Convergent Evolution in the Supercoiling of Prokaryotic Flagellar Filaments

Mark Andrew Boudreau Kreutzberger

Charlottesville Virginia

BS in Chemistry, University of North Carolina Wilmington, 2016

A Dissertation presented to the Graduate Faculty of the University of Virginia in Candidacy for the Degree of Doctor of Philosophy

Department of Molecular Physiology and Biological Physics

University of Virginia
August 2022

Table of Contents

Acknowledgements

I would like to acknowledge all of my advisors, colleagues, friends, and family members who have provided me with tremendous support during my graduate schooling at the University of Virginia. My advisor Dr. Egelman has been very helpful and supportive of me during my PhD and pushed me to make progress on my projects which I never thought possible. I would also like to thank the former Biophysics program DGS Dr. Bob Nakamoto who has offered me crucial help during the early stages of my PhD as well as all of the members of my committee who have been helpful throughout the process.

Chapter 1. Introduction

1.1 The bacterial flagellum

The bacterial flagellum has three main components the motor, the hook, and the filament[1,2]. All bacterial flagella have a rotary motor. For some species such as *Salmonella,* and *E. coli* the motor rotation is bi-direction[2,3] (counter-clockwise and clockwise rotation) while for others such as *S. meliloti* and *Agrobacterium lupini* it is unidirectional[4]. The motor rotation is driven by an ion gradient which can be either proton motive force or sodium ion motive force depending on the species of bacteria[5]. The hook functions as a universal joint and transmits toque to the filament. The filament functions as a propeller and its supercoiling is key to motility as straight flagellar filaments are nonmotile and do not produce thrust. For many bacteria polymorphism is key to motility[3,6] and changes in mechanical force such as a rotational switching of the motor result in a change in the super helical waveform of the flagellum.

Bacterial species from the *Escherichia* and *Salmonella* genera have peritrichous flagella[7]. During running mode, the flagellar motors at the base of the flagellum are rotating counterclockwise, which orients the filaments in a bundle toward one end of the cell[3,8]. When the direction of rotation of one or several flagella changes to clockwise, the affected flagellar filaments break out of the bundle, and the filament switches from the normal left-handed waveform to right-handed semi-coiled and curly forms[3,9]. This causes the cell to tumble, allowing the trajectory of the cell to change[3,8-11]. During chemotaxis, alternation between running and tumbling modes depends on the presence

of both positive and negative stimuli[12] and result in a “biased random-walk” swimming pattern[13].

1.2 Bacterial Flagellar filament supercoiling

The polymorphic switching of the flagellar filament has been attributed to the N- and C-terminal core domains D0 and D1 of the flagellin subunits (Fig 1.1) [14-17]. The central flagellin region contains the “hyper-variable” outer domains , which have been shown to have non-critical roles in motility for some bacterial species (Fig 1.1) [14,18-20]. Large deletions of up to 100 amino acid residues in the outer domains D2 and D3 in *Salmonella typhimurium* flagellin FliC had no effect on bacterial motility[18,19]. Since the outer domains are on the surface of the flagellar filaments the quaternary structures of bacterial flagella are different across species of bacteria (Fig 1.2). Additionally, many bacterial species have flagellar filaments assembled from flagellins with just the core domains D0 and D1 and no outer domains[14,21,22]. Rather than affecting motility, many outer domains are thought to provide their flagellar filaments with non-conserved species-specific functions such as adhesion[23-25], modulation of host immune responses[26-29], or protease activity[30].

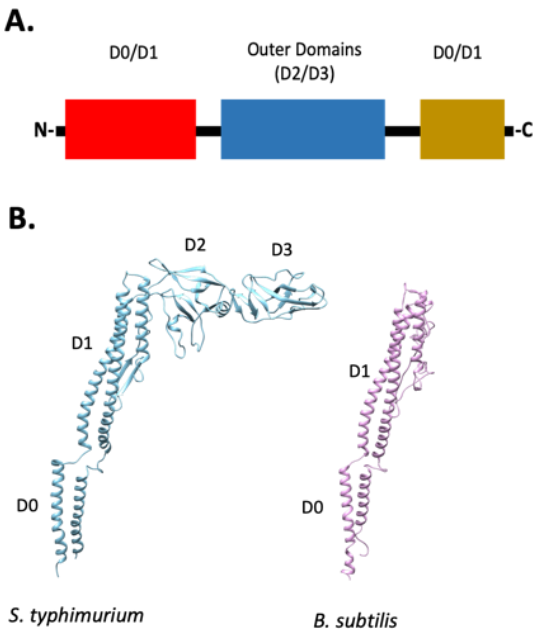

Figure 1.1. Architecture of the bacterial flagellin. (A). Domain organization of the domains of bacterial flagellins. The core domains, D0 and D1 are on the N and C-termini of the flagellin sequence. Outer domains are in the middle of the protein sequence. (**B).** Flagellin subunits from *S. typhimurium* (light blue) and *B. subtilis* (pink).

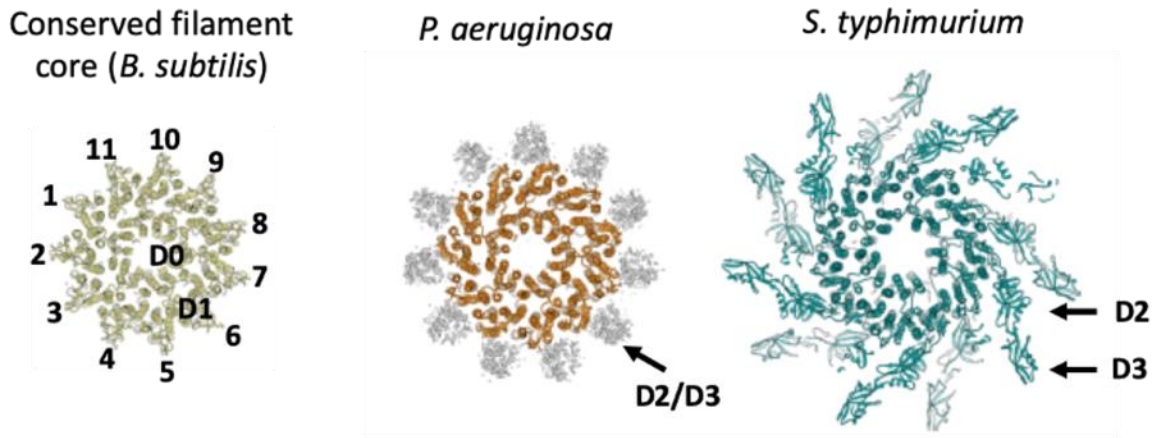

Figure 1.2. Flagellar filament models from various species of bacteria. The core domains (D0 and D1) of each filament model are conserved across bacterial species while the outer domains are not. These images were adapted from Wang et al[14].

The mechanism of flagellar filament supercoiling has been attributed to the existence of the 11 protofilaments in a mixture of two subunit conformations that differ very slightly in length[14,16]. The short protofilaments are located on the inside of the supercoil curve, while the long protofilaments are on the outside. It is this path length difference between the short and long protofilaments that causes the supercoiling. Most previous structural studies analyzed straight mutant flagellar filaments so that helical symmetry could be imposed for the purposes of averaging[14,31-33]. Recently, high resolution cryo-EM structures of supercoiled flagellar filaments were achieved with helical reconstruction, assuming all the protofilaments are in a single state[22,34]. These studies provided valuable information related to flagellar filament structure, but very few insights into the mechanism of polymorphic switching. However, they showed that a high-resolution cryo-EM structure can be generated without using straight mutant flagellar filaments. Two poorly areas of flagellar filament research include understanding the structure and function of flagellar filament outer domains as well as the structure of supercoiled flagellar filaments.

1.3 Archaeal flagellum

Similar to bacteria, archaeal flagellar motility is driven by rotation of the flagellar filament by a motor structure. However, the components of these two systems have no homology[35]. In fact archaeal flagellins have homology with bacterial and archaeal type IV pilins (T4P) and no homology with bacterial flagellins (Fig 1.3). Given that type IV pili do not supercoil and archaeal flagellar filaments do[36], a greater investigation into the mechanism of supercoiling in archaeal flagella will yield insights into the differences between these two filaments.

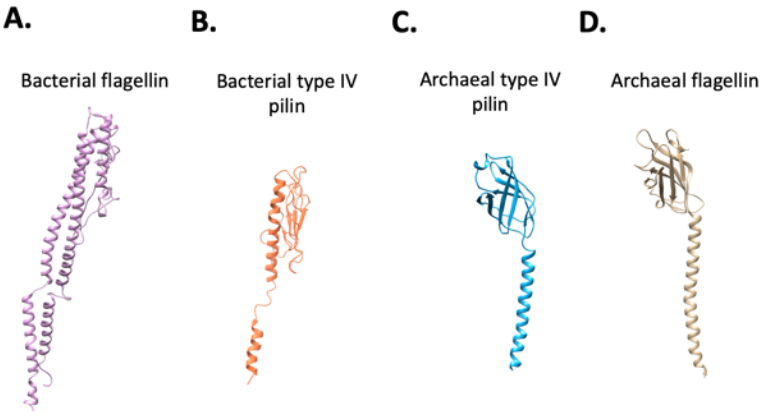

Figure 1.3. Structural comparison of bacterial and archaeal flagellins and type IV pilins. (A). Bacterial flagellin from *B. subtilis* (5WJY). (**B).** Bacterial type IV pilin from *P. aeruginosa* (5VXY). (**C).** Archaeal type IV pilin from *S. islandicus*. (**D).** Archaeal flagellin from *M. hungatei* (5TFY).

Chapter 2: Cryo-EM structure of the *C. jejuni* flagellar filament reveals how ε Proteobacteria escape Toll-like receptor 5 surveillance

This chapter was published in PNAS.

Kreutzberger, M. A. B., Ewing, C., Poly, F., Wang, F. & Egelman, E. H. Atomic structure of the Campylobacter jejuni flagellar filament reveals how epsilon Proteobacteria escaped Toll-like receptor 5 surveillance. *Proc Natl Acad Sci U S A* **117**, 16985-16991.

2.1 Summary

Vertebrates, from zebra fish to humans, have an innate immune recognition of many bacterial flagellins. This involves a conserved eight amino acid epitope in flagellin recognized by the Toll-like receptor 5 (TLR5). Several important human pathogens, such as *Helicobacter pylori* and *Campylobacter jejuni*, have escaped TLR5 activation by mutations in this epitope. When such mutations were introduced into *Salmonella* flagellin, motility was abolished. It was previously argued, using very low resolution cryo-electron microscopy (cryo-EM), that *C. jejuni* accommodated these mutations by forming filaments with 7 protofilaments, rather than the 11 found in other bacteria. We have now determined the atomic structure of the *C. jejuni* G508A flagellar filament from a 3.5 Å resolution cryo-EM reconstruction and show that it has 11 protofilaments. The residues in the *C. jejuni* TLR5 epitope have reduced contacts with the adjacent subunit compared to other bacterial flagellar filament structures. The weakening of the subunit-

subunit interface introduced by the mutations in the TLR5 epitope is compensated for by extensive interactions between the outer domains of the flagellin subunits. In other bacteria, these outer domains can be nearly absent or removed without affecting motility. Furthermore, we provide evidence for the stabilization of these outer domain interactions through glycosylation of key residues. These results explain the essential role of glycosylation in *C. jejuni* motility, and show how the outer domains have evolved to play a role not previously found in other bacteria.

2.2 Introduction

Innate immunity, as opposed to acquired immunity, involves a genetically pre-programmed potential for the recognition of pathogens that exists before an organism ever encounters the threat. TLR5 has been shown to recognize a highly conserved epitope in bacterial flagellin as part of this innate immunity [37], and various experiments have shown that this epitope is crucial for the assembly of bacterial flagellar filaments [38-41]. Since this conserved epitope is recognized by vertebrates from zebrafish to humans, the assumption has been that this epitope in many pathogenic bacteria has not diverged in more than 350 million years, and that TLR5 has evolved to recognize a region that is functionally constrained [39,40]. However, it is apparent that some bacteria, such as the *α* and *ϵ Proteobacteria*, have tolerated sequence changes within this epitope and escape TLR5 activation. Since such important human pathogens as *Helicobacter pylori* and *Campylobacter jejuni* are members of the *ϵ Proteobacteria*, it is of great interest to understand how they evolved to avoid TLR5 activation.

A number of studies have provided a very detailed picture of the interactions of bacterial flagellin with TLR5. Mutational analysis initially mapped the epitope to residues 79-117 and 408-439 of *Salmonella typhimurium* flagellin [41]. A subsequent study showed

that the evasion of TLR5 activation by *H. pylori* and *C. jejuni* was due to amino acid substitutions in the region 89-96. A crystal structure of a fragment of flagellin bound to TLR5 showed the structural basis for this recognition [39]. Bacterial flagellins have two largely coiled-coil domains that are called D0 and D1, and these are highly conserved across all bacteria. These domains assemble to form the core of the flagellar filament. In contrast, there are outer, globular domains, that are highly variable. In *S. typhimurium* these are D2 and D3 [32], while in *Bacillus subtilis* the outer domains do not exist [14]. The TLR5-flagellin crystal structure showed that the main interaction with TLR5 involved three helices in flagellin D1, one of which included the previously described 89-96 epitope. A subsequent crystallographic and mutational analysis extended these results [38] and showed that three residues in *B. subtilis* flagellin formed a “hot spot” that was the largest determinant of this interaction: R89, L93 and E114. These three residues are quite conserved among the bacterial flagellins that activate TLR5, while in *C. jejuni* the corresponding residues are T91, K95 and D115.

During the Dark Ages of cryo-EM, a very low resolution reconstruction of *C. jejuni* flagellar filaments containing the G508A mutation suggested that *C. jejuni* was able to accommodate the mutations in the TLR5 89-96 epitope by having a dramatically different quaternary structure, with seven protofilaments rather than the 11 seen in all other bacterial flagellar filaments that had been studied [42]. The G508A mutation was used as it led to straight flagellar filaments, rather than the supercoiled filaments that exist in the wild type. The use of straight filaments [15] had been crucial to the structural studies of *Salmonella* flagellar filaments [31,32] which have served as a model system for understanding flagellar structure.

The resolution revolution in cryo-EM [43-45] has led to the ability to reach near-atomic resolution almost routinely for many specimens. At less than near-atomic resolution, it is possible for ambiguities in helical symmetry to exist [46], and we now understand that for some specimens ambiguities can exist at even 5 Å resolution. We have reinvestigated the structure of the *C. jejuni* flagellar filaments using a direct electron detector and have generated a cryo-EM reconstruction at 3.5 Å resolution. There are clearly 11 protofilaments in this structure, not seven, and the resolution is sufficient to actually fit glycans into the density for the outer domains. We show that the amino acid changes in the TLR5 epitope would destabilize the filament, and stabilization of the filament occurs by large interactions among the outer domains rather than by simply compensatory mutations at the site on a neighboring subunit facing the TLR5 epitope. We further show the role of glycosylation in the interactions between neighboring outer domains, and thus explain why glycosylation is necessary for filament formation in *C. jejuni* [47]. In contrast, the outer domains in *Salmonella* do not interact with each other [31,32], and mutations have been found lacking most or all of the outer domains that assembled flagellar filaments and were motile [48]. As previously stated, in the *Bacillus subtilis* wt filament the outer domains are almost completely absent [14]. While the outer domains in *H. pylori* have diverged significantly from *C. jejuni*, we show that a similar mechanism of stabilization is likely to occur in *H. pylori*.

2.3 Results

We used *C. jejuni* strain 81-176, where the minor flagellin FlaB was knocked out, and the wt flagellin was replaced by the G508A mutant [42]. The *C. jejuni* G508A flagellar filaments were imaged by cryo-EM (Fig. 2.1A) and were revealed, as expected, to be

straight filaments. Using IHRSR [49], the helical structure of the *C. jejuni* filaments was solved. Analysis of the averaged power spectrum from thousands of images of the flagellar filament revealed several possible symmetries arising from either 7, 9 or 11 protofilaments. Trying each of these, we were able to get a high-resolution structure (Fig 2.1B-E) for the *C. jejuni* flagellar filaments showing that it has 11 protofilaments (Fig 2.1C). The correctly indexed power spectrum is shown in Fig 2.2A. Using Relion [50], we were able to reach a resolution of 3.5 Å as determined by the "gold-standard" map:map Fourier shell correlation (FSC) plot (Fig. 2.2B).

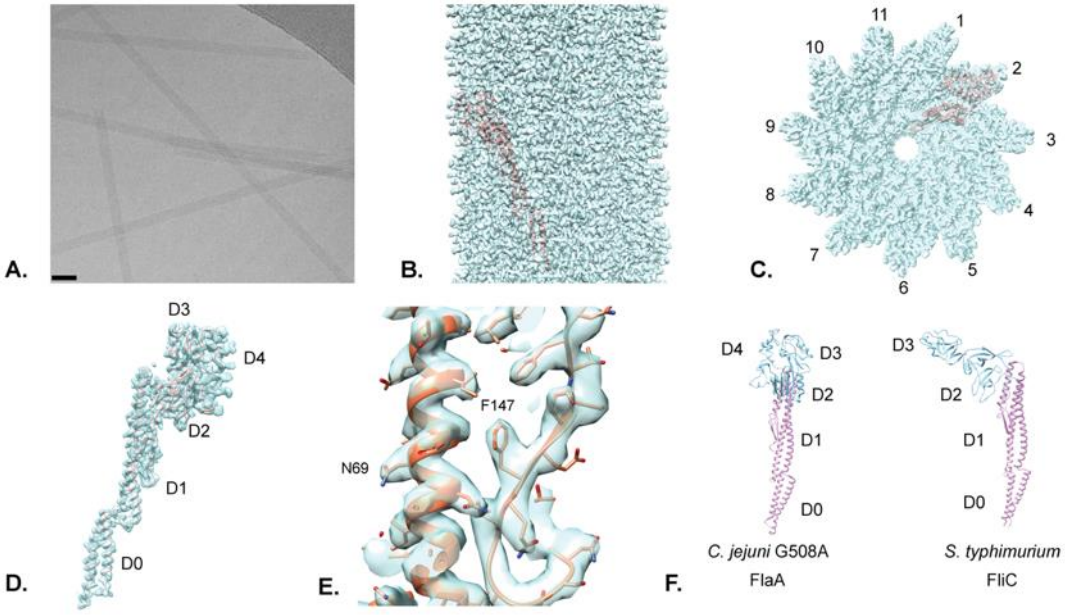

Figure 2.1. Cryo-EM structure of the *C. jejuni* G508A flagellar filament. (A). Cryo-electron micrograph of the flagellar filaments. Scale bar is 40 nm. **(B).** Side view of the 3.5 Å resolution *C. jejuni* G508A density map. **(C).** A top view of the density map shows 11 protofilaments. A model of a single flagellin subunit (red) is fit into the map. **(D).** The model of a single *C. jejuni* G508A flagellin fit into its corresponding electron density. **(E).** A blowup showing the quality of the map around residues 50 to 160. **(F).** Comparison of the *C. jejuni* and *S. typhimurium* flagellins. The inner core domains (D0 and D1) are in pink, and are shown in the same orientation. The outer domains are in blue (D2 and D3) in *S. typhimurium* and green (D2, D3 and D4) in *C. jejuni*.

The filaments have a canonical L-type flagellar symmetry with an axial rise of 4.8 Å and a twist of 65.32°. This means that the protofilaments supercoil in a left-handed

manner, due to their being very slightly more than 11 subunits in every two turns of the right-handed 1-start helix. A model of a single *C. jejuni* flagellin subunit fit into its corresponding electron density is shown in Fig 2.1D. Using the model:map FSC with a threshold of 0.5 a resolution of 3.5 Å was obtained (Fig 2.2C), consistent with the map:map FSC results from Relion. The flagellin subunit, as shown in Fig 2.1D, can be subdivided into five structural domains. These are the two conserved inner core domains D0 and D1, and three domains on the outside of the filaments. These outer domains are referred to as D2, D3, and D4 and are shown in greater detail in Fig 2.3A and 2.3B. A close up of the atomic model for D1 fit to the map is shown in Fig. 2.1E. While domains D0 and D1 of the *C. jejuni* flagellin look the same as those of all other existing bacterial flagellin structures, the outer domains are ~ 80 amino acids larger than those of the next largest solved flagellin structure, from *Salmonella typhimurium* [32]. A comparison of the *C. jejuni* and *S. typhimurium* flagellins (Fig 2.1F) shows the conservation of D0 and D1 (pink) and the absence of detectable homology in the outer domains (blue). Thus, domains D2 and D3 in *C. jejuni* have no homology with domains D2 and D3 in *S. typhimurium.*

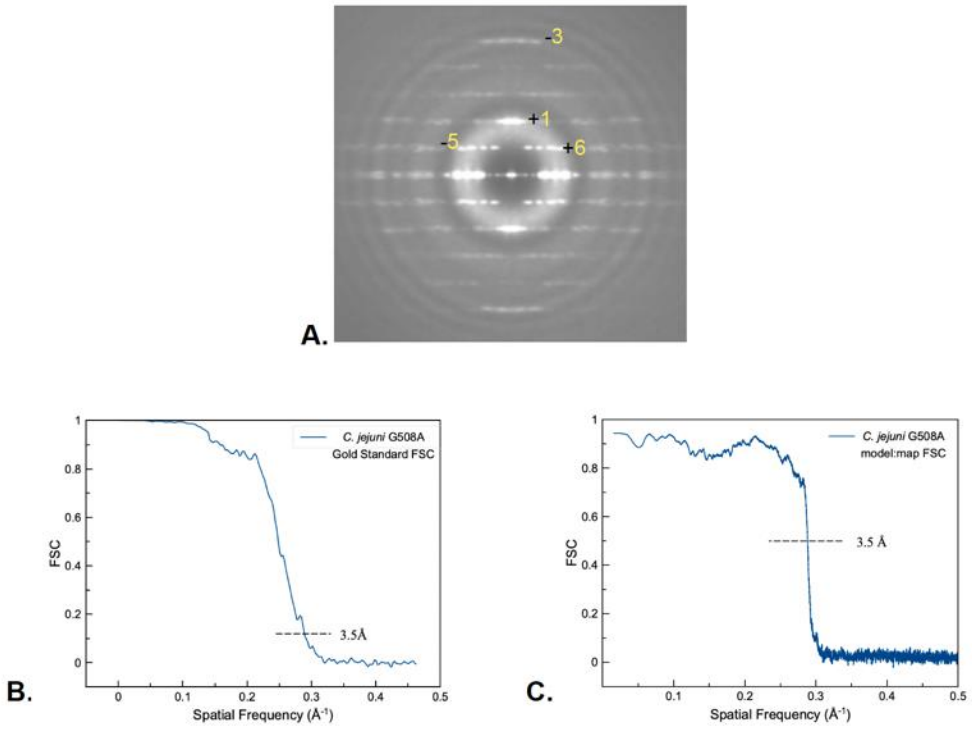

Figure 2.2 (A). The indexed power spectrum of the *C. jejuni* G508A flagellar filament. **(B).** FSC plot using the "gold-standard" map:map method. **(C).** FSC plot using the model:map method with a cutoff of 0.5.

The Dali server [51] was used to search for structural homologs of each individual *C. jejuni* FlaA outer domain. Domains D2 and D3 are in fact structural homologs of each other (Fig 2.3B) and searching for homologs to D2 and D3 yielded the same top hits but in different order. Interestingly, all of these top hit homologs are domains from various flagella-related proteins (Fig 2.3 C). Two of the top hits for D2 of *C. jejuni* were the outer domain from the *P. aeruginosa* A-type flagellin FliC [52] and filament cap protein, FliD, from *Bdellovibrio bacteriovorus* [53]. Some of *C. jejuni* FlaA D3's top hits were FliD from *H. pylori* [54] and D4 from the *C. jejuni* flagellar hook protein FlgE [55]. The top Dali server

hits for *C. jejuni* FlaA D2 and D3 are shown in SI Appendix, Table 1, along with the corresponding Z-score.

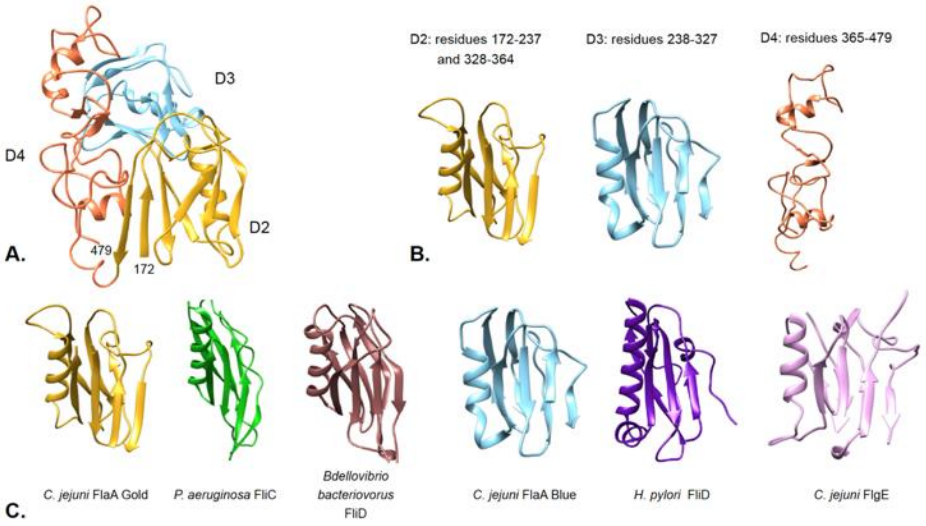

Figure 2.3. Domain duplication and structural homologs of the *C. jejuni* outer domain D2. **(A).** The outer domains D2 (gold), D3 (blue) and D4 (orange) of the *C. jejuni* flagellin). **(B).** The three outer domains shown separately. **(C).** Structural homology between the domain duplicated D2/D3 and domains from various flagella proteins.

Table 2.1: Structural homologs of *C. jejuni* FlaA D2 and D3.

***C. jejuni* FlaA D2 Homologues**	***C. jejuni* FlaA D3 Homologues**
P. aeruginosa FliC: 4NX9 z=7.5	*H. pylori* FliD: 6IWY z=6.7
E. coli FliD: 5H5V z=7.3	*E. coli* FliD: 5H5V z=5.9
B. bacteriovorus FliD: 6KTY z=7.3	*P. aeruginosa* FliC: 4NX9 z=5.4
H. pylori FliD: 6IWY z=6.0	*B. bacteriovorus* FliD: 6KTY z=5.4
C. jejuni FlgE: 5JXL z=5.0	*C. jejuni* FlgE: 5JXL z=5.1
C. jejuni FlgK: 5XBJ z=4.8	*C. jejuni* FlgK: 5XBJ z=4.1

During the flagellin model building process large extra densities were observed in the map contiguous with serine and threonine residues. These densities were all quite

uniform in size and are indicative of O-linked glycosylation. FlaA from *C. jejuni* strain 81-176 has been shown to be glycosylated entirely with pseudaminic acid (Pse) and its derivatives [56]. Fig. 2.4A shows that these glycans bound to their serine or threonine residues provide an excellent fit to the map. Due to the inability to distinguish between the slightly different-in-size Pse derivatives at ~3.5 Å resolution, only the basic Pse glycan (Fig 2.4B) was modeled at each site. Density corresponding to glycosylation was found for 17 different residues on the outer domain. These residues are listed in Fig 2.5A and shown in the sequence in Fig. 2.5B. The majority of these residues (15 of the 17) are in D4 (Fig. 2.4C). As might be expected, all of the glycosylation is present on the surface of the flagellar filament (Fig 2.4D).

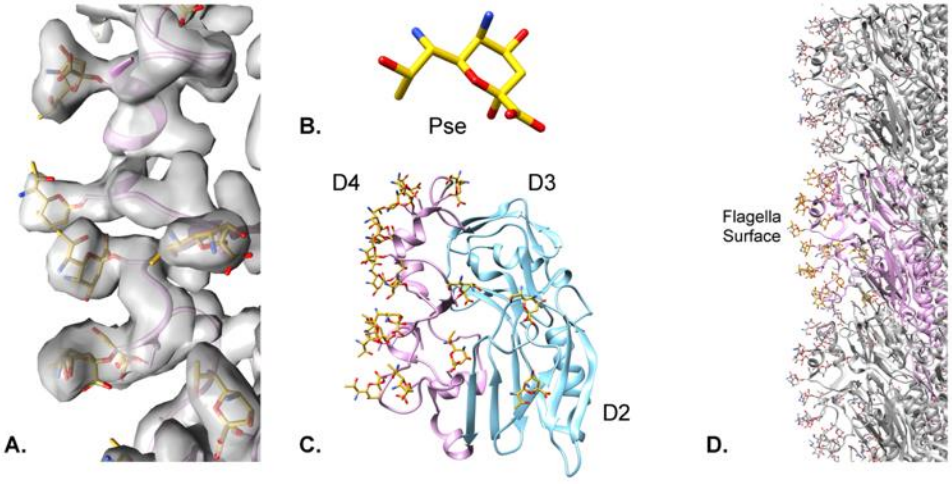

Figure 2.4. Extensive glycosylation of the *C. jejuni* outer domains with pseudaminic acid glycans. (A). The density map (transparent grey) with the fit of Pse glycans (gold). **(B).** The atomic model of a single pseudaminic acid glycan (Pse). **(C).** Flagellin outer domains are shown in pink (D4) and blue (D2, D3), with its glycan moieties shown in gold. **(D).** The surface exposed region of a single 11-start protofilament.

A.

Amino Acid residue
S343
S348
T394
S398
S401
S405
S412
S418
S426
S430
S437
S441
S443
S449
S455 *
S461*
T469

B.

1 MGFRINTNVAALNAKANSDLNAKSLDASLSRLSSGLRINSAAD 43
44 DASGMAIADSLRSQANTLGQAISNGNDALGILQTADKAMDEQL 86
87 KILDTIKTKATQAAQDGQSLKTRTMLQADINKLMEELDNIANT 129
130 TSFNGKQLLSGNFTNQEFQIGASSNQTVKATIGATQSSKIGVT 172
173 RFETGAQSFTSGVVGLTIKNYNGIEDFKFDNVVISTSVGTGLG 215
216 ALAEEINKSADKTGVRATYDVKTTGVYAIKEGTTSQEFAINGV 258
259 TIGKIEYKDGDGNGSLISAINAVKDTTGVQASKDENGKLVLTS 301
302 ADGRGIKITGDIGVGSGILANQKENYGRLSLVKNDGRDINISG 344
345 TNLSAIGMGTTDMISQSSVSLRESKGQISATNADAMGFNSYKG 387
388 GGKFVFTQNVSSISAFMSAQGSGFSRGSGFSVGSGKNLSVGLS 430
431 QGIQIISSAASMSNTYVVSAGSGFSSGSGNSQFAALKTTAANT 473
474 TDETAGVTTLKGAMAVMDIAETAITNLDQIRADIGSIQNQVTS 516
517 TINNITVTQVNVKAAESQIRDVDFASESANYSKANILAQSGSY 559
560 AMAQANSSQQNVLRLLQ 576

Figure 2.5: Glycosylation sites detected in the *C. jejuni* FlaA Flagellin. (A). List of glycosylated residues seen in the *C. jejuni* FlaA density map. **(B).** FlaA sequence showing each of the observed glycosylated residues in red

It was previously proposed [42] that changes in the TLR5 epitope that would destabilize the flagellar filament led to a different packing with only seven protofilaments. Since we now know that is wrong, the key question is: how does the *C. jejuni* accommodate mutations in D1 which are known to abrogate filament formation in *Salmonella* FliC? We considered amino acids 88-98 in *C. jejuni* which include four more amino acids than those focused on by Andersen-Nissen *et al*.[40]. First, we examined how these residues interacted with neighboring flagellin subunits. Fig. 2.6A shows all of the flagellin subunits (S+5, S+6, S+11, S+16) that make contacts with a reference subunit (S0, in pink). The most prominent interactions the *Campylobacter* TLR5 epitope makes are with the S+11 subunit, in blue. Fig. 2.6B shows a comparison of the TLR5 epitope interactions along the 11-start helix between *C. jejuni* and *Bacillus subtilis*, which has a similar TLR5 epitope to *S. typhimurium*. In the *C. jejuni* model amino acids that are significantly different from the ones conserved across most bacteria are colored gold. In

B. subtilis these corresponding residues are colored grey. Fig. 2.6C shows a sequence alignment of the TLR5 epitope in *C. jejuni* and *H. pylori* against several flagellins whose filament structure has been solved. In *C. jejuni* D90, T91, T94, and T97 are at the interface of the S0 subunit with S+11. In *B. subtilis* the corresponding residues are Q88, R89, E92, and V95.

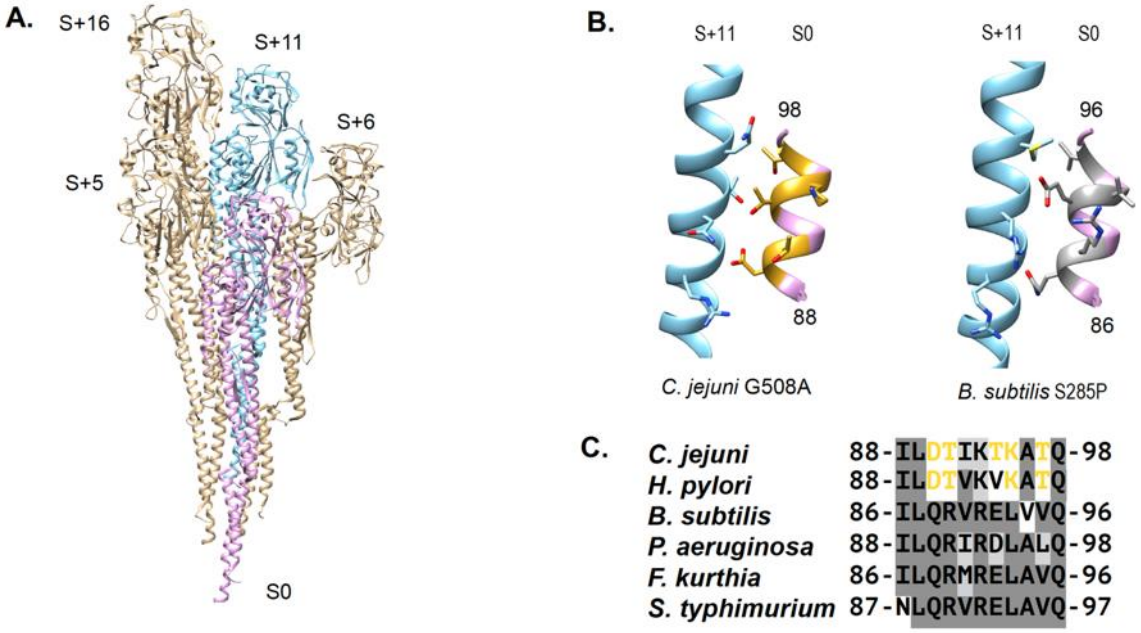

Figure 2.6. Interactions of the TLR5 recognition site (*C. jejuni* residues 88-98) with the next 11-start flagellin subunit (S+11). (A). Each subunit in the filament contacts four other subunits. The subunits that contact a single flagellin (S0) are S+5, S+6, S+11 and S+16. **(B).** Comparison of the contacts between the TLR5 site (pink) and the next 11-start subunit, S+11 (blue), in *C. jejuni* (left) and *B. subtilis* (right). Residues that are not identical between the two TLR5 sites are shown in gold. **(C).** Alignment of the TLR5 sequences from *C. jejuni* flagellin with those of five flagellins from different species whose high-resolution structures have been determined by cryo-EM.

Using PDBePISA[57] we calculated the interfacial surface area between the TLR5 epitope of the S0 subunit and the entire S+11 subunit (S0-TLR5/S+11) of several flagellins from different bacteria. The interfacial surface area between the entire S0 and S+11 subunits were also calculated (Table 2.2). The S0-TLR5/S+11 interface is much

smaller in *C. jejuni* than in the other flagellar filaments. Strikingly, the interface between the entire *C. jejuni* S0/S+11 subunit interface was greater than in any other flagellar filament. These results suggest that in *C. jejuni* the weakened interactions between the TLR5 interface and the adjacent subunit are compensated for by interactions between the subunits not found in the other structures. Most of this arises from interactions of the outer domains.

Table 2.2 Comparison of the interactions between the TLR5 recognition sequence and the next 11-start flagellin across different flagellar filament structures, with the PDB ID given.

Bacteria	**TLR5 Site 11-start interactions Interfacial Surface Area (Å²)**	**11-start Full Model Interfacial Area Å²**
C. jejuni G508A 6X80	300.6 (AA 88-98)	2869.9
B. subtilis S285P 5WJY	410.8 (AA 86-96)	1881.5
B. subtilis N226Y 5WJT	381 (AA 86-96)	1858.9
P. aeruginosa A443V 5WK5	369.5 (AA 88-98)	1845
*F.kurthia*6T17	371.4 (AA 86-96)	1964.9
S. typhymurium FljB A461V 6JY0	385.8 (AA 87-97)	1787.3

The outer domains of the *C. jejuni* flagellin strongly interact with the outer domains of adjacent subunits on the same protofilament. Fig. 2.7A shows the interactions of two adjacent flagellin outer domains along the same 11-start

protofilament. Previously, it has been shown that when any one of the glycosylation sites S426, S455, or S461 are mutated to alanine, *C. jejuni* 81-176 is unable to assemble full length flagellar filaments and exhibits greatly reduced motility[47]. In our filament model, two of these glycosylated serines (S455 and S461) are directly at the 11-start interface between the D4 domains of both subunits, and are indicated (Fig. 2.7A). Extensive contacts are made between D3 of S-11 and D1 and D2 of S0. These are highlighted by the red box and shown in greater detail in Fig 3.7B. When FlaA has been deleteriously mutated or deleted, *C. jejuni* has been known to form very short less-stable filaments using only FlaB, and the bacteria exhibit greatly reduced motility[58]. There is 95% sequence identity between the two flagellins, and most sequence differences between FlaA and FlaB occur in D3 (Fig. 2.7C), however there are some in D0 and D1 as well. The gold-colored residues in Fig. 2.7B correspond to residues in D3 at the 11-start outer domain interface that are significantly different between FlaA and FlaB. Thus, the instability of the FlaB filaments may be directly due to the loss of interactions involving D3.

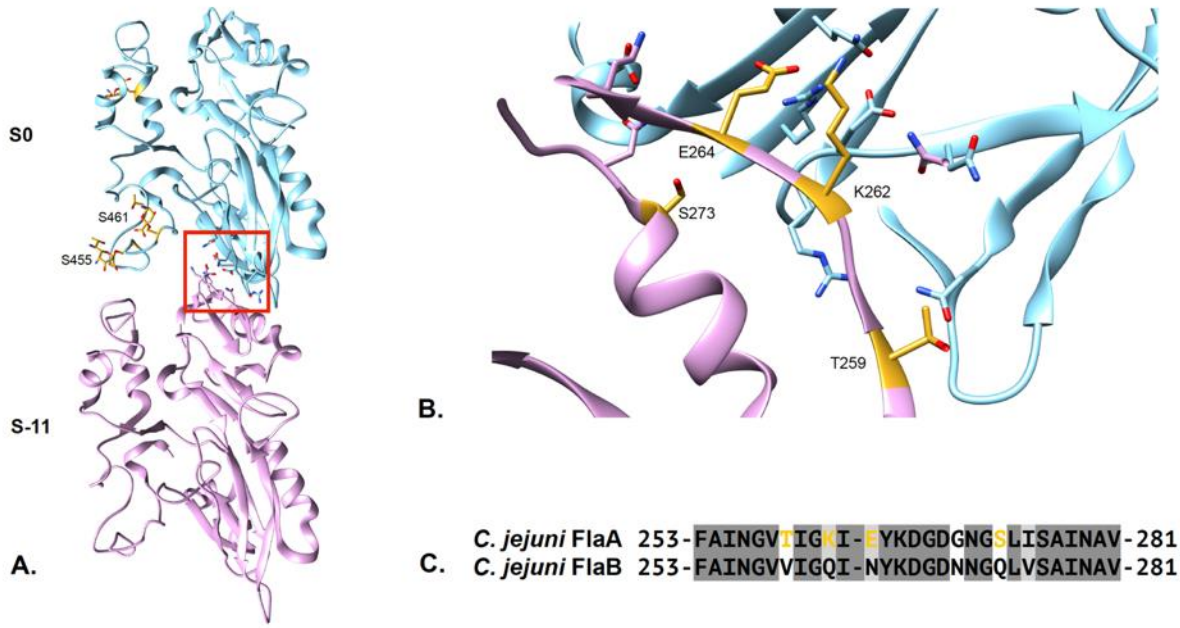

Figure 2.7. Outer domain interactions between adjacent subunits along the same 11-start protofilament. (A). Two neighboring subunits with the glycans attached to S455 and S461 in domain D3 shown in gold. The 11-start interaction site between D2 in one subunit and D3 in an adjacent subunit is boxed in red. **(B).** Close up of this D2/D3 11-start interface. Residues at the interface between the two subunits in the FlaA filament that are significantly different in the *C. jejuni* FlaB sequence are in gold.

2.4 Discussion

The structure of the *Campylobacter jejuni* flagellar filaments provides clear evidence for how *ϵ Proteobacteria* evolved to evade detection by TLR5. The amino acids in the *C. jejuni* TLR5 epitope display a dramatically weakened ability to interact with the adjacent subunit on the same protofilament. However, the entire *C. jejuni* flagellin has much greater interactions with its neighboring 11-start subunit. These are the result of outer domain interactions among the adjacent subunits that are entirely absent in all other existing flagellar filament structures. The two outer domains (D2 and D3) that make these contacts are actually structural homologs of not only each other,

but domains found in various flagellar structural proteins across different species of bacteria (Fig 2.3 and 2.8).

Two of the structural homologs are the flagellar hook protein FlgE domain D4 and the hook-filament junction protein FlgK D4, both from *C. jejuni.* Fig. 3.8 shows how both FlaA D3 and FlgE D4 have conserved interactions along the 11-start interface and a few conserved residues. The homology between these domains is likely due to FlaA, FlgE, and FlgK sharing a common ancestor protein which functioned as the protein subunit for the filament before the specialized hook evolved. The conservation of a few key residues in the 11-start interface in both the *C. jejuni* FlaA D3 and FlgE D4 suggests that there are similar stabilizing interactions amongst the outer domains of the hook, hook junction protein, and flagellar filament. The role of homologs playing a role at filament ends and in filament nucleation has been seen for both actin and tubulin. In the case of actin, we know that Arps (actin related proteins), which are structural homologs of actin, such as Arp2/3, play an important role in actin filament nucleation and branching[59]. In the case of tubulin, gamma-tubulin, a homolog of alpha- and beta-tubulin, plays a key role in nucleating microtubules[60].

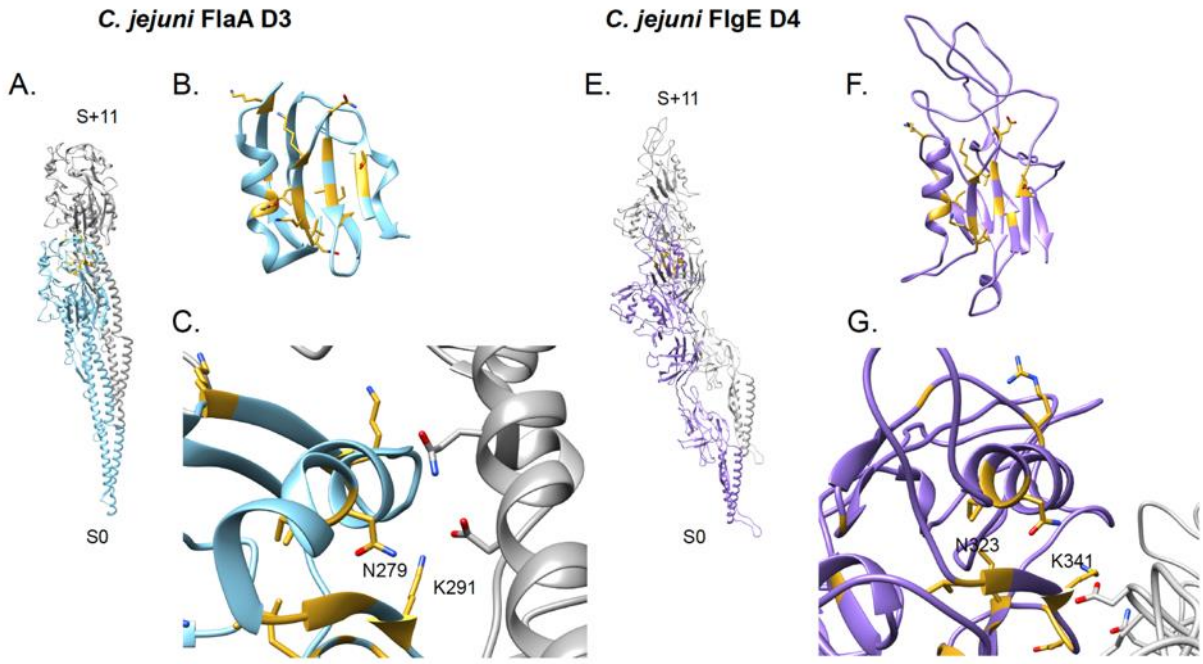

Figure 2.8. Comparison of *C. jejuni* FlaA D3 (A-C) and FlgE D4 (D-F) along their 11-start interfaces. **(A).** Two flagellins along the same protofilament are shown. **(B).** Residues that are conserved or similar in *C. jejuni* FlaA to *C. jejuni* FlgE are colored in gold. **(C).** A close up of the FlaA D3 11-start interface shows two key conserved residues (N279 and K291) between FlaA and FlgE. **(D-F).** As with A-C, two flagellar hook proteins along the same protofilament are shown. **(E).** Residues that are conserved or similar in *C. jejuni* FlgE to *C. jejuni* FlaA are colored in gold. **(F).** Two conserved residues along the 11-start interface in FlgE domain D4 are N323 and K341.

The glycosylation of Ser426, Ser455, and Ser 461 has been previously shown to be critical for filament formation and motility[47] The glycans attached to Ser455 and Ser461 in the 11-start interface are part of the extensive interactions of the flagellin outer domains. Ser426 is not at the interface with any neighboring flagellins and may play a role in the stabilization of the outer domains. Many of the other serine/threonine residues that are glycosylated in our structure have been shown to be critically important for autoagglutination (AAG), which is correlated with *C. jejuni* virulence[47,61]. It was previously observed that Ser207 in D2, Ser409 in D4, and Thr482 in D1 were glycosylated. We detect no significant glycosylation in our structure at these sites. This

is likely due to either disorder of the glycans, or a heterogeneity in glycosylation, such that the glycans are attached at these residues to only some subunits. For the glycosylated residues that we do see clearly (Fig. 2.4A), the glycans must be both highly ordered and stoichiometrically attached to these residues in almost every subunit. Most of the glycosylation was not seen in the cryo-EM map of an extensively glycosylated archaeal Type 4 pilus[62]. The glycosylation we see at Thr469 has not previously been reported.

All ϵ *Proteobacteria* possess an alanine at A493, the position corresponding to I411 in *S. typhimurium.* When I411 is mutated to alanine in *Salmonella typhimurium*, TLR5 activation is reduced and motility is completely lost[41]. Follow-up studies showed that when *Salmonella typhimurium* FliC I411A amino acids Arg58 and Gly59 were mutated to the corresponding *H. pylori* residues (Ser59 and Ser 60), motility and filament formation were restored but the TLR5 activation was still greatly reduced[40]. These results were mapped to the structure of the *S. typhimurium* flagellar filament to show that these three residues were predicted to be at the interface between two subunits along the same protofilament. We show (Fig 2.9A) that the predictions about the *H. pylori* and *C. jejuni* flagellin interactions were correct and that A493 in *C. jejuni* would be in the same position as *S. typhimurium* I411. *C. jejuni* has amino acids A59 and T60 instead of the two serine side chains in *H. pylori* at these positions.

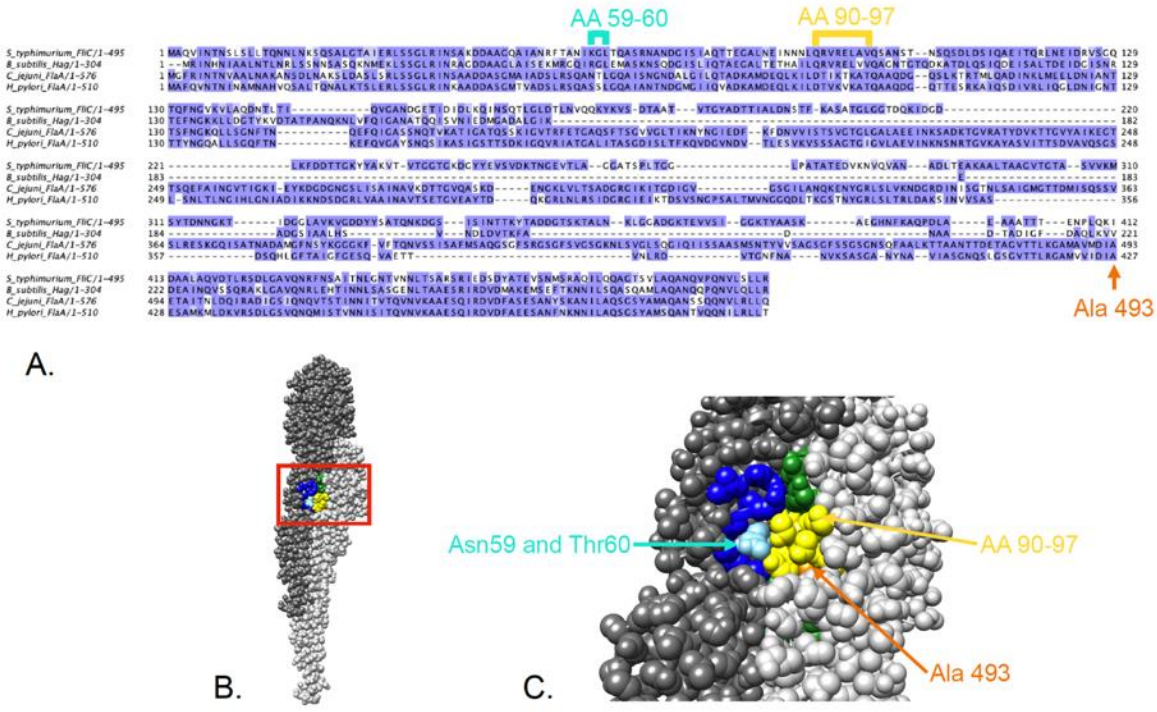

Figure 2.9 Sequence changes in the residues interfacing the TLR5 epitope previously suggested to allow for motility in *H. pylori* and *C. jejuni*. (A). Multiple sequence alignment of several key bacterial flagellins. **(B).** Space filling representation of two flagellins along the 11-start protofilament. The S0 flagellin is shown in light gray while the S+11 is in dark grey. **(C).** A close up of the interface site between the two flagellins confirming the previously reported predictions about the *H. pylori* and *C. jejuni* flagellins.

While we have only determined the structure of the *C. jejuni* flagellar filament, our results give new insights into the structure of the *H. pylori* filament, and it is certainly likely that extensive outer domain interactions in *H. pylori* also stabilize that flagellar filament. Fig 2.10A shows a comparison between the observed secondary structure of the outer domains of *C. jejuni* FlaA and the predicted [63] secondary structure of *H. pylori* FlaB. The predicted *H. pylori* structure closely resembles the secondary structure observed, except for a few places, most notably at the N-terminus of the outer domains. As a control, the predicted *C. jejuni* secondary structure agrees quite well with the observed secondary structure (Fig. 2.10B). Strikingly, *H. pylori* appears to lack most of

D4, but the expectation would be that the extensive interactions observed in *C. jejuni* between D2 in one subunit and D3 in an adjacent subunit would be preserved in *H. pylori*.

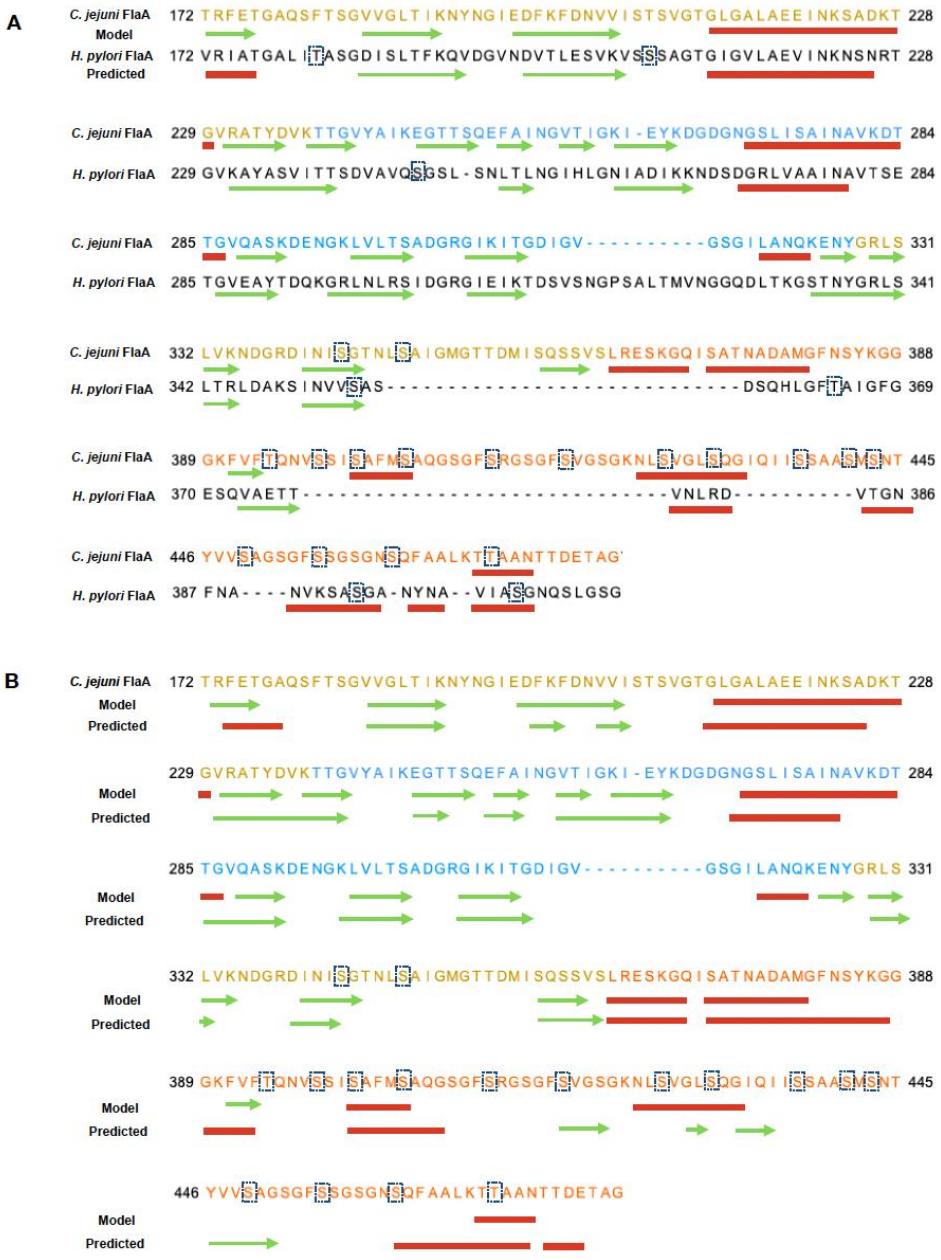

Figure 2.10. Sequence alignment and secondary structure of the *C. jejuni* and *H. pylori* FlaA flagellins. (A). Sequence alignment showing the observed secondary structure in *C. jejuni* FlaA and the predicted *H. pylori* FlaA secondary structure. **(B).** A comparison between the observed and predicted secondary structure of *C. jejuni* G508A.

Our results re-emphasize two points. First, at low-resolution, it may never be possible to determine the correct helical symmetry [46] for a filament or tube. Second, although the bacterial flagellum has been a model for Creationists of "irreducible complexity" showing that the flagellum could never have evolved [64], it actually serves as a beautiful system for studying evolution [65]. In response to the innate immune surveillance that vertebrates developed to recognize pathogenic bacteria, ***ϵ*** *Proteobacteria* such as *Campylobacter jejuni* and *Helicobacter pylori* compensated for mutations in the TLR5 recognition site that destabilize the flagellar filament by developing outer domain interactions that stabilize these filaments. We expect that future studies of bacterial flagella will further elucidate the evolutionary divergence of extant flagellar systems.

2.5 Methods

Preparation of *C. jejuni* G508A Flagellar Filaments

C. jejuni cells were obtained from a MH broth culture incubated at 37° C in a microaerobic atmosphere (5% O_2, 10% CO_2, 85% N_2). The culture was spun and the pellet was resuspended in a low salt buffer (20 mM Tris, 50 mM sodium chloride, pH 7.4). The bacterial suspension was subjected to bead beating, 6 x 30 seconds. Cell debris were collected by centrifugation, resuspended in the low salt buffer, and centrifuged at 45,000 rpm for 1 hour. The pellet was washed in the low salt buffer and centrifuged at 45,000 rpm for 1 additional hour. After the spin was complete, the pellet was resuspended in the low salt buffer and the sample was stored at 4°C until processing.

CryoEM Imaging

The sample was imaged at the Molecular Electron Microscopy Core (MEMC) at the University of Virginia on a Titan Krios equipped with a K3 direct electron detector with a nominal magnification of 81,000x (1.08 Å/pixel). Movies of each area imaged contained 40 frames with a total dose of 51 ($e^-/Å^2$).

Image Processing

Motion correction on the images was done using Motioncorr2 [66]. Motion correction was done using both the full frames and then frames 2-17 (~20 $e^-/Å^2$) using dose-weighting. CTFfind3 [67] was used to calculate the contrast transfer function for each image and images with poor CTFs were discarded. Flagellar filaments were boxed using the e2helixboxer program from the Eman2 software package [68].

Helical Reconstruction

Initial helical reconstruction on CTF-multiplied dose-weighted images was done using the Iterative Helical Real Space Reconstruction (IHRSR) implemented in Spider [49]. The power spectra of hundreds of non-overlapping images of flagellar filaments were calculated and from this an averaged power spectrum was created. Indexing this power spectrum revealed the possible helical symmetries of the *C. jejuni* G508A flagellar filaments [49,69]. The most likely symmetries were ones with 7, 9, and 11 protofilaments. Using IHRSR we tried each of the possible symmetries and the canonical flagellar symmetry with 11 protofilaments was determined to be correct given that we were able to clearly see the pitch and hand of the α-helices and some bulky side chains visible.

The 11-protofilament volume from IHRSR in Spider was used as the input volume for Refine3d in Relion[50,70]. With ~120,000 images (we started with 280,000

before Class2d) and Refine3d followed by CTFRefine, Particle Polishing, and Postprocessing [70] we were able to reach a resolution of 3.5 Å using the “gold-standard” map:map FSC.

Model Building

Using Chimera [71], the L-type flagellin from *B. subtilis* S285P, PDB:5WJY, [14] was fit into the D0/D1 region of the *C. jejuni* G508A density map. Using Coot [72], the amino acids in D0/D1 of the starting model were manually replaced with the corresponding ones from *C. jejuni* G508A and fit to the map. The outer domain region was then built manually, starting with a C_α trace. The structure of the Pse glycan was generated from PubChem CID 101137651. A single *C. jejuni* flagellin and its corresponding map density served as inputs into RosettaCM [73] which rebuilt the model with better geometries and secondary structure. The side chain placement and secondary structure of the model were further refined in Coot and Phenix [74]. Finally, a full filament model was generated with Rosetta scripts [75] and real-space refined in Phenix. The final model refinement statistics (Table 1) were validated with MolProbity [76].

Sequence alignments

All multiple sequence alignments in this study were done in Jalview 2 [77] using the Clustal Omega alignment [78]. Amino acids in a particular flagellin that were identical to the majority of the amino acids at the same position in other flagellins were shaded in dark gray. Similar amino acids were shaded in light gray. Amino acids with no similarity were not shaded. The alignment in SI Appendix, Fig S4A was done according to a Blosum62 scoring matrix.

Structural analysis of *C. jejuni* G508A

Structural homologs to the *C. jejuni* FlaA outer domains D2, D3, and D4 were identified with the DALI server [51,79]. All secondary structure predictions were made with JPred [63]. Interfacial analyses were made using the PDBePISA server [57]. For interfacial analysis at least one structure from each of the various bacteria whose flagellar filaments have been studied was chosen. These were *Salmonella typhimurium* [80]*, Firmicute kurthia* [22], *Pseudomonas aeruginosa* [14], and *Bacillus subtilis* [14].

2.6 Acknowledgements

Cheryl Ewing and Frederic Poly at the Naval Medical Research Center prepared the *C. jejuni* flagellar filament samples.

Chapter 3. Flagellin outer domain dimerization modulates motility in pathogenic and soil bacteria from viscous environments

This chapter was published in Nature Communications

Kreutzberger, M. A. B. *et al.* Flagellin outer domain dimerization modulates motility in pathogenic and soil bacteria from viscous environments. *Nat Commun* **13**, 1422, doi:10.1038/s41467-022-29069-y (2022).

3.1 Summary

Flagellar filaments function as the propellers of the bacterial flagellum and their supercoiling is key to motility. The outer domains on the surface of the filament are non-critical for motility in many bacteria and their structures and functions are not conserved. Here, we show the atomic cryo-electron microscopy structures for flagellar filaments from enterohemorrhagic *E. coli* O157:H7, enteropathogenic *E. coli* O127:H6, *Achromobacter*, and *Sinorhizobium meliloti*, where the outer domains dimerize or tetramerize to form either a sheath or a screw-like surface. These dimers are formed by 180° rotations of half of the outer domains. The outer domain sheath (ODS) plays a role in bacterial motility by stabilizing an intermediate waveform and prolonging the tumbling of *E. coli* cells. Bacteria with these ODS and screw-like flagellar filaments are commonly found in soil and human intestinal environments of relatively high viscosity suggesting a

role for the dimerization in these environments.

3.2 Introduction

Bacteria use their flagella to swim towards or away from various environmental signals[81]. The flagellum is divided into several parts: the motor (basal body), the rod, the hook, and the filament[2,82]. The rotating flagellar filament primarily functions as a propeller, acting as an Archimedean screw and supercoiling of the filament is essential as rotations of a straight filament do not generate thrust.

Bacterial species from the *Escherichia* and *Salmonella* genera have peritrichous flagella[7]. During running mode, the flagellar motors at the base of the flagellum are rotating counterclockwise, which orients the filaments in a bundle toward one end of the cell[3,8]. When the direction of rotation of one or several flagella changes to clockwise, the affected flagellar filaments break out of the bundle, and the filament switches from the normal left-handed waveform to right-handed semi-coiled and curly forms[3,9]. This causes the cell to tumble, allowing the trajectory of the cell to change[3,8-11]. During chemotaxis, alternation between running and tumbling modes depends on the presence of both positive and negative stimuli[12] and result in a “biased random-walk” swimming pattern[13].

The mechanism of flagellar filament supercoiling has been attributed to the existence of the 11 protofilaments in a mixture of two subunit conformations that differ very slightly in length[14,16]. The short protofilaments are located on the inside of the supercoil curve, while the long protofilaments are on the outside. It is this path length difference between the short and long protofilaments that causes the supercoiling. Most previous structural studies analyzed straight mutant flagellar filaments so that helical

symmetry could be imposed for the purposes of averaging[14,31-33]. Recently, high resolution cryo-EM structures of supercoiled flagellar filaments were achieved with helical reconstruction, assuming all the protofilaments are in a single state[22,34]. These studies provided valuable information related to flagellar filament structure, but very few insights into the mechanism of polymorphic switching. However, they showed that a high-resolution cryo-EM structure can be generated without using straight mutant flagellar filaments.

The polymorphic switching of the flagellar filament has been attributed to the N- and C-terminal core domains D0 and D1 of flagellin[14-17]. The central flagellin region contains the “hyper-variable” outer domains, which have been shown to have non-critical roles in motility for some bacterial species[14,18-20]. Large deletions of up to 100 amino acid residues in the outer domains D2 and D3 in *Salmonella typhimurium* flagellin FliC had no effect on bacterial motility[18,19]. Additionally, many bacterial species have flagellar filaments assembled from flagellins with just the core domains D0 and D1 and no outer domains[14,21,22]. Rather than affecting motility, many outer domains are thought to provide their flagellar filaments with non-conserved species-specific functions such as adhesion[23-25], modulation of host immune responses[26-29], or protease activity[30].

The flagellar filaments of certain soil bacteria such as *Sinorhizobium meliloti* have been called “complex” due to the pairing of outer domain subunits to produce a criss-cross pattern on the surface of the filament[83-85]. This pairing was regarded as a perturbation of the underlying symmetry[85], modeled in part on the work done with the Dahlemense strain of tobacco mosaic virus (TMV)[86]. This model proposed that pairing of subunits across the helical groove, involving displacements of ~ 3 Å, served to break

the helical symmetry. In TMV, subunits are rigidly locked into a helical lattice and major rearrangements of subunits would not be possible. We discovered the existence of non-canonical flagellar filaments in *Achromobacter* sp. MFA1 R4[87,88], a soil bacterium that was identified as cell culture contaminant. These flagellar filaments have a sheath surrounding the core of the filament similar to those previously described for some H-serotypes of *E. coli*[89,90].

Here, we show the high resolution cryo-EM structures of the *S. meliloti,* enterohemorrhagic *E. coli* O157:H7 (EHEC O157:H7)*, Achromobacter* and enteropathogenic *E. coli* O127:H6 (EPEC O127:H6) flagellar filaments. The three former structures would have been classified as “helically perturbed” while the latter as “non-helically perturbed”[91,92]. In the non-helical perturbation model, subunit pairing introduced a seam causing a discontinuity in the helical lattice[91], much as such a seam or discontinuity exists in microtubules[93]**.** In the present study, we show that the outer domains in these flagellar filaments have considerable freedom. In physics and mathematics, perturbations are regarded as small deviations, such as the influence of the gravitational field of the sun on the orbit of the moon around the earth[94]. Instead of the previously predicted small perturbations, half of the outer domain population rotated by 180° to form symmetrical dimers or tetramers with other outer domains. These outer domain dimers form either a screw-like surface surrounding the filament core domains (D0 and D1) as in *S. meliloti*, or an outer domain sheath (ODS) around the core as in EHEC O157:H7, *Achromobacter* sp*,* and EPEC O127:H6.

The ODS surrounding the EHEC H7 and *Achromobacter* flagellar filaments produces an intermediate waveform that is not adopted by the “canonical” *S.*

typhimurium and *E. coli* K-12 H48 flagellar filaments. In addition, we provide evidence that the ODS of the EHEC H7 flagellar filament prolongs the average time *E. coli* cells spent in the tumbling mode and suggest that this behavior is due to an additional step in flagellar polymorphism during tumbling created by the intermediate waveform. We hypothesize that extended tumbling mediated by ODS flagellar filaments offers an advantage to both intestinal pathogens and soil bacteria by allowing for better reorientation of cells in their respective environments.

3.3 Results

Cryo- electron microscopy structures of screw-like and ODS flagellar filaments

Cryo electron microscopy (cryo-EM) was used to determine the structures of flagellar filaments from *Sinorhizobium meliloti*, enterohemorrhagic *E. coli* O157:H7 (EHEC O157:H7), *Achromobacter* , and enteropathogenic *E. coli* O127:H6 (EPEC O127:H6). Differences in filament surface structures were observable in the cryo-electron micrographs by the naked eye (Fig. 3.1A-D).

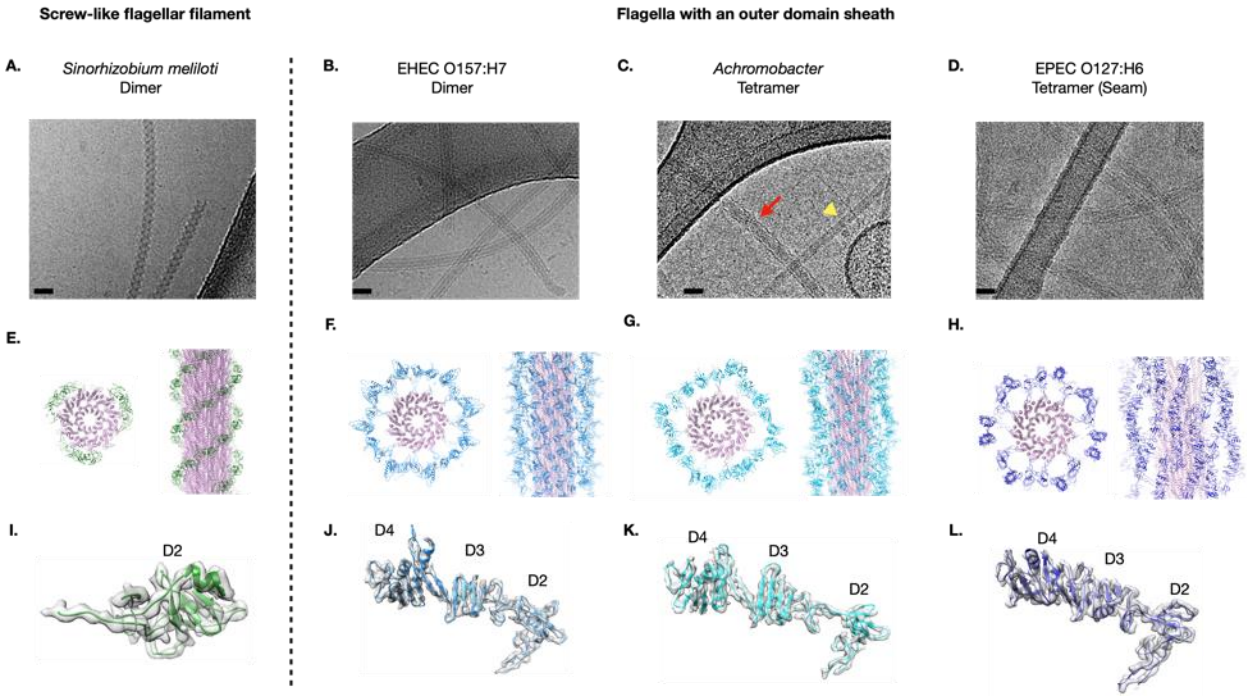

Figure 3.1. Cryo-EM structures of outer-domain sheath and screw-like flagellar filaments. a-d. Cryo-electron micrographs of *S. meliloti* (**A**), EHEC H7 (**B**), *Achromobacter* (**C**), and EPEC H6 (**D**) flagellar filaments. The scale bar in each micrograph is ~20 nm. In **c**, the red arrow points to an *Achromobacter* flagellar filament. The yellow triangle indicates an *Agrobacterium tumefaciens* flagellar filament. Scale bar is 200 Å. Atomic models of the *S. meliloti* (**E**), EHEC H7 (**F**), *Achromobacter* (**G**), and EPEC H6 flagellar filaments (**H**). The conserved D0/D1 core domains are shown in pink. Outer domains are colored green for the screw-like flagellar filament (**E)** and various shades of blue for those with an outer-domain sheath (**F-H**). Models and density maps for the single outer domain (D2) of *S. meliloti* (**I**), and for the three outer domains (D2, D3, D4) of EHEC H7 (**J**), *Achromobacter* (**K**), and EPEC H6 (**L**).

The *S. meliloti*, EHEC O157:H7, and *Achromobacter* flagellar filament structures were solved using helical reconstruction to 3.7 Å, 3.6 Å, and 3.7 Å resolution, respectively. Since a seam breaks the helical symmetry in the EPEC O127:H6 flagellar filament, reconstruction was performed asymmetrically to 4.0 Å. For the *S. meliloti* flagellar filament, a prominent 3-start helix is created by the outer domains surrounding the flagellar core (Fig. 3.1E), which gives the *S. meliloti* filament surface the appearance of a 3-start screw. A simple screw can be described by a 1-start helix, where there is a single continuous helix forming the threads on the surface of the screw. For a 3-start screw, there are three separate strands that form the surface threads. The outer domains of the EHEC O157:H7, *Achromobacter,* and EPEC O127:H6 flagellar filaments form a sheath-like structure surrounding the core domains D0 and D1 (Fig. 3.1F-H), which we named an Outer Domain Sheath (ODS). The *S. meliloti* filament possesses a single outer domain, D2 (Fig. 3.1I), while the ODS filaments have three outer domains, D2, D3 and D4 (Fig. 3.1J-L). The folds of the sheath-forming outer domains D2 and D3 have no homology or similarity to those of the D2 and D3 domains in *Salmonella* filaments[95]. The power spectrum of each flagellar filament (Fig. 3.2B-E) was very different from that of any published canonical flagellar filament (Fig. 3.2A)[14], indicative of the uniqueness of each structure. These unique power spectra were the result of

reductions in helical symmetry of the canonical flagellar symmetry where the symmetry of the outer domains corresponded to twice or four times the helical symmetry of a canonical flagellar filament (Table 3.1).

Table 3.1. Helical symmetry parameters for *S. meliloti*, EHEC H7, *Achromobacter sp.* MFA1 R4, and EPEC H6.

Organism	Core rise (Å)	Core twist (°)	Outer domain rise (Å)	Outer domain twist (°)	Outer domain symmetry
Sinorhizobium meliloti	4.75	65.46	9.5	130.92	Dimer
EHEC O157:H7	4.8	65.4	9.6	130.8	Dimer
Achromobacter	4.725	65.4	18.9	-98.4	Tetramer
EPEC O127:H6	4.813	65.534	N/A	N/A	Tetramer w/ seam

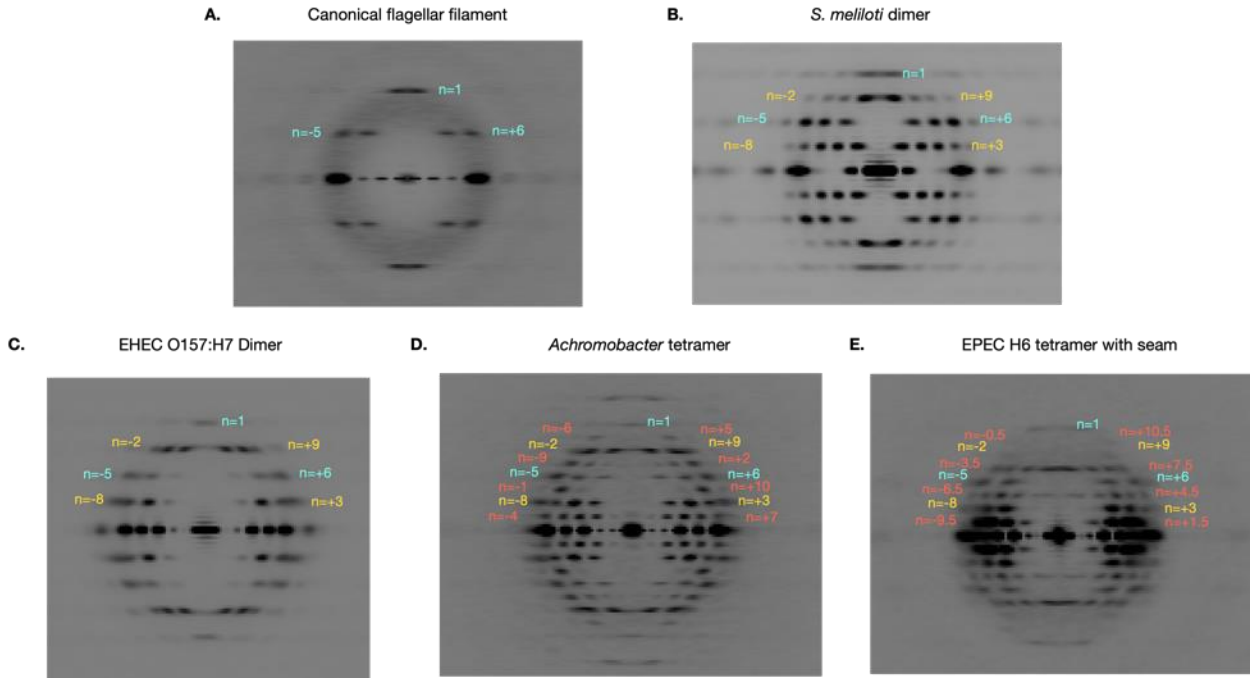

Figure 3.2. Power spectra of canonical (monomeric), dimeric, and tetrameric flagellar filaments. (A). Averaged power spectrum of canonical flagellar filament form *Agrobacterium tumefaciens*. The power spectrum matches those published for wild-type supercoiled flagellar filaments[13]. The n=-5 and n=+6 layer lines are not separated due to the fact that the average twist is close to 65.45°, which means that the 11-start protofilaments are nearly vertical. (**B).** Averaged power spectrum of the *S. meliloti* dimeric flagellar filament. There is a doubling of the number of observed layer lines compared to the canonical filament due to a reduction in symmetry along the 6-start helix, giving rise to the 3-start helices. (**C).** Averaged power spectrum of the dimeric EHEC O157:H7 flagellar filament. (**D).** Averaged power spectrum of the tetrameric *Achromobacter* flagellar filament. This *Achromobacter* filament has four times the number of observed layer lines compared to the canonical flagellar filament. A tetramer is created by a reduction of symmetry along the 8-start (n=8) helix. (**E).** Averaged power spectrum of the EPEC O127:H6 flagellar filament. This power spectrum also has four times the number of observed layer lines as the canonical flagella. However, its pattern is distinct from the *Achromobacter* due to a reduction in symmetry along the 9-start helix rather than the 8-start. This generates layer lines with half-integer Bessel orders.

A shared feature in all four structures is the formation of symmetrical homodimers between the outer domains of flagellin subunits, with one outer domain rotated 180° relative to the other (Fig. 3.3). These symmetrical dimers are the result of outer domains being able to adopt either an “up” or “down” conformation. The outer surface of all these filaments is therefore bipolar. This shared feature is surprising as while the structures and sequences of the EHEC H7, *Achromobacter,* and EPEC H7 flagellin outer domains are all similar, the *S. meliloti* outer domains shared no sequence homology (Fig. 3.4) or structural homology (Fig. 3.1I-L) with the others.

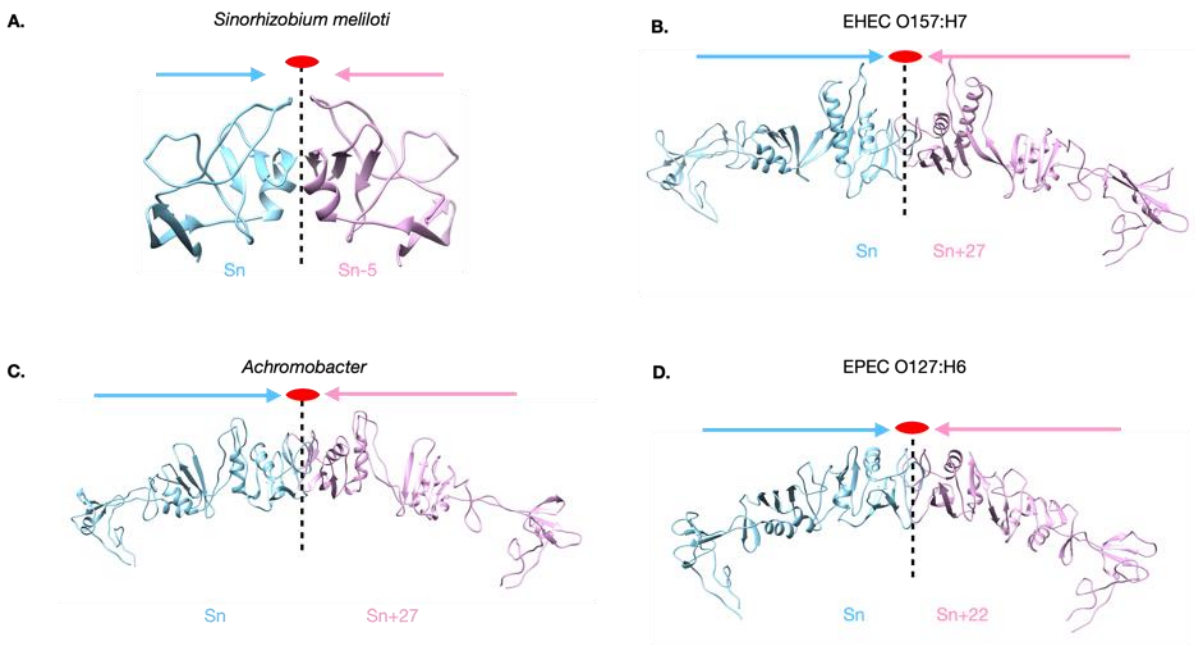

Figure 3.3 Outer domain dimers of EHEC O157:H7, EPEC O127:H6, *Achromobacter*, and *S. meliloti* flagellar filaments. Arrows pointing towards each other indicate that the outer domains have D1 symmetry and are identical to each other through a 180° rotation about the vertical 2-fold axis shown. Since these 2-fold axes intersect the helical axis of the filament, the outer surfaces of the flagellar filaments have a global D1 symmetry.

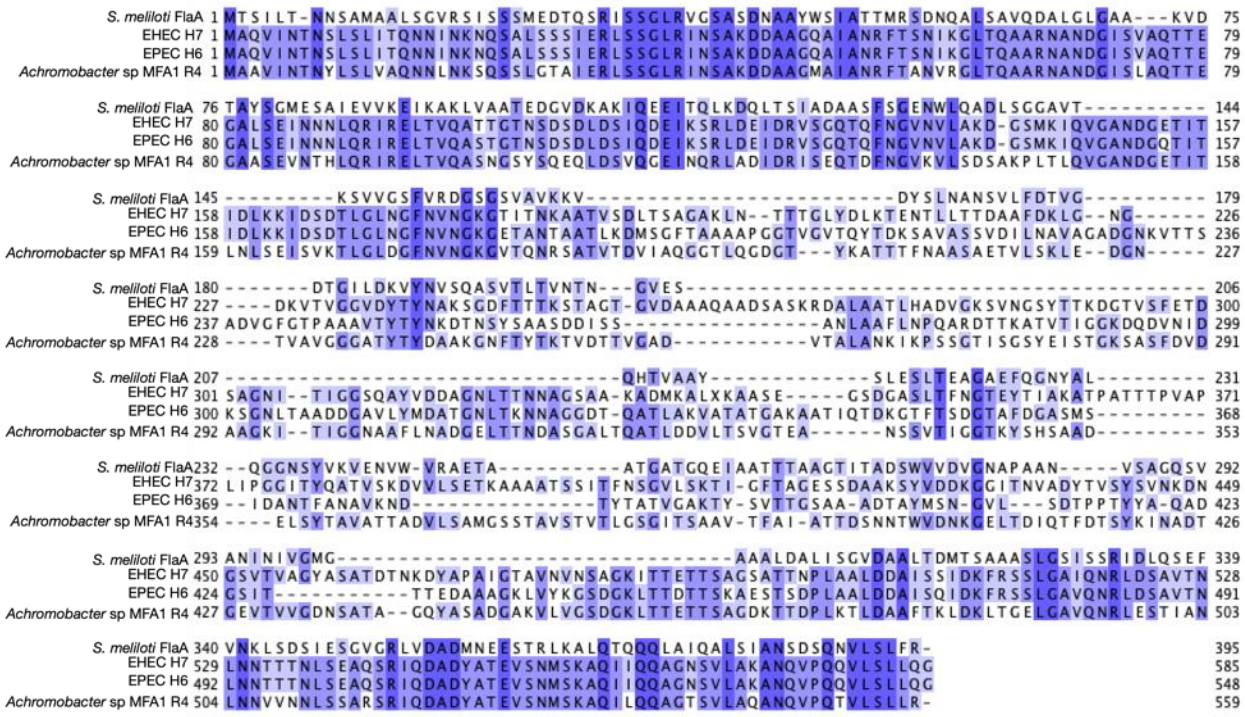

Figure 3.4. Multiple sequence alignment of the flagellin subunits from the structures in this study. The alignment is colored according to percentage identity.

Dimeric outer domain interactions of the *S. meliloti* screw-like filament

The asymmetric unit (ASU) of the *S. meliloti* flagellar filament consists of two flagellin conformations, which are identical in the core domains, but the outer domains are rotated 180° from each other resulting in an “up” and a “down” conformation (Fig. 3.5A). The D2 domains in the *S. meliloti* filament form dimers between an arbitrary “up” conformation subunit (S_n) and two “down” conformation subunits (S_{n+5}) and (S_{n+11}) as depicted in Fig. 3.5B. Since the outer surface is bipolar, a “down” D2 will also interact with two “up” D2s. These interactions form a 3-start helix surrounding the flagellar core, which resembles the grooves in a helical screw (Fig. 3.5C). They can be described as a reduction in helical symmetry along the 6-start helix, resulting in a 3-start helix (Fig. 3.6).

Due to this subunit pairing, the axial rise and helical twist of the outer domains are twice those of the flagellar core.

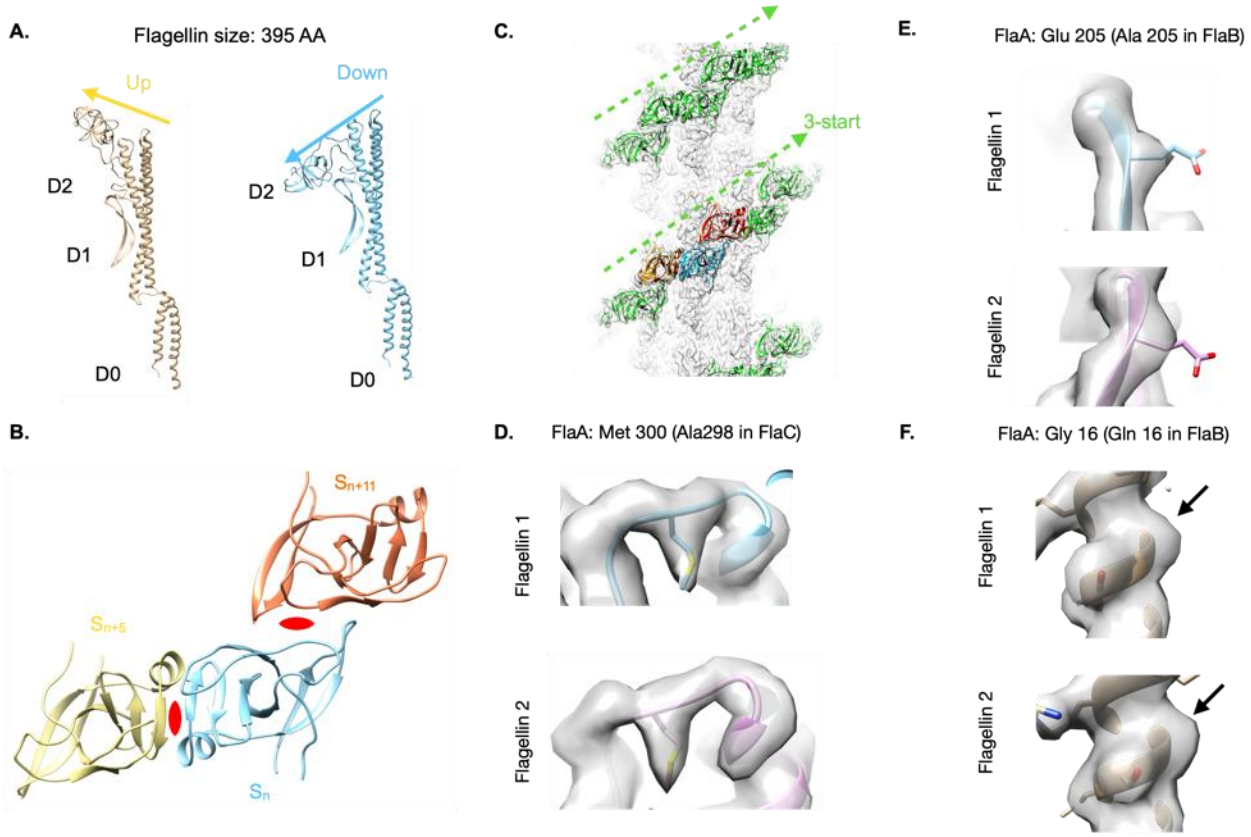

Figure 3.5. Structural details of *S. meliloti* screw-like flagellar filament. (A). The two flagellin conformations of the *S. meliloti* flagellar filament asymmetric unit. The core domains D0 and D1 are unchanged in either conformation. The outer domains however change and have either an up or down conformation. (**B).** Each *S. meliloti* outer domain from subunit S_n interacts with the outer domains of flagellins that are 5 (S_{n+5}) and 11 (S_{n+11}) subunits away. Two-fold axes are indicated by the red symbols. (**C).** The S_n: S_{n+5} and S_n: S_{n+11} dimers generate a right-handed 3-start helix (green) surrounding the filament core (light gray). (**D).** Model and map showing methionine 300 for the FlaA model fitting the map density well for both flagellin conformations. (**E).** FlaA model and map density of position 205 for both flagellins in the ASU. (**F).** FlaA model and map density of position 16 for both flagellins in the ASU.

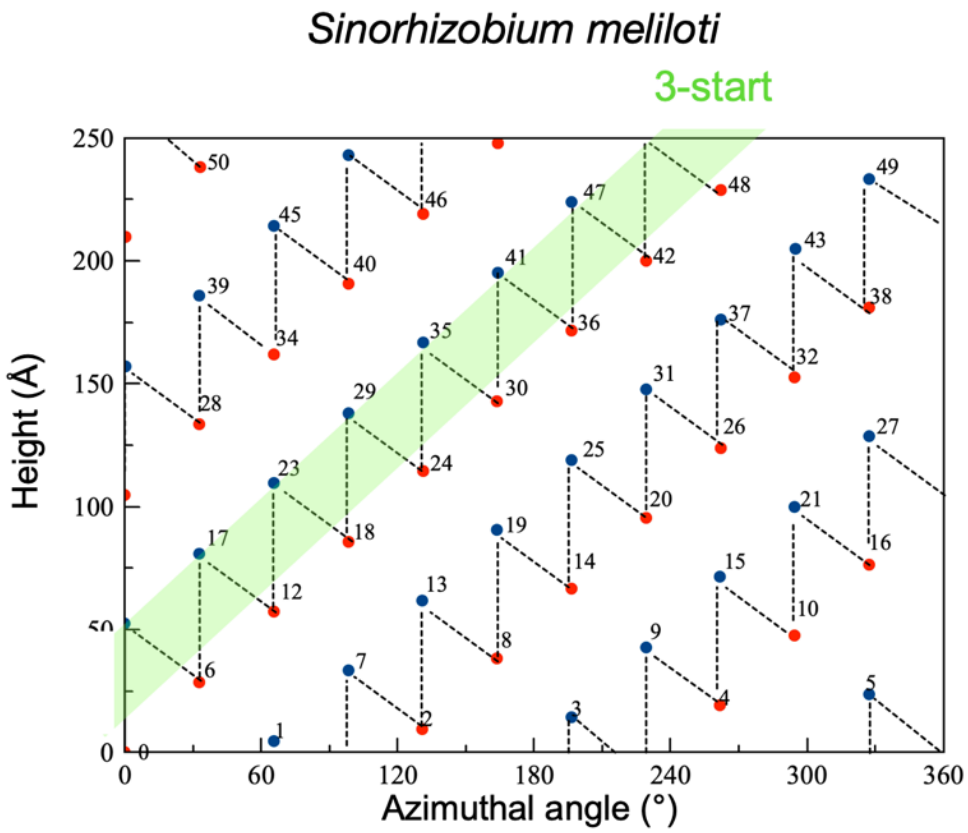

Figure 3.6. Helical net of the *S. meliloti* screw-like flagellar filament. We use the convention throughout that the surface has been unrolled and one is looking at it from the outside. Alternating blue and red dots indicate the two different flagellin conformations (up or down). The S_n: S_{n+5} and S_n: S_{n+11} dimers are indicated by dashed lines and generate the 3-start helix (green highlight), which is a pairing across the 6-start helix of the core symmetry.

It has been hypothesized that flagellin dimers in *S. meliloti* are comprised of one each of FlaA and FlaX where FlaX can be any of the other three flagellins[96]. A multiple sequence alignment of all four *S. meliloti* flagellins reveals key similarities and differences among them (Fig 3.7). FlaD is composed of 321 amino acid residues, and therefore is too short to fit the structure. In addition, it is transcribed at very low levels and expected to have a minor structural function[96]. FlaA, B, and C have the correct size, but a few key residues in FlaC clearly do not fit the density map (Fig. 3.5D). FlaA and FlaB are very similar in sequence. In positions with residue variations between FlaA and FlaB, there is support for FlaA at position 205 (Fig. 3.5E) and FlaB at position 16 (Fig.

3.5F) in both flagellins of the ASU. This suggest that there is likely a mixture of FlaA and FlaB in the flagellar filament segments of our images. The fact that FlaA is the primary component of the filaments is supported by the strongly reduced motility of a *flaA* mutant strain due to severely truncated flagellar filaments[96].

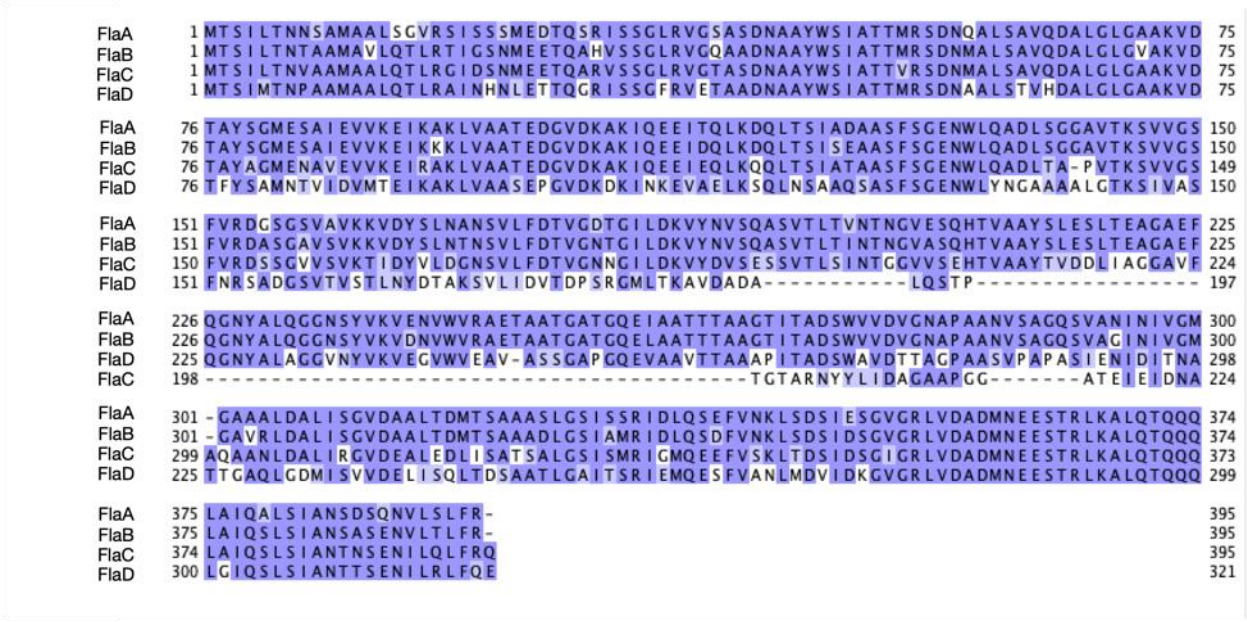

Figure 3.7. Multiple sequence alignment of the four *S. meliloti* flagellins. Alignment is colored by percent identity.

Outer domain dimerization of the EHEC H7 flagellar filament

Similar to the *S. meliloti* flagellar filament, the EHEC O157:H7 flagellar filament has a dimer symmetry (Table 3.1) with two flagellins in the asymmetric unit, one with an up and one with a down outer domain conformation (Fig. 3.8A). The much larger EHEC H7 outer domains (D2, D3 and D4) interact with outer domains from three other subunits forming a mesh-like ODS (Fig. 3.8B). In addition to forming dimers with outer domains that are five and 11 subunits away, every EHEC H7 outer domain also pairs

with an outer domain that is 27 subunits away (Fig. 3.8C). This long-range S_n:S_{n+27} dimer results in an observable 8-start helix on the surface of the sheathed filament in addition to a 3-start (Fig. 3.8D). This can be seen quite easily by drawing these interactions out on the EHEC H7 helical net (Fig 3.9).

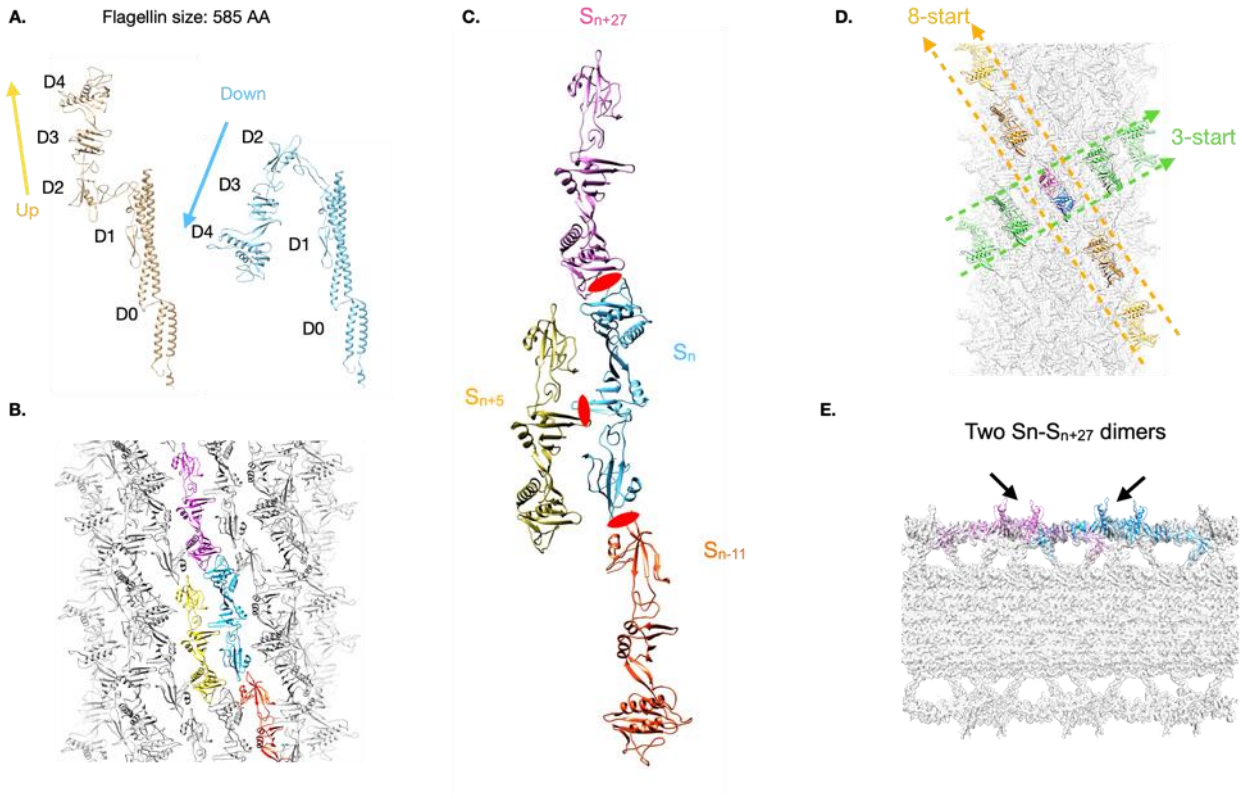

Figure 3.8: Structural details of the dimeric EHEC O157:H7 flagellar filament. (A). The two flagellin conformations of the EHEC H7 flagellar filament asymmetric unit. (**B).** Mesh-like sheath generated by interactions between four flagellin outer domains (blue, pink, yellow, and orange). (**C).** The EHEC H7 outer domains from subunit S_n interact with the outer domains of flagellins that are 5 (S_{n+5}), 11 (S_{n+11}), and 27 (S_{n+27}) subunits away. (**D).** The S_n:S_{n+27} dimer generates prominent right-handed 3-start and left-handed 8-start helices in the sheath surrounding the core. **(E).** View of the EHEC H7 flagellar filament cut halfway through the filament. The arrows are pointing to two S_n:S_{n+27} dimers, one blue and the other magenta, which are at the same radius

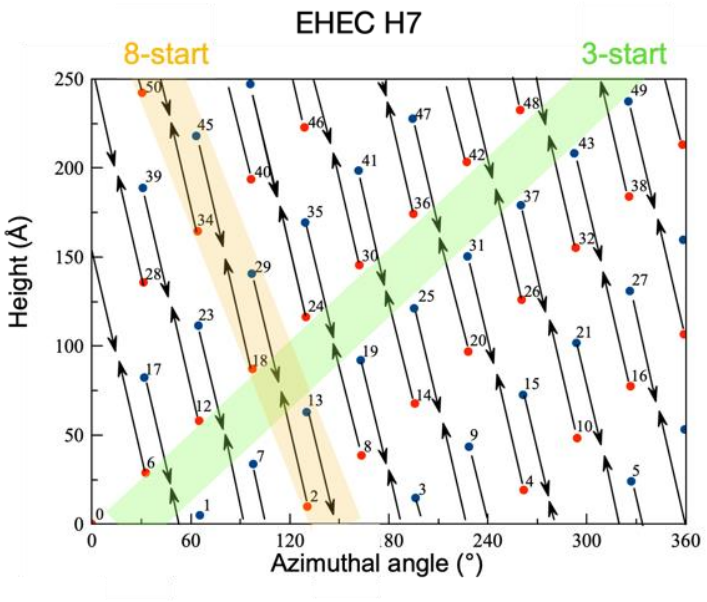

Fig 3.9. Helical net of the EHEC H7 flagellar filament. Arrows indicate "up" or "down" conformation and meet at the location of each S_n: S_{n+27} dimer. The 3-start helix formed by this dimer is shown in green. The 8-start is shown in yellow.

Outer domain tetramerization of the *Achromobacter* flagellar filament

Like the EHEC H7, the *Achromobacter* ODS is also formed by interactions between flagellins that are five, 11, and 27 subunits away (Fig. 3.8C). However, there are four flagellins in the *Achromobacter* ASU, with two up and two down conformations (Fig. 3.10A). The key distinction is that all of the EHEC H7 S_n:S_{n+27} dimers form at a single radius (Fig. 3.8E) while the *Achromobacter* S_n:S_{n+27} dimers occur at two different radii (Fig. 3.10B). Consequently, the *Achromobacter* outer domain exists in four conformations and the filament has four flagellins in the asymmetric unit with an axial rise and twist for the outer domains that is four times larger than that of the flagellar core (Table 3.1). This tetrameric symmetry is characterized by two 4-start helices that occur at different radii (Fig. 3.10C). This can be viewed as a reduction of the 8-start helix

observed in the EHEC H7 outer domains (Fig. 3.8D and Fig 3.9) to two 4-start helices occurring at different radii (Fig. 3.10C and Fig 3.11).

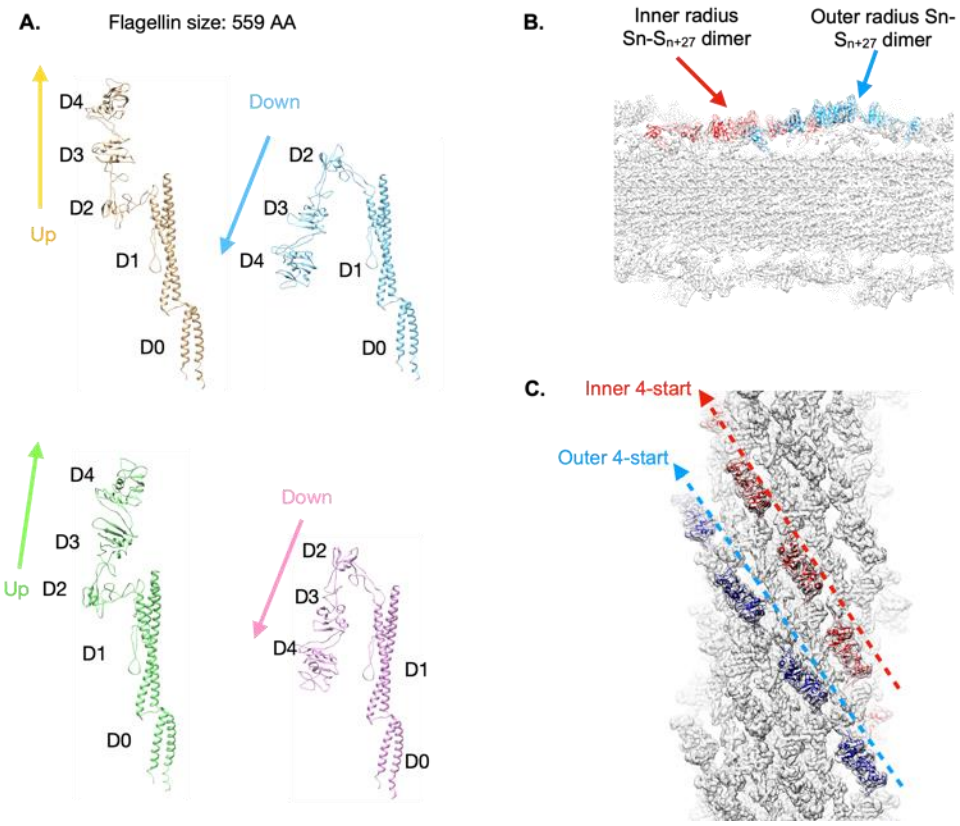

Figure 3.10 Structural details of the tetrameric *Achromobacter* flagellar filament. (A). The four flagellin conformations within the *Achromobacter* flagellar filament asymmetric unit. **(B).** *Achromobacter* outer domain interactions from subunits S_n: S_{n+27} occur at two different radii. **(C).** The two different dimer conformations result in two left-handed 4-start helices on the surface of the *Achromobacter* flagellar filament.

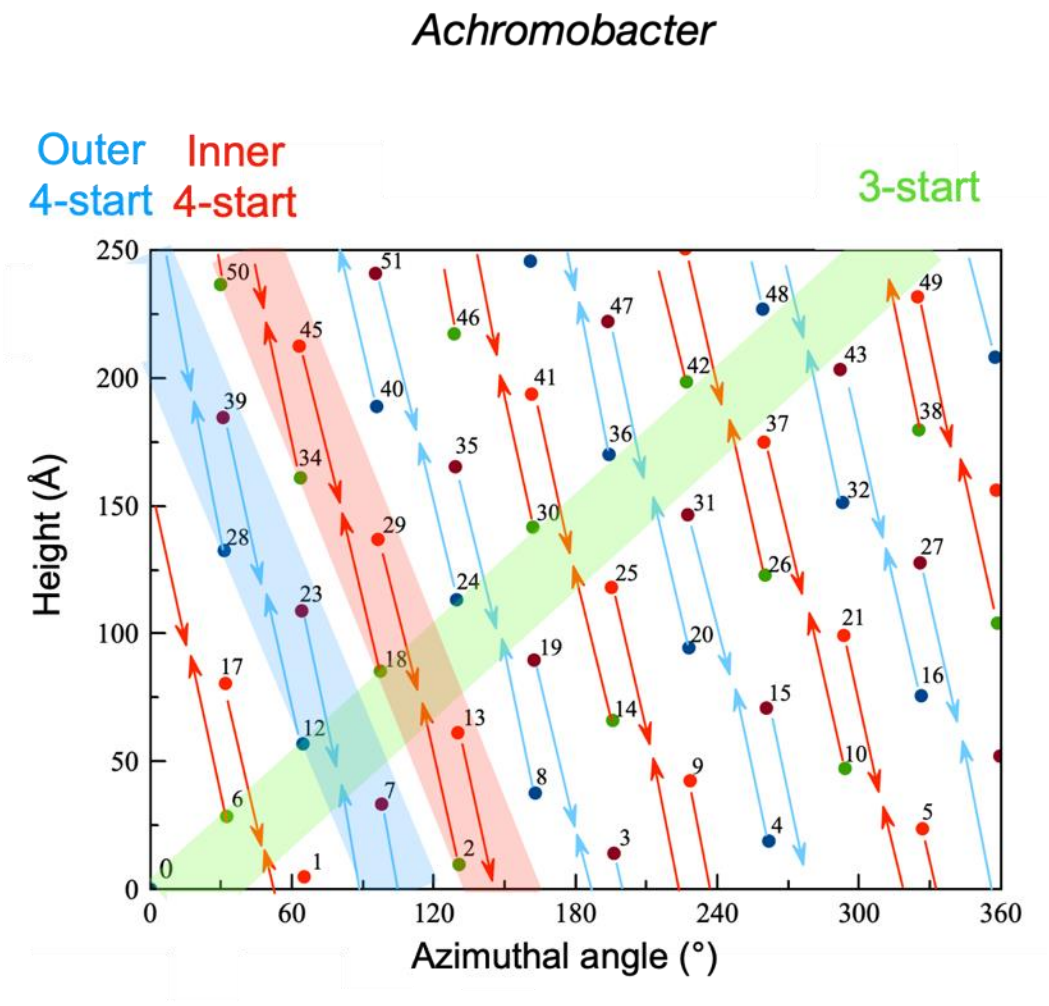

Figure 3.11. Helical net of the *Achromobacter* flagellar filament. Red and blue arrows and highlights show the two different radii for the S_n: S_{n+27} dimers creating the two different 4-start helices. The right-handed 3-start helix is shown in green.

Outer domain tetramerization of the EHEC H6 filament results in a seam

The power spectrum of the EPEC O127:H6 flagellar filament (Fig 3.2E) superficially looks similar to the tetrameric *Achromobacter* filament. However, the Bessel order of many of the layer lines must be different because the EPEC H6 outer domains cannot be reconstructed using the tetrameric flagella symmetry of *Achromobacter*. We have determined that this is a result of a seam. The EPEC H6 filament ODS is created by an outer domain tetramer with most flagellins being in one of

four flagellin outer domain conformations (Fig. 3.12A), but the details of this tetramerization are quite different from those in *Achromobacter*. Due to the presence of the seam, most but not all D2 domains are in one of these four conformations. In one dimer conformation, “D2 Dimer A”, domain D2 of one subunit forms a dimer with D2 from another flagellin that is 11 subunits away (Fig. 3.12B). In the other conformation, “D2 Dimer B”, flagellins that are 6-subunits away (S_n:S_{n+6}) form dimers with each other (Fig. 3.12C). These two dimer conformations are easily observed in the 4.0 Å EPEC H6 map (Fig 3.13). While a large region of the ODS has very poor density, the resolution of another portion allows the creation of an *ab initio* atomic model of the complete flagellin structure. A 6.7 Å resolution volume was reconstructed asymmetrically where all flagellin subunits could be fitted into the density (Fig. 3.12D-E). Analysis of this low-resolution structure revealed the seam, which involves the lack of D2 dimer formation between two flagellin subunits (Fig. 3.12E). This seam is a consequence of a reduction of symmetry along the 9-start helices of a flagellar filament composed of dimeric flagellin such as EHEC H7. This results in two 4.5-start helices (Fig. 3.12D) that are discontinuous along a line indicated by the symbol “*” (Fig. 3.12E). The two outer domain subunits in the seam still form S_n:S_{n+22} dimers in domain D4. However, rather than forming an S_n:S_{n+6} or S_n:S_{n+11} dimer in domain D2, the two seam subunits form S_n:S_{n+5} dimers with each other (Fig. 3.12F) similar to those forming the EHEC H7 and the *Achromobacter* ODSs.

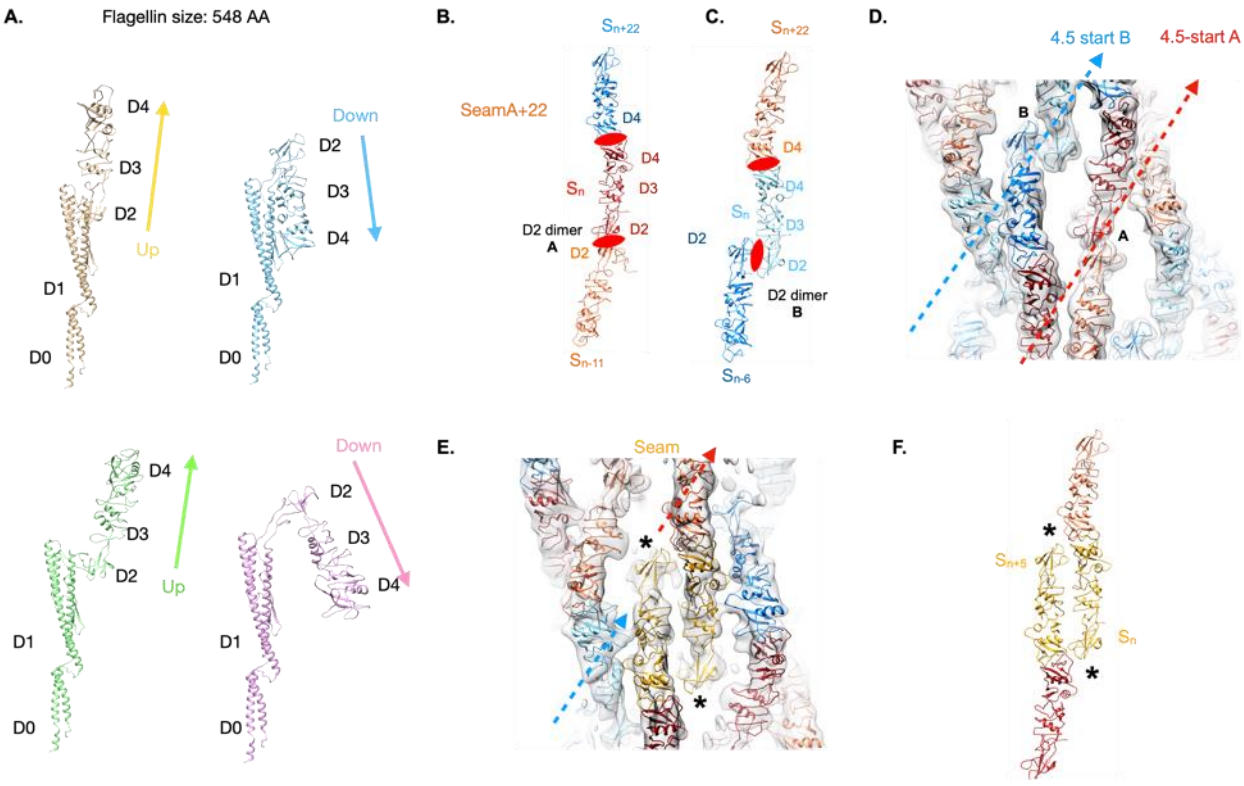

Figure 3.12. Structural details of the EPEC O127:H6 flagellar filament with a seam. (A). The four main flagellin conformations of the EPEC H6 flagellar filament. **(B).** An EPEC H6 outer domain dimer is generated by a symmetrical interaction between outer domain D2 from flagellin S_n with domain D2 of a flagellin 11 subunits away, S_{n+11}. This is named "D2 dimer A". Flagellin outer domains also form a symmetrical dimer between domain D4 in subunit S_n with domain D4 from a flagellin 22 subunits away (S_{n+22}). **(C).** The second EPEC H6 outer domain dimer, "D2 dimer B", is created between S_n and a flagellin six subunits away S_{n-6}. **(D).** Surface view of the EPEC H6 flagellar filament showing the fit of the outer domain models into the low-resolution map. The two dimer conformations are indicated by "A" and "B" and form two right-handed 4.5-start helices. **(E).** A surface view of the EPEC H6 flagellar filament, rotated by 180° around the filament axis from the view in (**D**), showing the seam, *, with the subunits forming the seam in gold. **(F).** Interactions between outer domain subunits (gold) along the seam. The two subunits in the seam are 5 subunits away from each other in the filament (S_n:S_{n+5}).

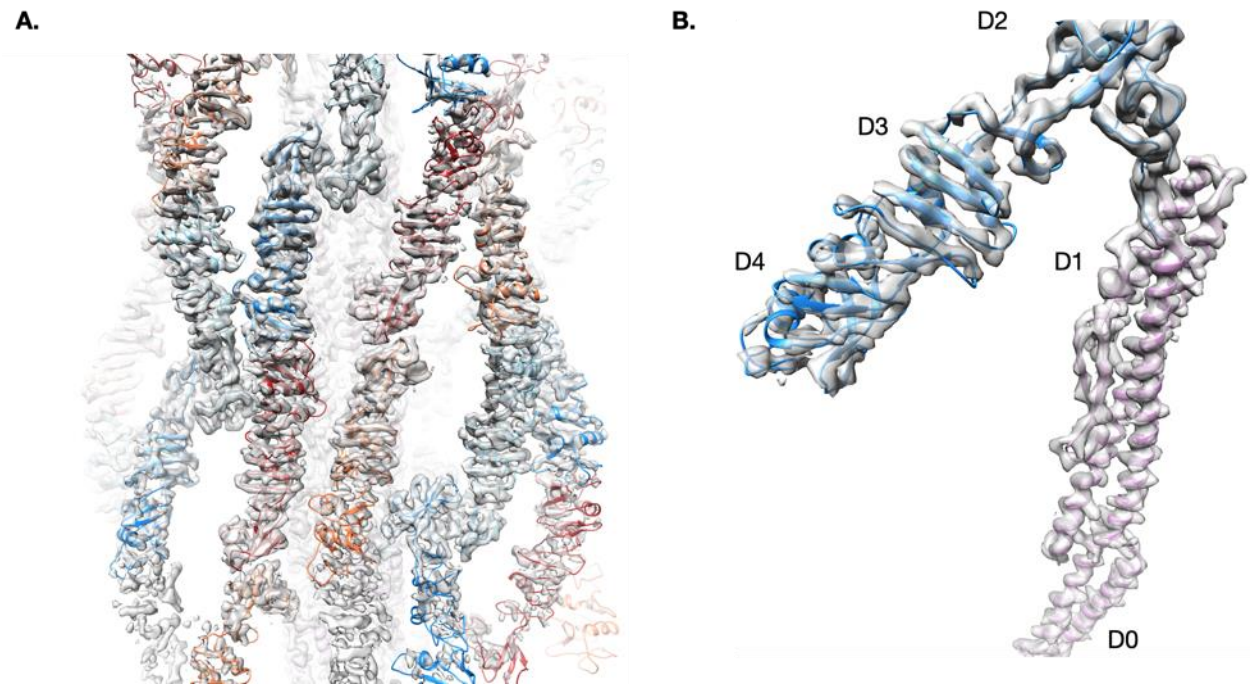

Figure 3.13 High resolution structure of the EPEC O127:H6 sheathed flagellar filament. (A). Surface of the EPEC H6 flagellar filament density map with fit of atomic models. **(B).** Full atomic model and corresponding density map for a single EPEC H6 subunit.

Disruption of the EHEC H7 ODS-forming dimer in domain D4 impairs motility

The surface of the screw-like flagellar filament of *S. meliloti* and related species is thought to be an adaptation to swimming in highly viscous environments[97,98], while the function of the ODS is unknown. To investigate ODS function, we aimed to disrupt its formation by mutation of several key residues in the EHEC H7 S_n:S_{n+27} domain D4 dimer interface (Fig. 3.14A) . Examination of the amino acid sequences of the prominent S_n:S_{n+27} and S_n:S_{n+22} D4 dimer interfaces of the EHEC H7, *Achromobacter*, and EPEC H6 flagellins revealed a highly conserved sequence (Fig. 3.14B). Residues N319 and N323 in the EHEC H7 flagellin were chosen for mutagenesis (Fig. 3.14A-B). Two double mutants were constructed: N319F N323F (FF) and N319R N323R (RR). Soft agar

(0.3%) motility assays revealed drastically reduced motility of cells with the FF and RR mutant flagellar filaments as compared to the wild type H7(Fig. 3.14C).

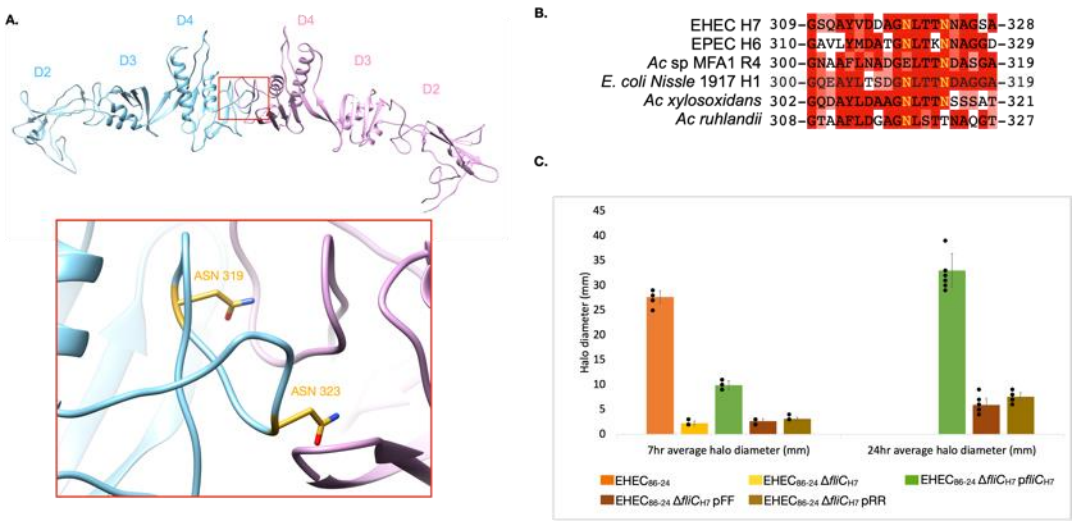

Figure 3.14. Mutagenesis of the EHEC H7 outer domain S_n: S_{n+27} dimer reduces motility. (A). The S_n: S_{n+27} dimer is shown with the same color scheme as Fig. 4.5E. The inset reveals a close-up view of the dimer site. Residues that were mutated are shown in gold in the cyan subunit. **(B).** Multiple sequence alignment of the D4 dimer site for EHEC H7, EPEC H6, *Achromobacter sp* MFA1 R4, *E. coli* Nissle 1917, *Achromobacter xylosoxidans* and *Achromobacter ruhlandii* flagellins. The two asparagine residues that were mutated in EHEC H7 (gold) are conserved with the exception of a glutamate residue in the second position in the *Achromobacter* flagellin. Sequences are colored by conservation of chemical properties. **(C).** Soft agar (0.3 %) motility assay for EHEC 86-24 strains after 7 and 24 hours. Data for the wild type ($EHEC_{86\text{-}24}$) and wildtype with fliC deletion ($EHEC_{86\text{-}24}$ Δ*fliC*$_{H7}$) are not shown for the 24-hour as the $EHEC_{86\text{-}24}$ covered the entire plate and $EHEC_{86\text{-}24}$ Δ*fliC*$_{H7}$ showed no change. The pGEN-*fliC* strain complements the Δ*fliC* deletion strain. The FF and RR mutants were generated with the pGEN-*fliC* strain. Each bar represents the average halo diameter on the plate for each condition. A total of nine soft agar plates was measured for each condition. The black dots overlayed over each graph represent the data individual data points from which the bar graph was made from. The error bars represent the standard deviation for each condition.

ODS flagellar filaments demonstrate non-canonical polymorphic transitions

For an extended look at the data and movies that led us to the conclusions of this section please see the paper that chapter 4's results are published in[99]. We labeled bacteria and their flagellar filaments with an amine-reactive dye and examined them

using fluorescence confocal light microscopy. For the *Achromobacter* sp. MFA1 R4 and EHEC H7 flagellar filaments we identified the presence of an intermediate waveform which was not observed in the canonical *E. coli* K-12 flagellar filaments (Fig 3.15). The intermediate waveform was in addition to the canonical polymorphic transitions (Fig 3.15B) and is created by the outer domain sheath of the *Achromobacter* and EHEC H7 flagellar filaments.

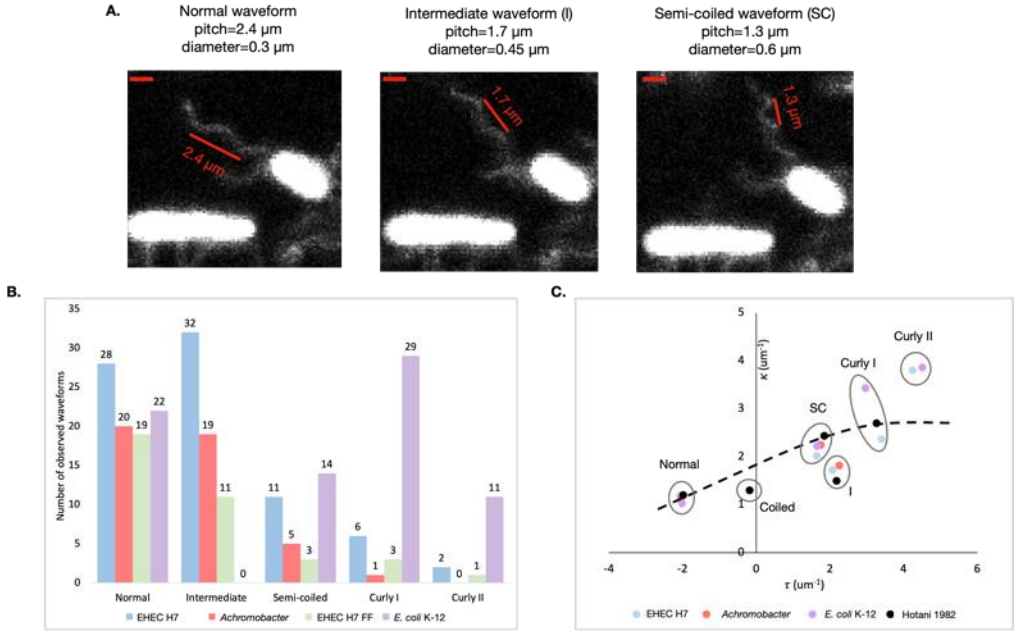

Figure 3.15. ODS flagellar filaments commonly adopt an "intermediate" waveform. (A). Images of fluorescently labeled *Achromobacter sp.* MFA1 R4 flagellar filaments with the three most common waveforms. Scale bars in the top left of each image correspond to ~1 µm. **(B).** Bar graph of waveforms observed in rotating EHEC H7, *Achromobacter sp* MFA1 R4, EHEC H7 FF mutant, and *E. coli* K-12 flagellar filaments. There were 55, 34, 25, and 54 unique filaments whose polymorphic transitions were analyzed for the EHEC H7, *Achromobacter sp* MFA1 R4, EHEC H7 FF mutant, and *E. coli* K-12 samples respectively. **(C).** Plot of curvature, κ, versus twist, τ, for the observed flagellar waveforms. The black dots are measured waveforms from Hotani (1982). The black dashed line is the predicted curvature κ (Calladine, 1978; Calladine et al., 2012) as a function of τ. "I" stands for intermediate waveform and also includes the previously characterized "medium" waveform. "SC" stands for semi-coiled.

The ODS prolongs *E. coli* tumbling

To compare the run-and-tumble motility behavior of the various EHEC strains to that of *E. coli* K-12 AW405 (K-12), we employed phase-contrast microscopy and recorded movies of swimming cells at 30 frames per second. The D0 and D1 domains of the *E. coli* K-12 flagellin exhibit over 90% sequence identity to the H7 flagellin (Fig. 3.16A). However, its outer domains show no significant sequence homology to H6 or H7 and do not form a sheath[100]. The *E. coli* K-12 cells exhibited the highest average swimming velocity at about 25 µm/s (Table 3.2), which agrees with previous published values[101,102]. Cells frequently displayed extended periods of straight swimming followed by very short tumbles with a duration of approximately 220 ms (Fig. 3.16B). EHEC cells expressing the wild-type H7 flagellar filament either from its native locus or from a plasmid had swimming velocities of about 19 µm/s (Table 3.2). Interestingly, these cells exhibited prolonged tumbles of around 500 ms, twice as long as that of K-12 wild type (Fig. 3.16B). EHEC strains with the FF mutant swam more slowly (Table 3.2) and stayed in the tumble mode for approximately 1,500 ms (Fig. 3.16B).

To investigate whether the H7 flagellar filament is the cause of the prolonged tumble duration, we complemented the *E. coli* K-12 *fliC* deletion strain HCB5 with EHEC *fliC* ($FliC_{H7}$) and the native K-12 *fliC* ($FliC_{H48}$) from plasmid pTrc99a using the native *fliC* promoters (p*P*$fliC_{H48}$ and p*P*$fliC_{H7}$). The *E. coli* K-12 Δ*fliC* strain regained swimming motility when complemented with either EHEC or K-12 *fliC* (Table 3.2). Interestingly, complementation with p*P*$fliC_{H7}$ resulted in tumbling durations similar to the wild-type EHEC H7 strains with approximately 500 ms and nearly twice of the *E. coli* K-12 strain

AW405 (Fig. 3.16C). These results suggest that the H7 flagellin is the cause of extended tumbles.

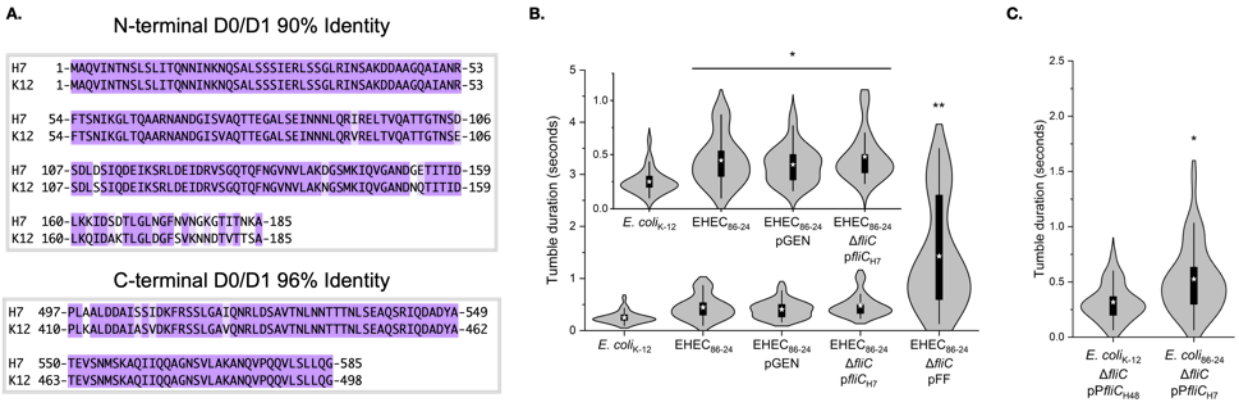

Figure 3.16. The H7 ODS prolongs *E. coli* tumbling. (A). Alignment of EHEC H7 and *E. coli* K-12 sequences in domains D0 and D1. Color indicates conservation of chemical properties. **(B).** Tumbling times for various strains of *E. coli.* The strains are *E. coli* K-12, WT EHEC O157:H7 ($EHEC_{86\text{-}24}$), WT EHEC O157:H7 with empty pGEN vector ($EHEC_{86\text{-}24}$ pGEN), EHEC O157 Δ*fliC* complemented with the H7 *fliC* in the pGEN vector ($EHEC_{86\text{-}24}$ Δ*fliC* $pFliC_{H7}$), and the EHEC H7 FF mutant ($EHEC_{86\text{-}24}$ Δ*fliC* pFF) are shown. The inset shows the tumbling times for all strains except the FF mutant. At least 50 cells were analyzed for each measurement. Asterisks represent statistically significant differences of the given strain(s) from other strain(s) (*, P<0.01; **, $P<10^{-6}$). **(C).** Tumbling times for *E. coli* K-12 HCB5 (Δ*fliC*) cells with the non-sheathed H48 flagellar filament (*E. coli* K-12 Δ*fliC* $pPfliC_{H48}$) and the sheathed H7 flagellar filament (*E. coli* K-12 Δ*fliC* $pPfliC_{H7}$). Asterisks represent statistically significant differences of the indicated strain(s) from other strain(s) (*, P<0.003).

Table 3.2. Swimming velocity measurements for *E. coli* K-12 and EHEC strains.

Strain	Velocity (μm/s)
E. coli$_{K\text{-}12}$	25.2 ± 0.2
E. coli$_{K\text{-}12}$ Δ*fliC* $pPfliC_{H48}$	19.9 ± 0.7
E. coli$_{K\text{-}12}$ Δ*fliC* $pPfliC_{H7}$	19.0 ± 0.1
$EHEC_{86\text{-}24}$ pGEN	19.5 ± 2.4
$EHEC_{86\text{-}24}$ Δ*fliC* $pfliC_{H7}$	19.0 ± 0.1
$EHEC_{86\text{-}24}$ Δ*fliC* pFF	13.8 ± 1.5
$EHEC_{86\text{-}24}$ Δ*fliC* pRR	13.2 ± 1.7

For each strain 50 cells were analyzed.

3.4 Discussion

Each of the four presented flagellar filament structures exhibit a different level of structural complexity: The *S. meliloti* and EHEC H7 structures are formed with outer domain dimers, while the *Achromobacter* and EPEC H6 structures are produced by outer domain tetramers. The effects of dimerization and tetramerization of flagellin outer domains has been seen before in diffraction patterns of flagellar filaments, but these studies were unable to determine the nature of the dimerization or tetramerization[85,91,103]. Flagellar filaments without a seam were defined as helically perturbed, and ones with a seam as non-helically perturbed. In the light of our findings, the use of the term perturbation should be re-evaluated, because the generation of these structures involves a full 180° rotation of every other flagellin's outer domain. Additionally, the outer domain sheaths we describe in this paper should not be confused with other flagellar sheaths which are either membranous[104] or created by proteins other than flagellin[105].

A structural homology search of the individual flagellin outer domains from each structure using the Dali server[51] yielded only a few potential structural homologs in unrelated proteins (Table 3.3). However, none of these hits were convincing in terms of sharing an overall fold. Since the outer domains of the four structures formed dimers about a two-fold axis in strikingly similar manners (Fig. 3.3), this led us to wonder if there was any commonality between the *S. meliloti* dimer site and that of the three ODS flagellar filaments' D4 dimer site. Interestingly, the electrostatic potential of the EHEC H7 and *Achromobacter* D4 dimers along with the *S. meliloti* D2 dimer all have a negative charge at the dimer site (Fig. 3.17A). In the case of the EHEC H7 dimer one

aspartate at position 316 from each subunit is at the interface (Fig. 3.17B) while for *S. meliloti* the negative charge is due to glutamates from each subunit (Fig. 3.17C). It has been suggested that divalent cations are critical to the stability of the *S. meliloti* outer domains. We found no evidence of cation density being preset in these negatively charged pockets in any of these structures, but this does not necessarily exclude the possibility. Another possibility is that the dimer interactions are strengthened at lower pH values is a remote possibility that the EHEC H7 dimer site functions as some sort of aspartic protease, but it is missing the classic Asp-Thr-Gly motif found in most aspartic proteases[106]. The discovery of proteolytic flagellin outer domains[30] might give some credibility to this possibility. The HIV-1 aspartic protease dimer[107] (Fig. 3.17D) bears some similarity to the negatively charged dimer sites in our models due to the interfacing aspartate residues from both subunits of the protease dimer.

The surface areas of the *S. meliloti*, EHEC H7, and *Achromobacter* Sn:Sn_{+5} and Sn:$Sn_{+11/-11}$ interfaces were calculated using PISA[57]. The values for these three structures are all comparable to each other and reveal increased contacts between the Sn:Sn_{+5} and Sn:$Sn_{+11/-11}$ compared to *B. subtilis* and *S. typhimurium* flagellar filaments[99], which is due to the dimers that form the screw-like surfaces and ODSs. Interestingly, the Sn:$Sn_{+11/-11}$ contacts are much less extensive than those of the *C. jejuni* flagellar filament, which outer domains make extensive contacts along the same protofilament to compensate for its destabilized toll-like receptor 5 (TLR5) sequence[33].

Table 3.3. Structural homologs to flagellin outer domains in this study.

Outer domain	Homolog 1 *	Homolog 1 PDB ID	Homolog 1 Z-score	Homolog 2 *	Homolog 2 PDB ID	Homolog 2 Z-score

S. meliloti D2	RRNA 2'-O-METHYLTRANSFERASE FIBRILLARIN	6zdt	3	2-5A-DEPENDENT RIBONUCLEASE	4oau	2.5
EHEC H7 D2	FRPC OPERON PROTEIN	5edf	3.8	CYTOTOXIC TRANSLATIONAL REPRESSOR OF TOXIN-ANTITO	5mje	3.6
EHEC H7 D3	HYPOTHETICAL PROTEIN YIIX	2if6	3.9	S-LAYER PROTEIN	3cvz	3.9
EHEC H7 D4	H1 RNA	6ahr	4.3	TYPE II SECRETION SYSTEM PROTEIN E, HEMOLYSIN-COR	4ksr	3.7
Achromobacter D2	FRPC OPERON PROTEIN	5edf	3	ADENYLATE CYCLASE EXOY	5xnw	2.9
Achromobacter D3	DE NOVO DESIGNED PROTEIN FOLDIT3	6msp	4.4	MACHADO-JOSEPH DISEASE PROTEIN 1	2aga	4.4
Achromobacter D4	STAGE II SPORULATION PROTEIN SA	3o6q	3.9	UNCHARACTERIZED PROTEIN	3fb9	3.5
EPEC H6 D2	PUTATIVE ENZYME RELATED TO ALDOSE 1-EPIMERASE	2hta	2.4	DNA PRIMASE/HELICASE	6n9x	2.1
EPEC H6 D3	BSHC	4wbd	3.5	GLUTATHIONE SYNTHETASE	1gsa	3.1
EPEC H6 D4	CDIA TOXIN	5t86	4.4	REPLICATION PROTEIN B	2h20	3.7

***** The first and second highest hits are shown for each domain.

Contacts made between adjacent *C. jejuni* flagellin outer domains might serve an additional role in stabilizing the unique motility associated with the two polar flagellar filaments of *C. jejuni,* one of which wraps itself around the cell body likely allowing for better penetration of host mucus[6]. The outer domain interactions of *S. meliloti*, EHEC H7, *Achromobacter*, and EPEC H6 flagellins are unique from all other high-resolution flagellar filament structures because they form dimers or tetramers on the surface of the filament, which generate different helical lattices around their flagellar cores. These new outer domain interactions then alter the polymorphism of the flagellar filament in the form of the constricted normal waveform for the *S. meliloti* screw-like flagellar filament[108] and the intermediate waveform for flagella with ODSs.

The intermediate waveform with an average pitch of 1.8 μm and diameter of 0.45 μm adopted by the ODS flagellar filaments is strikingly similar to the previously discovered unstable right-handed “medium” waveform with a pitch of 1.9 μm and diameter of 0.43 μm[109]. This medium waveform required the same magnitude of torque for the transition to curly or semi-coiled forms, while the torque needed for the transition from medium to normal was very small. Mutations in the EHEC H7 S_nS_{n+27} dimer site result in increased transitions between the normal and intermediate waveforms. Although FF mutant flagella still form the intermediate waveform, we believe that it is destabilized in a similar manner to the medium waveform[109]. The wild-type S_nS_{n+27} dimer site could stabilize the intermediate waveform, thus increasing the torque required for the intermediate-to-normal transition.

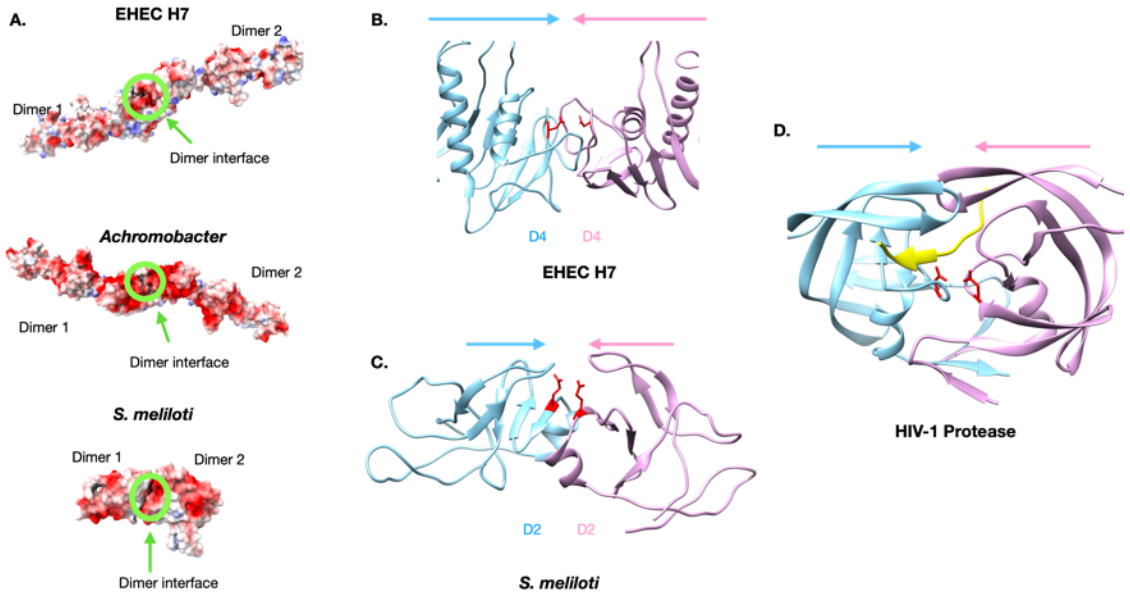

Figure 3.17. Negatively charged interfaces of the outer domain dimers. (A). Electrostatic potential surfaces of the EHEC H7, *Achromobacter,* and *S. meliloti* dimers. Red indicates negatively charged surfaces, white indicates neutral charged surfaces, and blue indicates positively charged surfaces. **(B).** Interfacing aspartate 316 residues of the EHEC H7 domain D4 dimer. The aspartate 316 residues are colored red. **(C).** Interfacing glutamate 224 residues

(red) of the *S. meliloti* domain D2 dimer. **(D).** The HIV-1 aspartic protease (1DMP). The catalytic residues are colored red.

The ODS and the intermediate waveform of the H7 filament might prolong tumbling time by creating an additional step in the canonical run-and-tumble scheme[3]. Upon switching to CW rotation, the EHEC H7 flagellar filament might first transition from normal to intermediate, then to semi-coiled, and finally to curly. Transitioning through an additional waveform might prolong the time spent in tumbling mode. Incomplete polymorphic transitions also occur during run-and-tumble motility[3,8,110]. We speculate that the additional intermediate waveform might not prolong tumbling in incomplete transitions when the motor rotation reverses back to CCW before the transition to curly[3]. This is consistent with our observation that cells with the H7 flagellar filament undergo quick tumbles similar to cells with the H48 filament. We observed fewer EHEC H7 and *Achromobacter* flagellar filaments that adopted the curly waveform, which could suggest that the curly waveforms are disfavored in ODS filaments.

The intermediate waveform of our ODS flagellar filaments is likely similar to the constricted normal waveform of flagellar filaments with screw-like surfaces[108,111], because both waveforms have shorter pitches and diameters than the normal waveform and are distinguishable from the semi-coiled and curly waveforms. The $Sn:Sn_{+5,}$ $Sn:Sn_{+6}$, and $Sn:Sn_{+11/-11}$ dimers and the 3-start helix formed on the surface likely allow for the formation of these unique shorter-pitch waveforms. However, without the additional $S_n:S_{n+27}$ or $S_n:S_{n+22}$ dimer, such waveforms might be rather unstable in *E. coli* cells with bidirectional flagellar rotation that require flagellar polymorphism. A different behavior is observed for the unidirectional flagellar rotation in *S. meliloti*, which does not require a change in flagellar waveform during run-tumble transitions[108]. The $S_n:S_{n+27}$

dimers of the EHEC H7 and *Achromobacter* flagellar filaments and the S_n:S_{n+22} dimers of the EPEC H6 flagellar filament may stabilize normal and intermediate waveforms. A weakened interaction in the EHEC H7 FF mutant flagellar filament with the disrupted S_n:S_{n+27} dimer site would explain why these mutant cells had drastically worse motility and seemed to constantly transition between waveforms.

Pathogens such as EHEC O157:H7 and EPEC O127:H6 must be able to swim through the intestinal mucus layers to reach the epithelium where they ultimately form attaching and effacing lesions[112-116]. The genus *Achromobacter* is found in water and soil[117] the latter which has a relatively high viscosity. Some *Achromobacter* species such as *A. xylosoxidans* and *A. ruhlandii* were also characterized as opportunistic pathogens in cystic fibrosis patients[118]. These two opportunistic pathogens have flagellins with high sequence homology in their outer domains with the ODS flagellins characterized in this manuscript and have the conserved S_n:S_{n+27} or S_n:S_{n+22} dimer interface (Fig. 3.10B). The common attribute between bacteria that produce these screw-like and ODS flagellar filaments is that they are all motile in environments that have higher viscosities than water. The mucosal environment is quite heterogenous and some pathogens prefer to swim through “gaps” in the mucus layers to reach the epithelium[119]. It has been shown that *E. coli* K-12 cells in porous environments exhibit long periods of being trapped followed by “hopping”[120]. The increased tumbling mediated by the outer domains of non-canonical flagellar filaments may allow for better reorientation in viscous and porous environments during a tumble, preventing the bacteria from being trapped while traversing these environments. These screw-like and ODS flagellar filaments are just one strategy that bacteria may use to cope with high-

viscosity environments as there are many bacteria with canonical flagellar filaments that inhabit these environments. The outer domains of canonical flagellar filaments likely provide other advantages in environments with higher viscosity. In *S. typhimurium*, for example, methylation of both phase-1 and phase 2 flagellins are implicated in adhesion and invasion into host cells[25].

In addition to their impact on motility, it has been shown that EHEC H7 and EPEC H6 ODS flagellar filaments function as adhesins[23,24], and ODS-forming H1, H6, and H7 flagellin monomers trigger increased TLR5 activation[26,27,121]. The flagellar filament mediated adhesion seems to be a general property of canonical and ODS flagellar filament outer domains as several studies show similar adhesive capabilities of both *S. typhimurium* and EHEC H7 flagellar filaments to the same surfaces[122]. The dimerization of the flagellin outer domains could also explain the increased TLR5 activation seen by the ODS-forming H1, H6, and H7 *E. coli* flagellins when they are bound to TLR5[26,27,121]. A crystal structure of the *S. typhimurium* flagellin FliC bound to TLR5 revealed a complex consisting of two TLR5:flagellin heterodimers[39] where domains D1 are bound by TLR5 and domains D2 of each flagellin are in close proximity (Fig. 3.18). For ODS-forming EHEC H7 and EPEC H6 flagellins these complexes could be stabilized by dimerization between the flagellin outer domains in the complex allowing for more robust TLR5-signalling. The outer domain dimers might also stabilize the flagellin subunits in detached flagellar filaments hindering the dissociation of the filament into monomers. A lower concentration of flagellin monomers might be advantageous during later stages as detached ODS flagellar filaments could possibly result in lower levels of TLR5 activation than canonical flagellar filaments.

Flagellin can be the single most abundant protein produced by bacteria and is thus under intense selection[123]. An example of such selection is the Macnab experiment: in stirred liquid culture, where flagellar motility provides no advantage, spontaneous mutations in flagellar genes lead to the complete loss of flagellar filament in only 10 days, because bacteria that fail to synthesize flagellin have a slightly increased fitness[124]. The outer domains required to produce ODS flagellar filaments are 50 to 100 amino acid residues larger than the *E. coli* K-12 and *S. typhimurium* flagellar filaments. Thus, producing these flagellar filaments with outer domain sheaths would be even more energetically costly than producing flagellar filaments without a sheath. Given the intense selective pressure on their synthesis, flagellin oligomerization and the subsequent effect on motility may be viewed as a mechanism for specific adaptations to environmental niches.

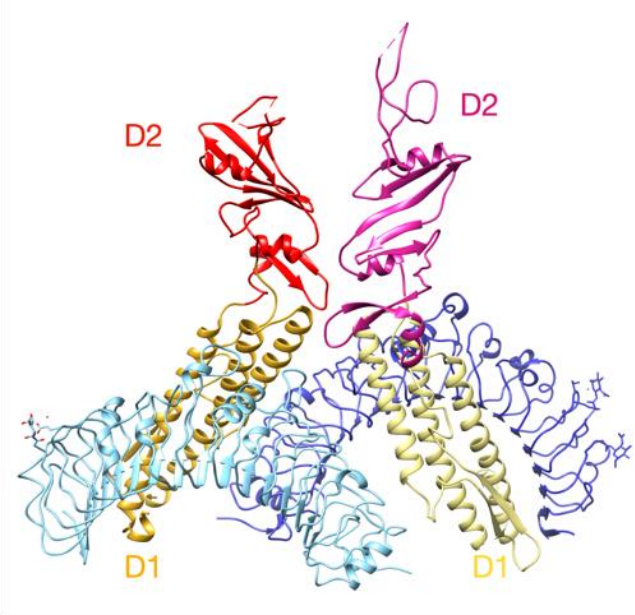

Figure 3.18. Binding of TLR5 to flagellin monomers. Crystal structure (3V47) showing a complex consisting of two TLR5:flagellin heterodimers that bind two each other[70]. The TLR5 ectodomains are colored blue. The *S. typhimurium* flagellins are colored gold or tan in domain D1 and red or violet in domain D2.

3.5 Methods

Generation of EHEC H7 flagellin mutants and EHEC and K-12 complementation plasmids

The H7 *fliC* gene was deleted from EHEC 86-24 using lambda Red recombination as published[125]. *fliC* was cloned into the pGEN-MCS vector using restriction digestion (NheI and HindIII) and ligation. Potential mutations were chosen based on changes in predicted stability calculated by the FoldX[126] and BindProfX[127] programs using the EHEC H7 flagellin outer domain structure in both monomer and dimer forms as input. Mutations were made in the pGEN-fliC vector using the Q5 mutagenesis kit according to the manufacturer's instructions.

For *fliC* complementation using native *fliC* promoters, *fliC* genes were amplified from genomic DNA using a primer pair that included a 280-bp region upstream of the *fliC* translation start site to generate 2,050 bp (EHEC) and 1,789 bp (K-12) fragments, respectively. Fragments were digested and cloned into the EcoRI and SalI sites of the pTrc99a plasmid.

Deletion, mutations, and complementation plasmids were all verified by Sanger sequencing.

EHEC O157:H7 soft agar motility assays

EHEC strain 86-24 was grown aerobically in LB overnight at 37 °C with shaking. The next day, cultures were diluted 1:100 into fresh LB and grown to mid-exponential phase (OD_{600} 0.4-0.6). From these, a 1 µL aliquot of culture was stab inoculated into motility plates (LB with 0.3 % agar). Plates were incubated for 7-24 hours at 37 °C and halo diameter was measured.

Phase contrast swimming behavior assays

EHEC strains were diluted from overnight LB broth cultures into fresh media with appropriate antibiotics and incubated at 37 °C in a roller drum to an OD_{600} of approximately 1.5. *E. coli* K-12 strains were incubated in tryptone broth at 37 °C in a roller drum to an OD_{600} of approximately 0.5. Cells were pelleted in round-bottom tubes for six minutes at 3,000 x g, the spent media removed, and cells were suspended in an equal volume of motility buffer (10 mM potassium phosphate, 10 mM lactate, and 70 mM sodium chloride, pH 7.0) in a roller drum at 20 rpm for approximately 20 minutes. The centrifugation and suspension steps were repeated once prior to visualization Videos were recorded at 30 fps using a Nikon Eclipse E600 phase contrast microscope equipped with a custom Nikon camera from Imaging Source. Swimming velocities were quantified using TumbleScore scripts written in MATLAB[128]. Tumble durations were measured manually using frame-by-frame playback of videos. Reported values are averages and standard deviations of at least 30 independent tumbling events.

Western blotting

Bacteria were grown and washed as described for swimming velocity and tumbling analyses. At each stage of the washing procedure, 10 µl of the corresponding sample was mixed with an equal volume of Laemmli buffer containing β-mercaptoethanol and samples were boiled for ten minutes prior to tank blot transfer as previously described[129]. Blots were incubated in blocking buffer consisting of PBS with 5% milk for at least one hour at room temperature. Anti-FliC polyclonal antibody serum and horseradish peroxidase linked goat anti-rabbit antibody were used at 1:10,000. Chemiluminescence signals were detected on ECL Hyperfilm.

Fluorescence labeling of flagellar filaments

Flagellar filaments were labeled using established methods[8] using the Alexa Fluor™ 546 NHS Ester Protein Labeling Kit from ThermoFisher Scientific. Overnight cultures of EHEC H7 and *Achromobacter sp*. MFA1 R4 cells were checked for motility using the DIC mode on Zeiss LSM 880. These cells were then pelleted at low centrifugation speeds of 1.2k x g for 5 minutes and washed 3 times by resuspension in 1x PBS and pelleting at the same low speed for 5 minutes. The final suspension volume of cells in 1x PBS was 500 µL and 50 uL of 1M bicarbonate was added to the suspension. This mixture was then transferred to the tube containing the AlexaFluor probe and rotated for 1 hour in the dark at room temperature. For *E. coli* K12 AW405 cells the steps were exactly the same however cells were grown to OD of 0.6 prior to labeling.

Fluorescence confocal microscopy

Fluorescently labeled cells were imaged using a laser scanning confocal microscope (Zeiss LSM 880) with a 40x water immersion objective (N.A. 1.2, Zeiss). The sample was excited at 561 nm and the emission was collected between 566 and 679 nm. For *E. coli* K-12 and *Achromobacter* sp. MFA1 R4 cells images of 512x512 pixels (53.14 µm x 53.14 µm) were collected with intervals of 240 milliseconds taken between each frame. For the experiments with the WT EHEC H7 and FF mutants, images of 256 x 256 pixels encompassing a larger physical area (83.78 um x 83.78 um) were obtained with 210 milliseconds between each image. Just prior to imaging, cells were diluted into 1:50 LB with or without 10^{-4} % tween.

Flagellar filament preparation

H7 flagellar filaments were prepared as published[23]. EPEC ICC526 (EPEC O127:H6 ΔBfpA + ΔEspA) was transformed with *flhDC* plasmid. To purify flagella, exponentially growing EPEC ICC526 culture (OD_{600} 0.6) carrying pflhDC was induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 3 h. Bacteria were then pelleted by centrifugation at 7000 x g and the pellet was suspended in 500 µl of cold, 1 M Tris/HCl, 100 mM NaCl buffer (pH 6.5). Bacteria were deflagellated by passing multiple times through a 25G needle until viscosity decreased. The deflagellated cells were pelleted by centrifugation at 10,000 x g for 15 minutes. The flagella in the supernatant were centrifuged at 10,000 x g at 4 °C to remove small debris. The resulting pellet was suspended in 500 µl of buffer. Purity of the flagella was analyzed by SDS PAGE.

Sinorhizobium meliloti flagellar filament purification was done essentially as described[96] with slight modifications. *S. meliloti* wild-type strain RU11/001 grown in TYC (0.5% tryptone, 0.3% yeast extract and 0.13% $CaCl_2$ x 6 H2O) at 30 °C was diluted to an OD_{600} of 0.05 in RB (6.1 mM K_2HPO_4, 3.9 mM KH_2PO_4, 1 mM $MgSO_4$, 1 mM $(NH_4)_2SO_4$, 0.1 mM $CaCl_2$, 0.1 mM NaCl, 0.01 mM Na_2MoO_4, 0.001 mM $FeSO_4$, 2 µg/l biotin). Twenty Bromfield plates (0.5% tryptone, 0.3% yeast extract and 0.13% $CaCl_2$ x 6 H2O) were over-laid with 10 mL of the diluted culture and grown for 15 hours at 30 °C to an OD_{600} of 0.6. Cells were harvested by centrifugation for 8 minutes at 8,000 x g and suspended in 100 mL motility buffer (0.5 mM $CaCl_2$, 0,1 mM EDTA, 20 mM HEPES [pH7.2]). Flagella were sheared by agitation in a mixer at full speed for 20 seconds, separated from cells by centrifugation at 8,000 x g for 7 min, 15,000 x g for 15 min, and 30,000 x g for 50 min at 4° C. Purified flagella were sedimented at 87,000 x g for 2 hours at 4° C, washed once

and suspended in 200 µL motility buffer. Purity of the flagella was analyzed by SDS PAGE.

Cryo-EM preparation

Flagella samples were prepared for cryoEM using established protocols[14,33]. Plunge freezing was done using a Vitrobot™ Mark II plunge freezer. Briefly, 3-4.5 µL of flagellar filament sample was applied to a lacey carbon grid. The droplet was blotted for 3.5 seconds with the blot force settings ranging from 3-6 and then plunged into liquid ethane.

Cryo-EM image acquisition

The image acquisition settings were the same as those previously published[33], with data acquisition using either EPU (ThermoFisher Scientific) or cryoSPARC.

Image processing and helical reconstruction

For the *Achromobacter* sp. MFA1 R4 and EHEC H7 flagellar filaments image processing and helical reconstruction were performed as published[14,33] using the Spider[130] and Relion 3[70] software packages. Motion correction was performed using MotionCorr2[66].

The *S. meliloti* and EPEC H6 image processing and structural determination were performed using cryoSPARC[131]. Images of *S. meliloti* and EPEC H6 flagellar filaments were motion corrected using the Patch Motion Correction process and contrast transfer function (CTF) estimation was done using Patch CTF Estimation. Initial subsets of flagellar filament segments (500-1000) were manually picked and underwent 2D classification. Selected class averages were then used as inputs for both Template Picker and Filament Tracer. More quality flagellar filament images were selected using

the Template Tracer function then Filament Tracer. The desired minimum separation between particles was determined based on the expected axial rise of the helical symmetry to be used during reconstruction. For *S. meliloti* a featureless cylindrical volume was used as the starting volume for reconstruction and an initial resolution reached 4.0 Å, and the helical symmetry converged to that shown in Table 1. Local CTF Refinement followed by helical reconstruction of these CTF refined particles resulted in a final volume with 3.5 Å resolution using the FSC=0.143 map:map criterion. For EPEC H6 the core domains were reconstructed with a monomeric symmetry and a resolution of ~4.5 Å was achieved. However, due to a seam in the outer domains, helical reconstruction could not reach high resolution for the full filament.

Asymmetric reconstruction of the EPEC H6 flagellar filament

The 4.5 Å resolution EPEC H6 core domain volume from helical reconstruction was filtered to 25 Å resolution and used as an input into the cryoSparc Homogenous Refinement program. A curved reconstruction of the H6 flagellar filament was obtained with core domain resolution around 3.5 Å and outer domains around 6 Å from ~270,000 particles. 3D variability analysis was performed, and the output was clustered. The particles were then subjected to CTF refinement. Using local refinement and a masking out of the core domains we achieved 4.3 Å resolution (FSC=0.143 map:map) for the EPEC H6 outer domains. A large portion of the outer domains on the inner most curved side of the reconstruction had very poor density likely due to misalignment of particles. Using cryoSPARC's 3D Variability analysis and solving for 3 different modes of heterogeneity we obtained several clusters (particles and their corresponding volumes) with differences in the overall quality of the outer domains as well as curvature. A single

cluster with ~60,000 particles showed low resolution (~8 Å) features corresponding to the outer domain dimers without the issue of non-uniform density in the inner curved region. Analysis of this region showed an apparent seam. Attempts to use the particles and volume from this 3D variability cluster for a reconstruction with extensive rotation and shift searches resulted in a final volume with no seam and very poor density in the corresponding inner curve. Local refinement of the reconstruction from these 60,000 particles starting from the cluster low-pass filtered to 12 Å and a mask around the outer domains resulted in a volume that retained the seam in the outer domain sheath with estimated 6.3 Å resolution for the outer domains (FSC=0.143 map:map).

Model building

Model building for the flagellar filaments in this study was as published for other flagellar filaments[14,132]. The corresponding density for an individual flagellin was isolated from the filament reconstruction using UCSF Chimera[71]. The D0/D1 model of an already deposited flagellin (PDB 5WJY) was fit into the corresponding region of each subunit's density map and then the residues were mutated to the appropriate ones for each structure using Coot[133]. For the outer domains the chains were traced manually in Coot and then refined using Rosetta CM[73]. Filament models were then generated in UCSF Chimera and refined and validated in Phenix[134] using real-space refinement[135] and real-space validation. For the H6 flagellar filament a homology model was initially generated using Swiss-Model[136]. This was subsequently refined in coot and much of the model was manually rebuilt using coot and the refined in phenix. For refinement of the H6 model to the high resolution map the model was fit into only the regions of good outer domain density in the map.

Achromobacter sp. MFA1 R4 structural determination

The *Achromobacter* sp. flagellar filament was a contaminant of a pili prep from *Agrobacterium tumefaciens.* Surprisingly, cryoEM images of the *A. tumefaciens* preparation revealed no pili at all but two very different flagellar filaments: thin flagellar filaments ~140 Å in diameter (yellow triangle Fig. 1c) with a canonical flagella power spectrum (Supplementary Fig. 1a), and much larger flagellar filaments (~220 Å diameter) with outer domain sheaths (red arrow Fig. 1c) and a tetrameric flagella power spectrum (Supplementary Fig. 1d). SDS-PAGE followed by silver staining was as published[62] and identified numerous gel bands arising from many contaminants. Multiple bands were excised from the gel and submitted for mass spectrometry analysis at the University of Virginia Biomolecular Analysis Facility.

The gel pieces from the band were transferred to a siliconized tube and washed in 200 µL 50% methanol. The gel pieces were dehydrated in acetonitrile, rehydrated in 30 µL of 10 mM dithiothreitol (DTT) in 0.1 M ammonium bicarbonate and reduced at room temperature for 0.5 hours. The DTT solution was removed, and the sample alkylated in 30 µL 50 mM iodoacetamide in 0.1 M ammonium bicarbonate at room temperature for 0.5 h. The reagent was removed, and the gel pieces dehydrated in 100 µL acetonitrile. The acetonitrile was removed, and the gel pieces rehydrated in 100 µL 0.1 M ammonium bicarbonate. The pieces were dehydrated in 100 µL acetonitrile, the acetonitrile removed, and the pieces completely dried by vacuum centrifugation. The gel pieces were rehydrated in 20 ng/µL trypsin in 50 mM ammonium bicarbonate on ice for 30 minutes. Any excess enzyme solution was removed and 20 µL 50 mM ammonium bicarbonate added. The sample was digested overnight at 37 ^{o}C and the peptides extracted from the

polyacrylamide in a 100 µL aliquot of 50% acetonitrile/5% formic acid. This extract was evaporated to 9 µL for MS analysis.

The LC-MS system consisted of a Thermo Electron Q Exactive HF mass spectrometer with an Easy Spray ion source connected to a Thermo 75 µm x 15 cm C18 Easy Spray column. 1-3 µL of the extract was injected and the peptides eluted from the column by an acetonitrile/0.1 M formic acid gradient at a flow rate of 0.3 µL/min over 1.0 hours. The nano spray ion source was operated at 1.9 kV. The digest was analyzed using the rapid switching capability of the instrument acquiring a full scan mass spectrum to determine peptide molecular weights followed by product ion spectra (10 HCD) to determine amino acid sequence in sequential scans.

Mass spectrometry identified the presence of peptides belonging to *A. tumefaciens* flagellin as well as potential peptide hits for the flagellin from the sheathed flagellar filament from various strains of *Achromobacter.* A BLAST search[137] of two peptide hits from the Mass Spec, FTANVRGLTQAAR and ISEQTDFNGVK, identified only two proteins in UniproKB with 100 % coverage in both sequences they were flagellin from *Achromobacter* sp. 2789STDY5608615 and *Achromobacter* sp. MFA1 R4. *Achromobacter* sp. 2789STDY5608615 was a partial sequence in the UniprotKB database missing a small part of its C-terminal D0 sequence and its outer domain sequence was about 15 amino acids too large, while the *Achromobacter* sp. MFA1 R4 was a better fit. Pure *Achromobacter* sp. MFA1 R4[88] cells were then obtained, and the flagellar filaments were confirmed to be sheathed using negative stain TEM.

Modeling of the *S. meliloti* flagellar filament

A sequence alignment of the four *S. meliloti* flagellins is shown in Supplementary Fig. 12. Both flagellin subunits in the asymmetric unit correspond to a flagellin that is ~395 amino acids in length which would exclude either subunit being FlaC. FlaA has Met300 which corresponds to Ala298 in FlaD. Both subunits in the asymmetric unit (ASU) show a large side chain density at that position (no dihedral symmetry has been imposed on the outer domains) which could correspond to a methionine but not alanine (Supplementary Fig. 13a). Thus, we can exclude FlaD from being one of the main components of the flagellar filament in our images. There are several regions where side chain density corresponding to the FlaA sequence would be difficult to explain with the FlaB sequence, such as those corresponding to FlaA residues Glu205 and Asn294 which correspond to Ala205 and Gly294 in FlaB (Supplementary Fig. 13b,c). Given that the densities of the two subunits in the ASU are nearly identical, it seems unlikely that the screw-like surface is formed specifically by a FlaA:FlaB heterodimer and most of the dimers must be formed by FlaA:FlaA homodimers. But it is still possible that some population of the segments used for reconstruction contained FlaB:FlaB homodimers or even FlaA:FlaB heterodimers that degrade the side chain density of residues like Glu205 and Asn294 when averaged with the predominant FlaA:FlaA dimers.

AlphaFold predictions

The *E. coli* K-12 flagellin model was predicted using the AlphaFold[138] option found in daily builds of UCSF ChimeraX[139].

Multiple sequence alignments

Figures for sequence alignment were made using Jalview-2[77]. All alignments were done using Clustal Omega[140].

Analysis of flagellar filament waveforms.

Movies of fluorescently labeled bacteria with attached flagellar filaments were analyzed in Fiji[141]. Rotating flagellar filaments were analyzed for polymorphic transitions. Each unique transition was measured and used to calculate the parameters in Fig. 7c and Supplementary Table 2.

3.6 Acknowledgements

Achromobacter sp. MFA1 R4 flagellar filaments were provided by Vince Conticello of Emory University. EHEC O157:H7 flagellar filaments were prepared by Jorge Giron at the University of Virginia. Additional EPEC O127:H6 flagellar filaments were prepared by Sharanya Chatterjee, Alejandro Pena, and Gadi Frankel. *S. meliloti* flagellar filaments were prepared by Birgit Scharf at Virginia Tech. Brooke Sauder and Melissa Kendall at the University of Virginia constructed the FF and RR mutants as well as performed soft agar motility assays, as well as grew EHEC cells for analysis by negative stain TEM and cryoEM. Richard Sobe from Virginia Tech performed western blot analysis and the phase contrast microscopy experiments to measure the velocity and tumbling of the various *E. coli* strains.

Chapter 4: Convergent evolution in the supercoiling of prokaryotic flagellar filaments

This chapter is submitted and under review at this time as

Kreutzberger, MAB et al.

4.1 Summary

Supercoiling of bacterial and archaeal flagellar filaments is required for motility, as the rotation of a straight or flexible filament will not generate thrust. The question of how a bacterial flagellar filament, composed of thousands of copies of identical subunits, can supercoil was addressed almost 50 years ago with two-state models for polymorphic switching in filaments comprised of 11 protofilaments, with different flagellar waveforms attributed to different mixtures of these two states. However, no direct visualization of mixed populations of protofilaments has ever been made. Archaeal flagellar filaments have no homology to bacterial ones and are instead homologs of bacterial Type IV pili. The bacterial model cannot simply be applied to the archaeal filaments, and no models exist for archaeal flagellar supercoiling. Great advances in cryo-electron microscopy (cryo-EM) make it now possible to directly see the basis for supercoiling, and here we show atomic structures of supercoiled bacterial and archaeal flagellar filaments. For the bacterial flagellar filament, we identified 11 protofilament conformations, rather than two. The archaeal flagellar filament is assembled from 10 protofilaments with 10 subunit conformations, with a seam on the inside of the curve. Our results reveal a process of convergent evolution where analogous functions have been achieved by completely different paths and mechanisms.

4.2 Introduction

Many organisms from the three domains of life, Archaea, Bacteria, and Eukarya, possess flagella which allow for motility[35]. The three systems evolved completely independently of one another with different components yet similar functions. This is an example of convergent evolution, just as bees, birds and bats have independently evolved wings for flight. Eukaryotic flagella and cilia are composed of microtubules, and ATP hydrolysis by dynein motors causes the appendages to beat back and forth[35]. Bacterial and archaeal flagella convey motility through the motor-driven rotation of supercoiled flagellar filaments which function as propellers[2,20,35,36]. While the eukaryotic flagella and cilia are quite complex and may contain as many as 600 different proteins[142], bacterial and archaeal flagellar filaments are both polymerized from a single kind of subunit, bacterial or archaeal flagellin.

The bacterial flagellum has been studied extensively[1,20] and the supercoiling of its flagellar filament and its ability to switch between supercoiled states (polymorphism) are key to the flagellar motility of bacteria[3,9]. The polymorphism of the flagellar filament allows for the switching between various waveforms under a mechanical force[9,143,144]. For the bacterial flagellar filament, the core domains D0 and D1 are key to supercoiling and polymorphism[17,19] while the outer domains can have diverse structures or be largely absent[21,99,103]. In some cases, the outer domains can dimerize or tetramerize to alter the motility of bacteria by allowing for different waveforms to be favorable[99].

Almost all structural studies of the bacterial flagellar filament have been performed using straight non-motile flagellar filaments produced by point mutations in

the flagellin subunit[14,21,95,145-147]. These point mutations lock the flagellar filament in one of two different kinds of helical lattices with either left-handed (L-type) or right-handed (R-type) protofilaments[15]. These straight filaments have strict helical symmetry, which was necessary at the time for structural studies using either x-ray fiber diffraction[146] or cryo-EM[31,32]. It was also observed that different kinds of mutant flagellins which only formed straight filaments on their own could polymerize into supercoiled flagella when mixed together in the presence of a filament seed[148]. These observations, combined with the all L- and all R-state structures, resulted in the development of a two-state model for the supercoiling of the bacterial flagellar filament[16,143,149] that has dominated thinking in this area for nearly 50 years. However, the supercoiling this model explains is of wild-type flagellar filaments, and until very recently[22,99,150] it was thought that high-resolution structural studies of such filaments were impossible. For many years the accepted model for the supercoiling of the bacterial flagellar hook, the universal joint connecting the flagellar filament to the motor, was also a two state model[151]. This was shown to be incorrect with asymmetric reconstructions of the supercoiled flagellar hook[152,153] where 11 different flagellar hook protein states were identified[152]. Complicating matters, recent high-resolution structures of wild-type supercoiled bacterial flagellar filaments were achieved using helical reconstruction[22,99,150], despite the fact that the subunits in a supercoiled flagellar filament cannot be related to each other by helical symmetry.

While the bacterial flagellum has been extensively studied, much of the work on the archaeal system is more recent[35,154-157]. The archaeal flagellar filament and archaeal flagellin have no homology to their bacterial analogs and are instead homologous to

bacterial and archaeal type IV pili (T4P)[158-160]. Despite their similarities, archaeal flagellar filaments must be able to supercoil[154] while T4P do not. Recently, it was suggested that the flexibility of the archaeal flagellar filament allowed for motility[156]. This cannot be true because rotation of a straight or flexible filament will not produce thrust, and the flagellar filament must be supercoiled. There have been no structural models for the supercoiling of archaeal flagella.

Here we show using near-atomic resolution cryo-EM structures of three supercoiled bacterial flagellar filaments from two different species of *Escherichia coli* (*E. coli*), Enteropathogenic *E. coli* O127:H6 (EPEC H6) and *E. coli* K-12. For the EPEC H6 we determined the supercoiled structure for two different waveforms, Normal and Curly I. For *E. coli* K-12 we solved the structure in the Normal state. We focus on only the core domain structures of these filaments as those are key to supercoiling. We show that there are 11 flagellin and therefore 11 protofilament states in both the Normal and Curly I waveform structures, which is contrary to the two-state model but matches recent results for the flagellar hook[152,153]. We also solve the cryo-EM structure of the supercoiled archaeal flagellar filament from *S. islandicus* Rey 15A and show that there are 10 protofilaments in the structure, each with a different conformation. As with the bacterial filaments, the supercoiling is governed by a core domain, containing a single α-helix in archaea compared with the coiled-coils in bacteria. In contrast to the bacterial filaments, a discrete seam exists between protofilaments on the inside of the archaeal supercoil. For both bacterial and archaeal flagellar filaments we are able to extrapolate microscopic atomic models built into short segments to correctly generate the

macroscopic properties of the supercoil, such as the pitch and diameter, verifying the model.

4.3 Results

Cryo-EM structures of curved flagellar filaments in the Curly I and Normal waveforms

We used cryo-EM to determine the structure of supercoiled bacterial flagellar filaments from EPEC O127:H6 and *E. coli* K-12. For this work, we focused on the core domains D0 and D1 and masked out the outer domains from the EPEC H6 reconstruction. While we have previously shown that the outer domains in H6 form tetramers that are dimers of dimers[99], the pairings of the outer domains do not introduce any perturbation of the core domains D0 and D1. The H6 flagellar were imaged under two different buffer conditions which results in differences curvature and supercoil pitch (Fig 4.1A, and 4.2A). We initially performed helical reconstruction refining a single global helical symmetry relating all subunits to each other (Figs 4.1, 4.2 and 4.3). In all cases, we were able to reconstruct the flagellar core domains D0 and D1 to ~4.0 Å resolution or higher. For each helical structure, the 0.143 map:map “gold standard” FSC (GS FSC) resolution estimations appeared to over-estimate the actual resolution of our maps (Figs 4.1-4.3). This is likely due to the fact that GS FSC is a measure of reproducibility and not of true resolution, particularly for structures with imposed[161]symmetry .

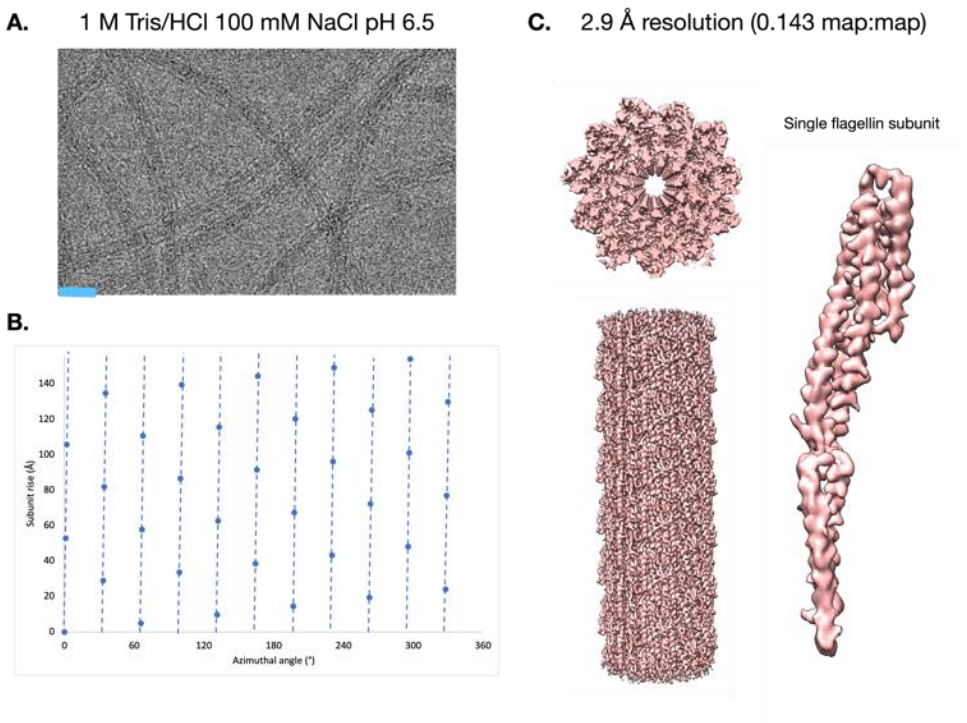

Figure 4.1. Helical reconstruction of the EPEC H6 flagellar core with Curly monomeric symmetry. (A). Micrograph of EPEC H6 flagellar filaments. **(B).** Helical net created from the final monomeric symmetry parameters of axial rise of 4.81 Å and helical twist of 65.54°. We use the convention that the surface has been unrolled and we are looking from the outside. The blue dashed lines represent the 11 protofilaments which have a tilt of + 0.94°/52.9 Å, where 52.9=11*4.81. This tilt is a twist of 2π radians/2.0 μm, or 3.1 μm^{-1}. **(C).** Density map from the monomeric symmetry helical reconstruction.

The average protofilament tilt for the various bacterial flagellar waveforms has previously been predicted from the number of L- and R-states expected to be present in the structure[16]. From this prediction, we expected EPEC H6 flagellar filaments with higher curvature in the micrographs (Fig 4.4A) to be in the Curly waveform. This is predicted due to the observed protofilament tilt of +0.94°/52.9 Å rise along a protofilament in the helical reconstruction (Fig S4.1B). The lower curvature EPEC H6 and *E. coli* K-12 flagellar filaments (Fig 4) were predicted to be in the Normal waveform due to their having a protofilament tilt of -0.6°/53.1 Å rise along a protofilament (Fig 4.2B, 4.3B). The flagellin densities in both domains D0 and D1 for these helical

structures were surprisingly poor for the GS FSC estimated resolution (Fig 4.2C, 4.2C, 4.3C). We hypothesized that this was due to the fact that we were imposing helical symmetry on structures that were not truly helical, but because the same symmetry is being imposed on both half-maps, it is reproducible and yields an overly optimistic estimate of resolution.

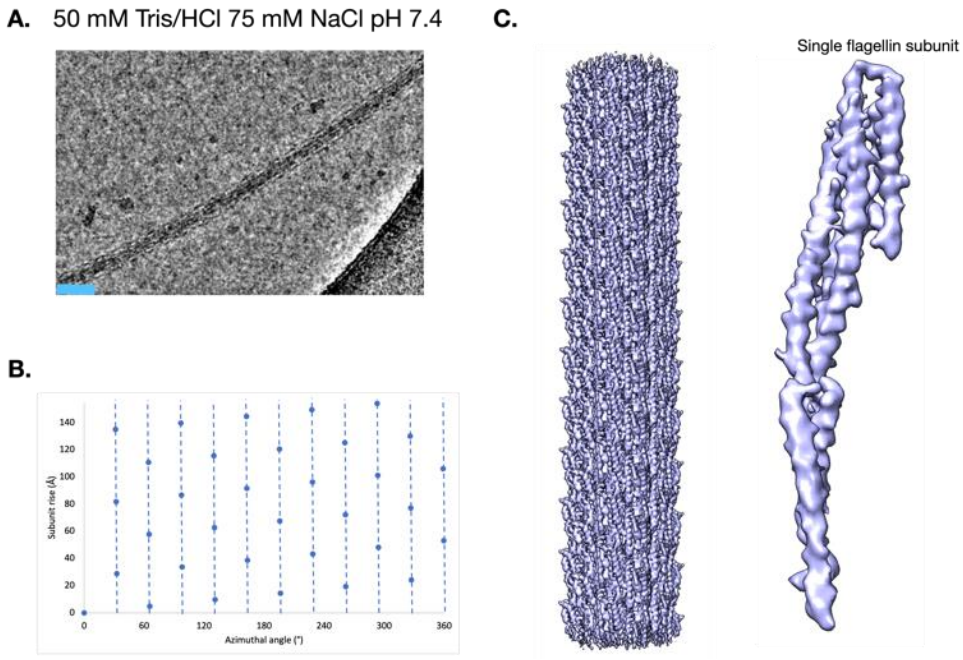

Figure 4.2. Helical reconstruction of the EPEC H6 flagellar core with Normal monomeric symmetry. (A). Electron micrograph of EPEC H6 flagellar filaments. **(B).** Helical net of from the final symmetry parameters of 4.82 Å axial rise and 65.403° helical twist, which is -0.567°/53.0 Å, or -1.9 μm^{-1}. **(C).** Density map from the helical reconstruction of the EPEC H6 flagellar filament in the normal waveform.

We then devised a scheme to reconstruct the bacterial flagellar filament asymmetrically (Fig 4.5A). From this reconstruction we were able to solve the curved flagellar filament structures of bacterial flagellar filaments in a higher curvature state or a lower curvature state (Fig 4.4). Based on both the helical symmetry as well as the 2.9 rad/µm curvature estimated from our map we are confident that the higher curvature EPEC H6 structure (Fig 4.4B and Fig 4.5B) is the Curly I waveform. Likewise, based on

the average protofilament tilt and curvature of ~1-1.4 rad/µm, the lower curvature EPEC H6 and *E. coli* K-12 curved structures (Fig 4.4E and Fig 4.5C-D) are of the Normal waveform.

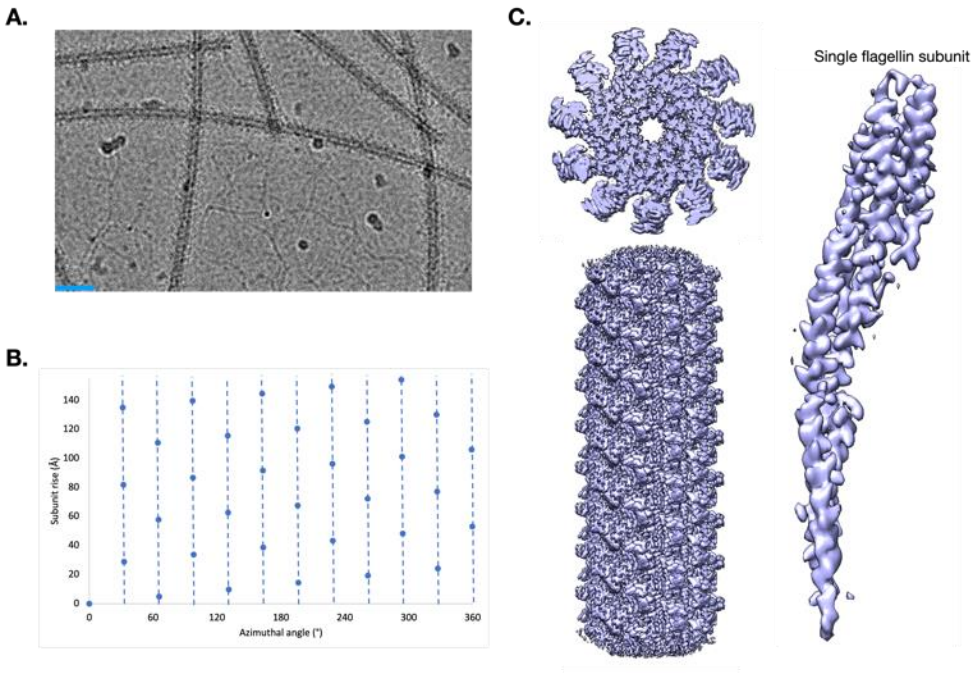

Figure 4.3. Helical reconstruction of the *E. coli* K-12 flagellar core with Normal monomeric symmetry. (A). Cryo electron micrograph of *E. coli* K-12 flagellar filaments. **(B).** Helical net for the *E. coli* K-12 flagellar filament. The axial rise is 4.83 Å and the twist is 65.406° which is -0.534°/53.1 Å or -1.8 μm^{-1}. **(C).** Density map of the *E. coli* K-12 monomeric symmetry helical reconstruction. Views of the filament are shown on the left. A density map of a single flagellin subunit is shown on the right.

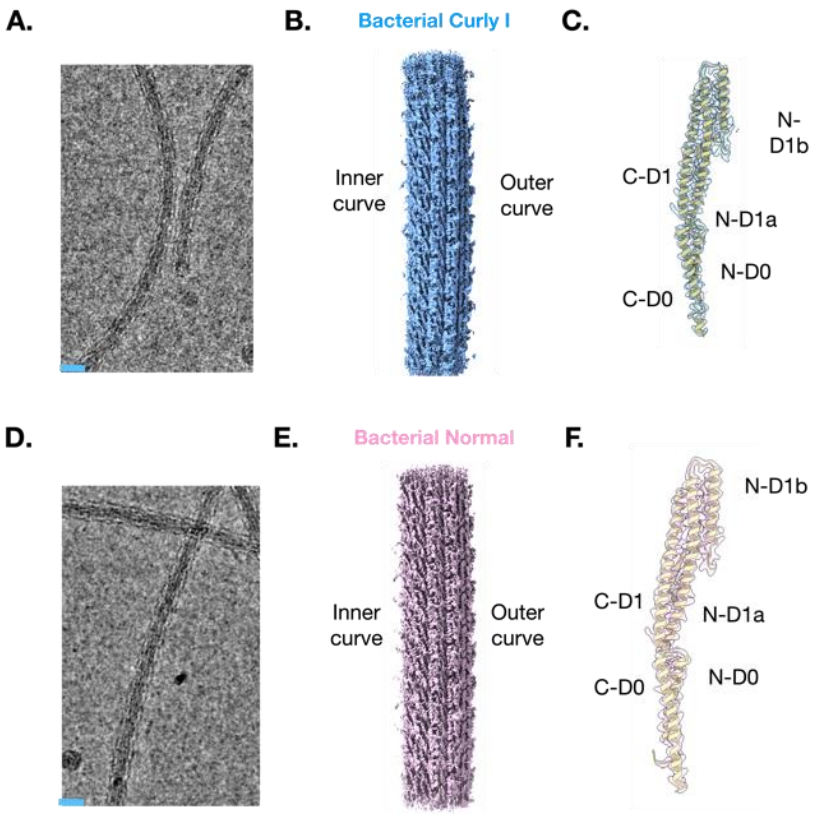

Figure 4.4. Cryo-EM structures of the bacterial flagellar filament core in the Curly I and Normal waveforms. (A). Cryo-electron micrograph of the EPEC H6 flagellar filaments with high degree of curvature. **(B).** Cryo-EM reconstruction of the H6 flagellar filament in the Curly I waveform state. **(C).** The 3.4 Å density map of a single Curly I subunit (light blue) with a model built in (gold). **(D).** Cryo-electron micrograph of the EPEC H6 flagellar filaments with less curvature. **(E).** The 3.9 Å cryo-EM reconstruction of the H6 flagellar core in the Normal waveform state. **(F).** The 3.2 Å density map of the core domains for a single *E. coli* K-12 flagellin in the Normal waveform state.

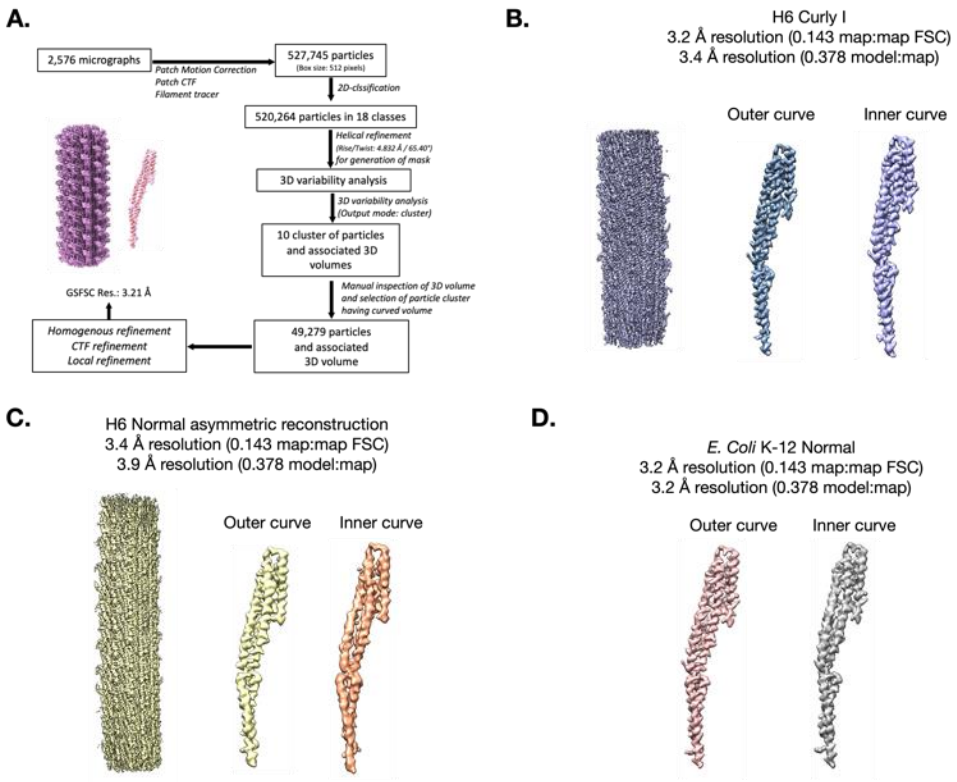

Figure 4.5: Asymmetric reconstruction of bacterial flagellar filaments. (A). Reconstruction scheme for generating a *E. coli* K-12 Normal flagellar filament map. **(B).** Density maps for inner and outer curve EPEC H6 normal flagellins. **(C).** Density maps for inner and outer curve EPEC H6 Normal flagellins. **(D).** Density maps for inner and outer curve *E. coli* K-12 flagellins.

We compared the subunit density from our EPEC H6 Curly I helical and asymmetric reconstructions to the nominal 2.8 Å resolution helically symmetric flagellar filament structure from *Firmicutes Kurthia*[22]. Both the *Kurthia* and EPEC H6 helical structures' density maps had discontinuous density in the loop region connecting the two helices of N-terminal domain D1 (Fig 4.6A-B). The flagellin subunit densities in the asymmetric Curly I structure did not have broken density in this loop (Fig 4.6C) and generally yielded a higher quality map despite having a lower resolution estimated by the "gold-standard" 0.143 map:map FSC. A similar observation has been made in the structure determination of microtubules, where the correct symmetry led to a lower resolution

FSC but a better map than an incorrect symmery[162]. We built models into our flagellar filament structures (Fig 4.4C and 4.4Ff) and used a map:model FSC threshold of 0.378 (which is √0.143) for our resolution estimate. The curved Curly I flagellar filaments had a better resolution of 3.4 Å than its corresponding helical reconstruction when using the 0.378 map:model criterion (Fig 4.6). We found a similar improvement in quality when we analyzed the Normal waveform asymmetric reconstructions from EPEC H6 and *E. coli* K-12, which had resolution estimates of 3.9 Å and 3.2 Å, respectively (Fig 4.5C and 4.5D).

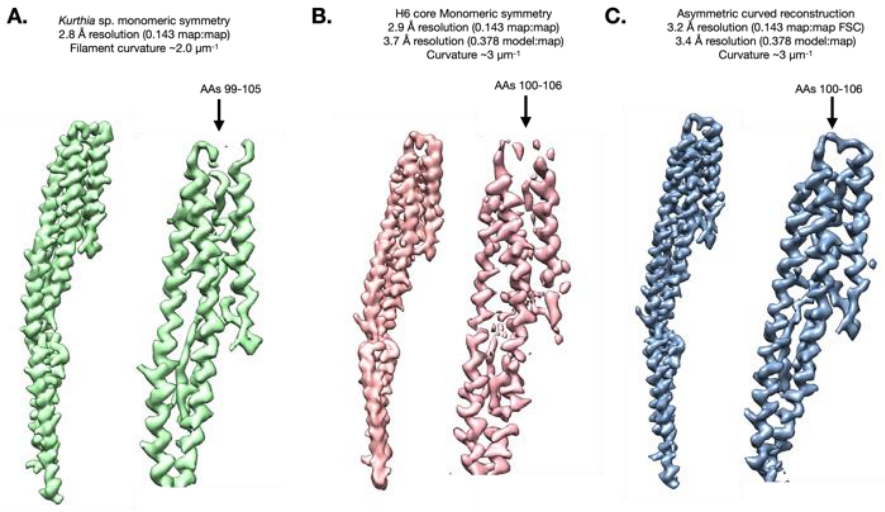

Figure 4.6. Comparison of *Kurthia* and EPEC H6 monomeric symmetry subunit density maps with the EPEC H6 asymmetric reconstruction subunit density. (A). The density map for a single *Kurthia* flagellin subunit. The left image shows the full flagellin and the right shows a close up of domain D1 density at high threshold. Broken density in domain D1 is indicated by the arrows. **(B).** Density map for flagellin core from the H6 helical reconstruction. The full domain D0/D1 core is shown on left. On the right is an image showing the broken density in domain D1 at high threshold. **(C).** Density map for flagellin core from the EPEC H6 asymmetric reconstruction. The left image shows the full D0/D1 core density. The density in the loop region (AAs 100-106) appears quite good and not broken in contrast to the helical reconstructions.

Building a model for supercoiled flagellar filaments

Since the diameter of supercoiled waveforms (~2,000-6,000 Å) is significantly larger than typical ice thickness in a cryo-EM grid[163], we suspected that the supercoiling of the flagellar filaments would be flattened in the z-dimension. To confirm this, we used cryo-electron tomography on wild-type *Campylobacter jejuni* flagellar filaments (Fig 4.7), which showed that the flagellar filaments were mostly flattened in the z-dimension (Fig 4.7D).

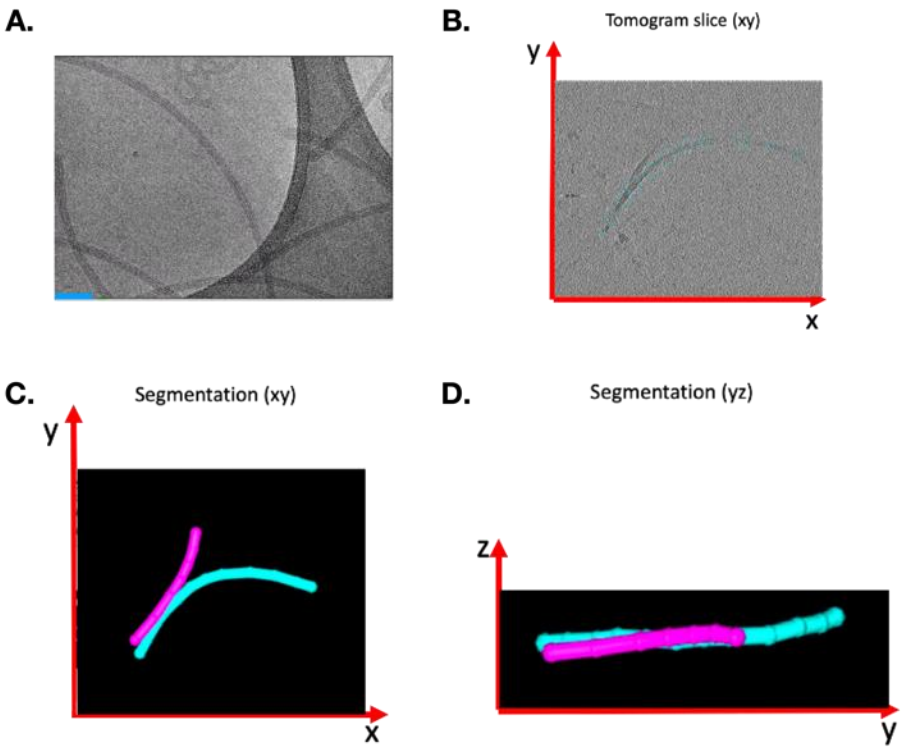

Figure 4.7. Cryo-electron tomography of the wild type *C. jejuni* flagellar filament. (A). Electron micrograph of wild type *C. jejuni* flagellar filaments. The scalebar is ~50 nm. (**B).** Tomogram slice of *C. jejuni* flagellar filaments. **(C).** Segmentation of the flagellar filaments showing them in the x- and y-dimensions. **(D).** Segmentation of the flagellar filaments in the y- and z-dimensions.

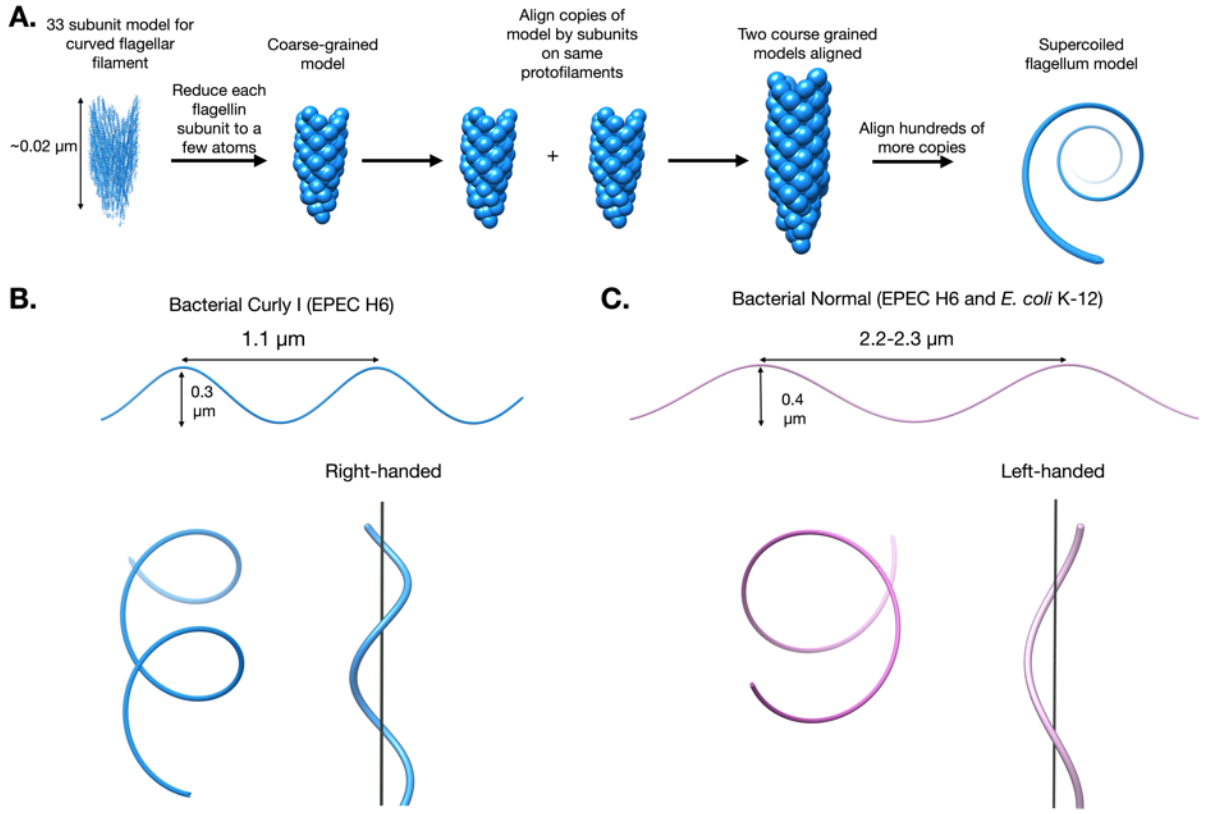

Figure 4.8. Supercoiled model of the bacterial flagellar filament. (A). Scheme for modeling the 3-dimensional supercoil of prokaryotic flagellar filaments starting from an atomic model built from cryo-EM data. For generation of the coarse-grained models several Cα atoms from several residues in each of the flagellin domains were selected while all other atoms were removed. These Cα atoms were then displayed with a much larger atomic radius, so the models were easier to visualize. **(B).** The supercoiled model of the Curly I waveform of the bacterial flagellar filament. The top image shows the "side-view of the filament" where the pitch and helix diameter are easily seen. The bottom left image shows a view "through" the supercoil. The bottom right image shows the right-handedness of the Curly I waveform with the helical axis of the supercoil represented by the black line. The 1.1 µm pitch, 0.3 µm diameter, and right-handedness of the model match published values. **(C).** Supercoiled model of the Normal bacterial flagellar waveform. The EPEC H6 model had a pitch of 2.3 µm and diameter of 0.4 µm. The *E. coli* K-12 model had a pitch of 2.2 µm and a diameter of 0.4 µm. The left-hand of the Normal supercoil is shown in the bottom right image. The black line represented the helical axis of the supercoil. These values as well as the left-handedness are consistent with published values.

It has previously been reasoned that bacterial flagellar filaments that are flattened by negative stain TEM preparative conditions retain their local curvature[143]. It has been suggested that filaments undergo a large twist at select locations to

accommodate the flattening without severe distortions elsewhere[143]. We anticipated that the same was true for flagellar filament structures that were reconstructed from very small segments (~0.04-0.07 μm) relative to the overall length of the flagellar filament (2-10 μm). To test this, we developed a method for aligning hundreds of copies of flagellar filament atomic models to generate a macroscopic model (Fig 4.8 A). We aligned models stepwise so that subunits from the same protofilament on the ends of two models were aligned). Using this scheme, we modeled the supercoil for the Curly I waveform (Figure 4.8 B)with a pitch of 1.1 μm and a diameter of 0.3 μm[8,164] for the Curly I atomic model. These parameters and the waveform's right-handedness[8] are consistent with the literature. We obtained a similar result for modeling the supercoil of the Normal structures (Figure 4.8C). For EPEC H6 the pitch was 2.3 μm and diameter was 0.4 μm. The *E. coli* K-12 supercoil model produced a similar result where the pitch was 2.2 μm and the diameter was 0.4 μm. Both of the Normal supercoils were left-handed. These Normal waveform parameters also agreed nicely with the literature[8,164]. These results show that despite the flattening of the overall supercoil in ice, the flagellar filament structures retained the local curvature and twist of a flagellum supercoiling in three-dimensions. Thus, the microscopic, short-curved bacterial flagellar filament structures were indeed representative of the supercoiled state with uniform curvature in three dimensions.

Supercoiling of the bacterial flagellar filament is created by 11 flagellin conformations

When assessing the flagellin conformations we used subunits on the inner- and outer-most regions of the curve as references (Fig 4.9A). When we aligned subunits from each of the 11 protofilaments by domain D0 or D1 we identified 11 flagellin conformations (Fig 4.9B-C). Curly I subunits aligned by D0 showed distinct differences between all 11 subunits (Fig 4.9B). Differences were also present for the Normal waveform subunits, but they were generally more similar to each other than the Curly I subunits (Fig 4.9 C). Importantly, we found that subunits from the same protofilament had nearly identical conformations when aligned by domain D0 compared to subunits from adjacent protofilaments. We concluded that the small differences between subunits on the same protofilament are cause by noise and/or modeling error.

We then compared the protofilaments of our Curly I and Normal waveforms by aligning each of the individual 11 protofilaments by the first subunit (S_n) of each protofilament model (Fig 4.9D). From this alignment we identified 11 protofilament states by examining the "top" view of subunit S_{n+22} from each protofilament. Differences between each of the Curly I protofilaments (Fig 4.9E were greater than those of the Normal (Fig 4.9F). Nevertheless, the protofilaments in both waveforms differed from each other according to the same circular pattern (Fig 4.9 E-F) where adjacent protofilaments were more similar to each other than non-adjacent ones. This pattern can be more easily explained considering their position in the supercoiled structure (Fig 4.9G) the same amino acid residues face the center of the pore in each structure. However, the tilt of each protofilament with respect to the center of the flagellar filament is different. Each of the protofilaments is "bending" into the curve (Fig 4.9H-I). The protofilament on the inner most curvature (blue in Fig 4.9G-I) is tilting away from the

center of central channel of the filament. The outer most protofilament (pink) is tilting towards the center and the protofilaments with intermediate distance between the outer and inner curves (gold protofilaments) are tilting normal to the filaments central axis (Fig 4.9I)

Historically, supercoiled flagellar filaments have been thought to be the product of a mixture of L- and R-states, where adjacent protofilaments segregate into one or the other state. This would generate discontinuities or seams at the two junctions between the L protofilaments and the R protofilaments. Instead, we find that each protofilament is in its own unique state, and there are no seams or discontinuities. However, for all three bacterial structures we discovered that the inner curve protofilaments were shorter than the outer curve ones, if we aligned each of the interfaces between adjacent protofilaments (5-start interface) in domain D1, (Fig 4.10A). For the Curly I waveform, we identified five longer protofilaments in the outer curve, four shorter ones in the inner curve, and two intermediate ones. These intermediate protofilaments were positioned in between the shorter and longer protofilaments (Fig 4.10B-C). Generally, the inner curve protofilaments were positioned downward by ~2.5 Å along this interface when compared to the outer curve ones, whereas the intermediate ones were displaced by ~1 Å. Protofilaments could be distinguished into three length (z-dimension) classes ,while each protofilament was unique in the x- and y-dimensions (Fig 4.10D). Shearing along the 5-start interface between the inner and outer curve protofilaments was ~1.6 Å for the two normal waveforms. Three shorter 5-start interfaces were found with no clear intermediates (Fig 4.11).

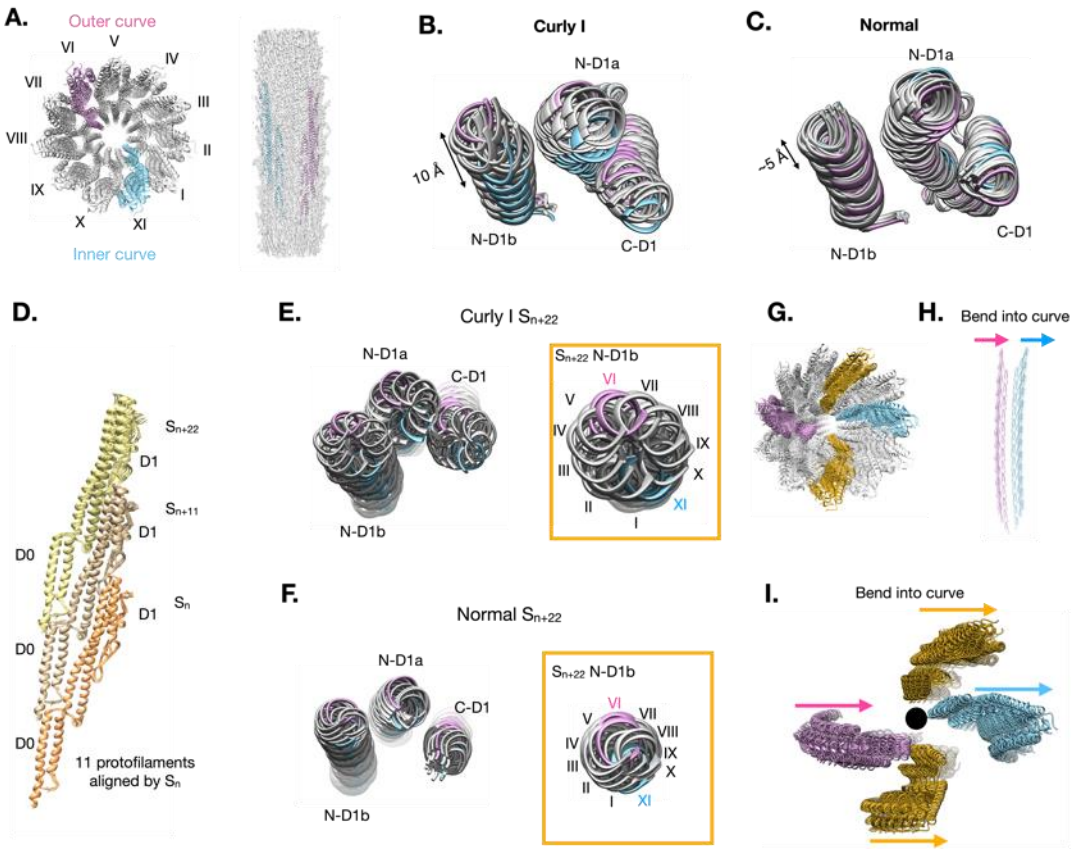

Figure 4.9. Subunit and protofilament states in the bacterial flagellar filament. (A). Top and side views of the flagellar core model (left) and map (right) with inner and outer protofilaments colored blue and pink respectively. **(B).** Alignment of Curly I flagellin subunits from each protofilament by domain D0. The double arrow represents the 10 Å end to end distance between an inner curve protofilament (light blue) and an outer curve protofilament (pink). **(C).** Alignment of Normal flagellin subunits from each protofilament by domain D0. The double arrow represents the 10 Å end to end distance between an inner curve protofilament (light blue) and an outer curve protofilament (pink). **(D).** Alignment of each of the 11 Curly 1 protofilaments by subunit S_n. **(E).** View of subunit S_{n+22} of each of the 11 Curly I protofilaments which have been aligned by subunit S_n. The orange inset shows a close up of domain N-D1b of S_{n+22} from this alignment. The pink and blue protofilaments correspond to protofilaments in the outer and inner curves respectively. **(F).** View of subunit S_{n+22} of each of the 11 Normal waveform (K-12) protofilaments which have been aligned by subunit S_n. The orange inset shows a close up of domain N-D1b of S_{n+22} from this alignment. **(G).** Top-down view of the Curly I filament model show inner (blue) and outer (pink) curve protofilaments as well as protofilaments in between the inner and outer most regions of the curve. **(H).** Side vies of the inner (light blue) and outer (pink) protofilaments. Both protofilaments bend into the curve. However, the outer curve protofilament tilts inward toward the flagellar lumen while the inner curve protofilament is bending away from the lumen. **(I).** Same view as in **G** but with only the colored protofilaments shown. The arrows indicate the direction in which the protofilaments bend.

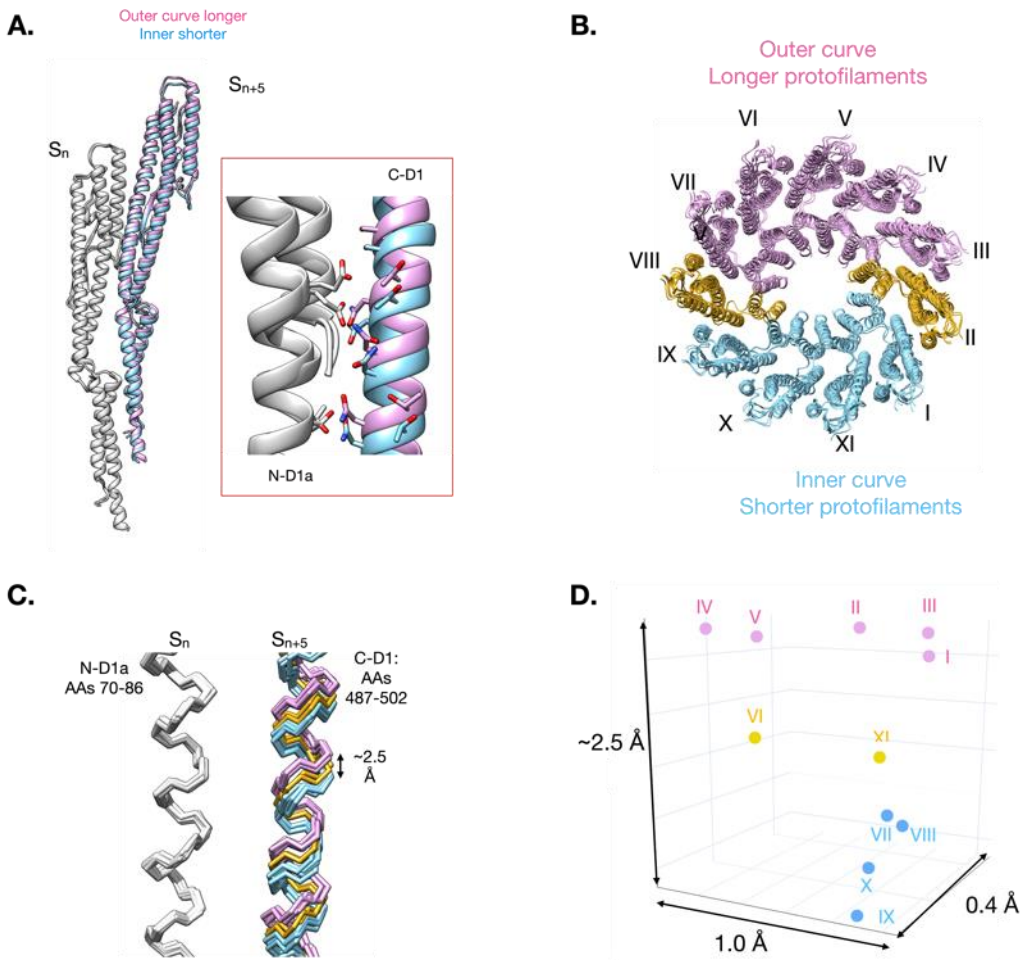

Figure 4.10. Differences between inner and outer curve protofilaments of the Curly I waveform.
(A). Alignment of the 5-start interface between adjacent protofilaments . The inner and outer curve interfaces are aligned by subunit S_n. From this, differences between the subunit S_{n+5} can be seen between inner and outer curve protofilaments. The red inset shows a closer view of the region of interest in domain D1. The differences are apparent between amino acids 487-503 in domain Cd1 of subunit S_{n+5}. **(B).** The three classes of protofilaments in the Curly I flagellar filament. Longer protofilaments on the outer curve are shown in pink while shorter ones on the inner curve are shown in blue. **(C).** The same alignment in **A** but for each of the 11 curly I protofilaments. The protofilaments are all aligned by N-D1a of subunit S_n. Longer protofilaments are pink, shorter are blue, intermediate are gold. **(D).** Three-dimensional plot of a single Cα atom from the alignment shown in **C.**

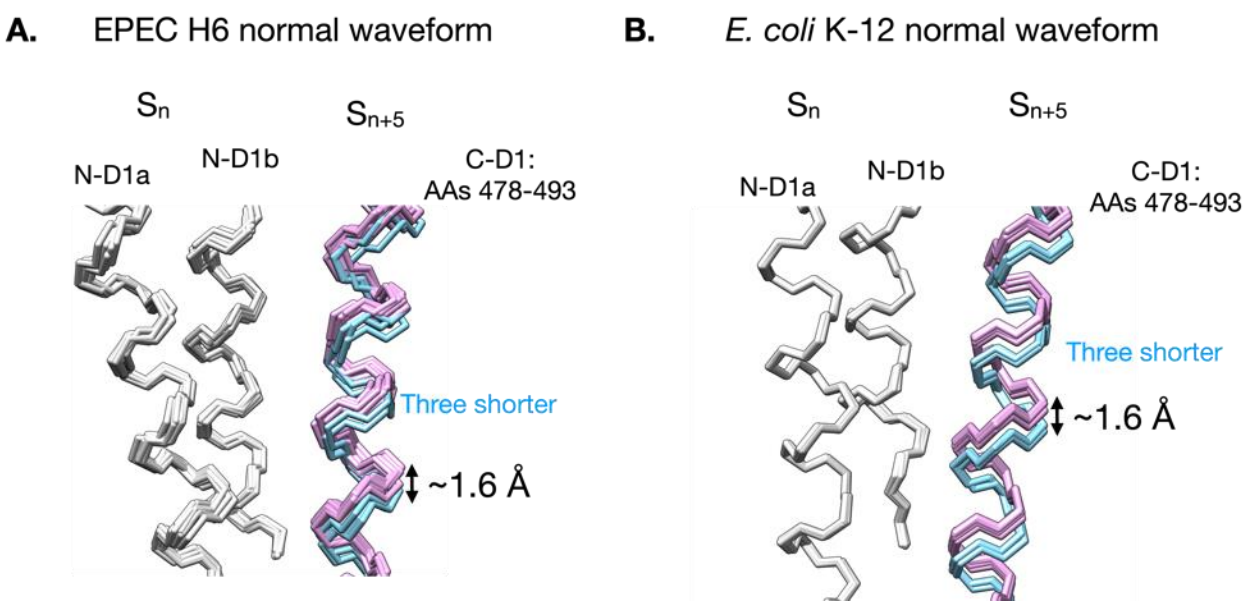

Figure 4.11. **5-start interface between adjacent protofilaments of the EPEC H6 and *E. coli* K-12 normal flagellar filament structures. (A).** 5-start interface between N-D1 domains of subunit, S_n and adjacent subunit S_{n+5} for the EPEC H6 normal flagellar filament structure. Shorter protofilaments are colored blue, longer ones are colored pink. The ~1.6 Å shearing between the shortest and longest protofilament is indicated. **(B),** 5-start interface between N-D1 domains of subunit, S_n and adjacent subunit S_{n+5} for the *E. coli* K-12 normal flagellar filament structure. Shorter protofilaments are colored blue, longer ones are colored pink. The ~1.6 Å shearing between the shortest and longest protofilament is indicated.

.

Cryo-EM structure of the supercoiled archaeal flagellar filament

We next sought to apply our method for reconstructing supercoiled flagellar filaments to archaeal flagellar filaments from *Sulfolobus islandicus* Rey15A. Both negative stain and cryo-electron micrographs revealed two distinct filament populations (Fig 4.12A-B). These were either supercoiled and identified as archaeal flagellar filaments or not supercoiled and identified as archaeal type IV pili (T4P). Differences in the surface structures of flagellar filaments and T4P were also observed in cryo-EM images (Fig 4.12B). Using the same workflow as used for the bacterial flagellar filament reconstruction, we solved the structure of the *S. islandicus* Rey15A flagellar filament imposing a global helical symmetry, which resulted in a density map at ~4 Å resolution

(Fig 4.13A). We then reconstructed the *S. islandicus* flagellar filament asymmetrically and obtained a 3.4 Å resolution map with clear curvature (Fig 4.12C) and modeled the flagellin subunits into the map (Fig 4.12D). From this short asymmetric reconstruction, we estimated the curvature to be ~2 radians/μm. Like many archaeal flagellar filaments and T4P[62,156,157,165], we detected extensive glycosylation with several N-linked glycans on the surface of the structure (Fig 4.14). We also found indication of tyrosine glycosylation (Fig 4.14C-D). To our knowledge, this would be the first evidence of tyrosine glycosylation in *S. islandicus*, although this type of modification has been detected for prokaryotes[166]. It should be noted that there is only one flagellin encoded in the *S. islandicus* Rey15A genome, in contrast to other archaea that encode and express multiple flagellins[156,167].

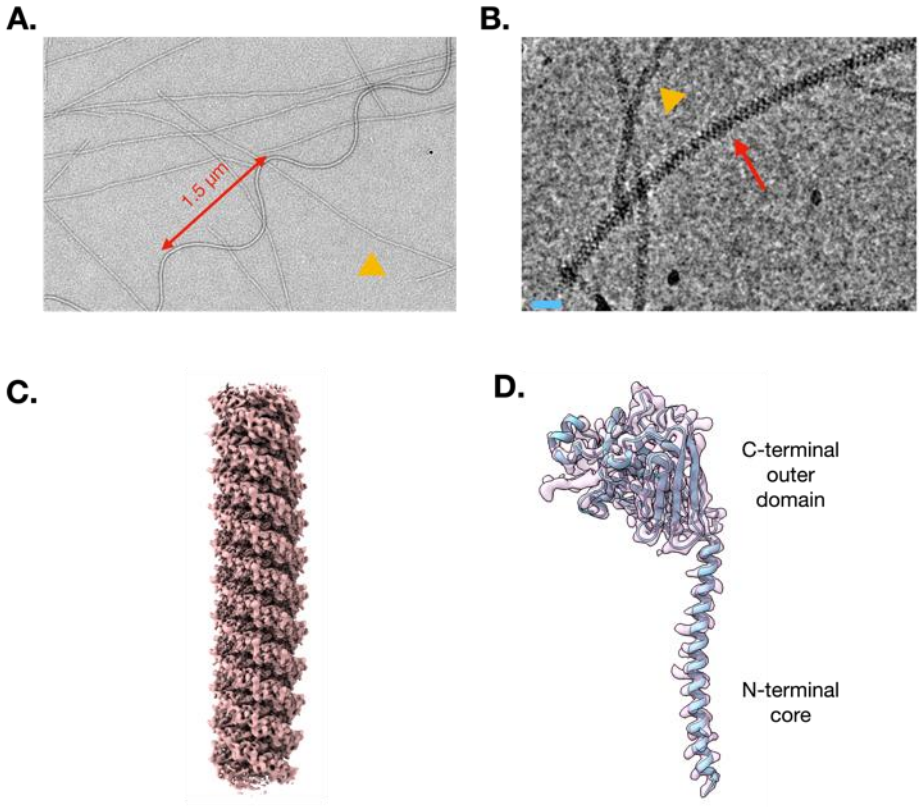

Figure 4.12. Cryo-EM structure and atomic model of the supercoiled *S. islandicus* Rey15A archaeal flagellar filament. (A). Negative stain electron micrograph of the *S. islandicus* Rey15A archaeal flagellar filament. The 1.5 µm pitch of the flagellum is shown with the red double arrow. An archaeal type IV pilus (T4P) is indicated with the yellow arrowhead. **(B).** Cryo-electron micrograph of the *S. islandicus* Rey15A flagellar filament (red arrow). The yellow arrowhead indicates an archaeal T4P. **(C).** The 3.4 Å resolution asymmetric cryo-EM reconstruction of the curved Rey15A archaeal flagellar filament. A surface view of the structure is shown on the left while a top view is shown on the right. **(D).** Density map (light pink) and model (light blue) of a single Rey15A archaeal flagellin.

The N-terminal regions of the preprocessed forms of archaeal flagellins and bacterial and archaeal type IV pilins exist as transmembrane domains[159]. After cleavage these mature subunits assemble into their respective polymers in the absence of a membrane. Interestingly, TMHMM[168] analysis of transmembrane regions for the Rey15A flagellin predicted additional transmembrane helices in the C-terminal outer domain (Fig 4.15A). These regions are either β-sheets or loops in the flagellar filament structure (Fig 4.15B). We took the sequences for each of the predicted C-terminal

transmembrane helices and analyzed their structure using AlphaFold[138]. For both ~20 amino acid residues spanning regions the predicted models were α-helical (Fig 4.15C). As a control, a region predicted to not be transmembrane and that was part of a β-sheet in the filament was predicted to be a β-hairpin by AlphaFold (Fig 4.15D). We speculate that these regions of the C-terminal outer domain are indeed transmembrane helices in the preprocessed flagellin monomers and insert in the archaeal membrane prior to assembly.

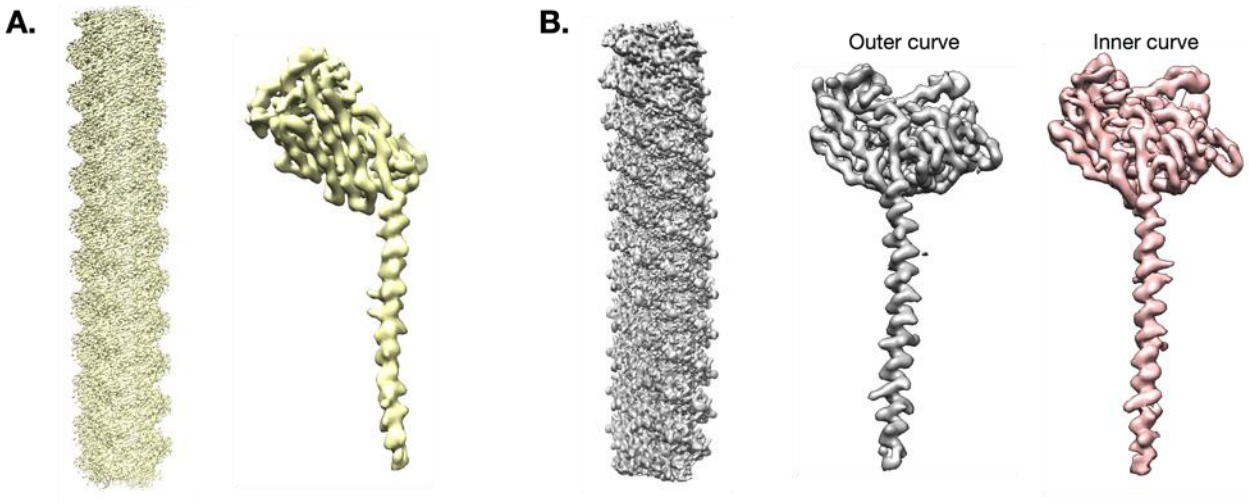

Figure 4.13. Density maps for the helical and asymmetric reconstructions of the *S. islandicus* Rey15A archaeal flagellar filament. (A). On the left is the helical reconstruction of the Rey15A flagellar filament. On the right is the density map for a single Rey15A subunit from the helical reconstruction. **(B).** The curved Rey15A flagellar filament asymmetric reconstruction is shown on the left and the density maps from that reconstruction for single subunits on the outer and inner curves are shown on the right.

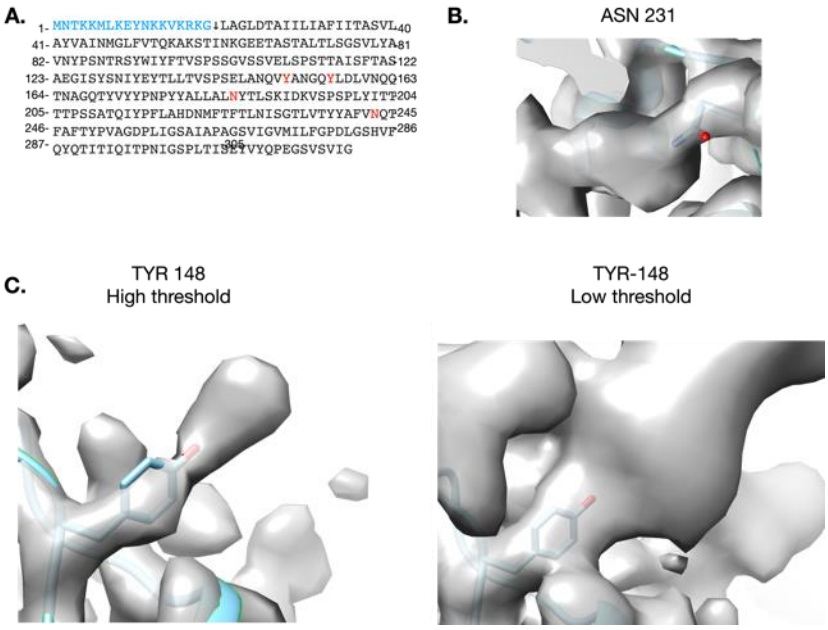

Figure 4.14 . Glycosylation of the Rey15A archaeal flagellar filament. (A). Sequence of the *S. islandicus* Rey15A archaeal flagellin. The N-terminal sequence removed in processing is shown in blue. Residues which appear glycosylated in the density map are colored red. **(B).** Glycosylation of asparagine 231. The residue is shown with its large glycan density. **(C).** Evidence of tyrosine glycosylation in the Rey15A flagellar filament. The left image shows Tyrosine 148 and its corresponding density at high threshold. The right image shows tyrosine 148 and its corresponding density at low threshold.

The supercoiled *S. islandicus archaeal flagellar filament* is assembled from 10 protofilaments each with a unique conformation

While 3-start helices composed of the outer domains are prominent on the surface of archaeal flagellar filaments[167,169,170], the Rey15A flagellar filament has a less prominent 10-start helix which. When comparing helical nets (Fig 4.16), the 10-start helix in archaea (Fig 4.16B) is reminiscent of the nearly vertical 11-start protofilaments of the bacterial flagellar filaments (Fig 4.16A). The 10-start interactions mainly occur for the N-terminal α-helical domain at a relatively small radius in the filament.

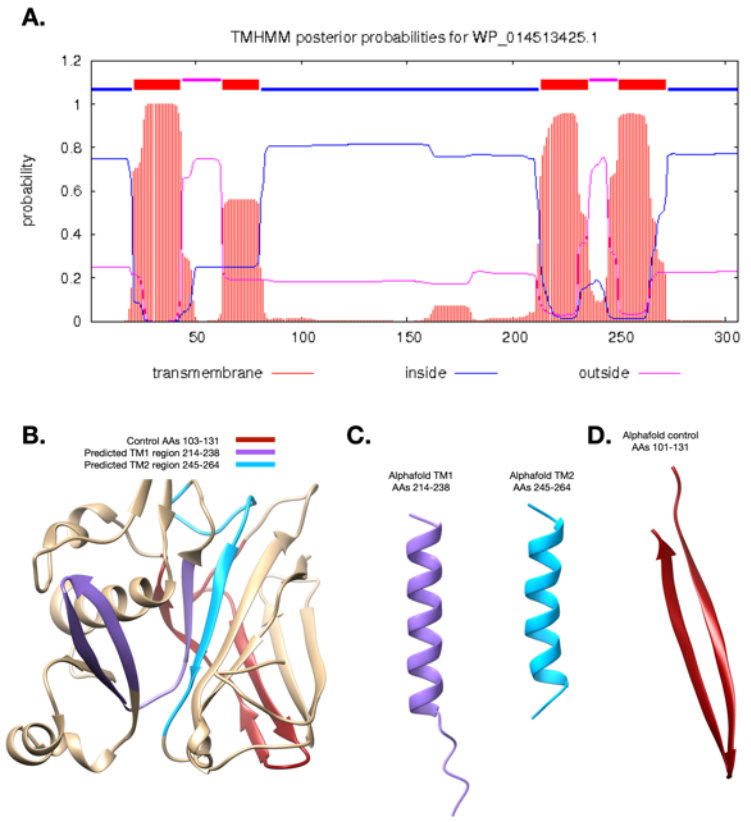

Figure S4.15. Putative transmembrane helices in the C-terminal domains of the preprocessed *S. islandicus* Rey15A archaeal flagellin. (A). TMHMM prediction of transmembrane helices in the Rey15A archaeal flagellin. **(B).** Model of the Rey15A C-terminal outer domain. The model was initially generated using AlphaFold prior to refinement. The regions in the model are colored according to the key at the top. **(C).** AlphaFold prediction of amino acids 214-238 (purple) and 245-264 (cyan). It should be noted that the extended loop at one of the ends of the AAs 214-238 prediction is close to being α-helical. **(D).** The AlphaFold predicted β-hairpin for control region of AAs 101-131 (dark red). This region is not predicted to be a transmembrane helix by TMHMM and is a β-hairpin in the filament model.

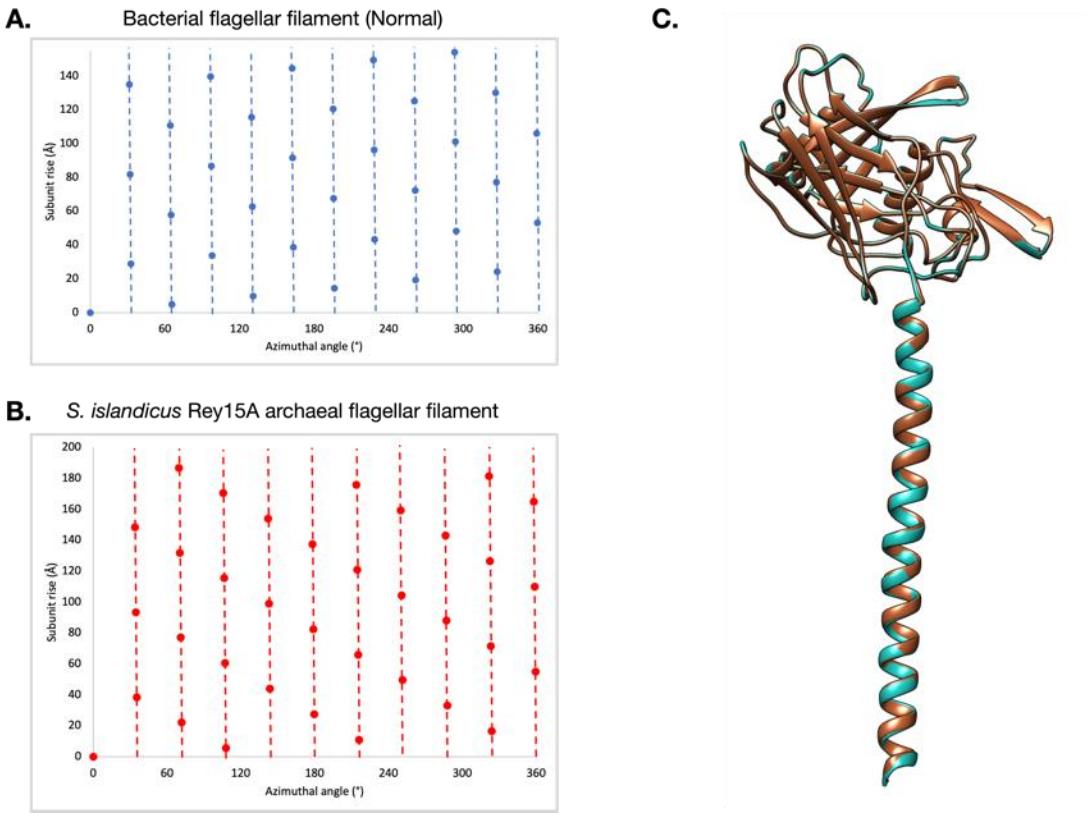

Figure 4.16. Helical net comparison of bacterial and archaeal flagellar filaments as well as alignment of subunit models from same Rey15A protofilament. (A). Helical net of the EPEC H6 normal waveform. The 11-start protofilaments are indicated by the blue dashed lines. **(B).** Helical net of the *S. islandicus* Rey15A archaeal flagellar filament. The 10-start protofilaments are indicated by the red dashed lines. **(C).** Alignment of two Rey15A archaeal flagellins from the same protofilament.

When we aligned a single subunit from each of these strands, there were 10 subunit states with one of these states being markedly different from any of the others (Fig 4.17A). We also found that subunits on the same 10-start strand were conformationally similar to each other (Fig 4.16C). This led us to conclude that the archaeal flagellar filaments are actually composed of 10 protofilaments (Fig 4.17B). We then aligned all of the 10 archaeal protofilaments by a subunit (S_n) and examined the differences that occur two subunits (S_{n+20}) up along the protofilaments (Fig 4.17C). Adjacent protofilaments were more similar to each other except for the two inner-most subunits in the curve (protofilaments I and IX Fig 4.17D). We then built a model for the

Rey15A archaeal flagellar filament in the same manner as for the bacterial filaments by aligning subunits from the same protofilament between two archaeal flagellar filament models (Fig 4.17E). The resulting supercoil parameters were in good agreement with our negative stain TEM observations (Fig 4.12A), and the waveform was left-handed.

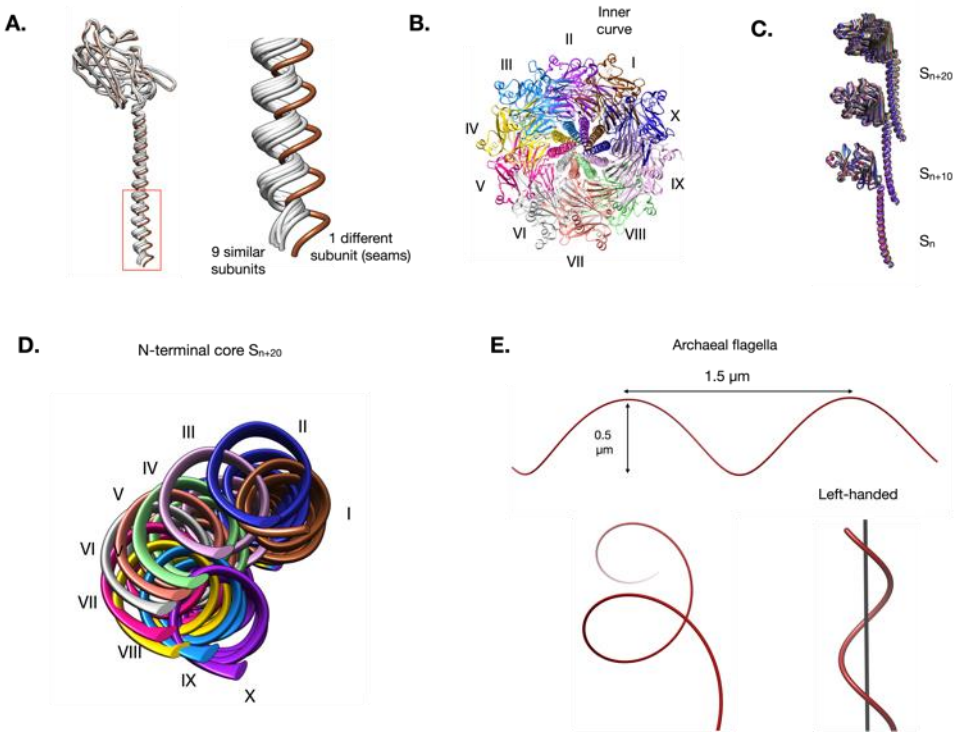

Figure 4.17. Subunit and protofilament conformations of the *S. islandicus* Rey15A archaeal flagellar filament. **(A).** Alignment of the 10 unique Rey15A flagellin subunit models. The left image shows the full flagellin models. The right image shows a close up of the red boxed region. There are 9 flagellin states that are very similar to each other (light grey) and a tenth state that is quite different (brown). **(B).** The 10 protofilaments of the archaeal flagellar filament. Each protofilament is a different color. **(C).** Alignment of the 10 archaeal flagellar protofilaments by subunit S_n. **(D).** Same alignment from c with a top view of the N-terminal core of S_{n+20}. This view stops at the end of the N-terminal helix (residue 65). **(E).** Supercoiled model of the *S. islandicus* Rey15A flagellar filament. The 1.5 µm pitch and 0.5 µm helix diameter agree with our negative stain TEM measurements. The supercoil is left-handed as shown in the bottom right image.

The *S. islandicus* Rey 15A archaeal flagellar filament contains two seams in its protofilament interfaces

The majority of the interactions between adjacent Rey15A protofilaments resemble what is shown in Fig 4.18A where isoleucine 26 of a subunit (S_{n+3}) is in close proximity to alanine 35 of a neighboring protofilament subunit (S_n). However, for the interface between two protofilaments labeled I and II (Fig 4.17B) in the inner region of the curve there is a seam where protofilament I is shifted ~4 Å down the interface (Fig 4.18B blue inset) when compared to the other nine standard interfaces. This results in isoleucine 26 of subunit S_{n+3} moving closer to isoleucine 32 of neighboring subunit S_n. These seam interactions do not occur in the outer domains (Fig 4.18A-B red insets), which are similar across all 3-start interfaces. The most conformationally different protofilament (I) is at the very inner region of the curve (Fig 4.18C), which is the location of the seam. We found that the major difference driving this ~4 Å shearing is the movement of the previously described archaeal flagellin "hinge"[156] (Fig 4.18D-E). This hinge is the linker between the N-terminal core and C-terminal outer domain and corresponds to threonine 66 and alanine 67. The hinge residues adopt different, but similar conformations in 9 protofilaments and a drastically different conformation in the subunits belonging to protofilament I (Fig 4.18D). The two most conformationally distinct subunits are thus from protofilaments I and II, which is caused by the movement of the hinge region downward in protofilament I relative to protofilament II (Fig 4.18E).

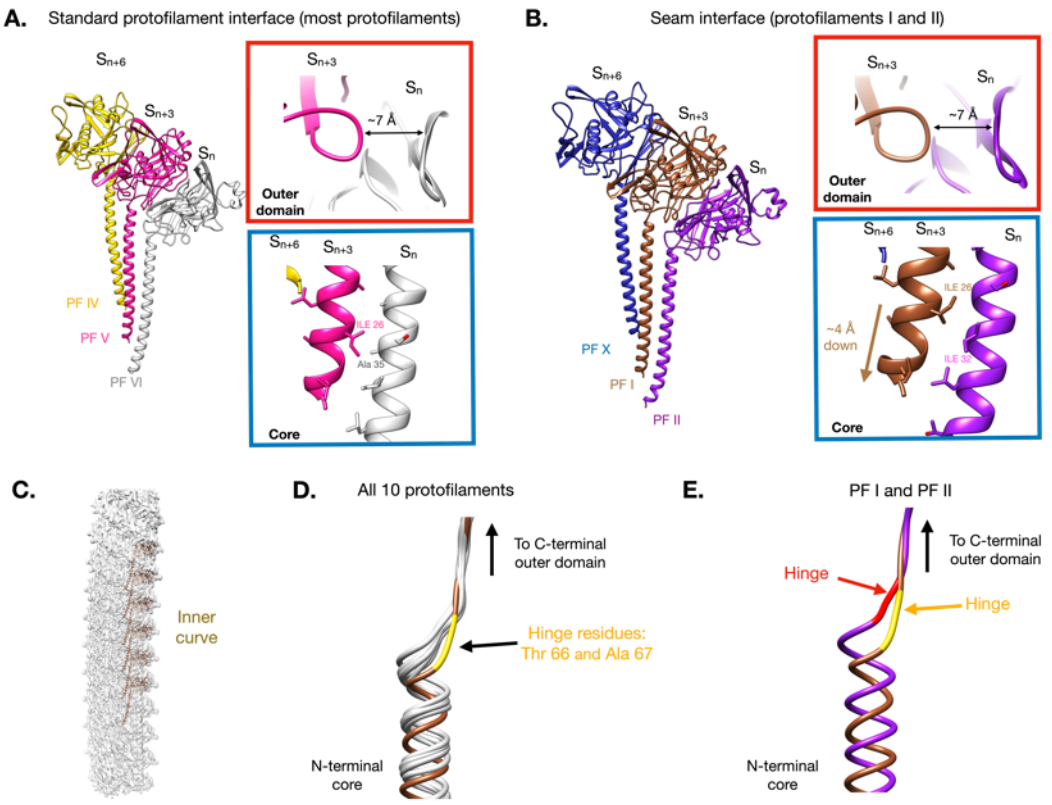

Figure 4.18. Subunit and protofilament interactions of the *S. islandicus* Rey15A flagellar filament. **(A). "**Standard" 3-start interface between Rey 15A protofilaments. Three adjacent subunits from three adjacent protofilaments are shown. They are colored as in Fig. 3f. The red inset shows the interaction between subunits S_{n+3} and S_n in their outer domains. A ~7 Å distance between the two backbones at this interface is indicated. The blue inset shows the interface between the subunits in the α-helical core. Here isoleucine 26 of subunit S_{n+3} is adjacent to neighboring alanine and not close to neighboring isoleucine 32. **(B).** The 3-start interface containing the seam subunit. The seam subunit is shown in brown as S_{n+3}. In the red inset a ~7 Å distance between the outer domain of the seam subunit (brown) and its standard neighbor (purple) is shown. In the blue inset the α-helical core of the seam protofilament (brown) has shifted along the 3-start interface by ~4 Å. This can be seen by the close proximity of isoleucine 26 of the seam subunit S_{n+3} to its neighbor subunit's isoleucine 32 (purple). **(C).** Density map (light grey) with the model for the seam subunit (brown) shown to be on the inner part of the curve. **(D).** View of the linker region between the N-terminal helical core and the C-terminal outer domains of all 11 archaeal flagellin conformations aligned together. **(E).** The same view as in d with only the two inner curve protofilaments numbered I (brown) and II (purple) shown. The linker hinge of residues 66-67 are colored gold and red for protofilaments I and II respectively.

4.4 Discussion

For the *S. islandicus* Rey15A archaeal flagellar filament, we determined the presence of 10 protofilaments, each having a different conformation. Initially, the 10

protofilaments in the Rey15A structure were difficult to identify, because the archaeal flagellar filament core is more compact than the bacterial one. In addition, the C-terminal outer domains make very few contacts along the 10-start helix. The presence of 10 protofilaments in the archaeal flagellar filament was speculated in a, previous a low-resolution structural study[160].

The Rey 15A flagellin has a larger C-terminal domain (~260 total residues) compared to other known archaeal flagellin structures[156,157,167]. This larger C-terminal domain results in extensive outer domain contacts, generating a very prominent 3-start helix on the surface of the filament (Fig 4.19A). The 3-start helix is even more noticeable when one examines the glycan density in the map (fig 4.19B) as the glycosylated residues form a ridged surface along the 3-start helix. This prominent 3-start helix creates a unique surface pattern, which somewhat resembles the screw-like surface of the *S. meliloti* bacterial flagellar filament[99]. It is feasible to speculate that the extensive outer domain 3-start interactions and glycosylations in the Rey15A archaeal flagellar filament is an adaptation to the natural habitat of *S. islandicus* by providing additional stability in hot and acidic environments. A similar degree of post-translational modification was observed for *S. islandicus* LAL14/1[167] T4P, when compared to non-thermoacidophilic archaeal T4P[165].

For the Rey15A archaeal flagellar filament, the linker hinge regions (residues 66-67) of all subunits exhibited a great degree of flexibility (Fig 4.18D). Strikingly, this hinge is also critical for the establishment of the seam between the interface of two inner curve protofilaments. Flexibility of this linker region has been reported for other archaeal flagellins[156] as well as T4P[165], however, the role that it plays in the supercoiling of the

archaeal flagellar filaments is unknown. Due to the supercoiling of the filament, the flexibility of the hinge region may allow for preservation of similar core domain contacts in 9 of the 10 interfaces and a seam in the 10th inner curvature interface (Fig 4.18).

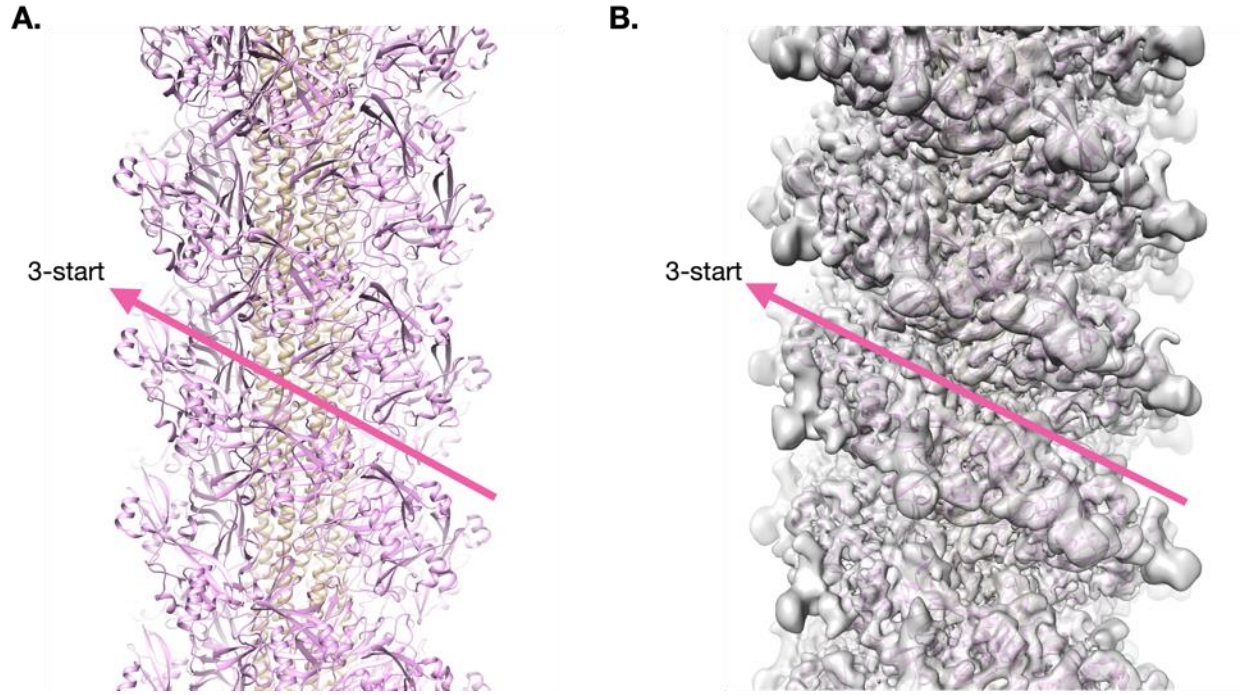

Figure 4.19. The 3-start helix on the surface of the Rey15A archaeal flagellar filament. (A). Surface of the Rey15A archaeal flagellar filament atomic model with the N-terminal core domains colored tan and the C-terminal outer domains colored pink. The 3-start helix formed by one set of these outer domain interactions is marked with the pink arrow. **(B).** The same view from **A** but with the density map (grey) shown as well. The extensive glycan densities make the 3-start helix even more pronounced.

While some archaea produce multiple flagellins that can assemble into an ordered pattern such as the recently described hexametric archaeal flagellar filament in *Methanocaldococcus villosus* [156], the *S. islandicus* Rey15A genome only encodes a single flagellin, which simplifies model building and interpretation. Altering buffer conditions such as pH and salt concentration change the waveforms of archaeal *Methanococcus voltae* flagella[154], but the visualization of rotating archaeal

Halobacterium salinarum flagellar filaments revealed little polymorphism or change in handedness upon motor rotational switching[36]. It is possible that the Rey15A flagellar filament is similar, and the seam coupled with the compact filament core hinder polymorphism of the archaeal flagellar filament. It should be noted that polymorphism and motor rotation display species-specific variations[4,171].Certain bacterial flagellar filaments exhibit no polymorphism at all during running and tumbling due to unidirectional rotation of the flagellar motor. Thus, the lack of flagellar filament polymorphism during swimming motility in one species of archaea does not exclude the possibility of polymorphism in others.

Like the archaeal flagellar filament, we found a single subunit conformation for each protofilament in the bacterial flagellar filament. This result contradicts the simple two-state model for flagellar filament protofilament switching[16,172]. There are several important differences between our work and previous studies leading to the two-state model. First, our structures are at much higher resolutions[14,95,146,147]. Second, our work was performed with wild type supercoiled flagellar filaments with no helical symmetry imposed, while previous studies were based on rather low-resolution structures of helically symmetric straight mutant flagellar filaments. As a consequence, we were able to take a short model of our wild-type flagellar filaments and - by combining many copies - built a supercoiled model for the flagellar filaments that matched the expected macroscopic pitch, diameter, and curvature for the supercoiled filaments.

If the supercoiling of bacterial flagella was the result of two discrete states, then strain would be exerted at the junctions between the two states[146,173]. This might lead to the situation where the all-L or all-R states were at an energy minimum. Since wild-type

filaments do not naturally exist in the all-L or all-R states, supercoiling might be a consequence of kinetic trapping that keeps these filaments away from these minimum-energy states. We devised a simple annealing experiment where supercoiled flagella (Fig 4.20A and 4.20C) were heated from room temperature to 55 °C and allowed to slowly cool back down to room temperature. If the L- and R-states were at minimum energy, then we expected to observe straight flagellar filaments. Instead, the bacterial flagellar filaments retained their supercoil (Fig 4.20B and 4.20D), suggesting that the supercoiled state is, in fact, the lower energy state.

Caspar[174] raised the possibility that the all-L or all-R states might not be found in wild-type flagella as there might be a symmetry mis-match between the inner and outer parts of the flagellar core preventing a perfectly helical structure from ever forming. We think that a simpler explanation may be possible, and an analogy might be made to the hexameric Rho transcription-termination helicase. While early low-resolution EM studies suggested that Rho formed a ring with a C6 symmetry[175], a high resolution crystal structure revealed that the ring was actually a “lock washer” with a discontinuity between two of the subunits[176]. It is tempting to view wild-type archaeal and bacterial flagellins in a similar manner. The single α-helix of the archaeal flagellin core (Fig 4.21A) might easily be able to assemble into a straight polymer of nine identical subunits forming nine identical protofilaments (Fig 4.21B). However, the addition of a tenth subunit to complete the AFF might be unable to assemble into a straight polymer due to clashes (Fig 4.21C). As a result, the minimum energy state of the AFF involves 10 protofilament states with a seam at the inner curve (Fig 4.21D). It would be a similar case for BFFs but with different mechanistic details due to the α-helices in the core domain forming

coiled-coils (Fig 4.21E). A straight filament with 10 identical coiled-coil subunits generating 10 identical protofilaments may be possible (Fig 4.21F). However, the subunits may not be able to generate a completed straight flagellar filament without considerable strain and clashes (Fig 4.21 G). Thus, the supercoiled BFF states may indeed also reflect a minimum energy state where there are 11 unique protofilaments, with shorter ones on the inner curve and longer ones on the outside (Fig 4.21H).

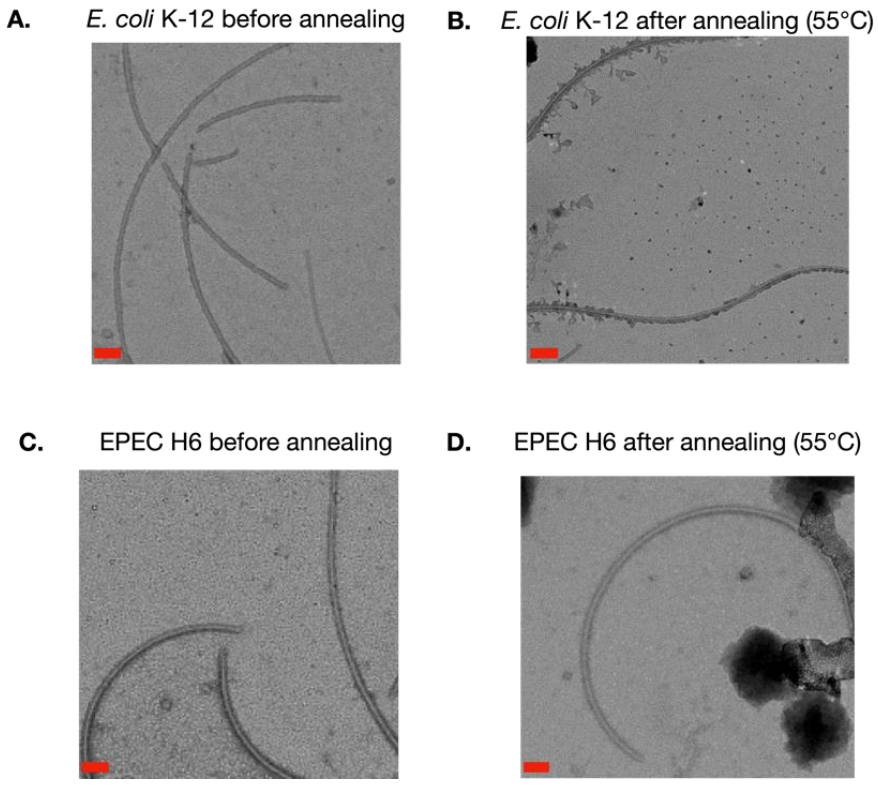

Figure 4.20. Annealing experiment of the *E. coli* K-12 and EPEC H6 bacterial flagellar filaments. (A). Negative stain TEM of *E. coli* K-12 flagellar filaments at room temperature in PBS pH 7.2. Filaments were stained with phosphotungstic acid (PTA) pH 7. **(B).** Negative stain TEM of *E. coli* K-12 flagellar filaments after annealing experiment. An 100µL aliquot of the filaments was heated up to 55°C for 30 minutes and then allowed to cool down to room temperature for where they remained for more than an hour. **(C).** Negative stain TEM of EPEC H6 flagellar filaments prior to the annealing experiment. **(D).** Negative stain TEM of EPEC H6 flagellar filaments after the annealing experiment.

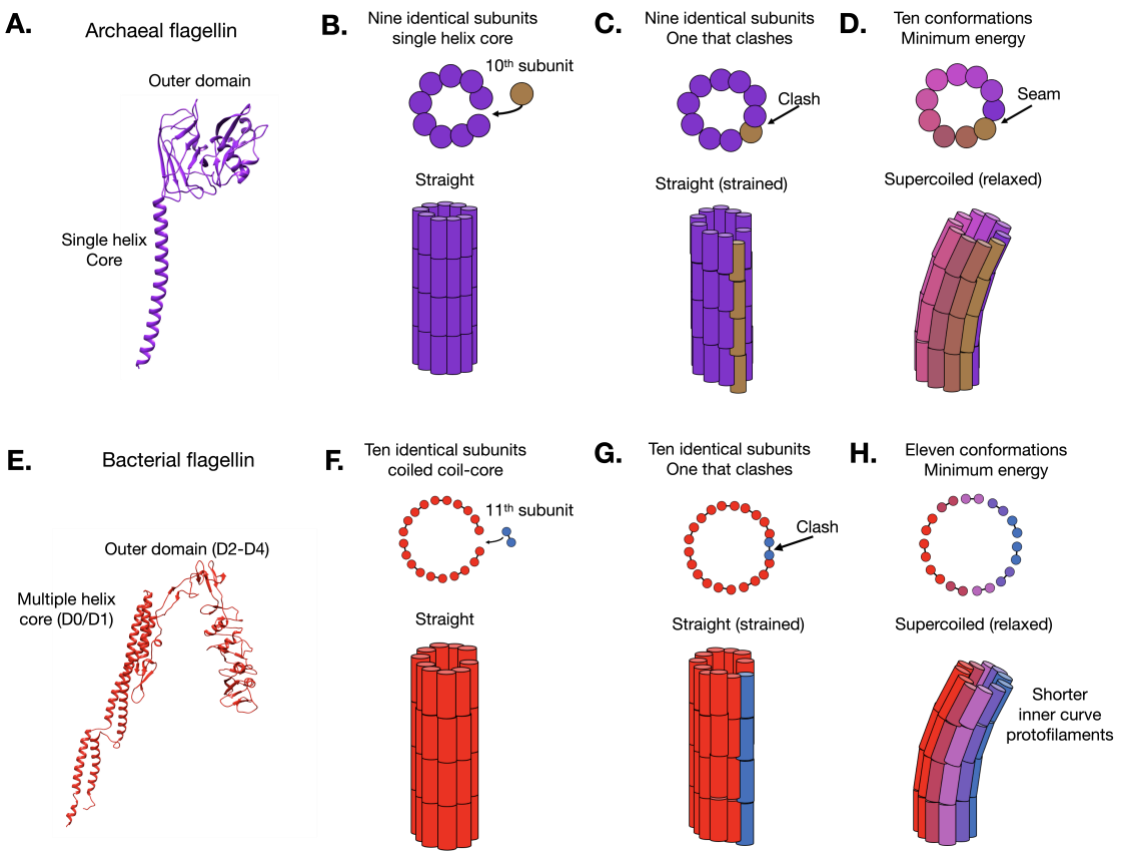

Figure 4.21. A comparison of the supercoiling of archaeal and bacterial flagellar filaments. A. A single REY15A archaeal flagellin is shown. **B.** Depiction showing how 9 identical subunits (purple) can assemble into a perfectly straight filament with nine protofilaments. Archaeal flagellar filaments, however, are assembled from 10 protofilaments. The top image is a top-down view of the filament, and circles represent the single α-helix for each subunit in the REY15A flagellar core. The bottom image shows the side-view of the filament with subunits represented as cylinders **C.** Depiction showing the addition of a 10th subunit (brown). If the archaeal flagellar filament cannot assemble into a straight flagellum as there would be serious clashes. **D.** Depiction of the minimum energy assembly of a supercoiled archaeal flagellum with 10 protofilaments each in a different conformation. Protofilaments are colored in order of similarity with the two extremes being purple and brown, which are the conformations most different from each other . These two protofilaments are adjacent to each other on the inner curve. **E.** A single EPEC H6 flagellin is shown. **F.** Similar concept as in **B** but for the bacterial flagellar filament. Here, there are 11-protofilaments and the coiled-coil subunits are each represented by two circles connected by a line. **G.** Depiction showing how the addition of the 11th bacterial protofilament would result in clashes if the filament was straight. **H.** The minimum energy state of the bacterial flagellar filament is supercoiled and has 11 protofilament conformations. These conformations vary in height. The shortest protofilaments are blue and the longest are red. The supercoiled state shown is that of Curly I.

While helically symmetric reconstructions of supercoiled flagellar filaments can be generated up to ~3.5 Å resolution, they are artifactual and the result of the different protofilament and subunit states being very similar to each other. This may be surprising because lower resolution structural studies suggested much larger differences between the L- and R-states[31]. Comparisons between L- and R-state bacterial flagellins[14] do

reveal larger differences between any of the flagellins in the supercoiled structures than those observed by us (Fig 4.22), but these differences are still much smaller than in the lower resolution structures. It is possible that the L- and R-state mutations prevent the shifting of 5-start interfacial interactions (Fig 4.23) or subtle conformational switching of the flagellin in other key regions[14]. The larger differences between the L- and R-state flagellins may therefore reflect the conformational shifts that these mutant subunits must undergo to assemble into a helically symmetric straight filament with minimal clashes.

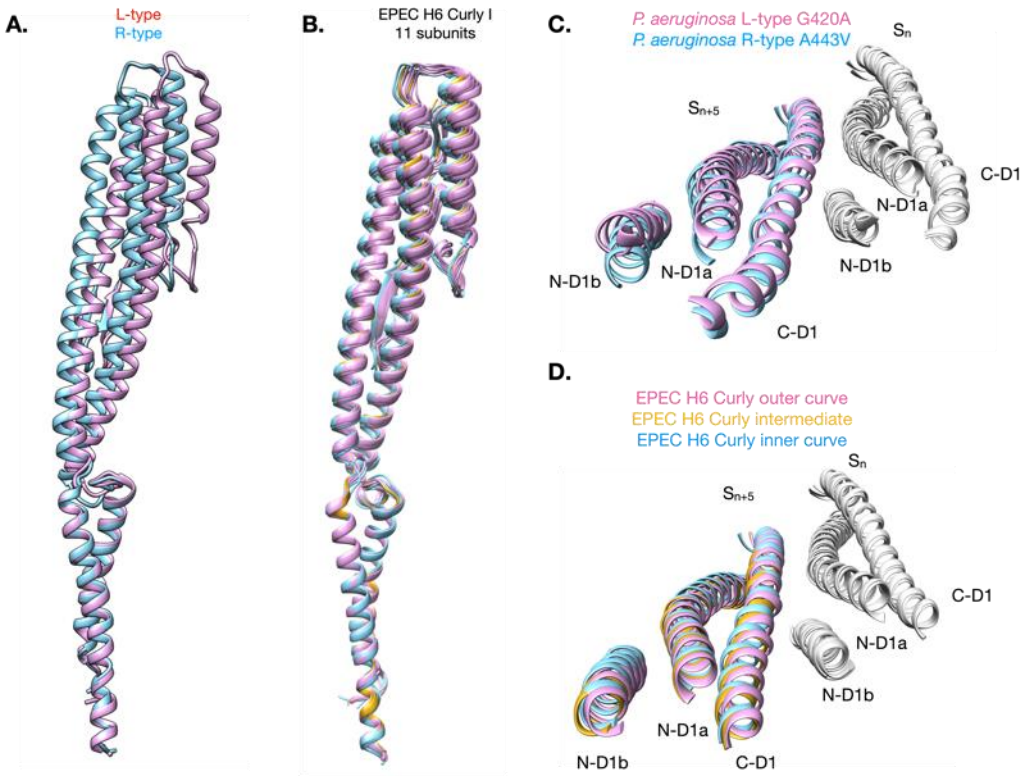

Figure 4.22. Comparison of L- and R-type flagellins with the flagellins form the supercoiled Curly I structure. **(A).** Alignment of the *P. aeruginosa* L-type (pink) and R-type flagellins by domain D0. **(B).** Alignment of the 11 EPEC H6 Curly I subunits by domain D0. **(C).** Alignment of the L-type (pink) and R-type (blue) 5-start interfaces by subunit S_n. **(D).** Alignment of the Curly I outer curve (pink), inner curve (blue) and intermediate 5-start interfaces by subunit S_n.

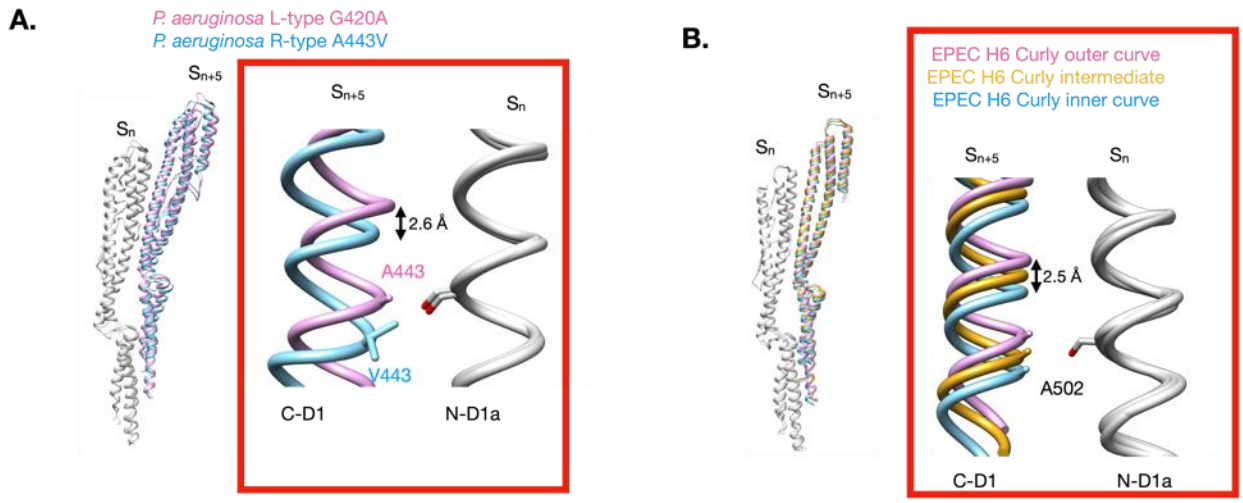

Figure 4.23. Comparison of the 5-start interface between L- and R-type flagellins and the EPEC H6 Curly I flagellins. (A). Alignment of the 5-start interface between L-type and R-type structures form *P. aeruginosa.* The two interfaces have been aligned by the subunit S_n. The left image shows the entire flagellins in the alignment. The right image shows the interface of C-D1 from subunit S_{n+5} and N-D1a from subunit S_n. **(B).** Alignment of the outer curve (pink), intermediate (gold), and inner curve (blue) 5-start interfaces in the H6 Curly I model.

Despite the 11 distinct bacterial flagellin states in the supercoiled structures, the two-state model does accurately predict the average twist of each protofilament[16]. The two-state model predicts that a Curly I flagellar filament would have 6 L-type (longer) and 5 R-type (shorter) protofilaments[16]. For our curved Curly I structure, 4 inner curve protofilaments are shifted downward by ~2-2.5 Å relative to five of the outer curve ones and there are two intermediate protofilaments that are shifted ~1 Å upward relative to the inner curve and ~1 Å downward relative to the outer curve. These protofilament interactions would be akin to having 6 L- and 5 R-type interactions, however, the intermediate protofilament shearing in the z-dimension (Fig 4.10) as well the 11 unique of protofilament conformations result in minimal strain compared to the combination of the two discrete L- and R-states.

When building the Curly I coarse-grained supercoiled model we found that the diameter of the supercoil varied depending on which protofilaments were used to generate the coarse-grained model. When just the inner- and outer-most curvature protofilaments were used a pitch of 1.1 μm and diameter of 0.34 μm were obtained. This produces a twist, τ, of 2.93 radians/μm and a curvature, κ of 2.85 radians/μm. When the "intermediate" protofilaments (gold in Figure 4.10) the supercoiled model had a pitch of 1.1 μm and diameter of 0.44 μm. This gives a twist, τ, of 2.22 radians/μm and a curvature, κ of 2.78 radians/μm. We therefore estimate that the margin of error in the parameters of our Curly I supercoiled model is a pitch of 1.1 +/- 0.1 μm and diameter of 0.34 +/- 0.1 μm. These values are similar to published values for the Curly I waveform with pitches ranging from ~0.8-1.2 μm and diameters from ~0.2-0.4 μm [8] . The slight-deviations from expected values can be attributed to many sources of small error including the atomic model building, selection of atoms for coarse-grained modeling, slight variations in resolution in the density map, as well as slight changes in filament structure due to compressional forces in cryo-EM grid preparation.

Bacterial and archaeal flagellar systems evolved completely independently from each other. This is reflected by their non-homologous subunits and their different, but analogous mechanism of supercoiling. The archaeal flagellin shares common ancestry with type IV pili[158], while the bacterial flagellar secretion system is evolutionarily related to the type III secretion system (T3SS)[177,178]. During assembly of the bacterial flagellar filament, flagellin monomers are secreted through the pore in the center of the filament[1] in a similar manner to effector proteins secreted through either the T3SS needle[179] or EspA needle extension filaments[180]. The archaeal flagellar protofilaments have a

striking difference in conformation between adjacent protofilaments in the inner curve (Fig 4.17,4.18). The conformations of the bacterial flagellar protofilaments are rather a continuum (Fig 4.9E-F), and each protofilament bends towards the inner curve while tilting differently with respect to the bacterial flagellar pore (Fig 4.9G-H). This preserves the circular shape of the lumen in a supercoiled bacterial flagellum allowing for the movement of monomers through the channel. Archaeal flagellar filaments do not secrete proteins through a lumen. This is another difference in the requirements for archaeal and bacterial flagellar filament supercoiling as the bacterial protofilaments must preserve the architecture of the flagellar lumen.

The 11 subunit and protofilament states we identify for the Curly I waveform are more different from each other (Fig 4.9B and 4.9E) than those from the Normal waveform (Fig 4.9C and 4.9F). This suggests that the 11 states that the protofilaments adopt depend on the curvature of the supercoiled structure. This likely keeps the flagellar pore intact as curvature increases or decreases due to the changes in the number of shorter or longer protofilaments.

For years any understanding of flagellar filament supercoiling was from a combination of low-resolution structural observations of supercoiled flagella using dark-field microscopy[10] and negative stain TEM[148] coupled with higher-resolution structural studies of straight mutant flagella[14,95,147]. The result of these studies was the two-state model[16], which could explain all of the L- and R-state structural data in the context of the supercoiling measurements from light microscopy[8,164]. With the data available, the two-state model was the most rational model for supercoiling of bacterial flagella. However, the major limitation of this model was the use of high-resolution structural information

from straight mutant rather than supercoiled wild-type flagellar filaments. Until recently, it was assumed that high-resolution structures of supercoiled flagellar filaments were impossible to obtain. With the cryo-EM structures of supercoiled bacterial and archaeal flagellar filaments we now proved that the atomic details of prokaryotic supercoiling are more complex than any simple two-state model.

The convergent evolution that has been described for these analogous structures in bacteria and archaea raises new questions about terminology. The archaeal flagellar system has been called the archaellum[181], based upon the notion that it has no homology with the bacterial one. However, the bacterial system has no homology to such eukaryotic flagella as the whipping tails of sperms[182] or of protists such as trypanosomes[183], yet these analogous structures are still called flagella in both, bacteria and eukaryotes. Wirth has commented on how the term “archaellum” leads to confusion, not clarity[184]. Despite the lack of homology between the structural components, the mechanisms of archaeal and bacterial flagellar supercoiling are analogous and examples of convergent evolution, much like how the wings of bees and bats have evolved independently.

4.5 Methods

Preparation of EPEC H6 flagellar filaments

The EPEC H6 flagellar filaments were prepared as previously published[99]. For the Curly I EPEC H6 sample the final buffer concentration was 1 M Tris/HCl 100 mM NaCl pH 6.5. For the Normal EPEC H6 sample the final buffer concentration was 50 mM Tris/HCl 75 mM NaCl pH 7.4.

Preparation of *E. coli* K-12 flagellar filaments

E. coli K12 strain MG1655 (motile variant) was grown overnight in LB (1% tryptone, 0.5% yeast extract and 0.5% NaCl) at 37ºC. The stationary culture was diluted 1:200 in 800 mL LB and grown to an OD_{600} between 1.0 and 1.4. Cells were harvested by centrifugation for 8 minutes at 8,000 x g and suspended in 100 mL motility buffer (0.5 mM $CaCl_2$, 0.1 mM EDTA, 20 mM HEPES [pH 7.2]). Flagella were sheared from cells by agitation in a mixer at full speed for 20 seconds and separated by three centrifugation steps: one at 8,000 x g for 7 minutes, followed by two at 15,000 x g for 15 minutes at 4ºC. Purified flagella were then sedimented at 87,000 x g for 2 hours at 4ºC, washed with motility buffer, and suspended in 200 mL motility buffer. Purity of flagella filaments was analyzed by SDS-PAGE.

Preparation of *S. islandicus* REY15A AFFs

S. islandicus REY15A was grown aerobically with shaking (145 rpm) at 75 °C in STV medium containing mineral salts (M), 0.2% (wt/vol) sucrose (S), 0.2% (wt/vol) tryptone (T), a mixed vitamin solution (V) (PMID:19513584). To prepare for the flagellar filaments, REY15A cells were inoculate into 300 ml MTSV medium with an initial OD_{600} of 0.05 and the cell culture was put outside of the shaker and cooled down naturally to room temperature after 18 h of cultivation; Then the cell culture was vortexed to shear off the flagellar filaments from the cell surface; The cells were removed by centrifugation at 8,000 rpm for 20 min and the flagellar filaments were pelleted down from the supernatant by ultra-centrifugation at 40,000 rpm (Type 45Ti rotor) at 10 °C for 2 h.

Preparation of *C. jejuni* flagellar filaments

Wild type *C. jejuni* 81-176 flagellar filaments were prepared as published[33].

Annealing experiments with BFFs

Diluted 100 µL samples of bacterial flagella were heated at 55 °C for 30 minutes. The samples were then allowed to cool at room temperature for one hour. Samples before and after the annealing experiment were stained on continuous carbon EM grids using 1 % phosphotungstic acid pH 7. The negatively stained samples were then imaged with a Technai T12 microscope operating at an accelerating voltage of 80 keV.

Cryo-EM sample and grid preparation

The cryo-EM grids were prepared for imaging as described previously[33,99]. Samples were plunge frozen using either a Vitrobot Mark IV or a Leica EM GP. For the Vitrobot 3.5-4 µL of sample was applied to the grid and all blot times were 3.5 seconds and the blot forces ranged from 3 to 6. For the Leica 2.5-3 µL of sample was applied to the grid and the blot times were 3.5 S while blot distances ranged from 214-217.

CryoEM image acquisition

Data collection settings were identical to previously published studies[99]. Images were acquired on a Titan Krios equipped with a K3 direct electron detector using a pixel size of 1.08 Å/pixel. The acerating voltage was 300 keV.

Helical reconstruction

Helical reconstruction of flagellar filaments was performed in Spider[130] using the iterative helical real space reconstruction[185] and also performed in cryoSPARC[131] to similar results.

Asymmetric reconstruction

Flagellar filaments were reconstructed asymmetrically using cryoSPARC[131] in a series of steps. First a low-resolution helical volume (~20-50 Å) was used as the initial

volume for homogenous refinement. For all structures this initial round of homogenous refinement showed no curvature in the volume and still appeared largely helical. The particles and mask from this job were then used as inputs for a 3D variability job. The results from this job were reconstructed into 3D variability clusters and the highest resolution cluster with a curved volume was used for subsequent steps. To the best of our knowledge the volumes reconstructed from the 3D variability job represented a sort of “constrained” ab-initio reconstruction, as no starting volume was used. However, a mask with general flagella features was as well as the particles containing the alignments parameters from the previous job. For the EPEC H6 reconstructions a static mask consisting of just the core domains D0 and D1 was sued for all steps, while for the *E. coli* K-12 and *S. islandicus* REY15A structures a mask consisting of the entire volume was used.

For some of our structures we found the curvature was present in the reconstruction from 3D variability but absent after taking these results as inputs for homogenous refinement. For best results, the chosen cluster of particles and its corresponding volume from 3D variability display was used as an input for a local refinement job with shifts constrained from 5-20 Å and rotational searches from 5-20° depending on the sample being imaged. These results were the used as inputs for local CTF correction followed by a second round of local refinement. For the Curly I EPEC H6 structure we were able to reconstruct curved volumes to 3.4 Å using a particles box size of 384 x 384 pixels and to ~4.0 Å (3.7 Å with 0.143 GFSC) using a box size of 624x624 pixels. The ~3 radians/µm curvature of the curly I was present in both structures,

however smaller 384 x 384 box sizes for the Normal BFFs as well as the *S. islandicus* REY15A AFF produced straight volumes with little to no evidence of curvature. The EPEC H6 Normal and *S. islandicus* REY15A volumes were reconstructed using 624x624 particles boxes. The *E. coli* K-12 structure was reconstructed with a box size of 512x512.

Model building

The core domains D0 and D1 of the previously published EPEC H6 model (7SN7) was used as the initial subunit model for the Curly I and Normal EPEC H6 structures. AlphaFold[138] was used to generate the starting models for the *E. coli* K-12 and *S. islandicus* REY15A flagellin models. For the REY15A AlphaFold model the secondary for each domain was predicted accurately, however the orientation of the C-terminal outer domain relative to the N-terminal core was incorrect. This makes sense because the model could not have accounted for the polymerization of the monomer into a filament. Modeling for each flagellar filament was initiated by fitting a single subunit into each protofilament using UCSF Chimera[71]. and then refining each model using Coot[133]. For each protofilament the subunit initially fit and refined into a protofilament was copied (at least 3 copies each) and each copy was then fit into adjacent subunits along the same protofilament. Each subunit from each protofilament was refined individually in Coot and then each protofilament was refined using real-space refinement in Phenix[135]. For each structure the 11 bacterial or 10 archaeal protofilament models were then combined together into a single model for that flagellar filament which was subsequently refined in Phenix.

Determination of modeling error in comparisons of independently built subunits

To determine the extent to which we could interpret differences between subunits on the same or adjacent protofilaments, we conducted an experiment to predict the error of fitting a subunit into a map. We took a small segment of the H6 curly model in domain D1 (Fig S9a) and created 10 copies. To each, a different set of shifts in x, y, and z was applied (Fig S9b), with a mean rms shift of 2.6 Å. We then took the shifted models and used Phenix real space refinement to fit them into the corresponding region of the density map (Fig S9c). The C_α backbones of the 10 models were all quite similar after refinement (Fig S9d), and the average backbone rmsd value for the shifted models compared to the starting model was 0.2 Å. This estimate of the uncertainty in our models provided confidence that the detected structural differences between protein subunits are true conformational differences.

Comparison of flagellin subunits

All comparisons between flagellin subunits through structural alignments were generated with the matchmaker function in UCSF Chimera. The 3-dimensional plot shown in Fig S10d was generated using the free version of Plotly Chart Studio.

Coarse-grained modeling of the supercoil of prokaryotic flagellar filaments

For each flagellar filament model only a few atoms from each domain were kept for the coarse-grained model. All other atoms were deleted. The atomic radii of the alpha carbons were increased by ~20 Å to represent the flagellin subunits as large spheres. The chains of inner-most (lowest resolution) and outer-most (highest resolution) curve protofilaments were aligned in UCSF Chimera.

Cryo electron tomography

A 5 µL aliquot of 10 nM BSA Gold Tracer (Electron Microscopy Sciences^{TM}) was added

to 10 µL of *C. jejuni* flagellar filaments. Grids were prepared in an identical fashion to all other flagellar filaments samples except that 300 copper mesh C-flat holey carbon grids were used with 2 µm x 2µm hole size. Tilt series were collected from -60° to +60° tilt with 1.5° degrees between each image and a dose of 4.2 $e^-/Å^2$. Tomograms were generated using Etomo[186] which is part of the IMOD processing suite[187]. Segmentation of the flagellar filaments was then performed to better visualize them in 3-dimensions[188].

4.6 Acknowledgments

REY15A filaments were prepared by Junfeng Liu of the Institut Pasteur. EPEC H6 filaments were prepared by Sharanya Chatterjee and Priyanka Biswas at the Imperial College of London. Ravi Sonani solved the *E. coli* K-12 flagella structure.

Chapter 5. References

1 Macnab, R. M. How bacteria assemble flagella. *Annu Rev Microbiol* **57**, 77-100, doi:10.1146/annurev.micro.57.030502.090832 (2003).

2 Morimoto, Y. V. & Minamino, T. Architecture and Assembly of the Bacterial Flagellar Motor Complex. *Subcell Biochem* **96**, 297-321, doi:10.1007/978-3-030-58971-4_8 (2021).

3 Darnton, N. C., Turner, L., Rojevsky, S. & Berg, H. C. On torque and tumbling in swimming Escherichia coli. *J Bacteriol* **189**, 1756-1764, doi:10.1128/JB.01501-06 (2007).

4 Götz, R. & Schmitt, R. Rhizobium meliloti swims by unidirectional, intermittent rotation of right-handed flagellar helices. *J Bacteriol* **169**, 3146-3150, doi:10.1128/jb.169.7.3146-3150.1987 (1987).

5 Biquet-Bisquert, A., Labesse, G., Pedaci, F. & Nord, A. L. The Dynamic Ion Motive Force Powering the Bacterial Flagellar Motor. *Front Microbiol* **12**, 659464, doi:10.3389/fmicb.2021.659464 (2021).

6 Cohen, E. J. *et al.* Campylobacter jejuni motility integrates specialized cell shape, flagellar filament, and motor, to coordinate action of its opposed flagella. *PLoS Pathog* **16**, e1008620, doi:10.1371/journal.ppat.1008620 (2020).

7 Leifson, E. Staining, shape and arrangement of bacterial flagella. *J Bacteriol* **62**, 377-389, doi:10.1128/jb.62.4.377-389.1951 (1951).

8 Turner, L., Ryu, W. S. & Berg, H. C. Real-time imaging of fluorescent flagellar filaments. *J Bacteriol* **182**, 2793-2801, doi:10.1128/jb.182.10.2793-2801.2000 (2000).

9 Macnab, R. M. Bacterial flagella rotating in bundles: a study in helical geometry. *Proc Natl Acad Sci U S A* **74**, 221-225, doi:10.1073/pnas.74.1.221 (1977).

10 Macnab, R. M. Examination of bacterial flagellation by dark-field microscopy. *J Clin Microbiol* **4**, 258-265, doi:10.1128/jcm.4.3.258-265.1976 (1976).

11 Wadhwa, N. & Berg, H. C. Bacterial motility: machinery and mechanisms. *Nat Rev Microbiol*, doi:10.1038/s41579-021-00626-4 (2021).

12 Wadhams, G. H. & Armitage, J. P. Making sense of it all: bacterial chemotaxis. *Nat Rev Mol Cell Biol* **5**, 1024-1037, doi:10.1038/nrm1524 (2004).

13 Berg, H. C. & Brown, D. A. Chemotaxis in Escherichia coli analyzed by three-dimensional tracking. *Antibiot Chemother (1971)* **19**, 55-78, doi:10.1159/000395424 (1974).

14 Wang, F. *et al.* A structural model of flagellar filament switching across multiple bacterial species. *Nat Commun* **8**, 960, doi:10.1038/s41467-017-01075-5 (2017).

15 Hyman, H. C. & Trachtenberg, S. Point mutations that lock Salmonella typhimurium flagellar filaments in the straight right-handed and left-handed forms and their relation to filament superhelicity. *J. Mol. Biol.* **220**, 79-88 (1991).

16 Calladine, C. R., Luisi, B. F. & Pratap, J. V. A "mechanistic" explanation of the multiple helical forms adopted by bacterial flagellar filaments. *J Mol Biol* **425**, 914-928, doi:10.1016/j.jmb.2012.12.007 (2013).

17 Kanto, S., Okino, H., Aizawa, S. & Yamaguchi, S. Amino acids responsible for flagellar shape are distributed in terminal regions of flagellin. *J Mol Biol* **219**, 471-480, doi:10.1016/0022-2836(91)90187-b (1991).

18 Malapaka, R. R., Adebayo, L. O. & Tripp, B. C. A deletion variant study of the functional role of the Salmonella flagellin hypervariable domain region in motility. *J Mol Biol* **365**, 1102-1116, doi:10.1016/j.jmb.2006.10.054 (2007).
19 Yoshioka, K., Aizawa, S. & Yamaguchi, S. Flagellar filament structure and cell motility of Salmonella typhimurium mutants lacking part of the outer domain of flagellin. *J Bacteriol* **177**, 1090-1093, doi:10.1128/jb.177.4.1090-1093.1995 (1995).
20 Nedeljković, M., Sastre, D. E. & Sundberg, E. J. Bacterial Flagellar Filament: A Supramolecular Multifunctional Nanostructure. *Int J Mol Sci* **22**, doi:10.3390/ijms22147521 (2021).
21 Trachtenberg, S. & DeRosier, D. J. Three-dimensional reconstruction of the flagellar filament of Caulobacter crescentus. A flagellin lacking the outer domain and its amino acid sequence lacking an internal segment. *J Mol Biol* **202**, 787-808, doi:10.1016/0022-2836(88)90559-1 (1988).
22 Blum, T. B., Filippidou, S., Fatton, M., Junier, P. & Abrahams, J. P. The wild-type flagellar filament of the Firmicute Kurthia at 2.8 Å resolution in vivo. *Sci Rep* **9**, 14948, doi:10.1038/s41598-019-51440-1 (2019).
23 Erdem, A. L., Avelino, F., Xicohtencatl-Cortes, J. & Giron, J. A. Host protein binding and adhesive properties of H6 and H7 flagella of attaching and effacing Escherichia coli. *J Bacteriol* **189**, 7426-7435, doi:10.1128/JB.00464-07 (2007).
24 Giron, J. A., Torres, A. G., Freer, E. & Kaper, J. B. The flagella of enteropathogenic Escherichia coli mediate adherence to epithelial cells. *Mol Microbiol* **44**, 361-379, doi:10.1046/j.1365-2958.2002.02899.x (2002).
25 Horstmann, J. A. *et al.* Methylation of Salmonella Typhimurium flagella promotes bacterial adhesion and host cell invasion. *Nat Commun* **11**, 2013, doi:10.1038/s41467-020-15738-3 (2020).
26 Steiner, T. S., Nataro, J. P., Poteet-Smith, C. E., Smith, J. A. & Guerrant, R. L. Enteroaggregative Escherichia coli expresses a novel flagellin that causes IL-8 release from intestinal epithelial cells. *J Clin Invest* **105**, 1769-1777, doi:10.1172/jci8892 (2000).
27 Zhou, X. *et al.* Flagellin of enteropathogenic Escherichia coli stimulates interleukin-8 production in T84 cells. *Infect Immun* **71**, 2120-2129, doi:10.1128/iai.71.4.2120-2129.2003 (2003).
28 Troge, A. *et al.* More than a marine propeller--the flagellum of the probiotic Escherichia coli strain Nissle 1917 is the major adhesin mediating binding to human mucus. *Int J Med Microbiol* **302**, 304-314, doi:10.1016/j.ijmm.2012.09.004 (2012).
29 Nisani-Bizer, K. & Trachtenberg, S. Unperturbing a non-helically perturbed bacterial flagellar filament: Salmonella typhimurium SJW23. *J Mol Biol* **416**, 367-388, doi:10.1016/j.jmb.2012.01.003 (2012).
30 Eckhard, U. *et al.* Discovery of a proteolytic flagellin family in diverse bacterial phyla that assembles enzymatically active flagella. *Nat Commun* **8**, 521, doi:10.1038/s41467-017-00599-0 (2017).
31 Maki-Yonekura, S., Yonekura, K. & Namba, K. Conformational change of flagellin for polymorphic supercoiling of the flagellar filament. *Nat.Struct.Mol.Biol.* **17**, 417-422 (2010).

32 Yonekura, K., Maki-Yonekura, S. & Namba, K. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. *Nature* **424**, 643-650 (2003).
33 Kreutzberger, M. A. B., Ewing, C., Poly, F., Wang, F. & Egelman, E. H. Atomic structure of the Campylobacter jejuni flagellar filament reveals how epsilon Proteobacteria escaped Toll-like receptor 5 surveillance. *Proc Natl Acad Sci U S A* **117**, 16985-16991, doi:10.1073/pnas.2010996117 (2020).
34 Montemayor, E. J. *et al.* Flagellar Structures from the Bacterium Caulobacter crescentus and Implications for Phage φ CbK Predation of Multiflagellin Bacteria. *J Bacteriol* **203**, doi:10.1128/jb.00399-20 (2021).
35 Khan, S. & Scholey, J. M. Assembly, Functions and Evolution of Archaella, Flagella and Cilia. *Curr Biol* **28**, R278-r292, doi:10.1016/j.cub.2018.01.085 (2018).
36 Kinosita, Y., Uchida, N., Nakane, D. & Nishizaka, T. Direct observation of rotation and steps of the archaellum in the swimming halophilic archaeon Halobacterium salinarum. *Nat Microbiol* **1**, 16148, doi:10.1038/nmicrobiol.2016.148 (2016).
37 Hayashi, F. *et al.* The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature* **410**, 1099-1103, doi:10.1038/35074106 (2001).
38 Song, W. S., Jeon, Y. J., Namgung, B., Hong, M. & Yoon, S. I. A conserved TLR5 binding and activation hot spot on flagellin. *Sci Rep* **7**, 40878, doi:10.1038/srep40878 (2017).
39 Yoon, S. I. *et al.* Structural basis of TLR5-flagellin recognition and signaling. *Science* **335**, 859-864, doi:10.1126/science.1215584 (2012).
40 Andersen-Nissen, E. *et al.* Evasion of Toll-like receptor 5 by flagellated bacteria. *Proc.Natl.Acad.Sci.U.S.A* **102**, 9247-9252 (2005).
41 Smith, K. D. *et al.* Toll-like receptor 5 recognizes a conserved site on flagellin required for protofilament formation and bacterial motility. *Nat.Immunol.* **4**, 1247-1253 (2003).
42 Galkin, V. E. *et al.* Divergence of quaternary structures among bacterial flagellar filaments. *Science* **320**, 382-385 (2008).
43 Egelman, E. H. The Current Revolution in Cryo-EM. *Biophys J* **110**, 1008-1012, doi:10.1016/j.bpj.2016.02.001 (2016).
44 Bai, X. C., McMullan, G. & Scheres, S. H. How cryo-EM is revolutionizing structural biology. *Trends Biochem Sci* **40**, 49-57, doi:10.1016/j.tibs.2014.10.005 (2015).
45 Kuhlbrandt, W. Cryo-EM enters a new era. *eLife* **3**, e03678, doi:10.7554/eLife.03678 (2014).
46 Egelman, E. H. Ambiguities in Helical Reconstruction. *eLife* **3:e04969 doi:10.7554/eLife.04969** (2014).
47 Ewing, C. P., Andreishcheva, E. & Guerry, P. Functional characterization of flagellin glycosylation in Campylobacter jejuni 81-176. *J Bacteriol* **191**, 7086-7093, doi:10.1128/JB.00378-09 (2009).
48 Yoshioka, K., Aizawa, S. & Yamaguchi, S. Flagellar filament structure and cell motility of Salmonella typhimurium mutants lacking part of the outer domain of flagellin. *Journal of Bacteriology* **177**, 1090-1093 (1995).
49 Egelman, E. H. A robust algorithm for the reconstruction of helical filaments using single-particle methods. *Ultramicroscopy* **85**, 225-234 (2000).
50 He, S. & Scheres, S. H. W. Helical reconstruction in RELION. *J Struct Biol* **198**, 163-176, doi:10.1016/j.jsb.2017.02.003 (2017).

51 Holm, L. & Rosenstrom, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38**, W545-549, doi:10.1093/nar/gkq366 (2010).
52 Song, W. S. & Yoon, S. I. Crystal structure of FliC flagellin from Pseudomonas aeruginosa and its implication in TLR5 binding and formation of the flagellar filament. *Biochem Biophys Res Commun* **444**, 109-115, doi:10.1016/j.bbrc.2014.01.008 (2014).
53 Cho, S. Y., Song, W. S. & Yoon, S. I. Crystal structure of the flagellar cap protein FliD from Bdellovibrio bacteriovorus. *Biochem Biophys Res Commun* **519**, 652-658, doi:10.1016/j.bbrc.2019.09.024 (2019).
54 Cho, S. Y. *et al.* Structural analysis of the flagellar capping protein FliD from Helicobacter pylori. *Biochem Biophys Res Commun* **514**, 98-104, doi:10.1016/j.bbrc.2019.04.065 (2019).
55 Matsunami, H., Barker, C. S., Yoon, Y. H., Wolf, M. & Samatey, F. A. Complete structure of the bacterial flagellar hook reveals extensive set of stabilizing interactions. *Nature communications* **7**, 13425, doi:10.1038/ncomms13425 (2016).
56 Thibault, P. *et al.* Identification of the carbohydrate moieties and glycosylation motifs in Campylobacter jejuni flagellin. *J Biol Chem* **276**, 34862-34870, doi:10.1074/jbc.M104529200 (2001).
57 Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-797, doi:10.1016/j.jmb.2007.05.022 (2007).
58 Wassenaar, T. M., Bleumink-Pluym, N. M. & van der Zeijst, B. A. Inactivation of Campylobacter jejuni flagellin genes by homologous recombination demonstrates that flaA but not flaB is required for invasion. *EMBO J* **10**, 2055-2061 (1991).
59 Stoddard, P. R., Williams, T. A., Garner, E. & Baum, B. Evolution of polymer formation within the actin superfamily. *Molecular biology of the cell* **28**, 2461-2469, doi:10.1091/mbc.E15-11-0778 (2017).
60 Kollman, J. M., Merdes, A., Mourey, L. & Agard, D. A. Microtubule nucleation by gamma-tubulin complexes. *Nature reviews. Molecular cell biology* **12**, 709-721, doi:10.1038/nrm3209 (2011).
61 Misawa, N. & Blaser, M. J. Detection and characterization of autoagglutination activity by Campylobacter jejuni. *Infect Immun* **68**, 6168-6175, doi:10.1128/iai.68.11.6168-6175.2000 (2000).
62 Wang, F. *et al.* An extensively glycosylated archaeal pilus survives extreme conditions. *Nat Microbiol* **4**, 1401-1410, doi:10.1038/s41564-019-0458-x (2019).
63 Drozdetskiy, A., Cole, C., Procter, J. & Barton, G. J. JPred4: a protein secondary structure prediction server. *Nucleic Acids Res* **43**, W389-394, doi:10.1093/nar/gkv332 (2015).
64 Dembski, W. A. *No Free Lunch: Why Specified Complexity Cannot Be Purchased without Intelligence*. (Rowman & Littlefield Publishers, Inc., 2001).
65 Egelman, E. H. Reducing irreducible complexity: divergence of quaternary structure and function in macromolecular assemblies. *Curr.Opin.Cell Biol.* **22**, 68-74 (2010).
66 Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332, doi:10.1038/nmeth.4193 (2017).

67 Mindell, J. A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol* **142**, 334-347, doi:10.1016/s1047-8477(03)00069-8 (2003).
68 Tang, G. *et al.* EMAN2: an extensible image processing suite for electron microscopy. *J Struct Biol* **157**, 38-46, doi:10.1016/j.jsb.2006.05.009 (2007).
69 Egelman, E. H. Reconstruction of helical filaments and tubes. *Methods Enzymol* **482**, 167-183, doi:10.1016/S0076-6879(10)82006-3 (2010).
70 Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, doi:10.7554/eLife.42166 (2018).
71 Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612, doi:10.1002/jcc.20084 (2004).
72 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132, doi:10.1107/S0907444904019158 (2004).
73 Song, Y. *et al.* High-resolution comparative modeling with RosettaCM. *Structure* **21**, 1735-1742, doi:10.1016/j.str.2013.08.005 (2013).
74 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography* **66**, 213-221, doi:10.1107/S0907444909052925 (2010).
75 Das, R. & Baker, D. Macromolecular modeling with rosetta. *Annu Rev Biochem* **77**, 363-382, doi:10.1146/annurev.biochem.77.062906.171838 (2008).
76 Williams, C. J. *et al.* MolProbity: More and better reference data for improved all-atom structure validation. *Protein Science* **27**, 293-315, doi:10.1002/pro.3330 (2018).
77 Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-1191, doi:10.1093/bioinformatics/btp033 (2009).
78 Sievers, F. & Higgins, D. G. Clustal omega. *Curr Protoc Bioinformatics* **48**, 3 13 11-16, doi:10.1002/0471250953.bi0313s48 (2014).
79 Holm, L. & Sander, C. Dali: a network tool for protein structure comparison. *Trends Biochem Sci* **20**, 478-480, doi:10.1016/s0968-0004(00)89105-7 (1995).
80 Yamaguchi, T. *et al.* Structural and Functional Comparison of Salmonella Flagellar Filaments Composed of FljB and FliC. *Biomolecules* **10**, doi:10.3390/biom10020246 (2020).
81 Grognot, M. & Taute, K. M. More than propellers: how flagella shape bacterial motility behaviors. *Current Opinion in Microbiology* **61**, 73-81, doi:https://doi.org/10.1016/j.mib.2021.02.005 (2021).
82 Armitage, J. P. & Berry, R. M. Assembly and Dynamics of the Bacterial Flagellum. *Annu Rev Microbiol* **74**, 181-200, doi:10.1146/annurev-micro-090816-093411 (2020).
83 Schmitt, R., Raska, I. & Mayer, F. Plain and complex flagella of Pseudomonas rhodos: analysis of fine structure and composition. *J Bacteriol* **117**, 844-857, doi:10.1128/jb.117.2.844-857.1974 (1974).
84 Krupski, G., Götz, R., Ober, K., Pleier, E. & Schmitt, R. Structure of complex flagellar filaments in Rhizobium meliloti. *J Bacteriol* **162**, 361-366, doi:10.1128/jb.162.1.361-366.1985 (1985).

85 Trachtenberg, S., DeRosier, D. J., Aizawa, S. & Macnab, R. M. Pairwise perturbation of flagellin subunits. The structural basis for the differences between plain and complex bacterial flagellar filaments. *J. Mol. Biol.* **190**, 569-576 (1986).
86 Caspar, D. L. & Holmes, K. C. Structure of dahlemense strain of tobacco mosaic virus: a periodically deformed helix. *J. Mol. Biol.* **46**, 99-133, doi:10.1016/0022-2836(69)90060-6 (1969).
87 Yuen, G. Y. & Schroth, M. N. Inhibition of Fusarium-Oxysporum F Sp Dianthi by Iron Competition with an Alcaligenes Sp. *Phytopathology* **76**, 171-176, doi:DOI 10.1094/Phyto-76-171 (1986).
88 Martinetti, G. & Loper, J. E. Mutational analysis of genes determining antagonism of Alcaligenes sp. strain MFA1 against the phytopathogenic fungus Fusarium oxysporum. *Can J Microbiol* **38**, 241-247, doi:10.1139/m92-041 (1992).
89 Lawn, A. M., Orskov, I. & Orskov, F. Morphological distinction between different H serotypes of Escherichia coli. *J Gen Microbiol* **101**, 111-119, doi:10.1099/00221287-101-1-111 (1977).
90 Lawn, A. M. Comparison of the flagellins from different flagellar morphotypes of Escherichia coli. *J Gen Microbiol* **101**, 112-130, doi:10.1099/00221287-101-1-121 (1977).
91 Trachtenberg, S., DeRosier, D. J., Zemlin, F. & Beckmann, E. Non-helical perturbations of the flagellar filament: Salmonella typhimurium SJW117 at 9.6 A resolution. *J Mol Biol* **276**, 759-773, doi:10.1006/jmbi.1997.1537 (1998).
92 Cohen-Krausz, S. & Trachtenberg, S. Helical perturbations of the flagellar filament: rhizobium lupini H13-3 at 13 A resolution. *J. Struct. Biol.* **122**, 267-282 (1998).
93 Debs, G. E., Cha, M., Liu, X., Huehn, A. R. & Sindelar, C. V. Dynamic and asymmetric fluctuations in the microtubule wall captured by high-resolution cryoelectron microscopy. *Proc Natl Acad Sci U S A* **117**, 16976-16984, doi:10.1073/pnas.2001546117 (2020).
94 Gutzwiller, M. C. Moon-Earth-Sun: The oldest three-body problem. *Reviews of Modern Physics* **70**, 589-639, doi:10.1103/RevModPhys.70.589 (1998).
95 Yonekura, K., Maki-Yonekura, S. & Namba, K. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. *Nature* **424**, 643-650, doi:10.1038/nature01830 (2003).
96 Scharf, B., Schuster-Wolff-Bühring, H., Rachel, R. & Schmitt, R. Mutational Analysis of the Rhizobium lupini H13-3 andSinorhizobium meliloti Flagellin Genes: Importance of Flagellin A for Flagellar Filament Structure and Transcriptional Regulation. *Journal of Bacteriology* **183**, 5334-5342, doi:10.1128/jb.183.18.5334-5342.2001 (2001).
97 Götz, R., Limmer, N., Ober, K. & Schmitt, R. Motility and chemotaxis in two strains of Rhizobium with complex flagella. *Microbiology* **128**, 789-798 (1982).
98 Trachtenberg, S., Fishelov, D. & Ben-Artzi, M. Bacterial Flagellar Microhydrodynamics: Laminar Flow over Complex Flagellar Filaments, Analog Archimedean Screws and Cylinders, and Its Perturbations. *Biophysical journal* **85**, 1345-1357, doi:10.1016/S0006-3495(03)74569-X (2003).

99 Kreutzberger, M. A. B. *et al.* Flagellin outer domain dimerization modulates motility in pathogenic and soil bacteria from viscous environments. *Nat Commun* **13**, 1422, doi:10.1038/s41467-022-29069-y (2022).
100 Kondoh, H. & Yanagida, M. Structure of straight flagellar filaments from a mutant of Escherichia coli. *Journal of Molecular Biology* **96**, 641-652, doi:https://doi.org/10.1016/0022-2836(75)90143-6 (1975).
101 Berg, H. C. & Brown, D. A. Chemotaxis in Escherichia coli analysed by three-dimensional tracking. *Nature* **239**, 500-504, doi:10.1038/239500a0 (1972).
102 Macnab, R. M. & Koshland, D. E., Jr. The gradient-sensing mechanism in bacterial chemotaxis. *Proc Natl Acad Sci U S A* **69**, 2509-2512, doi:10.1073/pnas.69.9.2509 (1972).
103 Kamiya, R., Hotani, H. & Asakura, S. Polymorphic transition in bacterial flagella. *Symp Soc Exp Biol* **35**, 53-76 (1982).
104 Chu, J., Liu, J. & Hoover, T. R. Phylogenetic Distribution, Ultrastructure, and Function of Bacterial Flagellar Sheaths. *Biomolecules* **10**, doi:10.3390/biom10030363 (2020).
105 Gibson, K. H. *et al.* An asymmetric sheath controls flagellar supercoiling and motility in the leptospira spirochete. *Elife* **9**, doi:10.7554/eLife.53672 (2020).
106 Liu, Y. & Yang, Q. Cloning and heterologous expression of aspartic protease SA76 related to biocontrol in Trichoderma harzianum. *FEMS Microbiology Letters* **277**, 173-181, doi:10.1111/j.1574-6968.2007.00952.x (2007).
107 Hodge, C. N. *et al.* Improved cyclic urea inhibitors of the HIV-1 protease: synthesis, potency, resistance profile, human pharmacokinetics and X-ray crystal structure of DMP 450. *Chem Biol* **3**, 301-314, doi:10.1016/s1074-5521(96)90110-6 (1996).
108 Scharf, B. Real-time imaging of fluorescent flagellar filaments of Rhizobium lupini H13-3: flagellar rotation and pH-induced polymorphic transitions. *Journal of bacteriology* **184**, 5979-5986, doi:10.1128/jb.184.21.5979-5986.2002 (2002).
109 Hotani, H. Micro-video study of moving bacterial flagellar filaments. III. Cyclic transformation induced by mechanical force. *J Mol Biol* **156**, 791-806, doi:10.1016/0022-2836(82)90142-5 (1982).
110 Mears, P. J., Koirala, S., Rao, C. V., Golding, I. & Chemla, Y. R. Escherichia coli swimming is robust against variations in flagellar number. *Elife* **3**, e01916, doi:10.7554/eLife.01916 (2014).
111 Fujii, M., Shibata, S. & Aizawa, S. Polar, peritrichous, and lateral flagella belong to three distinguishable flagellar families. *J Mol Biol* **379**, 273-283, doi:10.1016/j.jmb.2008.04.012 (2008).
112 Sauder, A. B. & Kendall, M. M. After the Fact(or): Posttranscriptional Gene Regulation in Enterohemorrhagic Escherichia coli O157:H7. *J Bacteriol* **200**, doi:10.1128/jb.00228-18 (2018).
113 Jubelin, G. *et al.* Modulation of Enterohaemorrhagic Escherichia coli Survival and Virulence in the Human Gastrointestinal Tract. *Microorganisms* **6**, 115, doi:10.3390/microorganisms6040115 (2018).
114 Frankel, G. *et al.* Enteropathogenic and enterohaemorrhagic Escherichia coli: more subversive elements. *Mol Microbiol* **30**, 911-921, doi:10.1046/j.1365-2958.1998.01144.x (1998).

115 Garmendia, J., Frankel, G. & Crepin, V. F. Enteropathogenic and enterohemorrhagic Escherichia coli infections: translocation, translocation, translocation. *Infect Immun* **73**, 2573-2585, doi:10.1128/IAI.73.5.2573-2585.2005 (2005).
116 Jarvis, K. G. *et al.* Enteropathogenic Escherichia coli contains a putative type III secretion system necessary for the export of proteins involved in attaching and effacing lesion formation. *Proc Natl Acad Sci U S A* **92**, 7996-8000, doi:10.1073/pnas.92.17.7996 (1995).
117 Garrity, G., Brenner, D., Kreig, N. & Staley, J. *Bergey's Manual of Systematic Bacteriology Vol. 2 Part C. The Alpha-, Beta-, Delta-, and Epsilonproteobacteria*. (2005).
118 Swenson, C. E. & Sadikot, R. T. Achromobacter respiratory infections. *Ann Am Thorac Soc* **12**, 252-258, doi:10.1513/AnnalsATS.201406-288FR (2015).
119 Furter, M., Sellin, M. E., Hansson, G. C. & Hardt, W. D. Mucus Architecture and Near-Surface Swimming Affect Distinct Salmonella Typhimurium Infection Patterns along the Murine Intestinal Tract. *Cell Rep* **27**, 2665-2678 e2663, doi:10.1016/j.celrep.2019.04.106 (2019).
120 Bhattacharjee, T. & Datta, S. S. Bacterial hopping and trapping in porous media. *Nat Commun* **10**, 2075, doi:10.1038/s41467-019-10115-1 (2019).
121 Steimle, A. *et al.* Flagellin hypervariable region determines symbiotic properties of commensal Escherichia coli strains. *PLoS Biol* **17**, e3000334, doi:10.1371/journal.pbio.3000334 (2019).
122 Wolfson, E. B. *et al.* The interaction of Escherichia coli O157 :H7 and Salmonella Typhimurium flagella with host cell membranes and cytoskeletal components. *Microbiology (Reading)* **166**, 947-965, doi:10.1099/mic.0.000959 (2020).
123 Beatson, S. A., Minamino, T. & Pallen, M. J. Variation in bacterial flagellins: from sequence to structure. *Trends in microbiology* **14**, 151-155, doi:10.1016/j.tim.2006.02.008 (2006).
124 Macnab, R. M. Genetics and biogenesis of bacterial flagella. *Annu.Rev.Genet.* **26**, 131-158 (1992).
125 Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**, 6640-6645, doi:10.1073/pnas.120163297 (2000).
126 Schymkowitz, J. *et al.* The FoldX web server: an online force field. *Nucleic Acids Res* **33**, W382-388, doi:10.1093/nar/gki387 (2005).
127 Xiong, P., Zhang, C., Zheng, W. & Zhang, Y. BindProfX: Assessing Mutation-Induced Binding Affinity Change by Protein Interface Profiles with Pseudo-Counts. *J Mol Biol* **429**, 426-434, doi:10.1016/j.jmb.2016.11.022 (2017).
128 Pottash, A. E., McKay, R., Virgile, C. R., Ueda, H. & Bentley, W. E. TumbleScore: Run and tumble analysis for low frame-rate motility videos. *Biotechniques* **62**, 31-36, doi:10.2144/000114493 (2017).
129 Rotter, C., Mühlbacher, S., Salamon, D., Schmitt, R. & Scharf, B. Rem, a new transcriptional activator of motility and chemotaxis in Sinorhizobium meliloti. *J Bacteriol* **188**, 6932-6942, doi:10.1128/jb.01902-05 (2006).
130 Frank, J. *et al.* SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* **116**, 190-199, doi:10.1006/jsbi.1996.0030 (1996).

131 Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296, doi:10.1038/nmeth.4169 (2017).
132 Kreutzberger, M. A. *et al.* Structural Studies of Straight and Supercoiled Flagellar Filaments from Campylobacter jejuni. *Biophysical Journal* **120**, 174a-175a, doi:10.1016/j.bpj.2020.11.1227 (2021).
133 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).
134 Zwart, P. H. *et al.* Automated structure solution with the PHENIX suite. *Methods Mol Biol* **426**, 419-435, doi:10.1007/978-1-60327-058-8_28 (2008).
135 Afonine, P. V. *et al.* Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol* **74**, 531-544, doi:10.1107/s2059798318006551 (2018).
136 Waterhouse, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**, W296-w303, doi:10.1093/nar/gky427 (2018).
137 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J Mol Biol* **215**, 403-410, doi:10.1016/s0022-2836(05)80360-2 (1990).
138 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589, doi:10.1038/s41586-021-03819-2 (2021).
139 Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci* **30**, 70-82, doi:10.1002/pro.3943 (2021).
140 Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539, doi:10.1038/msb.2011.75 (2011).
141 Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9**, 676-682, doi:10.1038/nmeth.2019 (2012).
142 Pazour, G. J., Agrin, N., Leszyk, J. & Witman, G. B. Proteomic analysis of a eukaryotic cilium. *J Cell Biol* **170**, 103-113, doi:10.1083/jcb.200504008 (2005).
143 Calldine, C. R. Change of waveform in bacterial flagella: The role of mechanics at the molecular level. *Journal of Molecular Biology* **118**, 457-479, doi:https://doi.org/10.1016/0022-2836(78)90285-1 (1978).
144 Berg, H. C. & Anderson, R. A. Bacteria swim by rotating their flagellar filaments. *Nature* **245**, 380-382, doi:10.1038/245380a0 (1973).
145 O'Brien, E. J. & Bennett, P. M. Structure of straight flagella from a mutant Salmonella. *J Mol Biol* **70**, 133-152, doi:10.1016/0022-2836(72)90168-4 (1972).
146 Yamashita, I. *et al.* Structure and switching of bacterial flagellar filaments studied by X-ray fiber diffraction. *Nature Structural Biology* **5**, 125-132, doi:10.1038/nsb0298-125 (1998).
147 Maki-Yonekura, S., Yonekura, K. & Namba, K. Conformational change of flagellin for polymorphic supercoiling of the flagellar filament. *Nat Struct Mol Biol* **17**, 417-422, doi:10.1038/nsmb.1774 (2010).
148 Asakura, S. Polymerization of flagellin and polymorphism of flagella. *Adv Biophys* **1**, 99-155 (1970).
149 Calladine, C. R. Construction of bacterial flagella. *Nature* **255**, 121-124 (1975).

150 Montemayor, E. J. *et al.* Flagellar Structures from the Bacterium Caulobacter crescentus and Implications for Phage varphi CbK Predation of Multiflagellin Bacteria. *J Bacteriol* **203**, doi:10.1128/JB.00399-20 (2021).
151 Kato, S., Okamoto, M. & Asakura, S. Polymorphic transition of the flagellar polyhook from Escherichia coli and Salmonella typhimurium. *J. Mol. Biol.* **173**, 463-476, doi:10.1016/0022-2836(84)90391-7 (1984).
152 Shibata, S., Matsunami, H., Aizawa, S. I. & Wolf, M. Torque transmission mechanism of the curved bacterial flagellar hook revealed by cryo-EM. *Nat Struct Mol Biol* **26**, 941-945, doi:10.1038/s41594-019-0301-3 (2019).
153 Kato, T., Makino, F., Miyata, T., Horváth, P. & Namba, K. Structure of the native supercoiled flagellar hook as a universal joint. *Nat Commun* **10**, 5295, doi:10.1038/s41467-019-13252-9 (2019).
154 Bardy, S. L., Mori, T., Komoriya, K., Aizawa, S. & Jarrell, K. F. Identification and localization of flagellins FlaA and FlaB3 within flagella of Methanococcus voltae. *J Bacteriol* **184**, 5223-5233, doi:10.1128/jb.184.19.5223-5233.2002 (2002).
155 Ghosh, A. & Albers, S. V. Assembly and function of the archaeal flagellum. *Biochem Soc Trans* **39**, 64-69, doi:10.1042/bst0390064 (2011).
156 Gambelli, L. *et al.* An archaellum filament composed of two alternating subunits. *Nat Commun* **13**, 710, doi:10.1038/s41467-022-28337-1 (2022).
157 Poweleit, N. *et al.* CryoEM structure of the Methanospirillum hungatei archaellum reveals structural features distinct from the bacterial flagellum and type IV pilus. *Nat Microbiol* **2**, 16222, doi:10.1038/nmicrobiol.2016.222 (2016).
158 Bayley, D. P. & Jarrell, K. F. Further evidence to suggest that archaeal flagella are related to bacterial type IV pili. *J Mol Evol* **46**, 370-373 (1998).
159 Thomas, N. A., Bardy, S. L. & Jarrell, K. F. The archaeal flagellum: a different kind of prokaryotic motility structure. *FEMS Microbiol Rev* **25**, 147-174, doi:10.1111/j.1574-6976.2001.tb00575.x (2001).
160 Cohen-Krausz, S. & Trachtenberg, S. The structure of the archeabacterial flagellar filament of the extreme halophile Halobacterium salinarum R1M1 and its relation to eubacterial flagellar filaments and type IV pili. *J Mol Biol* **321**, 383-395, doi:10.1016/s0022-2836(02)00616-2 (2002).
161 Subramaniam, S., Earl, L. A., Falconieri, V., Milne, J. L. & Egelman, E. H. Resolution advances in cryo-EM enable application to drug discovery. *Curr. Opin. Struct. Biol.* **41**, 194-202, doi:10.1016/j.sbi.2016.07.009 (2016).
162 Zhang, R. & Nogales, E. A new protocol to accurately determine microtubule lattice seam location. *J. Struct. Biol.* **192**, 245-254, doi:10.1016/j.jsb.2015.09.015 (2015).
163 Rice, W. J. *et al.* Routine determination of ice thickness for cryo-EM grids. *J Struct Biol* **204**, 38-44, doi:10.1016/j.jsb.2018.06.007 (2018).
164 Kamiya, R. & Asakura, S. Helical transformations of Salmonella flagella in vitro. *J Mol Biol* **106**, 167-186, doi:10.1016/0022-2836(76)90306-5 (1976).
165 Wang, F. *et al.* The structures of two archaeal type IV pili illuminate evolutionary relationships. *Nat Commun* **11**, 3424, doi:10.1038/s41467-020-17268-4 (2020).
166 Zarschler, K. *et al.* Protein tyrosine O-glycosylation--a rather unexplored prokaryotic glycosylation system. *Glycobiology* **20**, 787-798, doi:10.1093/glycob/cwq035 (2010).

167 Meshcheryakov, V. A. *et al.* High-resolution archaellum structure reveals a conserved metal-binding site. *EMBO Rep* **20**, doi:10.15252/embr.201846340 (2019).
168 Sonnhammer, E. L., von Heijne, G. & Krogh, A. A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol* **6**, 175-182 (1998).
169 Poweleit, N. *et al.* CryoEM structure of the *Methanospirillum hungatei* archaellum reveals structural features distinct from the bacterial flagellum and type IV pili. *Nat Microbiol* **2**, 16222, doi:10.1038/nmicrobiol.2016.222 (2016).
170 Daum, B. *et al.* Structure and in situ organisation of the *Pyrococcus furiosus* archaellum machinery. *eLife* **6**, doi:10.7554/eLife.27470, doi:10.7554/eLife.27470 (2017).
171 Thormann, K. M. & Paulick, A. Tuning the flagellar motor. *Microbiology (Reading)* **156**, 1275-1283, doi:10.1099/mic.0.029595-0 (2010).
172 Kamiya, R. & Asakura, S. Helical transformations of Salmonella flagella in vitro. *J. Mol. Biol.* **106**, 167-186 (1976).
173 Mimori-Kiyosue, Y., Vonderviszt, F., Yamashita, I., Fujiyoshi, Y. & Namba, K. Direct interaction of flagellin termini essential for polymorphic ability of flagellar filament. *Proc Natl Acad Sci U S A* **93**, 15108-15113, doi:10.1073/pnas.93.26.15108 (1996).
174 Caspar, D. L. Bacterial flagellar coiling explained by slip-and-click strand switching. *Nat Struct Biol* **5**, 92-94, doi:10.1038/nsb0298-92 (1998).
175 Yu, X., Horiguchi, T., Shigesada, K. & Egelman, E. H. Three-dimensional reconstruction of transcription termination factor rho: orientation of the N-terminal domain and visualization of an RNA- binding site. *J. Mol. Biol.* **299**, 1279-1287 (2000).
176 Skordalakes, E. & Berger, J. M. Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading. *Cell* **114**, 135-146 (2003).
177 Abby, S. S. & Rocha, E. P. The non-flagellar type III secretion system evolved from the bacterial flagellum and diversified into host-cell adapted systems. *PLoS Genet* **8**, e1002983, doi:10.1371/journal.pgen.1002983 (2012).
178 Zheng, W. *et al.* Cryoelectron-microscopy structure of the enteropathogenic Escherichia coli type III secretion system EspA filament. *Proceedings of the National Academy of Sciences* **118**, e2022826118, doi:10.1073/pnas.2022826118 (2021).
179 Deng, W. *et al.* Assembly, structure, function and regulation of type III secretion systems. *Nat Rev Microbiol* **15**, 323-337, doi:10.1038/nrmicro.2017.20 (2017).
180 Knutton, S. *et al.* A novel EspA-associated surface organelle of enteropathogenic Escherichia coli involved in protein translocation into epithelial cells. *Embo j* **17**, 2166-2176, doi:10.1093/emboj/17.8.2166 (1998).
181 Albers, S. V. & Jarrell, K. F. The archaellum: how Archaea swim. *Front Microbiol* **6**, 23, doi:10.3389/fmicb.2015.00023 (2015).
182 Gaffney, E. A., Ishimoto, K. & Walker, B. J. Modelling Motility: The Mathematics of Spermatozoa. *Front Cell Dev Biol* **9**, 710825, doi:10.3389/fcell.2021.710825 (2021).
183 Langousis, G. & Hill, K. L. Motility and more: the flagellum of Trypanosoma brucei. *Nature reviews. Microbiology* **12**, 505-518, doi:10.1038/nrmicro3274 (2014).
184 Wirth, R. Response to Jarrell and Albers: seven letters less does not say more. *Trends in microbiology* **20**, 511-512, doi:10.1016/j.tim.2012.07.007 (2012).

185 Egelman, E. H. A robust algorithm for the reconstruction of helical filaments using single-particle methods. *Ultramicroscopy* **85**, 225-234, doi:10.1016/s0304-3991(00)00062-0 (2000).
186 Mastronarde, D. N. & Held, S. R. Automated tilt series alignment and tomographic reconstruction in IMOD. *J Struct Biol* **197**, 102-113, doi:10.1016/j.jsb.2016.07.011 (2017).
187 Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer visualization of three-dimensional image data using IMOD. *J Struct Biol* **116**, 71-76, doi:10.1006/jsbi.1996.0013 (1996).
188 Noske, A. B., Costin, A. J., Morgan, G. P. & Marsh, B. J. Expedited approaches to whole cell electron tomography and organelle mark-up in situ in high-pressure frozen pancreatic islets. *J Struct Biol* **161**, 298-313, doi:10.1016/j.jsb.2007.09.015 (2008).