Type IV pili-associated secretion of a biofilm matrix protein from *Clostridium perfringens* that forms intermolecular isopeptide bonds

Running title: Isopeptide bonds link *C. perfringens* biofilm proteins

Sarah E. Kivimaki[1*], Samantha Dempsey[1#], Collette Camper[1†], Julia M. Tani[1#], Ian K. Hicklin[2‡], Crysten E. Blaby-Haas[3,4], Anne M. Brown[2], Stephen B. Melville[1]

1. Department of Biological Sciences, Virginia Tech, Blacksburg, VA 24061, USA
2. University Libraries and Department of Biochemistry, Virginia Tech, Blacksburg, VA 24061, USA
3. Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
4. DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

***Correspondence:**

Stephen Melville

melville@vt.edu

* Current address: Janelia Research Campus, 19700 Helix Dr., Ashburn VA, 20147

\# Current address: TechLab Inc., 2001 Kraft Drive, Blacksburg, VA 24060-6358

† Current address: University of South Florida, Tampa Campus, 4202 E. Fowler Avenue, Tampa, FL 33620

‡ Current address: University of California, Berkeley, 110 Sproul Hall #5800, Berkeley, CA 94720

Key words: *Clostridium perfringens*, biofilm matrix protein, protein secretion, isopeptide bonds

Abstract

Clostridium perfringens is a Gram-positive, anaerobic, spore-forming, bacterial pathogen of humans and animals. *C. perfringens* also produces type IV pili (T4P) and has two complete sets of T4P-associated genes, one of which has been shown to produce surface pili needed for cell adherence. One hypothesis about the second set of T4P genes is that they comprise a system analogous to the type II secretion systems (TTSS) found in Gram-negative bacteria, which is used to export folded proteins from the periplasm through the outer membrane to the extracellular environment. Gram-positive bacteria have a similar secretion barrier in the thick peptidoglycan (PG) layer, which blocks secretion of folded proteins >25 kD. To determine if the T4P-associated genes comprise a Gram-positive TTSS, the secretome of mutants lacking type IV pilins were examined and a single protein, a von Willebrand A domain containing protein, BsaC (CPE0517), was identified as being dependent on pilin PilA3 for secretion. The *bsaC* gene is in an operon with genes encoding a SipW signal peptidase and two putative biofilm matrix proteins BsaA and BsaB, both of which have remote homology to *Bacillus subtilis* biofilm protein TasA. Since BsaA forms long oligomers that are secreted, we analyzed BsaA monomer interactions with *de novo* modeling. These models projected that the monomers formed isopeptide bonds as part of a donor strand exchange process, in which an N-terminal disordered loop of one monomer intercalates into a beta sheet structure of an adjacent monomer and reforms into a beta sheet with subsequent isopeptide bond formation. Mutations in residues predicted to form the isopeptide bonds led to loss of oligomerization, supporting an exchange and lock mechanism. Phylogenetic analysis showed the BsaA family of proteins are widespread among bacteria and archaea but only a subset is predicted to form isopeptide bonds.

72 **Importance**

73 For bacteria to secrete folded proteins to the environment, they have to overcome the physical
74 barriers of an outer membrane in Gram-negative bacteria and the thick peptidoglycan layer in
75 Gram-positive bacteria. One mechanism to do this is the use of a Type II secretion system in
76 Gram-negative bacteria, which has a structure similar to type IV pili and is modeled to act as a
77 piston that pumps folded proteins through the outer membrane to the environment. *Clostridium*
78 *perfringens*, like all or most all of the clostridia, has type IV pili and, in fact, has two sets of
79 pilus-associated genes. Here we present evidence that *C. perfringens* uses one set of pilus genes
80 to secrete a biofilm associated protein and may be responsible for secreting the main biofilm
81 protein, BsaA. We show that BsaA monomers are, unlike most other biofilm matrix proteins,
82 linked by intermolecular isopeptide bonds, enhancing the physical strength of BsaA fibers.

83 **Introduction**

84 Bacterial protein secretion to the exterior environment involves transporting the protein across
85 several physical barriers. For Gram-negative bacteria these include the cytoplasmic membrane
86 (CM), a relatively thin peptidoglycan (PG) cell wall and the outer membrane. Some proteins, after
87 secretion through the CM, need to be folded in the periplasmic space, which can prevent them
88 from transiting the outer membrane. To get around this problem, Gram-negative bacteria have
89 evolved a Type II secretion systems (TTSS), which have many proteins similar to those found in
90 Type IV pili (T4P) (Fig. 1 and (1, 2)). TTSS are modeled to function as pistons that use a short
91 pseudopilus to pump folded periplasmic proteins through a protein channel (secretin) in the outer
92 membrane (Fig. 1 and (3, 4)). Protein secretion in Gram-positive bacteria involves transport across
93 the CM and a usually thick PG layer. There are many protein secretion mechanisms for transiting
94 the CM of Gram-positive bacteria, including the Sec system, Twin arginine translocation (Tat),
95 Type IV secretion systems, Esat-6 or type VII secretion system, flagella export apparatus (FEA),
96 fimbrilin (pilus) protein exporter (FPE), and holin-dependent translocation. The Type IV, Type
97 VII and FEA can translocate proteins across the PG layer (5, 6). However, proteins secreted
98 through the CM via the Sec system, used for the majority of proteins, need to fold and, once they
99 do, the PG layer of Gram-positive bacteria acts as a barrier to secretion to the environment because
100 the mesh size in the PG layer is too small to allow diffusion of globular proteins >25 kD, as
101 determined by measurements of the PG from the Gram-positive bacterium *Bacillus subtilis* (7).
102 One possible mechanism for secretion of folded proteins from the space between the CM and the
103 PG layer in Gram-positive bacteria would be a system analogous to TTSS in Gram-negative
104 bacteria (Fig. 1). Our discovery that T4P systems are ubiquitous and perhaps universal in all the
105 Clostridia (8) provided a means for a potential TTSS to exist in the Clostridia. We hypothesized in

106 a review (9) that in Clostridial species with two separate T4P systems, one of them could function
107 as a TTSS to transport proteins through the PG layer. Multiple species of *Clostridium* do carry two
108 copies of complete T4P systems, including *C. perfringens* (Fig. 2) and *Clostridioides difficile* (9).
109 The core proteins needed for functional T4P usually include a pilin, a pre-pilin peptidase (PilD),
110 an assembly ATPase (PilB), an inner membrane core protein (IMCP, PilC), and inner membrane
111 accessory proteins (IMAP) PilM, PilN, and PilO (9). *C. perfringens* carries multiple pilin-
112 encoding genes, two homologs of PilB, two homologs of PilC and single copies of PilM, PilN, and
113 PilO (Fig. 2), making it plausible that one of these systems could function as a TTSS.
114 In 2017 a paper by Obana et al (10) reported that among many other gene transcripts, RNase Y
115 regulated the transcript levels of an operon, *CPE0514-CPE0517*, containing genes encoding
116 extracellular proteins. A later report by Obana et al (11) was published in 2020 in which they
117 examined the regulation and several other features of the *CPE0514-CPE0517* operon that encoded
118 a SipW signal peptidase, two biofilm matrix proteins and a protein with a von Willebrand A
119 domain, the genes of which they designated as *sipW-bsaA-bsaB-bsaC*. BsaA, BsaB, and BsaC are
120 all predicted to have secretion signal sequences (Signal-P 6.0), which suggests they are secreted
121 via the Sec system and the signal sequence is then removed by the SipW signal peptidase. BsaA
122 and BsaB have low levels of sequence homology to the well-studied biofilm matrix protein TasA
123 of *B. subtilis*, (12-15), indicating BsaA and BsaB may also be biofilm matrix proteins. While much
124 of the report of Obana et al. from 2020 concerned the regulation of matrix protein synthesis under
125 different experimental conditions (11), they determined that the BsaA protein was the main
126 component of the *C. perfringens* biofilm matrix. They also discovered some important physical
127 features of BsaA: (1) it required SipW for signal processing and secretion; (2) expression in *E. coli*
128 of a form of BsaA lacking a signal sequence led to spontaneous oligomerization of BsaA, which

129 was resistant to denaturation by SDS or formic acid, indicting the oligomers were held together by
130 very strong intermolecular interactions; (3) a mutant lacking BsaB still formed biofilm matrix
131 oligomers, suggesting it was not important in BsaA oligomerization (11)
132 These two seemingly disparate areas of research described above overlapped when we undertook
133 experiments to determine if T4P-associated pilin proteins were involved in protein secretion. Prior
134 to the published observations of Obana et al, we discovered that the von Willebrand A domain
135 protein, CPE0517, designated as BsaC by Obana et al. (11), required the PilA3 pilin for secretion
136 (see Results section, below). This demonstrated for the first time that a Gram-positive bacterium
137 required T4P proteins for secretion, providing strong support for our hypothesis that a TTSS-like
138 system is functioning in a Gram-positive bacterium. This finding led us to test whether the main
139 biofilm matrix protein, BsaA, also required T4P-associated proteins for secretion. While our
140 findings indicate that several T4P proteins are needed for secretion of BsaA oligomers, these
141 experiments also led to our discovery that the oligomers of BsaA observed by Obana et al and our
142 group involved formation of isopeptide bonds between BsaA monomers. We then characterized
143 the nature of these bonds and conservation of isopeptide bond formation based on phylogeny of
144 the BsaA/TasA family of proteins. Identifying a potential TTSS-like system in a Gram-positive
145 bacterium and the novel finding of isopeptide bond formation in a well-studied family of biofilm
146 matrix proteins will lead to significant advances in both of these emerging fields of microbial cell
147 biology.
148

149 **Results**

150 **Secretome analysis of a *pilA3* mutant indicates a 70 kD protein with a Von Willebrand**
151 **domain A (CPE0517, BsaC) is absent.** To test our hypothesis that T4P proteins were needed for

152 protein secretion, we examined the secretome of strains with in-frame deletions in genes encoding
153 Type 4 pilins

154 We screened the secretome of in-frame deletion mutants in four pilin-encoding genes, *pilA1*,
155 *pilA2*, *pilA3*, *pilA4* for missing proteins, indicating they may be dependent on the pilin for
156 secretion. Only the *pilA3* mutant showed evidence of a distinct missing band in the secretome and
157 this band was restored by complementation of a WT copy of the *pilA3* gene (Fig. 3A). The band
158 was excised from a gel and subjected to trypsin digestion and analysis by mass spectrometry.
159 Peptide fragments matched those of the CPE0517 (BsaC) protein from the *C. perfringens* strain
160 13 genome. The protein has an unknown function but does contain a von Willebrand A domain,
161 which can potentially function in protein-protein interactions with host cells. Western blots using
162 anti-BsaC antibodies showed a similar loss of CPE0517 secretion by the *pilA3* mutant and
163 complementation by expression of a WT copy of the *pilA3* gene (Fig. 3B). This gene is the fourth
164 gene in a four gene operon containing a type I signal peptidase (CPE0514, SipW) and two genes
165 (*CPE0515-CPE0516*, *bsaA-bsaB*) encoding proteins with similarity to the TasA protein in *B.*
166 *subtilis* (Fig. 3C).

167 To determine if other Type IV pilus (T4P) related genes were needed for secretion of BsaC, we
168 expressed the *bsa* operon in a bank of in-frame deletion mutants encompassing the gene known be
169 required for synthesis and assembly of T4P (Fig. 2) using a vector we developed containing a
170 lactose-inducible promoter (16). The amounts of secreted BsaC were measured using quantitative
171 western blotting of TCA precipitated secreted proteins. We noted variation in day-to-day results
172 with the WT strain and, since these were western blots on TCA precipitated protein from culture
173 supernatants, we normalized all of the data from mutant strains to that of the WT strain on the
174 same blot (see legend in Fig. 4). Using quantitative densitometry on western blots of the secreted

175 oligomers, we measured statistically significant decreased secretion of BsaC in the following
176 mutants: *pilA3*, *pilB2*, *pilC1*, *pilD*, *pilM*, *pilN* and *pilO* (Fig. 3). The proteins encoded by these
177 genes have the ability to form a complete Type IV pilus. Mutations in the *pilT* gene (encodes a T4P
178 retraction ATPase), *cpe2277* (encodes a protein of unknown function) and *cpe2279* and *cpe2280*
179 (encode minor pilins) did not affect the levels of secretion of BsaC.

180 A major difference in the structures of T4P and TTSS is the type IV pilus extends beyond
181 the outer membrane (Gram-negative) or PG layer (Gram-positive, (8)) but in the TTSS in Gram
182 negative bacteria the pilus does not (Fig. 1 and (2)). Our original model for a TTSS in Gram-
183 positive bacteria is that the pilus (i.e., PilA3) does not extend past the PG layer (Fig. 1). To test
184 this model, we exposed *C. perfringens* cells to rabbit anti-PilA3 antibodies and a fluorescently
185 tagged goat anti-rabbit secondary antibody and found that we could detect binding localized to the
186 poles of the cells (Fig. S1). This refined the location of the PilA3 pili to the poles, and that they
187 extend far enough out past the PG layer to be exposed to antibodies, which suggests the model
188 shown in the right most panel in Fig. 1 may need to be modified to show extension of the pilus to
189 the outside environment.

190 To determine if there are generalized secretion defects in T4P mutants, we examined
191 whether each mutant in the T4P-associated genes was able to secrete another protein, the
192 phospholipase-sphingomyelinase, PLC. Secretion of PLC can be easily detected by a ring of
193 precipitation around a colony using egg yolk agar, and we found that none of the T4P mutants
194 showed a detectable decrease in PLC secretion (Fig. S2), suggesting the secretion defects seen
195 with BsaC were not due to generalized secretion defects.

196 **Expression of the *bsa* operon leads to synthesis of extracellular biofilm matrix material.** Since
197 the *bsa* operon encodes proteins with similar synteny and predicted structural homology to the

198 *sipW-tasA-tapA* biofilm matrix operon in *B. subtilis* (15), we placed the entire operon (*sipW-bsaA-*
199 *bsaB-bsaC*) under control of the lactose inducible promoter for high level expression. The cells
200 produced large quantities of an extracellular gel like material (Fig. S3A), which, after staining with
201 crystal violet, appeared as amorphous material between the bacterial cells (Fig. S3B). A His_6 tag
202 was added to the C-terminus of the BsaA protein and, when the operon was expressed, a distinct
203 high molecular weight ladder of BsaA was visible in anti-His_6 antibody western blots after running
204 on SDS-PAGE gels using samples boiled in SDS-PAGE buffer (Fig. S3C). This suggested BsaA
205 was in very heat-stable oligomers. This was in agreement with the results of Obana et al who found
206 that BsaA was the major component of the biofilm matrix and that BsaA oligomers were stable
207 after boiling in SDS-PAGE buffer and resistant to dissociation by high levels of SDS and formic
208 acid (11).

209 **Efficient secretion of the BsaA oligomer is dependent on some T4P-associated genes.** Since it
210 was possible that, like BsaC, BsaA in its oligomer form was secreted with the aid of T4P-associated
211 proteins, we expressed a clone containing just the *sipW* and *bsaA* genes in which a FLAG-tag was
212 added to the C-terminus of BsaA for quantification in western blots using densitometry. Only *sipW*
213 and *bsaA* were included to simplify the analysis of the resulting western blots. An example of a
214 typical blot is shown in Fig. 5A. Using this method, we found that the *pilA1*, *pilB1*, *pilB1/pilB2*,
215 *pilN* and *pilO* mutants exhibited decreased levels of BsaA oligomers in the supernatants (Fig. 5B).
216 However, since the genes required for secretion lacked an inner membrane core protein (*pilC*)
217 component and pre-pilin peptidase, which are essential for T4P functions (17), the results were not
218 definitive.

219 **Models of BsaA oligomers predicts an isopeptide bond is formed linking adjacent monomers.**
220 BsaA forms very heat and SDS resistant oligomers ((11) and Figs. S3C and 5A). To investigate the

221 structural nature of the interactions between monomers of BsaA that leads to this stability, a
222 structure would need to be developed of a homomultimer. Due to the difficulty of solving the
223 structure of long, thin chains by experimental methods, computational modelling was used. Models
224 were initially created using Phyre2 (18), with the top model being an x-ray crystal structure of
225 TasA from *B. subtilis* (13). That structure, and homology models from that structure, were
226 monomeric, and not useful for determining intermolecular interactions between monomers.
227 AlphaFold Multimer was then used to create a multimeric model to determine inter-chain
228 interactions. Using ColabFold 1.3.0 (19), which is based on AlphaFold 2.2.0 (20), a homotrimeric
229 structure of BsaA was created (Fig. 6A). The SipW-specific signal sequences (residues 1-27) were
230 included in the initial model, but only formed disordered loops which do not have any notable
231 interactions. Removing the signal sequences did not change the overall structure of the monomers
232 so this was done for subsequent analyses.

233 In monomeric structures, residues 28-38 would extend out as disordered loops (Fig. 6A,
234 monomer 1). In the multimeric structures, these residues incorporated themselves into the beta
235 sheet structure of adjacent monomers (Fig. 6A). Being fully incorporated into a nearby structure
236 indicates that residues 28-38 provide specific interactions with the receiving monomer, providing
237 the basis of a mechanism for which oligomers of BsaA can be created. For example, V33 and F37
238 of the interdigitated strand become part of the hydrophobic core of the receiving monomer (Fig.
239 6B). The interactions between specific side chains of residues from the interdigitated strand and
240 the receiving monomer were investigated further for less standard interactions.

241 The *de novo* created structures indicate a bond between two side chains: N35 of the
242 interdigitated strand and K74 of the receiving monomer (Fig. 6C). At first, this was thought to be
243 an error in AlphaFold's model generation, since the chemistry would not be possible. The bond

244 was created between the oxygen of N35 and the nitrogen of K74, resulting in three bonds in the
245 oxygen. However, this bond was created in nearly every model generated, so it was investigated
246 further. A crosslink between asparagine and lysine is a type of isopeptide bond, which has been
247 characterized through a few different mechanisms and can be found as an intermolecular crosslink,
248 as in ubiquitylation (21), or as an intramolecular crosslink found in Gram-positive pili (22). In
249 Gram-positive pili, this bond is formed between an asparagine and a lysine, autocatalyzed by a
250 glutamate within the hydrophobic core of the protein. Residues equivalent to these are found
251 surrounding the bond in BsaA (Fig. 6C). The hydrophobic region surrounding the N35-K74 pair
252 (Fig. 6B, brown colored residues) is believed to be necessary to help the catalytic cross-linking
253 reaction (23).
254 It is unclear why this bond was created by AlphaFold. In the documentation of AlphaFold Multimer
255 (24), there are no mentions of predicting crosslinks. One potential reason is that it generated the
256 side chains so close to each other there was a steric clash, but the rest of the model was so favorable
257 the structure was created regardless. Models created in ColabFold have an option to perform
258 energy minimization using Amber. For these models, this was not performed. Recreating this
259 model while enabling energy minimization eliminates the bond every time. This is expected given
260 the incorrect chemistry of the bond. AlphaFold3 (25) does not include a minimization step, and,
261 like AlphaFold2, creates the isopeptide bond.
262 This amyloid-like fold and interdigitation is seen in other Gram-positive bacteria, such as
263 TasA in *B. subtilis*. A cryo-EM structure has been created for a chain of TasA (12), which
264 demonstrated how hydrophobic interactions between donor and acceptor monomers drive
265 interdigitation. No crosslinks were found or mentioned, and the required triad of N-E-K residues
266 is not present in the structure (PDB 8aur). Therefore, the isopeptide bonds in BsaA oligomers

267 would be the first example of an intermolecular bond of this mechanism found in this family of
268 proteins.

269 **Mutations in isopeptide bond forming residues abolish oligomerization of BsaA.** The
270 AlphaFold Multimer models predicted an isopeptide bond was formed between N35 and K74 from
271 adjacent residues in a BsaA oligomer. To test this computational prediction, we made point mutant
272 substitutions of each residue to alanine, forming N35A and K74A mutants, along with a double
273 N35A/K74A mutant. His-tagged versions of these mutant forms were expressed along with *sipW*
274 and the cell extracts and culture supernatants were examined for the presence of the mutant
275 proteins in oligomeric form. Mutating N35 or K74 to alanine residues resulted in a loss of
276 oligomers in the supernatants and the appearance of monomers of the K74A mutant but not of the
277 N35A mutant, suggesting the N35A version was not effectively released by the cell (Fig. 7A).
278 Monomeric forms of both mutants appeared in the cell extract (Fig. 7A). In some experiments,
279 small amounts of dimers and higher mol wt forms of BsaA appeared in the cell pellet, but the large
280 majority were in monomeric form, which was absent in the WT form of BsaA (Fig. 7A), perhaps
281 due to the stability of oligomers even in the absence of an isopeptide bond. To test this, we
282 suspended cell pellet extracts from BsaA-His_6 and N35A-His_6 in SDS-PAGE buffer and heated at
283 95 C for 0, 10, 20 and 40 min before running the samples on SDS-PAGE gels and performing
284 western blots. The unheated samples were incubated for 40 min at room temp as a control. The
285 N35A-His mutant maintained stable oligomers in SDS-PAGE buffer and after being run through
286 an SDS-PAGE gel in the absence of heating, but heating the samples for 10-40 min led to
287 dissociation of the oligomers, in contrast to the WT BsaA protein, which was not affected by
288 heating (Fig. S4). This suggests that N35, and therefore the isopeptide bond, is essential for
289 forming detergent and heat resistant oligomers, and that the N35 monomer-monomer interactions

290 are stable even in the presence of 2% SDS (the concentration in the SDS-PAGE buffer), which is
291 in contrast to oligomers formed by the *B. subtilis* TasA protein, which were dissociated by SDS
292 (14).

293 We also tested the N35A/K74A mutant and observed that it showed a similar loss of
294 oligomerization as did the N35A mutant (Fig. 7B, left four lanes). These experiments were all done
295 in strain HN13 (WT) background where there was an intact *bsaA* gene. To determine if this
296 potential source of BsaA, even though it was not His-tagged, affected the results we observed with
297 the mutants, we expressed the N35A and N35A/K74A mutants in a strain with an in-frame deletion
298 of the *bsaA* gene, but did not observe any differences to the pattern seen in the strain with an intact
299 *bsaA* gene (Fig. 7B, right four lanes), suggesting the chromosomal source of BsaA was not forming
300 oligomers in combination with the N35A and N35A/K74A mutant forms of BsaA.

301 **The Glu56 residue of BsaA is essential for efficient isopeptide bond formation.** Spontaneous
302 or self-catalyzed isopeptide bond formation is catalyzed by an acidic Glu or Asp residue located
303 within the hydrophobic pocket where the Asn and Lys residues are located (23). In AlphaFold 2
304 BsaA models of dimer interfaces, Glu56 is predicted to be in the precise location to catalyze bond
305 formation (Fig. 6C). To test if this was actually occurring, we constructed an E56A substitution in
306 the BsaA protein and examined its ability to oligomerize by using western blotting on cell extracts
307 and concentrated supernatants from the mutant and WT strain (Fig. 7C). The E56A mutant showed
308 a clear lack of oligomer formation and an accumulation of the monomeric form of the protein
309 which was not observed with the WT form of BsaA (Fig. 7C), providing strong evidence that
310 Glu56 is responsible for catalyzing isopeptide bond formation in BsaA.

311 **A FAST fusion to BsaA displays surface exposed localization.** To investigate the cellular
312 location of the BsaA oligomers, a FAST gene fusion to the *bsaA* gene was constructed and, along

313 with the *sipW* gene, was placed under control of the lactose-inducible promoter. The FAST protein
314 binds a soluble dye which, upon binding, exhibits a large increase in fluorescence (26-28) allowing
315 localization of the fusion proteins. We expressed the *sipW-bsaA*-FAST fusion genes and exposed
316 the cells to two different dyes, Coral, which is membrane permeable, and Amber-NP which is
317 impermeable to membranes (both obtained from The Twinkle Factory). To test if the Coral and
318 Amber-NP permeability profiles were working in *C. perfringens* membranes, we expressed a *pilT*-
319 FAST gene fusion, which should be located only in the cytoplasm, and exposed them to the Coral
320 and Amber-NP dyes. Consistent with the predicted cytoplasmic location of the *pilT*-FAST fusion,
321 only the Coral dye showed significant fluorescence after treatment (Fig. 8A). However, when the
322 *bsaA*-FAST fusion was expressed along with *sipW*, both dyes were able to bind and fluoresce (Fig.
323 8B), indicating the BsaA-FAST fusion was competent for secretion in *C. perfringens* and was
324 located on the surface of the bacteria.

325 **BsaA oligomers are anchored to the membrane and not the peptidoglycan in strain HN13.**
326 Because the BsaA-FAST fusion experiments indicated BsaA oligomers were located on the cell
327 surface and SEM images of the BsaA matrix proteins suggested they were in the form of fibrils
328 associated with the bacterial surface (11), we wanted to determine if the oligomers were embedded
329 in the cytoplasmic membrane or covalently linked to the PG layer. To test this, we expressed *sipW*-
330 *bsaA*-His_6 constructs for the WT and each mutant in our lactose-inducible promoter system and
331 then boiled the cells in 10% SDS for 30 min, followed by digestion of the PG matrix by lysozyme.
332 Our experimental approach was based on the hypothesis that if BsaA was anchored in the
333 membrane, boiling in SDS would release it from the cell but if it was covalently anchored to the
334 PG, it would be freed only after digestion with lysozyme. We tested the WT, N35A, K74A and
335 N35/K74 mutants and found that all of the long wild-type oligomers were released after the SDS

336 boiling step, while the cells appeared to retain a small number of monomers and dimers of BsaA
337 (Fig. S5). However, after treatment with lysozyme, the monomers and dimers were no longer
338 detected. For the mutants that can't oligomerize, the monomer/dimers were retained after boiling
339 in SDS, but not after lysozyme treatment. We interpret these findings as indicating the oligomers
340 are anchored to the membrane but after SDS boiling they are released. A small number of
341 monomers and dimers remain trapped inside the sacculi produced by boiling in SDS but lysozyme
342 treatment releases these from the sacculi. However, they are then too dilute to detect on western
343 blots. This is consistent with an analysis of the *bsaA* and *bsaB* protein sequences, both of which
344 lack a recognizable LPXTG motif, despite the presence of a sortase gene, *srtB*, adjacent to the
345 operon (Fig. 3C).

346 **BsaB, an ortholog of BsaA, forms dimers but not long oligomers.** Previous results by Obana et
347 al (11) showed a *bsaB* mutant still made significant amounts of the BsaA oligomers but they did
348 not test to see if BsaB was incorporated into BsaA oligomers. Therefore, we added an HA-tag to
349 the BsaB protein and expressed the *sipW*-*bsaA*-*bsaB*-HA-*bsaC* genes in our inducible system and
350 examined the culture supernatants for BsaB-HA using western blots. Based on the predicted
351 molecular weight of the BsaB protein, we detected BsaB-HA only in a monomeric and a higher
352 molecular weight form which could be either a homodimer or a BsaA-BsaB heterodimer (Fig. S6),
353 but it is not covalently attached to the longer BsaA oligomers.

354 **Phylogeny of isopeptide bond formation in the TasA family.** Since a BlastP search with BsaA
355 only identifies homologs in closely related Clostridial species, an iterative search with profile
356 hidden Markov models was employed. This analysis resulted in the identification of 2,172 proteins,
357 including the previously described TasA from *B. subtilis* and CalY from *Bacillus cereus* (Fig. 9).
358 Of these, 632 hits belong to the "peptidase M73, camelysin" (InterPro: IPR022121) family; so-

359 called because of the likely misidentification of CalY as a peptidase. In addition to the use of
360 probabilistic models, the relatedness of these proteins is evident at the structural level. Like other
361 family members, BsaA is predicted to adopt a β-sheet–rich Ig-fold-like structure and is predicted
362 to participate in donor-strand exchange (Fig. 6), as described for TasA (12). These analyses suggest
363 that while the structures of these proteins and homopolymer assembly are highly conserved,
364 sequences are highly divergent.

365 In addition to BsaA, we identified three other uncharacterized proteins in *C. perfringens*
366 (BsaB, CPE2262 and CPE0554), but these paralogs are not expected to form an intermolecular
367 isopeptide bond, because the N-E-K triad is missing. Nevertheless, over 100 BsaA-like proteins in
368 other species were identified with the N-E-K triad, suggesting conservation of the donor-strand-
369 exchange-and-lock mechanism. These proteins are largely from Clostridia with some homologs
370 from *Coprobacillaceae*, *Enterococcaceae*, and *Thermococcus* (Fig. 9).

371 Based on the putative functions of gene neighbors, the BsaA/TasA family members appear
372 to be involved in the assembly of surface filament structures (Fig. 9C). They are often in putative
373 operons with genes encoding proteins containing the Gram-positive cell wall anchor motif LPxTG,
374 fibronectin domains, SpaA domains, choice-of-anchor A domains, as well as signal peptidase and
375 sortase domains. Multiple BsaA/TasA paralogs can be found in a single putative operon (Fig. 9C),
376 suggesting the possibility of donor-strand exchange assembly of heteropolymers.

377

378 **Discussion**

379 The original goal of this work to determine if Gram-positive bacteria have the functional
380 equivalent of TTSS that are found in Gram-negative bacteria. Since TTSS have a similar overall
381 architecture and mechanism as T4P (Fig. 1), it made sense to test a Gram-positive bacterium known

382 to have T4P. With a few notable exceptions (e.g., *Streptococcus sanguinis* (29-31)) amongst Gram-
383 positive bacteria, T4P are confined within the Clostridia (9) so *C. perfringens* was chosen as
384 representing a typical clostridial species. The overall concept was that the thick Gram-positive PG
385 layer represents a similar barrier to protein secretion to the environment as the Gram-negative
386 outer membrane. Since TTSS depend upon a pseudopilus to function, we tested mutants with
387 deletions in each of the four main pilin-encoding genes in *C. perfringens* for changes in the
388 secretome compared to a WT strain. We identified one protein BsaC (CPE0517) that was depleted
389 in the secretome of a *pilA3* deletion strain (Fig. 3A-B). We then examined a bank of in-frame
390 deletion mutants in all of pilin-associated genes (Fig. 2) and found that, for BsaC secretion, a
391 homolog of each of the essential proteins needed for a type IV pilus assembly, PilA3 (pilin), PilB2
392 (assembly ATPase), PilC1 (inner membrane core protein), PilD (prepilin peptidase), and PilM-N-
393 O, (inner membrane accessory proteins) (Fig. 2) were required for efficient secretion (Fig. 4). This
394 appears to be specific for BsaC secretion, since the secretion of the toxin PLC was not affected by
395 mutations in any of the T4P-associated genes (Fig. S2). While this may indicate the presence of a
396 TTSS like apparatus in *C. perfringens*, the lines separating T4P- and T2SS-dependent secretion
397 have been blurred by the discoveries that *Pseudomonas aeruginosa* and *Dichelobacter nodosus*,
398 Gram-negative bacteria, have T4P systems capable of mediating protein secretion (32, 33).
399 However, further analysis of the structure and components of these Gram-negative pilus-dependent
400 secretion systems have not been published. Since PilA3 pilin was detected on the surface of *C.*
401 *perfringens* (Fig. S1), it is possible the PilA3-dependent secretion system acts more like a T4P.
402 BsaC monomer secretion may appear to be independent of BsaA oligomers, but it is
403 possible that they are being secreted simultaneously, since we did not track BsaA and BsaC
404 secretion at the same time. The biological role that BsaC plays in biofilm matrix functions is

405 unknown, but it is intriguing that the protein contains a VWA domain, which is widespread in
406 eukaryotes, bacteria and archaea and is associated with protein-protein interactions (34).
407 Therefore, the role BsaC plays in biofilm functions is also currently under investigation in the
408 laboratory.

409 As noted above, BsaC is in an operon with genes encoding a signal peptidase, SipW and
410 two proteins (BsaA and BsaB) with homology to the main biofilm protein of *B. subtilis*, TasA (Fig.
411 3C). Because this gene synteny is similar to the *sipW-tasA-tasB* biofilm operon in *B. subtilis*, we
412 expressed the entire operon to see if a biofilm matrix was produced and noted that high levels of a
413 gel like matrix material was formed in the culture supernatant (Fig. S3A-B). A high mol wt
414 oligomer of BsaA was seen in SDS-PAGE gels (Fig. S3D), which was similar to that seen by Obana
415 et al in their analysis of the same operon (11). Because BsaC was dependent on T4P-associated
416 proteins for secretion, it was logical to determine if secretion of the BsaA oligomer was as well.
417 We used densitometry of western blots of the BsaA oligomers in culture supernatants that had been
418 concentrated by TCA precipitation. Although we saw significantly decreased levels of secretion
419 in several T4P mutants, it did not give conclusive evidence that a full T4P system was needed for
420 secretion due to the absence of a PilC or PilD component. This experimental approach was
421 hindered by relatively high amounts of experimental variability in trying to measure an oligomer
422 in solution; this can be seen by the high range of the SEM determined for most of the samples (Fig.
423 4) despite multiple repetitions of each experiments. However, the fact that several genes in each
424 type of experiment was shown to be necessary for efficient secretion does strongly suggest that
425 T4P proteins play a role in the process.

426 Although we observed that a *pilA1* deletion strain did not secrete the biofilm matrix
427 oligomers efficiently (Fig. 5), we did not observe this phenotype in our initial screening for

428 proteins absent in the secretome, as shown for PilA3 in Fig. 3A-B. We think this is likely due to
429 the fact that our growth conditions for the experiment, overnight culture in BHI liquid medium at
430 37 ºC, is not optimum for production of the biofilm matrix proteins, as described by Obana et al
431 (11). These findings point out the limitations of screening for T4P-dependent secretion: using only
432 a single type of environmental condition may not be the one in which a protein is made and secreted
433 so it would not be discovered. Even 2-D gels would be unlikely to have uncovered the presence of
434 the BsaA oligomers due to their extremely high mol weight. Therefore, it seems protein-based
435 mass spectrometry on samples from a wide variety of environmental conditions would be optimal
436 for detecting T4P-dependent secretion substrates in future experiments.

437 Since our results and that of Obana et al both showed that oligomers of BsaA were
438 extremely resistant to heat and other denaturants (Fig, S3C and (11)), this suggested that covalent
439 bonding might be involved. Subsequent analysis using AlphaFold 3 indicated there were likely
440 isopeptide bonds being formed by donor strand exchange involving an N35 of one monomer and
441 a K74 of another, catalyzed by an E56 residue in the recipient monomer (Fig. 6). This model was
442 confirmed by subsequent substitution of each residue involved with an Ala, which led to a
443 complete lack of oligomerization (Fig. 7). Isopeptide bonds are formed relatively slowly (23) so
444 we hypothesized that the donor strand exchange resulted in formation of a temporally stable
445 interaction that gave the isopeptide bonds time to form. However, based on experiments with the
446 N35A mutant, the oligomers are stable even without the isopeptide bond, as the mutant version
447 was resistant to denaturation in SDS-PAGE buffer and being run through an SDS-PAGE gel. Only
448 after heating in the SDS-PAGE buffer at 95 ºC did the N35 mutant depolymerize (Fig. S4), while
449 the WT oligomer withstood this treatment. We further characterized the formation of the oligomer
450 and showed, using FAST fluorescent protein fusions, that the oligomers were anchored to the

451 surface of the bacteria (Fig. 8) but via the cytoplasmic membrane and not directly to the PG layer
452 (Fig. S5).

453 BsaB, the other biofilm protein encoded in the operon, did not form oligomers, did not
454 become covalently linked to BsaA oligomers (Fig. S6) and both BsaA and BsaB have no apparent
455 LPXTG motif present in their protein sequence (data not shown). These structural features seem
456 somewhat at odds with the presence of a sortase gene (*srtB*) adjacent to, but oriented in the opposite
457 direction, of the *sipW* gene (Fig. 3C). Obana et al (11) tested the premise that SrtB was responsible
458 for oligomerization of BsaA, but found that a *srtB* mutant had normal levels of oligomer present
459 (11). However, examination of the *bsa* operon in other strains of *C. perfringens* reveals that a
460 different gene synteny is in place, where *sipW* is followed by a *bsaA* homolog and two *bsaB*-like
461 genes, the first of which encodes an identifiable LPXTG motif (Fig. 3D). The BsaB sequence from
462 strain 13, when subjected to a BLAST analysis (35) against other *C. perfringens* strains, shows
463 proteins with high levels of homology to either the N-terminal half or C-terminal half of the strain
464 13 BsaB, but not both. This suggests strain 13 underwent a genetic rearrangement event in which
465 the two *bsaB* homologs were recombined into a single gene with the loss of the LPXTG motif
466 found in the first gene (Fig. 3C-D). This brings up the distinct possibility that the SrtB sortase
467 could catalyze the cross-linking of the first BsaB ortholog to the PG layer in strains other than
468 strain 13. What role BsaB plays in BsaA oligomerization is still unknown, since a *bsaB* mutant
469 still formed similar amounts of oligomers of BsaA as the wild-type strain (11). However, based
470 on the differences in composition of BsaB homologs between strain 13 and other strains (Fig. 3C-
471 D), it is possible that BsaB is involved in anchoring the BsaA oligomers to the PG layer via the
472 activity of SrtB and the BsaB homolog with an LPXTG motif, but this is speculative.

473 Taking all these results into consideration, we have constructed a model showing the
474 mechanisms underlying the assembly of BsaA into cross-linked oligomers (Fig. 10). In this model,
475 BsaA monomers are secreted via the Sec pathway, fold on the outer surface of the CM but are
476 anchored in the membrane by their N-terminal signal sequence. An oligomer of BsaA with the
477 most recent monomer having its N-terminal disordered region exposed, contacts the newly
478 secreted monomer and undergoes a donor strand exchange in which the N-terminal disordered
479 region binds to the exposed beta-sheet fold and interdigitates as a newly formed beta sheet (Fig.
480 6A). This is followed by cleavage of the signal sequence by the signal peptidase SipW. We believe
481 the SipW acts after oligomer binding to the newly secreted monomer because the BsaA oligomers
482 appear to be anchored to the membrane (Fig. S5) and the oligomers are in a relatively stable
483 assembly even before the monomers are "locked" together by isopeptide bond formation (Fig. S4).
484 After SipW cleavage, the oligomer is released and can then be used to add another monomer or be
485 secreted through the PG layer to the external medium. The presence of significant amounts of
486 oligomers attached to the bacterial surface (Fig. 8) suggests the monomers assemble in the space
487 between the CM and PG rather than being secreted as monomers for oligomerization in the
488 environment. The mechanism for export of the oligomers through the PG layer is not complete but
489 we provided evidence that T4P play at least some role in the process (Fig. 5).

490 The phylogeny of BsaA-like proteins shows they are widespread amongst Gram-positive
491 bacteria but also present in some Gram-negative bacteria and archaea (Fig. 9). However, the
492 homologs that are predicted to form isopeptide bonds are more narrowly focused in the Clostridia
493 and Enterococcus genera (Fig. 9). This brings up the question as to what advantage isopeptide
494 bond formation bestows on the biofilm matrix properties of the bacteria that form them. Formation
495 of isopeptide bonds can greatly increase the mechanical resistance to distension and separation of

496 the BsaA fibers. As an example, the intramolecular isopeptide bonds formed in sortase-dependent
497 pili made by Gram-positive bacteria have been shown to greatly increase the amount of force that
498 needs to be applied to distend the individual subunits, thereby strengthening the entire fiber (36).
499 To trap macrophages, neutrophils or other phagocytic cells, *C. perfringens* and other species may
500 have evolved a mechanism to form isopeptide bonds to increase the mechanical strength of the
501 fibers, since macrophages, neutrophils and other phagocytic cells can impart a large locomotive
502 force (1.9 to 10.7 nN) during their motility (37). It is also worth noting that *C. perfringens* is an
503 extremely ubiquitous bacterium in nature, found in soil, freshwater sediments and even extreme
504 environments like Antarctica (38, 39) and likely forms biofilms in these natural environments. The
505 BsaA biofilm matrix fibers could protect them from killing by amoeba, protozoa and other
506 predators found in these environments. Analyzing the relative mechanical strength of fibers with
507 and without isopeptide bonds will help answer some of these questions.
508

509 **Materials and Methods**

510 **Bacterial strains and growth conditions.** Bacterial strains, plasmids, and primers used in this
511 study are listed in Table 1. *Escherichia coli* strain DH10B was grown in Luria Bertani broth at
512 37°C for all transformations. When necessary, kanamycin and chloramphenicol were added to the
513 media at a concentration of 100 μg/ml and 20 μg/ml, respectively. *C. perfringens* strain HN13, a
514 Δ*galKT* derivative of strain 13, was used as the wild type strain in this study. *C. perfringens* strains
515 were grown anaerobically in PGY (30 g proteose peptone #3, 20 g glucose, 10 g yeast extract, 1 g
516 sodium thioglycolate per liter) or brain-heart infusion (BHI) (Thermo Fisher) in an anaerobic
517 chamber (Coy Laboratory Products, Inc.). Anaerobic egg yolk agar medium was prepared as

518 previously described (https://www.fda.gov/food/laboratory-methods-food/bam-media-m12-
519 anaerobic-egg-yolk-agar#).

520 **Construction of in-frame deletions of T4P-associated genes.** In-frame deletions of T4P-
521 associated genes were made using the method of Nariya et al. (40), modified as described in
522 Hendrick et al (41). The primers used to amplify the flanking DNA for each gene are listed in
523 Table S1. All deletions were confirmed by PCR across the deleted region. Complementation of
524 the in-frame *pilA3* mutant (Fig. 3A-B) was done using the lactose-inducible promoter vector
525 pKRAH1 (16) to express the *pilA3* gene. The construct was made by PCR using the primers listed
526 in Table S7.

527 **SDS-PAGE and staining of secretome proteins.** Cultures of strain HN13 were grown for 16 h
528 in BHI medium (i.e., into stationary phase) and the bacteria were removed by centrifugation in a
529 table top centrifuge. The culture supernatants were removed and passed through a 0.45 μM filter.
530 Proteins in the culture medium were precipitated using 10% trichloroacetic acid (TCA)
531 precipitation and washed twice in acetone. After drying, sample pellets were resuspended in water
532 and then 4X SDS-PAGE sample buffer (42) was added to a final concentration of 1X. Secretome
533 samples were separated on an SDS-PAGE gel and then stained with the fluorescent protein stain,
534 Sypro-Ruby, according to the manufacturer's instructions (Thermo Fisher). Stained gels were
535 imaged with a Gel-Doc Imaging system (Bio-Rad, Inc).

536 **Cloning the *bsa* operon into the lactose-inducible expression vector, pKRAH1.** Because the
537 SipW signal peptidase is likely needed for processing the secretion signal on BsaC (11), to analyze
538 secretion of BsaC in T4P gene associated mutants, the entire *sipW-basA-bsaB-bcaC* operon was
539 expressed under control of the lactose inducible promoter in plasmid pKRAH1 (16) using primers
540 OHL145 (PstI site) and OHL146 (BamHI site) by PCR (Table 1 and Table S3). The PCR fragment

541 was ligated to the PCR cloning vector pGEM-T to make pHLL64. pHLL64 and pKRAH1 were
542 digested with PstI and BamHI and ligated to make pHLL65.

543 **Construction of a FAST protein fusion to BsaA.** The gene fusion was constructed using
544 overlapping PCR of the FAST coding region on plasmid pAG104 FAST (26) to the *bsaA* gene
545 linked to *sipW* (i.e., the *bsaB* and *bsaC* genes were absent). The FAST protein was fused to the C-
546 terminal end of BsaA because of the processing of the signal sequence on the N-terminal end. The
547 primers used to construct the gene fusion are listed in Table S5.

548 **Production of antibodies to BsaC, PilA3 and immunofluorescence microscopy.** Rabbit
549 polyclonal antibodies against the BsaC peptide sequence N-CNIEGDKISWDRGEN-C were
550 produced by GenScript USA, Inc. along with pre-immune serum as a negative control. Rabbit
551 polyclonal antibodies against the PilA3 peptide sequence N-CIKIDNSKDSIDLTR-C were
552 produced by Yenzyme Antibodies, LLC. along with pre-immune serum as a negative control. To
553 test for surface exposure of PilA3 pilins, the bacteria were fixed with 2.5% paraformaldehyde in
554 Dulbecco's phosphate buffer saline (DPBS), incubated with rabbit serum with anti-PilA3
555 antibodies, washed with DPBS and incubated with goat anti-rabbit antibodies fluorescently tagged
556 with Alexafluor 647 (Invitrogen). Stained cells were imaged on an Olympus IX71 fluorescent
557 inverted microscope equipped with a CoolSnap HQ2 CCD camera and DeltaVision deconvolution
558 and image analysis software.

559 **Quantitative immunoblotting of secreted BsaC.** Overnight cultures were grown in BHI
560 supplemented with chloramphenicol and diluted 1:50 into fresh BHI medium pre-incubated at
561 37°C. The cultures were grown for one hour then induced by adding lactose, to a final
562 concentration of 10 mM, to the growth media every 30 min for 3 h. Following induction, 1 mL of
563 culture supernatant was collected by centrifugation, and proteins precipitated using TCA, as

564 described above. Resulting precipitates were resuspended in 50 uL of SDS-PAGE loading buffer
565 with 100 mM DTT, the pH was adjusted using one μL of 1.5 M Tris-HCl, pH 9, if necessary. Five
566 μL of sample was loaded per well, and SDS-PAGE was performed. Following SDS-PAGE,
567 proteins were transferred to a polyvinylidene difluoride (PVDF) membrane utilizing the iBlot
568 system (Bio-Rad, Inc), and developed with serum containing anti-BsaC antibodies at 1:250
569 dilution, followed by a Dylight 550 labeled fluorescent secondary goat anti-rabbit antibody
570 (Thermo-Fisher). Imaging was performed on a Typhoon gel scanner, and mean pixel intensity
571 was measured for each band utilizing the Amersham Typhoon densitometry control software.
572 **Purification of BsaC.** The *bsaC* gene with a C-terminal His_6 tag was cloned into the expression
573 vector pET24a by PCR using the primers listed in Table S4. An overnight culture of *E. coli*
574 LOBSTR pET24a::BsaC-His_6 was used to inoculate 2 L of 37ºC pre-warmed terrific broth in a 4
575 L beveled flask. The culture was allowed to grow to 0.6 OD_{600} where it was induced with 1 mM
576 IPTG. After 6 hours of induction, the culture was centrifuged to remove cells (BsaC-His_6 is
577 secreted by *E. coli*), and the resulting supernatant was precipitated with 90% ammonium sulfate.
578 Precipitated protein was solubilized in PBS and used to load on to an ÄKTA purifier FPLC for Ni-
579 NTA affinity chromatography in PBS with 10 mM imidazole as loading buffer, and PBS with 1M
580 imidazole as elution buffer. Eluted fractions were concentrated with an Amicon Ultra 2 mL 10 K
581 cutoff centrifugation columns. Concentrated elution fractions were then further processed using a
582 Superose 6 gel filtration column linked to an ÄKTA purifier FPLC.
583 **Immunoblotting of BsaA oligomers.** BsaA oligomer secretion by *C. perfringens* was analyzed
584 using western blots. For the immunoblot experiments, the proteins were blotted onto a PVDF
585 membrane while submerged in cold buffer which contained 3 mM Na_2CO_3, 10 mM $NaHCO_3$ and
586 20% methanol.

587 **Immunoblotting Method One.** The membranes were placed in a Snap i.d. 2.0 system
588 (Millipore Sigma) which uses a vacuum system to pull solutions through the membrane (referred
589 to as washing steps below). Once set up, a solution containing Tris-buffered saline (TBS) (pH 7.4),
590 0.1% Tween 20 (TBST), and 1% bovine serum albumin (BSA) was added as a blocker. Both
591 primary and secondary antibodies were added to separate aliquots of the solution containing TBST
592 and 1% BSA. In the first step, the primary antibody solution was incubated for 10 minutes on the
593 membrane, followed by three washings with TBST. Then the secondary antibody solution was left
594 to sit for 10 minutes on the membrane, followed by three washings with TBST. The primary
595 antibody utilized was OctA-probe H-5 (1:250 dilution) used to detect the FLAG epitope (Santa
596 Cruz Biotechnology). The secondary antibody utilized for these experiments was StarBright700
597 goat anti-mouse (Bio-Rad) (1:10,000 dilution). The fluorescently labeled immunoblots were
598 imaged using a ChemiDoc MP imaging system (Bio-Rad), and the intensity of the fluorescence
599 was measured using the ChemiDoc MP imaging software for densitometry measurements.
600 **Immunoblotting method 2**. Membranes were transferred to a container and blocked with
601 EveryBlot blocking Buffer (Bio-Rad). Antibodies were diluted with EveryBlot blocking buffer and
602 applied to membranes, followed by a minimum of five washes with TBST.
603 The primary antibodies utilized included OctA-probe H-5 (1:227 dilution) used to detect
604 FLAG, and HA-Tag F-7 (1:227 dilution) used to detect HA (both from Santa Cruz Biotechnology,
605 Santa Cruz, CA). Membranes submerged in EveryBlot blocking buffer with primary antibody were
606 incubated at 4°C overnight. The secondary antibody utilized for these experiments was
607 StarBright700 goat anti-mouse (Bio-Rad) (1:13,333 dilution), incubated for one hour at room
608 temperature. Images of the immunoblots were taken using a ChemiDoc MP imaging system (Bio-
609 Rad), to measure the intensity of the fluorescence.

610 **Western blots on *bsaA* mutants.** Western blots to detect oligomerization of BsaA mutants were
611 performed by expressing the *sipW-bsaA-His*$_6$ genes under control of the lactose-inducible
612 promoter. Supernatants were separated from cell pellets by centrifugation. The proteins in the
613 supernatant were concentrated by TCA precipitation (as described above), suspended in SDS-
614 PAGE loading buffer with 100 mM DTT and heated at 95 °C for 20 min. The bacteria in the cell
615 pellet were suspended in PBS and disrupted by bead beating with 0.1 mm zirconia-silica beads.
616 The beads were then removed by centrifugation and the cell extract brought to 1 X SDS-PAGE
617 buffer with 100 mM DTT and heated at 95 °C for 20 min. After running on SDS-PAGE gels, the
618 proteins were transferred to a PVDF membrane and developed using immunoblotting method 2
619 above, except the primary antibody was mouse monoclonal anti-His6 antibody His.H8 (Santa Cruz
620 Biotechnology).
621 **Modeling BsaA Structure.** BsaA from *C. perfringens* (accession number WP_011009871) was
622 created as a homotrimer using ColabFold v1.3.0 (19) based on AlphaFold v2.2.0 (20). Amber
623 minimization was not enabled. In the top ranked model based on pLDDT PyMOL v2.5.0 was used
624 to analyze side chains of interdigitated beta strands for potential sidechain crosslinks. Images were
625 created using ChimeraX v1.2. (43, 44) The same trimer of BsaA was created in AlphaFold 3 (25)
626 to observe if the new version creates crosslinks, and all trimer models made did contain the
627 crosslinks. Structures of BsaA were created with AlphaFold 3 (25).
628 **Phylogeny of BsaA-like proteins.** BsaA-like protein sequences were identified with 3 iterations
629 of jackhmmer (45) (HmmerWeb version 2.41.2) (46) using BsaA (UniProt:Q8XN23) from *C.*
630 *perfringens* as the seed. The protein sequence similarity network of proteins identified either with
631 jackhmmer or as a match to IPR022121 was constructed using the EFI-EST tool
632 (https://efi.igb.illinois.edu/efi-est/) (47, 48) with an alignment score of 10. Nodes were collapsed

633 at a sequence identity of 100%. The network was visualized with Cytoscape
634 (https://www.cytoscape.org) (49) using the Prefuse Force Directed OpenCL Layout. Gene
635 neighborhoods were retrieved using the EFI-GNT tool (https://efi.igb.illinois.edu/efi-gnt/) (48).
636 The presence of a N-E-K triad was determined by manual inspection of the multiple sequence
637 alignment. For phylogenetic reconstruction, sequences were aligned with MAFFT on XSEDE
638 using CIPRES (50, 51), edited in Jalview (52), and analyzed with IQ-Tree (53-55). Maximum
639 likelihood trees were constructed using the LG+F+G4 substitution model and ultrafast bootstrap
640 (1000 replicates). Consensus trees were visualized, rooted at the midpoint, and annotated with
641 iTOL (56). Data files are available in Supplemental File Data Set 1.
642

643 **Acknowledgments**

644 We thank W. Keith Ray, Richard Helm and the Virginia Tech Proteomics Incubator for assistance
645 with mass spectrometry and Megan Rochford and Hualan Liu for construction of in-frame
646 deletions mutants. Research reported in this publication was supported by the National Institute of
647 Allergy and Infectious Disease of the National Institutes of Health under award
648 number R21AI109391 and R21AI149177 to S. B. M. The content is solely the responsibility of
649 the authors and does not necessarily represent the official views of the National Institutes of Health.
650 Work at the Molecular Foundry was supported by the Office of Science, Office of Basic Energy
651 Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. The work
652 conducted by the U.S. Department of Energy Joint Genome Institute (https://ror.org/04xm1d337),
653 a DOE Office of Science User Facility, is supported by the Office of Science of the U.S.
654 Department of Energy operated under Contract No. DE-AC02-05CH11231.
655

Author Contributions

Sarah E. Kivimaki: Investigation, Methodology, Visualization, and Writing – Review and Editing.

Samantha Dempsey: Investigation, Methodology, Visualization, and Writing.

Gary J. Camper: Investigation, Methodology, Visualization, and Writing.

Julia M. Tani: Investigation, Methodology, Visualization.

Ian K. Hicklin: Investigation, Methodology, Visualization, and Writing.

Crysten E. Blaby-Haas: Methodology, Visualization, and Writing – review and editing.

Anne M. Brown: Methodology, Visualization, and Writing – review and editing.

Stephen B. Melville: Conceptualization, Funding acquisition, Supervision, Investigation, Methodology, Project administration, Writing, Review and editing.

666 **References**

667 1. Singh PK, Little J, Donnenberg MS. Landmark Discoveries and Recent Advances in Type IV
668 Pilus Research. Microbiol Mol Biol Rev. 2022;86(3):e0007622.
669 2. Naskar S, Hohl M, Tassinari M, Low HH. The structure and mechanism of the bacterial
670 type II secretion system. Mol Microbiol. 2021;115(3):412-24.
671 3. Korotkov KV, Sandkvist M, Hol WG. The type II secretion system: biogenesis, molecular
672 architecture and mechanism. Nature reviews Microbiology. 2012;10(5):336-51.
673 4. Ayers M, Howell PL, Burrows LL. Architecture of the type II secretion and type IV pilus
674 machineries. Future microbiology. 2010;5(8):1203-18.
675 5. Desvaux M, Hebraud M, Talon R, Henderson IR. Secretion and subcellular localizations of
676 bacterial proteins: a semantic awareness issue. Trends Microbiol. 2009;17(4):139-45.
677 6. Anne J, Economou A, Bernaerts K. Protein Secretion in Gram-Positive Bacteria: From
678 Multiple Pathways to Biotechnology. Current topics in microbiology and immunology.
679 2017;404:267-308.
680 7. Demchick P, Koch AL. The permeability of the wall fabric of Escherichia coli and Bacillus
681 subtilis. Journal of bacteriology. 1996;178(3):768-73.
682 8. Varga JJ, Nguyen V, O'Brien DK, Rodgers K, Walker RA, Melville SB. Type IV pili-dependent
683 gliding motility in the Gram-p
684 ositive pathogen Clostridium perfringens and other Clostridia. Mol Microbiol. 2006;62(3):680-
685 94.
686 9. Melville S, Craig L. Type IV pili in Gram-positive bacteria. Microbiol Mol Biol Rev.
687 2013;77(3):323-41.

688 10. Obana N, Nakamura K, Nomura N. Role of RNase Y in Clostridium perfringens mRNA
689 Decay and Processing. J Bacteriol. 2017;199(2).
690 11. Obana N, Nakamura K, Nomura N. Temperature-regulated heterogeneous extracellular
691 matrix gene expression defines biofilm morphology in Clostridium perfringens. NPJ Biofilms
692 Microbiomes. 2020;6(1):29.
693 12. Bohning J, Ghrayeb M, Pedebos C, Abbas DK, Khalid S, Chai L, et al. Donor-strand
694 exchange drives assembly of the TasA scaffold in Bacillus subtilis biofilms. Nature
695 communications. 2022;13(1):7082.
696 13. Diehl A, Roske Y, Ball L, Chowdhury A, Hiller M, Moliere N, et al. Structural changes of
697 TasA in biofilm formation of Bacillus subtilis. Proc Natl Acad Sci U S A. 2018;115(13):3237-42.
698 14. Erskine E, Morris RJ, Schor M, Earl C, Gillespie RMC, Bromley KM, et al. Formation of
699 functional, non-amyloidogenic fibres by recombinant Bacillus subtilis TasA. Mol Microbiol.
700 2018;110(6):897-913.
701 15. Stover AG, Driks A. Secretion, localization, and antibacterial activity of TasA, a Bacillus
702 subtilis spore-associated protein. J Bacteriol. 1999;181(5):1664-72.
703 16. Hartman AH, Liu H, Melville SB. Construction and characterization of a lactose-inducible
704 promoter system for controlled gene expression in Clostridium perfringens. Applied and
705 environmental microbiology. 2011;77(2):471-8.
706 17. Nunn D, Bergman S, Lory S. Products of three accessory genes, pilB, pilC, and pilD, are
707 required for biogenesis of Pseudomonas aeruginosa pili. J Bacteriol. 1990;172(6):2911-9.
708 18. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for
709 protein modeling, prediction and analysis. Nature protocols. 2015;10(6):845-58.

710 19. Mirdita M, Schutze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold:
711 making protein folding accessible to all. Nature methods. 2022;19(6):679-82.
712 20. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate
713 protein structure prediction with AlphaFold. Nature. 2021;596(7873):583-9.
714 21. Kang HJ, Baker EN. Intramolecular isopeptide bonds: protein crosslinks built for stress?
715 Trends Biochem Sci. 2011;36(4):229-37.
716 22. Kang HJ, Coulibaly F, Clow F, Proft T, Baker EN. Stabilizing isopeptide bonds revealed in
717 gram-positive bacterial pilus structure. Science. 2007;318(5856):1625-8.
718 23. Baker EN, Squire CJ, Young PG. Self-generated covalent cross-links in the cell-surface
719 adhesins of Gram-positive bacteria. Biochemical Society transactions. 2015;43(5):787-94.
720 24. Evans RON, M.; Pritzel, A.; Antropova, N.; Senior, A.; Green, T.; Žídek, A.; Bates, R.;
721 Blackwell, S.; Yim, J.; et al. Protein Complex Prediction with AlphaFold-Multimer. bioRxiv
722 2022;2022.
723 25. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure
724 prediction of biomolecular interactions with AlphaFold 3. Nature. 2024;630(8016):493-500.
725 26. Plamont MA, Billon-Denis E, Maurin S, Gauron C, Pimenta FM, Specht CG, et al. Small
726 fluorescence-activating and absorption-shifting tag for tunable protein imaging in vivo. Proc Natl
727 Acad Sci U S A. 2016;113(3):497-502.
728 27. Tebo AG, Pimenta FM, Zhang Y, Gautier A. Improved Chemical-Genetic Fluorescent
729 Markers for Live Cell Microscopy. Biochemistry. 2018;57(39):5648-53.

730 28. Gautier A, Jullien L, Li C, Plamont MA, Tebo AG, Thauvin M, et al. Versatile On-Demand
731 Fluorescent Labeling of Fusion Proteins Using Fluorescence-Activating and Absorption-Shifting
732 Tag (FAST). Methods Mol Biol. 2021;2350:253-65.
733 29. Fives-Taylor PM, Thompson DW. Surface properties of Streptococcus sanguis FW213
734 mutants nonadherent to saliva-coated hydroxyapatite. Infect Immun. 1985;47(3):752-9.
735 30. Berry JL, Gurung I, Anonsen JH, Spielman I, Harper E, Hall AMJ, et al. Global biochemical
736 and structural analysis of the type IV pilus from the Gram-positive bacterium Streptococcus
737 sanguinis. J Biol Chem. 2019;294(17):6796-808.
738 31. Gurung I, Spielman I, Davies MR, Lala R, Gaustad P, Biais N, et al. Functional analysis of
739 an unusual type IV pilus in the Gram-positive Streptococcus sanguinis. Mol Microbiol.
740 2016;99(2):380-92.
741 32. Han X, Kennan RM, Parker D, Davies JK, Rood JI. Type IV fimbrial biogenesis is required
742 for protease secretion and natural transformation in Dichelobacter nodosus. J Bacteriol.
743 2007;189(14):5022-33.
744 33. Kirn TJ, Bose N, Taylor RK. Secretion of a soluble colonization factor by the TCP type 4
745 pilus biogenesis pathway in Vibrio cholerae. Molecular microbiology. 2003;49(1):81-92.
746 34. Whittaker CA, Hynes RO. Distribution and evolution of von Willebrand/integrin A
747 domains: widely dispersed domains with roles in cell adhesion and elsewhere. Molecular
748 biology of the cell. 2002;13(10):3369-87.
749 35. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and
750 PSI-BLAST: a new generation of protein database search programs. Nucleic acids research.
751 1997;25(17):3389-402.

752 36. Alegre-Cebollada J, Badilla CL, Fernandez JM. Isopeptide bonds block the mechanical
753 extension of pili in pathogenic Streptococcus pyogenes. J Biol Chem. 2010;285(15):11235-42.
754 37. Guilford WH, Lantz RC, Gore RW. Locomotive forces produced by single leukocytes in
755 vivo and in vitro. Am J Physiol. 1995;268(5 Pt 1):C1308-12.
756 38. Rood JI, Cole ST. Molecular genetics and pathogenesis of Clostridium perfringens.
757 Microbiological reviews. 1991;55(4):621-48.
758 39. Lisle JT, Smith JJ, Edwards DD, McFeters GA. Occurrence of microbial indicators and
759 *Clostridium perfringens* in wastewater, water column samples, sediments, drinking water, and
760 Weddell seal feces collected at McMurdo Station, Antarctica. Applied and environmental
761 microbiology. 2004;70(12):7269-76.
762 40. Nariya H, Miyata S, Suzuki M, Tamai E, Okabe A. Development and application of a
763 method for counterselectable in-frame deletion in Clostridium perfringens. Applied and
764 environmental microbiology. 2011;77(4):1375-82.
765 41. Hendrick WA, Orr MW, Murray SR, Lee VT, Melville SB. Cyclic Di-GMP Binding by an
766 Assembly ATPase (PilB2) and Control of Type IV Pilin Polymerization in the Gram-Positive
767 Pathogen Clostridium perfringens. J Bacteriol. 2017;199(10).
768 42. Sambrook J, Russel DW. Molecular Cloning, A Laboratory Manual, 3rd Edition. Cold
769 Spring Harbor, New York: Cold Spring Harbor Laboratory Press; 2001.
770 43. Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, et al. UCSF
771 ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci. 2018;27(1):14-
772 25.

773 44. Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, et al. UCSF ChimeraX:
774 Structure visualization for researchers, educators, and developers. Protein Sci. 2021;30(1):70-
775 82.
776 45. Johnson LS, Eddy SR, Portugaly E. Hidden Markov model speed heuristic and iterative
777 HMM search procedure. BMC Bioinformatics. 2010;11:431.
778 46. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. HMMER web server: 2018 update.
779 Nucleic Acids Res. 2018;46(W1):W200-W4.
780 47. Gerlt JA, Bouvier JT, Davidson DB, Imker HJ, Sadkhin B, Slater DR, et al. Enzyme Function
781 Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity
782 networks. Biochim Biophys Acta. 2015;1854(8):1019-37.
783 48. Oberg N, Zallot R, Gerlt JA. EFI-EST, EFI-GNT, and EFI-CGFP: Enzyme Function Initiative
784 (EFI) Web Resource for Genomic Enzymology Tools. J Mol Biol. 2023;435(14):168018.
785 49. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a
786 software environment for integrated models of biomolecular interaction networks. Genome
787 Res. 2003;13(11):2498-504.
788 50. Katoh K, Kuma K, Toh H, Miyata T. MAFFT version 5: improvement in accuracy of multiple
789 sequence alignment. Nucleic Acids Res. 2005;33(2):511-8.
790 51. M. A. Miller WPaTS. Creating the CIPRES Science Gateway for inference of large
791 phylogenetic trees. 2010 Gateway Computing Environments Workshop (GCE).1-8.
792 52. Clamp M, Cuff J, Searle SM, Barton GJ. The Jalview Java alignment editor. Bioinformatics.
793 2004;20(3):426-7.

794 53. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast
795 model selection for accurate phylogenetic estimates. Nat Methods. 2017;14(6):587-9.
796 54. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic
797 algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32(1):268-74.
798 55. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: Improving the
799 Ultrafast Bootstrap Approximation. Mol Biol Evol. 2018;35(2):518-22.
800 56. Letunic I, Bork P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic
801 tree display and annotation tool. Nucleic Acids Res. 2024;52(W1):W78-W82.
802 57. Grant SG, Jessee J, Bloom FR, Hanahan D. Differential plasmid rescue from transgenic
803 mouse DNAs into Escherichia coli methylation-restriction mutants. Proc Natl Acad Sci U S A.
804 1990;87(12):4645-9.
805 58. Andersen KR, Leksa NC, Schwartz TU. Optimized E. coli expression strain LOBSTR
806 eliminates common contaminants from His-tag purification. Proteins. 2013;81(11):1857-61.
807 59. Guzman LM, Belin D, Carson MJ, Beckwith J. Tight regulation, modulation, and high-level
808 expression by vectors containing the arabinose PBAD promoter. J Bacteriol. 1995;177(14):4121-
809 30.
810
811
812
813
814
815

816 **Figure legends**

817 **Figure 1.** Models showing proposed structures of Gram-positive T4P and Type II secretion
818 systems in comparison to those found in Gram-negative bacteria. Figure modified from Fig. 1 in
819 ref. (9).

820

821 **Figure 2.** Diagram showing the type IV pili-associated genes in *C. perfringens* strain 13 and other
822 species with two T4P systems. IMCP, inner membrane core protein; IMAP, inner membrane
823 accessory protein.

824

825 **Figure 3. (A).** SDS-PAGE gel showing the proteins in the secretome of the respective strains. The
826 gel was stained with the fluorescent protein dye Sypro-Ruby, the colors inverted to make the
827 protein bands in the samples appear dark but the pre-stained molecular weight markers appear as
828 bright bands. **(B).** Western blot showing the levels of secretion of the BsaC protein in different
829 strains. The secondary antibody was fluorescently tagged for visualization. Purified BsaC was
830 added in lane 5 as a positive control for detecting BsaC. **(C).** The gene synteny of the *bsa* operon
831 and adjacent sortase-encoding gene in *C. perfringens* strain 13. **(D).** The gene synteny of the *bsa*
832 operon and adjacent sortase-encoding gene in *C. perfringens* strain ATCC 13124 and most other
833 strains of *C. perfringens* that were observed. The sections of the *bsaB* gene from strain 13 that
834 were homologous to those in strain ATCC 13124 are labeled, along with the location of an LPXTG
835 sortase dependent sequence in one of the *bsaB*-encoding genes.

836

837 **Figure 4.** Efficient secretion of the BsaC protein requires T4P proteins. Results from densitometry
838 analysis of western blots from the secretome of strains with in-frame deletions of T4P-encoding

839 genes. Each strain contained plasmid pHLL65 for regulated expression of the *bsa* operon (*sipW-*
840 *bsaA-bsaB-bsaC*) in the wild type (HN13) and mutants deleted for T4P-associated genes. After
841 expression, secreted proteins were concentrated from culture supernatants using TCA and
842 quantitative western blotting was performed. The role of each gene product in pilus function is
843 shown in Fig. 2. For each experiment, the relative fluorescence intensity of the wild-type strain
844 (HN13) was set at 1.0 and the relative fluorescence intensity for each other strain in the same
845 experiment was then set relative to that value. Each strain was tested at least three times with
846 triplicate samples for each experiment and the mean and SEM are shown. All the experimental
847 runs were combined, and one sample t and Wilcoxon tests were run to determine if the mutant
848 strains were statistically different from 1. Asterisks indicate strains that were different ($P < 0.05$)
849 from strain HN13 (i.e., 1).
850
851 **Figure 5.** Use of western blots of secretomes to determine if T4P genes play a role in BsaA
852 oligomer secretion. After culture in BHI medium of strain HN13 with pSK4 (*sipW-CPE0515-*
853 FLAG), proteins were precipitated from culture supernatants by using TCA. **(A)** Sample blot
854 showing the presence of oligomers in the wild-type and some mutant strains. Densitometry was
855 used to determine the intensity of the fluorescent signal for all the oligomer forms that were
856 present. **(B)** Summary of results from all the T4P-associated mutants. Each strain was tested at
857 least six times in independent experiments and the mean and SEM are shown. For each experiment,
858 the relative fluorescent intensity of the wild-type strain (HN13) was set at 1.0 and the relative
859 fluorescence intensity for each other strain in the same experiment was then set relative to that
860 value. All the experimental runs were combined, and one sample t and Wilcoxon tests were run to

861 determine if the mutant strains were statistically different from 1. Asterisks indicate strains that
862 were different ($P < 0.05$) from 1.0 (i.e., strain HN13).
863
864 **Figure 6. (A)** AlphaFold model of a homotrimer of BsaA in cartoon representation. Each
865 independent monomer is colored separately. Note that residues 28-38 of a monomer incorporate
866 themselves into the beta strands of an adjacent monomer. **(B)** Close-up view of interdigitated beta
867 strands from two monomers in cartoon representation, showing relevant side chains as sticks
868 colored by hydrophobicity. N35 and the giving chain are colored purple, E56, K74, and the
869 receiving monomer are colored cyan. Side chains forming the hydrophobic pocket are colored
870 brown. The bond created by AlphaFold between N35 and K74 is between an oxygen and nitrogen
871 due to limitations in AlphaFold; in reality, an amide bond would form between the side chains with
872 ammonia as the leaving group. **(C)** Close-up view of the N35-E56-K74 catalytic triad, showing
873 the close proximity of the catalytic E56 residue to the bond forming N35 and K74. All images
874 were created with ChimeraX v1.2. (43, 44).
875
876 **Figure 7.** Western blots showing loss of oligomer formation in BsaA with mutations in residues
877 involved in isopeptide bond formation**. (A)** Oligomer formation seen in TCA precipitated culture
878 supernatants or in the disrupted whole cell extracts of strains producing BsaA with single
879 mutations. **(B)** Comparison of BsaA oligomer formation in the N35A and N35A/K74A mutants in
880 a wild-type background (left side) and one in which the chromosomal copy of the *bsaA* gene was
881 deleted. Note that the chromosomal *bsaA* mutation did not change the oligomerization levels of
882 the mutants. **(C)** Western blot showing mutation of the E56 residue to Ala abolished
883 oligomerization of BsaA. P, disrupted cell pellet; S, TCA precipitate of culture supernatant.

884

885 **Figure 8.** BsaA oligomers are found on the bacterial surface. **(A)** Expression of a FAST protein
886 fusion to the cytoplasmic PilT protein (pSM413) only shows fluorescence with the membrane
887 permeable Coral dye. **(B)** A BsaA-FAST fusion protein (pSM412) fluoresces in the presence of
888 both a membrane permeable dye (Coral) and a membrane impermeable dye (Amber-NP),
889 indicating it is surface exposed to these dyes. Scale bars = 5 μm.

890

891 **Figure 9.** BsaA phylogeny**.** (**A**) A consensus tree of representative BsaA-like proteins. Leaf labels
892 are colored by taxonomy according to the color key. BsaA, BsaB, and two other BsaA-like proteins
893 identified in *C. perfringens* are labeled, as are TasA from *B. subtilis* and CalY from *B. cereus*.
894 Bootstrap values are represented as scaled circles according to the inset key. (**B**) Sequence
895 similarity network of identified BsaA-like proteins. Nodes representing labelled *Clostridium* and
896 *Bacillus* proteins in panel A are labeled with an arrow. Nodes are colored by taxonomy according
897 to the color key. Proteins with conserved residues that would be expected to be involved in
898 formation of a spontaneous intermolecular isopeptide bond are represented by triangle nodes with
899 a red outline. (**C**) Example gene neighborhoods of genes encoding BsaA-like proteins and genes
900 encoding putative surface filament structures. Hatching indicates the presence of the conserved
901 NEK triad in the encoded protein. The complete matrix tree data set is included in a Supplemental
902 Information file.

903

904 **Figure 10.** Model of BsaA oligomerization. BsaA monomers (purple ovals) are oligomerized by
905 isopeptide bond formation after secretion via the Sec system in *C. perfringens*. SipW is the signal

906 peptidase for BsaA secretion (Fig. 3C and (11)). Secretion of the oligomers past the PG layer may
907 be due to the action of T4P-associated proteins (Fig. 5).
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930

931 **Table 1.** Strains and plasmids used in this report.

Strain or plasmid	Relevant characteristics	Reference or Source
Strains		
Clostridium perfringens type A strain HN13	*ΔgalKT*	(40)
Escherichia coli strain DH10B	F– *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *end*A1 *ara*D139 Δ(*ara leu*) 7697 *gal*U *gal*K *rps*L *nup*G λ–	(57)
Escherichia coli strain LOBSTR	Derivative of strain BL21, designed for low background binding to Ni-affinity columns	(58)
Plasmids		
pKRAH1	Cm^r; Lactose-inducible expression system plasmid for use in *C. perfringens*	(16)
pET24a	Protein expression vector	Novagen
pCRBluntIITOPO	PCR cloning vector	Invitrogen
pBAD30	Amp^r; Arabinose-inducible expression system plasmid for use in *E. coli*	(59)
pKRAH1-*pilA3*	Expression of *pilA3* for complementation	This study
pHLL65	The *bsa* operon in pKRAH1	This study
pGC1	pHLL65 with a His_6 tag added to the C-terminus of *bsaC*	This Study
pGC8	Vector based on pET24-a; used to express and purify VWA-His6 in *E. coli*	This Study
pSK1	The *bsa* operon in pKRAH1 with a FLAG-tag added to the C-terminus of *bsaA*	This Study
pSK2	Same as pSK1 with an HA tag at the C-terminus of *bsaB*	This Study
pSK4	The *sipW* and *bsaA* genes in pKRAH1 with a FLAG tag at the C-terminus of *bsaA*.	This Study
pSM407	The *sipW* and *bsaA* genes in pKRAH1 with a His_6 tag at the C-terminus of *bsaA*.	This Study

pSM408	pSM407 with a N35A mutation in *bsaA*	This Study
pSM409	pSM407 with a K74A mutation in *bsaA*	This Study
pSM410	pSM407 with N35A/K74A mutations in *bsaA*	This Study
pSM411	pSM407 with E56A mutations in *bsaA*	This Study
pSM412	pSM407 with a FAST gene fusion at the C-terminus of *bsaA* in place of the His_6 tag	This Study
pSM413	A *pilT*-FAST gene fusion in pKRAH1	This Study

Figure 1

Models of Type IV pili and Type II secretion systems

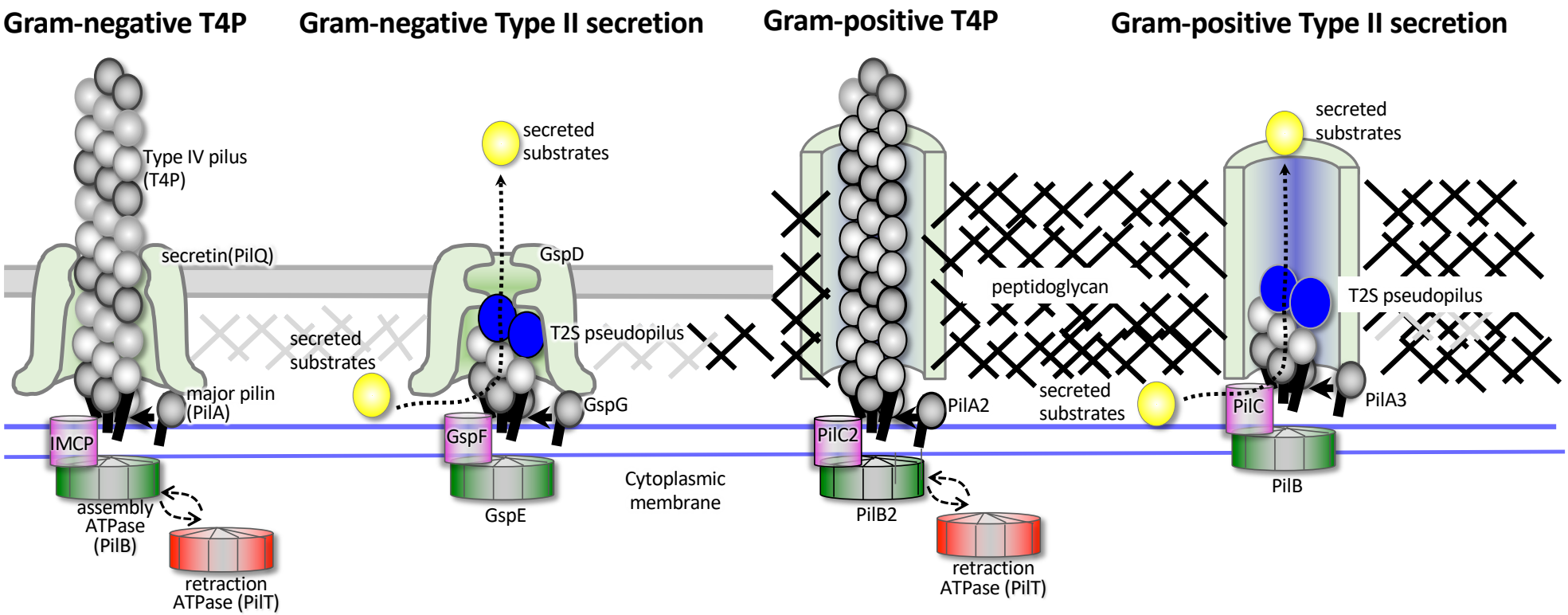


Figure 2

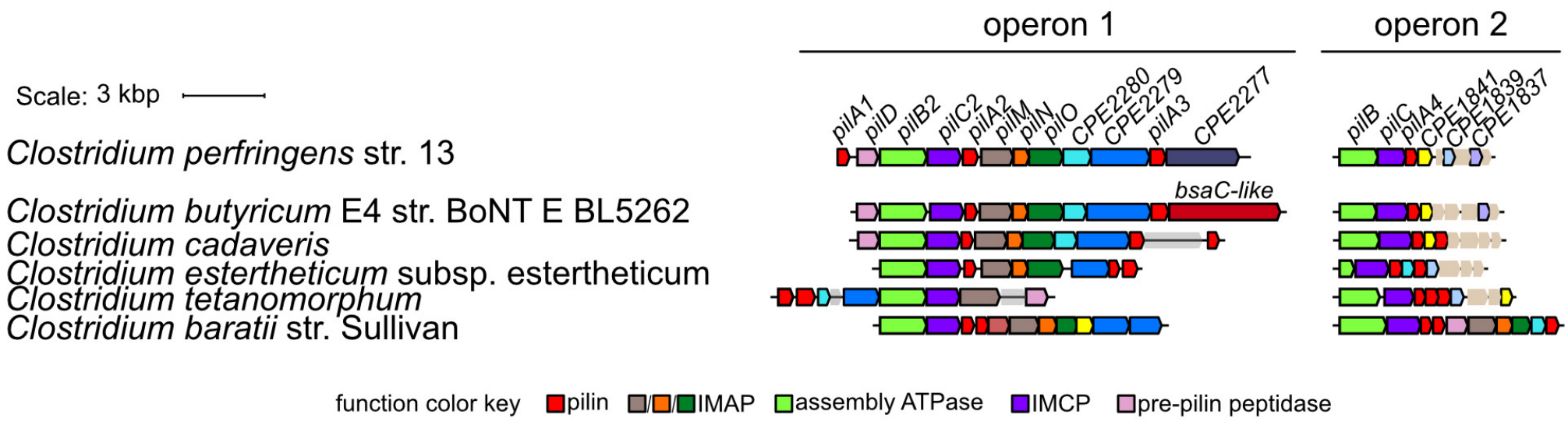

operon 1
operon 2
Scale: 3 kbp
pilA1
pilD
pilB2
pilC2
pilA2
pilM
pilN
pilO
CPE2280
CPE2279
pilA3
CPE2277
pilB
pilC
pilA4
CPE1841
CPE1839
CPE1837
bsaC-like
Clostridium perfringens str. 13
Clostridium butyricum E4 str. BoNT E BL5262
Clostridium cadaveris
Clostridium estertheticum subsp. estertheticum
Clostridium tetanomorphum
Clostridium baratii str. Sullivan
function color key
pilin
IMAP
assembly ATPase
IMCP
pre-pilin peptidase

Figure 3

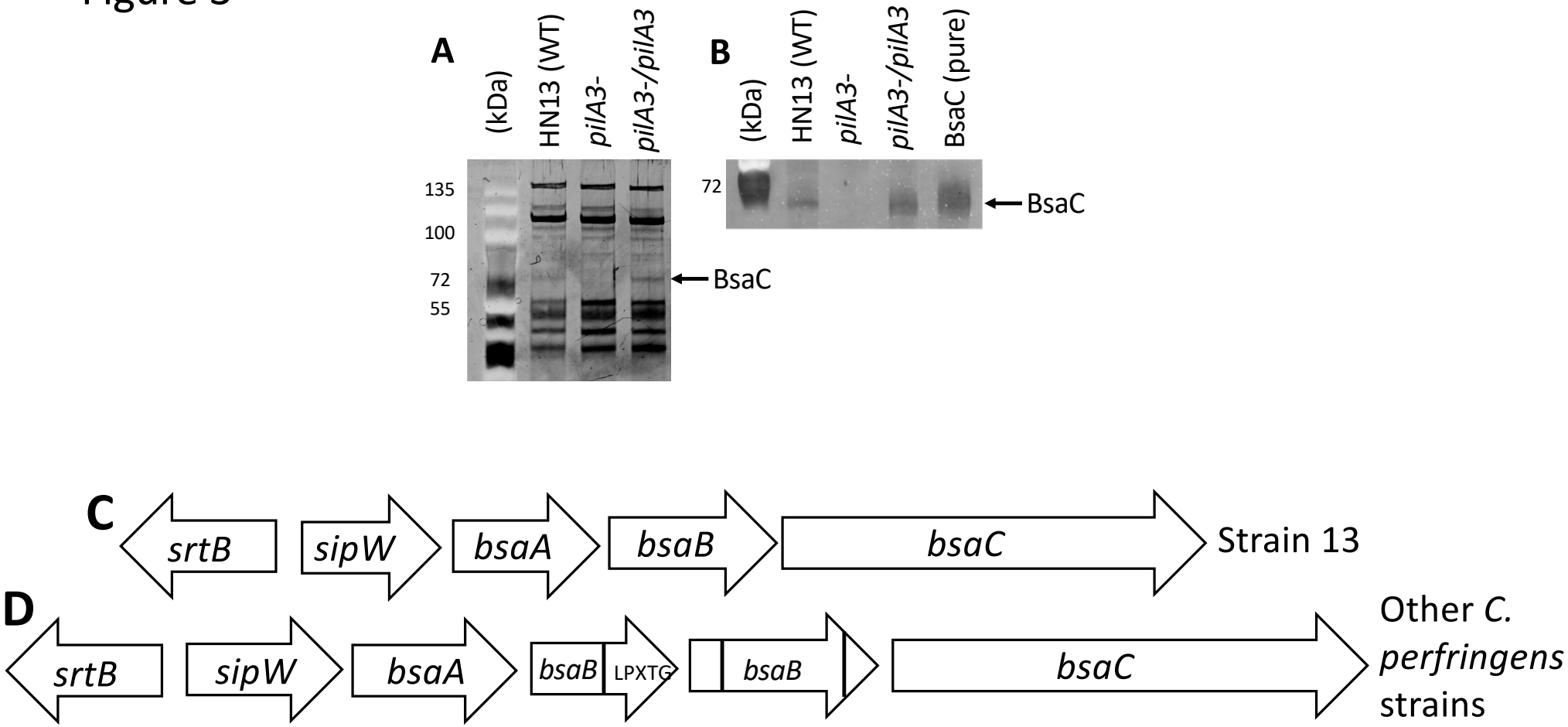

A
(kDa)
HN13 (WT)
pilA3-
pilA3-/pilA3
135
100
72
55
BsaC
B
(kDa)
HN13 (WT)
pilA3-
pilA3-/pilA3
BsaC (pure)
72
BsaC
C
srtB
sipW
bsaA
bsaB
bsaC
Strain 13
D
srtB
sipW
bsaA
bsaB
LPXTG
bsaB
bsaC
Other C. perfringens strains

Figure 4

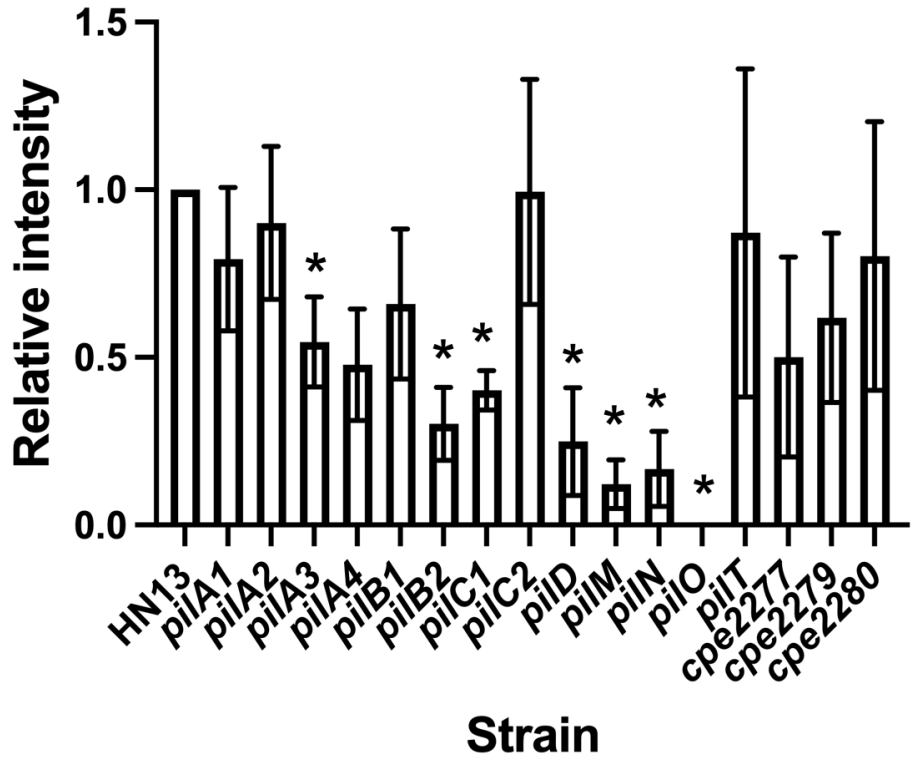

1.5
1.0
0.5
0.0
Relative intensity
*
*
*
*
*
*
*
HN13
pilA1
pilA2
pilA3
pilA4
pilB1
pilB2
pilC1
pilC2
pilD
pilM
pilN
pilO
pilT
cpe2277
cpe2279
cpe2280
Strain

Figure 5

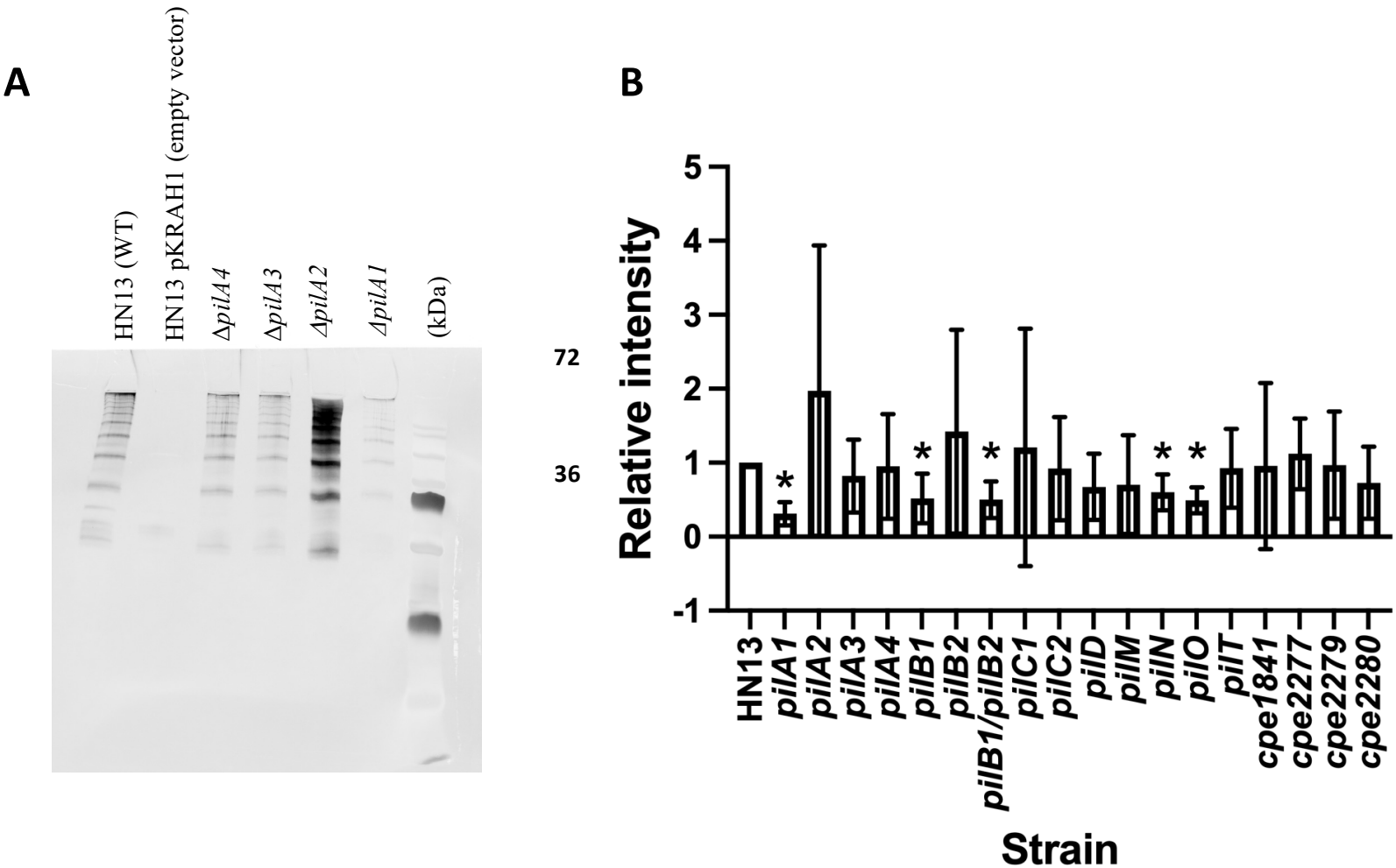
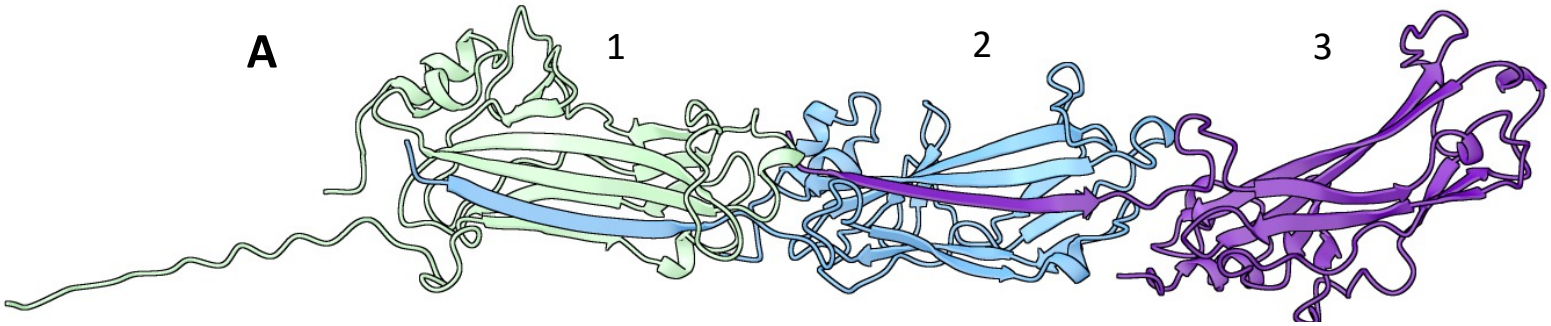
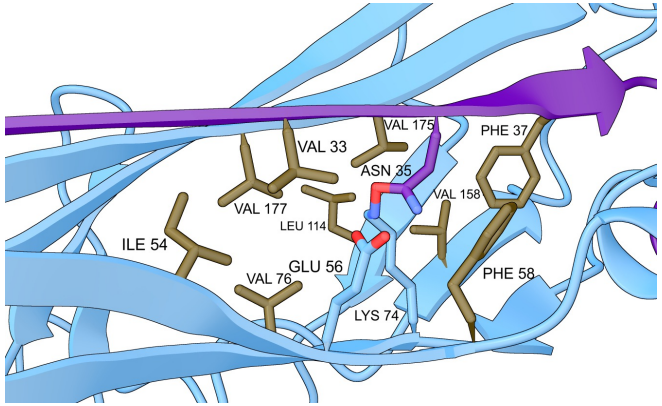

A
HN13 (WT)
HN13 pKRAH1 (empty vector)
ΔpilA4
ΔpilA3
ΔpilA2
ΔpilA1
(kDa)
72
36
B
Relative intensity
5
4
3
2
1
0
-1
HN13
pilA1
pilA2
pilA3
pilA4
pilB1
pilB2
pilB1/pilB2
pilC1
pilC2
pilD
pilM
pilN
pilO
pilT
cpe1841
cpe2277
cpe2279
cpe2280
Strain

Figure 6


A
1
2
3


B
VAL 175
PHE 37
VAL 33
ASN 35
VAL 158
VAL 177
LEU 114
ILE 54
GLU 56
VAL 76
PHE 58
LYS 74

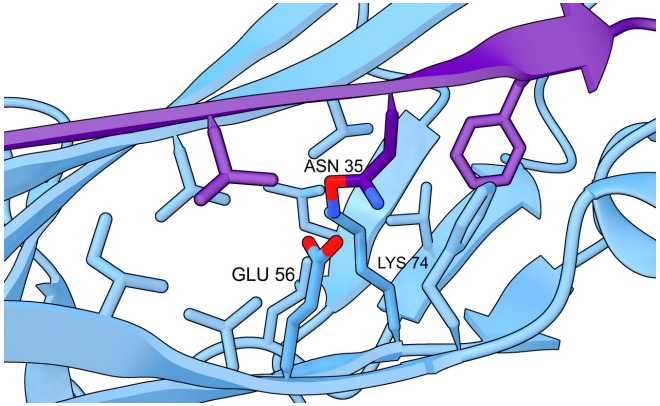

C
ASN 35
GLU 56
LYS 74

Figure 7

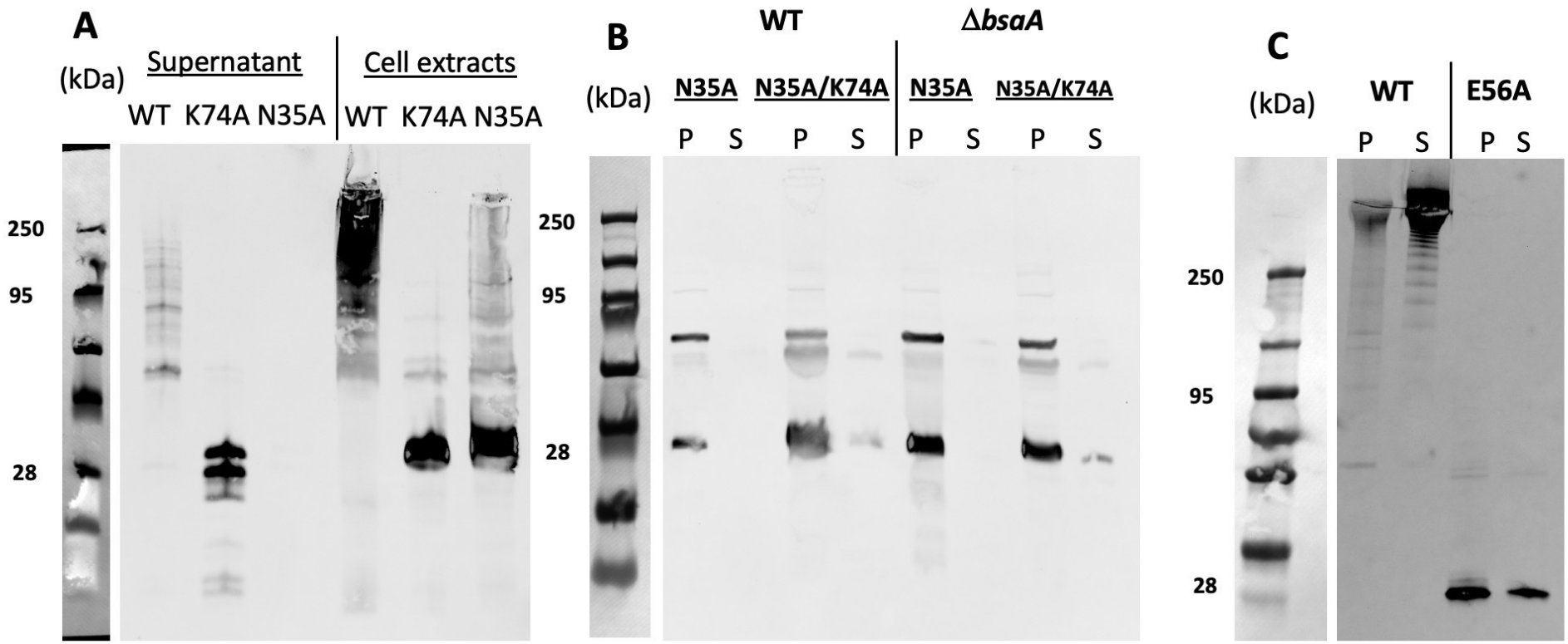

A
(kDa)
Supernatant
Cell extracts
WT K74A N35A
WT K74A N35A
250
95
28
B
(kDa)
WT
ΔbsaA
N35A
N35A/K74A
N35A
N35A/K74A
P S P S P S P S
250
95
28
C
(kDa)
WT
E56A
P S P S
250
95
28

Figure 8

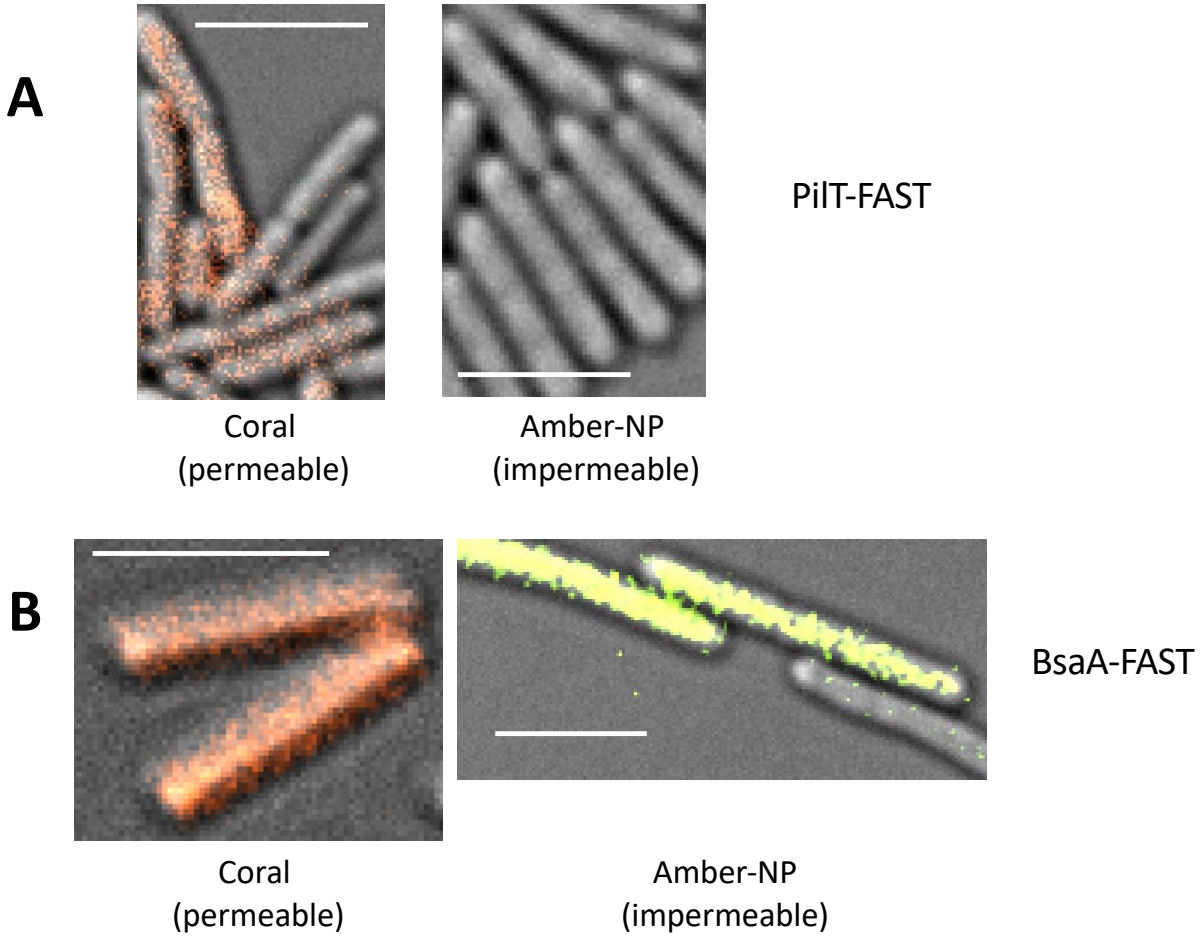

A
PilT-FAST
Coral
(permeable)
Amber-NP
(impermeable)
B
BsaA-FAST
Coral
(permeable)
Amber-NP
(impermeable)

Figure 9

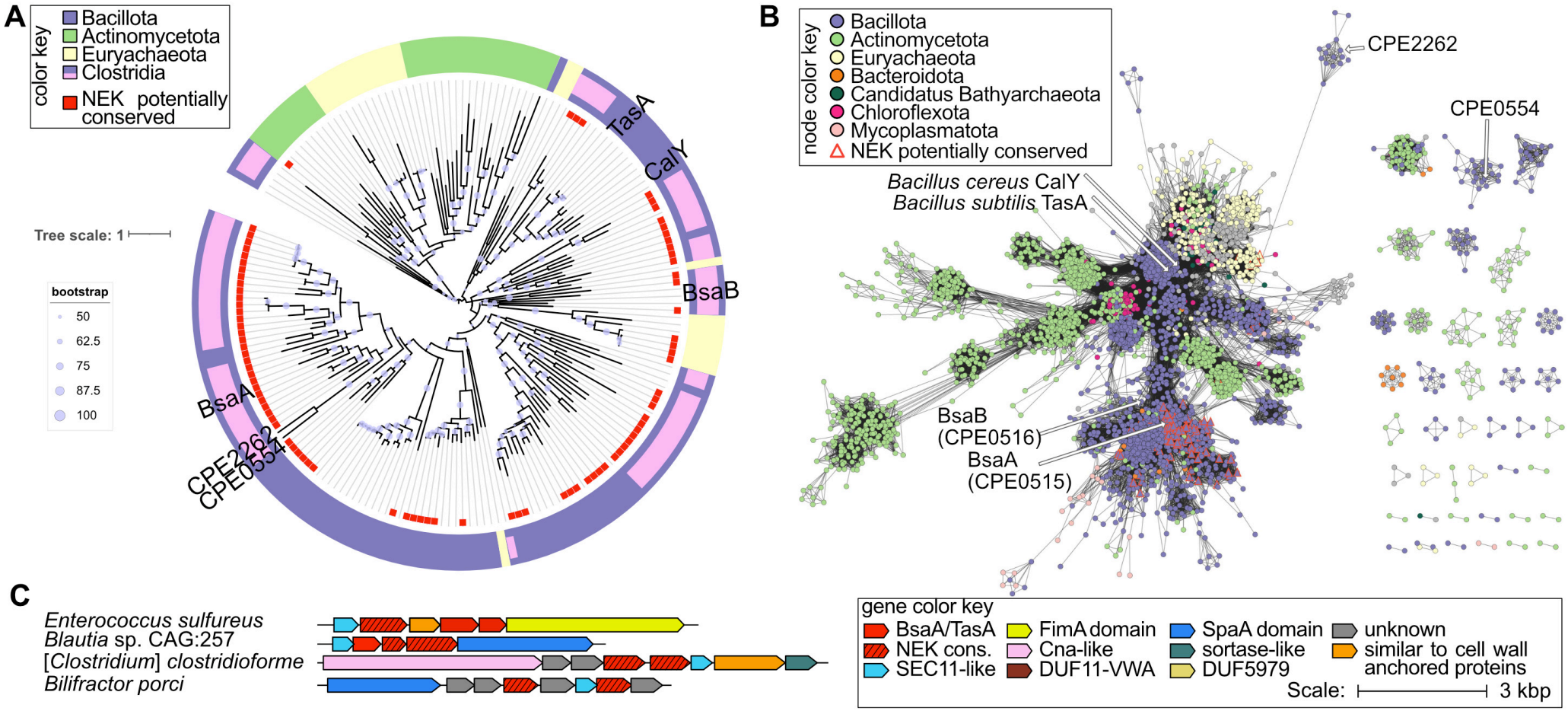
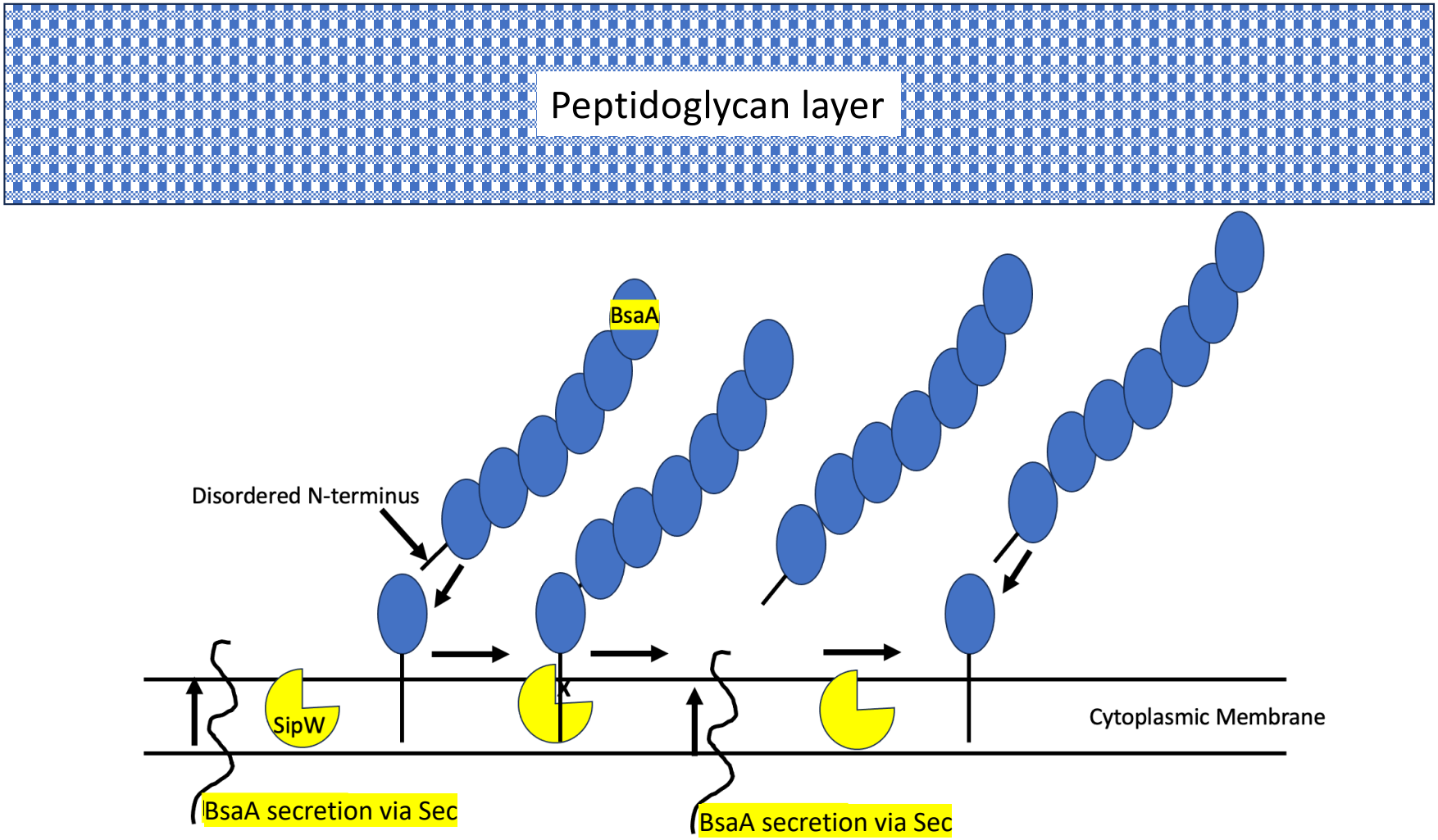

A
color key
Bacillota
Actinomycetota
Euryachaeota
Clostridia
NEK potentially conserved
Tree scale: 1
bootstrap
50
62.5
75
87.5
100
TasA
CalY
BsaB
BsaA
CPE2262
CPE0554
B
node color key
Bacillota
Actinomycetota
Euryachaeota
Bacteroidota
Candidatus Bathyarchaeota
Chloroflexota
Mycoplasmatota
NEK potentially conserved
CPE2262
CPE0554
Bacillus cereus CalY
Bacillus subtilis TasA
BsaB
(CPE0516)
BsaA
(CPE0515)
C
Enterococcus sulfureus
Blautia sp. CAG:257
[Clostridium] clostridioforme
Bilifractor porci
gene color key
BsaA/TasA
NEK cons.
SEC11-like
FimA domain
Cna-like
DUF11-VWA
SpaA domain
sortase-like
DUF5979
unknown
similar to cell wall anchored proteins
Scale:
3 kbp

Figure 10


Peptidoglycan layer
BsaA
Disordered N-terminus
SipW
Cytoplasmic Membrane
BsaA secretion via Sec
BsaA secretion via Sec