1 CRYOEM STRUCTURE OF THE VIBRIO CHOLERAE TYPE IV
2 COMPETENCE PILUS SECRETIN PILQ

3 Sara J. Weaver[1,5] , Matthew H. Sazinsky[2], Triana N. Dalia[3], Ankur B. Dalia[3], Grant
4 J. Jensen[4].
5 [1]Division of Chemistry and Chemical Engineering, California Institute of
6 Technology, Pasadena, CA, USA, [2]Department of Chemistry, Pomona College,
7 Pomona, CA, USA, [3]Department of Biology, Indiana University, Bloomington, IN,
8 USA, [4]Division of Biology and Biological Engineering and Howard Hughes
9 Medical Institute, California Institute of Technology, Pasadena, CA, USA, [5]Present
10 address Howard Hughes Medical Institute, David Geffen School of Medicine,
11 Departments of Biological Chemistry and Physiology, University of California, Los
12 Angeles, CA, USA

Abstract

14 Natural transformation is the process by which bacteria take up genetic material
15 from their environment and integrate it into their genome by homologous
16 recombination. It represents one mode of horizontal gene transfer and contributes to
17 the spread of traits like antibiotic resistance. In *Vibrio cholerae*, the Type IV
18 competence pilus is thought to facilitate natural transformation by extending from
19 the cell surface, binding to exogenous DNA, and retracting to thread this DNA
20 through the outer membrane secretin, PilQ. A lack of structural information has
21 hindered our understanding of this process, however. Here, we solved the first ever
22 high-resolution structure of a Type IV competence pilus secretin. A functional
23 tagged allele of VcPilQ purified from native *V. cholerae* cells was used to
24 determine the cryoEM structure of the PilQ secretin in amphipol to ~2.7 Å. This
25 structure highlights for the first time key differences in the architecture of the Type
26 IV competence pilus secretin from the Type II and Type III Secretin System
27 secretins. Based on our cryoEM structure, we designed a series of mutants to
28 interrogate the mechanism of PilQ. These experiments provide insight into the

29   channel that DNA likely traverses to promote the spread of antibiotic resistance via

30   horizontal gene transfer by natural transformation. We prove that it is possible to

31   reduce pilus biogenesis and natural transformation by sealing the gate, suggesting

32   VcPilQ as a new drug target.

Introduction

34   Horizontal gene transfer, or the ability of microorganisms to directly share DNA

35   with one another, facilitates rapid evolution, can contribute to the development of

36   antibiotic resistance, promote the spread of virulence factors, and allow bacterial

37   pathogens to rapidly evade host immune response [1]. A clear understanding of the

38   mechanisms of horizontal gene transfer can aid the development of tools in the fight

39   against antibiotic resistance.

40   One mechanism of horizontal gene transfer is natural transformation, where a

41   competent bacterium can take up DNA from its environment and then maintain this

42   exogenous genetic material, either as a plasmid or by integrating it into the genome

43   by homologous recombination [2]. During transformation, DNA is actively brought

44   through the membrane. The requirements for DNA uptake vary with species, but

45   there is evidence for transformation of DNA from a few hundred base pairs up to

46   tens of thousands base pairs [2–5]. Intergenus transformation can result in the

47   development of mosaic alleles that confer antibiotic resistance, and has been

48   demonstrated in a variety of genera, including *Streptococcus*, *Neisseria*, and

49   *Actinobacter*[6,7].

50   Additionally, natural transformation of large regions of DNA can induce serotype

51   switching in *Vibrio cholerae* [8,9] and in *Streptococcus pneumoniae* [10]. The serotype

52   of a bacterial strain describes the immunodominant surface antigen that it displays,

53   and adaptive immune responses in human populations are generally serotype

54   specific. This has major public health implications, as serotype switching can help

55   bacteria evade the immune system. For example, after the introduction of a new

56   vaccine against a serotype of *S. pneumoniae*, epidemiologists traced a vaccine-

57   escape serotype to horizontal gene transfer [10]. Thus, preventing horizontal gene

58   transfer represents a unique approach to mitigate the spread of antibiotic resistance

59   and virulence in bacterial pathogens.

60   Here, we focus our attention to the structural biology of natural transformation in

61   the gram-negative bacterium *Vibrio cholerae. V. cholerae* is the causative agent of

62   the diarrheal disease cholera [11]. Since 1817, cholera has spread globally in seven

63   pandemics that each feature strains of distinct characteristics. Natural

64   transformation in *V. cholerae* is tightly regulated and induced when these bacteria

65   are grown on chitin in their aquatic environment [12]. Chitin, a biopolymer found in

66   the exoskeletons of crustaceans, induces transcription of the chitin regulon and

67   expression of the Type IV competence pilus machinery [13]. The Type IV

68   competence pilus machinery requires four elements: an inner membrane pilus

69   assembly complex, cytoplasmic motors to extend and retract the pilus, an outer

70   membrane pore (the secretin), and the pilus itself [14]. The Type IV competence pilus

71   facilitates environmental DNA uptake by extending and retracting from the cell

72   surface through a large, outer membrane secretin pore called PilQ [15]. To mediate

73   DNA uptake the pilus must translocate DNA across the membrane through the PilQ

74   secretin. Thus, the PilQ secretin represents a potential novel target to thwart this

75   mechanism of horizontal gene transfer.

76   The Type IV competence secretin is a member of the bacterial Secretin superfamily

77   [16,17]. Secretins are found in the Type II Secretion System (T2SS), the T3SS, the

78   Type IV pilus machine, and filamentous phage [18]. Bacterial secretins are united by a

79   common C-terminal secretin domain that oligomerizes to form a large pore in the

80   outer membrane [19]. The N-terminal protein domains form rings in the periplasm

81   and are thought to mediate interactions with other proteins [20,21]. While the C-

82   terminal secretin domain is remarkably similar across these secretion systems, the

83   N-terminal region varies (**Figure 1A**). The T2SS secretins have N-terminal protein

84   domains N0 to N3, followed by the secretin domain and then the S-domain. The

85   T3SS secretins lack the N2 domain, while the Type IV competence pilus secretins

86   lack the N1, N2, and S-domains. Additionally, the N-terminus of Type IV

87   competence secretins typically include one or more AMIN domains that interact

88   with the peptidoglycan [22]. These variations in domain architecture are likely related

89   to the specialization of different secretins [23]. The T2SS exports periplasmic folded

90   proteins to the extracellular space. The T3SS uses a "needle and syringe" to export

91   cytosolic effector proteins outside of the cell, or directly into another cell. The Type

92   IV competence pilus extends and retracts a filament to take up DNA cargo.

93   The structure of the Type IV competence pilus secretin has been examined several

94   times, though no work has provided high-resolution details sufficient to model the

95   passage of DNA or to design inhibitors of this potential drug target. In 2012, a 26 Å

96   structure of *Neisseria meningitidis* PilQ was solved by cryoEM [20]. In 2016, Koo *et*

97   *al.* solved the structure of *Pseudomonas aeruginosa* PilQ to 7.4 Å (pink in **Figure**

98   **2**). PaPilQ appeared to have 14 peripheral spokes lining the beta barrel that did not

99   demonstrate clear C14 symmetry, so C7 symmetry was assumed. In a 2017 paper,

100 D'Imprima *et al.* described the structure of *T. thermophilus* PilQ to ~7 Å resolution

101 with reported C13 symmetry (orange in **Figure 2**). D'Imprima *et al.* observed

102 flexibility between the N-terminal domains in TtPilQ that complicated the cryoEM

103 data analysis. Several high resolution cryoEM structures of bacterial T2SS and

104 T3SS secretins have been published, including *E. coli* GspD [24–26], *V. cholerae*

105 GspD [24,26], *Pseudomonas aeruginosa* XcpQ [27], *Klebsiella pneumoniae* PulD [28], *A.*

106 *hydrophila* (pilotin-independent ExeD), *V. vulnificus* (pilotin-dependent EpsD)[29],

107 and *Salmonella typhimurium* (InvG)[30]. Thus, to date, structural information about

108 the Type IV competence pilus secretin has mainly been inferred from the related,

109 but distinct T2SS and T3SS secretins.

110 Here, we present the structure of the *Vibrio cholerae* Type IV competence pilus

111 secretin PilQ to ~2.7 Å using a fully functional, His-tagged allele that we expressed

112 and purified from the native bacterium. Our work highlights differences between

113 the Type IV competence pilus, the T2SS, and the T3SS secretins, and emphasizes

114 the need for structures of different secretin family members. In particular we

4

115 discuss differences and remaining puzzles, including how the pilus could be
116 accommodated within PilQ during natural transformation and what part/s of the
117 secretin if any penetrate the outer membrane. Finally we report structure-inspired
118 designs of cysteine pair mutants that allowed us to reversibly inhibit pilus assembly
119 and natural transformation, presumably by sealing the secretin gate. These
120 experiments support the designation of VcPilQ as a druggable target, and more
121 broadly demonstrate how cysteine pair mutations can be employed to study the
122 activity of bacterial secretins.

123

Results

Purification of *V. cholerae* PilQ in amphipol

126 To ensure properly-folded and fully-function Type IV competence pilus machinery,
127 we chose to purify PilQ from *Vibrio cholerae* rather than a recombinant system. A
128 chromosomal mutation was made to add a deca-histidine tag to the N-terminus of
129 PilQ. The bacteria retained wild-type levels of natural transformation
130 (**Supplemental Figure 1A**). Previous work demonstrated that similar N-terminal
131 tags allowed for functional Type IV competence pilus activity [15].

132 *V. cholerae* cells expressing PilQ were lysed in the presence of the surfactant n-
133 dodecyl β-D-maltoside (DDM). After affinity purification, PilQ was exchanged into
134 an amphipol environment. Secretins are detergent- and heat-resistant multimers
135 [31,32]. SDS page analysis of PilQ in amphipol showed a high molecular weight band
136 near the top of the gel (marked with an asterisk in **Supplemental Figures 1B, D,**
137 **and E**). Western blotting against the histidine tag on PilQ confirmed the high
138 molecular weight band as PilQ (marked with an asterisk in **Supplemental Figures**
139 **1C and F**). The gel electrophoresis also demonstrated some low molecular weight
140 species in the purified sample, some of which are detected on the Western blot and
141 presumed to be PilQ monomer or proteolyzed PilQ species that retained the N-
142 terminal histidine tag. Other bands on the SDS-PAGE not labeled in the Western,

143   likely represent PilQ fragments lacking the N-terminal histidine or contaminant

144   proteins. Regardless, the size difference between these contaminants (<100 kDa)

145   and the PilQ multimer (~860 kDa) made it easy to distinguish PilQ from the milieu

146   in electron micrographs (**Supplemental Figure 3A)**.

147   To confirm that full length PilQ was present in our samples (residues 30-571 after

148   the 29 amino acid N-terminal signal peptide is cleaved), gel band analysis with a

149   trypsin digest and mass spectrometry was used to analyze the multimeric PilQ in

150   amphipol from a SDS page gel (band marked with an * in **Supplemental Figure**

151   **1D** was analyzed). The results demonstrated 65% sequence coverage, with

152   fragments identified in each domain of the folded protein from residues 50 to 567

153   (**Supplemental Figure 2**).

154   **Single particle cryoEM of the Type IV competence pilus secretin PilQ**

155   Here we report the high-resolution structure of the purified Type IV competence

156   pilus secretin VcPilQ by single particle cryoEM (**Figure 1**). The cryoEM data

157   processing steps are summarized as a flow chart in **Supplemental Figure 4** and in

158   Methods.

159   Briefly, *V. cholerae* PilQ (VcPilQ) was plunge-frozen on Quantifoil Holey Carbon

160   Grids with R2/2 spacing. A Titan Krios operating in a three by three pattern with

161   beam-image shift [33] was used to collect 3 movies per hole, resulting in 3,808

162   movies. The movies were motion corrected and screened for quality, which left

163   2,510 micrographs.

164   The cryoSPARC blob picker was used to identify 3,100,353 putative particle

165   location, several micrographs were inspected to adjust the threshold parameters to

166   select the top 252,319 particles for two-dimensional (2D) classification. 2D

167   classification demonstrated a wide variety of orientations, the clear presence of

168   secondary structure (**Supplemental Figure 3B**), and C14 symmetry

169   (**Supplemental Figure 3C**). After the initial 2D classification runs, 66,507 particles

6

170 were excluded from the dataset and used to generate a nonsense initial model

171 representing contamination. The 185,812 quality particles ('include') were used to

172 generate an initial model. A CryoSPARC heterogeneous refinement using the high

173 quality initial model and the contamination initial model was used to further clean

174 the dataset. Several rounds of 3D classification and 3D refinement were used to

175 select 100,543 particles for further analysis in Relion.

176 After polishing and CTF Refinement, the cryoEM structure (C14 symmetry, overall

177 resolution FSC@0.143 2.7 Å, FSC@0.5 3 Å, **Figure 2**) reached near-atomic

178 resolution and demonstrated isotropic resolution in the X, Y, and Z axes (3DFSC

179 sphericity 0.96) (**Supplemental Figure 5-6**)[34,35].

180 We were able to recognize and model residues 160 to 571 of PilQ (**Figure 1C,**

181 **Supplemental Movie 1**). Noting that residues 1-29 are cleaved (the N-terminal

182 signal peptide), our mass spectrometric analysis of the purified protein

183 (**Supplemental Figure 2**) demonstrated the presence of residues between 50 and

184 567, including the AMIN domain (residues 54-125) that is thought to interact with

185 the peptidoglycan in the periplasm [22]. While the AMIN domain was not resolved in

186 the PilQ structure, hazy density is present in the 2D classification in the region near

187 the PilQ N-terminus where we would expect the AMIN domain (marked with

188 asterisks in **Figure 1B**). Following the AMIN domain, residues 126-159 are

189 predicted by homology modeling to be unstructured (**Supplemental Figure 7**), so

190 they are also likely included in the hazy density. If that region is unstructured, it is a

191 likely source of the conformational heterogeneity observed as a hazy density in

192 **Figure 1B**. Extensive focused classification calculations were unable to improve

193 the resolution in this region (data not shown). Because the cell wall does not have a

194 crystalline structure, and the 14 AMIN domains around the ring must bind to it in

195 different places around the barrel, the AMIN domains are probably not regularly

196 arranged *in situ* either. In *Neisseria meningitidis*, the lipoprotein PilP is thought to

197 bind the AMIN domain and act as a bridge between the inner and outer membrane

198 components of the Type IV pilus machinery [36].

The structure of VcPilQ demonstrates similarities to other secretins

In each PilQ monomer, four beta strands come together to form a beta sheet (**Figure 1C**). Once assembled, PilQ forms a 56-strand beta-barrel. Inside the barrel, two further beta hairpins (beta strand-turn-beta strand each) form a gate. These regions match the topology of other secretin structures (**Figure 3A-C**). Chimera MatchMaker was used to align several published secretin structures to VcPilQ and calculate the root-mean-square deviation (RMSD) between the C-alpha carbons of each pair [37] (**Supplemental Figure 8**). In each case, the RMSD calculated over the protein domains present in both structures (N3, secretin, and beta lip domains) was about 1 Å. The outer membrane regions differ significantly, however, in both angle and membrane spanning distance, as discussed in more detail below.

While their structures are similar, the electrostatic characteristics of the inner surfaces of the T4P, T2SS and T3SS secretins vary (**Figure 3D**). The inner beta-lip, N3, and coil regions of PilQ (**Figure 1C**) are all highly negatively charged (**Figure 3D**). In contrast, in the *V. cholerae* GspD and *E. coli* K12 GspD structures, there are alternating negatively- and weakly positively-charged regions. These differences are likely related to the function of PilQ in natural competence: the DNA cargo of the Type IV competence pilus is also negatively charged, so it is possible that this electrostatic repulsion will help the cargo pass through the cavity, rather than getting stuck. By comparison, for T2SS secretins, the charge alternates in the inner cavity.

The N0 domain in VcPilQ agrees with previous structures

The N0 domain of VcPilQ (residues 160 to 227) was resolved 4 to 7 Å local resolution. This allowed us to build a model based on homology to previously solved structures of N0 domains (**Figure 1C**), which includes a crystal structure of isolated N0 domain from *Neisseria meningitidis* (4AR0) [20] (**Supplemental Figure 9**).

8

226 Only one of the seven T2SS secretins, and one of the T3SS published structures
227 resolves the N0 domain (Worrall et al. 2016, Chernyatina and Low 2019). Here we
228 can resolve the N0 domain (residues 160 to 227) of PilQ to 4 to 7 Å local
229 resolution, which allowed us to build a model based on homology to previously
230 solved structures of N0 domains (the cross-linked *Klebsiella oxytoca* PulD [38] and *S.*
231 *typhimurium* InvG [30]), plus a crystal structure of an isolated N0 domain from
232 *Neisseria meningitidis* PilQ (4AR0)[20] (**Supplemental Figure 10**). Our VcPilQ N0
233 structure (**Figure 1C**) falls within 1.2 Å RMSD of each of the previous structures
234 (**Supplemental Figure 10E**).

VcPilQ uses a novel helical coil to transition into the N3 domain

236 In VcPilQ, a 32 Å alpha-helix follows the N0 domain (**Figure 1C**). None of the
237 T2SS or T3SS structures contain a helical coil to link N-terminal domains; instead
238 their periplasmic protein domains are linked by unstructured loops (**Figure 3**). This
239 difference was not anticipated, as homology models of VcPilQ based on other
240 structures more closely match the previously published T2SS and T3SS structures
241 (**Supplemental Figure 7**).

242 Following the end of this helix, the protein chain abruptly changes direction (~104°
243 angle) as the coil flows into the N3 domain (**Figure 1C**). This dramatically reduces
244 the channel diameter, from 90 Å at the bottom of the N0 domain to 60 Å across the
245 N3 domain (**Figure 4**). In the T2SS and T3SS structures, the diameter of the
246 channel is relatively constant.

The putative outer membrane region of VcPilQ is thicker than T2SS secretins

248 The secretin amphipathic helix lip (AHL) is thought to be a key determinant for
249 secretin outer membrane insertion and among Type IV pilus secretins, the AHL is
250 conserved [18]. The AHL is thought to mark the lower boundary of the outer
251 membrane region of secretins [18,29,30]. The upper boundary of the outer membrane is

252   less clear since the micelle density in cryoEM reconstructions does not always

253   match the positions of aromatic residues [25].

254   In our VcPilQ atomic model, the distances between the bottom of the AHL to the

255   top of the beta strands is about 3 nm. As seen in **Figure 3C**, this putative outer

256   membrane region of VcPilQ is substantially taller than the same region in the

257   previously published T2SS structures. To investigate more than just the residue

258   locations, an inverted mask based on the atomic model density of VcPilQ was

259   generated and subtracted from the empirical cryoEM density, which reflects

260   unmodeled density in the cryoEM map that is not accounted for by the atomic

261   model (**Figure 4C**). This presumed micelle density (in grey) blooms around the

262   putative outer membrane region of the protein. This density appears both on the

263   outside of the beta barrel, and inside, coating the inner lip of PilQ. It remains

264   unclear which residues of PilQ are embedded within the outer membrane *in situ*

265   (see Discussion).

266   **Cysteine mutants indicate gate must open for pilus biogenesis and natural**

267   **transformation**

268   PilQ is thought to mediate DNA uptake during natural transformation in *V.*

269   *cholerae*. Blocking DNA uptake by inhibiting PilQ could prevent *V. cholerae* from

270   undergoing horizontal gene transfer in chitin biofilms, which can promote the

271   spread of antibiotic resistance genes and virulence factors. Toward this end, we

272   designed cysteine pair mutants to reversibly lock the gate with disulfide bonds. The

273   structure of VcPilQ was analyzed using the Disulfide by Design 2.0 web tool to

274   identify residue pairs with geometries that could support a disulfide bond [39–41]. The

275   top hits from this screen were analyzed in Chimera and introduced into the *V.*

276   *cholerae* genome. Two cysteine pairs (S448C/S453C and L445C/T493C) were able

277   to assemble into PilQ multimers (**Figure 5**). The S448C/S453C pair is in the

278   proximal hairpin of the gate, likely crosslinking adjacent PilQ monomers(**Figure**

279   **5C, 5E**). The L445C/T493C pair likely crosslinks the upper and lower gate hairpins

10

280   of a single PilQ monomer(**Figure 5D, 5F**),. In both cases, under oxidizing

281   conditions, the cysteine pair mutants demonstrated lower transformation efficiency

282   than the control PilQ strain (**Figure 6A, Supplementary Figure 8**). Also, both

283   mutants exhibited reduced piliation (**Figure 6B**). High levels of the reducing agent

284   DTT were toxic to WT *V. cholerae* cells and inhibited piliation (**Figure 6B-D**) and

285   natural transformation (**Supplementary Figure 8**). However, when the cells were

286   grown with varying subinhibitory concentrations of DTT, the cysteine pair mutants

287   recovered transformation efficiency and piliation similar to the control parent strain

288   (**Figure 6C-D**). The S448C/S453C mutant is slightly more recalcitrant to DTT

289   rescue, which may be due to the fact that these cysteines are located further down in

290   the gate region making them less accessible to the reducing agent.

291   **Discussion**

292   Here we present the first high-resolution structure of a bacterial Type IV

293   competence pilus secretin. We observed key differences in the outer membrane

294   region and the periplasmic region among the different members of the secretin

295   family. These differences identify weaknesses in relying on homology models of

296   evolutionarily related secretins, like the T2SS secretin GspD, to understand PilQ.

297   **Homology modeling was insufficient to predict a Type IV Pilus Secretin**

298   **structure**

299   Before we solved the VcPilQ structure, we used homology modeling to predict the

300   structure of VcPilQ based on its sequence and the structures of previously solved

301   secretins [42]. The top five predictions are shown **Supplemental Figure 7**.

302   Comparing these predictions to our now-solved structure reveal significant

303   differences, including the thick outer membrane domain of VcPilQ and the extent

304   of the alpha helix between the N0 and N3 domains (**Supplemental Figure 7**).

305   Additionally, homology modeling cannot predict the three-dimensional

306   arrangements and orientations of the protein domains, the number of subunits, or

307   the inner barrel diameter. These features can be hypothesized based on T2SS or

11

308   T3SS secretin structures, but that strategy would neglect the clear variations in

309   selectivity and functionality across the Secretin superfamily. It is possible that

310   structural variability between secretins also indicates which regions of the protein

311   are more often repurposed by evolution to generate new function and/or selectivity.

312   **Outer membrane-spanning domain**

313   Curiously, in some previous secretin structures, the trans-outer-membrane region

314   appears to be only 2-3 nm thick (**Figure 3C**), which is much thinner than a typical

315   membrane. This has left it unclear which regions of secretin molecules are actually

316   embedded in the outer membrane, and whether any residues are exposed to the

317   extracellular surface. To investigate these questions, we used the Positioning of

318   Proteins in Membrane (PPM) web server to predict the location of the

319   transmembrane region of VcPilQ [43]. The PPM web server result delineates the

320   hydrophobic region of the atomic model, which is shown by the dark horizontal

321   lines parallel to the membrane in **Figure 7**. The full lipid bilayer membrane will

322   then extend above and below these hydrophobic boundaries to account for the polar

323   head groups of the lipids. As we predicted based on the micelle density (**Figure

324   4B**), the lower boundary of the hydrophobic region is the amphipathic helix lip. The

325   upper hydrophobic boundary encompasses the beta strands of the beta lip, so the

326   loops likely protrude into the polar head groups.

327   Next we docked the VcPilQ structure into a sub-tomogram average of a closely

328   related PilQ imaged *in situ*. Recently, Chang et al. undertook an exhaustive

329   cryoelectron tomography (cryoET) study of the T4aP machinery in *Myxococcus*

330   *xanthus* (MxT4aP) [44]. In the resulting sub-tomogram averages of the T4aP

331   machinery, the location of each component protein was identified. The domain

332   organization of MxPilQ is nearly identical to that of VcPilQ[45] (**Supplemental

333   Figure 11A**), but it has two more AMIN domains at its N-terminus (**Supplemental

334   Figure 11A**). The Phyre2 protein model of MxPilQ also looks remarkably similar

12

335   to the structure of VcPilQ solved here (**Supplemental Figure 11B**)[46], even though

336   the Phyre2 model was built independently.

337   We therefore docked the atomic model of VcPilQ into the sub-tomogram averages

338   of the non-piliated MxT4aP in *M. xanthus* Δ*pilP* (EMD-3249), the non-piliated

339   MxT4aP in *M. xanthus* Δ*pilB* (EMD-3260), and the piliated MxT4aP in wild-type

340   *M. xanthus* (EMD-3247) (**Figure 7A-C**)[44]. Here we focus on the Δ*pilP* MxT4aP

341   sub-tomogram average (EMD-3249) because in this mutant only PilQ and TsaP

342   localize correctly, and are therefore the only proteins likely to be present in the sub-

343   tomogram average (**Figure 7A**)[44]. The putative location of TsaP suggested in

344   Chang *et al.* 2016 is marked with an asterisk in **Figure 7**. In *M. xanthus*, TsaP is a

345   peptidoglycan-binding protein[47]. In *V. cholerae*, the TsaP homolog LysM has not

346   been implicated in Type IV competence pilus function. The Δ*pilB* MxT4aP sub-

347   tomogram average (EMD-3260), which lacks the extension ATPase *pilB*, is also

348   shown because it was used to build a model of the non-piliated MxT4aP system

349   (PDB 3JC9) in Chang *et al.* 2016 (**Figure 7B**).

350   The atomic model of VcPilQ was docked into the MxT4aP sub-tomogram averages

351   using the predicted hydrophobic boundaries. At a high cryoEM density threshold,

352   the lipid bilayer is clearly observed as two leaflets in the MxT4aP sub-tomogram

353   averages (left in **Figure 7A-C**), so there is little doubt where the hydrophobic

354   region of VcPilQ should be placed in the sub-tomogram average. The bilayer

355   produces stronger features in the MxT4aP sub-tomogram averages than the protein,

356   so a lower cryoEM density threshold is required to visualize the full MxT4aP

357   machinery (right in **Figure 7**). In the non-piliated sub-tomogram averages (**Figure

358   7A-B**), VcPilQ agrees well with the MxPilQ density. The piliated MxT4aP clearly

359   undergoes a conformational change to accommodate the pilus in which the distance

360   between the inner and outer membranes increases (**Figure 7C**). In that case the

361   VcPilQ N0 domain doesn't extend far enough into the periplasm to line up with the

362   MxPilQ N0 domain (labeled in **Figure 7C**).

363   The gate region of VcPilQ nicely superimposes onto the gate of the *M. xanthus*

364   secretin, which suggests the VcPilQ structure was positioned correctly. However,

365   the putative outer membrane region in VcPilQ does not extend across the entire

366   outer membrane of the MxT4aP sub-tomogram averages. The thickness of the outer

367   membrane in the sub-tomogram averages ranges from 52 to 85 Å depending on the

368   cryoEM density threshold (labeled in **Figure 7A-C**), whereas the amphipol micelle

369   region in VcPilQ seems to be between 25 and 50 Å thick (**Figure 4B**). While this

370   discrepancy may be partially explained by the observation that the outer membrane

371   is about 5 nm thick in *M. xanthus* and 4 nm thick in *V. cholerae*[48], the shortness of

372   the other predicted trans-membrane domains of secretins (**Figure 3C**) and the two

373   bumps outside the outer membrane in the sub-tomogram average directly above the

374   secretin barrel (arrows in **Figure 7A** and **B**) call into question whether and how

375   parts of PilQ may penetrate the outer membrane *in situ*.

376   **VcPilQ could accommodate a Type IV Competence Pilus with gate**

377   **rearrangement**

378   The Type IV competence pilus machinery extends and retracts the Type IV

379   competence pilus through the PilQ outer membrane pore. In our structure of

380   VcPilQ in the absence of the pilus and accessory proteins, the inner diameter of the

381   channel ranges from 25 to 108 Å (**Figure 4A**).

382   The *V. cholerae* Type IV competence pilus is a Type IVa pilus. The major pilin

383   subunit in the T4aP is typically smaller than the major pilin subunits in the T4bP [49].

384   CryoEM structures of several Type IVa Pili (T4aP) have been solved from *N.*

385   *meningitidis* (NmT4aP, 6 Å resolution, 6 to 7 nm diameter), *N. gonorrhoeae*

386   (NgT4aP, 5 Å resolution, 6 nm diameter), *E. coli* (EHEC strain, EcT4aP, 8

387   Å resolution, 6 nm diameter), and *P. aeruginosa* (PaT4aP, 8 Å resolution, 5 nm

388   diameter) [50–52]. Additionally, a recent preprint on bioRxiv studying Type IV pili in

389   *T. thermophilus* demonstrated the presence of narrow (4.5 nm) and wide (7 nm) pili

390   [53]. In **Figure 7D**, the EcT4aP cryoEM structure (yellow) is fit into our VcPilQ

391   structure (in grey)[52]. The gate region with diameters of 25 Å at the lower gate and

392   36 Å in the upper gate clashes with the pilus (**Figures 4A and 7D**). VcPilQ was

393   solved in the absence of the pilus, so this gate conformation likely represents a

394   "closed" state in cases where the pilus is absent or fully retracted. The next

395   narrowest region in the VcPilQ channel is across the periplasmic N3 ring (~60 Å in

396   **Figure 4A**). Some of the previously solved T4aP are 6 to 7 nm in diameter, which

397   would be a tight fit in this region. The N3 domain is connected to the secretin

398   domain by a loop, so it could possibly expand to accommodate a 70 Å pilus.

399   We used homology modeling to predict the VcT4aP monomer structure, but

400   because Neuhaus et al. reported both wide and narrow pili assemebled by pilins

401   whose subunit structure are almost identical, we do not feel confident guessing the

402   VcT4aP diameter[53]. Regardless, the gates in VcPilQ would have to move to

403   accommodate a pilus. Consistent with this notion, our disulfide locked cysteine pair

404   mutants could not extend pili in the absence of reducing agent, which strongly

405   suggests that the conformation adopted in the presented structure (in which these

406   cysteine pairs would be in close enough proximity to disulfide bond) represents the

407   closed gate conformation and cannot accommodate a pilus fiber. Additionally, the

408   sub-tomogram averaging of the MxT4aP demonstrates a secretin conformational

409   change with and without a pilus present (**Figure 7B-C**)[44].

410   In Yan *et al.*, a glycine in the VcGspD gate (G453) was identified as a putative

411   hinge point to facilitate gate opening, and showed that the G453A mutant trapped

412   the T2SS secretin in a partially open state[24]. In our VcPilQ structure, the inner

413   channel distance between the corresponding glycine (G439) alpha carbons is about

414   8 nm (**Figure 4A**). Thus, we hypothesize that a gate hinge mechanism could

415   accommodate pili up to 7 nm in diameter. In an open state the gate loops could flip

416   up toward the extracellular space to accommodate the pilus.

417   **Conclusion**

15

418   Here we report the first high-resolution structure of a Type IV competence pilus
419   secretin, *V. cholerae* PilQ. This protein complex facilitates DNA uptake into
420   diverse bacterial species to aid in their evolution, and thus, represents a potential
421   target for therapeutic intervention. The *V. cholerae* Type IV competence pilus is a
422   model system to study natural transformation in bacteria. We identify key
423   differences between VcPilQ and the previously published structures of T2SS and
424   T3SS secretins. We designed cysteine pair mutants to reversibly seal the VcPilQ
425   gate and inhibit natural transformation, which can be used as a tool to further
426   investigate the function of the Type IV competence pilus machinery *in situ*. We
427   suggest a structural rearrangement that would transition our closed VcPilQ into a
428   piliated state that could accommodate previously solved structures of Type IV pili.
429   We also compare our structure to previous T4aP sub-tomogram averaging results in
430   *M. xanthus* which call in question which parts if any of PilQ extend across the outer
431   membrane. Together, these results elucidate the structure of PilQ and provide
432   mechanistic insight into natural competence in *V. cholerae*.

433

Figure Captions

Figure 1: CryoEM structure of the *V. cholerae* Type IV Competence Pilus Secretin PilQ

(A) Protein domain organization in the Type IV competence pilus machine, the Type II Secretion System (T2SS) and the Type III Secretion System (T3SS). The AMIN domain (red), the N0 (or STN) domain (blue), the Secretin_N domains N1 (purple), N2 (brown), and N3 (green), the secretin domain (yellow) and the S domain (pink) are shown.

(B) Example 2D classes show PilQ from the side (left) and from the top (right). The outer membrane is indicated with OM. Asterisks indicate the putative AMIN domains.

(C) The atomic model of the *V. cholerae* PilQ multimer is shown (right). One chain is shown by itself (left). The outer membrane is represented in a light blue rectangle. Dashed lines represent the different domains of the PilQ structure: AMIN (not shown), N0 (grey/dark blue), coil (purple), N3 (green), Secretin (yellow), Beta-lip (sky blue). The gate is labeled on the monomer.

(D) Model and cryoEM density of symmetrized PilQ is shown from the side (left) and cut through the center (right).

Figure 2: Comparison of VcPilQ to *P. aeruginosa* and *T. thermophilus* PilQ structures

Comparison of cryoEM density of PilQ structures from *V. cholerae* (left, grey)(this study), *Pseudomonas aeruginosa* (magenta, center) (Koo et al. 2016), and *Thermus thermophilus* (orange, right) (D'Imprima et al. 2017). In **(A)**, the inside cavity of the cryoEM densities are shown, while **(B)** highlights the gate region with a cut through the center of the density. The membrane is blue.

17

Figure 3: Comparison of VcPilQ to T2SS and T3SS secretins

The structures of VcPilQ (this study, dark red), *V. cholerae* GspD (5WQ8) (green), *E. coli* K12 GspD (5WQ7) (blue), and *S. typhimurium* InvG (6DV3) (yellow) (Jurrus et al. 2018; Worrall et al. 2016; Yan et al. 2017) are compared. The putative outer membrane location is depicted as a blue rectangle. The structures are shown as multimers **(A)** or monomers **(B).** In **(C)**, the multimer is shown in tan with one subunit colored. The outer membrane region is highlighted to show that VcPilQ (left, red) has a ~30 Å membrane spanning distance, while VcGspD, EcGspD, and StInvG have about a 24 Å membrane spanning distance.

(D) Adaptive Poisson-Boltzmann Solver was used to calculate the electrostatic potential calculation of each secretin (Jurrus et al. 2018). The inner cavity of each secretin is shown. The scale varies from -5 (red) to +5 (blue) in units of K_bT/e_c.

Figure 4: Dimensions of the VcPilQ inner channel and outer membrane region

(A) A central slice of the atomic model of *V. cholerae* PilQ colored as in **Figure 1**. The inner cavity distances are depicted with dashed lines and labeled in ångstroms. The outer membrane is depicted as a blue rectangle.

(B) VcPilQ micelle density with putative outer membrane diameter labeled in Å. The non-protein cryoEM density of VcPilQ is shown in grey at four different Chimera thresholds (left to right). The atomic model of VcPilQ is shown in blue and yellow according to the color scheme of **Figure 1**. In each image, yellow spheres mark the approximate top and bottom of the micelle density. The distance between the yellow spheres (dashed line) is shown below each image in Ångstroms.

Figure 5: Design of cysteine pair mutants to lock the gate

The atomic model of VcPilQ is shown as one chain **(A)** or as a multimer **(B)**. In **(C)** and **(E)**, the S448C/S453C mutant, which locks the lower gate, is shown. In **(D)** and **(F)**, the L445C/T493C mutant, which links the upper and lower gates, is shown. In

485　**(A)**, the upper and lower gate are highlighted in a dashed line box. A zoomed in

486　region of this box is shown in **(C)** and **(D)** with the cysteine pair mutant residues

487　highlighted in red. In **(B)**, the dashed line rectangle highlights the gate region,

488　which is shown on the left after a 90˚ rotation about the x axis. The coordinate

489　system is shown. This gate region is shown in (E) and (F) with the cysteine pair

490　mutant residues highlighted in red.

491　**Figure 6: Cysteine pair mutants lock the gate, reduce piliation, and reversibly**

492　**inhibit transformation**

493　**(A)** The transformation frequency of two cysteine pair VcPilQ mutants

494　(S448C/S453C on the left and L445C/T493C on the right) is plotted. Data are

495　normalized to the parental strain that expresses wild-type VcPilQ. Natural

496　transformation assays were performed in the presence of varying concentrations of

497　DTT (0 to 1.0 mM). Data are from at least 4 independent biological replicates and

498　shown as the mean ± SD. The dashed line indicates the transformation frequency

499　expected if mutants are equivalent to the parental strain expressing wild-type

500　VcPilQ. Statistical comparisons were made by one way ANOVA with Tukey's post

501　test. * = $p < 0.05$, *** = $p < 0.001$. **(B)-(D)** Representative phase contrast (top) and

502　epifluorescence images (bottom) of *V. cholerae* pilA-Cys cells containing the

503　indicated mutations in PilQ when grown in the presence of (B) 0mM, (C) 1mM, or

504　(D) 2mM DTT prior to labeling with AF488-maleimide to visualize bacterial pili.

505　**Figure 7: Comparison of VcPilQ with sub-tomogram averages of *M. xanthus***

506　**Type IVa pilus machinery, and with the *E. coli* Type IVa pilus**

507　Sub-tomogram averages of the Type IVa pilus machinery in various *Myxococcus*

508　*xanthus* strains are compared with VcPilQ[44]. VcPilQ is compared with the non-

509　piliated strains Δ*pilP* (**A**, EMD-3249) and Δ*pilB* (**B**, EMD-3260), and with the

510　piliated wild-type strain (**C,** EMD-3247). In the Δ*pilP* strain, only PilQ and TsaP

511　are present in the complex. The putative TsaP location is marked with an asterisk in

512　the rightmost volume in **(A), (B)** and **(C).** The location of the PilQ N0 domain is

19

513   marked with arrowheads. In **(B)** and **(C)** the location of the MxPilN and MxPilO

514   proteins are marked with arrowheads. Each Mx sub-tomogram average structure is

515   depicted at a high cryoEM threshold (left) to delineate the inner and outer leaflets of

516   the outer membrane, and at a lower cryoEM threshold (right) to fully reveal the

517   protein features. For each cryoEM threshold, a dashed line is used to indicate the

518   apparent membrane thickness, which ranges from 52 Å to 85 Å. The structure of

519   the VcPilQ solved here is shown with coloring like in **Figure 1**. **(D)** A central slice

520   of the cryoEM density of VcPilQ (grey) is shown with the *E. coli* Type IVa pilus

521   structure (yellow) fit in [52]. The outer membrane is depicted as a blue rectangle.

522   **Supplemental Figure 1: Expression and purification of VcPilQ from *V.***

523   ***cholerae* cells**

524   **(A)** The functionality of N-terminally deca-histidine tagged native PilQ was

525   assessed via natural transformation assays. These data indicate that the tagged

526   variant supports equal rates of transformation compared to an isogenic parent

527   strain, thus indicating that it is fully functional. Data are from four independent

528   biological replicates and shown as the mean ± SD. Statistical comparisons were

529   made by unpaired Student's t-test. NS = not significant. **(B-F)** Gel electrophoresis

530   analysis of VcPilQ in amphipols by coomassie stain **(B)**, **(D), (E)**, and western

531   blotting against anti-His tag **(C)**, **(F)**. The multimer band is indicated with an

532   asterisk.

533   **Supplemental Figure 2: *V. cholerae* PilQ sequence with mass spectroscopy**

534   **results.**

535   Mass spectrometry coverage of purified PilQ: The sequence of VcPilQ (571 amino

536   acids) is shown in black with residue numbers marked ever 10 residues. Regions of

537   the sequence corresponding to protein motifs (Signal Peptide, AMIN Domain,

538   STN, Coil, Secretin_N and Secretin) or folded regions of the protein structure (N0,

539   N3) are highlighted with flags below the relevant sequence. The sequence coverage

540   achieved by mass spectrometry is highlighted with grey flags labeled "MS".

20

541 Approximately 65% of the sequence was represented in the peptides analyzed by

542 mass spectrometry. The protein sequence was visualized using the Benchling

543 Biology Software.

544 **Supplemental Figure 3: CryoEM data**

545 **(A)** Representative micrograph with a 10 Å low pass filter applied. A representative

546 Fourier transform and contrast transfer function (CTF) fit of the micrograph is

547 shown in the upper left corner. Scale bar, 200 Å. **(B)** Representative 2D classes

548 calculated in Relion. The 2D class box is 441.6 Å wide. **(C)** A top view 2D class

549 from demonstrates that VcPilQ has C14 symmetry. Yellow lines point to inner

550 spokes in the PilQ gate. Each monomer is numbered in white.

551 **Supplemental Figure 4: CryoEM data processing flow chart**

552 The major steps in data processing are summarized in a flow chart.

553 **Supplemental Figure 5: Fourier Shell Correlation and 3DFSC of VcPilQ**

554 **(A)** Fourier shell correlation (FSC) of VcPilQ cryoEM structure. **(B) to (D)** The

555 Salk institute 3DFSC server was used to examine the resolution in the X, Y and Z

556 dimensions in the VcPilQ structure(Tan et al. 2017). The sphericity is 0.962 out of

557 1.

558 **Supplemental Figure 6: Local resolution of VcPilQ**

559 ResMap was used to calculate the local resolution per voxel of the density map.

560 Resolution is plotted from 2.4 Å (cyan) to 8.4 Å (black). VcPilQ is shown from

561 the side **(A),** as a central slice **(B),** from the top (extracellular side, **C**) and from

562 the bottom (periplasmic side, **D**).

563

564

565 **Supplemental Figure 7: Homology modeling of VcPilQ**

566 I-TASSER was used to predict the structure of *V. cholerae* PilQ. I-TASSER
567 reported five potential models (labeled VcPilQ Homology model 1 to 5 and shown
568 with gray background). The ITASSER models are based on structures deposited in
569 the PDB, so the five models can be sorted by their similarity to the T2SS with the
570 cap (Vibrio-type), without the cap (Klebsiella-type) or the T3SS, so the models are
571 shown next to the corresponding secretin type. On the right the structure of VcPilQ
572 solved in this paper is shown in dark red. The membrane is blue.

573 **Supplemental Figure 8: Raw transformation frequency data for cysteine pair**
574 **mutants.**

575 Transformation assays were performed with the indicated strains in the presence of
576 DTT as indicated above the bars. Data are from at least 4 independent biological
577 replicates and shown as the mean ± SD.

578 **Supplemental Figure 9: RMSD of VcPilQ versus published secretin structures.**

579 The monomer structure of VcPilQ (dark red) was aligned to the *E. coli* K12 GspD
580 (5WQ7) **(A)**, the *V. cholerae* GspD (5WQ8) **(B)**, and the *S. typhimurium* InvG
581 (6DV3 and 5TCQ) **(C-D)** (Worrall et al. 2016; Yan et al. 2017; J. Hu et al. 2018).
582 Only the N3 and Secretin domains were compared. RMSD was calculated with
583 MatchMaker in Chimera. The RMSD results are summarized in Table **(E).**

584 **Supplemental Figure 10: Alignment of VcPilQ N0 domain compared to**
585 **others.**

586 The VcPilQ coil and N0 domain (red, residues 160 to 228) were compared to
587 previously published N0 structures. The *N. meningitidis* PilQ N0 NMR structure
588 (purple, residues 350-417) is shown fit to the VcPilQ N0 (red) with the smallest
589 RMSD structure (0.9 Å) in **(A)** and with all structures (RMSD 0.9 to 1.2 Å) in
590 **(B). (C)** *S. typhimurium* InvG (PDB 6DV3, residues 34-104), RMSD 1.1 Å. **(D)**

591   The N0 domain of *Klebsiella oxytoca* PulD (PDB 6HCG, residues 27-100) RMSD

592   1.1 Å (Berry et al. 2012; Chernyatina and Low 2019; Worrall et al. 2016). Scale,

593   10 Å. **(E)** Table of RMSD results comparing VcPilQ N0 domain to other N0

594   domains.

595   **Supplemental Figure 11: Comparison of VcPilQ with putative *M. xanthus***

596   **PilQ structure**

597   **(A)** The domain organization of *V. cholerae* PilQ (top) and *M. xanthus* PilQ

598   (bottom) calculated in CDVist (Adebali, Ortega, and Zhulin 2015). **(B)** The

599   structure of VcPilQ (this study) in dark red is compared to the Phyre2 homology

600   modeling prediction of the MxPilQ structure (right, pink)(Kelley et al. 2015).

601   **Supplemental Movie 1: Atomic model of VcPilQ fit into the cryoEM density**

602   A single subunit of VcPilQ (labeled and colored by domain as in **Figure 1**) is

603   shown first by itself, then in the multimer. Next the model is displayed in the

604   cryoEM density to demonstrate agreement.

605

606 **Table 1.** Summary of single-particle data collection, 3D reconstruction, and model
607 refinement

Imaging parameters and 3D reconstruction	
Acceleration voltage (kV)	300
Magnification (X)	81,000
Pixel size (Å)	1.104
Frame rate (s^{-1})	0.092
Exposure time (s)	3.7
Total exposure (e$^-$ / Å)	60
Particles	
Micrographs used for selection	2,510
Defocus range (μm)	-0.5 to -3.5
Windowed	252,319
In final 3D reconstruction	100,543
Resolution	
'Gold-standard' at FSC 0.5 (Å)	3.0 Å
'Gold-standard' at FSC 0.143 (Å)	2.7 Å
Map-sharpening B factor (Å^2)	-69
Model refinement	
Resolution in phenix.real_space_refine (Å)	3.0
Model-to-map fit (CC_mask)	0.745
Number of atoms/residues/molecules	
NCS restrained chains	14
Protein atoms, residues (per chain)	43736, 412
Ramachandran angles (%)	
Favored	92.21
Allowed	7.79
Outliers	0
r.m.s. deviations	
Bond lengths (Å)Bond angles (°)	0.006
Bond angles (°)	0.847
Molprobity	
Score	2.81
Clashscore	10.53
Rotamer outliers (%)	10.91
EMRinger score	3.13

608

609

24

Methods

Bacterial Strains and culture conditions

All *V. cholerae* strains were derived from the El Tor strain E7946[54]. Strains were routinely grown in LB Miller broth and agar.

VcPilQ was purified from *V. cholerae* with a deca-histidine tag added to the N-terminus of PilQ at the native locus. The full genotype of this strain (TND1751) is 10xHis-PilQ, ΔVC1807::SpecR, lacZ::lacIq, comEA-mCherry, ΔluxO, Ptac-tfoX, ΔTCP::ZeoR, ΔMSHA::CarbR, ΔCTX::KanR.

For testing the impact of cysteine pair mutants on natural transformation, *V. cholerae* strains were generated where the native copy of PilQ was deleted and the corresponding PilQ allele was expressed at an ectopic site. The full genotype of the parent strain (TND2140) was ΔlacZ::Pbad-10XHis-PilQ CmR, ΔpilQ::TetR, ΔCTX::KanR, ΔMSHA::CarbR, ΔluxO, ΔTCP::ZeoR, comEA-mCherry, Ptac-tfoX. The cysteine pair mutants were isogenic other than the cysteine mutations introduced into the Pbad-10XHis-PilQ construct, which were PilQ S448C S453C (TND2169) and PilQ L445C T493C (TND2170).

To test the impact of cysteine pair mutants on pilus biogenesis, *V. cholerae* strains were generated akin to those described above, with the exception that the retraction ATPase PilT was deleted and the strains contained a cysteine substitution mutation in the major pilin that allows for competence pilus labeling as previously described[15]. The full genotype of the parent strain (TND2244) was ΔlacZ::Pbad-10XHis-PilQ CmR, ΔpilT::TmR, ΔpilQ::TetR, ΔCTX::KanR, ΔMSHA::CarbR, ΔluxO, ΔTCP::ZeoR, pilA S67C, comEA-mCherry, Ptac-tfoX. The cysteine pair mutants were isogenic other than the cysteine mutations introduced into the Pbad-10XHis-PilQ construct, which were PilQ S448C S453C (TND2242) and PilQ L445C T493C (TND2243).

25

636   All strains were generated by natural transformation and cotransformation exactly

637   as previously described [55].

Natural transformation assays

639   Natural transformation assays were performed exactly as described in [15]. For

640   reactions where strains harbored Pbad-10XHis-PilQ constructs, arabinose was

641   added to a final concentrations of 0.2%. Where indicated, DTT was added at the

642   indicated concentrations throughout the assay.

Competence pilus labeling and microscopy

644   Cells were labeled with AlexaFluor 488-maleimide dye and imaged to visualize

645   competence pili exactly as previously described [15,56]. All strains were grown with

646   arabinose added to a final concentrations of 0.2% to induce expression of the Pbad-

647   10XHis-PilQ construct. Where indicated, cells were grown in the presence of the

648   indicated concentration of DTT prior to labeling.

Expression

650   *Vibrio cholerae* expressing His-tagged PilQ (10xHis-PilQ, ΔVC1807::SpecR,

651   lacZ::lacIq, comEA-mCherry, ΔluxO, Ptac-tfoX, ΔTCP::ZeoR, ΔMSHA::CarbR,

652   ΔCTX::KanR) was streaked on Luria Broth agar plates and grown overnight at

653   30˚C. Small cultures were seeded (5 mL) and grown over night at 30˚C. The next

654   day, 500 mL cultures were seeded with the 5 mL culture. LB broth was

655   supplemented with 20 mM MgCl$_2$, 10 mM CaCl$_2$ and 100 μM IPTG. The large

656   cultures were grown overnight at 30˚C in beveled flasks. The following day,

657   cultures were spun down (6,000 rpm, 4˚C, 20 minutes) and the cell paste was

658   weighed, aliquoted, and stored at -80˚C.

659

660

Purification

Cell pellet (15 g) was resuspended in lysis buffer (50 mM Tris HCl, pH 8, 300 mM NaCl, 1% DDM, 20 mM imidazole) supplemented with lysozyme (40 mg/mL in 50% glycerol/water), DNAse I (4 mg/mL in 50% glycerol/water), and EDTA-free Protease Inhibitor tablet (Roche, 11697498001). Lysis proceeded with stirring at 4˚C for 20 hours. Lysate was clarified by ultracentrifugation (Beckman L8-M ultracentrifuge, Rotor Type 45 Ti, 35,000rpm, 1 hour). The supernatant was mixed with Ni NTA agarose beads (Anatrace, SUPER-NINTA25) and incubated with stirring (4˚C, 8 hours). In a gravity column at 4˚C, proteins conjugated to Ni NTA agarose beads were washed (50 mM Tris HCl, pH 8, 300 mM NaCl, 0.05% DDM, 70 mM imidazole), (50 mM Tris HCl, pH 8, 300 mM NaCl, 0.05% DDM, 300 mM imidazole), and eluted (50 mM Tris HCl, pH 8, 300 mM NaCl, 0.05% DDM, 1 M imidazole). Eluant was concentrated to ~1 mg/L (EMD Millipore Amicon Ultra-15, 30 kDa cutoff, UFC903024). Concentrated PilQ (150 µL of ~1 mg/mL protein) was exchanged into Amphipol A8-35 (0.585 mg for a 3:1 ratio, Anatrace, A835) and allowed to incubate at 4 C for 1 hr. Excess DDM was removed using Bio-Beads SM2 (Bio-Rad, 1523920) by incubating overnight at 4˚C. The protein was concentrated.

Electron microscopy

For cryoEM, Quantifoil R2/2 300 Mesh grids (EMS, Q33100CR2) were glow discharged (Pelco EasiGlow, 20 mA, 60 seconds). PilQ in amphipol (3 µL of ~0.8 mg/L) was frozen on a Mark IV Vitrobot (FEI, 20 ˚C, 100% relative humidity, blot force -6, blot time 4 s). Micrographs were collected on a 300 kV Titan Krios microscope (FEI) with energy filter (Gatan) and equipped with a K3 direct electron detector (Gatan). Data was collected using Serial EM software with a pixel size of 1.104 Å (81,000x magnification) and a defocus range from -1.0 µm to -3.0 µm [57]. A fluence of 19.8 electrons/pixel/second was used with a 3.7 s exposure time to collect 60 e-/ $Å^2$.

27

Image Processing

690 The cryoEM image processing workflow is summarized in **Supplemental Figure**
691 **4**. MotionCor2 was used for motion correction and dose weighting of 3,808
692 movies[58]. CTF correction was used to evaluate micrograph quality(Rohou and
693 Grigorieff 2015). CryoSPARC blob picking on 2,510 micrographs yielded
694 3,100,353 potential particles[60]. After inspection, the 252,319 particles were
695 analyzed by several rounds of 2D classification and 3D classification to yield
696 100,543 particles. These particles were moved to Relion using the UCSF PyEM
697 package (*https://github.com/asarnow/pyem*) script[61]. In Relion, several rounds of
698 3D refinement, polishing and CTF refinement were used[62–64]. ResMap was used to
699 calculate local resolution[35].

Model Building and Refinement

701 The initial model (residues 230-571) was auto-built using Buccaneer[65]. Subsequent
702 building and model adjustments were performed by hand using COOT[66]. A
703 homology model of the N0 domain (residues 160-229) was created using I-
704 TASSER and manually docked using COOT[42,67,68]. Coulombic potential density for
705 residues 1-159 were not observed. The model was refined in PHENIX version 1.16-
706 dev3549 using phenix.real_space_refine with the resolution set to 3 Å [69]. NCS
707 constraints were applied for the 14 subunits and were automatically detected and
708 refined. Automatically determined secondary structure restraints, rotamer restraints,
709 and Ramachandran restraints were applied as well. The quality of the model was
710 evaluated using EMRinger[70] and Molprobity [71] (Table 1).

Mass Spectrometry

712 After running a BioRad Stain Free gel and performing a coomassie staining, the
713 band of interest was excised with a clean razor blade. The gel piece was destained
714 with ammonium bicarbonate and reduced with DTT (50˚C, 30 minutes). Next the
715 sample was alkylated with iodoacetamide (room temperature, dark, 20 minutes).

28

716   The gel pieces were then dehydrated. Trypsin was used to digest the protein in the

717   gel (37˚C, overnight). Peptides were extracted from the gel matrix, dried, and de

718   salted with a zip tip.

719   The in-gel-digested samples were subjected to LC-MS/MS analysis on a nanoflow

720   LC system, EASY-nLC 1200, (Thermo Fisher Scientific) coupled to a QExactive

721   HF Orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany)

722   equipped with a Nanospray Flex ion source.

723   Samples were directly loaded onto a C18 Aurora series column (Ion Opticks,

724   Parkville, Australia). The 25cm x 50µm ID column (1.6 µm) was heated to 45° C.

725   The peptides were separated with a 60 min gradient at a flow rate of 350 nL/min.

726   The gradient was as follows: 2–6% Solvent B (3.5 min), 6-25% B (42.5 min), and

727   25-40% B (14.5min), to 100% B (1min) and 100% B (12min). Solvent A consisted

728   of 97.8% H2O, 2% ACN, and 0.2% formic acid and solvent B consisted of 19.8%

729   H2O, 80% ACN, and 0.2% formic acid.

730   The QExactive HF Orbitrap was operated in data dependent mode. Spray voltage

731   was set to 1.8 kV, S-lens RF level at 50, and heated capillary at 275 °C. Full scan

732   resolution was set to 60,000 at m/z 200. Full scan target was 3×10^6 with a

733   maximum injection time of 15 ms (profile mode). Mass range was set to 300−1650

734   m/z. For data dependent MS2 scans the loop count was 12, target value was set at 1

735    $\times 10^5$, and intensity threshold was kept at 1×10^5. Isolation width was set at 1.2 m/z

736   and a fixed first mass of 100 was used. Normalized collision energy was set at 28.

737   Peptide match was set to off, and isotope exclusion was on. Ms2 data was collected

738   in centroid mode.

739   Raw data were analyzed using MaxQuant (v. 1.6.5.0)[72,73]. Spectra were searched

740   against UniProt *V. cholerae* entries (3784 sequences) and a contaminant protein

741   database (246 sequences). Trypsin was specified as the digestion enzyme and up to

742   two missed cleavages were allowed. Precursor mass tolerance was 4.5 ppm after

743   recalibration and fragment mass tolerance was 20 ppm. Variable modifications

29

744   included oxidation of methionine and protein N-terminal acetylation.

745   Carbamidomethylation of cysteine was specified as a fixed modification. A decoy

746   database was used to set score thresholds to ensure a 1% false discovery rate at the

747   protein and peptide level. Protein abundances were estimated using iBAQ and the

748   fractional abundance was calculated as the protein abundance divided by the sum of

749   all non-contaminant protein abundances[74].

Author Contributions

751   S.J.W. conceptualized the project, expressed and purified the protein, prepared

752   samples for cryoEM, collected cryoEM data, processed cryoEM data, assisted

753   in atomic model building, interpreted results, designed figures, and wrote the

754   paper.

755   M.S. purified protein, assisted with cryoEM sample prep and data collection,

756   built the atomic model, interpreted results, and provided feedback on the

757   paper.

758   T.D. engineered the *V. cholerae* constructs, performed microbial assays, and

759   interpreted results.

760   A.D. conceptualized the project, obtained funding, engineered the *V. cholerae*

761   constructs, performed microbial assays, interpreted results, designed figures,

762   and provided feedback on the paper.

763   G.J.J. conceptualized the project, obtained funding, and provided feedback on

764   the paper.

Competing interests statement

766   The authors declare no competing interests.

Acknowledgements

References

1. von Wintersdorff, C. J. H. *et al.* Dissemination of Antimicrobial Resistance in Microbial Ecosystems through Horizontal Gene Transfer. *Front. Microbiol.* **7**, (2016).

2. Blokesch, M. In and out—contribution of natural transformation to the shuffling of large genomic regions. *Current Opinion in Microbiology* **38**, 22–29 (2017).

3. Goodgal, S. H. DNA uptake in Haemophilus transformation. *Annu. Rev. Genet.* **16**, 169–192 (1982).

4. Mell, J. C., Shumilina, S., Hall, I. M. & Redfield, R. J. Transformation of natural genetic variation into Haemophilus influenzae genomes. *PLoS Pathog.* **7**, e1002151 (2011).

5. Mell, J. C. & Redfield, R. J. Natural Competence and the Evolution of DNA Uptake Specificity. *Journal of Bacteriology* **196**, 1471–1483 (2014).

6. Maiden, M. C. J. Horizontal Genetic Exchange, Evolution, and Spread of Antibiotic Resistance in Bacteria. *Clinical Infectious Diseases* **27**, S12–S20 (1998).

7. Domingues, S. *et al.* Natural Transformation Facilitates Transfer of Transposons, Integrons and Gene Cassettes between Bacterial Species. *PLoS Pathog* **8**, (2012).

8. Blokesch, M. & Schoolnik, G. K. Serogroup conversion of Vibrio cholerae in aquatic reservoirs. *PLoS pathogens* **3**, (2007).

9. Miller, M. C., Keymer, D. P., Avelar, A., Boehm, A. B. & Schoolnik, G. K. Detection and transformation of genome segments that differ within a coastal population of Vibrio cholerae strains. *Appl. Environ. Microbiol.* **73**, 3695–3704 (2007).

10. Croucher, N. J. *et al.* Rapid Pneumococcal Evolution in Response to Clinical Interventions. *Science* **331**, 430–434 (2011).

11. Hu, D. *et al.* Origins of the current seventh cholera pandemic. *PNAS* **113**, E7730–E7739 (2016).

12. Meibom, K. L., Blokesch, M., Dolganov, N. A., Wu, C.-Y. & Schoolnik, G. K. Chitin Induces Natural Competence in *Vibrio cholerae*. *Science* **310**, 1824–1827 (2005).

13. Meibom, K. L. *et al.* The Vibrio cholerae chitin utilization program. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 2524–2529 (2004).

14. Piepenbrink, K. H. DNA Uptake by Type IV Filaments. *Frontiers in Molecular Biosciences* **6**, 1 (2019).

15. Ellison, C. K. *et al.* Retraction of DNA-bound type IV competence pili initiates DNA uptake during natural transformation in Vibrio cholerae. *Nature Microbiology* **3**, 773–780 (2018).

16. Sandkvist, M. Biology of type II secretion. *Molecular Microbiology* **40**, 271–283 (2001).

17. Korotkov, K. V., Gonen, T. & Hol, W. G. Secretins: dynamic channels for protein transport across membranes. *Trends in biochemical sciences* **36**, (2011).

18. Majewski, D. D., Worrall, L. J. & Strynadka, N. C. Secretins revealed: structural insights into the giant gated outer membrane portals of bacteria. *Current Opinion in Structural Biology* **51**, 61–72 (2018).

19. Genin, S. & Boucher, C. A. A superfamily of proteins involved in different secretion pathways in gram-negative bacteria: modular structure and specificity of the N-terminal domain. *Molec. Gen. Genet.* **243**, 112–118 (1994).

20. Berry, J.-L. *et al.* Structure and Assembly of a Trans-Periplasmic Channel for Type IV Pili in Neisseria meningitidis. *Plos Pathog* **8**, e1002923 (2012).

21. Burkhardt, J., Vonck, J., Langer, J. D., Salzer, R. & Averhoff, B. Unusual N-terminal ααβαββα fold of PilQ from Thermus thermophilus mediates ring formation and is essential for piliation. **287**, 8484–94 (2012).

22. de Souza, R. F., Anantharaman, V., de Souza, S. J., Aravind, L. & Gueiros-Filho, F. J. AMIN domains have a predicted role in localization of diverse periplasmic protein complexes. *Bioinformatics* **24**, 2423–2426 (2008).

23. Green, E. R. & Mecsas, J. Bacterial Secretion Systems – An overview. *Microbiology spectrum* **4**, 215–239 (2016).

24. Yan, Z., Yin, M., Xu, D., Zhu, Y. & Li, X. Structural insights into the secretin translocation channel in the type II secretion system. **24**, 177–183 (2017).

25. Hay, I. D., Belousoff, M. J., Dunstan, R. A., Bamert, R. S. & Lithgow, T. Structure and Membrane Topography of the Vibrio-Type Secretin Complex from the Type 2 Secretion System of Enteropathogenic Escherichia coli. *Journal of bacteriology* **200**, (2018).

26. Yin, M., Yan, Z. & Li, X. Structural insight into the assembly of the type II secretion system pilotin–secretin complex from enterotoxigenic Escherichia coli. *Nature Microbiology* **3**, 581–587 (2018).

27. Hay, I. D., Belousoff, M. J. & Lithgow, T. Structural Basis of Type 2 Secretion System Engagement between the Inner and Outer Bacterial Membranes. *mBio* **8**, e01344–17 (2017).

28. Chernyatina, A. A. & Low, H. H. Architecture of a bacterial type II secretion system. *bioRxiv* 397794 (2018) doi:10.1101/397794.

29. Howard, P. S. *et al.* Structure and assembly of pilotin-dependent and -independent secretins of the type II secretion system. *PLOS Pathogens* **15**, e1007731 (2019).

30. Worrall, L. *et al.* Near-atomic-resolution cryo-EM analysis of the Salmonella T3S injectisome basal body. *Nature* **540**, 597 (2016).

31. Chami, M. *et al.* Structural Insights into the Secretin PulD and Its Trypsin-resistant Core. *Journal of Biological Chemistry* **280**, 37732–37741 (2005).

32. Guilvout, I., Nickerson, N. N., Chami, M. & Pugsley, A. P. Multimerization-defective variants of dodecameric secretin PulD. *Research in Microbiology* **162**, 180–190 (2011).

33. Cheng, A. *et al.* High resolution single particle cryo-electron microscopy using beam-image shift. *J. Struct. Biol.* **204**, 270–275 (2018).

34. Tan, Y. *et al.* Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nature Methods* **14**, 793–796 (2017).

35. Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nature Methods* **11**, 63–65 (2014).

36. Balasingham, S. V. *et al.* Interactions between the Lipoprotein PilP and the Secretin PilQ in Neisseria meningitidis▽ †. *Journal of Bacteriology* **189**, 5716–5727 (2007).

37. Pettersen, E. F. *et al.* UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25**, 1605–1612 (2004).

38. Chernyatina, A. A. & Low, H. H. Core architecture of a bacterial type II secretion system. *bioRxiv* 397794 (2019) doi:10.1101/397794.

39. Craig, D. B. & Dombkowski, A. A. Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. *BMC Bioinformatics* **14**, 346 (2013).

35

40. Dombkowski, A. A. Disulfide by Design: a computational method for the rational design of disulfide bonds in proteins. *Bioinformatics* **19**, 1852–1853 (2003).

41. Dombkowski, A. A., Sultana, K. Z. & Craig, D. B. Protein disulfide engineering. *FEBS Lett.* **588**, 206–212 (2014).

42. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, 40 (2008).

43. Lomize, M. A., Pogozheva, I. D., Joo, H., Mosberg, H. I. & Lomize, A. L. OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Res.* **40**, D370-376 (2012).

44. Chang, Y.-W. *et al.* Architecture of the type IVa pilus machine. *Science* **351**, aad2001 (2016).

45. Adebali, O., Ortega, D. R. & Zhulin, I. B. CDvist: a webserver for identification and visualization of conserved domains in protein sequences. *Bioinformatics* **31**, 1475–1477 (2015).

46. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* **10**, 845–858 (2015).

47. Siewering, K. *et al.* Peptidoglycan-binding protein TsaP functions in surface assembly of type IV pili. *Proceedings of the National Academy of Sciences* **111**, E953–E961 (2014).

48. Phillips, R. Membranes by the Numbers. *arXiv:1703.02066 [cond-mat, physics:physics, q-bio]* (2017).

49. Craig, L., Pique, M. E. & Tainer, J. A. Type IV pilus structure and bacterial pathogenicity. *Nature Reviews Microbiology* **2**, 363 (2004).

50. Kolappan, S. *et al.* Structure of the Neisseria meningitidis Type IV pilus. *Nature Communications* **7**, 13015 (2016).

51. Wang, F. *et al.* Cryoelectron Microscopy Reconstructions of the Pseudomonas aeruginosa and Neisseria gonorrhoeae Type IV Pili at Sub-nanometer Resolution. *Structure* **25**, 1423–1435.e4 (2017).

52. Bardiaux, B. *et al.* Structure and Assembly of the Enterohemorrhagic Escherichia coli Type 4 Pilus. *Structure (London, England : 1993)* (2019) doi:10.1016/j.str.2019.03.021.

53. Neuhaus, A. *et al.* A new twist on bacterial motility – two distinct type IV pili revealed by cryoEM. *bioRxiv* 720938 (2019) doi:10.1101/720938.

54. Miller, V. L., DiRita, V. J. & Mekalanos, J. J. Identification of toxS, a regulatory gene whose product enhances toxR-mediated activation of the cholera toxin promoter. *J. Bacteriol.* **171**, 1288–1293 (1989).

55. Dalia, A. B. Natural Cotransformation and Multiplex Genome Editing by Natural Transformation (MuGENT) of Vibrio cholerae. in *Vibrio Cholerae: Methods and Protocols* (ed. Sikora, A. E.) 53–64 (Springer, 2018). doi:10.1007/978-1-4939-8685-9_6.

56. Ellison, C. K., Dalia, T. N., Dalia, A. B. & Brun, Y. V. Real-time microscopy and physical perturbation of bacterial pili using maleimide-conjugated molecules. *Nat Protoc* **14**, 1803–1819 (2019).

57. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *Journal of Structural Biology* **152**, 36–51 (2005).

58. Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods* (2017) doi:10.1038/nmeth.4193.

59. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. **192**, 216–221 (2015).

60. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods* **14**, (2017).

61. Daniel Asarnow, Eugene Palovcak & Yifan Cheng. *asarnow/pyem: UCSF pyem v0.5*. (Zenodo, 2019). doi:10.5281/zenodo.3576630.

62. Kimanius, D., Forsberg, B. O., Scheres, S. & Lindahl, E. Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. 059717 (2016) doi:10.1101/059717.

63. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).

64. Zivanov, J., Nakane, T. & Scheres, S. H. W. Estimation of High-Order Aberrations and Anisotropic Magnification from Cryo-EM Datasets in RELION-3.1. *bioRxiv* 798066 (2019) doi:10.1101/798066.

65. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallographica Section D: Biological Crystallography* **62**, 1002–11 (2006).

66. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Cryst D* **66**, 486–501 (2010).

67. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* **5**, 725–738 (2010).

68. Yang, J. *et al.* The I-TASSER Suite: protein structure and function prediction. *Nature Methods* **12**, 7–8 (2015).

69. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst D* **66**, 213–221 (2010).

70. Barad, B. A. *et al.* EMRinger: side chain-directed model and map validation for 3D cryo-electron microscopy. (2015) doi:10.1038/nmeth.3541.

71. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Cryst D* **66**, 12–21 (2010).

72. Cox, J. *et al.* Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *J. Proteome Res.* **10**, 1794–1805 (2011).

73. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367–1372 (2008).

74. Schwanhäusser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* **473**, 337–342 (2011).

Figure 1

A

| T4P | AMIN | N0 | N3 | Secretin | |
| T2SS | N0 | N1 | N2 | N3 | Secretin | S |
| T3SS | | N0 | N1 | N3 | Secretin | S |

B

OM · OM · *

C

Beta-lip

AHL

Gate · Secretin

N3

Coil

N0

AMIN (flexible)

D

Figure 2

A

B

Figure 3

Vc PilQ This study	Vc GspD 5WQ8	Ec GspD 5WQ7	St InvG 6DV3

A

B

C

~30 Å

~24 Å

D

Figure 4

A

~70 Å

~108 Å

~36 Å

~80 Å

~25 Å

~60 Å

~76 Å

~85 Å

~90 Å

B

48 Å

43 Å

32 Å

26 Å

Figure 5

A

B

~90° rotate x

y
x
z
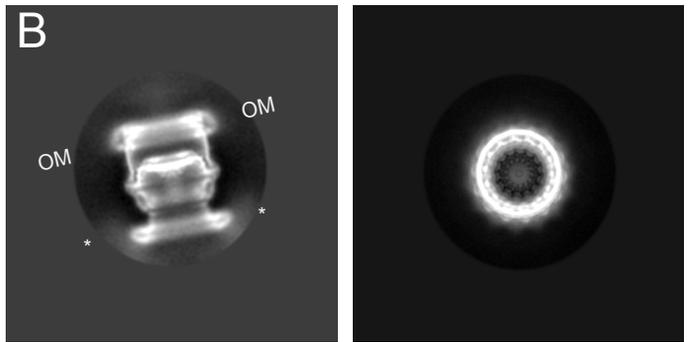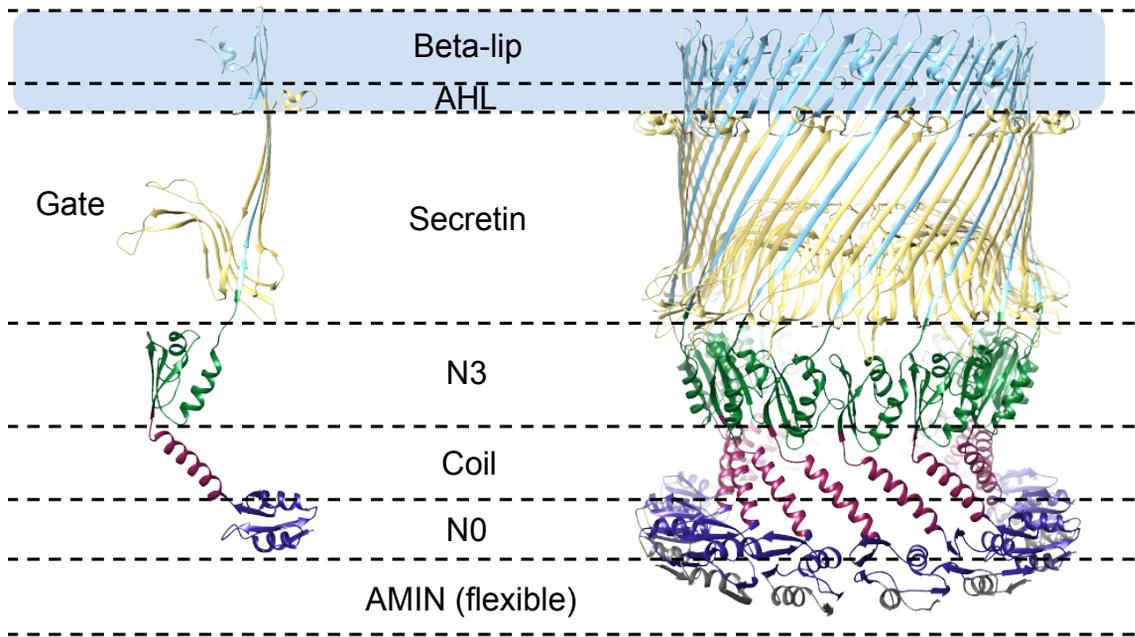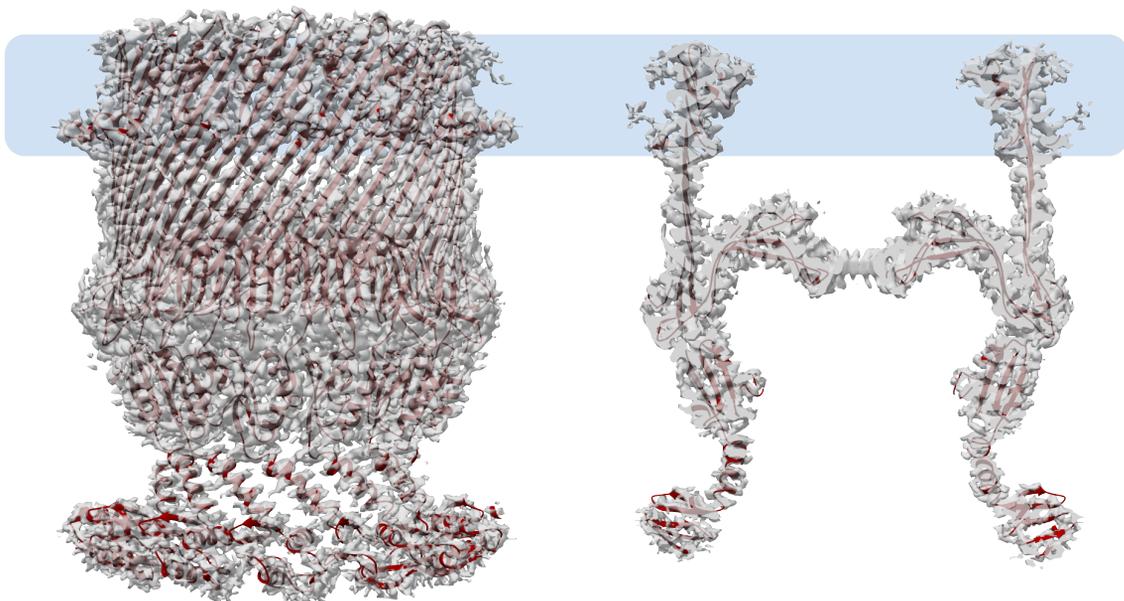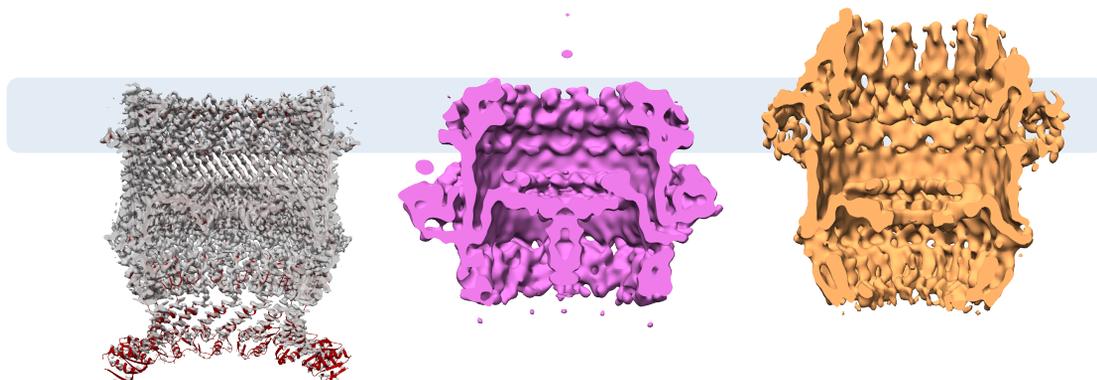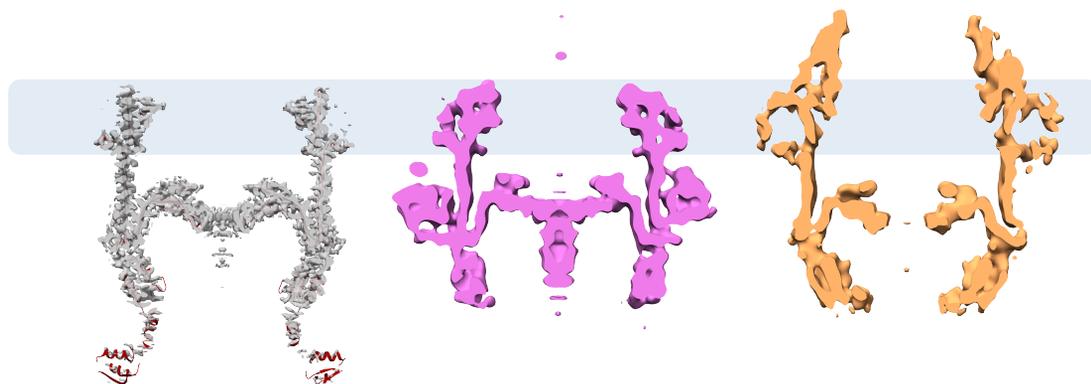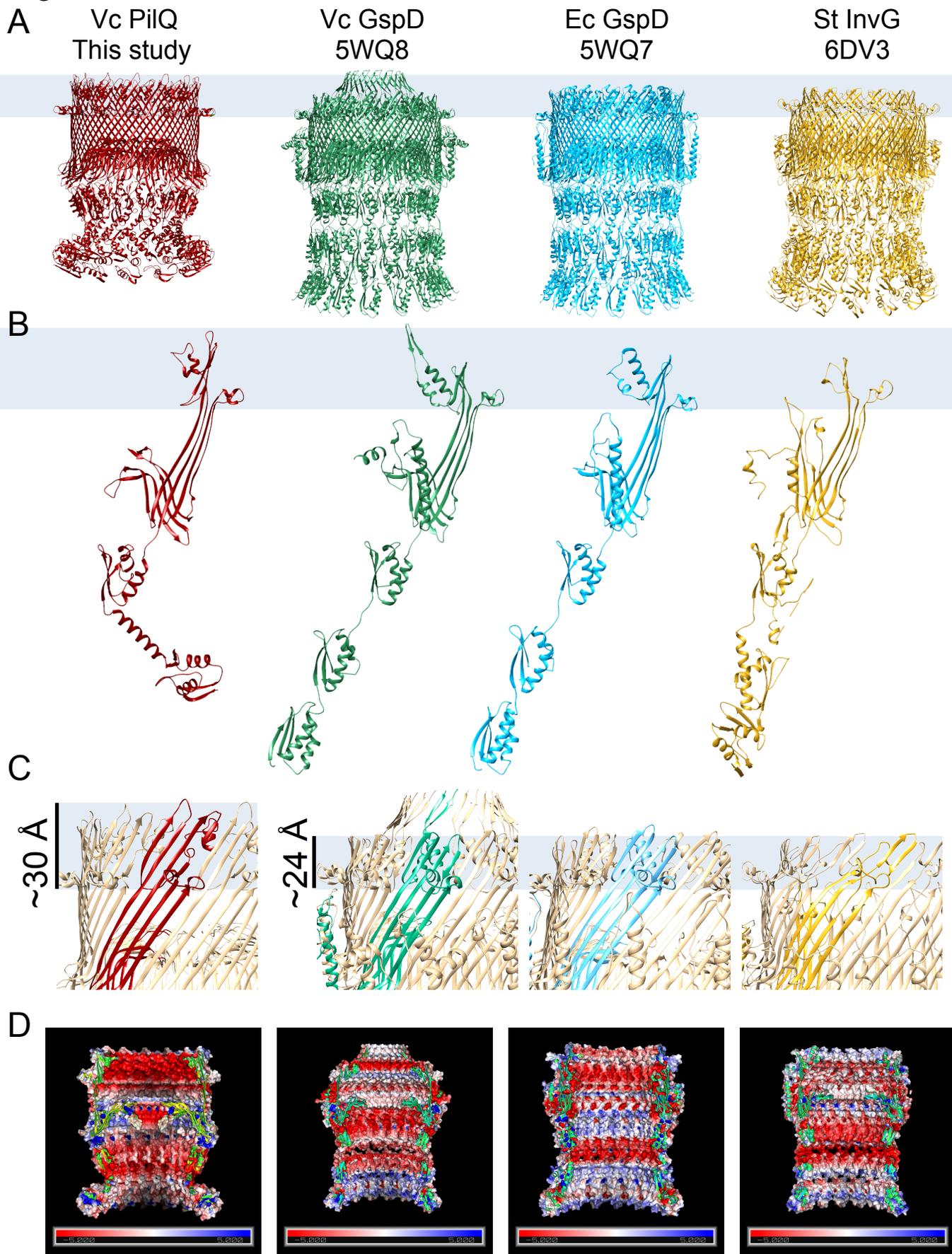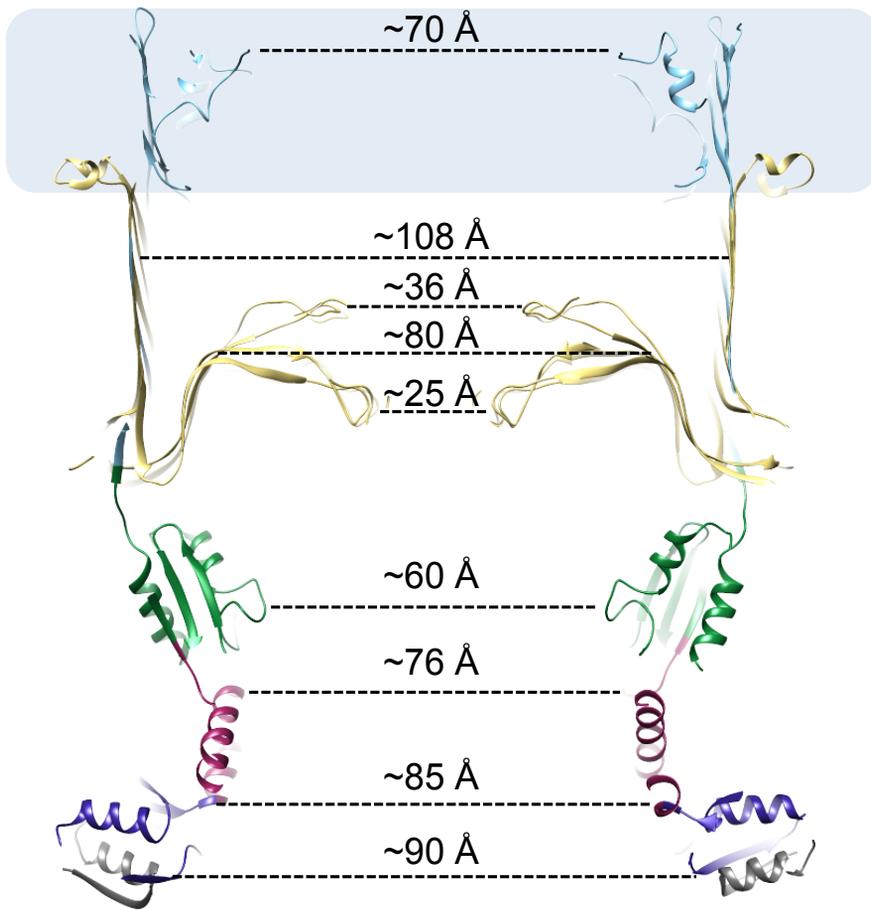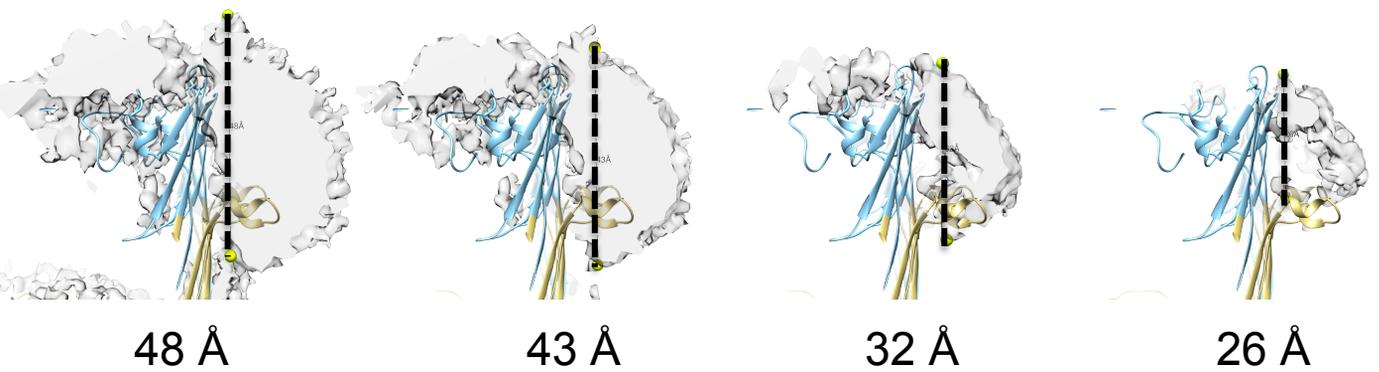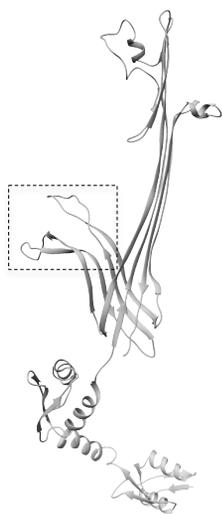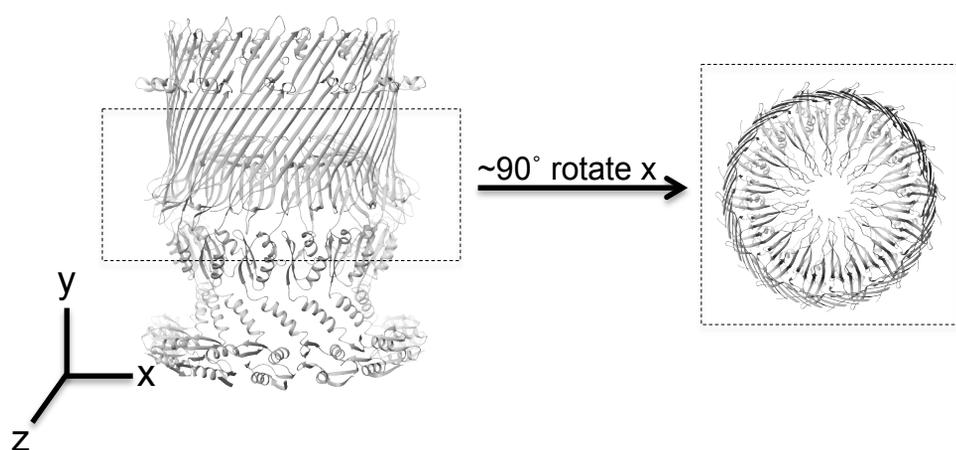
C
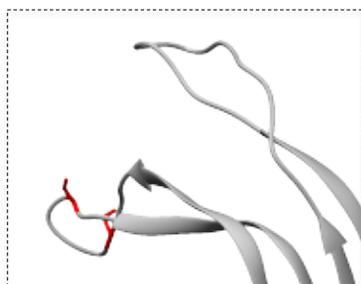
S448C/S453C

D

L445C/T493C
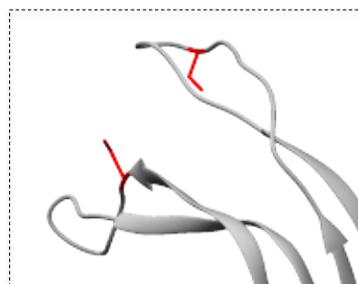
E

S448C/S453C

F

L445C/T493C

Figure 6

A

B

0mM DTT

C

1mM DTT

D

2mM DTT

Figure 7

A
Non-piliated, Δ*pilP*
EMD-3249

OM

53 Å

72 Å

< PilQ (N0) >

*

IM

B
Non-piliated, Δ*pilB*
EMD-3260

OM

52 Å

77 Å

< PilQ (N0) >

< PilN/PilO >

*

IM

C
Piliated, wild-type
EMD-3247

OM

53 Å

85 Å

< PilQ (N0) >

< PilN/PilO >

*

IM

D