAG
LapG
6HIS
AG
B
Intact *
Cleaved
LapG
pMQ72
empty
LapA
LapA^TRRG
Pfl01_1463
C
VC0395_0388
D
PA1874
anti-6HIS


A
LapG
Poly-G
B
S
WC
C
CS
anti-HA
WT
lapG
lapA
ΔS02-107A
ΔD31-95A
lapG
D
OD550
1.0
0.8
0.6
0.4
0.2
0.0
WT
lapG
lapA
ΔS02-107A
ΔD31-95A
lapG


A
LapA
HA-C235
N
C
Retention Module
LapA
PA1874
VC0395_0388
B
C
D
pMQ72
empty
C235
C235-lapE
WT
lapE
lapD
lapG
lapGΔD31-95A
S
WC
anti-HA (C235)
Growth
Wash
-EGTA
+EGTA
E
lapAlapG
F
G
CS
anti-HA
Background
WT
lapG
% WT
LapA
ΔS02-107A
VC-Swap
PA-Swap
H


A
bits
N
C
LapB-like Transporters
SO4318
Mar181_1244
D782_4135
LapB
BB1189
PA1876
LssB
0.95
PaxB
ApxIIIB
0.82
0.97
HlyB
AqxB
LktB
RtxB
CyaB
0.97
CvaB
LtnT
ClyB
NukT
LagD
ComA
HlyB-like Transporters
Bacteriocin Transporters
B
5211
C
5098
Ca2+
4976
4833
VWA
4703
4503
Repeat 40
4400
38x
287
Repeat 1
194
73
3
N
OM
LapE
(TolC,
PDB:
1ek9)
159GG160
117GGxxGGG123
107TAAG110
xx
L
xxx
DP
LapG
86Helix 196
Retention
domain model
(template PDB: 5ix9)