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 USA

† 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 PilA3 for secretion. The *bsaC* gene is in an operon with genes encoding a signal peptidase and two putative biofilm matrix proteins BsaA and BsaB, with remote homology to *Bacillus subtilis* TasA. We analyzed BsaA monomer interactions with *de novo* modeling, which projected that the monomers formed isopeptide bonds as part of a donor strand exchange process. 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 are predicted to form isopeptide bonds.

Importance

For bacteria to secrete folded proteins to the environment, they have to overcome the physical barriers of an outer membrane in Gram-negative bacteria and the thick peptidoglycan layer in

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

Introduction

Bacterial protein secretion to the exterior environment involves transporting the protein across several physical barriers. For Gram-negative bacteria these include the cytoplasmic membrane (CM), a relatively thin peptidoglycan (PG) cell wall and the outer membrane. Gram-negative bacteria have evolved several different mechanisms to achieve this, including Type II secretion systems (TTSS), which have many proteins similar to those found in Type IV pili (T4P) (Fig. S1 and (1, 2)). TTSS are modeled to function as pistons that use a short pseudopilus to pump folded periplasmic proteins through a protein channel (secretin) in the outer membrane (Fig. S1 and (3, 4)). Protein secretion in Gram-positive bacteria involves transport across the CM and a usually thick PG layer. There are many protein secretion mechanisms for transiting the CM known for Gram-positive bacteria, including the Sec system, Twin arginine translocation (Tat), Type IV secretion systems, Esat-6 or type VII secretion system, flagella export apparatus (FEA), fimbrilin (pilus) protein exporter (FPE), and holin-dependent translocation. The Type IV, Type VII and FEA are able to translocate proteins across the PG layer (5, 6). However, proteins secreted through the CM via the Sec system, used for the majority of proteins, need to fold and, once they do, the PG layer of Gram-positive bacteria acts as a barrier to secretion to the environment because the mesh size in the PG layer is too small to allow diffusion of globular proteins >25 kD, as determined by measurements of the PG from the Gram-positive bacterium *Bacillus subtilis* (7).

One possible mechanism for secretion of folded proteins from the space between the CM and the PG layer in Gram-positive bacteria would be a system analogous to TTSS in Gram-negative bacteria. Our discovery that T4P systems are ubiquitous and perhaps universal in all of the Clostridia (8) provided a means for a potential TTSS to exist in the Clostridia. We hypothesized in a review (9) that in Clostridial species with two separate T4P systems, one of them could function

103 as a TTSS to transport proteins through the PG layer. Multiple species of Clostridia do carry
104 multiple copies of complete T4P systems, including *C. perfringens* (Fig. 1) and *Clostridioides*
105 *difficile* (9). The core proteins needed for functional T4P usually include pili, an assembly ATPase
106 (PilB), an inner membrane core protein (PilC), and associated assembly proteins PilM, PilN, and
107 PilO (9). *C. perfringens* carries two homologs of PilB, two homologs of PilC and single copies of
108 PilM, PilN, and PilO (Fig.1), making it plausible that one of these systems could function as a
109 TTSS. To test this hypothesis, we examined the secretome of strains with in-frame deletions in
110 genes encoding Type 4 pilins; in *C. perfringens* these are designated as PilA1-PilA4 (Fig. 1). We
111 noticed a protein of ~70 kD was present in the WT secretome but absent in the secretome of one
112 deletion strain, *pilA3* (Fig. 2A). Mass spectrometry identified the 70 kD protein as CPE0517,
113 which was encoded by a gene in an operon with three other genes, including a SipW signal
114 peptidase homolog and two biofilm matrix proteins, CPR0515 (BsaA) and CPE0516 (BsaB) (Fig.
115 2C). We expressed the entire operon in *C. perfringens* and found that it produced large amounts of
116 biofilm matrix proteins (Fig. S2B), and that the biofilm matrix protein (CPE0515) was present as
117 a long oligomer and was resistant to boiling in SDS-PAGE buffer (Fig. S2D). We also noted that
118 CPE0515 had some predicted structural homology to the TasA protein of *B. subtilis*, which is also
119 a biofilm matrix protein that uses a SipW signal peptidase for secretion (10-12). At around this
120 same time, a paper by Obana et al (13) was published in which they examined many features of
121 the same operon, which they designated as *sipW-bsaA-bsaB-bsaC*. While the majority of their
122 work concerned the regulation of matrix protein synthesis, they discovered some important
123 physical features of the biofilm matrix protein (BsaA, Fig. 2C): it is the main component of the
124 biofilm matrix; it required SipW for signal processing and secretion; expression in *E. coli* of a
125 form of BsaA lacking a signal sequence led to spontaneous oligomerization of BsaA, which was

126 resistant to denaturation by SDS or formic acid, indicting the oligomers were held together by very
127 strong intermolecular interactions (13).
128 In this report we describe the process of identifying T4P proteins involved in secreting CPE0517
129 (BsaC), the 70 kD protein described above, as well as pilin proteins involved in secreting the BsaA
130 oligomers. We also propose an "exchange and lock" mechanism involving isopeptide bond
131 formation that is responsible for the stability of the BsaA oligomer. We characterized the nature
132 of these bonds and conservation of isopeptide bond formation based on phylogeny of the
133 BsaA/TasA family of proteins. Identifying a potential TTSS-like system in a Gram-positive
134 bacterium and the novel finding of isopeptide bond formation in a well-studied family of proteins
135 will lead to significant advances in both of these emerging fields of microbial cell biology.
136

137 **Results**

138 **Secretome analysis of a *pilA3* mutant indicates a 70 kD protein with a Von Willebrand**
139 **domain A (CPE0517, BsaC) is absent.** We screened the secretome of in-frame deletion mutants
140 in four pilin-encoding genes, *pilA1*, *pilA2*, *pilA3*, *pilA4* for missing proteins, indicating they may
141 be dependent on the pilin for secretion. Only the *pilA3* mutant showed evidence of a distinct
142 missing band in the secretome and this band was restored by complementation of a WT copy of
143 the *pilA3* gene (Fig. 2A). The band was excised from a gel and subjected to trypsin digestion and
144 analysis by mass spectrometry. Peptide fragments matched those of the CPE0517 (BsaC) protein
145 from the *C. perfringens* strain 13 genome. The protein has an unknown function but does contain
146 a von Willebrand A domain, which can potentially function in protein-protein interactions with
147 host cells. This gene is the fourth gene in a four gene operon containing a type I signal peptidase
148 (CPE0514, SipW) and two genes (*CPE0515-CPE0516*, *bsaA-bsaB*) encoding proteins with

149 similarity to the TasA protein in *B. subtilis* (Fig. 2C). Western blots using anti-BsaC antibodies
150 showed a similar loss of CPE0517 secretion by the *pilA3* mutant and complementation by
151 expression of a WT copy of the *pilA3* gene (Fig. 2B).
152 To determine if other Type IV pilus (TFP) related genes were needed for secretion of BsaC, we
153 expressed the gene in a bank of in-frame deletion mutants encompassing each gene related to
154 synthesis and assembly of TFP (Fig. 1) using a vector we developed containing a lactose-inducible
155 promoter (14). We noted variation in day-to-day results with the WT strain and, since these were
156 western blots on TCA precipitated protein from culture supernatants, we normalized all of the data
157 from mutant strains to that of the WT strain on the same blot (Fig. 3). Using quantitative
158 densitometry on western blots of the secreted oligomers, we measured statistically significant
159 decreased secretion of BsaC in the following mutants: *pilA3*, *pilB2*, *pilC1*, *pilD*, *pilM*, *pilN* and
160 *pilO* (Fig. 3). With the exception of *pilC1*, these results suggest efficient secretion of BsaC requires
161 components of the TFP system present in the large *pilD-CPE2277* operon in *C. perfringens* (Fig.
162 1 and (8, 9)). Whether these form a type IV pilus with PilA3 as the pilin or a Gram-positive TTSS
163 remains to be determined. However, a major difference in the structures of T4P and TTSS is the
164 type IV pilus extends beyond the outer membrane (Gram-negative) or PG layer (Gram-positive,
165 (8)) but in the TTSS in Gram negative bacteria the pilus does not (2). Our model for a TTSS in
166 Gram-positive bacteria is that the pilus (i.e., PilA3) does not extend past the PG layer (Fig. S1).
167 To test this model, we exposed *C. perfringens* cells to rabbit anti-PilA3 antibodies and a
168 fluorescently tagged goat anti-rabbit secondary antibody and found that we could detect binding
169 localized to the poles of the cells (Fig. S2). This refined the location of the pili, and that they extend
170 far enough out past the PGY layer to be exposed to antibodies.

171 To determine if there are generalized secretion defects in TFP mutants, we examined
172 whether each mutant in the TFP-associated genes were able to secrete another protein, the
173 phospholipase-sphingomyelinase, PLC. Secretion of PLC can be easily detected by a ring of
174 precipitation around a colony using egg yolk agar, and we found that none of the TFP mutants
175 showed a detectable decrease in PLC secretion (Fig. S3), suggesting the secretion defects seen
176 with BsaC were not due to generalized secretion defects.

177 **Expression of the *bsa* operon leads to synthesis of extracellular biofilm matrix material.** Since
178 the *bsa* operon encodes proteins with similar synteny and predicted structural homology to the
179 *sipW-tasA-tapA* biofilm matrix operon in *B. subtilis* (12), we placed the entire operon (*sipW-bsaA-*
180 *bsaB-bsaC*) under control of the lactose inducible promoter for high level expression. The cells
181 produced large quantities of an extracellular gel like material (Fig. S4B), which, after staining with
182 crystal violet, appeared as amorphous material between the bacterial cells (Fig. S4C). A His_6 tag
183 was added to the C-terminus of the BsaA protein and, when the operon was expressed, a distinct
184 high molecular weight ladder of BsaA was visible in anti-His_6 antibody western blots after running
185 on SDS-PAGE gels using samples boiled in SDS-PAGE buffer (Fig. S4D). This suggested BsaA
186 was in very heat-stable oligomers. This was confirmed by the results of Obana et al who found
187 that BsaA was the major component of the biofilm matrix and that BsaA oligomers were stable
188 after boiling in SDS-PAGE buffer and resistant to dissociation by high levels of SDS and formic
189 acid (13). These results suggested BsaA oligomers were in an extremely stable conformation which
190 may include covalent bonds. However, these stable oligomers were not due to the activity of the
191 sortase-encoding gene *srtB* (Fig. 2B), since a mutant lacking that gene still produced highly stable
192 oligomers of BsaA (13).

193 **Efficient secretion of the BsaA oligomer is dependent on some TFP-associated genes.** Since
194 it was possible that, like BsaC, BsaA in its oligomer form was secreted with the aid of TFP-
195 associated proteins, we expressed a clone containing just the *sipW* and *bsaA* genes in which a
196 FLAG-tag was added to the C-terminus of BsaA for quantification in western blots using
197 densitometry. An example of a typical blot is shown in Fig. S5A. The WT strain showed significant
198 levels of secretion of the BsaA oligomer (Fig. S5B). Using this method, we found that the *pilA1*,
199 *pilB1*, *pilB1/pilB2*, *pilN* and *pilO* mutants exhibited decreased levels of BsaA oligomers in the
200 supernatants (Fig. S5B). However, since this profile of T4P-associated genes lacked an inner
201 membrane core protein (*pilC*) component, which is essential for TFP functions (15) the results
202 were not definitive. Therefore, we used a quantitative slot blot method as an alternative. In this
203 experiment, culture supernatants were serially diluted in medium and passed through a membrane,
204 and probed with an anti-FLAG antibody (Fig. S6A). The densitometry readings were then plotted
205 to identify an amount of signal that was in the linear range for quantification (Fig. S6B). However,
206 using this method only the *pilA1*, *pilA2* (higher), and *cpe1841* (higher) mutants showed a statistical
207 difference from the WT strain (Fig. S6C), which does not comprise a full T4P system.
208 **Models of BsaA oligomers predicts an isopeptide bond is formed linking adjacent monomers.**
209 BsaA forms very heat and SDS resistant oligomers ((13) and Figs. S4D and S5A). To investigate
210 the structural nature of the interactions between monomers of BsaA that leads to this stability, a
211 structure would need to be developed of a homomultimer. Due to the difficulty of solving the
212 structure of long, thin chains by experimental methods, computational modelling was used. Models
213 were initially created using Phyre2 (16), with the top model being an x-ray crystal structure of
214 TasA from *B. subtilis* (17). That structure, and homology models from that structure, were
215 monomeric, and not useful for determining intermolecular interactions between monomers.

216 AlphaFold Multimer was then used to create a multimeric model to determine inter-chain
217 interactions. Using ColabFold 1.3.0 (18), which is based on AlphaFold 2.2.0 (19), a homotrimeric
218 structure of BsaA was created (Fig. 4A). The SipW signal sequences (residues 1-27) were included
219 in the initial model, but only formed disordered loops which do not have any notable interactions.
220 Removing the signal sequences did not change the overall structure of the monomers so this was
221 done for subsequent analyses.

222 In monomeric structures, residues 28-38 would extend out as disordered loops (Fig. 4A,
223 monomer 1). In the multimeric structures, these residues incorporated themselves into the beta
224 sheet structure of adjacent monomers (Fig. 4A). Being fully incorporated into a nearby structure
225 indicates that residues 28-38 provide specific interactions with the receiving monomer, providing
226 the basis of a mechanism for which oligomers of BsaA can be created. For example, V33 and F37
227 of the interdigitated strand become part of the hydrophobic core of the receiving monomer (Fig.
228 4B). The interactions between specific side chains of residues from the interdigitated strand and
229 the receiving monomer were investigated further for less standard interactions.

230 The *de novo* created structures indicate a bond between two side chains: N35 of the
231 interdigitated strand and K74 of the receiving monomer (Fig. 4C). At first, this was thought to be
232 an error in AlphaFold's model generation, since the chemistry would not be possible. The bond
233 was created between the oxygen of N35 and the nitrogen of K74, resulting in three bonds in the
234 oxygen. However, this bond was created in nearly every model generated, so it was investigated
235 further. A crosslink between asparagine and lysine is a type of isopeptide bond, which has been
236 characterized through a few different mechanisms and can be found as an intermolecular crosslink,
237 as in ubiquitylation (20), or as an intramolecular crosslink found in Gram-positive pili (21). In
238 Gram-positive pili, this bond is formed between an asparagine and a lysine, autocatalyzed by a

239 glutamate within the hydrophobic core of the protein. Residues equivalent to these are found
240 surrounding the bond in BsaA (Fig. 4C). The hydrophobic region surrounding the N35-K74 pair
241 (Fig. 4B, brown colored residues) is believed to be necessary to help the catalytic cross-linking
242 reaction (22).

243 It is unclear why this bond was created by AlphaFold. In the documentation of AlphaFold Multimer
244 (23), there are no mentions of predicting crosslinks. One potential reason is that it generated the
245 side chains so close to each other there was a steric clash, but the rest of the model was so favorable
246 the structure was created regardless. Models created in ColabFold have an option to perform
247 energy minimization using Amber. For these models, this was not performed. Recreating this
248 model while enabling energy minimization eliminates the bond every time. This is expected given
249 the incorrect chemistry of the bond. AlphaFold3 (24) does not include a minimization step, and,
250 like AlphaFold2, creates the isopeptide bond.

251 This amyloid-like fold and interdigitation is seen in other Gram-positive bacteria, such as
252 TasA in *B. subtilis*. A cryo-EM structure has been created for a chain of TasA (10), which
253 demonstrated how hydrophobic interactions between donor and acceptor monomers drive
254 interdigitation. No crosslinks were found or mentioned, and the required triad of N-E-K residues
255 is not present in the structure (PDB 8aur). Therefore, the isopeptide bonds in BsaA oligomers
256 would be the first example of an intermolecular bond of this mechanism found in this family of
257 proteins.

258 **Mutations in isopeptide bond forming residues abolish oligomerization of BsaA.** The
259 AlphaFold Multimer models predicted an isopeptide bond was formed between N35 and K74 from
260 adjacent residues in a BsaA oligomer. To test this computational prediction, we made point mutant
261 substitutions of each residue to alanine, forming N35A and K74A mutants, along with a double

262 N35A/K74A mutant. His-tagged versions of these mutant forms were expressed along with *sipW*
263 and the cell pellets and culture supernatants were examined for the presence of the mutant proteins
264 in oligomeric form. Mutating N35 or K74 to alanine residues resulted in a loss of oligomers in the
265 supernatants and the appearance of monomers of the K74A mutant but not of the N35A mutant,
266 suggesting the N35A version was not effectively released by the cell (Fig. 5A). Monomeric forms
267 of both mutants appeared in the cell pellet of some samples (Fig. 5A). In some experiments, small
268 amounts of dimers and higher mol wt forms of BsaA appeared in the cell pellet, but the large
269 majority were in monomeric form, which was absent in the WT form of BsaA (Fig. 5A), perhaps
270 due to the stability of oligomers even in the absence of an isopeptide bond. To test this, we
271 suspended cell pellet extracts from BsaA-His_6 and N35A-His_6 in SDS-PAGE buffer and heated at
272 95 C for 0, 10, 20 and 40 min before running the samples on SDS-PAGE gels and performing
273 western blots. The N35A-His mutant maintained stable oligomers in SDS-PAGE buffer and after
274 being run through an SDS-PAGE gel in the absence of heating, but heating the samples for 10-40
275 min led to dissociation of the oligomers, in contrast to the WT BsaA protein, which was not
276 affected by heating (Fig. S5). This suggests that N35, and therefore the isopeptide bond, is essential
277 for forming detergent and heat resistant oligomers, and that the monomer-monomer interactions
278 are stable even in the presence of 2% SDS (the concentration in the SDS-PAGE buffer), which is
279 in contrast to oligomers formed by the *B. subtilis* TasA protein, which were dissociated by SDS
280 (11).

281 We also tested the N35A/K74A mutant and observed that it showed a similar loss of
282 oligomerization as did the N35A mutant (Fig. 5B, left four lanes). These experiments were all done
283 in strain HN13 (WT) background where there was an intact *bsaA* gene. To determine if this
284 potential source of BsaA, even though it was not His-tagged, affected the results we observed with

285 the mutants, we expressed the N35A and N35A/K74A mutants in a strain with an in-frame deletion
286 of the *bsaA* gene, but did not observe any differences to the pattern seen in the strain with an intact
287 *bsaA* gene (Fig. 7B, right four lanes), suggesting the chromosomal source of BsaA was not forming
288 oligomers in combination with the N35A and N35A/K74A mutant forms of BsaA.

289 **The Glu56 residue of BsaA is essential for efficient isopeptide bond formation.** Spontaneous
290 or self-catalyzed isopeptide bond formation is catalyzed by an acidic Glu or Asp residue located
291 within the hydrophobic pocket where the Asn and Lys residues are located (22). In AlphaFold 2
292 BsaA models of dimer interfaces, Glu56 is predicted to be in the precise location to catalyze bond
293 formation (Fig. 4B). To test this was actually occurring, we constructed an E56A substitution in
294 the BsaA protein and examined its ability to oligomerize by using western blotting on cell extracts
295 and concentrated supernatants from the mutant and WT strain (Fig. 5C). The E56A mutant showed
296 a clear lack of oligomer formation and an accumulation of the monomeric form of the protein
297 which was not observed with the WT form of BsaA (Fig. 5C), providing strong evidence that
298 Glu56 is responsible for catalyzing isopeptide bond formation in BsaA.

299 **A FAST fusion to BsaA displays surface exposed localization.** To investigate the cellular
300 location of the BsaA oligomers, a FAST gene fusion to the *bsaA* gene was constructed and, along
301 with the *sipW* gene, was placed under control of the lactose-inducible promoter. The FAST protein
302 binds a soluble dye which, upon binding, exhibits a large increase in fluorescence (25-27) allowing
303 localization of the fusion proteins. We expressed the *sipW-bsaA*-FAST fusion genes and exposed
304 the cells to two different dyes, Coral, which is membrane permeable, and Amber-NP which is
305 impermeable to membranes (both obtained from The Twinkle Factory). To test if the Coral and
306 Amber-NP permeability profiles were working in *C. perfringens* membranes, we expressed a *pilT*-
307 FAST gene fusion, which should be located only in the cytoplasm, and exposed them to the Coral

308 and Amber-NP dyes. Consistent with the predicted cytoplasmic location of the *pilT*-FAST fusion,
309 only the Coral dye showed significant fluorescence after treatment (Fig. 6A). However, when the
310 *bsaA*-FAST fusion was expressed along with *sipW*, both dyes were able to bind and fluoresce (Fig.
311 6B), indicating the BsaA-FAST fusion was competent for secretion in *C. perfringens* and was
312 located on the surface of the bacteria.

313 **BsaA oligomers are anchored to the membrane and not the peptidoglycan in strain HN13.**
314 Because the BsaA-FAST fusion experiments indicated BsaA oligomers were located on the cell
315 surface and SEM images of the BsaA matrix proteins suggested they were in the form of fibrils
316 associated with the bacterial surface (13), we wanted to determine if the oligomers were embedded
317 in the cytoplasmic membrane or covalently linked to the PG layer. To test this, we expressed *sipW*-
318 *bsaA*-His_6 constructs for the WT and each mutant in our lactose-inducible promoter system and
319 then boiled the cells in 10% SDS for 30 min, followed by digestion of the PG matrix by lysozyme.
320 Our experimental approach was based on the hypothesis that if BsaA was anchored in the
321 membrane, boiling in SDS would release it from the cell but if it was covalently anchored to the
322 PG, it would be freed only after digestion with lysozyme. We tested the WT, N35A, K74A and
323 N35/K74 mutants and found that all of the long wild-type oligomers were released after the SDS
324 boiling step, while the cells appeared to retain a small number of monomers and dimers of BsaA
325 (Fig. S4). However, after treatment with lysozyme, the monomers and dimers were no longer
326 detected. For the mutants that can't oligomerize, the monomer/dimers were retained after boiling
327 in SDS, but not after lysozyme treatment. We interpret these findings as indicating the oligomers
328 are anchored to the membrane but after SDS boiling they are released. A small number of
329 monomers and dimers remain trapped inside the sacculi produced by boiling in SDS but lysozyme
330 treatment releases these from the sacculi. However, they are then too dilute to detect on western

331 blots. This is consistent with an analysis of the *bsaA* and *bsaB* protein sequences, both of which
332 lack a recognizable LPXTG motif, despite the presence of a sortase gene, *srtB*, adjacent to the
333 operon (Fig. 1C).

334 **BsaB, an ortholog of BsaA, forms dimers but not long oligomers.** Previous results by Obana et
335 al (13) showed a *bsaB* mutant still made significant amounts of the BsaA oligomers but they did
336 not test to see if BsaB was incorporated into BsaA oligomers. Therefore, we added an HA-tag to
337 the BsaB protein and expressed the *sipW*-*bsaA*-*bsaB*-HA-*bsaC* genes in our inducible system and
338 examined the culture supernatants for BsaB-HA using western blots. Based on the predicted
339 molecular weight of the BsaB protein, we detected BsaB-HA only in a monomeric and a higher
340 molecular weight form which could be either a homodimer or a BsaA-BsaB heterodimer (Fig. S6),
341 but it is not covalently attached to the longer BsaA oligomers.

342 **Phylogeny of isopeptide bond formation in the TasA family.** Since a BlastP search with BsaA
343 only identifies homologs in closely related Clostridial species, an iterative search with profile
344 hidden Markov models was employed. This analysis resulted in the identification of 2,172 proteins,
345 including the previously described TasA from *B. subtilis* and CalY from *Bacillus cereus* (Fig. 7).
346 Of these, 632 hits belong to the "peptidase M73, camelysin" (InterPro: IPR022121) family; so-
347 called because of the likely mis-identification of CalY as a peptidase. In addition to the use of
348 probabilistic models, the relatedness of these proteins is evident at the structural level. Like other
349 family members, BsaA is predicted to adopt a β-sheet–rich Ig-fold-like structure and is predicted
350 to participate in donor-strand exchange (Fig. 4), as described for TasA (10). These analyses suggest
351 that while the structures of these proteins and homopolymer assembly are highly conserved,
352 sequences are highly divergent.

353 In addition to BsaA, we identified three other uncharacterized proteins in *C. perfringens*
354 (BsaB, CPE2262 and CPE0554), but these paralogs are not expected to form an intermolecular
355 isopeptide bond, because the N-E-K triad is missing. Nevertheless, over 100 BsaA-like proteins in
356 other species were identified with the N-E-K triad, suggesting conservation of the donor-strand-
357 exchange-and-lock mechanism. These proteins are largely from Clostridia with some homologs
358 from *Coprobacillaceae*, *Enterococcaceae*, and *Thermococcus* (Fig. 7).

359 Based on the putative functions of gene neighbors, the BsaA/TasA family members appear
360 to be involved in the assembly of surface filament structures (Fig. 7C). They are often in putative
361 operons with genes encoding proteins containing the Gram-positive cell wall anchor motif LPxTG,
362 fibronectin domains, SpaA domains, choice-of-anchor A domains, as well as signal peptidase and
363 sortase domains. Multiple BsaA/TasA paralogs can be found in a single putative operon (Fig. 7C),
364 suggesting the possibility of donor-strand exchange assembly of heteropolymers.

365

366 **Discussion**

367 The original goal of this work to determine if Gram-positive bacteria have the functional
368 equivalent of TTSS that are found in Gram-negative bacteria. Since TTSS have a similar overall
369 architecture and mechanism as T4P (Fig. S1), it made sense to test a Gram-positive bacterium
370 known to have T4P. With a few notable exceptions (e.g., *Streptococcus sanguinis* (28-30))
371 amongst Gram-positive bacteria, T4P are confined within the Clostridia (9) so *C. perfringens* was
372 chosen as representing a typical clostridial species. The overall concept was that the thick Gram-
373 positive PG layer represents a similar barrier to protein secretion to the environment as the Gram-
374 negative outer membrane. Since TTSS depend upon a pseudopilus to function, we tested mutants
375 with deletions in each of the four main pilin-encoding genes in *C. perfringens* for changes in the

376 secretome compared to a WT strain. We identified one protein BsaC (CPE0517) that was depleted
377 in the secretome of a *pilA3* deletion strain (Fig. 2A-B). We then examined a bank of in-frame
378 deletion mutants in all of pilin-associated genes (Fig. 1) and found that, for BsaC secretion, a
379 homolog of each of the essential proteins needed for a type IV pilus assembly, PilA3 (pilin), PilB2
380 (assembly ATPase), PilC1 (inner membrane core protein), PilD (prepilin peptidase), and PilM-N-
381 O, (inner membrane accessory proteins) (Fig. 1) were required for efficient secretion (Fig. 3). This
382 appears to be specific for BsaC secretion, since the secretion of the toxin PLC was not affected by
383 mutations in any of the T4P-associated genes (Fig. S2). While this may indicate the presence of a
384 TTSS like apparatus in *C. perfringens*, the lines separating T4P- and T2SS-dependent secretion
385 have been blurred by the discoveries that *Pseudomonas aeruginosa* and *Dichelobacter nodosus*,
386 Gram-negative bacteria, have T4P systems capable of mediating protein secretion (31, 32).
387 However, further analysis of the structure and components of these Gram-negative pilus-dependent
388 secretion systems have not been published. PilA3 pilin was detected on the surface of *C.*
389 *perfringens,* so it is possible the PilA3-dependent secretion system acts more like a T4P.

390 BsaC monomer secretion may appear to be independent of BsaA oligomers, but it is
391 possible that they are being secreted simultaneously, since we did not track BsaA and BsaC
392 secretion at the same time. The biological role that BsaC plays in biofilm matrix functions is
393 unknown, but it is intriguing that the protein contains a VWA domain, which is widespread in
394 eukaryotes, bacteria and archaea and is associated with protein-protein interactions (33).
395 Therefore, the role BsaC plays in biofilm functions is also currently under investigation in the
396 laboratory.

397 BsaC is in an operon with genes encoding a signal peptidase, SipW and two proteins (BsaA
398 and BsaB) with homology to the main biofilm protein of *B. subtilis*, TasA (Fig. 2C). Because this

399 gene synteny is similar to the *sipW-tasA-tasB* biofilm operon in *B. subtilis*, we expressed the entire
400 operon to see if a biofilm matrix was produced and noted that high levels of a gel like matrix
401 material was formed in the culture supernatant (Fig. S3B-C). A high mol wt oligomer of BsaA
402 was seen in SDS-PAGE gels (Fig. S3D), which was similar to that seen by Obana et al in their
403 analysis of the same operon (13). Because BsaC was dependent on T4P-associated proteins for
404 secretion, it was logical to determine if secretion of the BsaA oligomer was as well. Two different
405 methods were tried, using densitometry of western blots of the BsaA oligomers in culture
406 supernatants and densitometry to measure the amount of BsaA in the culture supernatants using a
407 slot blot device. Although we saw significantly decreased levels of secretion in mutants using both
408 methods (Fig. S4 and S5), neither one gave conclusive evidence that a full T4P system was needed
409 for secretion. Both experimental approaches were hindered by relatively high amounts of
410 experimental variability in trying to measure an oligomer in solution; this can be seen by the high
411 range of the SEM determined for most of the samples (Fig. S4 and S5) despite multiple repetitions
412 of each experiments. However, the fact that several genes in each type of experiment was shown
413 to be necessary for efficient secretion, does strongly suggest that T4P proteins play a role in the
414 process.

415 Although we observed that a *pilA1* deletion strain did not secrete the biofilm matrix
416 oligomers efficiently (Fig. S4 and S5), we did not observe this phenotype in our initial screening
417 for proteins absent in the secretome, as shown for PilA3 in Fig. 2A-B. We think this is likely due
418 to the fact that our growth conditions for the experiment, overnight culture in BHI liquid medium
419 at 37 ºC, is not optimum for production of the biofilm matrix proteins, as described by Obana et al
420 (13). These findings point out the limitations of screening for T4P-dependent secretion: using only
421 a single type of environmental condition may not be the one in which a protein is made and secreted

422 so it would not be discovered. Even 2-D gels would be unlikely to have uncovered the presence of
423 the BsaA oligomers due to their extremely high mol weight. Therefore, it seems protein-based
424 mass spectrometry on samples from a wide variety of environmental conditions would be optimal
425 for detecting T4P-dependent secretion substrates.

426 Since our results and that of Obana et al both showed that oligomers of BsaA were
427 extremely resistant to heat and other denaturants (Fig, S3D and (13)), this suggested that covalent
428 bonding might be involved. Subsequent analysis using AlphaFold 3 indicated there were likely
429 isopeptide bonds being formed by donor strand exchange involving an N35 of one monomer and
430 a K74 of another, catalyzed by an E56 residue in the recipient monomer (Fig. 4). This model was
431 confirmed by subsequent substitution of each residue involved with an Ala, which led to a
432 complete lack of oligomerization (Fig. 5). However, isopeptide bonds are formed relatively slowly
433 (22) so we hypothesized that the donor strand exchange resulted in formation of a temporally stable
434 interaction that gave the isopeptide bonds time to form. However, based on experiments with the
435 N35A mutant, the oligomers are stable even without the isopeptide bond, as the mutant version
436 was resistant to denaturation in SDS-PAGE buffer and being run through an SDS-PAGE gel. Only
437 after heating in the SDS-PAGE buffer at 95 ºC did the N35 mutant depolymerize (Fig. S6), while
438 the WT oligomer withstood this treatment. We further characterized the formation of the oligomer
439 and showed, using FAST fluorescent protein fusions, that the oligomers were anchored to the
440 surface of the bacteria (Fig. 6) but via the cytoplasmic membrane and not directly to the PG layer
441 (Fig. S7).

442 BsaB, the other biofilm protein encoded in the operon, did not form oligomers, did not
443 become covalently linked to BsaA oligomers (Fig. S8) and both BsaA and BsaB have no apparent
444 LPXTG motif present in their protein sequence (data not shown). These structural features seem

445 somewhat at odds with the presence of a sortase gene (*srtB*) adjacent to, but oriented in the opposite
446 direction, of the *sipW* gene (Fig. 2C). Obana et al (13) tested the premise that SrtB was responsible
447 for oligomerization of BsaA, but found that a *srtB* mutant had normal levels of oligomer present
448 (13). However, examination of the *bsa* operon in other strains of *C. perfringens* reveals that a
449 different gene synteny is in place, where *sipW* is followed by a *bsaA* homolog and two *bsaB*-like
450 genes, the first of which carries an identifiable LPXTG motif (Fig. 2D). The BsaB sequence from
451 strain 13, when subjected to a BLAST analysis (34) against other *C. perfringens* strains, shows
452 proteins with high levels of homology to either the N-terminal half or C-terminal half of the strain
453 13 BsaB, but not both. This suggests strain 13 underwent a genetic rearrangement event in which
454 the two *bsaB* homologs were recombined into a single gene with the loss of the LPXTG motif
455 found in the first gene (Fig. 2D). This brings up the distinct possibility that the SrtB sortase could
456 catalyze the cross-linking of the first BsaB ortholog to the PG layer in strains other than strain 13.
457 What role BsaB plays in BsaA oligomerization is still unknown, since a *bsaB* mutant still formed
458 similar amounts of oligomers of BsaA as the wild-type strain (13). However, based on the
459 differences in composition of BsaB homologs between strain 13 and other strains (Fig. 2C-D), it
460 is possible that BsaB is involved in anchoring the BsaA oligomers to the PG layer via the activity
461 of SrtB and the BsaB homolog with an LPXTG motif, but this is speculative.

462 Taking all of these results into consideration, we have constructed a model showing the
463 mechanisms underlying the assembly of BsaA into cross-linked oligomers (Fig. 8). In this model,
464 BsaA monomers are secreted via the Sec pathway, fold on the outer surface of the CM but are
465 anchored in the membrane by their N-terminal signal sequence. An oligomer of BsaA with the
466 most recent monomer having its N-terminal disordered region exposed, contacts the newly
467 secreted monomer and undergoes a donor strand exchange in which the N-terminal disordered

468 region binds to the exposed beta-sheet fold and interdigitates as a newly formed beta sheet (Fig.
469 4A). This is followed by cleavage of the signal sequence by the signal peptidase SipW. We believe
470 the SipW acts after oligomer binding to the newly secreted monomer because the BsaA oligomers
471 appear to be anchored to the membrane (Fig. S7) and the oligomers are in a relatively stable
472 assembly even before the monomers are "locked" together by isopeptide bond formation (Fig. S6).
473 After SipW cleavage, the oligomer is released and can then be used to add another monomer or be
474 secreted through the PG layer to the external medium. The presence of significant amounts of
475 oligomers attached to the bacterial surface (Fig. 6) suggests the monomers assemble in the space
476 between the CM and PG rather than being secreted as monomers for oligomerization in the
477 environment. The mechanism for export of the oligomers through the PG layer is not complete but
478 there is evidence that T4P play at least some role in the process (Fig. S4 and S5).

479 The phylogeny of BsaA-like proteins shows they are widespread amongst Gram-positive
480 bacteria but also present in some Gram-negative bacteria and archaea (Fig. 7). However, the
481 homologs that are predicted to form isopeptide bonds are more narrowly focused in the Clostridia
482 and Enterococcus genera (Fig. 7). This brings up the question as to what advantage isopeptide
483 bond formation bestows on the biofilm matrix properties of the bacteria that form them. Formation
484 of isopeptide bonds can greatly increase the mechanical resistance to distension and separation of
485 the BsaA fibers. As an example, the intramolecular isopeptide bonds formed in sortase-dependent
486 pili made by Gram-positive bacteria have been shown to greatly increase the amount of force that
487 needs to be applied to distend the individual subunits, thereby strengthening the entire fiber (35).
488 To trap macrophages, neutrophils or other phagocytic cells, *C. perfringens* and other species may
489 have evolved a mechanism to form isopeptide bonds to increase the mechanical strength of the
490 fibers, since macrophages, neutrophils and other phagocytic cells can impart a large locomotive

491 force (1.9 to 10.7 nN) during their motility (36). It is also worth considering that *C. perfringens* is
492 an extremely ubiquitous bacterium in nature, found in soil, freshwater sediments and even extreme
493 environments like Antarctica (37, 38) and likely forms biofilms in these natural environments. The
494 BsaA biofilm matrix fibers could protect them from killing by amoeba, protozoa and other
495 predators found in these environments. Analyzing the relative mechanical strength of fibers with
496 and without isopeptide bonds will help answer some of these questions.
497

498 **Materials and Methods**

499 **Bacterial strains and growth conditions.** Bacterial strains, plasmids, and primers used in this
500 study are listed in Table 1. *Escherichia coli* strain DH10B was grown in Luria Bertani broth at
501 37°C for all transformations. When necessary, kanamycin and chloramphenicol were added to the
502 media at a concentration of 100 µg/ml and 20 µg/ml, respectively. *C. perfringens* strain HN13, a
503 Δ*galKT* derivative of strain 13, was used as the wild type strain in this study. *C. perfringens* strains
504 were grown anaerobically in PGY (30 g proteose peptone #3, 20 g glucose, 10 g yeast extract, 1 g
505 sodium thioglycolate per liter) or brain-heart infusion (BHI) (Thermo Fisher) in an anaerobic
506 chamber (Coy Laboratory Products, Inc.). Anaerobic egg yolk agar medium was prepared as
507 previously described (https://www.fda.gov/food/laboratory-methods-food/bam-media-m12-
508 anaerobic-egg-yolk-agar#).

509 **Construction of in-frame deletions of TFP-associated genes.** In-frame deletions of TFP-
510 associated genes were made using the method of Nariya et al. (39), modified as described in
511 Hendrick et al (40). The primers used to amplify the flanking DNA for each gene are listed in
512 Table S1. All deletions were confirmed by PCR across the deleted region. Complementation of
513 the in-frame *pilA3* mutant (Fig. 2A-B) was done using the lactose-inducible promoter vector

514 pKRAH1 (14) to express the *pilA3* gene. The construct was made by PCR using the primers listed
515 in Table S7.
516 **SDS-PAGE and staining of secretome proteins.** Cultures of strain HN13 were grown for 16 h
517 in BHI medium (i.e., into stationary phase) and the bacteria were removed by centrifugation in a
518 table top centrifuge. The culture supernatants were removed and passed through a 0.45 μM filter.
519 Proteins in the culture medium were precipitated using 10% trichloroacetic acid (TCA)
520 precipitation and washed twice in acetone. After drying, sample pellets were resuspended in water
521 and then 4X SDS-PAGE sample buffer (41) was added to a final concentration of 1X. Secretome
522 samples were separated on an SDS-PAGE gel and then stained with the fluorescent protein stain,
523 Sypro-Ruby according to the manufacturer's instructions (Thermo Fisher). Stained gels were
524 imaged with a Gel-Doc Imaging system (Bio-Rad, Inc).
525 **Cloning the *bsa* operon into the lactose-inducible expression vector, pKRAH1.** Because the
526 SipW signal peptidase is likely needed for processing the secretion signal on BsaC (13), to analyze
527 secretion of BsaC in TFP gene associated mutants, the entire *sipW-basA-bsaB-bcaC* operon was
528 expressed under control of the lactose inducible promoter in plasmid pKRAH1 (14) using primers
529 OHL145 (PstI site) and OHL146 (BamHI site) by PCR (Table 1 and Table S3). The PCR fragment
530 was ligated to the PCR cloning vector pGEM-T to make pHLL64. pHLL64 and pKRAH1 were
531 digested with PstI and BamHI and ligated to make pHLL65.
532 **Construction of a FAST protein fusion to BsaA.** The gene fusion was constructed using
533 overlapping PCR of the FAST coding region on plasmid pAG104 FAST (25) to the *bsaA* gene
534 linked to *sipW* (i.e., the *bsaB* and *bsaC* genes were absent). The FAST protein was fused to the C-
535 terminal end of BsaA because of the processing of the signal sequence on the N-terminal end. The
536 primers used to construct the gene fusion are listed in Table S5.

537 **Production of antibodies to BsaC, PilA3 and immunofluorescence microscopy.** Rabbit
538 polyclonal antibodies against the BsaC peptide sequence N-CNIEGDKISWDRGEN-C were
539 produced by GenScript USA, Inc. along with pre-immune serum as a negative control. Rabbit
540 polyclonal antibodies against the PilA3 peptide sequence N-CIKIDNSKDSIDLTR-C were
541 produced by Yenzyme Antibodies, LLC. along with pre-immune serum as a negative control. To
542 test for surface exposure of PilA3 pilins, the bacteria were fixed with 2.5% paraformaldehyde in
543 Dulbecco's phosphate buffer saline (DPBS), incubated with rabbit serum with anti-PilA3
544 antibodies, washed with DPBS and incubated with goat anti-rabbit antibodies fluorescently tagged
545 with Alexafluor 647 (Invitrogen). Stained cells were imaged on an Olympus IX71 fluorescent
546 inverted microscope equipped with a CoolSnap HQ2 CCD camera and DeltaVision deconvolution
547 and image analysis software.
548 **Quantitative immunoblotting of secreted BsaC.** Overnight cultures were grown in BHI
549 supplemented with chloramphenicol, and diluted 1:50 into fresh BHI medium pre-incubated at
550 37ºC. The cultures were grown for one hour then induced by adding lactose, to a final
551 concentration of 10 mM, to the growth media every 30 min for 3 h. Following induction, 1 mL of
552 culture supernatant was collected by centrifugation, and proteins precipitated using TCA, as
553 described above. Resulting precipitates were resuspended in 50 uL of SDS-PAGE loading buffer
554 with 100 mM DTT, the pH was adjusted using one µL of 1.5 M Tris-HCl, pH 9, if necessary. Five
555 µL of sample was loaded per well, and SDS-PAGE was performed. Following SDS-PAGE,
556 proteins were transferred to a polyvinylidene difluoride (PVDF) membrane utilizing the iBlot
557 system (Bio-Rad, Inc), and developed with serum containing anti-BsaC antibodies at 1:250
558 dilution, followed by a DyLight 550 labeled fluorescent secondary goat anti-rabbit antibody

559 (Thermo-Fisher). Imaging was performed on a Typhoon gel scanner, and mean pixel intensity
560 was measured for each band utilizing the Amersham Typhoon densitometry control software.
561 **Purification of BsaC.** The *bsaC* gene with a C-terminal His_6 tag was cloned into the expression
562 vector pET24a by PCR using the primers listed in Table S4. An overnight culture of *E. coli*
563 LOBSTR pET24a::BsaC-His_6 was used to inoculate 2 L of 37ºC pre-warmed terrific broth in a 4
564 L beveled flask. The culture was allowed to grow to 0.6 OD_{600} where it was induced with 1 mM
565 IPTG. After 6 hours of induction, the culture was centrifuged to remove cells (BsaC-His_6 is
566 secreted by *E. coli*), and the resulting supernatant was precipitated with 90% ammonium sulfate.
567 Precipitated protein was solubilized in PBS, and used to load on to an ÄKTA purifier FPLC for
568 Ni-NTA chromatography in PBS with 10 mM imidazole as loading buffer, and PBS with 1M
569 imidazole as elution buffer. Eluted fractions were concentrated with an Amicon Ultra 2 mL 10 K
570 cutoff centrifugation columns. Concentrated elution fractions were then further processed using a
571 Superose 6 gel filtration column linked to an ÄKTA purifier FPLC.
572 **Immunoblotting of BsaA oligomers.** BsaA oligomer secretion by *C. perfringens* was analyzed
573 using western blots and slot blots. For the immunoblot experiments, the proteins were blotted onto
574 a PVDF membrane while submerged in cold buffer which contained 3 mM Na_2CO_3, 10 mM
575 $NaHCO_3$ and 20% methanol.
576 **Immunoblotting Method One.** The membranes were placed in a Snap i.d. 2.0 system
577 (Millipore Sigma) which uses a vacuum system to pull solutions through the membrane (referred
578 to a swashing steps below). Once set up, a solution containing Tris-buffered saline (TBS) (pH 7.4),
579 0.1% Tween 20 (TBST), and 1% bovine serum albumin (BSA) was added as a blocker. Both
580 primary and secondary antibodies were added to separate aliquots of the solution containing TBST
581 and 1% BSA. In the first step, the primary antibody solution was incubated for 10 minutes on the

582 membrane, followed by three washings with TBST. Then the secondary antibody solution was left
583 to sit for 10 minutes on the membrane, followed by three washings with TBST. The primary
584 antibody utilized was OctA-probe H-5 (1:250 dilution) used to detect the FLAG epitope (Santa
585 Cruz Biotechnology). The secondary antibody utilized for these experiments was StarBright700
586 goat anti-mouse (Bio-Rad) (1:10,000 dilution). The fluorescently labeled immunoblots were
587 imaged using a ChemiDoc MP imaging system (Bio-Rad), and the intensity of the fluorescence
588 was measured using the ChemiDoc MP imaging software for densitometry measurements.

589 **Immunoblotting method 2**. Membranes were transferred to a container and blocked with
590 EveryBlot blocking Buffer (Bio-Rad). Antibodies were diluted with EveryBlot blocking buffer and
591 applied to membranes, followed by a minimum of five washes with TBST.

592 The primary antibodies utilized included OctA-probe H-5 (1:227 dilution) used to detect
593 FLAG, and HA-Tag F-7 (1:227 dilution) used to detect HA (both from Santa Cruz Biotechnology,
594 Santa Cruz, CA). Membranes submerged in EveryBlot blocking buffer with primary antibody were
595 incubated at 4°C overnight. The secondary antibody utilized for these experiments was
596 StarBright700 goat anti-mouse (Bio-Rad) (1:13,333 dilution), incubated for one hour at room
597 temperature. Images of the immunoblots were taken using a ChemiDoc MP imaging system (Bio-
598 Rad), to measure the intensity of the fluorescence.

599 **Slot blots to detect secretion of BsaA oligomers**. For slot blots, supernatants samples were
600 prepared by diluting supernatants with sterile BHI. For initial experiments, samples of 100% pure
601 culture supernatant, 50% culture supernatant and 50% blank media, 25% culture supernatant and
602 75% blank media, 10% culture supernatant and 90% blank media, 1% culture supernatant and 99%
603 blank media, and 100% blank media were prepared. For later experiments, samples of 50% culture
604 supernatant and 50% blank media, 20% culture supernatant and 80% blank media, 10% culture

605 supernatant and 90% blank media, 5% culture supernatant and 95% blank media, 1% culture
606 supernatant and 99% blank media, and 100% blank media were prepared. The 10% culture
607 supernatant samples were most often used to collect the raw intensity data for quantitative results.
608 A PVDF membrane was soaked in TBST prior to being loaded into the slot blot device
609 while still damp. The slot blot was loaded with TBST in each slot for 10 minutes. The TBST was
610 pulled through by vacuum, and then the diluted supernatants for the slot blots were loaded into the
611 individual slots of the device, and allowed to sit for 10 minutes. The supernatants were pulled
612 through by vacuum, and then the slot blot were loaded with TBST in each slot for another 10
613 minutes. The TBST was vacuumed off as well, and the membrane was developed using
614 immunoblotting method 2.
615 **Western blots on *bsaA* mutants.** Western blots to detect oligomerization of BsaA mutants were
616 performed by expressing the *sipW-bsaA-His*$_6$ genes under control of the lactose-inducible
617 promoter. Supernatants were separated form cell pellets by centrifugation. The proteins in the
618 supernatant were concentrated by TCA precipitation (as described above), suspended in SDS-
619 PAGE loading buffer with 100 mM DTT and heated at 95 °C for 20 min. The bacteria in the cell
620 pellet were suspended in PBS and disrupted by bead beating with 0.1 mm zirconia-silica beads.
621 The beads were then removed by centrifugation and the cell extract brought to 1 X SDS-PAGE
622 buffer with 100 mM DTT and heated at 95 °C for 20 min. After running on SDS-PAGE gels, the
623 proteins were transferred to a PVDF membrane and developed using immunoblotting method 2
624 above, except the primary antibody was mouse monoclonal anti-His6 antibody His.H8 (Santa Cruz
625 Biotechnology).
626 **Modeling BsaA Structure.** BsaA from *C. perfringens* (accession number WP_011009871) was
627 created as a homotrimer using ColabFold v1.3.0 (18) based on AlphaFold v2.2.0 (19). Amber

628 minimization was not enabled. In the top ranked model based on pLDDT PyMOL v2.5.0 was used
629 to analyze side chains of interdigitated beta strands for potential sidechain crosslinks. Images were
630 created using ChimeraX v1.2. (42, 43) The same trimer of BsaA was created in AlphaFold 3 (24)
631 to observe if the new version creates crosslinks, and all trimer models made did contain the
632 crosslinks. Structures of BsaA were created with AlphaFold 3 (24).
633 **Phylogeny of BsaA-like proteins.** BsaA-like protein sequences were identified with 3 iterations
634 of jackhmmer (44) (HmmerWeb version 2.41.2) (45) using BsaA (UniProt:Q8XN23) from *C.*
635 *perfringens* as the seed. The protein sequence similarity network of proteins identified either with
636 jackhmmer or as a match to IPR022121 was constructed using the EFI-EST tool
637 (https://efi.igb.illinois.edu/efi-est/) (46, 47) with an alignment score of 10. Nodes were collapsed
638 at a sequence identity of 100%. The network was visualized with Cytoscape
639 (https://www.cytoscape.org) (48) using the Prefuse Force Directed OpenCL Layout. Gene
640 neighborhoods were retrieved using the EFI-GNT tool (https://efi.igb.illinois.edu/efi-gnt/) (47).
641 The presence of a N-E-K triad was determined by manual inspection of the multiple sequence
642 alignment. For phylogenetic reconstruction, sequences were aligned with MAFFT on XSEDE
643 using CIPRES (49, 50), edited in Jalview (51), and analyzed with IQ-Tree (52-54). Maximum
644 likelihood trees were constructed using the LG+F+G4 substitution model and ultrafast bootstrap
645 (1000 replicates). Consensus trees were visualized, rooted at the midpoint, and annotated with
646 iTOL (55). Data files are available in Supplemental File Data Set 1.
647
648
649
650

Acknowledgments

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

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.

674 **References**

675 1. **Singh, P. K., J. Little, and M. S. Donnenberg.** 2022. Landmark Discoveries and Recent
676 Advances in Type IV Pilus Research. Microbiol Mol Biol Rev **86:**e0007622.
677 2. **Naskar, S., M. Hohl, M. Tassinari, and H. H. Low.** 2021. The structure and mechanism of
678 the bacterial type II secretion system. Mol Microbiol **115:**412-424.
679 3. **Korotkov, K. V., M. Sandkvist, and W. G. Hol.** 2012. The type II secretion system:
680 biogenesis, molecular architecture and mechanism. Nature reviews. Microbiology
681 **10:**336-351.
682 4. **Ayers, M., P. L. Howell, and L. L. Burrows.** 2010. Architecture of the type II secretion and
683 type IV pilus machineries. Future microbiology **5:**1203-1218.
684 5. **Desvaux, M., M. Hebraud, R. Talon, and I. R. Henderson.** 2009. Secretion and
685 subcellular localizations of bacterial proteins: a semantic awareness issue. Trends
686 Microbiol **17:**139-145.
687 6. **Anne, J., A. Economou, and K. Bernaerts.** 2017. Protein Secretion in Gram-Positive
688 Bacteria: From Multiple Pathways to Biotechnology. Curr Top Microbiol Immunol
689 **404:**267-308.
690 7. **Demchick, P., and A. L. Koch.** 1996. The permeability of the wall fabric of Escherichia coli
691 and Bacillus subtilis. Journal of bacteriology **178:**768-773.
692 8. **Varga, J. J., V. Nguyen, D. K. O'brien, K. Rodgers, R. A. Walker, and S. B. Melville.** 2006.
693 Type IV pili-dependent gliding motility in the Gram-positive pathogen *Clostridium*
694 *perfringens* and other Clostridia. Mol Microbiol **62:**680-694.
695 9. **Melville, S., and L. Craig.** 2013. Type IV pili in Gram-positive bacteria. Microbiol Mol Biol
696 Rev **77:**323-341.
697 10. **Bohning, J., M. Ghrayeb, C. Pedebos, D. K. Abbas, S. Khalid, L. Chai, and T. a. M. Bharat.**
698 2022. Donor-strand exchange drives assembly of the TasA scaffold in Bacillus subtilis
699 biofilms. Nat Commun **13:**7082.
700 11. **Erskine, E., R. J. Morris, M. Schor, C. Earl, R. M. C. Gillespie, K. M. Bromley, T.**
701 **Sukhodub, L. Clark, P. K. Fyfe, L. C. Serpell, N. R. Stanley-Wall, and C. E. Macphee.** 2018.
702 Formation of functional, non-amyloidogenic fibres by recombinant Bacillus subtilis TasA.
703 Mol Microbiol **110:**897-913.
704 12. **Stover, A. G., and A. Driks.** 1999. Secretion, localization, and antibacterial activity of
705 TasA, a Bacillus subtilis spore-associated protein. J Bacteriol **181:**1664-1672.
706 13. **Obana, N., K. Nakamura, and N. Nomura.** 2020. Temperature-regulated heterogeneous
707 extracellular matrix gene expression defines biofilm morphology in *Clostridium*
708 *perfringens*. NPJ Biofilms Microbiomes **6:**29.
709 14. **Hartman, A. H., H. Liu, and S. B. Melville.** 2011. Construction and characterization of a
710 lactose-inducible promoter system for controlled gene expression in *Clostridium*
711 *perfringens*. Appl Environ Microbiol **77:**471-478.
712 15. **Nunn, D., S. Bergman, and S. Lory.** 1990. Products of three accessory genes, pilB, pilC,
713 and pilD, are required for biogenesis of Pseudomonas aeruginosa pili. J Bacteriol
714 **172:**2911-2919.
715 16. **Kelley, L. A., S. Mezulis, C. M. Yates, M. N. Wass, and M. J. Sternberg.** 2015. The Phyre2
716 web portal for protein modeling, prediction and analysis. Nat Protoc **10:**845-858.

17. **Diehl, A., Y. Roske, L. Ball, A. Chowdhury, M. Hiller, N. Moliere, R. Kramer, D. Stoppler, C. L. Worth, B. Schlegel, M. Leidert, N. Cremer, N. Erdmann, D. Lopez, H. Stephanowitz, E. Krause, B. J. Van Rossum, P. Schmieder, U. Heinemann, K. Turgay, U. Akbey, and H. Oschkinat.** 2018. Structural changes of TasA in biofilm formation of Bacillus subtilis. Proc Natl Acad Sci U S A **115:**3237-3242.
18. **Mirdita, M., K. Schutze, Y. Moriwaki, L. Heo, S. Ovchinnikov, and M. Steinegger.** 2022. ColabFold: making protein folding accessible to all. Nat Methods **19:**679-682.
19. **Jumper, J., R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Zidek, A. Potapenko, A. Bridgland, C. Meyer, S. a. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, and D. Hassabis.** 2021. Highly accurate protein structure prediction with AlphaFold. Nature **596:**583-589.
20. **Kang, H. J., and E. N. Baker.** 2011. Intramolecular isopeptide bonds: protein crosslinks built for stress? Trends Biochem Sci **36:**229-237.
21. **Kang, H. J., F. Coulibaly, F. Clow, T. Proft, and E. N. Baker.** 2007. Stabilizing isopeptide bonds revealed in gram-positive bacterial pilus structure. Science **318:**1625-1628.
22. **Baker, E. N., C. J. Squire, and P. G. Young.** 2015. Self-generated covalent cross-links in the cell-surface adhesins of Gram-positive bacteria. Biochem Soc Trans **43:**787-794.
23. **Evans, R. O. N., M.; Pritzel, A.; Antropova, N.; Senior, A.; Green, T.; Žídek, A.; Bates, R.; Blackwell, S.; Yim, J.; Et Al.** 2022. Protein Complex Prediction with AlphaFold-Multimer. bioRxiv **2022**.
24. **Abramson, J., J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, S. W. Bodenstein, D. A. Evans, C. C. Hung, M. O'neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Zemgulyte, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A. I. Cowen-Rivers, A. Cowie, M. Figurnov, F. B. Fuchs, H. Gladman, R. Jain, Y. A. Khan, C. M. R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E. D. Zhong, M. Zielinski, A. Zidek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis, and J. M. Jumper.** 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature **630:**493-500.
25. **Plamont, M. A., E. Billon-Denis, S. Maurin, C. Gauron, F. M. Pimenta, C. G. Specht, J. Shi, J. Querard, B. Pan, J. Rossignol, K. Moncoq, N. Morellet, M. Volovitch, E. Lescop, Y. Chen, A. Triller, S. Vriz, T. Le Saux, L. Jullien, and A. Gautier.** 2016. Small fluorescence-activating and absorption-shifting tag for tunable protein imaging in vivo. Proc Natl Acad Sci U S A **113:**497-502.
26. **Tebo, A. G., F. M. Pimenta, Y. Zhang, and A. Gautier.** 2018. Improved Chemical-Genetic Fluorescent Markers for Live Cell Microscopy. Biochemistry **57:**5648-5653.
27. **Gautier, A., L. Jullien, C. Li, M. A. Plamont, A. G. Tebo, M. Thauvin, M. Volovitch, and S. Vriz.** 2021. Versatile On-Demand Fluorescent Labeling of Fusion Proteins Using Fluorescence-Activating and Absorption-Shifting Tag (FAST). Methods Mol Biol **2350:**253-265.

760 28. **Fives-Taylor, P. M., and D. W. Thompson.** 1985. Surface properties of *Streptococcus*
761 *sanguis* FW213 mutants nonadherent to saliva-coated hydroxyapatite. Infect Immun
762 **47:**752-759.
763 29. **Berry, J. L., I. Gurung, J. H. Anonsen, I. Spielman, E. Harper, A. M. J. Hall, V. J. Goosens,**
764 **C. Raynaud, M. Koomey, N. Biais, S. Matthews, and V. Pelicic.** 2019. Global biochemical
765 and structural analysis of the type IV pilus from the Gram-positive bacterium
766 *Streptococcus sanguinis*. J Biol Chem **294:**6796-6808.
767 30. **Gurung, I., I. Spielman, M. R. Davies, R. Lala, P. Gaustad, N. Biais, and V. Pelicic.** 2016.
768 Functional analysis of an unusual type IV pilus in the Gram-positive *Streptococcus*
769 *sanguinis*. Mol Microbiol **99:**380-392.
770 31. **Han, X., R. M. Kennan, D. Parker, J. K. Davies, and J. I. Rood.** 2007. Type IV fimbrial
771 biogenesis is required for protease secretion and natural transformation in *Dichelobacter*
772 *nodosus*. J Bacteriol **189:**5022-5033.
773 32. **Kirn, T. J., N. Bose, and R. K. Taylor.** 2003. Secretion of a soluble colonization factor by
774 the TCP type 4 pilus biogenesis pathway in *Vibrio cholerae*. Molecular microbiology
775 **49:**81-92.
776 33. **Whittaker, C. A., and R. O. Hynes.** 2002. Distribution and evolution of von
777 Willebrand/integrin A domains: widely dispersed domains with roles in cell adhesion and
778 elsewhere. Mol Biol Cell **13:**3369-3387.
779 34. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J.**
780 **Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database
781 search programs. Nucleic Acids Res **25:**3389-3402.
782 35. **Alegre-Cebollada, J., C. L. Badilla, and J. M. Fernandez.** 2010. Isopeptide bonds block
783 the mechanical extension of pili in pathogenic Streptococcus pyogenes. J Biol Chem
784 **285:**11235-11242.
785 36. **Guilford, W. H., R. C. Lantz, and R. W. Gore.** 1995. Locomotive forces produced by single
786 leukocytes in vivo and in vitro. Am J Physiol **268:**C1308-1312.
787 37. **Rood, J. I., and S. T. Cole.** 1991. Molecular genetics and pathogenesis of *Clostridium*
788 *perfringens*. Microbiol Rev **55:**621-648.
789 38. **Lisle, J. T., J. J. Smith, D. D. Edwards, and G. A. Mcfeters.** 2004. Occurrence of microbial
790 indicators and *Clostridium perfringens* in wastewater, water column samples, sediments,
791 drinking water, and Weddell seal feces collected at McMurdo Station, Antarctica. Appl
792 Environ Microbiol **70:**7269-7276.
793 39. **Nariya, H., S. Miyata, M. Suzuki, E. Tamai, and A. Okabe.** 2011. Development and
794 application of a method for counterselectable in-frame deletion in *Clostridium*
795 *perfringens*. Applied and environmental microbiology **77:**1375-1382.
796 40. **Hendrick, W. A., M. W. Orr, S. R. Murray, V. T. Lee, and S. B. Melville.** 2017. Cyclic Di-
797 GMP Binding by an Assembly ATPase (PilB2) and Control of Type IV Pilin Polymerization
798 in the Gram-Positive Pathogen *Clostridium perfringens*. J Bacteriol **199**.
799 41. **Sambrook, J., and D. W. Russel.** 2001. Molecular Cloning, A Laboratory Manual, 3rd
800 Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
801 42. **Goddard, T. D., C. C. Huang, E. C. Meng, E. F. Pettersen, G. S. Couch, J. H. Morris, and T.**
802 **E. Ferrin.** 2018. UCSF ChimeraX: Meeting modern challenges in visualization and
803 analysis. Protein Sci **27:**14-25.

43. **Pettersen, E. F., T. D. Goddard, C. C. Huang, E. C. Meng, G. S. Couch, T. I. Croll, J. H. Morris, and T. E. Ferrin.** 2021. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci **30:**70-82.
44. **Johnson, L. S., S. R. Eddy, and E. Portugaly.** 2010. Hidden Markov model speed heuristic and iterative HMM search procedure. BMC Bioinformatics **11:**431.
45. **Potter, S. C., A. Luciani, S. R. Eddy, Y. Park, R. Lopez, and R. D. Finn.** 2018. HMMER web server: 2018 update. Nucleic Acids Res **46:**W200-W204.
46. **Gerlt, J. A., J. T. Bouvier, D. B. Davidson, H. J. Imker, B. Sadkhin, D. R. Slater, and K. L. Whalen.** 2015. Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. Biochim Biophys Acta **1854:**1019-1037.
47. **Oberg, N., R. Zallot, and J. A. Gerlt.** 2023. EFI-EST, EFI-GNT, and EFI-CGFP: Enzyme Function Initiative (EFI) Web Resource for Genomic Enzymology Tools. J Mol Biol **435:**168018.
48. **Shannon, P., A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker.** 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res **13:**2498-2504.
49. **Katoh, K., K. Kuma, H. Toh, and T. Miyata.** 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res **33:**511-518.
50. **M. A. Miller, W. P. a. T. S.** Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gateway Computing Environments Workshop (GCE)**:**1-8.
51. **Clamp, M., J. Cuff, S. M. Searle, and G. J. Barton.** 2004. The Jalview Java alignment editor. Bioinformatics **20:**426-427.
52. **Kalyaanamoorthy, S., B. Q. Minh, T. K. F. Wong, A. Von Haeseler, and L. S. Jermiin.** 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods **14:**587-589.
53. **Nguyen, L. T., H. A. Schmidt, A. Von Haeseler, and B. Q. Minh.** 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol **32:**268-274.
54. **Hoang, D. T., O. Chernomor, A. Von Haeseler, B. Q. Minh, and L. S. Vinh.** 2018. UFBoot2: Improving the Ultrafast Bootstrap Approximation. Mol Biol Evol **35:**518-522.
55. **Letunic, I., and P. Bork.** 2024. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Res **52:**W78-W82.
56. **Grant, S. G., J. Jessee, F. R. Bloom, and D. Hanahan.** 1990. Differential plasmid rescue from transgenic mouse DNAs into Escherichia coli methylation-restriction mutants. Proc Natl Acad Sci U S A **87:**4645-4649.
57. **Andersen, K. R., N. C. Leksa, and T. U. Schwartz.** 2013. Optimized E. coli expression strain LOBSTR eliminates common contaminants from His-tag purification. Proteins **81:**1857-1861.
58. **Guzman, L. M., D. Belin, M. J. Carson, and J. Beckwith.** 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J Bacteriol **177:**4121-4130.

847 **Figure legends**

848 **Figure 1.** Diagram showing the type IV pili-associated genes in *C. perfringens* strain 13. Figure
849 derived from reference (9).

850

851 **Figure 2. (A).** SDS-PAGE gel showing the proteins in the secretome of the respective strains. The
852 gel was stained with the fluorescent protein dye Sypro-Ruby, the colors inverted to make the
853 protein bands in the samples appear dark but the pre-stained molecular weight markers appear as
854 bright bands. **(B).** Western blot showing the levels of secretion of the BsaC protein in different
855 strains. The secondary antibody was fluorescently tagged for visualization. Purified BsaC was
856 added in lane 5 as a positive control for detecting BsaC. **(C).** The gene synteny of the *bsa* operon
857 and adjacent sortase-encoding gene in *C. perfringens* strain 13. **(D).** The gene synteny of the *bsa*
858 operon and adjacent sortase-encoding gene in *C. perfringens* strain ATCC 13124 and most other
859 strains of *C. perfringens* that were observed. The sections of the *bsaB* gene from strain 13 that
860 were homologous to those in strain ATCC 13124 are labeled, along with the location of an LPXTG
861 sortase dependent sequence in one of the *bsaB*-encoding genes.

862

863 **Figure 3.** Results from densitometry analysis of western blots from the secretome of strains with
864 in-frame deletions of T4P-encoding genes. Each strain contained plasmid pGC1 for regulated
865 expression of the *bsa* operon with a His_6 tag added to the *bsaC* gene. The role of each gene product
866 in pilus function is shown in Fig. 1. For each experiment, the relative fluorescence intensity of the
867 wild-type strain (HN13) was set at 1.0 and the relative fluorescence intensity for each other strain
868 in the same experiment was then set relative to that value. Each strain was tested at least three
869 times with triplicate samples for each experiment. All of the experimental runs were combined and

870 one sample t and Wilcoxon tests were run to determine if the mutant strains were statistically
871 different from 1. Asterisks indicate strains that were different ($P < 0.05$) from strain HN13 (i.e.,
872 1).
873
874 **Figure 4. (A)** AlphaFold model of a homotrimer of BsaA in cartoon representation. Each
875 independent monomer is colored separately. Residues 28-38 of a monomer incorporate themselves
876 into the beta strands of an adjacent monomer. **(B)** Close-up view of interdigitated beta strands from
877 two monomers in cartoon representation, showing relevant side chains as sticks colored by
878 element. N35 and the giving chain are colored purple, E56, K74, and the receiving monomer are
879 colored cyan. Side chains forming the hydrophobic pocket are colored brown. The bond created
880 by AlphaFold between N35 and K74 is between an oxygen and nitrogen due to limitations in
881 AlphaFold; in reality, an amide bond would form between the side chains with ammonia as the
882 leaving group. **(C)** Close-up view of the N35-E56-K74 catalytic triad, showing the close proximity
883 of the catalytic E56 residue to the bond forming N35 and K74. All images were created with
884 ChimeraX v1.2. (42, 43).
885
886 **Figure 5.** Western blots showing loss of oligomer formation in BsaA with mutations in residues
887 involved in isopeptide bond formation**. (A)** Oligomer formation seen in TCA precipitated culture
888 supernatants or in the disrupted whole cell extracts of strains producing BsaA with single
889 mutations. **(B)** Comparison of BsaA oligomer formation in the N35A and N35A/K74A mutants in
890 a wild-type background (left side) and one in which the chromosomal copy of the *bsaA* gene was
891 deleted. Note that the chromosomal *bsaA* mutation did not change the oligomerization levels of

892 the mutants. **(C)** Western blot showing mutation of the E56 residue to Ala abolished
893 oligomerization of BsaA. P, disrupted cell pellet; S, TCA precipitate of culture supernatant.
894
895 **Figure 6.** BsaA oligomers are found on the bacterial surface. **(A)** Expression of a FAST protein
896 fusion to the cytoplasmic PilT protein (pSM413) only shows fluorescence with the membrane
897 permeable Coral dye. **(B)** A BsaA-FAST fusion protein (pSM412) fluoresces in the presence of
898 both a membrane permeable dye (Coral) and a membrane impermeable dye (Amber-NP),
899 indicating it is surface exposed to these dyes.
900
901 **Figure 7.** BsaA phylogeny**.** (**A**) A consensus tree of representative BsaA-like proteins. Leaf labels
902 are colored by taxonomy according to the color key. BsaA, BsaB, and two other BsaA-like proteins
903 identified in *C. perfringens* are labeled, as are TasA from *B. subtilis* and CalY from *B. cereus*.
904 Bootstrap values are represented as scaled circles according to the inset key. (**B**) Sequence
905 similarity network of identified BsaA-like proteins. Nodes representing labelled *Clostridium* and
906 *Bacillus* proteins in panel A are labeled with an arrow. Nodes are colored by taxonomy according
907 to the color key. Proteins with conserved residues that would be expected to be involved in
908 formation of a spontaneous intermolecular isopeptide bond are represented by triangle nodes with
909 a red outline. (**C**) Example gene neighborhoods of genes encoding BsaA-like proteins and genes
910 encoding putative surface filament structures. The complete matrix tree data set is included in a
911 Supplemental Information file.
912
913 **Figure 8.** Model of BsaA oligomerization. BsaA monomers (purple ovals) are oligomerized by
914 isopeptide bond formation after secretion via the Sec system in *C. perfringens*. SipW is the signal

915 peptidase for BsaA secretion (Fig. 2C and (13)). Secretion of the oligomers past the PG layer may

916 be due to the action of T4P-associated proteins (Fig. S4).

917

918 **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*	(39)
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–	(56)
Escherichia coli strain LOBSTR	Derivative of strain BL21, designed for low background binding to Ni-affinity columns	(57)
Plasmids		
pKRAH1	Cm^r; Lactose-inducible expression system plasmid for use in *C. perfringens*	(14)
pET24a	Protein expression vector	Novagen
pCRBluntIITOPO	PCR cloning vector	Invitrogen
pBAD30	Amp^r; Arabinose-inducible expression system plasmid for use in *E. coli*	(58)
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

919

920

Figure 1

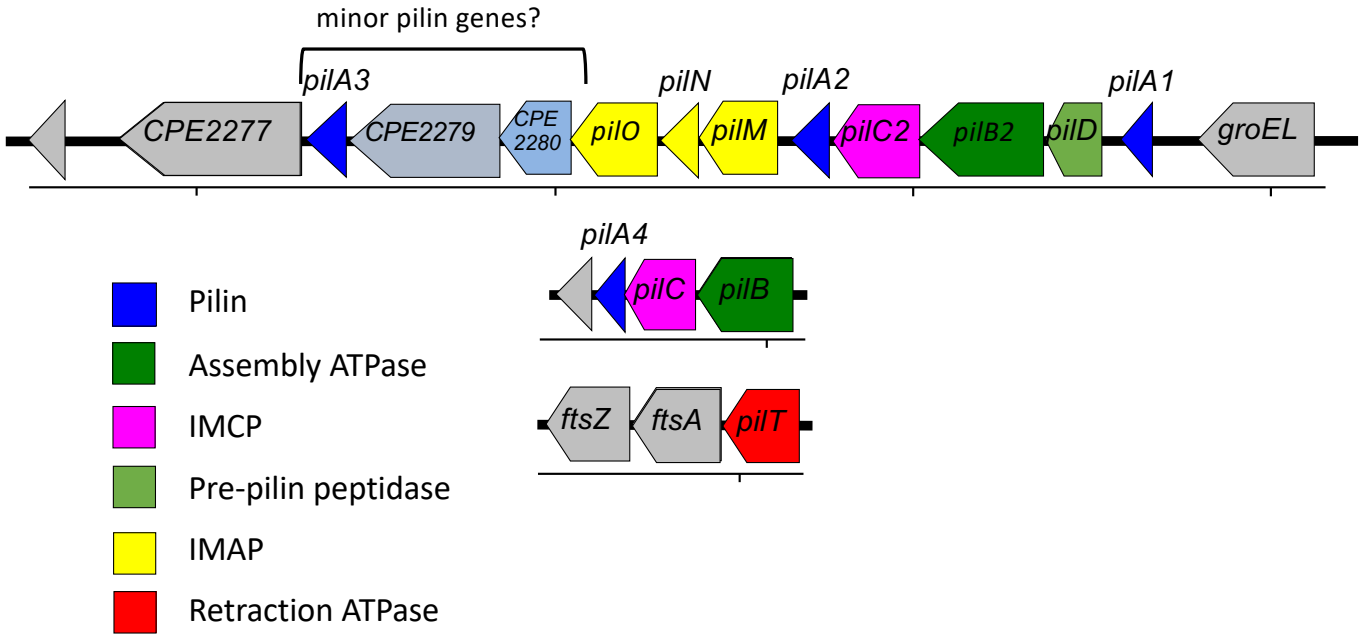

minor pilin genes?
pilA3
pilN
pilA2
pilA1
CPE2277
CPE2279
CPE 2280
pilO
pilM
pilC2
pilB2
pilD
groEL
pilA4
pilC
pilB
ftsZ
ftsA
pilT
Pilin
Assembly ATPase
IMCP
Pre-pilin peptidase
IMAP
Retraction ATPase

Figure 2

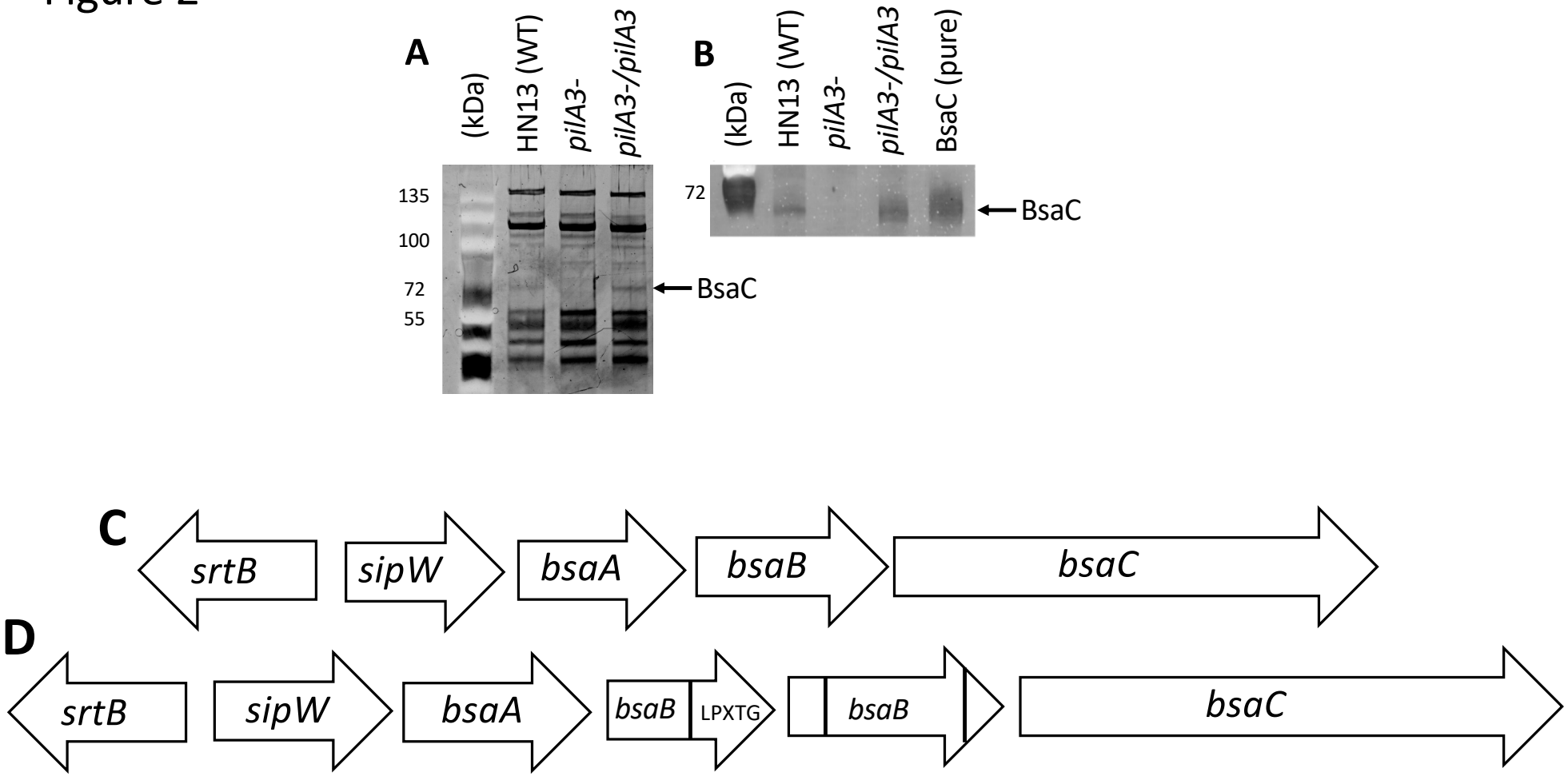

A
(kDa)
HN13 (WT)
pilA3-
pilA3-/pilA3
135
100
72
55
BsaC
B
(kDa)
HN13 (WT)
pilA3-
pilA3-/pilA3
BsaC (pure)
72
BsaC
C
srtB
sipW
bsaA
bsaB
bsaC
D
srtB
sipW
bsaA
bsaB
LPXTG
bsaB
bsaC

Figure 3

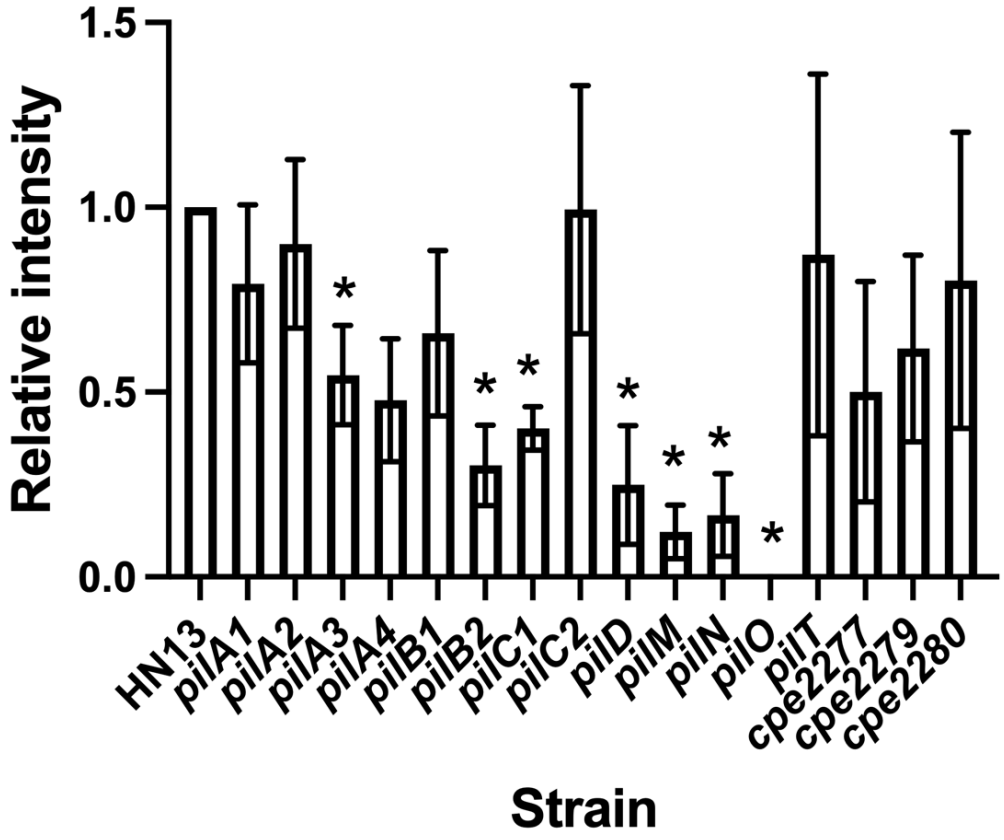
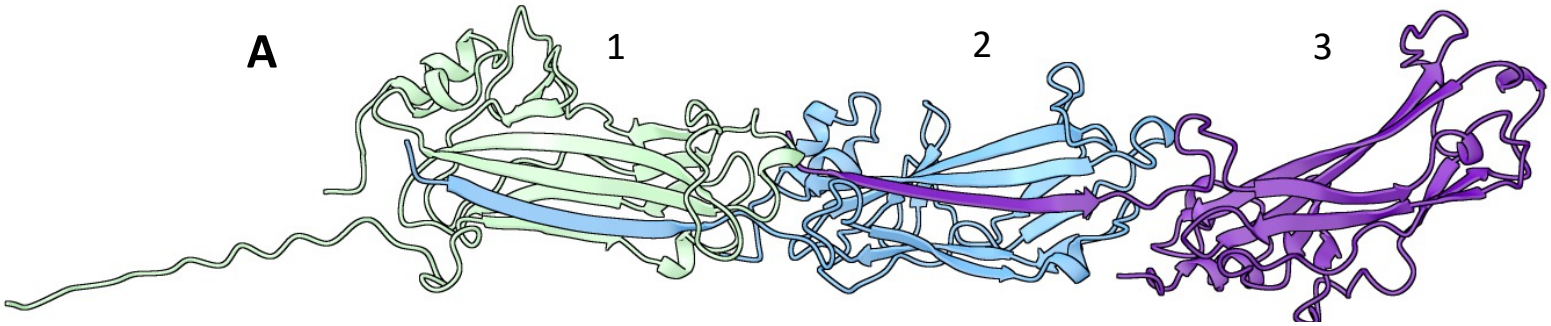
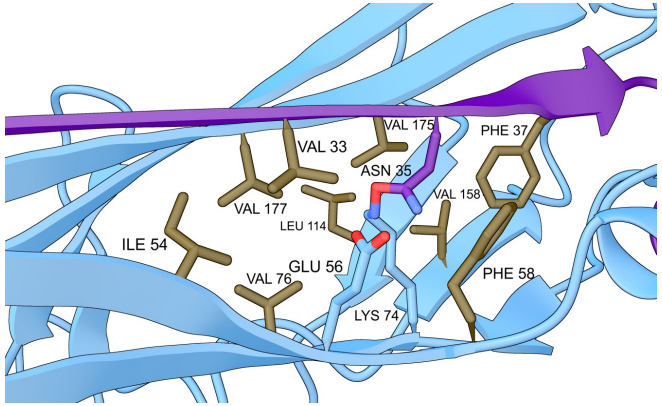

1.5
1.0
0.5
0.0
Relative intensity
*
*
*
*
*
*
*
HN13
pilA1
pilA2
pilA3
pilA4
pilB1
pilB2
pilC1
pilC2
pilD
pilM
pilN
pilO
pilT
cpe2277
cpe2279
cpe2280
Strain

Figure 4


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

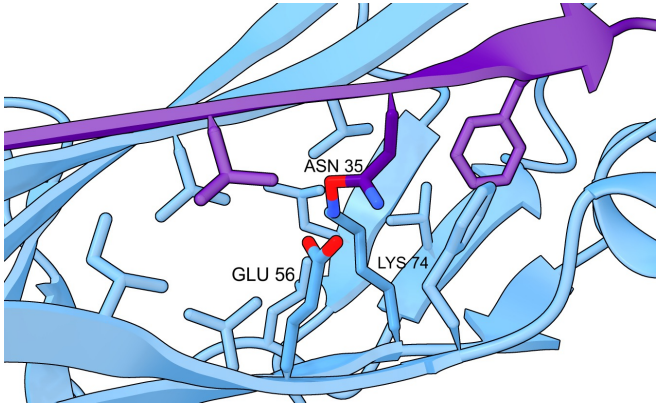

C
ASN 35
GLU 56
LYS 74

Figure 5

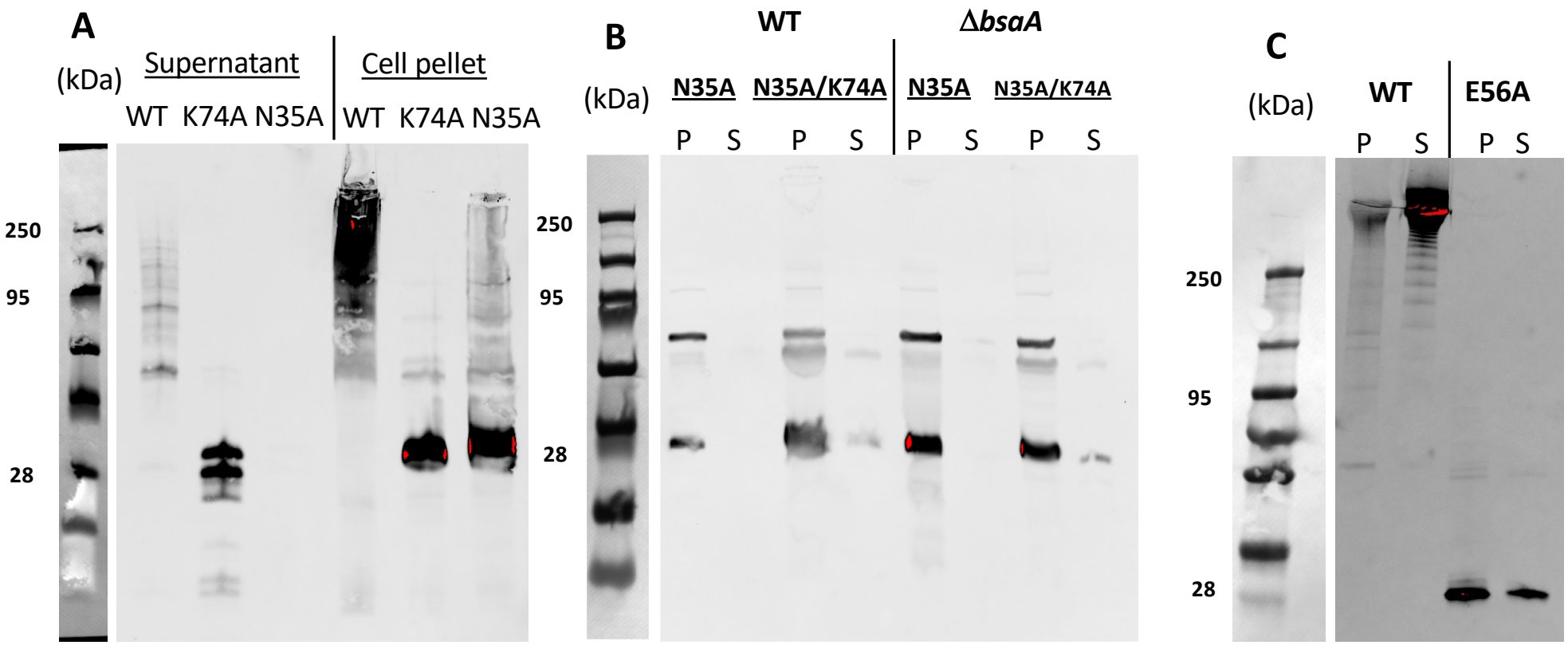

A
(kDa)
Supernatant
Cell pellet
WT K74A N35A
WT K74A N35A
250
95
28
B
(kDa)
WT
ΔbsaA
N35A
N35A/K74A
N35A
N35A/K74A
P S P S P S P S
250
95
28
C
(kDa)
WT
E56A
P S P S
250
95
28

Figure 6

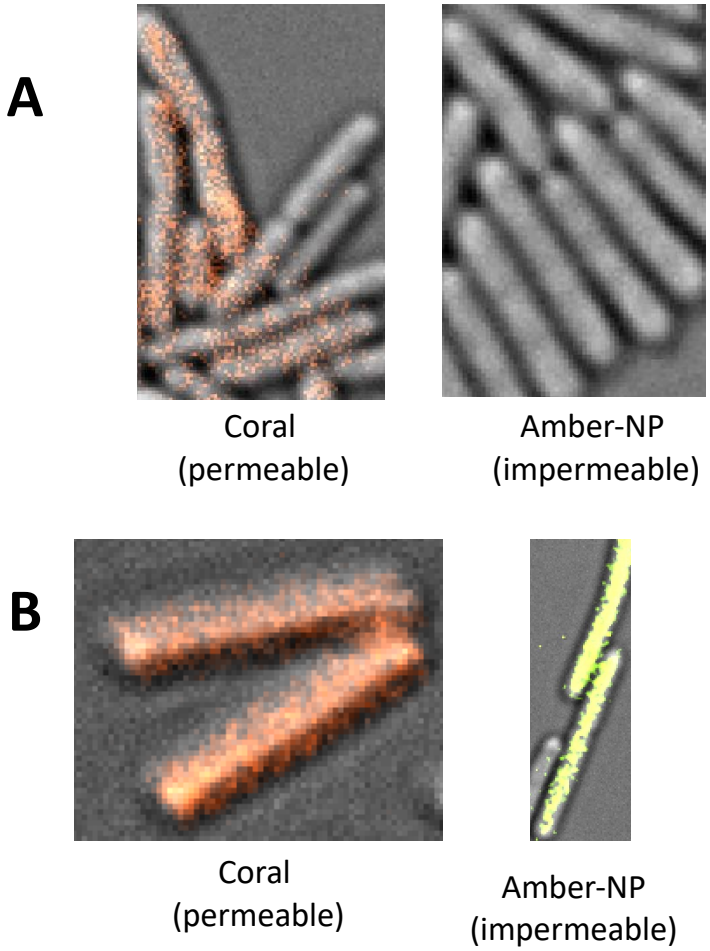

A
Coral
(permeable)
Amber-NP
(impermeable)
B
Coral
(permeable)
Amber-NP
(impermeable)

Figure 7

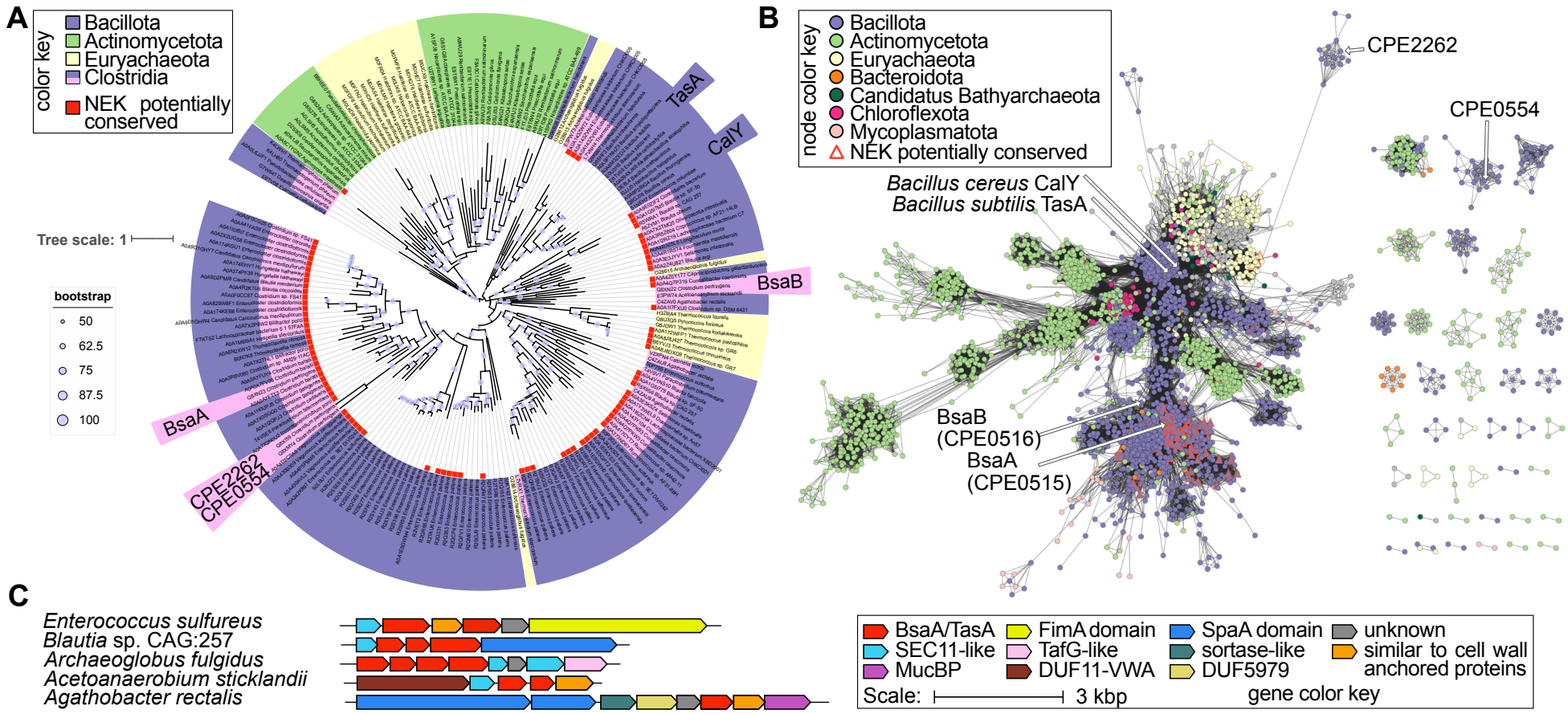
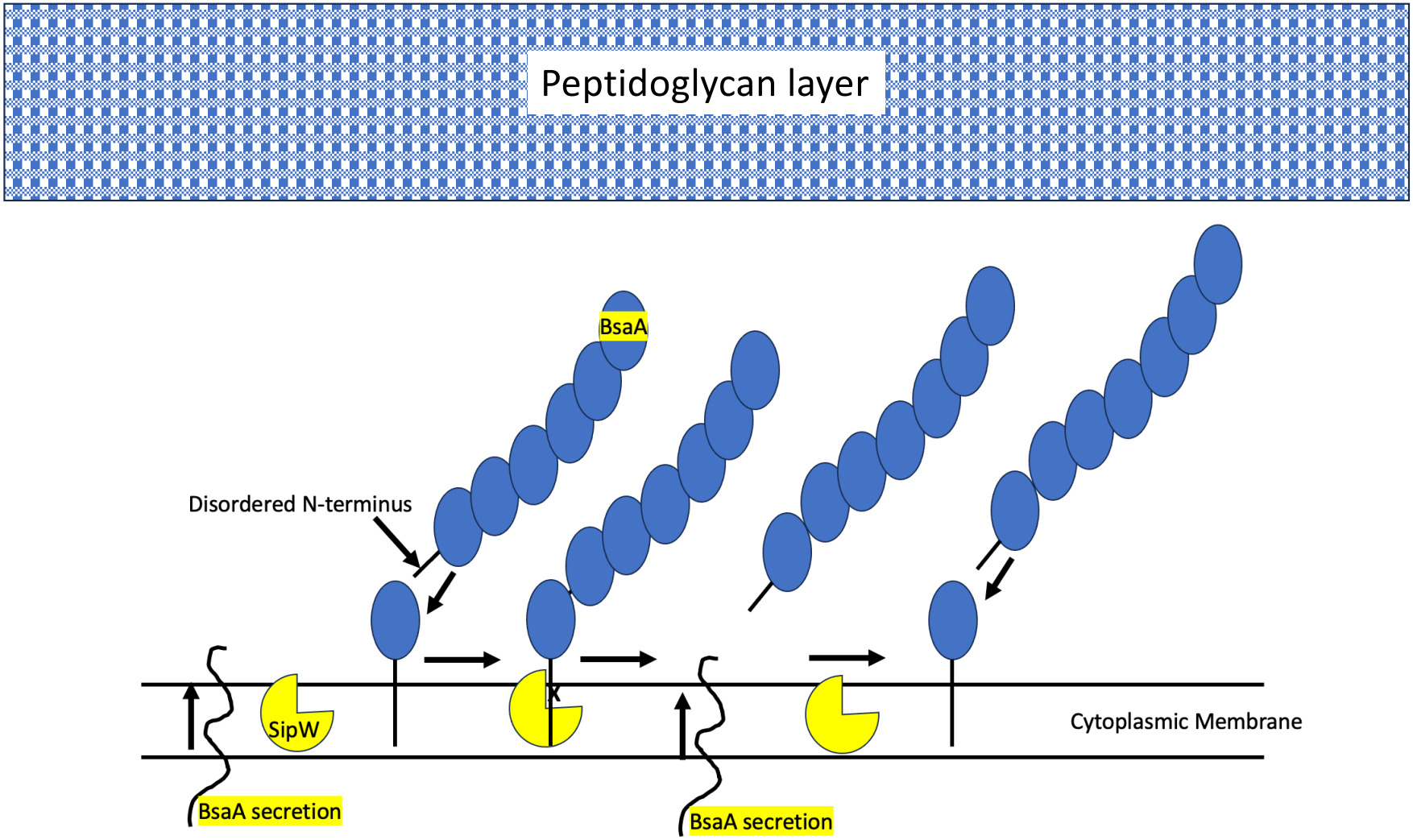

A
color key
Bacillota
Actinomycetota
Euryachaeota
Clostridia
NEK potentially conserved
Tree scale: 1
bootstrap
50
62.5
75
87.5
100
TasA
CalY
BsaB
BsaA
CPE2262
CPE0554
B
node color key
Bacillota
Actinomycetota
Euryachaeota
Bacteroidota
Candidatus Bathyarchaeota
Chloroflexota
Mycoplasmatota
NEK potentially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

Figure 8


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