PNAS

Supporting Information for

Tad and toxin coregulated pilus structures reveal unexpected diversity in bacterial type IV pili

Ravi R. Sonani[1], Juan Carlos Sanchez[2], Joseph K. Baumgardt[2], Shivani Kundra[3], Elizabeth R. Wright[2], Lisa Craig[3] and Edward H. Egelman[1]

[1]Department of Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, VA 22903, U.S.A.
[2]Department of Biochemistry, University of Wisconsin-Madison, Madison, WI 53706, USA
[3]Department of Molecular Biology and Biochemistry, Simon Fraser University
Burnaby, British Columbia, BC V5A 1S6, Canada

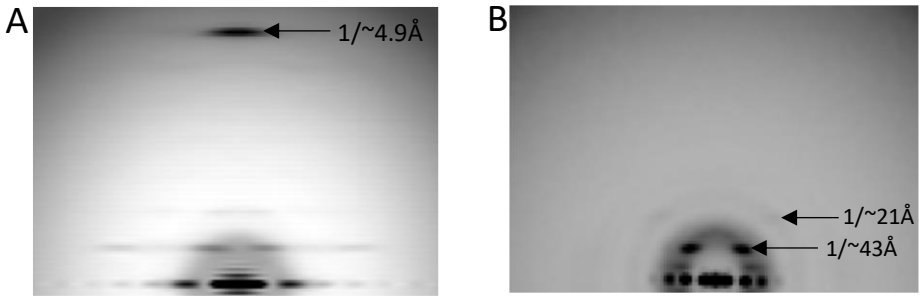
*Corresponding Author: Edward H. Egelman
Email: ehe2n@virginia.edu

This PDF file includes:

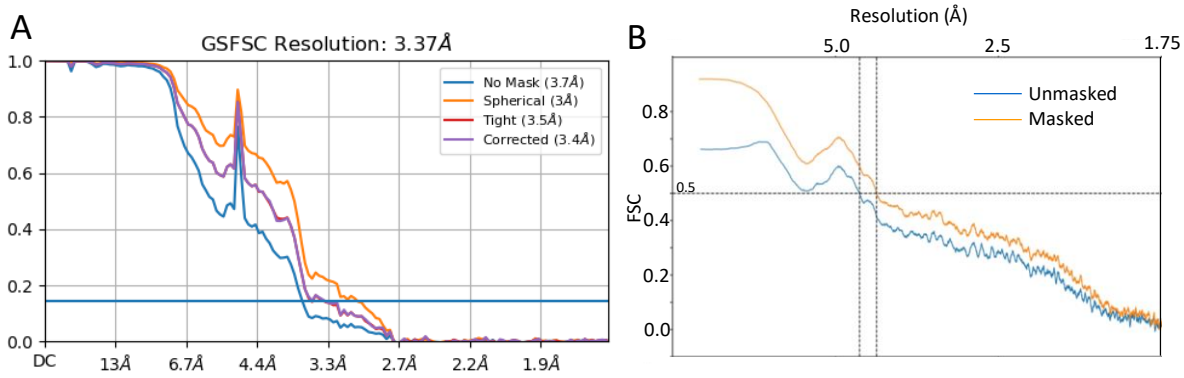
Figures S1 to S10
Tables S1 to S2
Legend for Movie S1

Other supporting materials for this manuscript include the following:

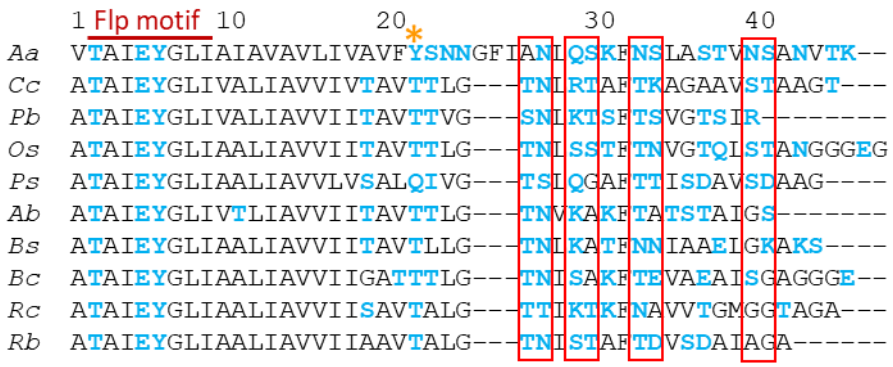
Movie S1



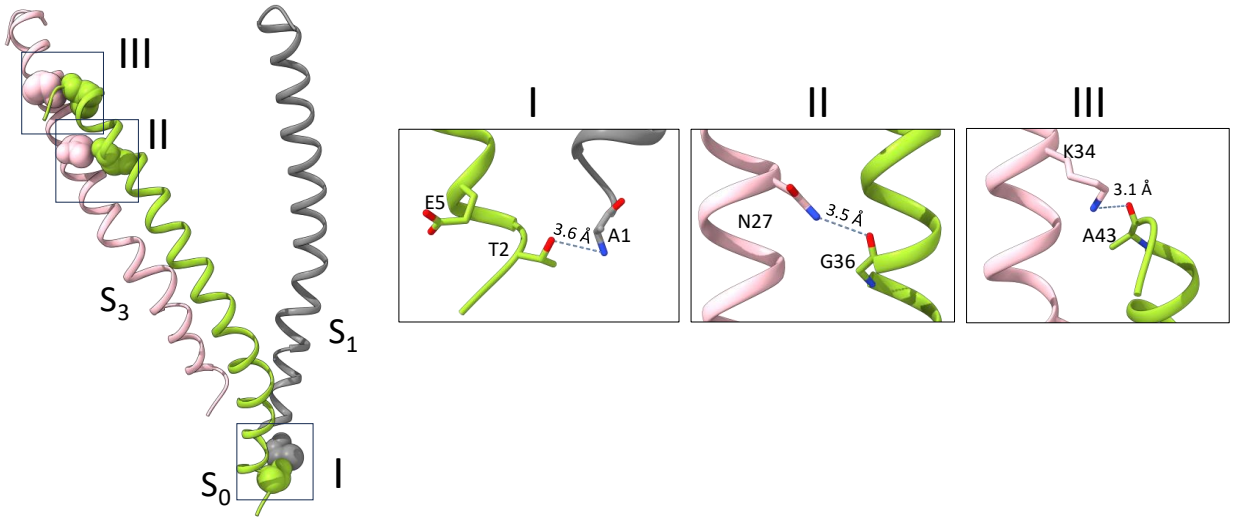
Supplementary Figure 1: Averaged power spectrum from raw segments of **(A)** *C. crescentus* tad pilus and **(B)** *V. cholerae* TCP.



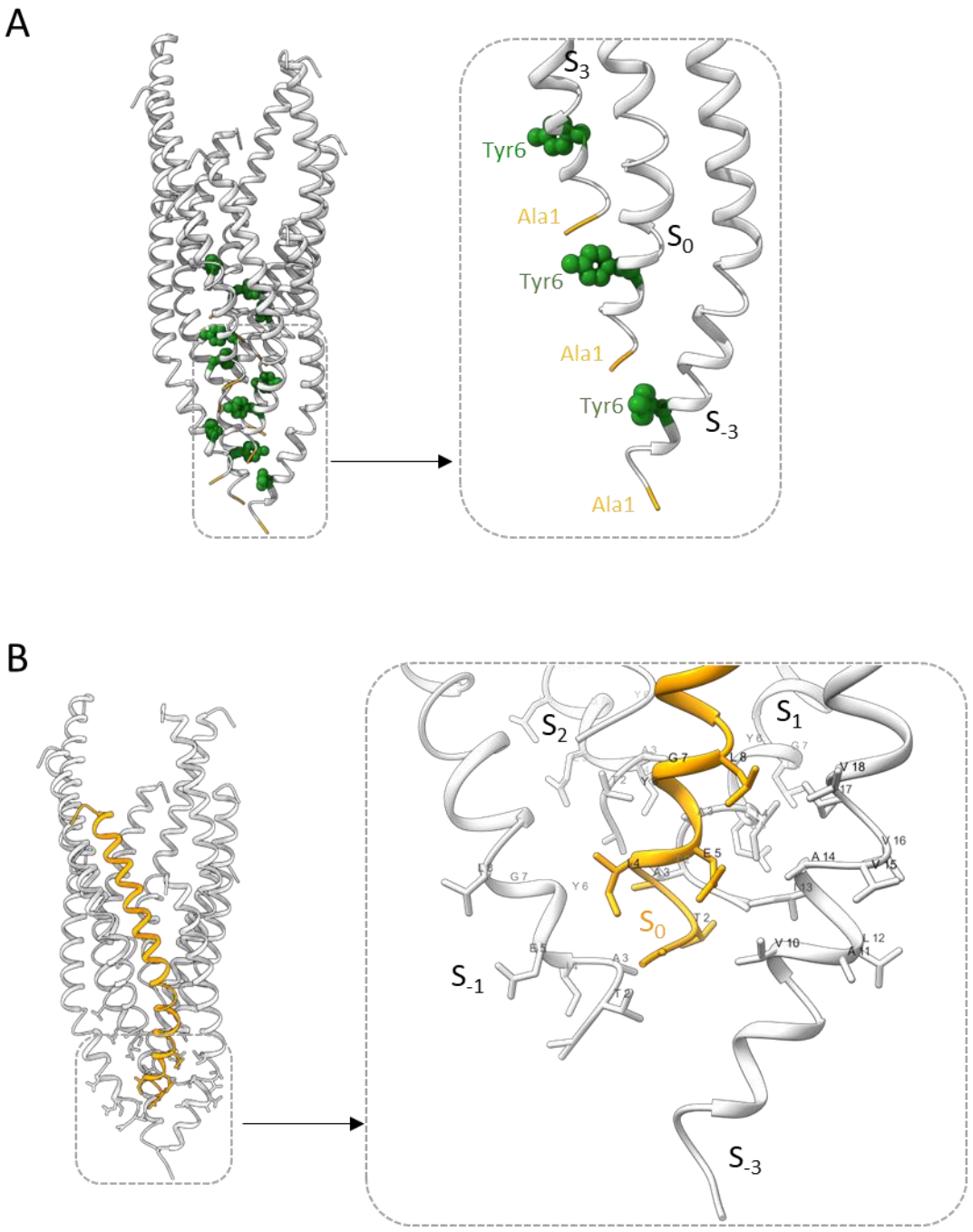
Supplementary Figure 2: (A) Map:map Gold Standard Fourier Shell Correlation (GSFSC) curves (threshold 0.143) of *C. crescentus* tad pilus **(B)** Map:model Fourier Shell Correlation (FSC) curves (threshold 0.4=√.143) for *V. cholerae* TCP.



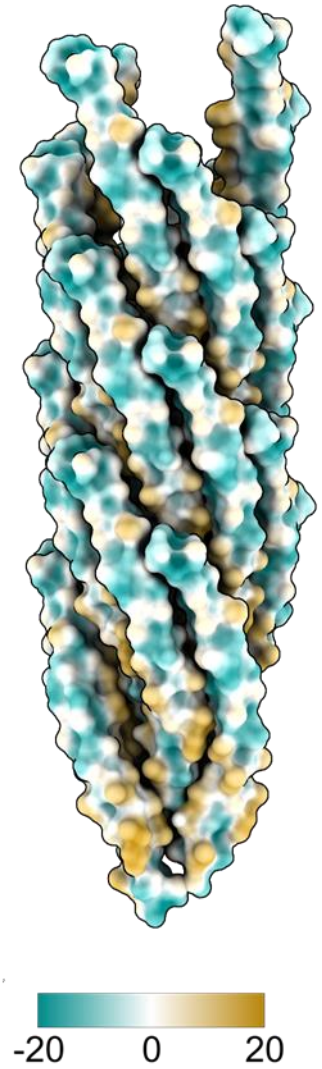
Supplementary Figure 3: Multiple sequence alignment of bacterial tad/flp pilins. The residues containing the polar side chain are shown in blue. Red boxes indicate the conservation of the solvent accessible polar residues present at C-terminal end of helix. Position 22 is indicated by * to show the absence of prolines in the tad/flp pilins. The conserved 'flp motif' is indicated by a red line. Abbreviations, *Aa*: *Aggregatibacter actinomycetemcomitans; Cc*: *Caulobacter crescentus*; *Pb*: *Pseudomonadota bacterium*; *Os*: *Oceanicaulis* sp.; *Ps*: *Phenylobacterium* sp.; *Ab*: *Alphaproteobacteria bacterium*; *Bs*: *Brevundimonas* sp.; *Bc*: *Brucella cytisi*; *Rc*: *Rhizobium* sp. CRIBSB; *Rb*: *Roseibacterium beibuensis*.



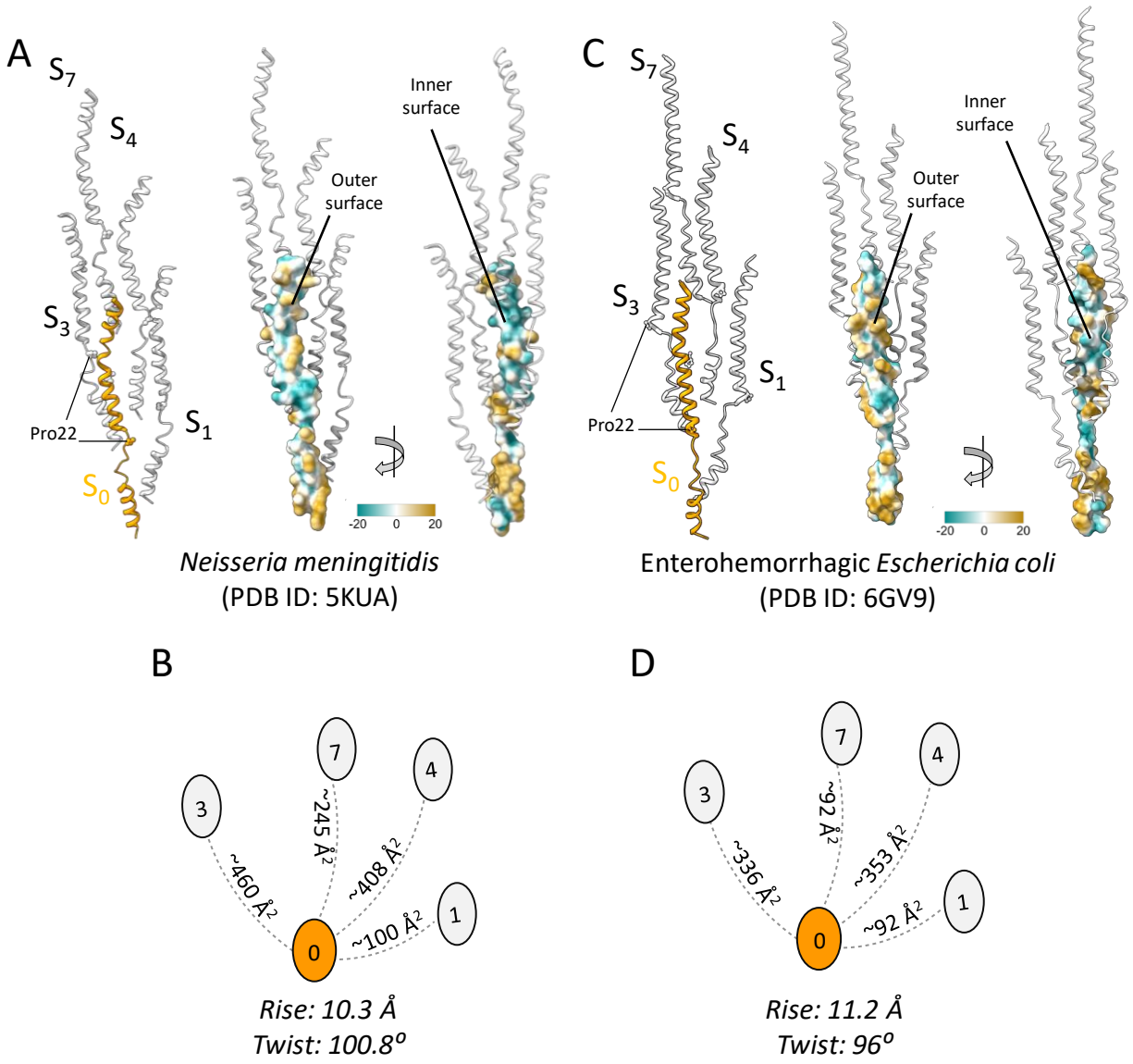
Supplementary Figure 4: Specific polar interactions between *C. crescentus* tad pilins. Interaction I: S_0 Thr2 - S_1 Ala1; Interaction II: S_0 Gly36 - S_3 Asn27; Interaction III: S_0 Ala43 - S_3 Lys34.



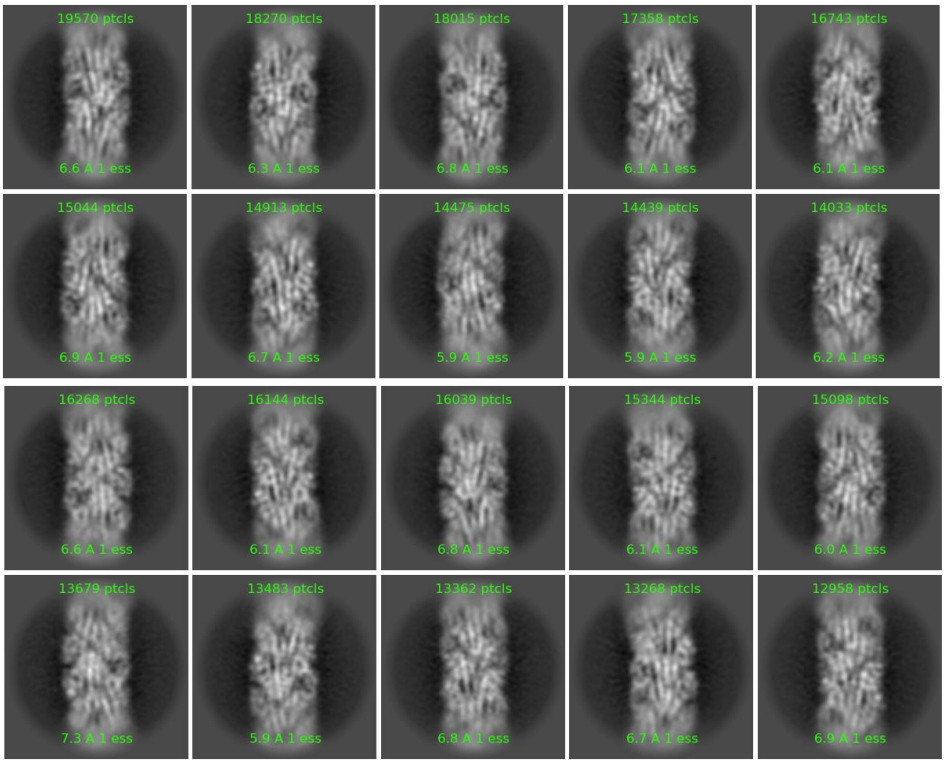
Supplementary Figure 5: Nonpolar interactions between tad pilins. **(A)** Tyr6 of S_0 subunit interacts with the Ala1 of the S_3 along with the interactions shown in Supplementary Figure 4 keeping tad pilins in helical register. **(B)** The Flp motif hydrophobic residues (Val1, Ala3, Ile4, Gly7, Leu8 and Ile9) of S_0 form extensive hydrophobic contacts with that of S_{-1}, S_1 and S_2, and to the residues 10-18 of S_{-3} stabilizing the core of tad filament.



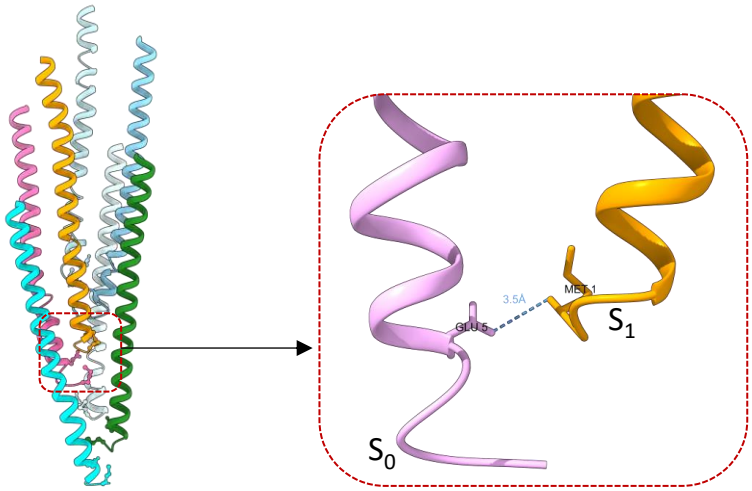
Supplementary Figure 6: Surface lipophilicity of *C. crescentus* tad pilus. Polar and hydrophobic surfaces are represented by green and gold color respectively.



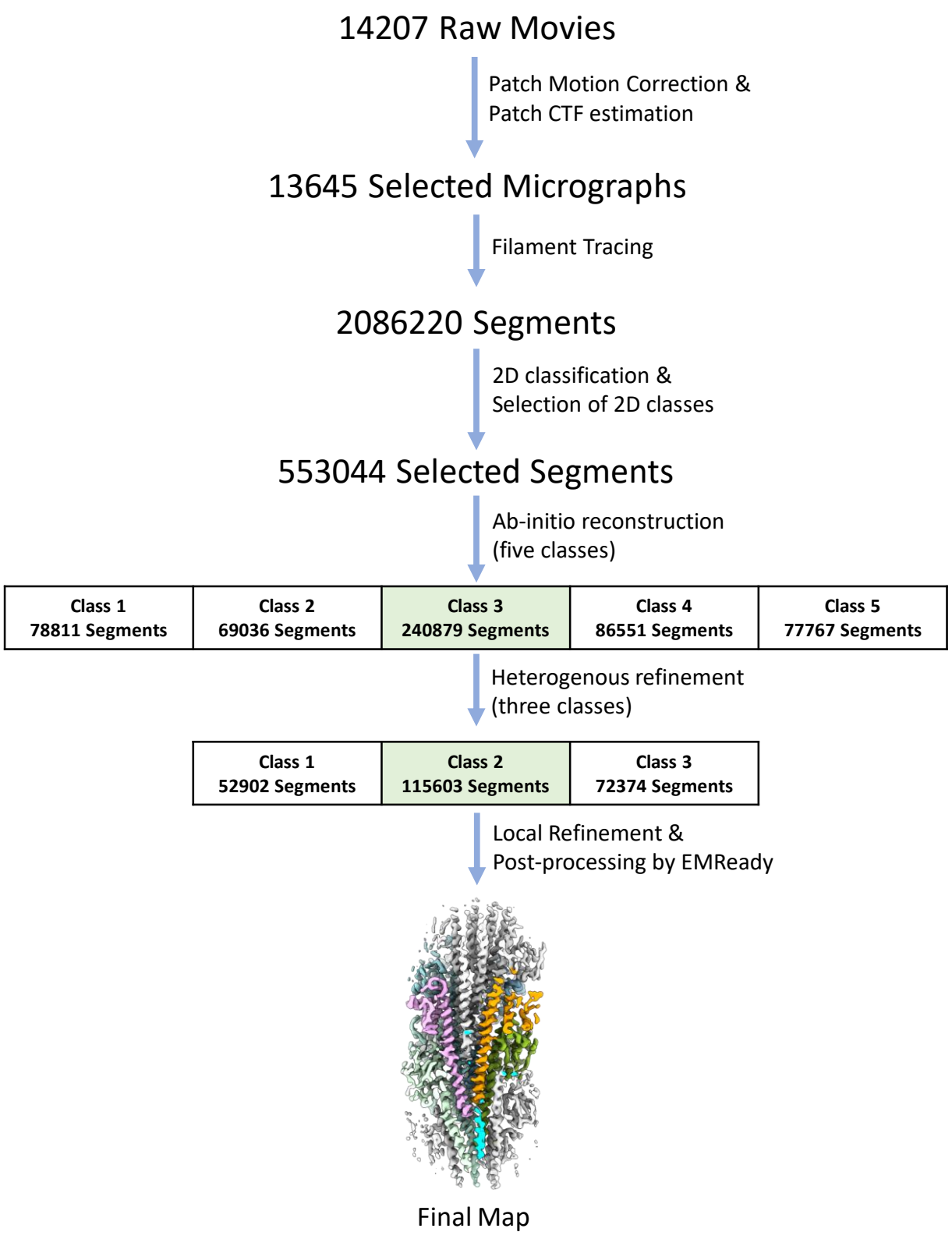
Supplementary Figure 7: Helical packing and surface lipophilicity of bacterial T4a pili. (A, B) *Neisseria meningitidis*, (C,D) EHEC. A and C show the helical packing with surface lipophilicity of the S_0 pilin (middle and right panels). The helix breaker Pro22 is shown as an atomic sphere. Polar-to-hydrophobic surfaces are represented by color gradient green-to-gold as shown in lipophilicity scale. B and D show the interfacial area between the helical symmetry mates calculated by PISA, and the helical symmetry parameters.



Supplementary Figure 8: Two-dimensional class averages of *V. cholerae* TCP filament showing its flexible morphology.



Supplementary Figure 9: Salt-bridge between the Glu5 of S_0 subunit and N-terminus of the S_1 subunit governs the helical registering of TCP pilins in the filament.



Supplementary Figure 10: Data processing workflow for the cryo-EM reconstruction of *V. cholerae* TCP filament.

Supplementary Table 1. Common features of members of the Type IV filament family

Features		T4aP	T4bP	T4cP (tad/flp)	Archaeal T4P	T2S endopili
Major pilin features	Signal peptide length (residues)	~ 6 aa	~25	7-40	~4-18	6-8
	Overall length of mature pilin (residues)	140 – 170	~200	40 – 60	125-300	~135
	Residue 1	Phe	variable	variable	variable	Phe
	Residue 2	Ser/Thr	Ser/Thr	Thr	Ser/Thr/Pro	Ser/Thr
	Residue 5	Glu	Glu	Glu	variable	Glu
	Residue 22	Pro	variable	variable	variable	Pro
Pilus machinery	Operons	multiple	single	single	multiple	single
	Retraction ATPase	yes	no	no	no	no
	No. of minor pilins	4	1	1-3	variable	4
	Examples	*Neisseria gonorrhoeae*, *Pseudomonas aeruginosa*, *Myxococcus xanthus* T4P, *V. cholera*e competence pilus	*Vibrio cholerae* Toxin Coregulated Pilus (TCP), enterotoxigenic *E. coli* (ETEC, CFA/III, Longus pili), *Citrobacter rodentium, Salmonella enterica* serovar Typhi, *Yersinia pseudotuberculosis* T4P, enteropathogenic *E. coli* bundle forming pilus (BFP)	*Caulobacter crescentus* Cpa pilius, *Aggregatibacter actinomycetemcomitans* Flp pili, *Salmonella enterica* R64 thin pilus, *Haemophilus ducreyi*, *Ralstonia solanacearum*	*Ignicoccus hospitalis*, *Sulfolobus solfataricus*, *Sulfolobus islandicus*, *Haloferax volcanii*, *Methanococcus maripaludis*	*V. cholerae*, ETEC, *P. aeruginosa*, *Klebsiella* oxytoca, *Legionella pneumophila*, *Dickeya dadantii*

Supplementary Table 2: Cryo-EM data collection and refinement statistics of *C. crescentus* tad pilus and *V. cholerae* TCP reconstructions and models

	***C. crescentus* tad pilus**	***V. cholerae* TCP (Asymmetric)**	***V. cholerae* TCP (Helical)**
PDB entry	8U1K	8UHF	-
EMBD entry	41815	42279	41968
Voltage (kV)	300	300	300
Electron exposure ($e^-/Å^2$)	50	50	50
Pixel size (Å/pixel)	0.834	0.83	0.83
Particle images (no.)	179116	115603	33045
Helical parameters (axial rise / twist)	4.9 Å / 102.8^o	-	~8.3 Å / ~97.0^o
Map global resolution (Å)			
Map:map FSC (0.143)	3.4 Å	3.8 Å	5.8 Å
Model:map FSC (0.5)	3.3 Å	4.0 Å	-
Refinement and model validation			
Clash score	14.42	9.99	
RMSD			
Bond Length (Å)	0.004	0.002	
Bond Angle (°)	0.603	0.515	
Ramachandran plot statistics (%)			
Favoured	97.67	96.96	
Allowed	2.33	3.04	
Outliers	0.00	0.00	
Model composition	**Chain ID: Residues** A: 1-45 B: 1-45 C: 1-45 D: 1-45 E: 1-45 F: 1-45 G: 1-45 H: 1-45 I: 1-45 J: 1-45	**Chain ID: Residues** A: 1-53, 61-199 B: 1-198 C: 1-198 D: 1-55, 60-196 E: 1-52, 61-198 F: 1-54, 61-198 G: 1-198 H: 2-52, 62-198 I: 2-55, 62-198	

Legend for Supplementary Movie 1: Animation of the averaged power spectra of segments from two different classes. The change in the spacing of the strongest layer line (n=-3) corresponds to a change in the pitch of the 3-start helices from ~116 Å to ~127 Å. A reference-based classification of the TCP segments using models with different axial rise values was employed to obtain these classes.