

# Supplementary Information for

## **Architecture of the *Vibrio cholerae* toxin-coregulated pilus machine revealed**

### **by electron cryotomography**

Yi-Wei Chang, Andreas Kjær, Davi R. Ortega, Gabriela Kovacicova, John A. Sutherland, Lee A.

Rettberg, Ronald K. Taylor, Grant J. Jensen\*

\*Correspondence to: [jensen@caltech.edu](mailto:jensen@caltech.edu).

#### **This PDF file includes:**

Supplementary Discussion

Supplementary Figures 1 to 7

Supplementary Tables 1 to 3

Supplementary References

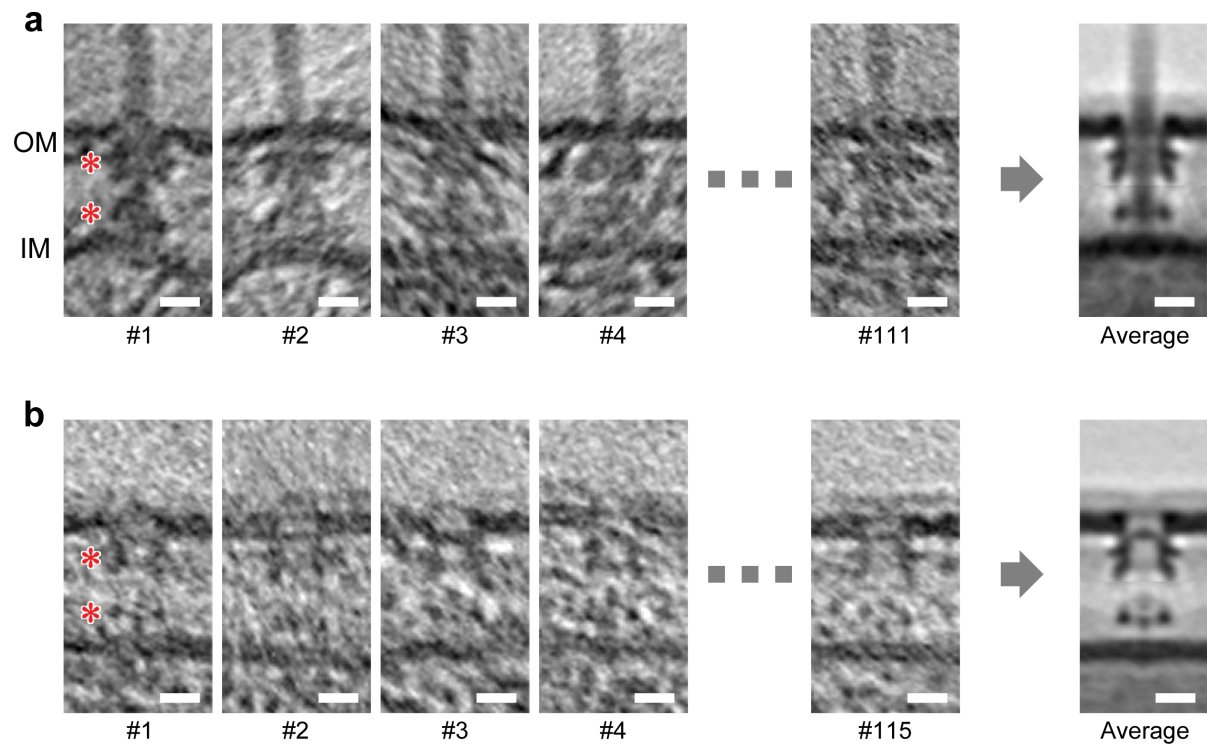
#### **Other supplementary material for this manuscript includes the following:**

Supplementary Video 1

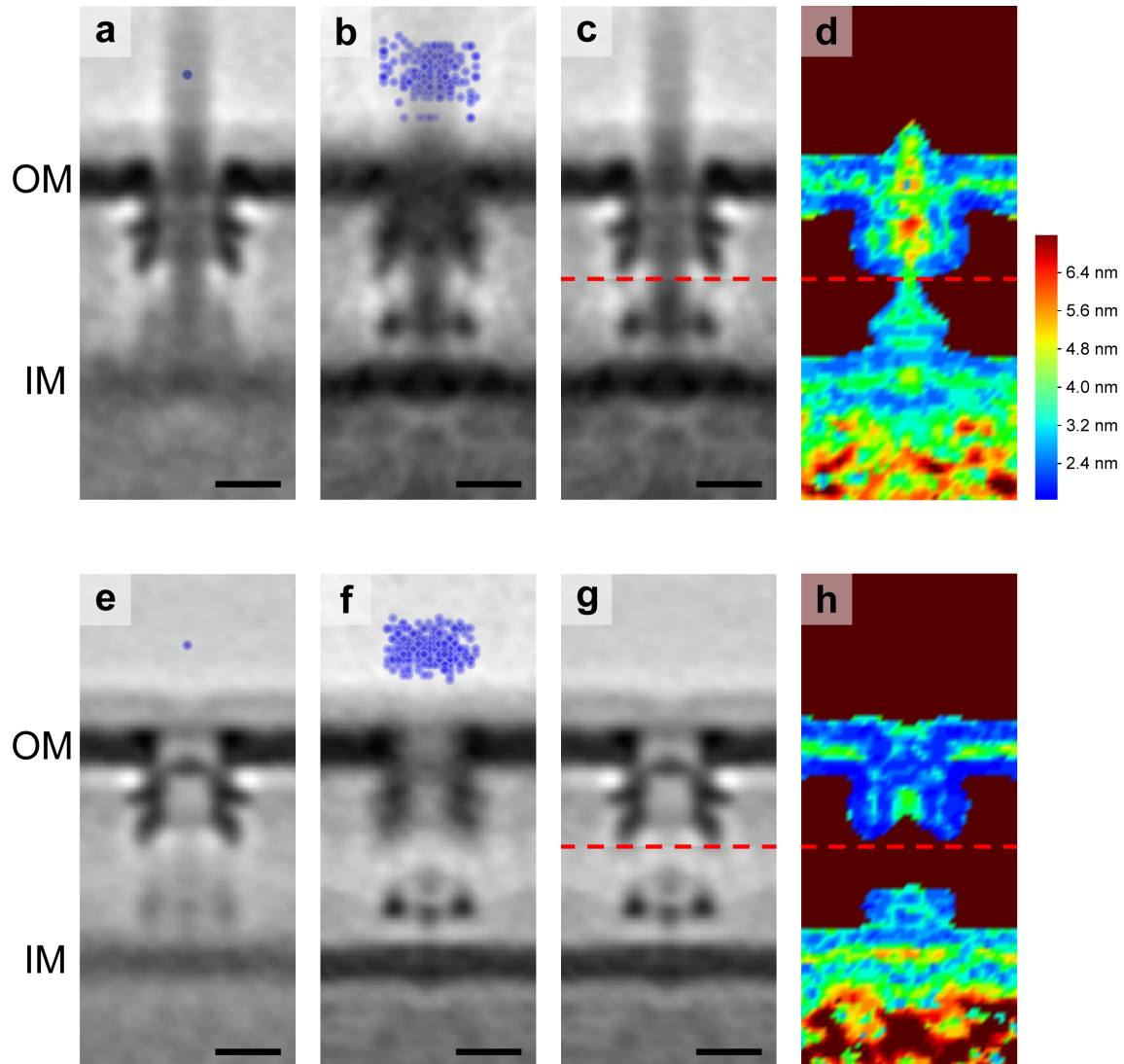
## Supplementary Discussion:

### *Evidences supporting that the structures analyzed in this study were TCPMs*

In addition to the TCP, *V. cholerae* O395 N1 also encodes two T4aP systems, the mannose-sensitive hemagglutinin (MSHA) pilus and the chitin-regulated competence pilus (ChiRP), as well as a T2SS. To test whether the structures observed were TCPMs or one of these other secretin-containing systems, we imaged a strain missing TcpQ, a component which is known to be crucial for stabilizing TcpC channels<sup>1</sup>. No similar structures were seen in 21 cryotomograms (Supplementary Table 2), indicating that the other secretin-containing systems are either not expressed under our laboratory conditions, or look substantially different from the putative TCPM structures. Additionally, the pilus fibers visualized in our cryotomograms have a diameter of 8 nm, which matches a three dimensional EM reconstruction of purified *V. cholerae* TCP<sup>2</sup>. Based on the numbers of pilin residues present, MSHA and ChiRP pili are expected to be thinner (~6 nm) similar to the T4aP of *Neisseria* spp.<sup>3,4</sup>, and the T2SS is not expected to assemble an extracellular pilus<sup>5</sup>. Furthermore, pili bundling has never been observed for MSHA and ChiRP pili, overexpression of ToxT downregulates expression of MSHA pili<sup>6-8</sup>, and chitin is required for expression of ChiRP under laboratory conditions<sup>9,10</sup>. Finally, as will be described later in the main text, in the course of this project we imaged mutants lacking the TCPM components TcpB, TcpD and TcpS and calculated sub-tomogram average structures. In each case we observed particles that resembled the putative TCPMs but lacked specific densities (Fig. 3), and the pore size and vestibule diameter of the mutants and wild-type were consistently larger than the cryo-EM single particle reconstruction of a purified T2SS secretin channel from the same species<sup>11</sup>. For these reasons, we conclude that the particles analyzed here were in fact TCPMs rather than any other known or unknown similar structure.



**Supplementary Figure 1:** Examples of slices through sub-tomograms containing (a) piliated and (b) non-piliated wild-type TCPM basal body structures. Red asterisks: OM- and IM-associated densities. Images are representative of 111 piliated and 115 non-piliated particles. Scale bars, 10 nm.

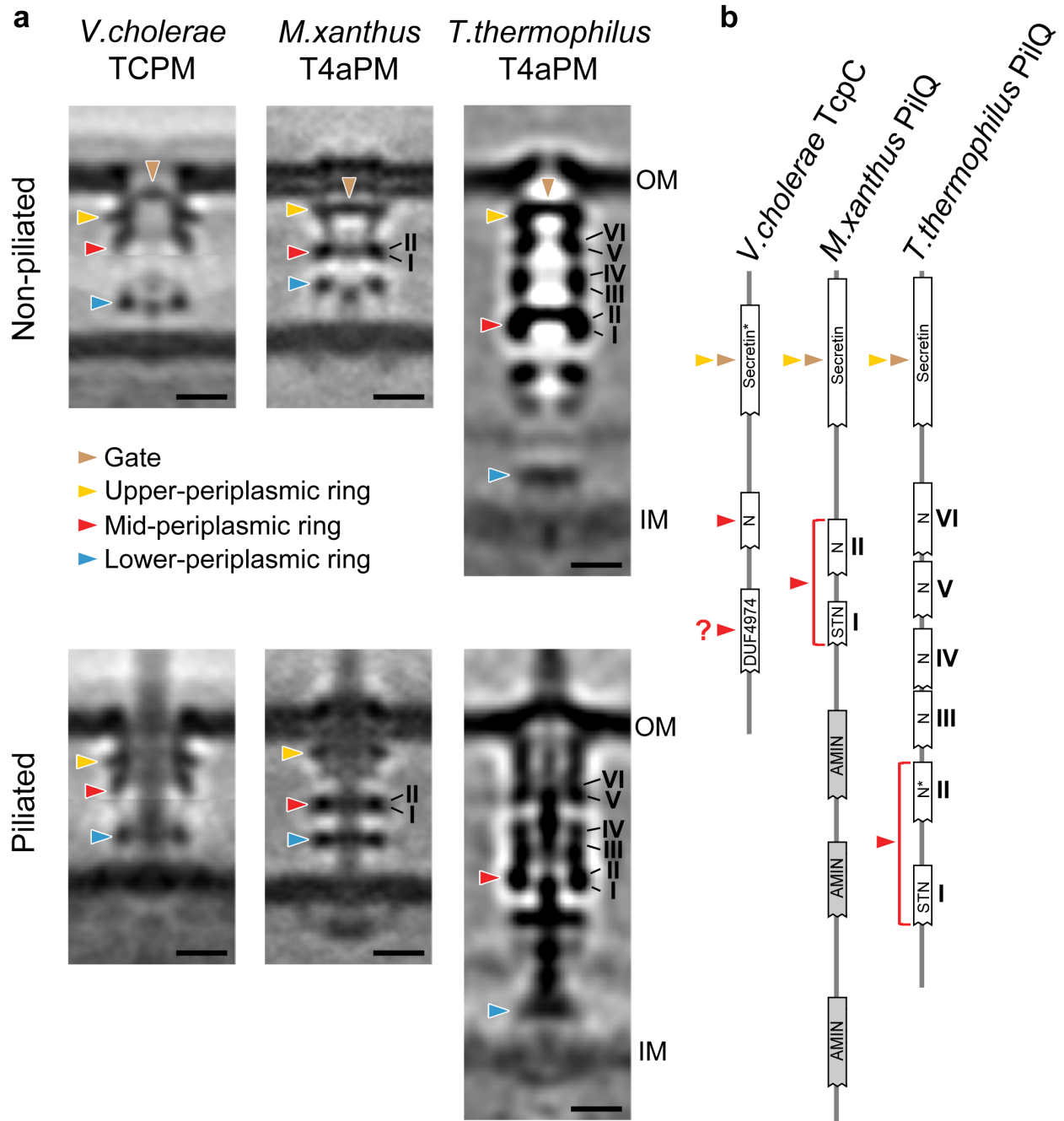


**Supplementary Figure 2:** Generation of the composite TCPM sub-tomogram averages.

(**a, b**) Sub-tomogram averages of wild-type piliated TCPM independently aligned on the OM and IM complexes, respectively. The distribution of the blue dots in (**b**) indicates the translations imposed on the OM complexes to align the IM complexes. We did not observe any pattern in the flexibility between the two complexes. (**c**) Composite sub-tomogram average using the upper and lower halves of (**a**) and (**b**), respectively, with the red dashed line indicating the interface.



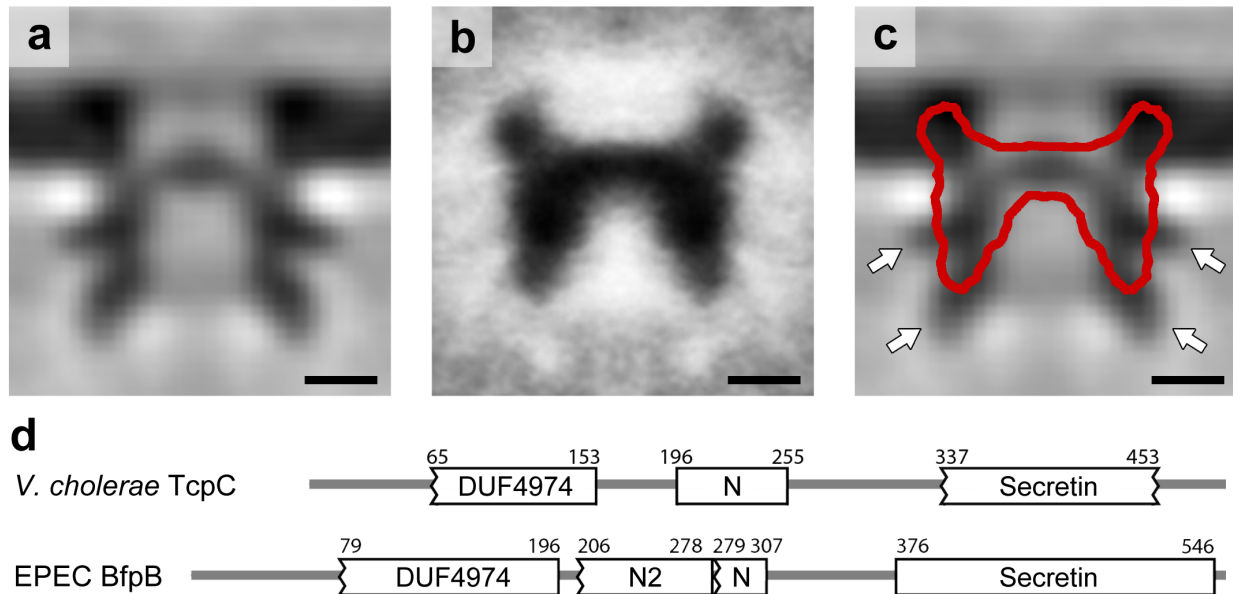
**(d)** Local resolution of **(c)** calculated by Resmap<sup>12</sup>. **(e-h)** As in **(a-d)** but of the non-piliated TCPM. Scale bars, 10 nm.



**Supplementary Figure 3:** Comparison of T4PM *in situ* structures.

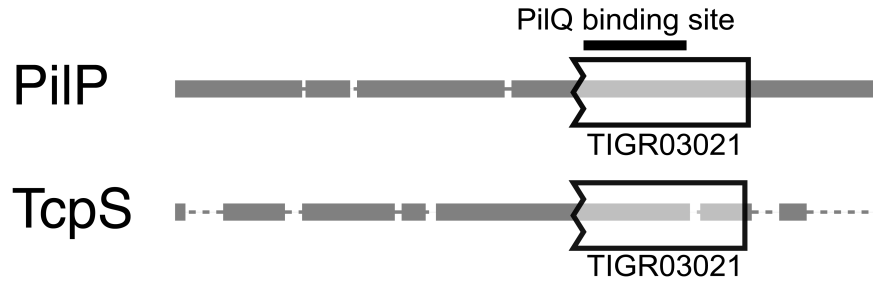
(a) From left to right: sub-tomogram averages of the *V. cholerae* wild-type TCPM non-piliated basal body (this work), the *M. xanthus*  $\Delta pilB$  T4aPM non-piliated basal body (EMD-3260)<sup>13</sup>, and the *T. thermophilus* T4aPM non-piliated basal body (EMD-3022)<sup>14</sup>. The sub-tomogram averages are aligned by the OM. The conserved gate and periplasmic ring structures are indicated by

colored arrowheads. The known locations of *M. xanthus* and *T. thermophilus* secretin N domains are labeled with I to VI. Scale bars, 10 nm. **(b)** Domain architectures of the secretin protein in *V. cholerae* TCPM (TcpC), *M. xanthus* T4aPM (PilQ), and *T. thermophilus* T4aPM (PilQ) determined based on their protein sequences using CDvist<sup>15</sup> with HHMER and HHsearch against Pfam 29.0. The protein regions corresponding to different density features in **(a)** are indicated by the arrowheads with corresponding colors. The AMIN domains in *M. xanthus* PilQ are known to have irregular binding on the peptidoglycan layer and not contribute to densities in the sub-tomogram average<sup>13</sup>. Asterisk indicates the secretin N domain described in the previous report<sup>14</sup> but could not be identified by CDvist.



**Supplementary Figure 4:** The upper- and mid-periplasmic rings of the TCPM are composed of more than the secretin channel protein TcpC.

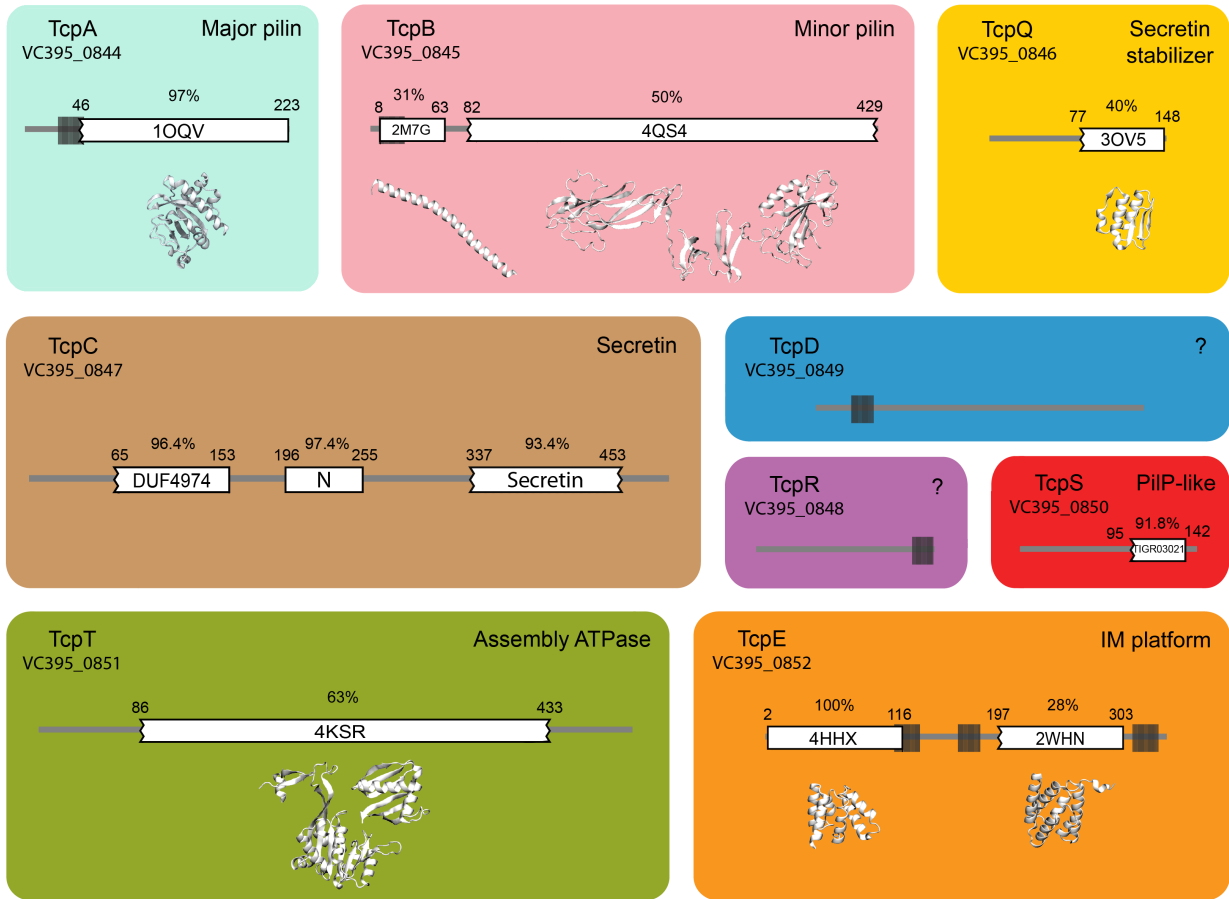
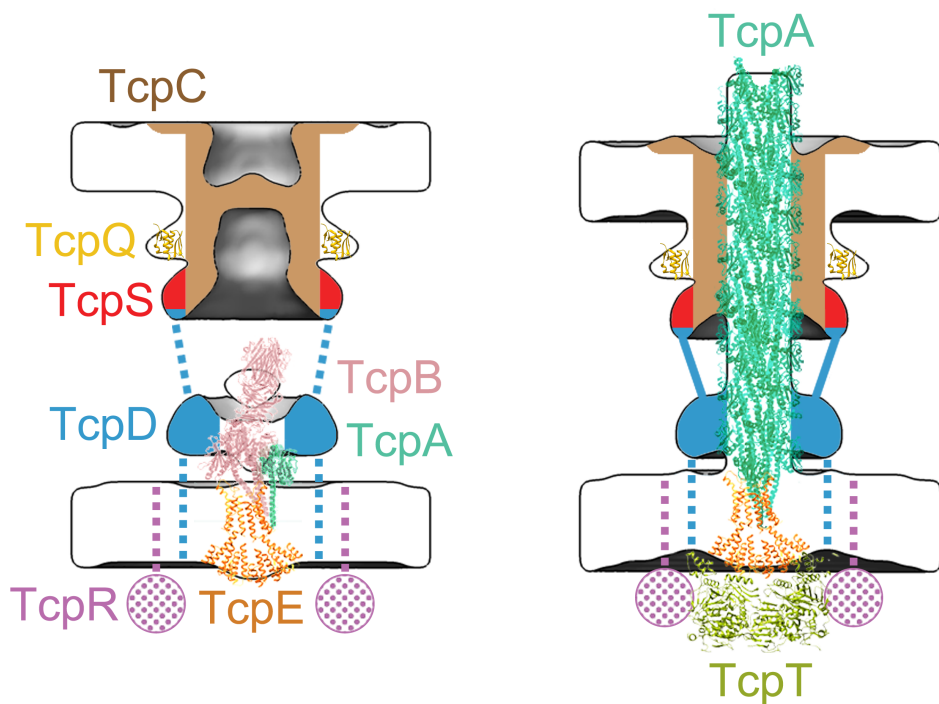
(a) Outer membrane portion of the *V. cholerae* TCPM sub-tomogram average. (b) 2D projection average of purified EPEC BfpB<sup>16</sup>. (c) Superposition of TCPM density in (a) with the outline of BfpB density in (b) on the gate shows extra densities in the TCPM (arrows). (d) Domain architecture of the secretin proteins *V. cholerae* TcpC and EPEC BfpB as predicted by HMMER and HHsearch using the Pfam 29.0 database via CDvist. The domain models of Secretin\_N and Secretin\_N\_2 are denoted as N and N2 respectively. Scale bars, 5 nm.



**Supplementary Figure 5:** The conserved domain between *M. xanthus* PilP and *V. cholerae*

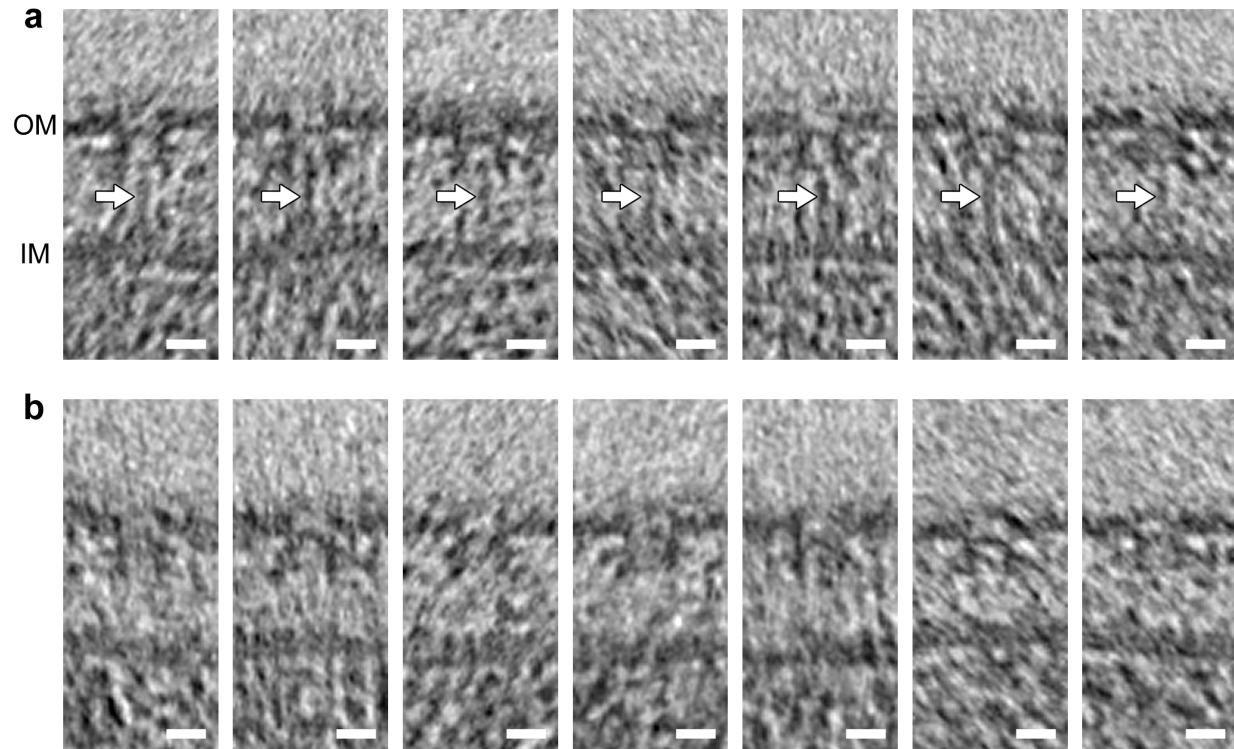
TcpS suggests that TcpS likely interacts with TcpC in a similar fashion as PilP interacts with PilQ.

PilP and TcpS were used as query sequences against sequences in the MiST database<sup>17</sup> using blastp search with an E-value cutoff of 1E-5 to retrieve their homologs. All homologs were aligned using L-INS-I from MAFFT package. The sequences of PilP and TcpS were then extracted from the alignment result and submitted to CDvist. Both sequences show significant homology to the CDD domain model TIGR03021 (pilP\_fam). The identified sequence region also coincides with the binding site of T2SS GspC (homolog to PilP) to GspD (homolog to PilQ) in a complex crystal structure<sup>18</sup>, suggesting that PilP and TcpS likely bind the secretin channel in their systems in a similar fashion. The gaps of the alignment (dotted lines) and aligned regions (solid lines) are shown for each sequence.

**a****b**

**Supplementary Figure 6:** Identifying domain architectures and available atomic homology models of TCPM components and placing the models into the TCPM *in vivo* molecular envelopes.

(a) Each panel shows a TCPM component with labels indicating regions conserved either in proteins with determined structures deposited in the Protein Data Bank database, or in domains in the Pfam or CDD database found in other T4P systems. The gene name and locus number of each TCPM component are shown in the upper-left corner of each panel. The component function is shown in the upper-right corner of each panel. Each domain box contains the PDB code of the most significant hit to the Protein Data Bank database with the exception of the TcpC and TcpS panels in which the domain boxes contain the names of the most significant hit to the Pfam 29.0 and CDD databases, respectively. The Secretin\_N Pfam domain model is denoted as N. The sequence identity is displayed above each domain when there is a PDB structure available for homology modeling. The probability of true positive is displayed above the identified domains from Pfam and CDD databases. The sequence coverage of the identified homologous structures and domains of the TCP proteins is denoted by residue numbers above the edges of each domain box. Atomic models are shown under the domain boxes. Predicted transmembrane regions are represented by grey shaded boxes. (b) Available atomic models of TcpA, TcpE, TcpT, TcpQ and TcpB are placed in the envelopes of non-piliated (left) and pilated (right) TCPMs.



**Supplementary Figure 7:** Examples of slices through TCPM structures in  $\Delta tcpR$  cells. (a) piliation-stalled and (b) non-piliated  $\Delta tcpR$  TCPM basal body structures. Images are representative of 115 particles. White arrows in (a) indicate stem densities extending from the IM to the OM vestibule. Scale bars, 10 nm.



**Supplementary Table 1:** ECT data collection parameters

---

Microscope	FEI Polara (FEG, Gatan energy filter)
Voltage (kV)	300
Camera	Gatan K2 summit
Magnification (x)	27,500
Defocus ( $\mu\text{m}$ )	-6
Pixel size (after binning for analysis) ( $\text{\AA}$ )	7.8
Tilt series angle coverage ( $^\circ$ )	-60 to +60
Tilt series increment angle ( $^\circ$ )	1
Electron dose per tomogram ( $\text{e}^-/\text{\AA}^2$ )	160

---

**Supplementary Table 2:** Number of cryotomograms collected on different strains and TCPM structures used for generating the sub-tomogram averages.

	Wild-type	$\Delta tcpB$	$\Delta tcpD$	$\Delta tcpQ$	$\Delta tcpR$	$\Delta tcpS$
Tomograms	74	73	42	21	35	41
Piliated TCPM	111	np	np	np	np	np
Non-piliated TCPM	115	240	325	np	115	222

np: none present

**Supplementary Table 3:** Relationships between the components of TCPM and other machines.

TCPM (known function)	T4aPM analog	Other analogs	Evidence of analogy
TcpA (major pilin)	PilA		Functional
TcpB (minor pilin)		CofB (ETEC T4bPM minor pilin)	HHsearch - Protein Data Bank structures
TcpC (secretin pore)	PilQ		HHsearch - Pfam 29.0 domain models
TcpQ (secretin pore stabilization)		VirB7 ( <i>X. citri</i> T4SS OM-complex component)	HHsearch - Protein Data Bank structures
TcpT (assembly ATPase)	PilB		BLAST - 2nd best hit (E-value: 5E-35)
TcpR (TcpT IM-tethering)			
TcpD (unknown)			
TcpS (unknown)	PilP		HHsearch - CDD domain models
TcpE (IM platform)	PilC		BLAST - 2nd best hit (E-value: 2E-7)
TcpF (colonization factor)			
TcpN/ ToxT (expression regulator)			
TcpJ (prepilin peptidase)		GspO ( <i>M. xanthus</i> T2SS prepilin peptidase)	BLAST - Best hit (E-value: 5E-15)

### Supplementary References:

- 1 Bose, N. & Taylor, R. K. Identification of a TcpC-TcpQ outer membrane complex involved in the biogenesis of the toxin-coregulated pilus of *Vibrio cholerae*. *J Bacteriol* **187**, 2225-2232 (2005).
- 2 Li, J., Egelman, E. H. & Craig, L. Structure of the *Vibrio cholerae* Type IVb Pilus and stability comparison with the *Neisseria gonorrhoeae* type IVa pilus. *Journal of molecular biology* **418**, 47-64 (2012).
- 3 Craig, L. *et al.* Type IV pilus structure by cryo-electron microscopy and crystallography: implications for pilus assembly and functions. *Molecular cell* **23**, 651-662 (2006).
- 4 Kolappan, S. *et al.* Structure of the *Neisseria meningitidis* Type IV pilus. *Nature communications* **7**, 13015 (2016).
- 5 Korotkov, K. V., Sandkvist, M. & Hol, W. G. The type II secretion system: biogenesis, molecular architecture and mechanism. *Nat Rev Microbiol* **10**, 336-351 (2012).
- 6 Hsiao, A., Toscano, K. & Zhu, J. Post-transcriptional cross-talk between pro- and anti-colonization pili biosynthesis systems in *Vibrio cholerae*. *Molecular microbiology* **67**, 849-860 (2008).
- 7 Hsiao, A., Xu, X., Kan, B., Kulkarni, R. V. & Zhu, J. Direct regulation by the *Vibrio cholerae* regulator ToxT to modulate colonization and anticolonization pilus expression. *Infection and immunity* **77**, 1383-1388 (2009).
- 8 Hsiao, A., Liu, Z., Joelsson, A. & Zhu, J. *Vibrio cholerae* virulence regulator-coordinated evasion of host immunity. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 14542-14547 (2006).

- 9 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).
- 10 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).
- 11 Reichow, S. L., Korotkov, K. V., Hol, W. G. & Gonen, T. Structure of the cholera toxin secretion channel in its closed state. *Nature structural & molecular biology* **17**, 1226-1232 (2010).
- 12 Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nature methods* **11**, 63-65 (2014).
- 13 Chang, Y. W. *et al.* Architecture of the type IVa pilus machine. *Science* **351**, aad2001 (2016).
- 14 Gold, V. A., Salzer, R., Averhoff, B. & Kuhlbrandt, W. Structure of a type IV pilus machinery in the open and closed state. *Elife* **4** (2015).
- 15 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).
- 16 Lieberman, J. A. *et al.* Outer membrane targeting, ultrastructure, and single molecule localization of the enteropathogenic *Escherichia coli* type IV pilus secretin BfpB. *J Bacteriol* **194**, 1646-1658 (2012).
- 17 Ulrich, L. E. & Zhulin, I. B. The MiST2 database: a comprehensive genomics resource on microbial signal transduction. *Nucleic Acids Res* **38**, D401-407 (2010).
- 18 Korotkov, K. V. *et al.* Structural and functional studies on the interaction of GspC and GspD in the type II secretion system. *PLoS Pathog* **7**, e1002228 (2011).