Structures of the Type IX Secretion/Gliding Motility Motor from across the Phylum *Bacteroidetes*

Rory Hennell James,[a,b*] Justin C. Deme,[a,c,d] Alicia Hunter,[a] Ben C. Berks,[b] Susan M. Lea[a,c,d]

[a]Sir William Dunn School of Pathology, University of Oxford, Oxford, United Kingdom
[b]Department of Biochemistry, University of Oxford, Oxford, United Kingdom
[c]The Central Oxford Structural Molecular Imaging Centre (COSMIC), University of Oxford, Oxford, United Kingdom
[d]Center for Structural Biology, Center for Cancer Research, National Cancer Institute, Frederick Maryland, USA

ABSTRACT Gliding motility using cell surface adhesins, and export of proteins by the type IX secretion system (T9SS) are two phylum-specific features of the *Bacteroidetes*. Both of these processes are energized by the GldLM motor complex, which transduces the proton motive force at the inner membrane into mechanical work at the outer membrane. We previously used cryo-electron microscopy to solve the structure of the GldLM motor core from *Flavobacterium johnsoniae* at 3.9-Å resolution (R. Hennell James, J. C. Deme, A. Kjaer, F. Alcock, et al., Nat Microbiol 6:221–233, 2021, https://dx.doi.org/10.1038/s41564-020-00823-6). Here, we present structures of homologous complexes from a range of pathogenic and environmental *Bacteroidetes* species at up to 3.0-Å resolution. These structures show that the architecture of the GldLM motor core is conserved across the *Bacteroidetes* phylum, although there are species-specific differences at the N terminus of GldL. The resolution improvements reveal a cage-like structure that ties together the membrane-proximal cytoplasmic region of GldL and influences gliding function. These findings add detail to our structural understanding of bacterial ion-driven motors that drive the T9SS and gliding motility.

IMPORTANCE Many bacteria in the *Bacteroidetes* phylum use the type IX secretion system to secrete proteins across their outer membrane. Most of these bacteria can also glide across surfaces using adhesin proteins that are propelled across the cell surface. Both secretion and gliding motility are driven by the GldLM protein complex, which forms a nanoscale electrochemical motor. We used cryo-electron microscopy to study the structure of the GldLM protein complex from different species, including the human pathogens *Porphyromonas gingivalis* and *Capnocytophaga canimorsus*. The organization of the motor is conserved across species, but we find species-specific structural differences and resolve motor features at higher resolution. This work improves our understanding of the type IX secretion system, which is a virulence determinant in human and animal diseases.

KEYWORDS *Bacteroidetes*, type IX secretion, motility, protein secretion, structural biology

The type IX secretion system (T9SS) is a protein export system found exclusively in the *Bacteroidetes* phylum of Gram-negative bacteria (1, 2). Substrates of the T9SS are transported to the periplasm by the Sec system, following which a conserved C-terminal domain (CTD) directs export through an outer membrane T9SS translocon (Fig. 1a) (2, 3). In most cases, the CTD is then removed and the substrate protein is either released into the environment or anchored to the outer membrane as a lipoprotein (4). The human oral pathogen *Porphyromonas gingivalis* uses the T9SS to secrete gingipain proteases and other virulence factors to evade the host immune system (5). The T9SS has also been identified as essential to the virulence of several economically relevant fish and poultry pathogens (6–8). In commensal and environmental *Bacteroidetes* species, the T9SS is characteristically used to secrete enzymes that enable the organisms to utilize complex polysaccharides as a food
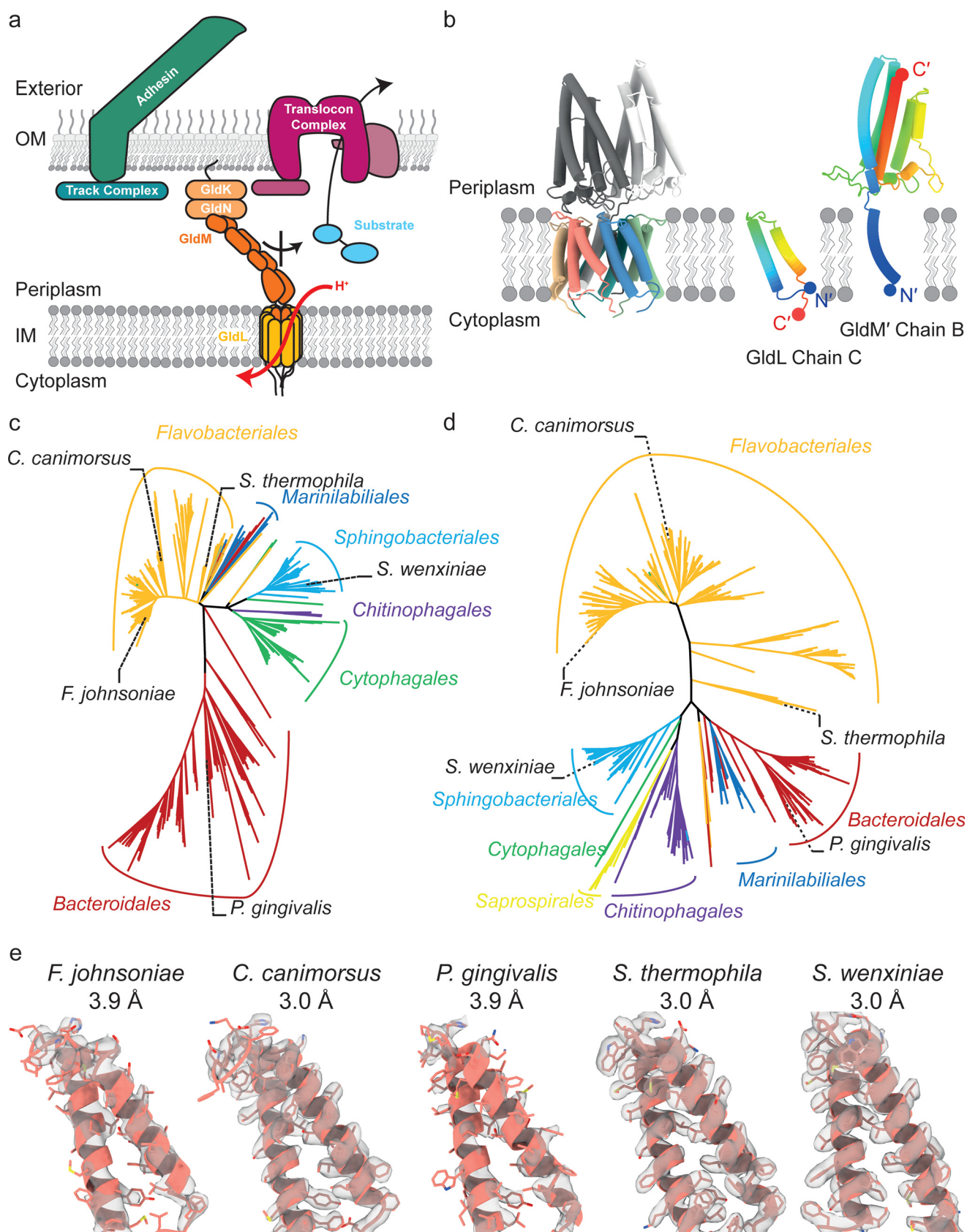
FIG 1 Role and phylogenetic diversity of the GldLM motor complex. (a) Cartoon illustrating the involvement of the GldLM motor complex in theT9SS and gliding motility. The GldLM motor converts electrochemical potential energy from the proton-motive force across the inner membrane (IM) into mechanical

source (1, 9, 10). Many *Bacteroidetes* species with a T9SS also exhibit gliding motility, in which cells travel rapidly across surfaces (2). This motility depends on the movement of cell surface adhesin molecules that are secreted to the cell surface by the T9SS (Fig. 1a).

The T9SS and the gliding motility apparatus share a motor complex that uses the proton motive force (PMF) across the inner membrane to drive both protein transport and gliding adhesin movement at the outer membrane (11–13). The motor complex is formed from the integral inner membrane proteins GldL and GldM (11, 14). GldL has two transmembrane helices (TMHs) and a cytoplasmic domain. GldM has one TMH and a large periplasmic region, which crystal structures have shown forms an extended dimer of four domains (D1 to D4) (11, 15, 16). The periplasmic region of GldM is long enough to span the periplasm to contact the outer membrane components of the T9SS and gliding motility apparatus.

We previously solved the structure of the core of the GldLM motor complex from the gliding bacterium *Flavobacterium johnsoniae* (11). This structure contains the full transmembrane region of the motor complex, but the cytoplasmic domain of GldL was not visible, and the periplasmic region of GldM had been genetically truncated after the first (D1) domain, forming a construct termed *Fjo*GldLM′ (11). The structure reveals that the T9SS/gliding motor is a GldL$_5$GldM$_2$ heteroheptamer in which the 10 GldL TMHs surround the two GldM TMHs (Fig. 1b). The symmetry mismatch between the total number of GldL and GldM TMHs results in an inherently asymmetric relationship between the two types of subunit around the GldL ring. Amino acid substitutions showed that several conserved protonatable residues in the transmembrane helices of GldL and GldM are important for T9SS and gliding motility function. These residues are likely to be involved in coupling proton flow across the inner membrane to mechanical motions in the motor. Based on the organization of the GldLM transmembrane helices and on the structural homology of the transmembrane part of GldLM to the ion-driven motor complexes that drive bacterial flagella (17, 18), we proposed that GldLM forms a rotary motor in which the GldM subunits rotate within the ring of GldL helices (11, 17). The periplasmic domain of GldM is then envisaged to transmit this rotary motion across the periplasm to the outer membrane components of the T9SS and gliding motility systems.

The structure of the *Fjo*GldLM′ complex was determined to a resolution (3.9 Å) at which only limited information can be inferred about the position of mechanistically important amino acid side chains. In addition, the protein was captured in a single conformational state, providing only one snapshot of the mechanism of the motor complex. Here, we have sought to overcome these limitations in the structural characterization of the GldLM motor by determining the structures of the motor complexes from a phylogenetically diverse range of organisms. We anticipated that some of these complexes would allow structure determination at improved resolution and in alternative conformational states. Here, we present structures of the T9SS/gliding motor core from the human pathogens *P. gingivalis* and *Capnocytophaga canimorsus* and the environmental bacteria *Schleiferia thermophila* and *Sphingobacterium wenxiniae*.

RESULTS AND DISCUSSION

The architecture of GldLM′ complexes is conserved across the *Bacteroidetes* phylum. To structurally survey the diversity of gliding motility/T9SS motor complexes, we selected proteins from a range of *Bacteroidetes* bacteria for recombinant expression in *Escherichia coli*. These proteins were chosen to maximize phylogenetic spread, to include proteins from organisms growing at a range of temperatures and from different environments (marine, fresh water,

FIG 1 Legend (Continued)

work that the periplasmic portion of GldM transfers across the periplasm to the outer membrane (OM). This mechanical energy is used to drive gliding adhesin movement (left) and protein transport through the T9SS (right). Coupling between these processes and GldM is thought to be mediated by a GldKN lipoprotein complex. (b) Cartoon representation of the structure of the *F. johnsoniae* GldLM′ complex solved previously (11) (PDB no. 6SY8 and EMDB no. EMD-10893). (*Left*) Whole structure. The five GldL chains are colored salmon, blue, green, teal, and tan and the two GldM chains are colored dark gray and white. (*Right*) Individual GldL and GldM chains are shown and rainbow colored from the N terminus (blue) to the C terminus (red). The most N-terminal (N′) and C-terminal (C′) modeled residues of each chain are marked with a sphere. (c, d) Maximum-likelihood phylogenetic tree of GldL (c) and GldM (d) sequences in the *Bacteroidetes* phylum. Branches are colored by taxonomic order and the positions of proteins for which structures were determined are indicated. (e) Increased resolution of the new T9SS/gliding motor complex structures shows improved side chain density. Chain GldL$_c$ is shown for each species with EM density displayed at the same contour level.

TABLE 1 Cryo-EM data collection, refinement, and validation statistics for the *Pgi*PorLM'$_C$, *Sth*GldLM', *Cca*GldLM''$_{peri}$, GldLM''$_{TMH}$, and *Swe*GldLM' structures

Statistic	P. gingivalis PorLM' (PDB no. 7SAT, EMDB no. EMD-24956)	S. thermophila GldLM' (PDB no. 7SAU, EMDB no. EMD-24957)	C. canimorsus GldLM''$_{peri}$ (PDB no. 7SB2, EMDB no. EMD-24961)	C. canimorsus GldLM''$_{TMH}$ (PDB no. 7SAZ, EMDB no. EMD-24959)	S. wenxiniae GldLM' (PDB no. 7SAX, EMDB no. EMD-24958)
Data collection and processing					
Magnification (×)	81,000	105,000	105,000	105,000	105,000
Voltage (kV)	300	300	300	300	300
Electron exposure (e$^-$ Å^{-2})	55.6	62.4 (without fOM), 61.2 (with fOM)	59.1	59.1	56.9
Defocus range (μm)	1.0–3.0	1.0–3.0	1.0–3.0	1.0–3.0	1.0–3.0
Pixel size (Å)	0.832	0.832	0.832	0.832	0.832
Symmetry imposed	C1	C1	C1	C1	C1
Initial particle images (no.)	8,205,503	13,743,455	9,197,926	9,197,926	7,167,266
Final particle images (no.)	649,359	394,678	595,559	77,223	111,727
Map resolution (Å)	3.9	3.0	3.4	3.0	3.0
FSC threshold	0.143	0.143	0.143	0.143	0.143
Map resolution range (Å)	3.7–4.9	2.8–3.7	3.3–6.0	2.8–4.3	3.0–4.3
Refinement					
Initial model used (PDB code)	None	None	None	None	None
Model resolution (Å)	3.9	3.0	3.4	3.0	3.0
FSC threshold	0.143	0.143	0.143	0.143	0.143
Model resolution range (Å)	3.7–4.9	2.8–3.7	3.3–6.0	2.8–4.3	3.0–4.3
Map sharpening B factor (Å^2)	−200	−93	−122	−58	−83
Model composition					
No. of nonhydrogen atoms	6,581	6,317	4,728	5,708	5,301
No. of protein residues	822	791	600	722	689
Ligands	0	0	0	0	0
B factors (Å^2)					
Protein	72.16	51.61	93.02	55.39	63.20
Ligand	NA[b]				
RMS[a] deviation					
Bond length (Å)	0.004	0.005	0.005	0.005	0.009
Bond angle (°)	0.747	0.600	0.781	0.675	0.700
Validation					
MolProbity score	2.16	1.68	2.39	1.82	1.50
Clashscore	15.42	5.98	22.84	9.65	5.09
Poor rotamers (%)	0.28	0.15	0.00	0.00	0.00
Ramachandran plot					
Favored (%)	92.57	94.85	90.34	95.48	96.44
Allowed (%)	7.43	5.15	9.66	4.38	3.56
Disallowed (%)	0.00	0.00	0.00	0.14	0.00

[a]RMS, root mean square.
[b]NA, not applicable.

terrestrial, and commensals/pathogens), and to include examples both from gliding bacteria and from nongliding bacteria with a T9SS. All constructs included the full-length GldL homologue. For the GldM homologue, we trialed constructs that included different numbers of periplasmic domains. However, as with the earlier *F. johnsoniae* GldLM' structure, we were only able to obtain structures by cryo-electron microscopy (cryo-EM) from complexes in which GldM was truncated after the first periplasmic D1 domain (GldM'), with the exception of one construct in which GldM was truncated after the D2 domain (GldM''). However, even in the latter case, this effectively produced a GldLM' structure, as the D2 domain was not resolved, as discussed below.

Following screening for expression, purification, and cryo-freezing, we determined structures for the T9SS/gliding motor complexes of *P. gingivalis* (*Pgi*PorLM'), *Capnocytophaga canimorsus* (*Cca*GldLM''$_{peri}$ and *Cca*GldLM''$_{TMH}$), *Schleiferia thermophila* (*Sth*GldLM'), and *Sphingobacterium wenxiniae* (*Swe*GldLM') (Fig. 1 and Table 1; see also Fig. S1 to S6 in the

supplemental material). *P. gingivalis* is a nongliding oral pathogen, the gliding bacterium *C. canimorsus* is a dog commensal and opportunistic human pathogen, *S. thermophila* is a thermophile isolated from a hot spring (optimum growth temperature, 50°C), and *S. wenxiniae* was isolated from a wastewater treatment plant. Note that in *P. gingivalis* the motor proteins are termed PorLM rather than GldLM. Figure 1c and d show the phylogenetic positions of these motor proteins within the diversity of GldL and GldM proteins. The GldL and GldM sequence similarity matches the bacterial phylogenetic tree in most cases, indicating that GldL has been predominantly vertically rather than horizontally transmitted, in agreement with an earlier analysis (19). The resolutions of the new structures ranged from 3.0 to 3.9 Å, with the higher-resolution structures allowing for more confident positioning of side chains than our previous 3.9-Å resolution *Fjo*GldLM' structure (Fig. 1e).

The overall architecture of the *Fjo*GldLM' complex is conserved in the four new motor complex structures, with five copies of GldL surrounding two copies of GldM (Fig. 2a to c). As in the *Fjo*GldLM' structure, only the transmembrane helices and periplasmic loop of GldL were fully resolved, with almost all of the C-terminal cytoplasmic domain not seen in the structures (11). The periplasmic D1 domains of the GldM dimer were visible in all structures. The precise angles between the two copies of the D1 domain varied from 30° to 45° between structures. However, in all cases the D1 pairs adopted the splayed arrangement seen in the previous *Fjo*GldLM' structure. This splayed arrangement contrasts with the closed arrangement of the D1 pair seen in crystal structures of the isolated periplasmic domain of GldM/PorM (15, 16), but is consistent with the low-resolution cryo-EM structure of full-length *Pgi*PorLM (11).

Structural data were obtained from a *C. canimorsus* construct (*Cca*GldLM″), which includes both the D1 and D2 periplasmic domains of GldM. During three-dimensional (3D) classification, classes could be identified by either the D2 domain or the transmembrane portion of the complex being well-resolved (the latter corresponding to the GldLM' complex structures determined from other organisms), but never both (Fig. S2 and S3 and Fig. S7a and b). This suggests that the *Cca*GldLM″ complex did not adopt a conformation in which both the transmembrane helices and D2 domains are simultaneously ordered. In the *Cca*GldLM″ structure with the ordered transmembrane helices (*Cca*GldLM″$_{TMH}$), the D1 domains are splayed as in the other GldLM'/PorLM' structures and in the low-resolution structure of the complete PorLM complex (Fig. S7a and d, 11). In contrast, in the *Cca*GldLM″ structure where the transmembrane domain is unresolved (*Cca*GldLM″$_{peri}$), the D1 domains adopt a parallel orientation (Fig. S7b, c, and e) similar to the isolated *Fjo*GldM periplasmic domain crystal structure (15), even though no crystal contacts are present. These observations suggest that splaying of the D1 domains is the most stable arrangement of the D1 domains in the intact motor complex, but leaves open the possibility that a parallel arrangement of the D1 domains might occur transiently during operation of the motor.

Mutagenesis was previously used to identify residues within the transmembrane domain of the *F. johnsoniae* GldLM motor that are important for function (11). These residues are well conserved in the new motor structures, with the exception that *Fjo*GldL Tyr13 is replaced by isoleucine in *S. wenxiniae* GldLM. The side chain positions of these residues can be assigned with more confidence due to the improved resolution of the *Cca*GldLM″$_{TMH}$, *Sth*GldLM', and *Swe*GldLM' structures. Notably, the orientation of these side chains is similar between the different motor structures where the resolution of the structures allows this judgement (Fig. 2e). The side chain position is least well defined for the *Fjo*GldL Glu49 equivalent, which we previously proposed forms a salt bridge between one copy of GldL and Arg9 in one of the two *Fjo*GldM molecules (11). The poor definition of this residue is not surprising, as glutamate sidechains are susceptible to damage by the electron beam during cryo-EM experiments, meaning that they are often not visible in EM density maps and cannot be accurately modeled (20). Nevertheless, in all the motor protein structures, the side chain of the *Fjo*GldM Arg9 equivalent is oriented toward the *Fjo*GldL Glu49 equivalent, suggesting that a salt bridge interaction between these residues is a conserved feature of the GldLM/PorLM complex.

The membrane-proximal region of GldL forms a cage-like structure. The majority of the cytoplasmic region of GldL was not visible in any of our structures. However, the part
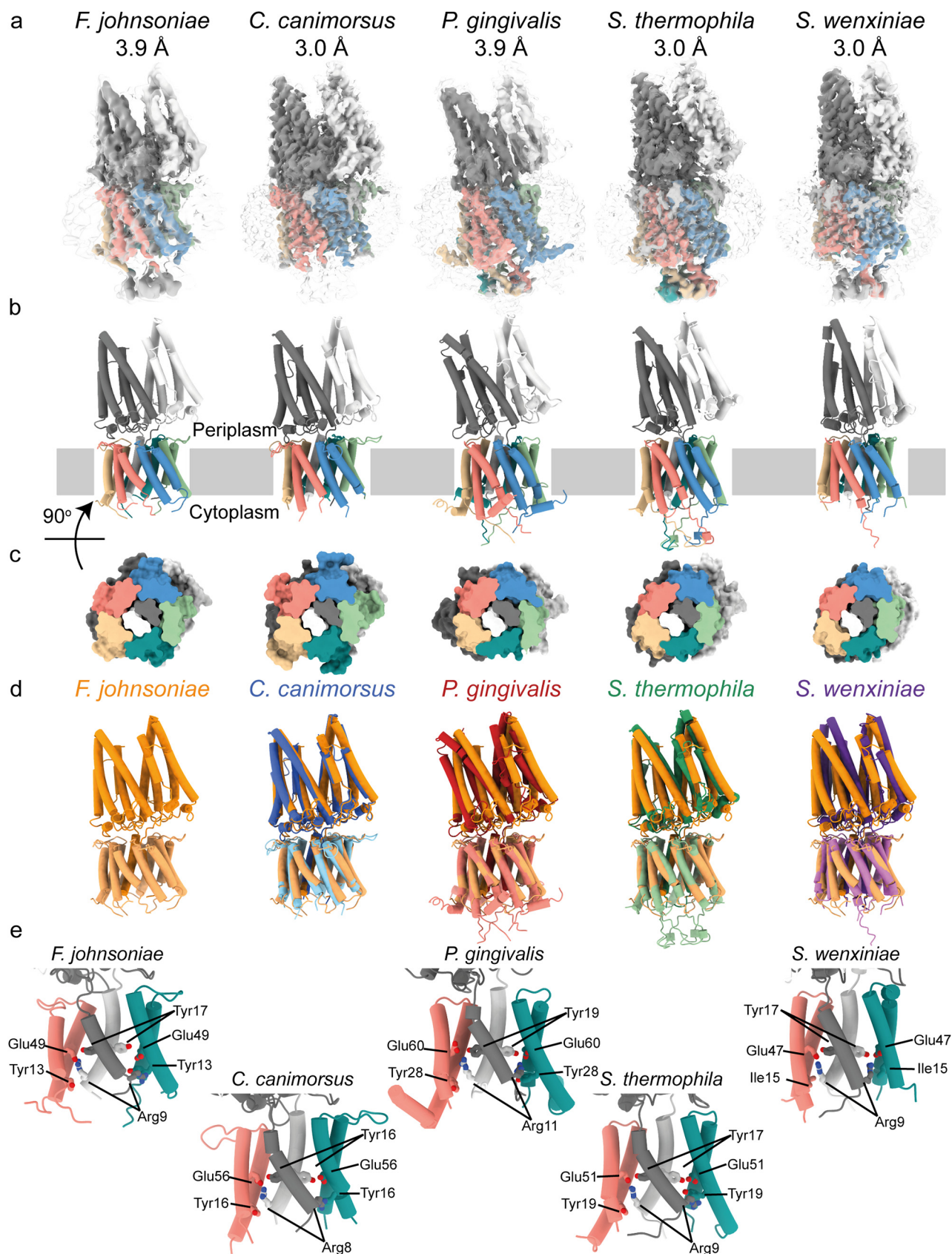
FIG 2 GldLM′ has conserved architecture across the *Bacteroidetes* phylum. (a) EM density maps for GldLM″_TMH/GldLM′/PorLM′ complexes from the indicated species at high (colored by protein chain) and low (transparent) contour. The structure of *F. johnsoniae* GldLM′ was solved previously (11).

(Continued on next page)

that is immediately adjacent to the inner face of the cytoplasmic membrane was much better resolved in the *Pgi*PorLM′, *Sth*GldLM′, and *Swe*GldLM′ structures than in the previously determined *Fjo*GldLM′ structure. It can now be seen that extended coils from the C-terminal end of TMH2 form a cage-like structure below the detergent micelle (Fig. 3a). In *Swe*GldLM′, the coil was only fully resolved for chain GldL$_C$, which is involved in the putative salt bridge with chain GldM$_B$. This coil is braced by the N terminus of the adjacent chain GldL$_D$, an interaction not seen in the other structures. The cage structure is resolved best in the *Sth*GldLM′ complex, revealing that the constituent coils are held together by a network of hydrogen bonds and hydrophobic packing interactions between aspartate, tryptophan, and tyrosine residues (Fig. 3c and d). The cage structure exhibits high sequence conservation, suggesting that it is of structural and/or functional importance (Fig. 3b and e). The cytoplasmic interactions between the GldL chains in the cage structure could help coordinate movements between subunits that do not contact each other in the TMH bundle. The surface of the cage is acidic (Fig. 3f) creating a region of negative charge that may assist in the release of protons flowing through the transmembrane part of the motor complex. Alternatively, it may help maintain separation between the cytoplasmic domain of GldL and the negatively charged phospholipid head groups of the cytoplasmic membrane.

We investigated the importance of the cage structure to motor function through mutagenesis of the chromosomal *gldL* gene in the genetically tractable organism *F. johnsoniae*. Mutant cells in which the cage was completely deleted (removal of residues 64 to 74; *gldL*$_{\Delta cage}$) had a small but reproducible gliding defect as measured by colony spreading on agar plates, while cells in which the cage sequence was replaced by a GSS repeat linker of the same length (*gldL*$_{CageToGSS}$) showed no gliding defect (Fig. 3g). The fact that changing the cage sequence to a GSS repeat had no effect on motility on agar indicates that the length of this cage region, rather than its precise sequence or structure, is most important to motor function.

***P. gingivalis* PorL has a N-terminal helix that is absent in other structurally characterized GldL/PorL proteins.** The *P. gingivalis* PorL protein has a N-terminal extension relative to the other motor proteins that we have structurally characterized (Fig. 4a). This extension forms a helix that is not present in the other structures (Fig. 4b). The helix points away from the transmembrane GldL helix bundle at an angle of approximately 100° from TMH1 (Fig. 4c and d) and approximately tangential to the circumference of the bundle (Fig. 4b). The helix lies against the curved surface of the detergent micelle (Fig. 4e) suggesting that *in vivo* it is likely to lie along the membrane surface with the positively charged N terminus (Fig. 4b and c) interacting with the negatively charged phospholipid head groups. This helix may, therefore, play a role in stabilizing the position of the PorLM complex in the cytoplasmic membrane. The presence of this N-terminal helix in *Pgi*PorLM′ and the bracing interaction seen between the N terminus of GldL$_D$ and the cytoplasmic region of GldL$_C$ in the *Swe*GldLM′ structure noted in the last section suggests a role for the N terminus of GldL in species-specific functional tuning of the motor complex.

Conclusion. This work, together with our previous study (11), provides the structures of the transmembrane cores of five T9SS/gliding motor complexes from species across three orders of the *Bacteroidetes*. These structures show that the architecture of the GldLM motor complex is well conserved and imply that the mechanism by which the motor converts proton flow to mechanical movement is the same across the *Bacteroidetes* phylum. The yield of the recombinant *S. wenxiniae* GldLM′ complex is much higher than that of the previously purified *Fjo*GldLM′ protein (Fig. S1) (11), which should expedite future *in vitro* mechanistic studies of the T9SS/gliding motor. Future work should also explore how GldL

FIG 2 Legend (Continued)

The resolution of the structures is indicated above the panels. (b) Cartoon representations of the structures with chains colored as in panel a. (c) Slab through the protein density from panel a viewed from the cytoplasm and sliced approximately half-way through the membrane region. (d) The new GldLM″$_{TMH}$/GldLM′/PorLM′ complex structures (colored as indicated) overlaid on *F. johnsoniae* GldLM′ (orange). GldM″$_{TMH}$/GldM′/PorM′ subunits are shown in a darker shade than the GldL/PorL subunits. (e) Conservation of residues that are functionally important in *F. johnsoniae* GldLM (11) (top left panel) in other GldLM″$_{TMH}$/GldLM/PorLM complexes. The proposed intersubunit salt bridge is between the labeled Glu residue in the salmon GldL chain and the labeled Arg residue in the white GldM chain. For clarity chains GldL$_D$, GldL$_E$ and GldL$_G$ are hidden for each structure.
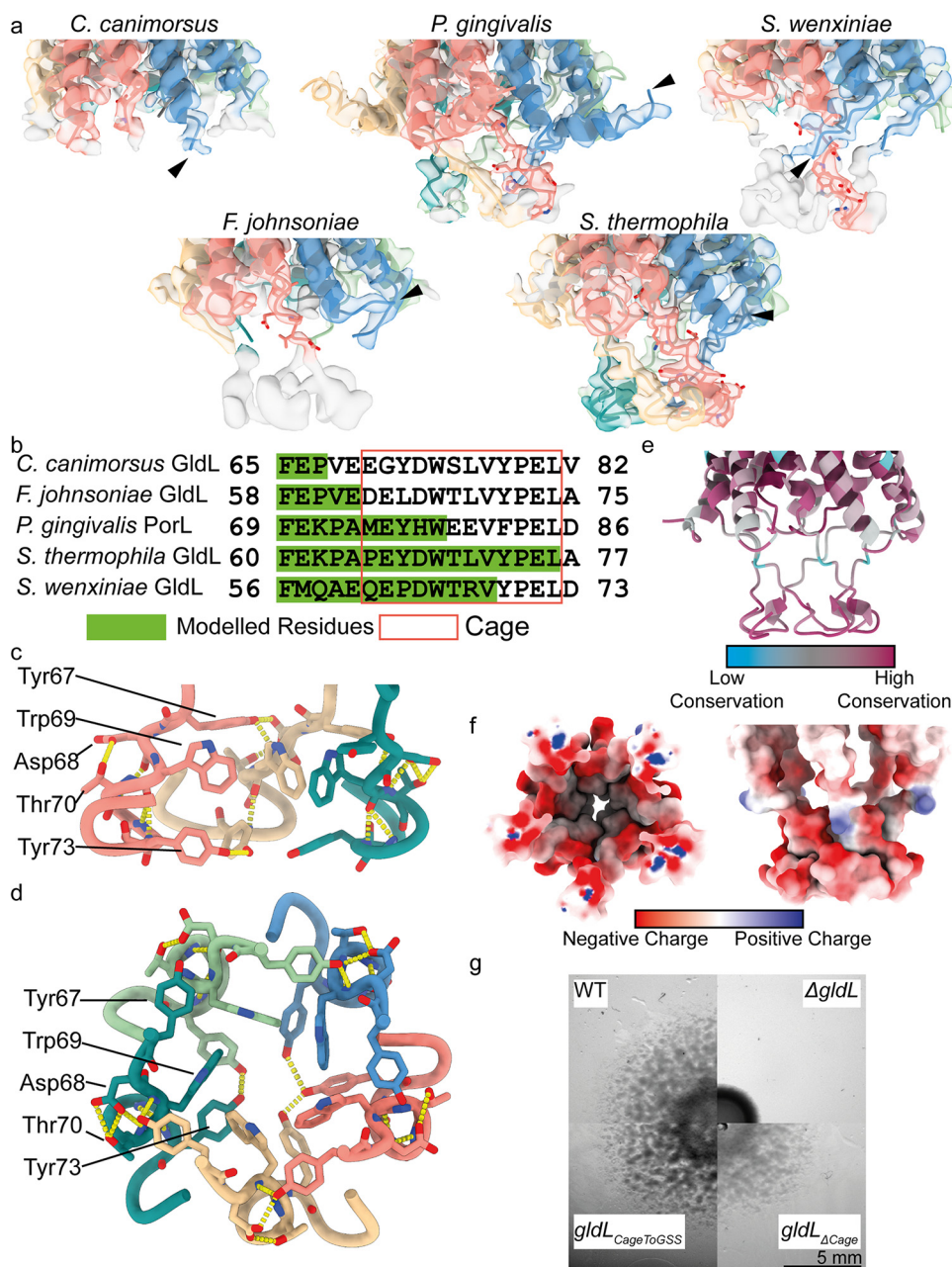
FIG 3 The membrane-proximal part of the GldL cytoplasmic domain forms a cage. (a) Overlay of EM density and the built model for the cytoplasmic region of each GldL structure. EM density is shown at the same contour level for all species. Side chains are shown for Chain C in the cage region. The most N-terminal residue modeled for Chain D is indicated with an arrowhead to highlight the bracing interaction between Chains C and D of *Swe*GldLM'. (b) Sequence alignment of the cage region for the five GldL/PorL sequences. Residues that could be modeled for each structure are highlighted in green. Residues constituting the cage region are boxed in red. (c, d) Interaction network at the base of the cage-like structure in *Sth*GldLM'. The view direction is parallel to the membrane (c) or from within the TMH bundle (d). For clarity chains D and F are hidden in panel c. Hydrogen bonds are shown as yellow dashes. Selected side chains are displayed for the same residues in each chain and are labeled on (c) Chain C (salmon) or (d) Chain F (teal). For other residues, backbone atoms are shown if they form hydrogen bonds. (e) Sequence conservation analysis of the cage region of *S. thermophila* GldL using the program Consurf (40, 41). (f) Coulombic surface potential representation for the cage region of *S. thermophila* GldL. The view direction is from within the TMH bundle (left) or parallel to the membrane (right). For clarity, the first TMH and N-terminal residues of GldL are hidden. (g) Effects of modifications in the GldL cytoplasmic cage on *F. johnsoniae* gliding motility (spreading) on plates. The region of the *F. johnsoniae* sequence enclosed by the red box in panel b (residues E64-L74) was either substituted with GSSGSSGSSGS ($gldL_{CagetoGSS}$) or deleted ($gldL_{\Delta Cage}$). The results are representative of three independent experiments.
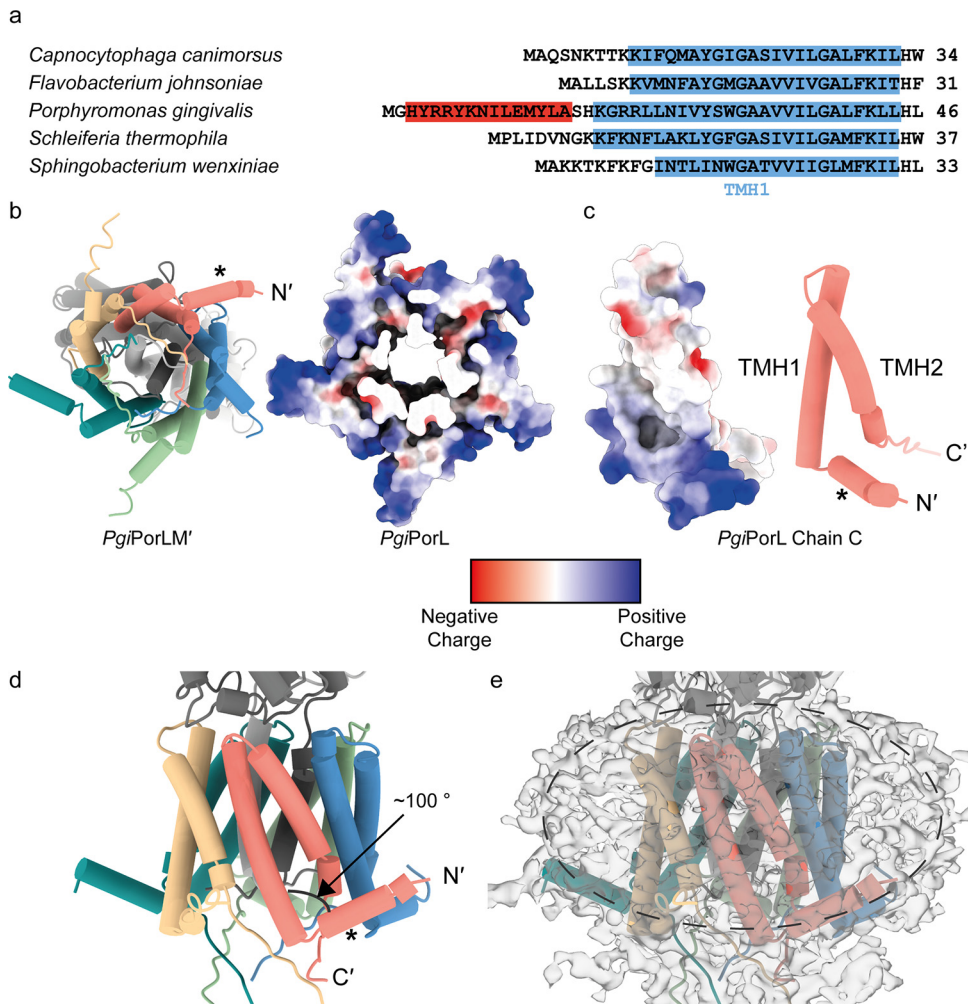
a

Capnocytophaga canimorsus	MAQSNKTTK**KIFQMAYGIGASIVILGALFKIL**HW	34
Flavobacterium johnsoniae	MALLSK**KVMNFAYGMGAAVVIVGALFKIT**HF	31
Porphyromonas gingivalis	MG**HYRRYKNILEMYLA**SH**KGRRLLNIVYSWGAAVVILGALFKLL**HL	46
Schleiferia thermophila	MPLIDVNGK**KFKNFLAKLYGFGASIVILGAMFKIL**HW	37
Sphingobacterium wenxiniae	MAKKTKFKFG**INTLINWGATVVIIGLMFKIL**HL	33
	TMH1	

FIG 4 *P. gingivalis* PorL has a N-terminal membrane surface-associated helix. (a) Sequence alignment of the N-terminal regions of the structurally characterized GldL homologues. The first transmembrane helix of each sequence is highlighted in blue. The additional N-terminal helix of *Pgi*PorL is highlighted in red. (b) View from the cytoplasm of *P. gingivalis* PorLM′ in cartoon (left) and coulombic surface (right) representation. For clarity PorM′ is hidden in the coulombic view. (c) Coulombic surface (left) and cartoon (right) representations of *P. gingivalis* PorL Chain C viewed from within the membrane. (d) Side view of *Pgi*PorLM′ in cartoon representation. (e) Side view of the *Pgi*PorLM′ complex model overlaid with the EM density map displayed at low contour level. The approximate boundary of the detergent micelle is marked with a dashed line. (b to d) An asterisk (*) indicates the N-terminal helix, and C′ indicates the most C-terminal modeled residue.

and GldM interact with other components of the T9SS and gliding motility machinery to convert motor motions into the useful work of protein translocation and adhesin propulsion.

MATERIALS AND METHODS

Bioinformatics analysis. A phylogenetic tree of GldL sequences was generated as follows. GldL sequences were obtained by BLAST searches against the UniRef90 database using the *F. johnsoniae* GldL, *P. gingivalis* PorL, and *S. wenxiniae* GldL sequences as queries (21, 22). Each sequence in the UniRef90 database represents a cluster of sequences with more than 90% identity to the representative sequence. Searching against the UniRef90 database reduces the number of highly similar sequences in the results compared to searching against an unfiltered database. Sequences duplicated between the three BLAST searches, sequences representing clusters where the lowest common taxon was higher than order (and thus likely incorrectly phylogenetically assigned), and sequences from organisms outside the *Bacteroidetes* phylum were removed. The sequences were then aligned using Clustal Omega (23). A phylogenetic tree was inferred using the maximum likelihood function of the MEGA X program with a Jones-Taylor-Thornton (JTT) matrix-based model with default settings (24, 25). A phylogenetic tree of GldM sequences was generated using the same approach.

Bacterial strains and growth conditions. Strains and plasmids used in this study are listed in Table S1 in the supplemental material. For cloning procedures, *E. coli* cells were routinely grown in

Luria-Bertani (LB) medium (26) at 37°C with shaking. *F. johnsoniae* cells were routinely grown in Casitone yeast extract (CYE) medium (27) at 30°C with shaking. PY2 medium (28) was used to assess motility on agar plates. When required, kanamycin was added to LB medium at 30 μg · mL^{-1} or to Terrific Broth (TB) medium at 50 μg · mL^{-1}. When required, erythromycin was added at 100 μg · mL^{-1}.

Genetic constructs. Primers used in this work are described in Table S2. All plasmid constructs were verified by sequencing.

GldL and C-terminally truncated and twin-Strep-tagged GldM proteins (GldM'-TS/GldM"-TS) were expressed from vectors derived from the plasmid pT12 (29) under the control of rhamnose-inducible promoters.

Suicide vectors to genetically modify *F. johnsoniae* were produced using the vectors pRHJ012 (11) and pYT354 (30), then introduced into the *F. johnsoniae* Δ*gldL* strain FJ_082 (11) using *Escherichia coli* strain S17-1 (31) as previously described (11).

Full details of all genetic constructs are given in Text S1.

Purification of protein complexes. Briefly, protein complexes were overexpressed in BL21(DE3) cells and then extracted from cell membranes using lauryl maltose neopentyl glycol (LMNG; Anatrace). Protein complexes were then affinity purified using Strep-Tactin XT resin (IBA) and then further purified using size exclusion chromatography. Full details of the purification scheme are given in Text S1.

Typical yields per liter of cell culture were as follows: 20 μg (*Cca*GldLM"); 50 μg (*Pgi*PorLM'); 6 μg (*Sth*GldLM'); 350 μg (*Swe*GldLM').

Cryo-EM sample preparation and imaging. Aliquots (4 μL) of purified samples at an A_{280} of 1 were applied onto glow-discharged holey carbon-coated grids (Quantifoil 300 mesh, Au R1.2/1.3), then adsorbed for 10 s, blotted for 2 s at 100% humidity at 4°C, and plunge-frozen in liquid ethane using a Vitrobot Mark IV (FEI). To prepare samples with fluorinated octyl maltoside (fOM; Anatrace), proteins were concentrated to an A_{280} of 3 and 13.5 μL was mixed with 1.5 μL of 7 mM fOM in buffer W (100 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA) plus 0.01% LMNG. All samples were centrifuged at 18,400 × *g* for 10 min at 4°C immediately before grid preparation.

Data were collected using a Titan Krios G3 instrument (FEI) operated at 300 kV and fitted with either a GIF energy filter (Gatan) and a K2 Summit detector (Gatan) or a BioQuantum imaging filter (Gatan) and a K3 direct detection camera (Gatan). Full details of the data collection strategy are given in Text S1.

Cryo-EM data processing. Motion correction, dose weighting, contrast transfer function determination, particle picking, and initial particle extraction were performed using SIMPLE 3.0 (32). Gold-standard Fourier shell correlations (FSC) using the 0.143 criterion, and local resolution estimations were calculated within RELION 3.1 (33).

In general, extracted particles were subjected to reference-free two-dimensional (2D) classification in SIMPLE, followed by 3D classification in RELION using either the previously solved *Fjo*GldLM' map (11) or another map produced in this study as a reference. Classes with clear secondary structure detail were selected and used for 3D autorefinement. Successive rounds of Bayesian particle polishing, 3D classification, and 3D autorefinement in RELION were used to generate the final maps for each data set. A full description of the data processing strategy for each data set is given in Text S1.

Model building and refinement. The Phyre^2 server was used to generate homology models for each new sequence from the structure of *Fjo*GldLM' (PDB no. 6SY8) using one-to-one threading (34). These models were rigid-body fitted into the cryo-EM volume using Coot, and residues were built *de novo* or removed as necessary in Coot (35). Rebuilding in globally sharpened and local-resolution-filtered maps was combined with real-space refinement in Phenix using secondary structure, rotamer, and Ramachandran restraints to give the final models described in Table 1 (36, 37). Validation was done in Molprobity (38). Structures were analyzed using ChimeraX (39), PyMOL 2.3.3 (Schrodinger), and the Consurf Server (40, 41).

Measurement of gliding motility on agar. Strains were grown overnight in PY2 medium, washed once in PY2 medium, then resuspended in PY2 medium to an optical density at 600 nm (OD$_{600}$) of 0.1. A 2-μL sample was then spotted onto PY2 agar plates. Plates were incubated at 25°C for 48 h before imaging with a Zeiss Axio Zoom MRm charge-coupled device (CCD) camera and Zeiss software (ZenPro 2012 v. 1.1.1.0).

Data availability. The cryo-EM volumes and atomic coordinates presented in this paper have been deposited in the Electron Microscopy Data Bank (EMDB) and the Protein Data Bank (PDB), respectively, with the following accession codes: *Cca*GldLM"$_{TMH}$ (EMDB no. EMD-24959 and PDB no. 7SAZ), *Cca*GldLM"$_{peri}$ (EMDB no. EMD-24961 and PDB no. 7SB2), *Pgi*PorLM' (EMDB no. EMD-24956 and PDB no. 7SAT), *Sth*GldLM' (EMDB no. EMD-24957 and PDB no. 7SAU), and *Swe*GldLM' (EMDB no. EMD-24958 and PDB no. 7SAX).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

TEXT S1, DOCX file, 0.03 MB.
FIG S1, TIF file, 1.9 MB.
FIG S2, TIF file, 1.6 MB.
FIG S3, TIF file, 2.9 MB.
FIG S4, TIF file, 1.7 MB.
FIG S5, PDF file, 0.08 MB.
FIG S6, PDF file, 0.06 MB.
FIG S7, TIF file, 2.1 MB.
TABLE S1, DOCX file, 0.02 MB.
TABLE S2, DOCX file, 0.02 MB.

REFERENCES

1. Sato K, Naito M, Yukitake H, Hirakawa H, Shoji M, McBride MJ, Rhodes RG, Nakayama K. 2010. A protein secretion system linked to bacteroidete gliding motility and pathogenesis. Proc Natl Acad Sci U S A 107:276–281. https://doi.org/10.1073/pnas.0912010107.

2. McBride MJ. 2019. *Bacteroidetes* gliding motility and the type IX secretion system. Microbiol Spectr 7:e0002-18. https://doi.org/10.1128/microbiolspec.PSIB-0002-2018.

3. Lauber F, Deme JC, Lea SM, Berks BC. 2018. Type 9 secretion system structures reveal a new protein transport mechanism. Nature 564:77–82. https://doi.org/10.1038/s41586-018-0693-y.

4. Lasica AM, Goulas T, Mizgalska D, Zhou X, De Diego I, Ksiazek M, Madej M, Guo Y, Guevara T, Nowak M, Potempa B, Goel A, Sztukowska M, Prabhakar AT, Bzowska M, Widziolek M, Thøgersen IB, Enghild JJ, Simonian M, Kulczyk AW, Nguyen KA, Potempa J, Gomis-Rüth FX. 2016. Structural and functional probing of PorZ, an essential bacterial surface component of the type-IX secretion system of human oral-microbiomic *Porphyromonas gingivalis*. Sci Rep 6:37708. https://doi.org/10.1038/srep37708.

5. Nakayama K. 2015. *Porphyromonas gingivalis* and related bacteria: from colonial pigmentation to the type IX secretion system and gliding motility. J Periodontal Res 50:1–8. https://doi.org/10.1111/jre.12255.

6. Guo Y, Hu D, Guo J, Wang T, Xiao Y, Wang X, Li S, Liu M, Li Z, Bi D, Zhou Z. 2017. *Riemerella anatipestifer* type IX secretion system is required for virulence and gelatinase secretion. Front Microbiol 8:2553. https://doi.org/10.3389/fmicb.2017.02553.

7. Yuan H, Huang L, Wang M, Jia R, Chen S, Liu M, Zhao X, Yang Q, Wu Y, Zhang S, Liu Y, Zhang L, Yu Y, You Y, Chen X, Zhu D, Cheng A. 2019. Role of the *gldK* gene in the virulence of *Riemerella anatipestifer*. Poult Sci 98:2414–2421. https://doi.org/10.3382/ps/pez028.

8. Li N, Zhu Y, LaFrentz BR, Evenhuis JP, Hunnicutt DW, Conrad RA, Barbier P, Gullstrand CW, Roets JE, Powers JL, Kulkarni SS, Erbes DH, García JC, Nie P, McBride MJ. 2017. The type IX secretion system is required for virulence of the fish pathogen *Flavobacterium columnare*. Appl Environ Microbiol 83:e01769-17. https://doi.org/10.1128/AEM.01769-17.

9. Cabral L, Persinoti GF, Paixão DAA, Martins MP, Morais MAB, Chinaglia M, Domingues MN, Sforca ML, Pirolla RAS, Generoso WC, Santos CA, Maciel LF, Terrapon N, Lombard V, Henrissat B, Murakami MT. 2022. Gut microbiome of the largest living rodent harbors unprecedented enzymatic systems to degrade plant polysaccharides. Nat Commun 13:629. https://doi.org/10.1038/s41467-022-28310-y.

10. Naas AE, Solden LM, Norbeck AD, Brewer H, Hagen LH, Heggenes IM, McHardy AC, Mackie RI, Paša-Tolić L, Arntzen MØ, Eijsink VGH, Koropatkin NM, Hess M, Wrighton KC, Pope PB. 2018. "Candidatus Paraporphyromonas polyenzymogenes" encodes multi-modular cellulases linked to the type IX secretion system. Microbiome 6:44. https://doi.org/10.1186/s40168-018-0421-8.

11. Hennell James R, Deme JC, Kjaer A, Alcock F, Silale A, Lauber F, Johnson S, Berks BC, Lea SM. 2021. Structure and mechanism of the proton-driven motor that powers type 9 secretion and gliding motility. Nat Microbiol 6:221–233. https://doi.org/10.1038/s41564-020-00823-6.

12. Dzink-Fox JAL, Leadbetter ER, Godchaux W. 1997. Acetate acts as a protonophore and differentially affects bead movement and cell migration of the gliding bacterium *Cytophaga johnsonae* (*Flavobacterium johnsoniae*). Microbiology 143:3693–3701. https://doi.org/10.1099/00221287-143-12-3693.

13. Nakane D, Sato K, Wada H, McBride MJ, Nakayama K. 2013. Helical flow of surface protein required for bacterial gliding motility. Proc Natl Acad Sci U S A 110:11145–11150. https://doi.org/10.1073/pnas.1219753110.

14. Vincent MS, Canestrari MJ, Leone P, Stathopulos J, Ize B, Zoued A, Cambillau C, Kellenberger C, Roussel A, Cascales E. 2017. Characterization of the *Porphyromonas gingivalis* type IX secretion trans-envelope PorKLMNP core complex. J Biol Chem 292:3252–3261. https://doi.org/10.1074/jbc.M116.765081.

15. Leone P, Roche J, Vincent MS, Tran QH, Desmyter A, Cascales E, Kellenberger C, Cambillau C, Roussel A. 2018. Type IX secretion system PorM and gliding machinery GldM form arches spanning the periplasmic space. Nat Commun 9:429. https://doi.org/10.1038/s41467-017-02784-7.

16. Sato K, Okada K, Nakayama K, Imada K. 2020. PorM, a core component of bacterial type IX secretion system, forms a dimer with a unique kinked-rod shape. Biochem Biophys Res Commun 532:114–119. https://doi.org/10.1016/j.bbrc.2020.08.018.

17. Deme JC, Johnson S, Vickery O, Aron A, Monkhouse H, Griffiths T, James RH, Berks BC, Coulton JW, Stansfeld PJ, Lea SM. 2020. Structures of the stator complex that drives rotation of the bacterial flagellum. Nat Microbiol 5:1553–1564. https://doi.org/10.1038/s41564-020-0788-8.

18. Santiveri M, Roa-Eguiara A, Kühne C, Wadhwa N, Hu H, Berg HC, Erhardt M, Taylor NMI. 2020. Structure and function of stator units of the bacterial flagellar motor. Cell 183:244–257. https://doi.org/10.1016/j.cell.2020.08.016.

19. Shrivastava A, Johnston JJ, Van Baaren JM, McBride MJ. 2013. *Flavobacterium johnsoniae* GldK, GldL, GldM, and SprA are required for secretion of the cell surface gliding motility adhesins SprB and RemA. J Bacteriol 195:3201–3212. https://doi.org/10.1128/JB.00333-13.

20. Hattne J, Shi D, Glynn C, Zee C-T, Gallagher-Jones M, Martynowycz MW, Rodriguez JA, Gonen T. 2018. Analysis of global and site-specific radiation damage in cryo-EM. Structure 26:759–766. https://doi.org/10.1016/j.str.2018.03.021.

21. Altschul SF, Miller W, Gish W, Myers EW, Lipman DJ. 1990. Basic Local Alignment Search Tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

22. Suzek BE, Wang Y, Huang H, McGarvey PB, Wu CH, the UniProt Consortium. 2015. UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. Bioinformatics 31:926–932. https://doi.org/10.1093/bioinformatics/btu739.

23. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, Lopez R. 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res 47:W636–W641. https://doi.org/10.1093/nar/gkz268.

24. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549. https://doi.org/10.1093/molbev/msy096.

25. Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8:275–282. https://doi.org/10.1093/bioinformatics/8.3.275.

26. Miller JH. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

27. McBride MJ, Kempf MJ. 1996. Development of techniques for the genetic manipulation of the gliding bacterium *Cytophaga johnsonae*. J Bacteriol 178:583–590. https://doi.org/10.1128/jb.178.3.583-590.1996.

28. Agarwal S, Hunnicutt DW, McBride MJ. 1997. Cloning and characterization of the *Flavobacterium johnsoniae* (*Cytophaga johnsonae*) gliding motility gene, *gldA*. Proc Natl Acad Sci U S A 94:12139–12144. https://doi.org/10.1073/pnas.94.22.12139.

29. Dietsche T, Tesfazgi Mebrhatu M, Brunner MJ, Abrusci P, Yan J, Franz-Wachtel M, Schärfe C, Zilkenat S, Grin I, Galán JE, Kohlbacher O, Lea S, Macek B, Marlovits TC, Robinson CV, Wagner S. 2016. Structural and functional characterization of the bacterial type III secretion export apparatus. PLoS Pathog 12:e1006071. https://doi.org/10.1371/journal.ppat.1006071.

30. Zhu Y, Thomas F, Larocque R, Li N, Duffieux D, Cladière L, Souchaud F, Michel G, McBride MJ. 2017. Genetic analyses unravel the crucial role of a horizontally acquired alginate lyase for brown algal biomass degradation by *Zobellia galactanivorans*. Environ Microbiol 19:2164–2181. https://doi.org/10.1111/1462-2920.13699.

31. Simon R, Priefer U, Pühler A. 1983. A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in Gram-negative bacteria. Nat Biotechnol 1:784–791. https://doi.org/10.1038/nbt1183-784.

32. Caesar J, Reboul CF, Machello C, Kiesewetter S, Tang ML, Deme JC, Johnson S, Elmlund D, Lea SM, Elmlund H. 2020. SIMPLE 3.0. Stream single-particle cryo-EM analysis in real time. J Struct Biol X 4:100040. https://doi.org/10.1016/j.yjsbx.2020.100040.

33. Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJH, Lindahl E, Scheres SHW. 2018. New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife 7:e42166. https://doi.org/10.7554/eLife.42166.

34. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10:845–858. https://doi.org/10.1038/nprot.2015.053.

35. Emsley P, Lohkamp B, Scott WG, Cowtan KD. 2010. Features and development of *Coot*. Acta Crystallogr D Biol Crystallogr 66:486–501. https://doi.org/10.1107/S0907444910007493.

36. Afonine PV, Poon BK, Read RJ, Sobolev OV, Terwilliger TC, Urzhumtsev A, Adams PD. 2018. Real-space refinement in *PHENIX* for cryo-EM and crystallography. Acta Crystallogr D Struct Biol 74:531–544. https://doi.org/10.1107/S2059798318006551.

37. Liebschner D, Afonine PV, Baker ML, Bunkoczi G, Chen VB, Croll TI, Hintze B, Hung LW, Jain S, McCoy AJ, Moriarty NW, Oeffner RD, Poon BK, Prisant MG, Read RJ, Richardson JS, Richardson DC, Sammito MD, Sobolev OV, Stockwell DH, Terwilliger TC, Urzhumtsev AG, Videau LL, Williams CJ, Adams PD. 2019. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in *Phenix*. Acta Crystallogr D Struct Biol 75:861–877. https://doi.org/10.1107/S2059798319011471.

38. Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, Verma V, Keedy DA, Hintze BJ, Chen VB, Jain S, Lewis SM, Arendall WB, Snoeyink J, Adams PD, Lovell SC, Richardson JS, Richardson DC. 2018. MolProbity: more and better reference data for improved all-atom structure validation. Protein Sci 27:293–315. https://doi.org/10.1002/pro.3330.

39. Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, Morris JH, Ferrin TE. 2021. UCSF ChimeraX: structure visualization for researchers, educators, and developers. Protein Sci 30:70–82. https://doi.org/10.1002/pro.3943.

40. Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, Ben-Tal N. 2016. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res 44:W344–W350. https://doi.org/10.1093/nar/gkw408.

41. Landau M, Mayrose I, Rosenberg Y, Glaser F, Martz E, Pupko T, Ben-Tal N. 2005. ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. Nucleic Acids Res 33:W299–W302. https://doi.org/10.1093/nar/gki370.