Abstract

Structural Basis of Flagellar Filament Asymmetry and Supercoil Templating by *Leptospira* Sheath Factors

Megan Rose Brady

2022

The spirochete family of bacteria, including pathogens such as *Borrelia burgdorferi* (Lyme disease), *Treponema pallidum* (syphilis), and *Leptospira interrogans* (leptospirosis), generates a unique, corkscrew-like form of motility that is crucial for pathogenicity. This motility is driven by flagella that are periplasmic rather than extracellular in nature and are composed of a FlaB core and a FlaA sheath. Purified *Leptospira* flagellar filaments form tight coils and contain the coiling proteins FcpA and FcpB, in addition to two FlaA isoforms. Loss of either the FlaA or Fcp proteins results in straighter flagella and bacteria that are non-motile and non-pathogenic.

I have used cryo-electron microscopy to solve the structure of three *Leptospira* flagellar mutants to near-atomic resolution. First, I resolved the structure of an *fcpB*$^{-}$ mutant, allowing for identification of the FcpB density in the sheath region of the asymmetric wild-type structure. I also developed a method to quantify the curvature of purified flagellar filaments, allowing us to investigate how the flagellar forms are affected by various mutations.

Next, I determined the structure of a *fcpA*$^{-}$ mutant, and found that the FlaA sheath components strictly localize to the inner curvature. We also discovered a previously uncharacterized flagellar sheath protein, which we have named FlaAP (FlaA-associated protein). We found that these sheath proteins interact with the core through interactions with glycans, and that the FlaB core appears to have ten 'L' protofilaments and one 'R' protofilament, with the 'R' protofilament directly underlying the sheath proteins and disrupting the helical symmetry of the core.

Third, I solved the structure of an *flaA2*$^{-}$ mutant. In doing so, I found that whereas the Fcp coiling factors bind solely to the outer curvature of the wild-type filament, forming an extensive helical lattice, this Fcp lattice extends symmetrically around the entire FlaB core when the FlaA sheath factors are missing. These three mutant flagellar structures provide crucial insight into the role of the flagellar sheath factors and highlight the importance of the supercoiled wild-type form for motility and virulence. This work has implications for other spirochetes as well as other flagellated bacteria, where structural work has mainly been focused on flagellar filaments rendered straight by specific mutations, instead of the naturally occurring supercoiled forms as we have described here.

Structural Basis of Flagellar Filament Asymmetry and Supercoil Templating by *Leptospira* Sheath Factors

A Dissertation

Presented to the Faculty of the Graduate School

Of

Yale University

In Candidacy for the Degree of

Doctor of Philosophy

By

Megan Rose Brady

Dissertation Director: Charles V. Sindelar

May 2022

Table of Contents

Index of Figures

Index of Tables

Acknowledgments

I have had the tremendous fortune of having an incredible support system during my PhD journey. Thank you to my advisor, Dr. Charles Sindelar, for your support and encouragement, and for always believing in my projects. Thank you to Dr. Albert Ko, for all of your support over the years. Thank you to Drs. Jonathon Howard and Jun Liu, for their incredibly helpful encouragement and advice during committee meetings. And thanks to Dr. Enrique de la Cruz for your infectious enthusiasm and love of science.

Thank you also to all of the lab members through the years. A special thank you to Kimberley Gibson, who taught me everything I know about cryo-EM, and whose project really paved the way for all of my work. Thank you to Fabiana San Martin and Dr. Alejandro Buchiazzo, for an amazing collaboration. Thank you to Garrett Debs for your help and patience with everything computational, and to Elsio Wunder and Cate Muenker for helping me with all things *Leptospira*. And thank you to Andrew, Karukriti, Xueqi, Mike, Shashank, and all the other lab members over the years, who always made going to work enjoyable.

I truly believe that I would not be where I am without my experiences at UConn. Thank you to Dr. Andrei Alexandrescu for giving me a start in research, and to Annie, Becca, and Richard for showing me how fun a lab could be.

I have such an amazing and supportive group of friends and family. To all of my classmates and friends in New Haven, thank you for the pizza trips, hikes, and game nights. To Megan A. and Zoryana, thank you for our weekly Catan games and catch-up sessions, and I hope that we can have more trips to Delaware in the future. To Gram,

thank you for our monthly chats that always brighten my day. To all my grandparents, thank you for your unending love, support, and cookies! To the 'Sean Cole crew' (including honorary members Parker and Andy), thank you for making me laugh every day, and for your continuous support and encouragement. And a special shoutout to Uncle Dave- from train rides with Kevin to life lessons at Dusty's, you remind me to always find the fun in life.

To my sister Kayla, my best friend in the world- thank you so much for always being there. Thank you for always understand my references when no one else does. You are just so incredibly supportive, and always know what I need to hear. I'm excited to be in the same city again soon, and I can't wait for your Water Safari wedding this summer!

And to my parents- thank you for absolutely everything. You taught me to always work hard and encouraged me to reach for the stars, never doubting for a second that I would reach them. Dad, thank you for our sports chats, for sending buddy pictures when I need them most, and for being just the best dad in the world. You taught me how to love learning, and your support in all things math and science really started me down this path. Mom, you have always been there for me in every possible way. Your unending kindness and support have shaped every aspect of my life in the best way possible. I know that if I ever need anything, you will be there in a second. Mom, Dad, and Kayla, just thank you- for everything.

Background

Bacterial flagellar composition

Bacteria are incredibly diverse, with cell volumes that range from less than 0.1 μm^3 (Duda, Suzina et al. 2012) to 3,000,000 μm^3 (Schulz, Brinkhoff et al. 1999), and inhabiting environments from hot springs, to the soil, the ocean, and the human body (Lozupone and Knight 2007, Human Microbiome Project 2012). Motility is crucial for many bacteria, allowing for the organism to move towards nutrients and away from toxins, and is often required for pathogenic bacteria to infect their host; many bacteria achieve this motility by using rotation of the flagella to propel through the environment (Leifson 1960, Harshey 2003).

Components of the flagellum

The flagellum is composed of three components: a flagellar motor, a hook, and a filament. The flagellar motor (the basal body), composed of over a dozen proteins, is a large macromolecular complex that utilizes an ion motive force to drive rotation (Macnab 2003, Mandadapu, Nirody et al. 2015, Minamino and Imada 2015, Chang, Carroll et al. 2021). The flagellar filament is ~10 μm long and is connected to the motor via the joint-like hook (Figure 1A) (Macnab 2003, Nakamura and Minamino 2019, Yamaguchi, Toma

et al. 2020). Rotation of the motor causes the hook and filament to rotate, allowing the bacteria to 'swim' (Macnab 2003, Nakamura and Minamino 2019, Chang, Carroll et al. 2021). The motor, hook, flagellar filament, and flagellar motility are described in more detail further below.

Types of flagellar arrangements

Bacteria such as *Escherichia coli*, *Bacillus subtilis*, *Kurthia spp.*, and *Salmonella enterica* have peritrichous flagella, with dozens of external flagellar filaments that decorate the entire cell surface (Fujii, Shibata et al. 2008, Roux, El Karkouri et al. 2012). Other bacteria, such as *Vibrio cholerae* and *Caulobacter crescentus*, have monotrichous polar flagella, and contain a single flagellum extruding from the cell (Lele, Roland et al. 2016, Echazarreta and Klose 2019), while species such as *Campylobacter jejuni* (Burnham, Kolar et al. 2020) have polar amphitrichous flagella with one flagellum on polar ends. Bacteria such as *Helicobacter pylori* have polar lophotrichous flagella, containing a tuft of flagella at either one or both ends of the cell (Burnham, Kolar et al.

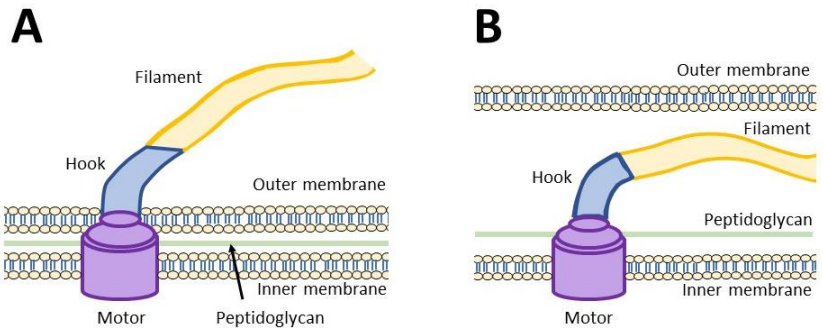

Figure 1. Three components comprise the bacterial flagellum. **A,** General arrangement present in an externally flagellated bacterium. The flagellar motor is embedded in the inner membrane; this connects to a hook, which is then connected to the flagellar filament. This filament extrudes from the outer membrane of the cell. **B,** Arrangement of a spirochete endoflagella. In these bacteria, the filament is located solely within the periplasmic space between the inner and outer membranes, and does not extrude from the cell.

2020, Chu, Liu et al. 2020). The spirochete family of bacteria, including *Leptospira spp.*, *Borrelia burgdorferi*, and *Treponema spp.*, contain one or more internal, periplasmic flagella which do not extend from the cell body (Wolgemuth 2015) (Figure 1B). This family is described in detail further below.

Architecture of *Salmonella* flagella

Salmonella has been used as a model for flagellar studies. Similar to the flagellar filaments of many other bacteria, *Salmonella* filaments are composed of an 11-fold repeat of the protein flagellin (FliC) (Yonekura, Maki-Yonekura et al. 2003, Maki-Yonekura, Yonekura et al. 2010). FliC is 491 residues in length, and contains 4 domains: the D0 and D1 domains, which each form a coiled-coil, and the globular D2 and D3 domains, with short linkers connecting each of the domains (Yonekura, Maki-Yonekura et al. 2003). A majority of flagellar filament proteins from other species, including the FlaB proteins found in the flagellar filament core of all spirochetes, have a high homology to this FliC flagellin molecule, particularly in the D0/D1 coiled-coil domains (Norris, Charon et al. 1988, Pallen, Penn et al. 2005).

All known bacterial flagellar filaments exhibit a similar 11-fold symmetric structure (Beatson, Minamino et al. 2006, Kreutzberger, Ewing et al. 2020). In *Salmonella*, there are two forms of flagellin: a ‘L’ form and an ‘R’ form, for which each ‘R’ monomer is ~1.5% shorter than an ‘L’ monomer (Yamashita, Hasegawa et al. 1998, Calladine, Luisi et al. 2013). When all 11 protofilaments are comprised of the ‘L’ or ‘R’ type, the filament is straight with either a left- or right-handed helix, respectively (Yonekura, Maki-Yonekura et al. 2003, Maki-Yonekura, Yonekura et al. 2010, Wang,

Jiang et al. 2012, Calladine, Luisi et al. 2013, Wang, Burrage et al. 2017). These 'all-L' and 'all-R' filaments can be a result of mutations in particular regions of the flagellum; for example, the mutation A449V causes each protofilament to be in an 'R' form, while the mutation G426A causes each protofilament to adopt an 'L' orientation (Yonekura, Maki-Yonekura et al. 2003, Maki-Yonekura, Yonekura et al. 2010). However, a combination of 'L' and 'R' protofilaments forces the filament to adopt a range of different supercoil forms (Leifson 1960, Kamiya and Asakura 1976), with any shorter 'R' protofilaments located together along the shorter side of the curved filament (Calladine 1975, Calladine, Luisi et al. 2013). The '9+2' form of the filament (nine 'L' protofilaments and two 'R' protofilaments) is considered the 'normal' form (Leifson 1960, Fujii, Shibata et al. 2008), with an expected helical diameter and pitch of ~0.6 μm and ~2.55 μm^{-1}, respectively (Figure 2) (Fujii, Shibata et al. 2008).

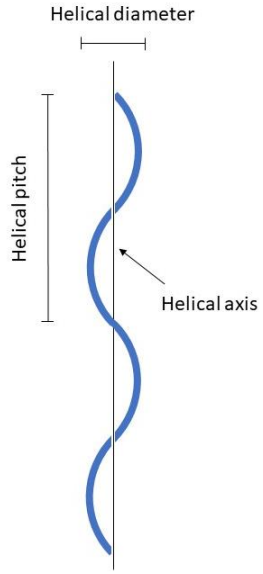

Figure 2. Helical parameters of a filament. The filament is oriented longitudinally with respect to the helical axis. The helical diameter is measured as the largest lateral distance of the filament. The helical pitch is the distance required to complete one turn of the helix.

Bacteria with multi-protein flagellar filaments

Many bacterial flagella are similar to *Salmonella*, with only one protein (flagellin) making up the filament. However, several species, including *V. cholerae*, *C. crescentus*, *H. pylori*, and *C. jejuni*, as well as all spirochetes, have flagella that are composed of multiple proteins (Guerry, Alm et al. 1991, Josenhans, Labigne et al. 1995, Klose and Mekalanos 1998, Ely, Ely et al. 2000, Chu, Liu et al. 2020). In *Campylobacter*, there are two flagella genes (*flaA* and *flaB*), which are 95% identical; in these bacteria, FlaA is the 'major' flagellar protein, as deletion of *flaB* does not affect the overall filament structure or motility of the bacteria, while the loss of FlaA results in shorter flagella and decreased bacterial motility (Guerry, Alm et al. 1991, Alm, Guerry et al. 1993). However, it was recently observed that FlaB is important for defense against bacteriophages, as mutants with a *flaB* deletion were significantly more susceptible to the phages than wild-type bacteria (Lis and Connerton 2016). Similar to *Campylobacter*, *Helicobacter* flagella also contain two highly homologous proteins (FlaA and FlaB, with 56% sequence identity), of which FlaA is required for proper motility of the organism and FlaB is located only at the flagellar base (Josenhans, Labigne et al. 1995, Eaton, Suerbaum et al. 1996). However, unlike in *Campylobacter* where the loss of FlaB does not affect motility or flagellar filament formation (Guerry, Alm et al. 1991), *H. pylori flaB* deletion mutants had decreased motility: a mutant lacking both FlaA and FlaB was motility deficient, with no observed flagella, and was unable to colonize the host organism (Josenhans, Labigne et al. 1995, Eaton, Suerbaum et al. 1996).

Other bacteria contain additional complexity beyond just two flagellin homologs. *C. crescentus* has six flagellin molecules, FliJ-O, all of which contain a similar coiled-

coil structure to the D0/D1 domains of flagellin (Ely, Ely et al. 2000, Montemayor, Ploscariu et al. 2021), and *V. cholera* has five flagellin molecules (FlaA-E), all highly homologous to each other, of which only FlaA is essential for filament formation (Klose and Mekalanos 1998). Spirochete flagella consist of FlaB and FlaA isoforms (of which only FlaB is a flagellin homolog) (Norris, Charon et al. 1988, Charon, Greenberg et al. 1992, Li, Motaleb et al. 2000). Several species also have additional flagellar proteins (Wunder, Figueira et al. 2016, Wunder, Slamti et al. 2018, Kurniyati, Liu et al. 2019, Nakamura 2020); spirochete flagella are described in more detail further below.

Flagellar glycosylation

Many bacterial flagella are often modified through glycosylation (Logan 2006, Chu, Liu et al. 2020); either via O-linkage (at Ser/Thr residues) or N-linkage (Asn-X-Ser/Thr, with the glycan attached at Asn) (Spiro 2002). The flagella of *C. jejuni* are glycosylated through O-linkage, with 17-19 sites of modification on each FlaA monomer (Thibault, Logan et al. 2001, Kreutzberger, Ewing et al. 2020). *H. pylori*, which has a similar flagellar composition, has 7 sites of O-linked glycosylation on each FlaA and 10 sites on each FlaB (Schirm, Soo et al. 2003). Both *Campylobacter* and *H. pylori* are modified with the monosaccharide pseudaminic acid (Pse) (Goon, Kelly et al. 2003, Schirm, Soo et al. 2003, Salah Ud-Din and Roujeinikova 2018). However, not every species exhibits flagellar glycosylation; for example, in *Salmonella enteria* serovar Typhimurium, surface-exposed lysine residues on the flagellum instead undergo methylation, which helps to promote adhesion onto the host (Horstmann, Lunelli et al. 2020).

In *Campylobacter*, mutation of three of the glycan sites on the FlaA flagllin resulted in shorter flagellar filaments and reduced cell motility, though mutations to the remaining glycan sites had no effect on motility (Ewing, Andreishcheva et al. 2009). In addition, when Pse biosynthesis genes were deleted, the cells were nonmotile and had no visible flagellar filaments, indicating the importance of glycosylation for flagellar motility (Goon, Kelly et al. 2003). An *H. pylori* mutant that exhibited a ~3-fold increase in the number of flagellar glycans was hyper-motile (Asakura, Churin et al. 2010), while mutations in Pse synthesis genes (Schirm, Soo et al. 2003) and a Pse transferase gene (Yang, Kao et al. 2021) resulted in motility-deficient cells lacking flagellar filaments, further suggesting a link between flagellar glycosylation and bacterial motility.

All six flagellins (FliJ-O) of *C. crescentus* are glycosylated in an O-linked manner (Ardissone, Kint et al. 2020, Montemayor, Ploscariu et al. 2021). The spirochete *Treponema denticola* exhibits flagellar glycosylation, as described below (Kurniyati, Kelly et al. 2017). The *Leptospira* flagellar core is predicted to be glycosylated based on a sequence similarity with two of the *T. denticola* glycosylation sites, though this glycosylation in *Leptospira* have not been experimentally observed (Holzapfel, Bonhomme et al. 2020). The crucial role that these nearly ubiquitous post-translational flagellar modifications play in the motility and virulence of the cells indicates the importance in understanding the entire flagellar network.

Bacterial motility

Flagellar motor

The rotation of the flagellar motor causes the flagellar filament (attached via the hook) to rotate, driving the motility of the cell. Most bacterial motors are large machines, with dozens of conserved protein components that are crucial for proper rotation, including the rod and C-, MS-, and LP-rings (Chang, Carroll et al. 2021). The rod (composed of FlgB, FlgC, FlgF, and FlgG) spans the length of the periplasmic space, and functions as the shaft of the motor (Minamino and Imada 2015, Chang, Carroll et al. 2021). The MS-ring (composed of FliF) spans the inner membrane, with the M- ring embedded in the membrane and the S-ring on the periplasmic side (Chang, Carroll et al. 2021). The C-ring (composed of FliG, FliM, and FliN) attaches to the cytoplasmic side of the MS-ring; together, the MS- and C-rings form the 'rotor' of the motor, with the C-ring crucial for the switch between clockwise and counter-clockwise rotation (Chen, Beeby et al. 2011, Chang, Carroll et al. 2021). The MS-ring encircles the proximal part of the rod within the inner membrane (Kawamoto, Miyata et al. 2021, Tan, Zhang et al. 2021). The 'stator' of the motor is an ion channel composed of MotA and MotB, which are embedded in the inner membrane above the C-ring and span to the periplasmic peptidoglycan layer (Minamino and Imada 2015, Chang, Carroll et al. 2021).

Upon interactions between the stator and the rotor, ions travel across the cytoplasmic membrane, generating torque (Chang, Carroll et al. 2021). The LP-ring functions as a bushing, with the P-ring (composed of FlgI and located in the peptidoglycan layer) and the L-ring (composed of FlgH and located adjacent to the P-ring and embedded in the outer membrane) both surrounding the rod (Minamino and Imada

2015, Tan, Zhang et al. 2021). In addition to these motor components, there is also an export apparatus (composed of FlhA, FlhB, FliO, FliP, FliQ, and FliR) located within the MS-ring, required for the proper flagellar assembly (Minamino and Imada 2015). Several species have additional motor components: *Vibrio alginolyticus*, for example, also has an H- and T-ring located in the periplasm, as well as an additional O-ring, adding further complexity to this important protein complex (Chang, Carroll et al. 2021).

Flagellar hook

Each flagellar filament in the bacterium is connected to a flagellar motor via a hook region; this hook is a 'joint' between the two and is composed of the protein FlgE (Courtney, Cozy et al. 2012, Matsunami, Barker et al. 2016, Kato, Makino et al. 2019). This hook consists of ~120-130 subunits of FlgE, arranged with 11-fold symmetry, and is ~55 ± 7 nm long (Kato, Makino et al. 2019, Lynch, Miller et al. 2019). When *flgE* is mutated, flagella do not form and the bacteria are nonmotile, highlighting the importance of this structure for proper filament formation (Li, Ruby et al. 1996, Courtney, Cozy et al. 2012, Miller, Miller et al. 2016). The hook extensively interacts with the distal rod (FlgG) of the flagellar motor, allowing for torque to be transmitted between the two (Tan, Zhang et al. 2021). In spirochetes, FlgE forms an inter-subunit lysinoalanine (Lal) crosslink (Lynch, Miller et al. 2019); this crosslink is self-catalyzed and is required for full motility, as bacteria deficient in the Lal crosslink produce intact flagella but are motility deficient (Miller, Miller et al. 2016).

‘Run’ and ‘tumble’ motility of *E. coli*

E. coli motility has been well studied: the bacteria move in a ‘run-and-tumble’ manner, in which counter-clockwise rotation of the flagellar motor causes the flagella to bundle behind the cell, driving the bacterium forward (‘run’). When some of the motors switch to a clockwise rotation, the flagella no longer bundle, and the cell ‘tumbles’ (Berg and Brown 1972, Berg 1975, Sourjik and Wingreen 2012). This abrupt switch to a clockwise rotation occurs in response to the binding of a phosphorylated CheY protein to the base of the motor (Sarkar, Paul et al. 2010, Sourjik and Wingreen 2012). CheY phosphorylation occurs upon the binding of an attractant or repellent to chemoreceptors that control the activity of the corresponding kinase, allowing for the cell to ‘run’ more frequently and ‘tumble’ less often in the presence of an attractant (Webre, Wolanin et al. 2003, Sourjik and Wingreen 2012, Chang, Carroll et al. 2021).

Spirochete flagella

The spirochete family of bacteria includes the pathogenic species *Borrelia burgdorferi* (Lyme disease), *Treponema pallidum* (syphilis), *Treponema denticola* (periodontal disease), *Brachyspira hyodysenteriae* (swine dysentery), and *Leptospira interrogans* (leptospirosis) (Li, Motaleb et al. 2000). Leptospirosis is a neglected tropical disease that can cause fever and malaise; 5-10% of cases are ‘severe’ and can cause Weil’s disease and pulmonary hemorrhage (Costa, Hagan et al. 2015). There are over one million cases and 58,000 deaths caused by leptospirosis reported every year (Costa, Hagan et al. 2015). *Leptospira* infects humans and animals through contact with either

contaminated water or urine from an affected animal host (most commonly rats) (Ko, Goarant et al. 2009, Evangelista and Coburn 2010). The *Leptospira* genus also contains saprophytic species such as *L. biflexa*, which shares ~61% of its genes (including all known flagellar genes) with the pathogenic *Leptospira* strains (Picardeau, Bulach et al. 2008, Evangelista and Coburn 2010, Fouts, Matthias et al. 2016).

Spirochetes are characterized by their spiral-shaped cell body and a unique, 'corkscrew' mode of motility (Holt 1978, Li, Motaleb et al. 2000, Wolgemuth 2015). Unlike most bacteria, spirochetes have flagellar filaments that are composed of multiple protein components that form a core and sheath; these flagellar filaments are located completely within the periplasmic space (endoflagella), and do not extend externally from the cell (Charon, Greenberg et al. 1992, Li, Motaleb et al. 2000).

Spirochetes differ in the number and length of flagella, ranging from one non-overlapping flagellum on each end (*Leptospira spp.*) to hundreds of filaments on each end that overlap for much of the cell body (*Cristispira*) (Holt 1978, Paster 2010). Furthermore, spirochetes differ in the number of proteins that comprise the flagellum: all spirochetes contain the flagellar proteins FlaB, which is homologous to flagellin, and FlaA, thought to comprise the flagellar sheath in most spirochetes (Norris, Charon et al. 1988, Charon, Greenberg et al. 1992, Li, Motaleb et al. 2000). FlaB and FlaA are strongly conserved amongst the various spirochetes, though different species contain different numbers of FlaA and FlaB proteins (Charon, Greenberg et al. 1992, Li, Motaleb et al. 2000, Picardeau, Brenot et al. 2001). Furthermore, several spirochetes contain additional flagellar factors that are not found in every spirochete species (Wunder, Figueira et al. 2016, Wunder, Slamti et al. 2018, Kurniyati, Liu et al. 2019).

Borrelia burgdorferi

B. burgdorferi contains seven to eleven flagella extending from either end of the filament; these flagella overlap throughout much of the length of the bacterium (Ge, Li et al. 1998, Sterba, Vancova et al. 2008, Kudryashev, Cyrklaff et al. 2009). In *Borrelia*, these flagella are crucial for maintaining the helical nature of the cell body, as cells lacking flagellar filaments are rod-shaped rather than spiral shaped (Motaleb, Corum et al. 2000, Charon, Cockburn et al. 2012). Purified *Borrelia* flagella have an average diameter of ~18.1 nm (Ge, Li et al. 1998), similar to the diameter of ~17.6 nm observed *in situ* (Kudryashev, Cyrklaff et al. 2009). Both the FlaB and FlaA flagellar proteins of *Borrelia* do not contain N-linked glycans (Sterba, Vancova et al. 2008), though other groups report glycosylation of FlaA (Ge, Li et al. 1998).

Borrelia contain one FlaB and one FlaA (Charon, Cockburn et al. 2012); *flaA* is expressed at ~10% of the level of *flaB*, in contrast to most spirochetes which express similar levels of these two flagellar proteins (Ge, Li et al. 1998, Motaleb, Sal et al. 2004). Purification of these flagella requires a gentle Triton-based procedure, as harsher purifications using Sarkosyl detergent resulted in a loss of FlaA in the purified filaments (Ge, Li et al. 1998, Motaleb, Sal et al. 2004). The Sarkosyl-purified filaments that lack FlaA have a smaller helical diameter than the Triton-purified filaments (~0.28 μm and ~0.33 μm respectively), though filaments both lacking and containing FlaA have a left-handed supercoil and a similar helical pitch (~1.48 μm and ~1.5 μm respectively) (Charon, Goldstein et al. 1992, Ge, Li et al. 1998, Motaleb, Sal et al. 2004). In a mutant cell that does not produce flagellar filaments (HB19Fla-), *flaA* is still transcribed to a similar level as in wild-type cells, though no FlaB or FlaA is present (Ge, Li et al. 1998).

When *flaB* was mutated, FlaA levels were reduced to just ~13% of the wild-type FlaA levels, even though the *flaA* expression was similar to wild-type; flagellar filaments failed to form in these mutants, highlighting the importance of FlaB for the formation of the flagellum (Motaleb, Corum et al. 2000, Motaleb, Sal et al. 2004).

Treponema pallidum

T. pallidum filaments are ~20 nm in diameter; typically, cells have three flagella on each end that extend towards the center and overlap for much of the cell body, forming a parallel array with the neighboring flagella (Izard, Renken et al. 2009, Liu, Howell et al. 2010). *T. pallidum* contains three FlaB proteins (FlaB1-3) (Norris, Paster et al. 2006) and two FlaA paralogs (McGill, Edmondson et al. 2010, Kubanov, Runina et al. 2017, Xie, Xu et al. 2017); the FlaB core is glycosylated, though the FlaA sheath is not (Norris 1993, Wyss 1998). The FlaB core is ~11 nm in diameter and surrounds a ~6 nm central channel; the FlaA sheath then forms an outer sheath around this assembly (Liu, Howell et al. 2010). However, *T. pallidum* was only recently successfully cultured *in vitro* (Edmondson, Hu et al. 2018); this has made the genetic manipulation of this species challenging (Romeis, Tantalo et al. 2021), with no flagellar mutants known to exist.

Treponema denticola

T. denticola typically has 2-3 flagellum inserted into each end of the cell, which overlap along much of the cell body (Izard, Hsieh et al. 2008, Kurniyati, Liu et al. 2019). These filaments are ~3 to 6 μm long (Kurniyati, Liu et al. 2019), are ~20 nm in diameter (Izard, Hsieh et al. 2008), and consist of three FlaB isoforms and at least one (and

potentially three) FlaA proteins (Ruby, Li et al. 1997, Kurniyati, Kelly et al. 2017, Ng, Slakeski et al. 2019).

In addition to FlaA and FlaB, *T. denticola* flagella also contain the protein FlaG; this protein is not found in spirochetes outside of the *Treponema* and *Spirochaeta* genus, though it is found in bacteria such as *C. jejuni*, where it is involved in negatively regulating flagellar length but is not present in the filament itself (Inoue, Barker et al. 2018, Kurniyati, Liu et al. 2019). *T. denticola* FlaG, at only 14 kDa, is predicted (via AlphaFold2) to form a three-strand beta-sheet with helices at both the N- and C-termini (data not shown) (Jumper, Evans et al. 2021). *T. denticola flaG*$^{-}$ mutants have shorter flagella than wild-type cells (~0.3 μm and ~4 μm respectively), though the filament diameter (~21 nm) remains the same (Kurniyati, Liu et al. 2019). FlaG likely helps to support the flagellar filament, as *flaG*$^{-}$ mutants had ~30% less FlaA and ~60% less FlaB present in cell lysates after 24 hours, though the amounts of the motor and hook proteins remained unchanged (Kurniyati, Liu et al. 2019). Purified *flaA*$^{-}$ flagella did not have any FlaG present, suggesting that FlaG closely associates with FlaA to form the filament sheath (Kurniyati, Liu et al. 2019).

The flagellum of *T. denticola* is glycosylated by a novel ~450 Da glycan, which decorates all three FlaB isoforms in an O-linked manner; FlaA, however, is not glycosylated (Kurniyati, Kelly et al. 2017). Mass spectrometry analysis revealed that the number of glycan modifications was different between the three FlaB isoforms: FlaB1 has five sites of glycosylation, FlaB2 has 4-5 sites, and FlaB3 has 3-4 sites of glycosylation (Kurniyati, Kelly et al. 2017). This glycosylation likely occurs through a putative Pse synthase gene (PseI); deletion of this gene not only decreases the levels of

the *flaB* and *flaA* transcripts, but also greatly reduces the presence of the FlaB and FlaA proteins in cell lysates and abolishes all FlaB glycosylation (Kurniyati, Kelly et al. 2017). As the quantity of the flagellar hook protein FlgE is not affected in the PseI-mutant, this suggests that PseI-driven glycosylation is important for proper flagella assembly in *T. denticola* (Kurniyati, Kelly et al. 2017). This is further supported through a visual analysis of PseI-mutant flagella, which are significantly shorter than wild-type filaments (~96 nm versus ~1500 nm); the mutant bacteria themselves are non-motile, emphasizing the importance of glycosylation in bacterial motility (Kurniyati, Kelly et al. 2017).

Brachyspira hyodysenteriae

B. hyodysenteriae has 11-14 flagella at each cell end that overlap for much of the length of the cell (Paster 2010). Their flagella contain one FlaA and three FlaB (FlaB1-3) proteins (Li, Corum et al. 2000), and the flagellar sheath is glycosylated (Li, Dumas et al. 1993). This is the same FlaA/FlaB composition as *Borrelia* and *T. pallidum*; however, unlike the latter two, *B. hyodysenteriae* can be easily cultured and genetically manipulated, making it an ideal model for spirochete flagellar composition (Li, Corum et al. 2000, Li, Motaleb et al. 2000). Unlike in *Borrelia*, Sarkosyl-based flagellar purification methods do not cause the loss of the FlaA sheath protein (Ge, Li et al. 1998). Purified wild-type filaments have a diameter of ~25.5 nm, and the filaments appear to have a slight curvature (Li, Corum et al. 2000).

Mutants of each of the three *flaB* genes were developed (*flaB1*$^{-}$, *flaB2*$^{-}$, *flaB3*$^{-}$); purified flagellar filaments from each of these mutants had a similar diameter to the wild-type filaments, and each produced similar amounts of FlaA (Li, Corum et al. 2000). Each

flaB^{-} mutant was also motile, though with smaller swarm diameters (~14 nm, ~9 nm, ~6 nm, and ~12 nm, for wild-type, *flaB1*$^{-}$. *flaB2*$^{-}$, and *flaB3*$^{-}$ filaments respectively) and slower swimming speed (with the *flaB*$^{-}$ mutants reaching just 62-87% of the relative speed of the wild-type), suggesting a deficiency in motility (Li, Corum et al. 2000, Li, Wolgemuth et al. 2008). This motility deficiency led to a decrease in pathogenicity, as mice infected with *flaB1*$^{-}$ mutants showed no colonization of the bacteria (Kennedy, Rosey et al. 1997). The helical pitch and helical diameter of the *flaB*$^{-}$ mutants were slightly smaller than the wild-type filaments (which have a helical pitch of ~2.8 μm and a helical diameter of ~0.8 μm), though the differences were not statistically significant (Li, Corum et al. 2000); likewise, the flagellar diameters (~18 – 20 nm) were similar to those seen in the wild-type (~21 nm) (Li, Wolgemuth et al. 2008).

In a *flaA*$^{-}$ mutant, the purified flagellar filaments had a smaller diameter than wild-type filaments (~14 nm versus ~21 nm), and the helical pitch (~2.4 μm) and helical diameter (~0.6 μm) were also significantly smaller than in the purified wild-type flagella, suggesting that the FlaA sheath protein is important for maintaining the proper helical form of the flagellar filaments (Li, Corum et al. 2000, Li, Wolgemuth et al. 2008). However, this change does not impact the motility of the bacteria, as *flaA*$^{-}$ cells remain motile, though with a similar decrease in swarm diameter (~9 nm compared to the ~14 nm in wild-type) and relative speed (reaching just ~72% of the relative speed of the wild-type) as was seen in the *flaB*$^{-}$ mutants (Li, Wolgemuth et al. 2008). Unlike the *flaB1*$^{-}$ mutant, however, the *flaA*$^{-}$ mutant remains pathogenic, with colonization of infected mice occurring at four days post infection at levels similar to wild-type (Kennedy, Rosey et al. 1997).

Double mutants of both FlaA and FlaB (i.e., *flaA*$^{-}$*flaB1*$^{-}$) were developed; all three *flaA*$^{-}$*flaB*$^{-}$ mutants were motility deficient, with a velocity that decreased 50-60% from the wild-type, compared to the 13-38% decrease seen with the *flaA*$^{-}$ or *flaB*$^{-}$ single mutants (Li, Wolgemuth et al. 2008). While the *flaB1*$^{-}$*flaB3*$^{-}$ mutant also had a ~50% decrease in velocity, the *flaB1*$^{-}$*flaB2*$^{-}$ mutant was nonmotile, suggesting that FlaB1 and FlaB2 have functional overlap that is crucial for motility, while FlaB3 is less essential (Li, Wolgemuth et al. 2008). The *flaA*$^{-}$*/flaB*$^{-}$ mutants had a smaller helical pitch and helical diameter than the wild-type and single mutants, though the *flaB1*$^{-}$*flaB3*$^{-}$ mutant had helical values and a filament diameter similar to wild-type, supporting the importance of FlaA in forming the helical sheath (Li, Wolgemuth et al. 2008). The *flaB1*$^{-}$*flaB2*$^{-}$ filaments were short, straight, and nonhelical, though with a diameter similar to that of wild-type filaments, emphasizing the importance of the underlying core for proper helical sheath assembly (Li, Wolgemuth et al. 2008).

Leptospira spp.

Leptospira spp. have one flagellum on each end; these filaments extend only ~3 μm towards the center, and do not overlap (Holt 1978, Raddi, Morado et al. 2012, Picardeau 2017). The flagellum on each end can distort the ends of the cell body into a ‘hook’ or ‘spiral’ shape; the central region, without any flagella, remains undistorted (Figure 3) (Holt 1978, Li, Motaleb et al. 2000, Kan and Wolgemuth 2007, Wolgemuth 2015).

All *Leptospira* species have four isoforms of FlaB (FlaB1-4) and two isoforms of FlaA (FlaA1 and FlaA2) (Li, Motaleb et al. 2000, Picardeau 2017). Purified *Leptospira* wild-type flagellar filaments form tight coils, in contrast to the slightly curved filaments observed in other spirochetes (Berg, Bromley et al. 1978, Bromley and Charon 1979, Li, Corum et al. 2000). *Leptospira* also differs from other spirochetes in terms of their sheath composition: FlaA alone does not form a sheath around FlaB, as cells lacking both FlaA1 and FlaA2 have filament diameters similar to that seen in wild-type filaments (though the purified flagella are straight instead of coiled) (Lambert, Picardeau et al. 2012). Cells lacking only FlaA1 also have no loss in filament diameter, and these filaments remain coiled when purified (Lambert, Picardeau et al. 2012). The *flaA2*$^{-}$ mutants (lacking both FlaA isoforms) are non-motile and non-pathogenic, and are unable to infect hamsters, while the *flaA1*$^{-}$ mutants (which are only lacking FlaA1) remain both motile and pathogenic (Lambert, Picardeau et al. 2012). When *flaB* is deleted, no flagellar filaments are formed, and the bacteria are nonmotile (Picardeau, Brenot et al. 2001).

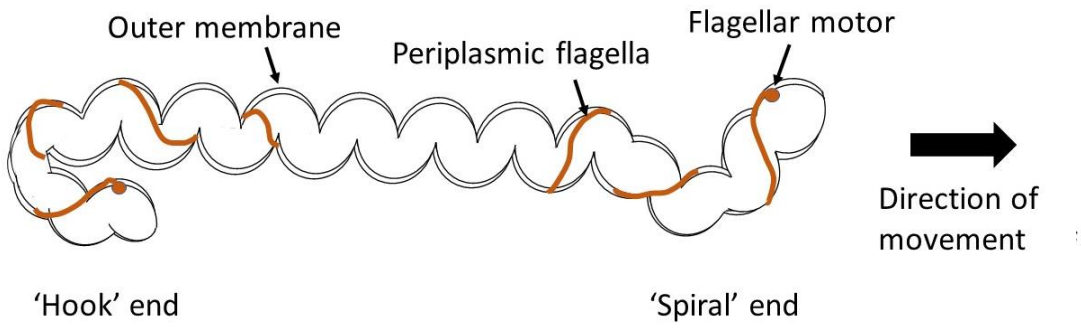

Figure 3. *Leptospira* have a spiral-shaped cell body and require a 'hook' and 'spiral' end for translational motility. *Leptospira* has one periplasmic flagellum on each end that do not overlap in the middle of the cell.

In addition to these conserved spirochete flagellar proteins, *Leptospira spp.* contain the flagellar coiling proteins FcpA and FcpB, which are not present in any other spirochete (Wunder, Figueira et al. 2016, Wunder, Slamti et al. 2018). In an *fcpA*$^{-}$ mutant, which lacks both FcpA and FcpB, the purified flagella are straighter, with a diameter of only ~16 nm (compared to the ~21 nm wild-type filaments) (Wunder, Figueira et al. 2016). This mutation renders the bacteria non-motile and non-pathogenic: all hamsters infected with an *fcpA*- strain of bacteria survive, with no colonization of the bacteria (Wunder, Figueira et al. 2016), making this mutation an intriguing vaccine candidate (Wunder, Adhikarla et al. 2021). When only FcpB is missing, the purified flagella remain coiled with a diameter similar to that of wild-type filaments (~20 nm), and these bacteria remain motile, though they cannot achieve the high velocity of the wild-type cells (Wunder, Slamti et al. 2018).

Spirochete motility

Spirochete flagellar motors operate in a similar manner to the well-studied *Salmonella* motors: spirochete motors also contain a rod (connected to the hook), rotor (MS- and C-ring), stator, and export apparatus, though a few important differences are present (Chen, Beeby et al. 2011, Carroll and Liu 2020). Most spirochete motors do not contain any L-ring proteins (FlgI), the exception being *Leptospira* (though in this species the L-ring is not located in the outer membrane as it is in non-spirochetes) (Murphy, Leadbetter et al. 2006, Chen, Beeby et al. 2011, Raddi, Morado et al. 2012). All spirochetes have a large ‘collar’- ~79 nm in diameter and composed of FlbB, FlcC, and FlcA in *Borrelia*- which likely helps to stabilize the stator (Murphy, Leadbetter et al.

2006, Chen, Beeby et al. 2011, Moon, Zhao et al. 2016, Xu, He et al. 2020, Chang, Carroll et al. 2021, Chang, Xu et al. 2021). This large collar, combined with a larger C-ring diameter (~57 nm in *Borrelia* and *Treponema primitia* versus the ~45 nm in *Salmonella*) potentially allows for spirochetes to generate greater amounts of torque (Murphy, Leadbetter et al. 2006, Liu, Lin et al. 2009, Chen, Beeby et al. 2011, Beeby, Ribardo et al. 2016, Carroll and Liu 2020, Nakamura 2020, Chang, Xu et al. 2021). Mutational studies in the *Borrelia* motor have revealed the sequential steps of flagellar motor formation, where the export apparatus, C- and MS-ring, stator, and collar are first to be assembled, followed by assembly of the rod, then the hook, and finally the flagellar filament (Zhao, Zhang et al. 2013).

Spirochetes move in a unique, 'corkscrew-like' fashion, allowing them to move through viscous environments that are not normally accessible to other bacteria (Li, Motaleb et al. 2000, Wolgemuth 2015). This motility is crucial to their pathogenicity, as motility-deficient mutants are unable to infect their hosts (Kennedy, Rosey et al. 1997, Lambert, Picardeau et al. 2012, Fontana, Lambert et al. 2016, Wunder, Figueira et al. 2016). This motility is due to the rotation of their periplasmic flagella, which in turn causes rotation and undulation of the cell body (Wolgemuth 2015). *Borrelia* has a 'flat-wave' motility, a result of opposing forces between the bending of the periplasmic flagella within the periplasmic space and the force that the bent flagella exert on the overlying protoplasmic cell cylinder (Charon, Cockburn et al. 2012, Wolgemuth 2015). When the flagella rotate, this flat wave undulates, producing thrust which propels the cell through the media (Wolgemuth 2015).

Leptospira can distort the ends of their cell body (corresponding to the locations of their two ~3 μm long periplasmic flagella) into 'hook' and 'spiral' shapes, an ability not found in other spirochetes (Holt 1978, Kan and Wolgemuth 2007, Wolgemuth 2015, Picardeau 2017). The flagellum at the 'hook' end rotates clockwise, while the flagellum at the 'spiral' end rotates counterclockwise; in order for translational motility to occur, one end must be in a 'hook' shape while the opposite end is in a 'spiral' shape, in which case the bacteria will move in the direction of the spiral end (Goldstein and Charon 1988, Li, Motaleb et al. 2000, Wolgemuth 2015). If the ends are either both 'hook' shaped or both 'spiral' shaped, motility will not occur (Goldstein and Charon 1988). The flagella can quickly change rotational directions, resulting in a switching of the 'hook' and 'spiral' ends that allows the cell to reverse its direction (Kan and Wolgemuth 2007). The underlying rotation and bending of the flagellum causes deformations in the stiffer cell cylinder, causing the cell cylinder to rotate in the opposite direction, producing thrust (Kan and Wolgemuth 2007). Some *Leptospira* flagellar mutations, such as the loss of FcpA or both FlaA proteins, affects the ability of the cell end to distort into the hook form, preventing the proper interplay between the flagellum and cell cylinder that is required for translational motility of the bacterium (Lambert, Picardeau et al. 2012, Wunder, Figueira et al. 2016). The link between flagellar structure, bacterial motility, and bacterial pathogenicity highlights the importance of understanding the spirochete flagellum, including what proteins comprise the filament, how it is assembled, and how filament rotation is achieved.

Structural studies of flagellar filaments

Cryo-electron microscopy (cryo-EM) is a method to study proteins or protein complexes at atomic-level detail. In this method, the sample of interest is placed on a grid which is plunge-frozen into liquid ethane, preserving the proteins in a native environment in vitreous ice without the need for crystallization. The grids are placed into a transmission electron microscope, in which electrons pass through the specimen and form 2D projections. These 2D projections, taken of the sample in various orientations, can be used to generate a 3D reconstruction of the complex of interest (Nogales and Scheres 2015).

High-resolution cryo-EM structures have been obtained for many bacterial flagella. Studies of *Salmonella* flagella that revealed that they are composed of an 11-fold symmetric pattern of the protein flagellin (FliC), which is 494 amino acids in length and contains two coiled-coil domains (D0 and D1), as well as two globular domains (D2 and D3); the coiled-coil domains surround a central pore (Yonekura, Maki-Yonekura et al. 2003, Maki-Yonekura, Yonekura et al. 2010). Structures were solved of straight *Salmonella* flagellar filaments with all 11 protofilaments in either the 'R' or 'L' forms, reaching ~4 Å resolution in both cases (Yonekura, Maki-Yonekura et al. 2003, Maki-Yonekura, Yonekura et al. 2010). Interactions between neighboring D0 and D1 coiled-coil domains were observed in both 'all-R' and 'all-L' structures, though the specific interacting residues within the assembly differ between the two forms (Yonekura, Maki-Yonekura et al. 2003, Maki-Yonekura, Yonekura et al. 2010).

A structure of a curved hook, which connects the motor to the flagellum, was recently obtained in *Salmonella*, reaching 3.6 Å in resolution, revealing the protein FlgE

helically arranged into a left-handed supercoiled filament, and highlighting important inter-subunit interactions that occurs between neighboring FlgE upon filament bending (Kato, Makino et al. 2019). In addition, a high-resolution structure of the *Salmonella* flagellar hook attached to the flagellar motor apparatus was recently obtained, highlighting interactions between the two complexes (Tan, Zhang et al. 2021).

A recent 3.5 Å resolution structure of a flagellum from *Campylobacter* reveals a mutationally straightened flagellum with a typical 11-fold symmetry, and highlights glycans that decorate the filament at 17 locations per flagellin monomer (Kreutzberger, Ewing et al. 2020). Likewise, an *in vivo* structure of the filament from *Kurthia* was recently resolved to 2.8 Å resolution, showing 11-fold symmetry and six sites of glycosylation on each flagellin (Blum, Filippidou et al. 2019); a structure of FliK flagellar filaments from *C. crescentus* resolved to 3.2 Å resolution also revealed 11-fold symmetry and glycosylation along the outer domains of the flagellin (Montemayor, Ploscariu et al. 2021). Near-atomic resolution was achieved in both 'R' and 'L' flagellar structures of *B. subtilis* (3.8 and 4.5 Å, respectively) and *Pseudomonas aeruginosa* (4.2 and 4.3 Å, respectively), highlighting the subtle structural differences between the two forms (Wang, Burrage et al. 2017).

Flagellar studies of *Leptospira*

While much is understood about the exoflagellar filaments from bacteria such as *Salmonella*, *B. subtilis*, and *Campylobacter*, many unknowns remain with regards to the structure and composition of spirochete flagella. The recent discoveries of the *Leptospira* flagellar factors FcpA and FcpB (Wunder, Figueira et al. 2016, Wunder, Slamti et al.

2018) and the *T. denticola* factor FlaG (Kurniyati, Liu et al. 2019) emphasizes that even the composition of these filaments may not be fully understood, and highlights the importance in studying such a complex and important structure.

While most studies have understandably been done on pathogenic *Leptospira*, the saprophytic *L. biflexa* is more easily transformable while conserving all of the known flagellar genes (Picardeau, Bulach et al. 2008, Evangelista and Coburn 2010, Fouts, Matthias et al. 2016, Wunder, Figueira et al. 2016, Wunder, Slamti et al. 2018). The use of *L. biflexa* in flagellar studies provides the opportunity to generate flagellar mutants, aiding in the understanding of the complex. The *fcpB*$^{-}$ mutant, for example, was first developed in *L. biflexa*, allowing for the determination of FcpB as a component of the flagellum (Wunder, Slamti et al. 2018). *L. biflexa* is easier to grow than the pathogenic *L. interrogans* (Picardeau, Brenot et al. 2001, Picardeau, Bulach et al. 2008), allowing for large-scale flagellar filament purification to be achieved more easily, while still providing crucial insight into flagellar composition.

Here I present the cryo-EM structures of three *Leptospira* flagellar mutants: *fcpB*$^{-}$ (Chapter 1), *fcpA*$^{-}$ (Chapter 2), and *flaA2*$^{-}$ (Chapter 3). I also show an improved single particle reconstruction of the tightly coiled wild-type filament (Chapter 2). I explore how these structural studies have provided new insights into *Leptospira* flagellar composition, flagellar assembly, filament supercoiling, and motility.

Chapter I: Loss of *Leptospira* flagellar coiling proteins affects curvature of purified flagellar filaments

Contribution statement: The work in this chapter is adapted from a recent publication (Gibson, Trajtenberg et al. 2020). (This eLife article is distributed under the terms of a Creative Commons Attribution License). Dr. Elsio Wunder purified the wild-type, *fcpB*$^{-}$, and *fcpA*$^{-}$ flagellar filaments used in this chapter. Kimberley Gibson froze grids with each of the samples, and collected tomograms of each. Kimberley Gibson also performed the reconstruction of the wild-type filament, using scripts adapted by Dr. Charles Sindelar. Dr. Zhiguo Shang also froze grids with each of the three samples, and collected micrographs for single particle analysis.

Introduction

Wild-type *Leptospira* flagellar filaments adopt tight coils when purified, in contrast to other spirochetes, whose purified flagella are curved but not coiled (Holt 1978). These flagella are approximately 2 – 4 μm in length, with a single filament on each of the cell that each extends only ~1/3 of the way up the cell body (Malmstrom, Beck et al. 2009, Fontana, Lambert et al. 2016, Wunder, Figueira et al. 2016, Picardeau 2017). *Leptospira* flagellar filaments contain a FlaB core (FlaB1-4), the sheath protein FlaA (FlaA1 and FlaA2), as well as additional coiling factors (FcpA and FcpB) (Li, Motaleb et al. 2000, Lambert, Picardeau et al. 2012, Wunder, Figueira et al. 2016, Wunder, Slamti et al. 2018).

Wild-type flagellar filaments have a diameter of ~20 – 26 nm, as measured in various studies with both *in situ* and purified filaments (Lambert, Picardeau et al. 2012, Wunder, Figueira et al. 2016, Sasaki, Kawamoto et al. 2018, Wunder, Slamti et al. 2018). Filaments from an *fcpB*$^{-}$ mutant (which only lacks FcpB) do not have a significant change in filament diameter, with an average diameter of ~19.9 nm in purified samples (Wunder, Slamti et al. 2018). Filaments from *fcpA*$^{-}$ mutants (which lack both FcpA and FcpB) are much thinner than wild-type ones, with purified filaments ~17.6 nm in diameter and *in situ* filaments ~15.7 nm in diameter; however, the length of the flagellar filaments remains unchanged (at ~2.14 μm and ~2.17 μm, respectively) (Wunder, Figueira et al. 2016).

For the *fcpA*$^{-}$ and *fcpB*$^{-}$ mutants, complementation restores the wild-type filament diameters and compositions (Wunder, Figueira et al. 2016, Wunder, Slamti et al. 2018). Interestingly, loss of the FlaA sheath proteins does not affect the filament diameter, as purified filaments from *flaA1*$^{-}$ mutants (lacking only FlaA1) are ~24.1 nm in diameter, and filaments from *flaA2*$^{-}$ mutants (which lack both FlaA1 and FlaA2) are ~25.2 nm in diameter (Lambert, Picardeau et al. 2012). Occasionally an abrupt loss of filament diameter is observed in negative stain images of purified flagella, likely reflecting a loss of the various sheath components that leaves only the ~12 nm inner FlaB core (Lambert, Picardeau et al. 2012, Wunder, Figueira et al. 2016).

Loss of both *Leptospira* FlaA1 and FlaA2 sheath proteins results in flagellar filaments that are straight instead of coiled. Purified filaments that have only lost FlaA1, however, remain tightly coiled (Lambert, Picardeau et al. 2012). A similar effect is observed upon loss of the *Leptospira*-specific coiling proteins: loss of both FcpA and

FcpB (*fcpA*$^{-}$ mutants) results in curved, but not coiled, filaments, whereas flagella that have only lost FcpB (*fcpB*$^{-}$ mutants) remain coiled (Wunder, Figueira et al. 2016, Wunder, Slamti et al. 2018). As both the FlaA and Fcp sheath factors appear to play a role in the supercoiling of the purified filaments, the essential sheath components required for the coiled wild-type geometry remains unknown.

The coiled nature of the purified flagellar filaments from wild-type *Leptospira* and the *fcpB*$^{-}$ and *flaA1*$^{-}$ mutants prevents the use of typical single particle cryo-EM techniques, as the planar, coiled filaments do not adopt the different orientations required for a high-resolution 3D reconstruction. Instead, cryo-electron tomography is used: the grids containing the purified filaments are tilted from (-60)° to (+60)°, and a 2D projection is obtained every 3°. However, the inability to tilt to 90° results in a 'missing wedge' of information, causing artefacts in the resultant structure; additionally, the resolution of the reconstruction is lower, as the dose per image applied to the sample is decreased (Turk and Baumeister 2020). Here I present the structure of a purified flagellar filament from an *fcpB*$^{-}$ mutant, determined through cryo-ET, which allowed for unambiguous docking of FcpB into the wild-type density. I also present quantifications of purified flagellar length and curvature, providing insight into how flagellar mutations affect the filament properties of the purified flagellum.

Results

FcpB is localized to the filament outer curvature

Kimberley Gibson had collected tomograms of flagella from wild-type, *fcpB*$^{-}$, and *fcpA*$^{-}$ samples, with the data sets containing 79, 26, and 25 tomograms respectively. I

began to process the *fcpB*$^{-}$ tomograms using the same pipeline that K. Gibson was simultaneously applying to the wild-type ones: using IMOD for tomogram reconstruction and filament selection (Mastronarde and Held 2017), RELION for initial sub-tomogram averaging (Scheres 2012), in-house scripts for filament smoothing and sub-volume coordinate interpolation (Huehn, Cao et al. 2018), and emClarity for higher-resolution 3D reconstruction (Himes and Zhang 2018) (see Methods). During this process, *fcpB*$^{-}$ filaments with ambiguous polarity were discarded, leaving 14 filaments from 10 tomograms, which corresponded to 1163 particles. The resultant *fcpB*$^{-}$ reconstruction was resolved to ~18 Å resolution, revealing a curved, asymmetric filament (Figure 4A).

The wild-type reconstruction performed by K. Gibson also revealed an asymmetric and curved flagellum; this reconstruction reached ~10 Å resolution (Gibson, Trajtenberg et al. 2020). A distorted central pore was present in the core of both the wild-type and *fcpB*$^{-}$ reconstructions, with the smeared density surrounding it likely corresponding to the FlaB core. Densities along the outer curvature appeared vastly

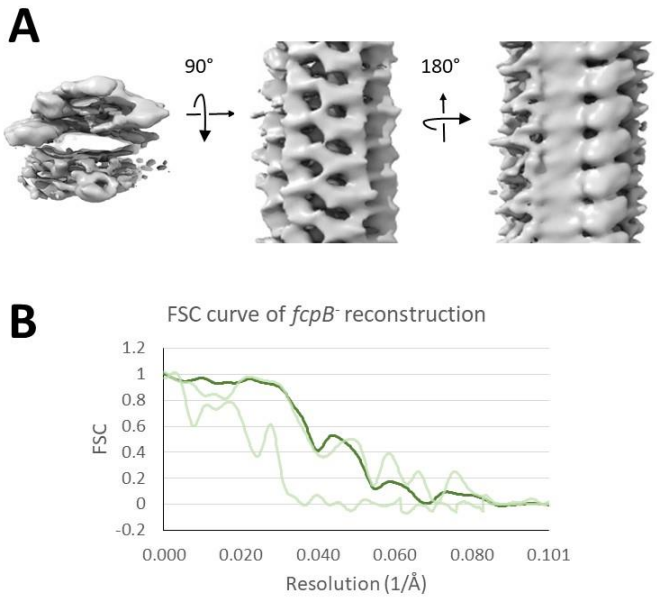

Figure 4. Reconstructed filament from an *fcpB*$^{-}$ mutant reveals a tightly coiled asymmetric structure. **A,** Density of the *fcpB*$^{-}$ structure; cross section is shown on the left, and the side views are shown in the center (looking at the outer curvature densities) and right (looking at the inner curvature densities). **B,** FSC curves of the reconstruction. Dark green is the average FSC (resolution of 18.41 Å at an FSC of 0.143), and the lighter green lines reflect the range of anisotropic resolutions.

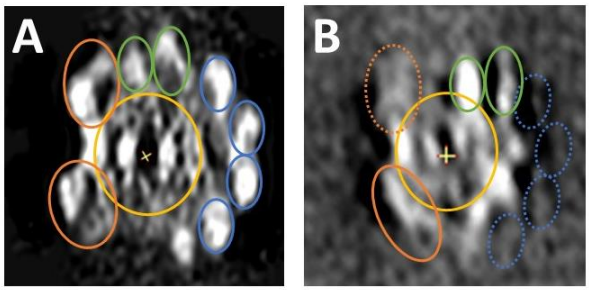

Figure 5. FcpB is localized to four lobes of the outer curvature of the wild-type filament. **A,** Cross sectional density of the wild-type filament. The core region is circled in yellow, and strong densities are circled in blue, green, and orange. **B,** Cross sectional density of the *fcpB*$^{-}$ mutant. While the core region, the two green densities, and one of the orange densities are present in this mutant, there is no density present for the four outermost lobes (blue, dashed lines) or one of the inner curvature densities (orange, dashed line).

different from the densities observed along the inner curvature, with no apparent symmetry along any direction. Comparison of the *fcpB*$^{-}$ reconstruction with the wild-type reconstruction revealed that four densities along the outer curvature of the wild-type filament were missing in the mutant filament, along with one of the densities along the wild-type inner curvature (Figure 5). FcpB was presumed to localize along the outer curvature of the flagellum, based on previous immunogold labeling studies (Wunder, Slamti et al. 2018), making the four outer curvature lobes the likely location of FcpB. Situs fittings of the FcpB crystal structure in the wild-type reconstruction, performed by K. Gibson, further supported this localization of FcpB, as the FcpB crystal structure fit into the outer curvature density, but not the missing inner curvature density. K. Gibson was able to dock the FcpA crystal structure (San Martin, Mechaly et al. 2017) into six densities beneath the FcpB lobes along the outer curvature, and 11-fold symmetrization of the region around the pore revealed the location of FlaB (Figure 6). The remaining inner curvature densities may correspond to the FlaA sheath proteins, which are known to

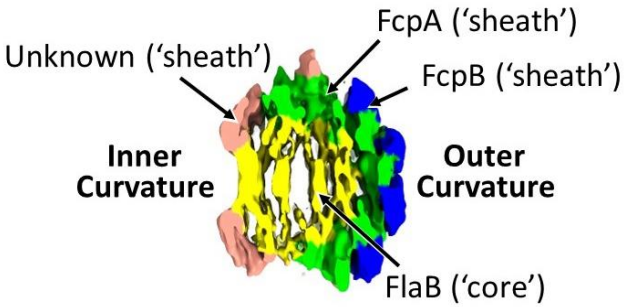

Figure 6. The flagellar proteins are asymmetrically distributed around the *Leptospira* flagellar filament. The FlaB core surrounds the central pore (yellow). FcpB is localized to the four outer curvature densities (blue), and FcpA is localized to the six densities underlying FcpB (green). Unidentified densities, presumably the FlaA sheath factors, are colored pink.

be present in the sample but have not been crystallized for density docking. The identity of these inner curvature densities is addressed in detail in Chapter 2.

Loss of Fcp proteins impacts filament curvature

Reconstructed tomograms provide a three-dimensional space in which the samples of interest are located. In the case of the purified flagellar filaments from wild-type, *fcpB*$^{-}$, and *fcpA*$^{-}$ samples, this results in the filaments, manually selected in the program IMOD (see Methods), having specific x, y, and z coordinates determined along their entire length. This allows for the determination of filament length: for wild-type flagella where the entire filament was located in the tomogram, the average length was 1.4 μm ± 1.2 μm. Of these 80 flagella, the majority (45) were < 1.0 μm, likely reflecting filaments that were broken during the purification process, whereas only seven filaments were > 3.0 μm, likely reflecting nearly intact flagella; the longest flagella observed from this sample was ~5.5 μm (Figure 7A). In contrast, of the 30 *fcpB*$^{-}$ mutant flagella that were located completely within the tomograms, 25 of the filaments were < 1.0μm, and no

filaments were > 2.0 μm (Figure 5B). Likewise, 37 of the 38 *fcpA*$^{-}$ flagella remaining entirely in the field of view were < 1.0 μm, with no filaments > 2.0 μm (Figure 7C).

Occasionally in the tomograms, an abrupt loss of filament diameter could be observed, likely reflecting a loss of the various sheath proteins that leaves only the ~12 nm diameter FlaB core. This change was also evident in the 2D single particle micrographs collected by Z. Shang (Figure 8). This abrupt decrease in diameter did not

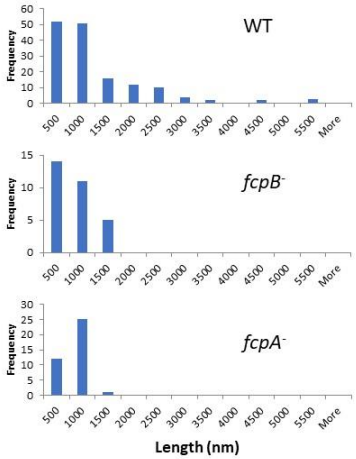

Figure 7. Flagella are disrupted during the purification process, resulting in shorter filaments. **A,** Distribution of wild-type flagellar lengths, in nm. A few wild-type filaments reached > 5 nm in length, likely representing nearly intact flagella. **B,** Filaments from the *fcpB*$^{-}$ mutants are shorter, with no flagella longer than 2 nm. **C,** Filaments from the *fcpA*$^{-}$ mutants are also broken and shorter than would be expected.

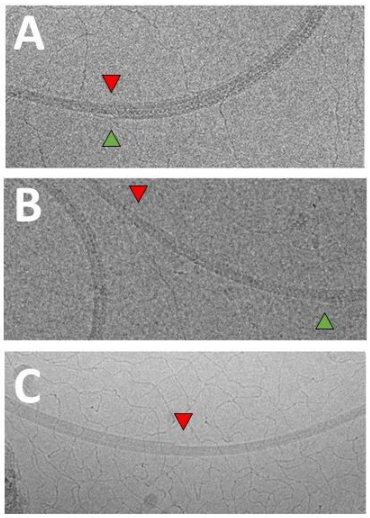

Figure 8. The flagellar sheath is occasionally shed from the filament. **A,** A 2D micrograph of a wild-type filament, showing an abrupt loss of sheath density along both the inner curvature (red arrow) and outer curvature (green arrow). **B,** A micrograph of a filament from an *fcpB*$^{-}$ mutant, highlighting a flagellum that loses both the inner and outer curvature sheath proteins at different locations on the filament. Loss of inner and outer curvature densities are represented by a red and green arrow, respectively**. C,** An *fcpA*$^{-}$ micrograph, showing a loss of density along the inner curvature (red arrow).

occur uniformly across the three different purified samples (wild-type, *fcpB*$^-$, and *fcpA*$^-$), with the *fcpA*$^-$ mutant filaments appearing to exhibit this loss more frequently than filaments from either the wild-type or *fcpB*$^-$ sample.

I examined the micrographs for evidence of the diameter change, as well as the location of the loss (i.e., whether it occurred on the outer curvature or inner curvature face of the purified filament). In 162 wild-type micrographs, a loss of diameter was observed 11 times (corresponding to 6.7% of the images)- four instances where the diameter was lost only along the outer curvature, and seven instances where the diameter was decreased in both the inner curvature and outer curvature sides of the flagellum. In the 139 *fcpB*$^-$ micrographs, the sheath abruptly decreased in diameter six times (representing 4.3% of images), four cases when the loss was only along the outer curvature, and two cases where the diameter decreased on both the inner and outer curvature. Of the 83 *fcpA*$^-$ micrographs, there were 20 instances of a decrease in filament diameter (corresponding to 24.1% of the images)- 16 cases where the decrease was observed along the outer curvature only, three cases where the decrease was observed

	Inner Curvature	**Outer Curvature**	**Both**	**Total # Images**
WT	0	4	7	162
***fcpB*$^-$**	0	4	2	139
***fcpA*$^-$**	16	3	1	83

Table 1. Loss of filament sheath density is seen more frequently in the *fcpA*$^-$ mutant than in the wild-type filaments. 2D micrographs of wild-type, *fcpB*$^-$, and *fcpA*$^-$ filaments were used. The number of instances where a filament lost density at either the inner curvature, the outer curvature, or both locations was recorded.

only along the inner curvature, and one case where a loss was observed along both sides of the filament (Table 1).

The determination of the x, y, and z coordinates for each of the filaments through the tomogram reconstruction (performed by myself for the *fcpB*$^{-}$ and *fcpA*$^{-}$ tomograms, and by K. Gibson for the wild-type tomograms) can also be used to quantify the flagellar curvature. While the loss of curvature of purified filaments can clearly be observed in the negative staining of *fcpA*$^{-}$ and *flaA2*$^{-}$ mutants (Lambert, Picardeau et al. 2012, Wunder, Figueira et al. 2016, Sasaki, Kawamoto et al. 2018), the exact curvature of these flagella had not been calculated. Similarly, while the *fcpB*$^{-}$ and *flaA1*$^{-}$ mutants appear to remain coiled (Lambert, Picardeau et al. 2012, Wunder, Slamti et al. 2018), there could be a subtle but significant difference from the observed wild-type curvature.

Therefore, I implemented a method to quantify the curvature of each flagellum, utilizing the determined filament coordinates (see Methods) (Crenshaw, Ciampaglio et al. 2000). Four consecutive points (P1, P2, P3, P4) are sufficient to characterize a helix, from which the curvature of the segment can be calculated (Figure 9A). Sliding the filament along to the next four points (P2, P3, P4, P5) gives the curvature at that location (Figure 9B); this can be repeated to determine how the curvature changes along the entire length of the filament, and a histogram can show the curvatures at each location sampled along each of the flagellum in the sample. Occasionally, filaments in these samples appeared to have an abrupt loss of filament diameter; as these filaments likely result in a sudden loss of sheath proteins, these flagellar segments were not included in the wild-type and *fcpB*$^{-}$ curvature calculations, and were separately included in the *fcpA*$^{-}$ curvatures when specified.

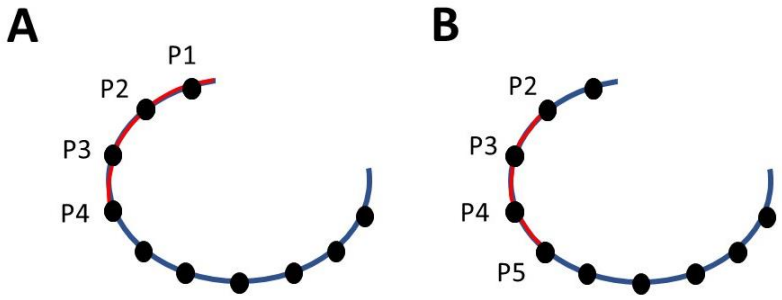

Figure 9. Four consecutive points are used to calculate the curvature of a filament segment. **A,** Curvature is calculated for the segment containing four points P1, P2, P3, and P4. **B,** Curvature is calculated for the curve defined by P2, P3, P4, and P5. This pattern is continued for the entire length of the filament.

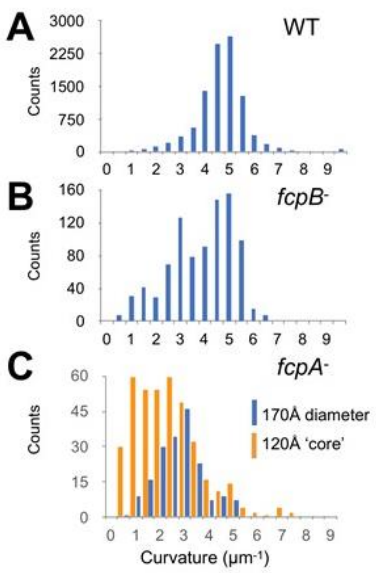

Figure 10. Loss of the outer curvature coiling proteins is correlated with a loss of filament curvature. **A,** Histogram of curvature values from wild-type filaments. Curvature is reported in μm^{-1}, and is the inverse of the radius of curvature. **B,** Histogram of *fcpB*$^-$ mutant filament curvatures reveals two distinct peaks, one at the wild-type value and one at the ~170 Å *fcpA*$^-$ peak. **C,** Histogram of curvatures from *fcpA*$^-$ mutant filaments. Values from the ~170 Å filaments are in blue, and the ~120 Å 'core' filaments are in orange. Loss of FcpA results in a decrease in filament curvature, and the core-only filaments (that have lost the additional inner curvature sheath proteins) have an even greater loss of curvature.

The curvatures from the wild-type sample reflect a tightly coiled filament, with a mode curvature of 5 μm^{-1} (which corresponds to a radius of curvature of 0.2 μm) in the near-Gaussian distribution (Figure 10A). The *fcpA*$^-$ sample also gave a Gaussian distribution, with a mode of 3 μm^{-1} (corresponding to a radius of curvature of 0.33 μm) (Figure 10C). However, the *fcpB*$^-$ sample appeared to be mixed, with peaks at both 5 μm^{-1} and 3 μm^{-1}, likely reflecting the ability of these filaments to adopt either a coiled, wild-type-like, or a straighter, *fcpA*$^-$-like form (Figure 10B). Many of the *fcpA*$^-$ mutant filaments appeared to experience the abrupt loss of the remaining sheath (reflected by a decrease in flagellar diameter to ~12 nm); curvatures for these filaments were not included above but were instead calculated separately. The curvature of these thinner (~120 Å) 'bare' filaments was smaller than the ~170 Å thick *fcpA*$^-$ filaments, with a peak from 1.0 – 2.5 μm^{-1} (Figure 10C). However, there was significant overlap between the 'thicker' and 'thinner' *fcpA*$^-$ curvatures, with many segments from the 'thicker' *fcpA*$^-$ mutant filaments having curvatures ≤ 2 μm^{-1}, and with many 'thinner' *fcpA*$^-$ mutant filaments having curvatures ≥ 3 μm^{-1}.

Discussion

Stability of purified *Leptospira* flagellar filaments

The length of purified filaments was significantly different between the wild-type and the *fcpB*$^-$ or *fcpA*$^-$ sample, with no filaments > 2.0 μm observed in either mutant sample. However, FcpB or FcpA are likely not crucial for full filament assembly, as *in situ* flagella from *fcpA*$^-$ bacteria are the same length as wild-type filaments (Wunder,

Figueira et al. 2016), suggesting that the difference in length observed in our samples is likely due to individual variations during the purification and freezing process.

It is possible that wild-type filaments containing all sheath proteins are more stable, and that loss of some sheath proteins makes the remaining sheath proteins more likely to 'fall off' the FlaB core. Indeed, the decrease in the amount of FlaA1 and FlaA2 as quantified through Western blots of flagella from *fcpA*$^{-}$ mutants (Wunder, Figueira et al. 2016, Sasaki, Kawamoto et al. 2018), suggests that the loss of the Fcp proteins may affect the presence of the remaining sheath, causing the abrupt changes in the diameters observed in the purified flagella. As these diameter decreases were observed in ~24% of imaged flagellar filaments of *fcpA*$^{-}$ mutants, though in < 10% of imaged filaments from wild-type and *fcpB*$^{-}$ mutant bacteria, FcpA likely plays a greater role in this stabilization than FcpB. In the wild-type structure, FcpA forms a sheath around the outer curvature of the filament, contacting the FlaA proteins that are solely located along the inner curvature (Figure 6) (Gibson, Trajtenberg et al. 2020); it is feasible, therefore, that the loss of FcpA could reduce the stability of the remaining inner curvature FlaA proteins, causing increased loss of this component during the flagellar purification process. FcpB, however, only contacts FcpA (Figure 6) (Gibson, Trajtenberg et al. 2020), and the loss of this protein therefore shouldn't significantly affect the stability of the FlaA region of the sheath. The diameter decreases observed in the filaments from wild-type and *fcpB*$^{-}$ mutant bacteria occur mainly along the outer curvature (suggesting a random loss of FcpB or FcpA) or along both sides of the filament. Few instances were seen of just an inner curvature loss in these coiled filaments, further supporting the theory that the loss of the FcpA proteins destabilizes the FlaA sheath. Our reconstruction from an *fcpB*$^{-}$

mutant is also missing one of the inner curvature densities (Figure 5), suggesting that a full outer sheath (containing both FcpA and FcpB) is important for full filament stability.

Relationship between filament curvature and motility

Leptospira flagella differ from other spirochete filaments, which are curved but are not tightly coiled (Holt 1978). In addition, *Leptospira* is unique amongst the spirochetes in terms of flagellar number and length, with only one periplasmic filament extending from each end of the cell that does not overlap in the center; the ability of these flagella to deform the cell body itself into the characteristic 'hook' and 'spiral' shapes is also unique to this species (Li, Motaleb et al. 2000, Kan and Wolgemuth 2007, Wolgemuth 2015).

The ability of the wild-type *Leptospira* flagellum to form a tight coil when purified likely plays an important role in motility, as loss of this coiling ability is correlated with a loss in virulence in the *fcpA*$^{-}$ and *flaA2*$^{-}$ mutants (Lambert, Picardeau et al. 2012, Wunder, Figueira et al. 2016). These straighter flagella are unable to distort the cell ends into the 'hook' shape and are thus unable to engage the cell cylinder in the torque-generating counter-rotation (Kan and Wolgemuth 2007, Wolgemuth 2015). It would be interesting to quantify the flagellar curvatures of wild-type and mutant filaments *in situ*, to determine the degree to which the native flagellar curvature is affected by the loss of these sheath proteins. It would also be interesting to quantify the curvature of purified flagellar filaments from other spirochetes; these other species do not have the coiling factors FcpA and FcpB (Wunder, Figueira et al. 2016, Wunder, Slamti et al. 2018), and have flagella that are curved (but not tightly coiled) when purified (Ge, Li

et al. 1998, Li, Corum et al. 2000, Motaleb, Sal et al. 2004, Li, Wolgemuth et al. 2008). It may be that these other filaments have curvatures similar to the *fcpA*$^{-}$ mutants, where both of these *Leptospira*-specific sheath factors are lost.

Conclusion

The asymmetric nature of the purified wild-type *Leptospira* flagellum is striking and unexpected. By reconstructing a flagellar filament from an *fcpB*$^{-}$ mutant and comparing it to the wild-type structure, we were able to localize the FcpB factor to the outer-most density along the outer curvature and dock FcpA into the layer beneath it, thus confirming that FcpA and FcpB are key novel, key components of the *Leptospira* flagellar sheath. I was able to quantify the curvature of purified flagellar filaments from wild-type as well as *fcpB*$^{-}$and *fcpA*$^{-}$ mutant *Leptospira*, confirming that the loss of both Fcp sheath factors results in significantly straighter flagella, while loss of only FcpB results in filaments with mixed curvatures, similar to both the wild-type and *fcpA*$^{-}$ values. Examination of micrographs revealed the importance of FcpA in filament stability, as filaments from *fcpA*$^{-}$ mutants experienced a marked increase in sheath-loss events along the filament inner curvature. The filaments from these *fcpA*$^{-}$ mutants, straighter than the coiled wild-type ones and containing the FlaA sheath a majority of the time, can therefore be used to determine the structure of the FlaA factors presumably located along the inner curvature; structural determination of these filaments is the main focus of Chapter 2.

Methods

Flagella purification

L. biflexa serovar Patoc strain Patoc I (Paris) wild-type, *fcpA*$^{-}$, and *fcpB*$^{-}$ strains (Wunder, Slamti et al. 2018) were grown in Ellinghausen-McCullough-Johnson-Harris (EMJH) media at 30°C, and were purified by Dr. Elsio Wunder as previously described (Trueba, Bolin et al. 1992, Wunder, Figueira et al. 2016). Briefly, 300 mL of cells were spun at 8000xg for 20 min at 4°C, and the pellet was washed in 28 mL of PBS. The mixture was re-centrifuged, and the pellet re-suspended in 30 mL of sucrose solution (0.5 M sucrose, 0.15 M Tris, pH 8.0); this was centrifuged at 8000xg for 15 min, the pellet resuspended in 8 mL of sucrose solution, and stirred on ice for 10 min. 0.8 mL of 10% Triton X-100 was added, and the mixture was stirred at room temperature for 30 min. 80 μL of Lysozyme (10 mg/mL) was added dropwise to the solution, which was stirred on ice for 5 min. 0.8 mL of 20 mM EDTA (pH 8.0) was slowly added to the solution, and the mixture was stirred at room temperature for 2 h. 160 μL of 0.1 M $MgSO_4$ was then added; after 5 min of stirring at room temperature, 160 μL of 0.1 M EDTA (pH 8.0) was added, and was further mixed for 5 min at room temperature. The sample was then centrifuged at 17000xg for 15 min; 2 mL PEG 8000 was added to the supernatant, and the mixture was kept on ice for 30 min before centrifugation at 27000xg for 30 min. The resulting pellet was resuspended in 3 mL water; this was spun at 80000xg for 45 min, and the final pellet was resuspended in 1 mL water.

Cryo-EM sample preparation

Samples were prepared for cryo-EM by K. Gibson, as described previously (Gibson, Trajtenberg et al. 2020). Purified filaments were mixed with fiducial markers; 6x concentrated Protein A conjugated with 10 nm colloidal Au (EMS Aurion) was used

for the wild-type sample, and 6x concentrated 10 nm colloidal Au BSA Gold Tracer beads (EMS Aurion) were used for the *fcpB*$^-$ and *fcpA*$^-$ samples.

For the wild-type filaments, 300 mesh Cu Quantifoil grids with 1.2 μm holes (Ted Pella, Inc, Redding, CA) were used; these grids were plasma discharged with H_2/O_2 for 30 sec in a Model 950, Solarus Advanced Plasma System (Gatan) before ~3 μL of the 2:1 (v/v) flagella: 6x gold fiducials was added to the grid. A Mark III Vitrobot (FEI Company, Eindhoven, the Netherlands) was used at 18°C and 100% humidity to plunge freeze the grids into liquid ethane, with a 2 min incubation time, a 6-7.5 s blot time, and a -2 mm blot offset.

The *fcpA*$^-$ and *fcpB*$^-$ samples were each plunge frozen by K. Gibson in a similar manner. ~3 μL of 2:1 (v/v) flagella: 6x gold fiducials were added to 200 mesh Cu C-flat grids with 1.2 μm spacing (EMS) with a layer of carbon 3 nm thick. These grids had been plasma discharged with a 75% argon/ 25% oxygen mixture for 6 sec in a Solarus Advanced Plasma System, Model 950 (Gatan). The same Vitrobot, incubation time, blotting time, and blotting discharge that was used with the wild-type sample was used with these mutant filaments, which were each plunge frozen separately.

Cryo-ET data collection

Tilt series from the wild-type grids were acquired on a 300 kV Polara cryo-TEM (UT Houston, Houston, TX) with a K2 direct electron detector (GATAN Inc, Pleasanton, CA) with SerialEM (University of Colorado, Boulder, CO). 120 tilt series were collected from -54° to +54° in 3° increments, with a dose rate of ~8 e^-/pixel/s and 12 frames

collected per tilt (0.1 s exposure per frame). A defocus of -2.0 to -4.0 μm was used; the magnification was 15500x, giving a 2.604 Å pixel size (at a binning of 1).

Tilt series from the *fcpB*$^-$ and *fcpA*$^-$ grids were collected on a 200 kV Tecnai F20 (CCMI, Yale University, New Haven, CT) with a K2 direct electron detector (GATAN) in counting mode. These tilt series were acquired through the program SerialEM (Mastronarde 2005) (Boulder, CO). Tilt series were collected from -60° to 60° at 3° increments, with 1.5 s of exposure per tilt angle, ~2.25 $e^-/Å^2$, and a dose fractionation of 7 frames (each which received 0.2 s of exposure), giving a final dose of ~90 $e^-/Å^2$; a defocus of -2.5 μm and a spot size of 7 were used. A magnification of 14500x was used, giving a final pixel size of 2.49 Å. 26 tilt series were collected for the *fcpB*$^-$ sample, and 25 were collected for the *fcpA*$^-$ sample.

Data processing, tilt series reconstruction, and subtomogram analysis

Data were processed as previously described (Gibson, Trajtenberg et al. 2020); a brief description of all processing steps is also included below.

Wild-type: Initial processing steps (motion correction, tilt series alignment, CTF estimation, and gold subtraction) were carried out with IMOD (Mastronarde and Held 2017). The resultant tilt series were binned by 2, giving a pixel size of 5.208 Å. Tomo3D (Agulleiro and Fernandez 2011) was then used for weighted back projection reconstructions of the tilt series. IMOD and PEET (Cope, Heumann et al. 2011) were used for filament selection with the proper repeat spacing of points (52 Å) and conversion for import into RELION v2.1 (Scheres 2012). Initial subtomogram alignment was performed with RELION (Bharat, Russo et al. 2015), utilizing CTFFind4 for CTF

correction and 3D autorefinement for unsupervised sub-tomogram averaging without helical symmetry. 62 tomograms underwent this RELION refinement, resulting in 10,851 particles. These RELION coordinates were then exported into emClarity.

CTF estimation and correction were performed in emClarity (Himes and Zhang 2018). Averaging and alignment were iterated in emClarity, with stricter search angles and translational shifts used for each iteration. Filaments were then smoothed using an in-house script (Huehn, Cao et al. 2018) to remove any outliers and ensure a consistent 52 Å repeat spacing; after several rounds of this smoothing protocol, the coordinates were re-imported into emClarity. Additional averaging and alignment cycles were performed: cycles 0-3 at a binning of 2, and cycles 4-8 at a binning of 1. The final reconstruction gave an average resolution of 9.83 Å, with the anisotropic resolutions ranging from 8.89 – 16.17 Å.

fcpB^{-}: Initial processing and subtomogram averaging was performed as above, utilizing IMOD (Mastronarde and Held 2017), Tomo3D (Agulleiro and Fernandez 2011), PEET (Cope, Heumann et al. 2011) and RELION v2.1 (Scheres 2012, Bharat, Russo et al. 2015). The particles were binned by 2, giving a pixel size of 4.97 Å. A total of 24 tomograms were used in the RELION refinement, yielding 5917 particles. Further refinement with emClarity (Himes and Zhang 2018), including the smoothing of particles with in-house scripts (Huehn, Cao et al. 2018), was performed as above. After particles with ambiguous polarity were discarded during the smoothing steps, 1163 particles (from 10 tomograms) remained. The final emClarity refinement used five cycles of averaging and alignment, with a binning of 2 used for each. The final resolution was 18.41 Å, with an anisotropic range of 14.58 – 31.67 Å.

fcpA$^-$: Initial processing and subtomogram averaging of these particles was also performed as described above, utilizing IMOD (Mastronarde and Held 2017), Tomo3D (Agulleiro and Fernandez 2011), PEET (Cope, Heumann et al. 2011), and RELION v2.1 (Scheres 2012, Bharat, Russo et al. 2015). The 25 tomograms yielded 4327 particles, though due to issues with sample heterogeneity and signal quality, no reconstruction was obtained. 2D classification was performed in RELION, revealing classes with two distinct diameters. A class with a diameter of ~120 Å contained 2958 particles, and a class with a diameter of ~170 Å contained 323 particles. The coordinates for each of these classes was used separately in determining the filament curvatures.

Curvature analysis

The smoothed 3D coordinates obtained from the wild-type, *fcpB*$^-$, and *fcpA*$^-$ reconstructions were used for curvature quantification, utilizing the method of Crenshaw et al. (Crenshaw, Ciampaglio et al. 2000). A more detailed description of the method has been previously described (Gibson, Trajtenberg et al. 2020). Briefly, consider the consecutive points P_1, P_2, P_3, …, P_N. Each of these points is spaced 52 Å apart, representing the repeat spacing of the filament. To minimize noise, points with a spacing of 10 are used for the calculations. Let $C = P_i - P_{(i-10)}$ and $D = P_{(i+10)} - P_i$, from i = 11 to i = (N-10). Curvature (κ) at point i can then be determined by:

$$\kappa_i^* = \left(\frac{C \cdot D}{|C||D|}\right)\left(\frac{2}{|C| + |D|}\right)$$

To determine the final curvature value, the curvature at two consecutive points is averaged, as shown below:

$$\kappa^*_{<i>} = \frac{\kappa^*_i + \kappa^*_{i+1}}{2}$$

The resultant curvature values are then multiplied by 1/(pixel size * binning), to give the curvature in units of 1/Å, for the points 11 to N-10 along the filament. Curvatures for 84 wild-type filaments, 14 *fcpB*$^-$ mutant filaments, and 8 *fcpA*$^-$ mutant filaments were used for the curvature histograms included above; short filaments that contained fewer than 20 points were not included in the analysis.

Chapter II: Structural basis of flagellar filament asymmetry and supercoil templating by the *Leptospira* spirochete sheath factor FlaA2

Contribution statement: Dr. Garrett Debs helped with some of the computational techniques used in the reconstruction and analysis. Fabiana San Martin (in the lab of Dr. Alejandro Buchiazzo) performed mass spectrometry on the purified flagella, with the help of Azalia Rodriguez and Rosario Duran. Dr. Elsio Wunder purified the wild-type filaments included in this study, and Kimberley Gibson prepared and imaged the wild-type grids. Dr. Charles Sindelar performed the cryoSPARC reconstruction of the wild-type flagellar filaments and aided in the analysis of the structures. This work is included in a bioRxiv preprint (doi: https://doi.org/10.1101/2022.03.03.482903).

Introduction

Salmonella flagellin can adopt two different forms: the typical 'L' form, and an 'R' form that is ~1.5% shorter and tilted by ~5.5° when compared to the 'L' form (Yamashita, Hasegawa et al. 1998, Calladine, Luisi et al. 2013). Exoflagella, such as those from *Salmonella*, are predicted to adopt 12 different flagellar supercoil forms, depending on the number of 'L' and 'R' protofilaments within the flagellar filament; the 'R' protofilaments are expected to cluster along the inner curvature side of the filament

(Leifson 1960, Calladine 1975, Kamiya and Asakura 1976, Calladine, Luisi et al. 2013). The 'normal' filament contains nine 'L' and two 'R' protofilaments (Leifson 1960, Fujii, Shibata et al. 2008, Calladine, Luisi et al. 2013), though challenges associated with solving supercoiled structures have resulted in the structural determination of only mutationally straightened 'all-L' and 'all-R' flagellar forms (Yonekura, Maki-Yonekura et al. 2003, Maki-Yonekura, Yonekura et al. 2010, Wang, Burrage et al. 2017, Kreutzberger, Ewing et al. 2020).

In contrast to the curved 'normal' *Salmonella*, the wild-type *Leptospira* flagellum is tightly coiled and has an asymmetric composition of its eight flagellar proteins within its structure. FlaB (of which there are four isoforms) comprises the filament core; the flagellar coiling proteins FcpA and FcpB form the sheath and are localized to the outer curvature face of the coiled filament, while FlaA1 and FlaA2 are also components of the sheath but are presumed to localize to the inner filament curvature (Lambert, Picardeau et al. 2012, Wunder, Figueira et al. 2016, Wunder, Slamti et al. 2018, Gibson, Trajtenberg et al. 2020).

The saprophyte *L. biflexa* and the pathogen *L. interrogans* share ~61% of genes, including all known flagellar genes (Picardeau, Bulach et al. 2008). Each of the flagellar proteins has a relatively high degree of sequence identity between *L. biflexa* and *L. interrogans*, ranging from 53% in FcpB to 92% in FlaB4 (Gibson, Trajtenberg et al. 2020). As *L. biflexa* is easier to grow and genetically manipulate (Picardeau, Brenot et al. 2001, Picardeau, Bulach et al. 2008), and as the flagellar structure is likely to be similar between the different species, several recent studies have used the saprophyte to study

Leptospira flagella (Picardeau, Brenot et al. 2001, Sasaki, Kawamoto et al. 2018, Wunder, Slamti et al. 2018, Gibson, Trajtenberg et al. 2020).

Within each species, the four FlaB isoforms of *Leptospira* have a high degree of sequence identity to each other, with the identity in *L. interrogans* ranging from 57 - 72% (Norris, Charon et al. 1988, Li, Motaleb et al. 2000, Holzapfel, Bonhomme et al. 2020). However, the isoforms are predicted to be present in vastly different quantities, ranging from ~12,000 copies per cell of FlaB1 to ~300 copies per cell of FlaB3 in *L. interrogans* (Malmstrom, Beck et al. 2009). FlaB is homologous to the coiled-coil domain of flagellin (Norris, Charon et al. 1988), and is assembled in the 11-fold symmetry present in all known bacterial flagella (Beatson, Minamino et al. 2006, Kreutzberger, Ewing et al. 2020). Many bacterial flagella are glycosylated (Logan 2006, Blum, Filippidou et al. 2019, Chu, Liu et al. 2020, Kreutzberger, Ewing et al. 2020, Montemayor, Ploscariu et al. 2021), including the FlaB core of the spirochete *T. denticola*, which is modified with a novel glycan (Kurniyati, Kelly et al. 2017). Due to the similarity of the FlaB sequences of *T. denticola* and *Leptospira*, it is predicted that the FlaB core of *Leptospira* is glycosylated as well (Holzapfel, Bonhomme et al. 2020).

Filaments lacking FcpA (*fcpA*$^{-}$ mutants) are motility deficient and non-virulent, and their purified flagella are not tightly coiled (Wunder, Figueira et al. 2016). The flagellum from *L. interrogans fcpA*$^{-}$ mutants contains FlaB, FlaA1, and FlaA2, though the levels of FlaA1 and FlaA2 are decreased by ~26% and ~57% respectively (Wunder, Figueira et al. 2016), likely due to the interaction of FcpA with these additional sheath proteins and the subsequent destabilization of FlaA upon the loss of FcpA (Gibson, Trajtenberg et al. 2020). However, as neither the FlaA1 nor the FlaA2 structures have

been solved, the specific locations of these FlaA sheath factors within the inner curvature of the wild-type flagellum could not be determined (Gibson, Trajtenberg et al. 2020). Though the FlaA sheath factor is present in all spirochetes (Charon, Greenberg et al. 1992, Li, Motaleb et al. 2000) and is crucial for motility and virulence in *Leptospira* (Lambert, Picardeau et al. 2012), little is known about its structure and function.

Here I present a near-atomic resolution reconstruction of a flagellum from an *L. biflexa fcpA*$^{-}$ mutant, revealing the structure of FlaA2 and a novel, previously uncharacterized *Leptospira* flagellar sheath protein, both localized solely to the filament's inner curvature. The structure also reveals high-resolution details of the asymmetric FlaB core, showing glycosylation along the core and revealing a 'seam' supporting a 10L/1R supercoiled filament.

Results

Near-atomic structure of an *fcpA*$^{-}$ mutant flagellum decorated with a FlaA2 sheath

While purified wild-type flagella form tight planar coils, purified flagellar filaments from *fcpA*$^{-}$ mutants are significantly straighter (Wunder, Figueira et al. 2016), and are therefore amenable to single particle cryo-EM. The filaments are curved and comprise of two populations in the sample: thinner, straighter filaments (~12 nm) that we hypothesize correspond to just the FlaB core, and curved, slightly thicker filaments (~16 nm) which contain an additional density along the inner curvature face of these filaments (Figure 11A). To account for the apparent heterogeneity and asymmetry of the filaments

in the sample, focused classification was used (see Methods). The resulting reconstruction, resolved to 3.4 – 4.4 Å, reveals a highly asymmetric filament containing a

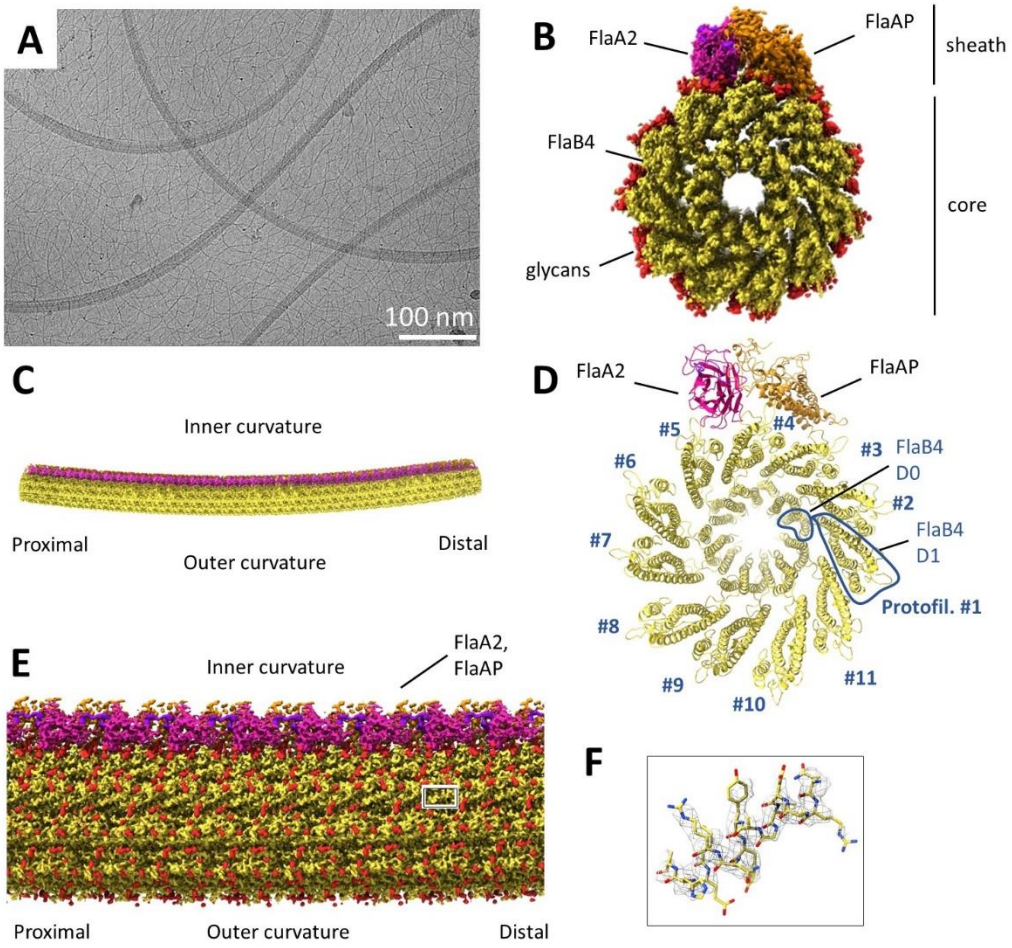

Figure 11. FlaA2 co-localizes to the filament inner curvature in a single row together with FlaAP, a previously uncharacterized sheath protein. **A,** Representative micrograph, showing the heterogeneity and slight curvature of the filaments from the *L. biflexa fcpA*$^{-}$ mutant. Scale bar is 100 nm. **B,** A 3D isosurface rendering of a flagellar filament decorated with FlaA2 and FlaAP. Density for the FlaB core is colored yellow, density identified as bound FlaA2 molecules is colored pink, and density corresponding to FlaAP is colored orange. Putative glycosylation site densities are colored red. **C,** A lower magnification view of an extended filament generated from the reconstruction in A, revealing supercoiling. Estimated supercoil pitch and diameter values are 2.01 μm and 0.43 μm respectively. **D,** Models of the FlaB4 core and FlaA2 and FlaAP sheath proteins, using the same coloring as in B. The initial models were generated with AlphaFold2, and were fit into the density using Isolde. **E,** Side view of the reconstructed *fcpA*$^{-}$ mutant filament, as in B. **F,** A zoomed-in view of the box in E, showing a segment of the helical density and the modeled FlaB4.

core region surrounding a central pore, as well as two distinct sheath densities each localized to a row along the inner curvature (Figure 11B).

Two rows of localized sheath density lie adjacent to each other overlying the 'core' (locally resolved to 3.4 Å; Figure 12B): one sheath row has a strong beta-sheet character, while the other sheath row is composed of numerous helices. AlphaFold2 models (Jumper, Evans et al. 2021) of FlaA1 and FlaA2, the known flagellar sheath proteins, predict that both proteins form a jellyroll-fold. While the FlaA1 model (which is 87 residues larger than FlaA2) did not fit into the beta-sheet sheath density, the FlaA2 model fit well into the density, which in turn allowed us to tentatively dock FlaA2 into the sheath region. Many of the AlphaFold-predicted side chains of the jellyroll fold fit nearly perfectly into the density, with only minor adjustments required (Figure 13A): the N-terminal 38 residues of FlaA2 could not be resolved in the density, and the C-terminal 15 residues were moved to fit into unoccupied density bridging adjacent monomers (as described below). This FlaA2 sheath protein overlays two FlaB protofilaments at the inner curvature face of the filament: protofilaments #4 and #5.

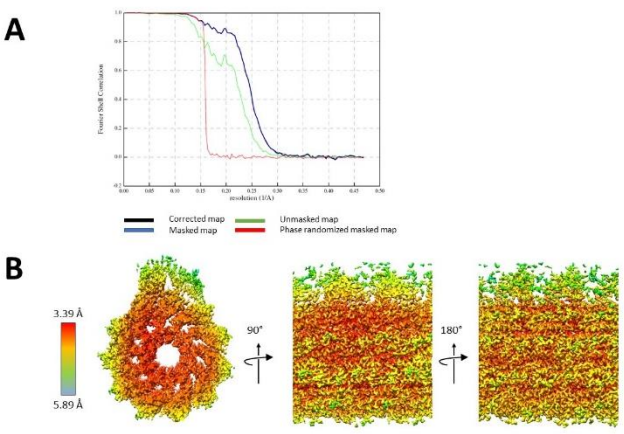

Figure 12. The core and sheath regions of the *fcpA*⁻ mutant flagellar filament both reached near-atomic resolution. **A,** FSC curve of the reconstruction, giving a resolution of 3.8 Å at an FSC of 0.143. **B,** Local resolution of the filament, showing the core was resolved to a higher resolution than either of the two sheath densities.

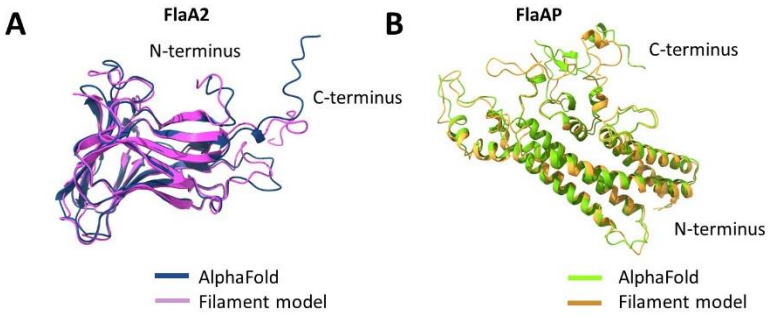

Figure 13. The AlphaFold2 structural predictions of FlaA2 and FlaAP show strong similarities to the models built into the density. **A,** Comparison of the FlaA2 structure, with the AlphaFold2 prediction in blue and the filament model in pink. Only the C-terminal region differed between the two. **B,** Comparison of the FlaAP AlphaFold2 (green) and filament model (orange). While the helices were consistent, a few loops differed between the structures. A disordered loop region (264-284) is not shown in either structure.

Locus	#Spectra (Replicate 1)	#Spectra (Replicate 2)	#Spectra (Replicate 3)	Σ #Spectra	Replicate Count	Protein name
B0SQZ5_LEPBP	675	190	246	1111	3	FlaB4
B0SSZ5_LEPBP	535	203	269	1007	3	FlaB1
B0SJC6_LEPBP	322	138	163	623	3	FlaAP
B0SSZ4_LEPBP	322	93	128	543	3	FlaB2
B0SKT5_LEPBP	182	92	99	373	3	FlaA2
B0SKT4_LEPBP	210	67	46	323	3	FlaA1
B0SS86_LEPBP	93	19	28	140	3	FlaB3

Table 2. Mass spectrometry of purified flagellar filaments from *L. biflexa fcpA*$^{-}$ mutants reveals the components of the complex. All of the expected FlaB core and FlaA sheath are present; the novel sheath protein (FlaAP) was the fourth most abundant protein in the sample. Non-flagellar genes were excluded for clarity. The analysis was performed in triplicate.

Identification of a novel *Leptospira* flagellar sheath protein

To identify the 'helical' sheath protein, which did not fit the beta-sheet models of the known flagellar sheath proteins FlaA1 and FlaA2, purified flagella from the *fcpA*$^{-}$ mutant were sent to Fabiana San Martin for mass spectrometry analysis. This analysis revealed several previously uncharacterized proteins that were in large quantities across various flagella preparations (Table 2). AlphaFold2 (Jumper, Evans et al. 2021) was used to obtain models of these potential candidate proteins: one protein, corresponding to the product of LEPBI_I0551, was predicted to form a helical bundle similar to the observed sheath density. Indeed, this novel protein, which we have named FlaAP (FlaA-associated Protein) fit nearly perfectly into the remaining sheath density (Figure 14), allowing us to identify it as a component of the *Leptospira* flagellar sheath. As with FcpA and FcpB, FlaAP is present across all *Leptospira* species, but is not present in other spirochetes (Wunder, Figueira et al. 2016, Wunder, Slamti et al. 2018). *flaAP* is not located near any other flagellar genes in the genome.

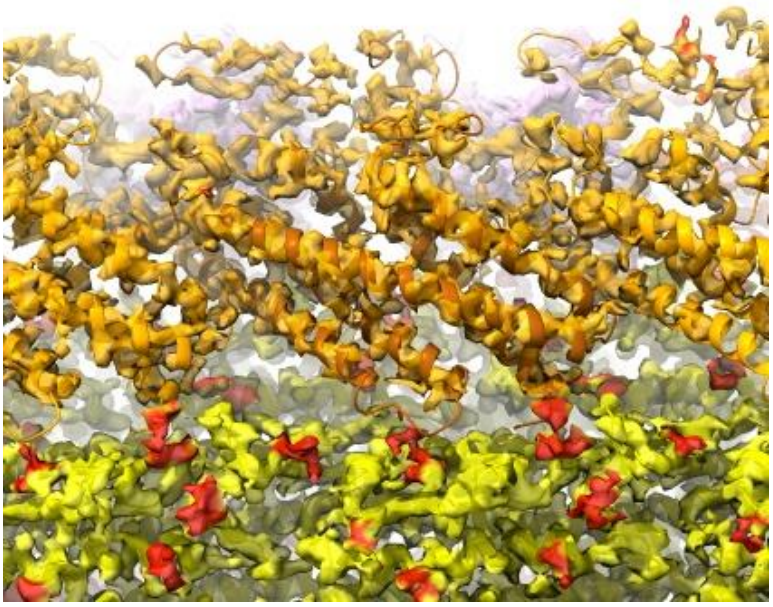

Figure 14. The helical sheath density is identified as FlaAP. The FlaAP structure (orange) fits into the helical sheath density. The FlaB core is in yellow, and the glycans are shown in red. The FlaA2 density (pink) is located behind FlaAP.

FlaAP consists of five helices connected by long loop regions; as with FlaA2, FlaAP molecules interact laterally to form a long row (see below) overlying the FlaB core at protofilaments #3 and #4 (Figure 11D). Crosslink mass spectrometry experiments on purified wild-type flagellar filaments (performed by F. San Martin) consistently identify an interaction between the FlaA2 and FlaAP sheath proteins, a contact supported by our built models that show the two peptides in close proximity (Figure 15).

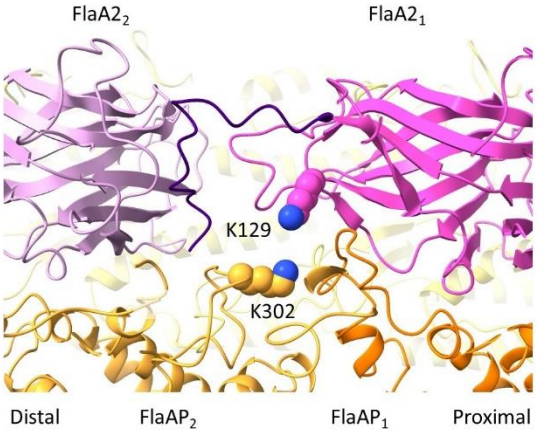

Figure 15. Crosslink identified in mass spectrometry is preserved in our sample. A crosslink between Lys129 (FlaA2) and Lys302 (FlaAP) was consistently found in mass spectrometry experiments (data not shown); our built models show these residues within 10 Å of each other.

FlaA2 and FlaAP interact with the glycosylated FlaB4 core

The core region of the filament consists of coiled-coils assembled with 11-fold symmetry, similar to the core regions of other bacterial flagella (Yonekura, Maki-Yonekura et al. 2003, Maki-Yonekura, Yonekura et al. 2010, Blum, Filippidou et al. 2019, Kreutzberger, Ewing et al. 2020, Montemayor, Ploscariu et al. 2021). Each of the four potential FlaB isoforms was modeled (with AlphaFold2) into this density (Figure 16) (Jumper, Evans et al. 2021). Additional oblong, elongated densities were observed at seven locations in each FlaB monomer (Figure 11E); this density is consistent with the glycosylation observed in other cryo-EM flagellar structures (Blum, Filippidou et al.

2019, Kreutzberger, Ewing et al. 2020, Montemayor, Ploscariu et al. 2021). Overlaying the modeled FlaB4 isoform with these presumed glycans resulted in all seven elongated densities overlaying Ser and Thr residues, suggesting that these locations correspond to sites of O-linked glycosylation. Other FlaB isoforms had Ser and Thr at only 4-5 of these locations (Figure 16A-C), suggesting that FlaB4 (which is the only isoform to contain a Ser/Thr at the glycan-containing residues of Thr137 and Thr182) is likely the dominant isoform in the structure. The same glycosylation pattern was seen in each of the 11 protofilaments, with no glycosylation difference between the protofilaments adjacent to the sheath region and the protofilaments on the outer curvature side of the filament where no sheath proteins are present (Figure 11B).

While this glycosylation fingerprint is suggestive of a filament core in the *L. biflexa fcpA*$^{-}$ mutant that is predominantly composed of FlaB4, previous studies (in *L. interrogans*) report a FlaB1-dominant flagellum that has ~3.5x more FlaB1 than FlaB4 (Malmstrom, Beck et al. 2009). Mass spectrometry on our purified *L. biflexa* flagellar

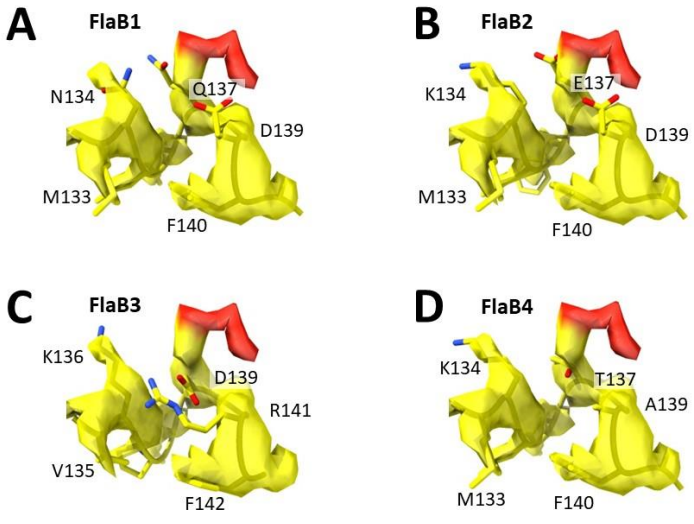

Figure 16. Glycosylation pattern identifies FlaB4 as the predominant isoform. **A,** The AlphaFold2 prediction for FlaB1, fit into the density. Visible residues are labeled. **B,** The FlaB2 prediction, as in A. **C,** The FlaB3 prediction, as in A. **D,** The FlaB4 prediction, as in A. This is the only isoform where a residue amenable to glycosylation underlies the glycan density (Thr137).

samples (performed by F. San Martin) supports our finding, with similar levels of FlaB4 and FlaB1 observed across all three replicates (Table 2).

Several residues on both FlaA2 and FlaAP appeared to be interacting with these glycans. In the case of FlaA2, pairs of positively charged residues interact with four glycans (Figure 17A), including an especially prominent interaction between Arg117 and Lys151 (FlaA2) and the glycan at Ser126 (FlaB4). Though FlaA2 is adjacent to two FlaB4 protofilaments (#4 and #5), all the glycans interacting with these FlaA2 residues are located on protofilament #5. FlaAP also appears to interact with four glycans through a single positively charged residue (instead of the pair of residues seen in the FlaA2-glycan interactions). These FlaAP glycan contacts span both FlaB4 protofilaments underlying FlaAP (protofilaments #3 and #4) (Figure 17C).

There are also protein-protein contacts observed between the FlaB4 core and both FlaA2 and FlaAP. Surprisingly, these contacts are sparse in nature, with only three sites of contact in FlaA2 (which span both protofilament #4 and #5) and three sites in FlaAP (spanning protofilaments #3 and #4) (Figure 17B,D). The majority of these contacts involve hydrophobic interactions, with few charged contacts between the core and sheath residues.

Several of these sheath-core interactions, for both the FlaA2 and FlaAP sheath proteins, are FlaB4-specific, supporting our identification of FlaB4 as a major core

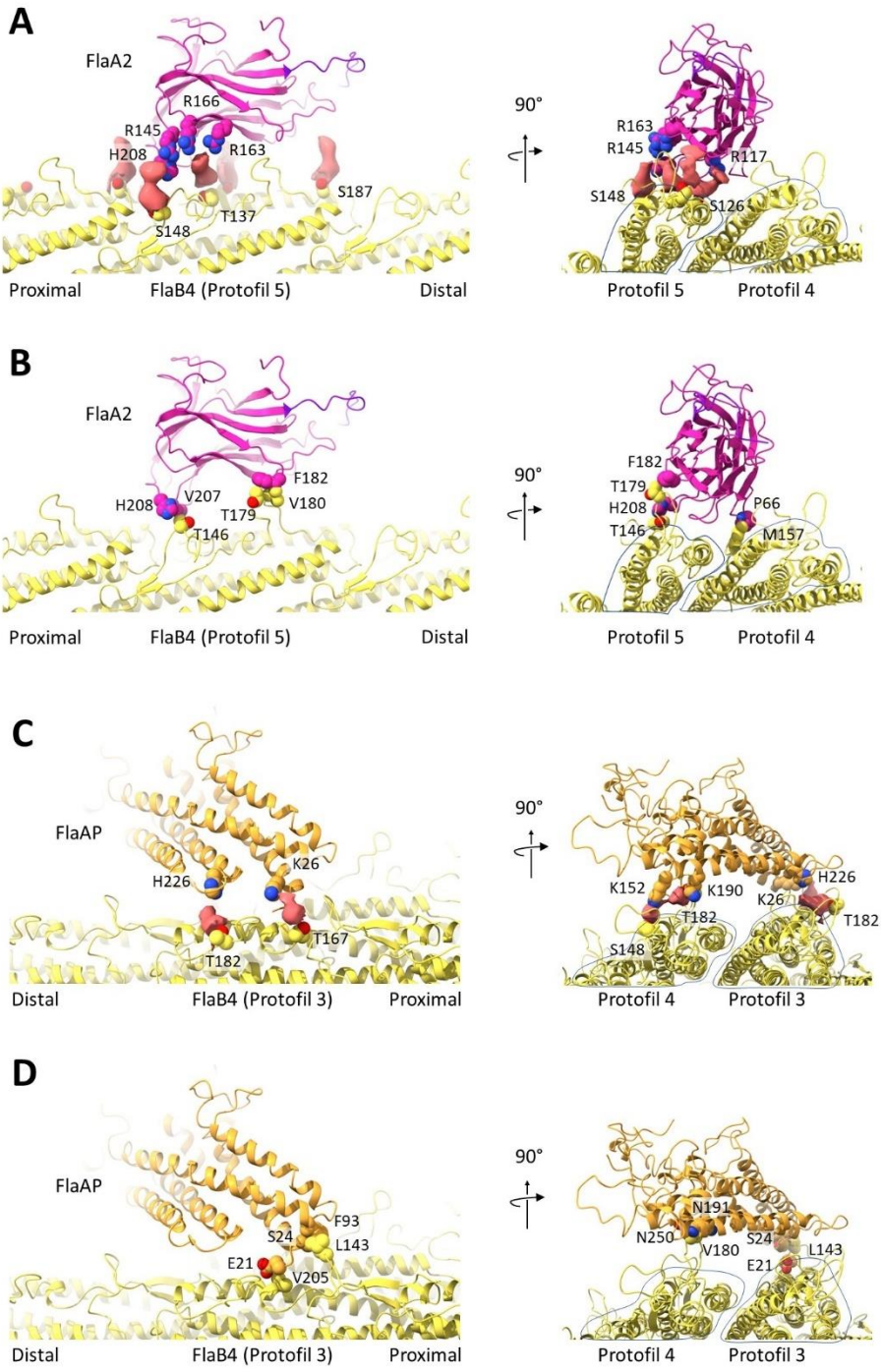

Figure 17. Core interactions of FlaA2 and FlaAP are primarily mediated by glycosylated FlaB4 side chains. **A,** Fitted atomic model of FlaA2 reveals probable sugar binding sites. Left: The glycan of Ser148 (FlaB4) interacts with Arg145 and His208 (FlaA2), and the glycan on Thr137 (FlaB4) interacts with Arg163 and Arg166 (FlaA2). Right: End-on view of the filament shows that the glycans are all located along one protofilament (protofilament #5). **B,** Protein-protein interactions between FlaA2 and the FlaB4 core. Left: His208 and Val207 (FlaA2) interact with Ser146 (FlaB4), and Phe182 (FlaA2) interacts with Thr179 and Val180 (FlaB4). End-on view, showing that these protein contacts bridge between two adjacent FlaB4 protofilaments (protofilament #4 and protofilament #5). **C,** Interactions between the FlaAP sheath protein and the glycans associated with the FlaB4 core. Left: The interaction between the presumed glycan on Thr182 (FlaB4) and His226 (FlaAP), as well as interactions between the presumed glycan of Thr167 (FlaB4) and Lys126 (FlaAP). Right: End-on view of the filament shows that these glycan interactions bridge between FlaB protofilaments, with two glycan interactions along both protofilaments #3 and #4. **D,** Protein-protein contacts between FlaAP and the FlaB4 core. Left: Ala172 (FlaB4) interacts with Glu21 (FlaAP), and Leu143 (FlaB4) interacts with Phe93 and Ser24 (FlaAP). Right: End-on view shows that the protein-protein contacts also help to bridge FlaAP between protofilaments #3 and #4.

component. The FlaA2 residues Arg163 and Arg166 interact with the glycan at Thr137, a residue which is not amenable to O-linked glycosylation in the other FlaB isoforms. FlaAP interacts with the glycan of Thr182 on both protofilaments #3 and #4; this is also a FlaB4-specific glycosylation site.

Val180 is located at the tip of a divergent loop (175-184) that is 1-2 residues shorter in the other FlaB isoforms; this residue is involved in both a hydrophobic contact with Phe182 of FlaA2, as well as contacting Asn191 and Asn250 in FlaAP. The interaction between His208 (FlaA2) and Thr146 (FlaB4) also would preferentially interact with FlaB4, as Thr146 is either absent or replaced by a lysine or an alanine in the other isoforms. Met157 is a glycine or alanine in FlaB3 and FlaB2 respectively, which would hinder its interaction with the FlaA2 residue Pro66. In total, three of the eight sheath-glycan interactions and four of the six protein-protein contacts between the sheath and core in our *fcpA*$^{-}$ structure are FlaB4-specific, suggesting that both FlaA2 and FlaAP show specificity for a FlaB4 core.

Lateral interactions of the FlaA2 and FlaAP sheath proteins

The FlaA2 and FlaAP sheath proteins form adjacent rows decorating the inner curvature of the FlaB4 core. Each FlaA2 monomer within the row is visibly distinct, with the jellyroll segments of each FlaA2 spaced ~15 – 20 Å from the adjacent monomer. However, there is one 'bridging loop' (corresponding to residues 178-196) that is able to bridge this gap between FlaA2 molecules; Arg188 in this loop forms a probable cation-pi interaction with Tyr144, located on a short loop on the adjacent distal FlaA2 monomer (Figure 18A). Glycan interactions are also involved in lateral FlaA2 associations, as the glycan at Ser187 interacts with both Lys156 from one FlaA2 monomer and Val110 and Lys113 from the next distal FlaA2 molecule (Figure 18A). A striking feature of FlaA2 is the C-terminal 'tentacle', corresponding to residues 226-241, which bridges the ~15 Å gap between FlaA2 monomers and occupies density on the adjacent FlaA2; this C-terminal region was the only segment of FlaA2 that was not well-predicted by AlphaFold2 (Figure 13A). The terminal Trp241 of the 'tentacle' is located in a prominent hydrophobic pocket formed by the distal FlaA2 (Figure 18C). The localization of this 'tentacle' to the adjacent monomer is supported by cross-link mass spectrometry studies (performed by F. San Martin), which identify a crosslink between Lys231 of one monomer (located on the 'tentacle') and Lys82 of the adjacent monomer. Together, these loop interactions may contribute to the formation of a stable FlaA2 row along the filament inner curvature.

In contrast to neighboring FlaA2 molecules, neighboring FlaAP monomers do not have a significant gap between them, as the helical bundles of adjacent monomers directly abut each other (Figure 18B). However, few protein-protein contacts are made

between adjacent helices from neighboring FlaAPs; these contacts are mostly bulky and hydrophobic in nature. The long, flexible loop regions of adjacent FlaAP molecules also

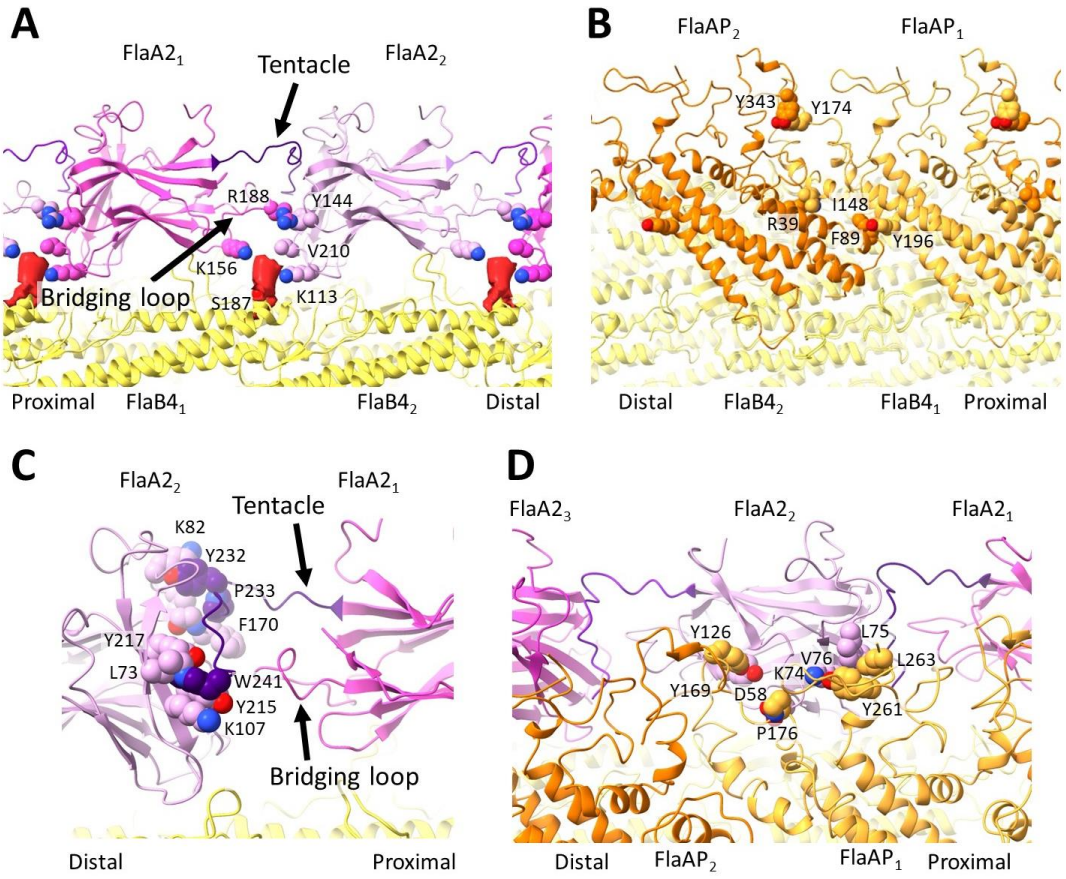

Figure 18. Interactions between neighboring FlaA2 and FlaAP molecules are mediated by flexible loops. **A,** The glycan on Ser187 bridges two FlaA2 monomers. The C-terminal 'tentacle' is colored in purple. A cation-pi interaction between Arg188 of one monomer and Tyr144 of an adjacent FlaA2 is also shown. **B,** Lateral interactions of FlaAP. Tyr174 of one monomer interacts with Tyr343 of an adjacent monomer, and Phe89 of one monomer interacts with an adjacent Tyr196. Arg39 also interacts with the residue Ile148 on the adjacent FlaAP monomer. **C,** Zoomed-in view of the C-terminal 'tentacle' of one monomer reveals interactions with the adjacent monomer. Trp241 of the tentacle is nestled in a hydrophobic pocket formed by the adjacent FlaA2, and hydrophobic interactions also occur with the tentacle residues Y232 and P233. **D,** Interactions between FlaA2 and FlaAP. A hydrophobic interaction occurs between Leu75 and Val76 (FlaA2) and Leu263 (FlaAP). A probable pi-pi interaction is formed between Tyr126 (FlaA2) and Tyr169 (FlaAP). Charged interactions are present between Lys74 (FlaA2) and Tyr261 (FlaAP), and between Asp58 (FlaA2) and Pro176.

approach each other, pairing Tyr343 of one FlaAP with Tyr174 of the proximal neighbor (Figure 18B).

FlaA2 and FlaAP are closely associated with each other, as predicted by a crosslink identified in mass spectrometry of purified wild-type flagella (performed by F. San Martin) (Figure 15). Both FlaA2 and FlaAP contact protofilament #4 of the FlaB4 core, though each interacts with different loops of the protofilament: FlaA2 with loop 152-160, and FlaAP with loop 175-184 (Figure 17A,C). Nearly all protein-protein contacts between FlaA2 and FlaAP are hydrophobic in nature, involving long, flexible loop regions of the sheath proteins (Fig. 18D). Together, FlaA2 and FlaAP form a lattice overlaying the inner curvature of the FlaB core. Many of the interactions within this lattice involve long, flexible loops, suggesting that this assembly may be able to accommodate various flagellar curvatures.

Asymmetry of the FlaB4 core reveals a 'seam' associated with sheath binding

A striking feature of our flagellar filament structure from the *fcpA*$^{-}$ mutant is the strict localization of both the FlaA2 and FlaAP proteins to the inner curvature of the FlaB core, with no sheath decoration elsewhere along the filament (Figure 11B). Filaments from our *fcpA*$^{-}$ mutants are visibly less coiled than the wild-type filaments, with a helical pitch of ~2.0 μm and a helical diameter of ~0.4 μm, (compared to the ~0.22 μm and ~0.45 μm of the wild-type purified flagellar filaments) (Gibson, Trajtenberg et al. 2020). Of note, these *fcpA*$^{-}$ helical parameters are similar to the 'normal' form observed in *Salmonella*, *Escherichia coli*, and *Bacillus subtilis*, which have helical pitches of ~2.06 –

2.55 μm and helical diameters of ~0.4 – 0.6 μm (Fujii, Shibata et al. 2008). Though the 'normal' form is the state used for the 'run' motion, it has not been observed structurally, as most flagellar cryo-EM structures are of mutationally straightened filaments (Yonekura, Maki-Yonekura et al. 2003, Maki-Yonekura, Yonekura et al. 2010, Calladine, Luisi et al. 2013, Wang, Burrage et al. 2017, Kreutzberger, Ewing et al. 2020). These 'normal' flagella are predicted to contain a mixture of nine 'L' and two 'R' protofilaments, with both shorter 'R' protofilaments located adjacent to each other along the inner curvature of the filament (Calladine, Luisi et al. 2013).

Our asymmetric FlaB core was resolved to 3.4 Å, allowing for individual fitting of FlaB4 into each of the 11 core protofilaments, and allowing us to observe any helical

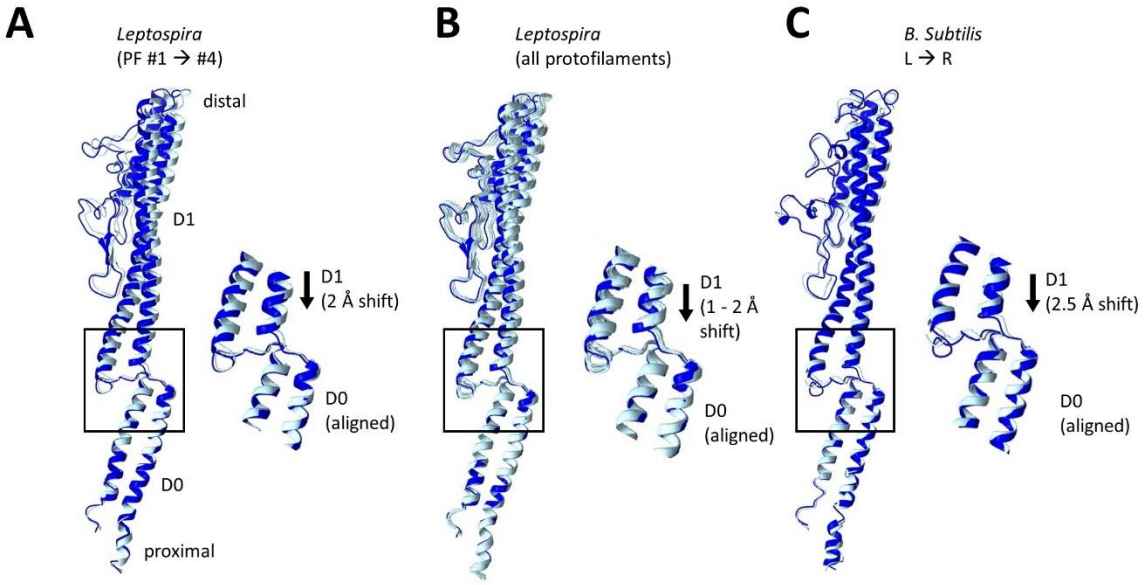

Figure 19. Protofilament #4 undergoes a lateral shift in the D1 domain. **A,** Comparison of the fit models of protofilament #1 and protofilament #4. Insert on the right is a zoomed-in view of the D0-D1 linkers, highlighting the ~2 Å lateral shift of protofilament #4. All of the monomers are aligned to the D0 domain. **B,** Comparison with all 11 individually modeled FlaB4 protofilaments, as in A. **C,** Comparison of the *B. subtilis* L and R protofilaments, revealing a ~2.5 Å lateral shift of 'R' relative to 'L'.

symmetry breaking associated with our supercoiled flagellar form. While neither the glycan pattern nor FlaB isoform identity varied in the sheath-binding region (as previously described), subtle but important differences were observed between the different FlaB4 monomers. Overlaying all 11 FlaB monomers and aligning them to the D0 region revealed that 10 of the 11 protofilaments seem to cluster together and follow similar atomic coordinates, with one protofilament- protofilament #4- appearing to be laterally shifted (by ~2 Å) in the D1 region with respect to the other protofilaments (Fig. 19B). However, all 11 protofilaments are unique, with slight adjustments between each as required for assembly of a curved filament. Comparison between the Z-coordinates of our

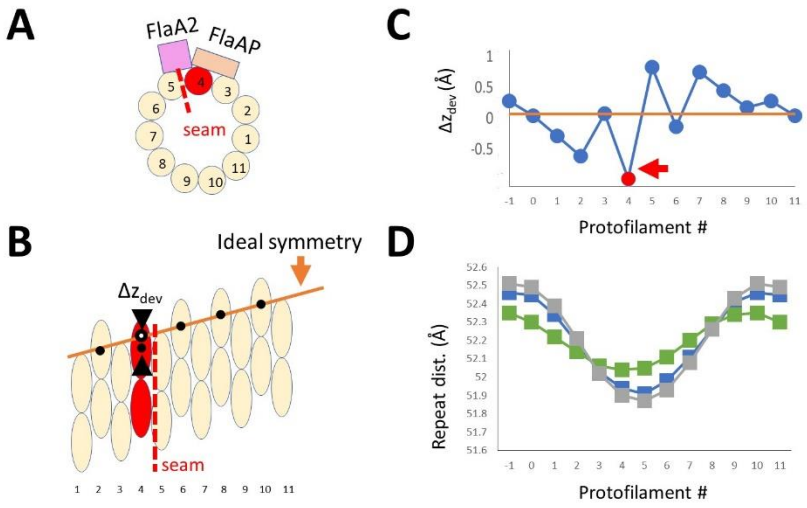

Figure 20. Lateral sliding of protofilament #4 leads to a break in the helical symmetry at the 'seam'. **A,** Diagram showing the numbered protofilaments, as well as the placement of the FlaA2 and FlaAP sheath proteins. Protofilament #4 is colored red, and the seam is marked by a dashed line. **B,** Longitudinal view of two repeats of the FlaB core. Numbering and coloring as in A. The orange line represents what a helically symmetric filament would look like; deviations from this value (Δz_{dev}) are shown for protofilament #4 as the difference between the observed center of mass (in white) and the symmetric center of mass (black).**C,** A graph of Δz_{dev} at each protofilament. This value is the difference between the observed center of mass z values, and the values that would be expected for a helically symmetric filament. **D,** Repeat spacing (in Å) of the D0 domain (green), D1 domain (gray), and the entire monomer (blue), calculated from the center of mass of each domain of all protofilaments.

asymmetric *fcpA*$^{-}$ FlaB model and an idealized symmetric structure, a measurement which we call ΔZ_{dev}, shows that while the largest difference (~ -0.7 Å) is observed in the axially shifted protofilament #4, differences of ~0.6 Å are seen in other protofilaments as well (Fig. 20C). The repeat spacing of each protofilament also varies by ~0.6 Å, with the smallest values occurring at the inner curvature (at protofilaments #4 and #5) and the largest values along the outer curvature (at protofilaments #10 and #11) (Fig. 20D). This

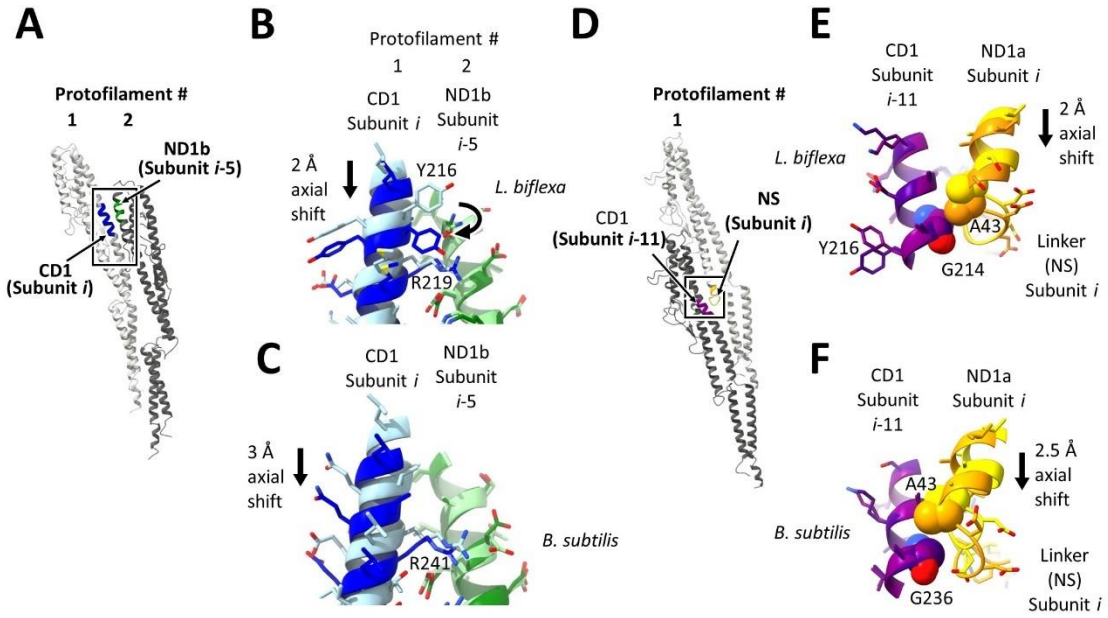

Figure 21. Lateral sliding of protofilament #4 affects contacts with adjacent monomers. **A,** Overview showing a region of contact between the D1 domains of two adjacent protofilaments. **B,** The contact between CD1 and ND1b, comparing protofilaments #1 (light blue) and #2 (light green) to #4 (dark blue) and #5 (dark green). All protofilaments are aligned to the ND1b segment of protofilament #1, revealing that protofilament #4 undergoes a lateral shift. Additionally, Tyr216 of protofilament #4 undergoes a conformational change. **C,** Same alignment as in A, but comparing this interaction in the 'all-L' filament (light blue and light green) to the 'all-R' filament (dark blue and dark green). The 'R' protofilament undergoes a similar shift. **D,** Within a protofilament, the D1 helix interacts with the D0/D1 linker of the next proximal monomer in the row. **E,** The alignment between these regions in protofilament #1 (light purple and yellow) and protofilament #4 (dark purple and orange), aligned to the CD1 helix. Protofilament #4 undergoes an axial shift, allowing Ala43 of one monomer to contact Gly214 of the next. **F,** As in E, but comparing this interaction in the 'all-L' protofilament (light purple and yellow) to the 'all-R' protofilament (dark purple and orange), showing sliding of the D0 domain, but without the hydrophobic contact between the alanine and glycine.

difference is largest in the D1 region, as the repeat spacing in D0 only varies by ~0.3 Å, suggesting that D1 undergoes a larger degree of compression along the inner curvature protofilaments and expansion along the outer curvature protofilaments.

The lateral sliding of protofilament #4 is further apparent when looking at interactions between adjacent protofilaments. One region of contact is between the D1 region of adjacent protofilaments, corresponding to a canonically numbered *i*/*i*-5 connection (Fig. 21A). When comparing to an 'unshifted' pair, the D1 domain of protofilament #4 undergoes a ~2 Å axial shift in the proximal direction; in this shifted protofilament #4, Tyr216 undergoes a significant conformational change, swinging down to engage in hydrogen bond interactions with the D1 domain of protofilament #5 (Fig. 21B). This tyrosine placement is only observed in protofilament #4; in the 10 other monomers, the tyrosine remains pointed distally. This ~2 Å lateral sliding of protofilament #4 is similar in magnitude and direction to the ~3 Å sliding observed in the *B. subtilis* 'R' structure (when compared to the 'L' structure) (Wang, Burrage et al. 2017). However, there was no significant change in side chain orientation in this region between the 'L' and 'R' structures of *B. subtilis* (Fig. 21C).

The axial shift in protofilament #4 is also observed when looking at contacts between adjacent monomers within a protofilament (*i*/*i*-11 contacts); specifically, at the contact region between the D0/D1 linker of one monomer and the D1 of the proximal monomer (Fig. 21D). A ~2 Å lateral shift is also observed in this connection; this shift allows the linker region and start of the D1 helix to come into contact with D1 from the proximal monomer, with a ~3.7 Å distance between Ala43 and Gly214 in protofilament #4 (compared with the 4.3 – 6.1 Å distance seen between these atoms in the other 10

protofilaments) (Fig. 21E). A similar shift of ~2.5 Å is observed between the *B. subtilis* 'R' and 'L' structures (Wang, Burrage et al. 2017); both the alanine and glycine are conserved between the two species, though the distance between the two in both structures (~5.1 Å and ~6.7 Å respectively) is too large for a hydrophobic contact to occur. Interestingly, a mutation in both Ala39 and Asn133 (to a Val and His, respectively) locks the entire *B. subtilis* filament into an R-type conformation (Wang, Burrage et al. 2017).

Due to these similarities with the laterally shifted *B. subtilis* R-type structure, protofilament #4 can be identified as 'R' type, with the remaining 10 protofilaments considered to be 'L' type. However, polymorphic switching of spirochete flagellar filaments has not been observed, and a different flagellar mechanism may instead be present in these organisms. Nonetheless, this difference in flagellar form and the relative 'sliding' of protofilament #4 results in the formation of a seam between protofilaments #4 and #5, breaking the helical symmetry.

Both FlaA2 and FlaAP contact protofilament #4, with FlaA2 overlaying the seam. However, there are few contacts between these sheath proteins and the core; for FlaA2, only nine residues contact the adjacent FlaB glycans, and only four residues directly contact FlaB4 residues. The scarcity of these contacts, the majority of which are hydrophobic in nature, highlights their strength and importance in binding of the sheath to the core. It is likely that the lateral sliding observed along this seam provides the unique interface required for the binding of FlaA2, allowing for the specialized contacts between FlaA2 and the core residues and glycans. FlaAP also binds to the shifted protofilament #4, and likely experiences a similar unique binding interface. Together this

results in the localization of both of these sheath proteins strictly along the inner curvature face of the filament.

FlaA2 and FlaAP are docked into the filament inner curvature

In the ~10 Å wild-type structure obtained through cryo-ET, the FlaA proteins were presumed to be located along the inner curvature, though without any crystal structures, the precise location of each could not be determined (Gibson, Trajtenberg et al. 2020). However, improved structural refinement methods have allowed for a single particle approach to the reconstruction of the tightly coiled wild-type filaments, regardless of a strong preferred orientation. Wild-type filaments were purified by Dr. Elsio Wunder, and grids were prepared by Kimberley Gibson. Dr. Charles Sindelar used cryoSPARC (Punjani, Rubinstein et al. 2017) to reconstruct the wild-type filament, revealing a highly curved asymmetric filament that greatly resembles the structure obtained from cryo-ET (Gibson, Trajtenberg et al. 2020) (Figure 22A).

The resolution of the core is greatly improved, revealing the expected 11-fold symmetry of FlaB. The FcpA and FcpB crystal structures are easily fit into the outer-curvature densities, with six FcpA monomers overlaying the core and four FcpB monomers located in the outermost layer overlaying the FcpA lattice, as was seen in the cryo-ET reconstruction (Figure 22D). Though the unidentified densities along the inner curvature face were resolved to a lower resolution than the FlaB core and the outer sheath region, the resolution is still greatly improved from the corresponding regions of the tomographic structure.

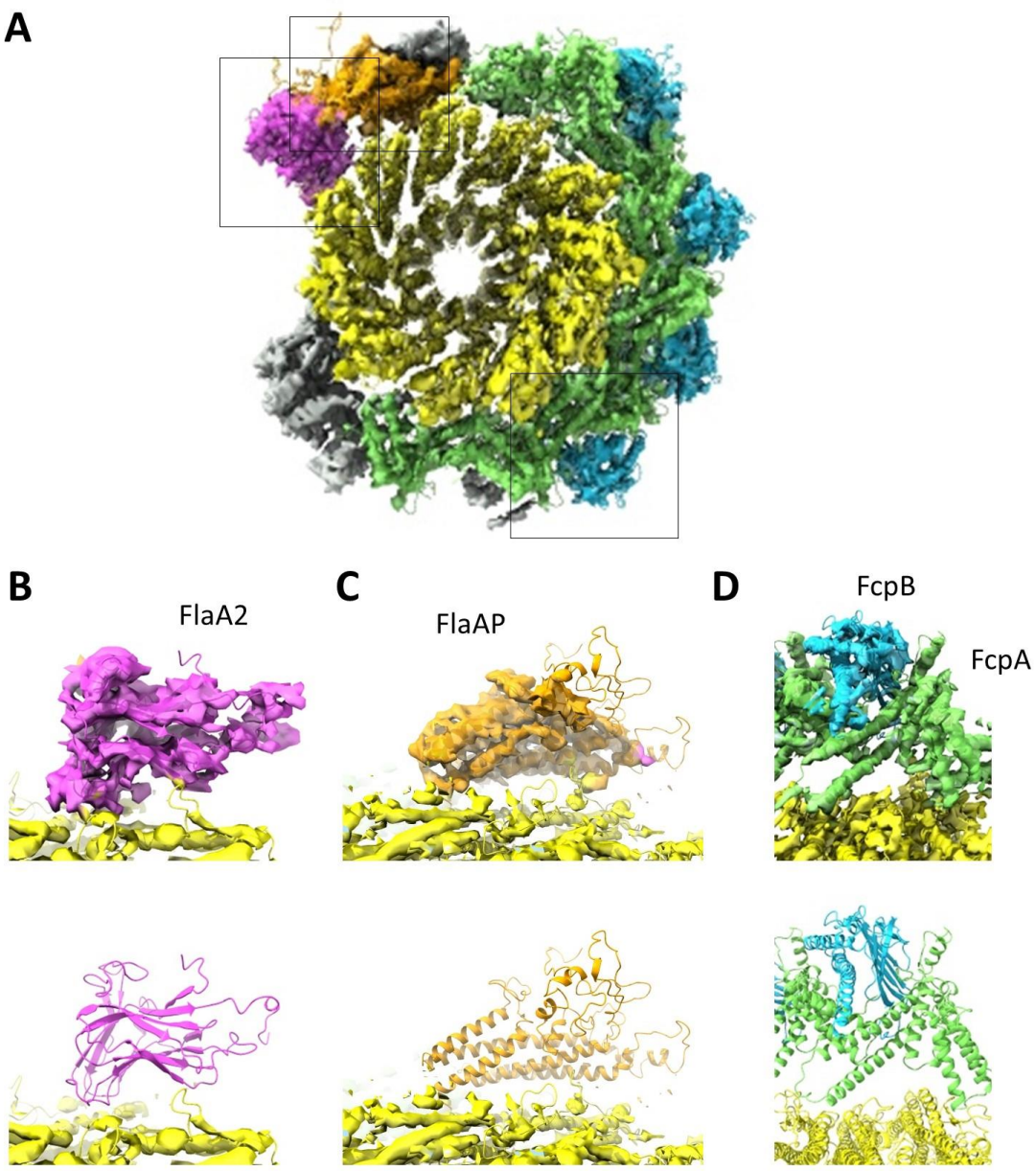

Figure 22. Lattice of FcpA and FcpB coiling proteins is interrupted by FlaA2/FlaAP in the wild-type *Leptospira* filament. **A,** Cross-sectional view of the wild-type structure. Density corresponding to the core is colored yellow, to FlaA2 is pink, to FlaAP is orange, to FcpA is green, and to FcpB is blue. Additional sheath density that does not correspond to these proteins is in gray. The three boxes highlight the regions featured in B, C, and D. **B,** Zoomed-in view of the fit FlaA2 model into the wild-type density. The top panel shows the density, the bottom shows the fit of the model within the density. FlaA2 is colored pink, and the FlaB model is colored yellow. **C,** FlaAP (orange) fit into the density, as in B. **D,** The fit of FcpA (green) and FcpB (blue) into the density, as in B.

As we have now obtained structural model of FlaA2 and FlaAP, we can place both of these sheath proteins into this wild-type density (see Methods). We can assume that FlaA2 and FlaAP closely associate with each other in the wild-type flagellum, as suggested through crosslink mass spectrometry analysis of the wild-type samples (performed by Fabiana San Martin); for the Situs fittings, I therefore used a model containing both one FlaA2 monomer and one FlaAP monomer. This model was fit into one of the inner curvature density lobes that were unidentified in the wild-type structure, with FlaAP closest to an FcpA monomer and with FlaA2 adjacent to the 'groove' containing no sheath density (Figure 22A). The FlaAP structure was nicely fit into a bundle of helices (Figure 22C), and the space between the beta-sheets is apparent in the FlaA2 density (Figure 22B), further supporting this placement in the wild-type structure. Unfortunately, the AlphaFold2 FlaA1 structure could not be unambiguously placed into the remaining, uncharacterized inner curvature sheath density.

Discussion

Glycosylation of the FlaB core is likely common to all spirochete flagella

Three of the residues glycosylated in this flagellum from the *fcpA*$^-$ sample (Ser115, Ser126, and Thr137) were observed to be glycosylated in all three FlaB isoforms of the spirochete *T. denticola* (Kurniyati, Kelly et al. 2017); while Ser115 and Ser126 are present in all four FlaB isoforms of *L. biflexa*, Thr137 is only amenable to O-linked glycosylation in FlaB4, and was one of the glycan 'fingerprints' that suggested our filament had a majority FlaB4 composition. This FlaB4 majority, supported through mass spectrometry on our purified *fcpA*$^-$ samples, differs from the strong FlaB1 majority seen

in mass spectrometry of wild-type *L. interrogans* (Malmstrom, Beck et al. 2009), suggesting that there may be differences in the FlaB quantities either between pathogenic and nonpathogenic species or as a result of the flagellar mutations.

The seven total glycans observed in our structure is in line with the amounts observed in *T. denticola*, where FlaB1 has five sites of glycosylation and FlaB3 has 3-4 (Kurniyati, Kelly et al. 2017). We did not observe any glycosylation of either FlaA2 or FlaAP; likewise, in *T. denticola*, the FlaA sheath was not glycosylated (Kurniyati, Kelly et al. 2017). Mass spectrometry analysis performed by Fabiana San Martin indicated that the glycan observed in our *L. biflexa* samples is different from the novel glycan decorating the *T. denticola* flagellin (Kurniyati, Kelly et al. 2017), though unfortunately the exact identity of the sugar remains unknown (data not shown).

The region of FlaB that is glycosylated in *T. denticola* is not only conserved amongst spirochete FlaBs but is also a TLR5 (Toll-like receptor 5) interaction site in the *Salmonella* FliC (Yoon, Kurnasov et al. 2012, Kurniyati, Kelly et al. 2017). TLR5, an immune system protein, is able to recognize this highly conserved D1 domain of flagella, and can thus initiate an innate immune response against the flagellated pathogen (Andersen-Nissen, Smith et al. 2005). However, *Leptospira* is able to avoid TLR5 recognition, likely due to the periplasmic location of their flagella and the presence of a protein sheath surrounding the FlaB core (Holzapfel, Bonhomme et al. 2020); indeed, the absence of the sheath proteins (in the *fcpA*$^{-}$ or *flaA2*$^{-}$ mutants) did not affect TLR5 signaling, while the absence of FlaB (in the *flaB4*$^{-}$ mutant) did, suggesting FlaB is the only flagellar protein with a role in TLR5 signal induction (Holzapfel, Bonhomme et al. 2020). It is conceivable that the glycosylation seen in D1 in both our *fcpA*$^{-}$ *Leptospira*

mutant and all three FlaB isoforms of *T. denticola*, the locations of which are in the TLR5 interaction domain, could thus provide another method of avoiding TLR5 recognition (Kurniyati, Kelly et al. 2017, Holzapfel, Bonhomme et al. 2020).

Glycosylation of flagella is observed across many different bacterial species, and is often required for proper assembly of the filament and motility of the bacteria (Goon, Kelly et al. 2003, Schirm, Soo et al. 2003, Ewing, Andreishcheva et al. 2009, Kurniyati, Kelly et al. 2017, Yang, Kao et al. 2021). Though no glycosylation mutants currently exist for *Leptospira*, it is conceivable that mutations to the glycosylated residues or in glycan-synthesizing genes would result in a similarly affected flagellum; though there are some protein-protein contacts between the FlaB core and the FlaA2/FlaAP sheath proteins, the interactions between the core glycans and the positively charged sheath residues are more numerous and seem to play an important role in the filament structure. It is likely, therefore, that a loss of these glycans would render the sheath less stable, potentially affecting motility of the bacteria.

The FlaB core in the *fcpA*$^{-}$ mutant filament appears to have both L and R protofilaments

While most flagellar structures are of filaments forced into straight 'all-L' or 'all-R' configurations, our structure represents a filament in its supercoiled state. This supercoiled form appears to resemble a 10L/1R form, with the 'R' protofilament (#4) located along the inner curvature, underlying both the FlaA2 and FlaAP sheath factors. This protofilament is shifted ~2 Å laterally with respect to the other 10 protofilaments, resulting in altered interactions with the neighboring FlaBs. Within protofilament #4, this

causes Ala43 of one monomer to contact Gly214 of the adjacent monomer; while these residues are both conserved in *B. subtilis*, the corresponding L to R lateral shift does not result in the same hydrophobic contact (Wang, Burrage et al. 2017), suggesting that there may be slight differences between our 'R' form and the canonical 'R' forms.

In our FlaB core, Tyr216 changes its orientation in protofilament #4 (the 'R' protofilament), affecting the hydrogen bonding network with the adjacent protofilament (#5); however, a shift in side chain orientations is not observed in the corresponding residue (Val238) in the L and R structures of *B. subtilis* (Wang, Burrage et al. 2017). This tyrosine is conserved amongst all FlaB isoforms in *L. biflexa*, *L. interrogans*, *T. denticola*, and *T. pallidum*, while in *Borrelia* it is a phenylalanine and in the exoflagellates (*B. subtilis*, *Campylobacter*, *Salmonella*) it is a short, hydrophobic isoleucine or valine. It would be interesting to see how a mutation to this tyrosine would affect the packing in the spirochete core, as it appears to be important for proper interactions with the neighboring protofilament and seems to play a role in the L to R transition within the supercoiled filament.

Role of FlaA in other spirochetes

While FcpA, FcpB, and FlaAP are specific to *Leptospira* (Wunder, Figueira et al. 2016, Wunder, Slamti et al. 2018), all spirochetes contain FlaA isoforms (Charon, Greenberg et al. 1992, Li, Motaleb et al. 2000). Of the other spirochetes, only *T. denticola* is currently known to contain an additional flagellar sheath factor, FlaG (which is not present in *Leptospira*) (Kurniyati, Liu et al. 2019). The FlaA proteins of the

spirochetes range from 239 residues (*L. interrogans* FlaA2) to 350 residues (*T. pallidum* FlaA1), and share significant sequence homology (Figure 23).

Glycosylation of flagella is commonly seen in many bacterial species and is observed or predicted to occur in many spirochetes (Ge, Li et al. 1998, Wyss 1998, Logan 2006, Kurniyati, Liu et al. 2019). It is conceivable, therefore, that FlaA sheath proteins from all spirochetes may also interact closely with the glycosylated FlaB core, as occurs with FlaA2 and FlaAP in our *fcpA*$^{-}$ structure. Nine FlaA2 residues interact with the underlying FlaB glycans in *L. biflexa*; these residues are usually positively charged (Lys, Arg, His), though aromatics are also commonly seen in glycan binding, as they have the ability to pi-stack with the sugars (Samanta and Chakrabarti 2001). Upon sequence alignment of the spirochete FlaA proteins, few of these nine glycan binding

Figure 23. Sequence alignment of spirochete FlaA proteins, performed with Clustal Omega. The nine glycan-interacting residues of *L. biflexa* FlaA2 are highlighted in orange, and the C-terminal residue involved in a hydrophobic interaction is highlighted in blue.

residues are conserved with either a bulky, positively charged or aromatic residue: one residue in *Borrelia* FlaA, three in *T. denticola* FlaA, three in *T. pallidum* FlaA1, two in *T. pallidum* FlaA2, one in *L. biflexa* FlaA1, and two in *L. interrogans* FlaA1. However, as interaction with glycans requires the residues to be solvent accessible, it may be that the overall fold of these other FlaA proteins presents different residues for glycan binding.

Interestingly, however, the C-terminal Trp of FlaA2, which is located on the 'tentacle' and interacts with a hydrophobic pocket on the adjacent FlaA2 monomer, is highly conserved amongst the spirochetes. This residue is a hydrophobic aromatic in all spirochete FlaA, except in *Leptospira* FlaA1 (where it is instead a Glu) (Figure 23). While this residue forms the C-terminus in *Leptospira* FlaA2 and *T. pallidum* FlaA2, it is ~30 residues from the terminus in *Borrelia* and *T. denticola* FlaA and the *Leptospira* and *T. pallidum* FlaA1. In the AlphaFold2 predicted structure of FlaA1 (Jumper, Evans et al. 2021), this residue is located along a helix and is predicted to face the beta-sheet core of the structure, making it unlikely to be involved in any longitudinal contacts. However, its strong conservation amongst the rest of the spirochete FlaA sheath factors suggests that this residue may play an important role in hydrophobic protein-protein contacts in these other species as it does in the *Leptospira* FlaA2 protein.

Tenuous nature of sheath-core interactions

One concern with structural analysis of purified flagella is whether all of the flagellar proteins present *in situ* can be accounted for in the reconstruction. Loss of flagellar proteins during the purification process has been previously observed in *Borrelia*, where a Sarkosyl-based purification resulted in flagellar filaments that lacked

FlaA (Ge, Li et al. 1998). In the *fcpB*$^{-}$ tomographic reconstruction (described in Ch. 1), one of the inner curvature densities is missing when compared to the wild-type (Gibson, Trajtenberg et al. 2020); as FcpB immunogold labeling was specific to the outer curvature (Wunder, Slamti et al. 2018), it was assumed that this missing protein corresponded to one of the FlaA proteins. Indeed, this missing lobe corresponds to our FlaAP/FlaA2 placement in the wild-type density; as previous Western blots did not indicate a loss of FlaA2 in the *fcpB*$^{-}$ mutant (Wunder, Slamti et al. 2018), and as our placement of FlaAP/FlaA2 into the WT structure would not interact with FcpB, the lack of FlaAP/FlaA2 density may be due to the purification procedure itself rather than a disruption in stability caused by the missing FcpB sheath layer.

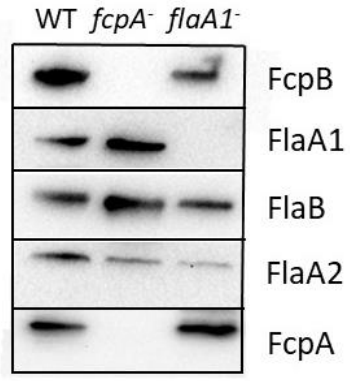

Figure 24. Western blot of purified *L. biflexa* flagella. Wild-type sample is in the left lane, the sample from the *fcpA*$^{-}$ mutant is in the middle lane, and the sample from the *flaA1*$^{-}$ mutant is in the right lane. Each is blotted against FcpB, FlaA1, FlaB, FlaA2, and FcpA. The coiling sheath factors are missing from the *fcpA*$^{-}$ sample and the FlaA1 protein is missing from the *flaA1*$^{-}$ sample, as expected; all other flagellar proteins are present.

Western blots on the *fcpA*$^{-}$ samples used in our structure determination indicated that both FlaA1 and FlaA2 were present in our sample (Figure 24), a conclusion supported by mass spectrometry (Table 2); however, our resultant structure only had FlaA2 and the unexpected FlaAP, and was lacking FlaA1. As our final structure represented just ~2% of the particles, it is possible that one of the other classes from the focused 3D classification may have FlaA1 present; however, it is also possible that the core-FlaA1 interaction is more flexible, making a high-resolution filament structure

containing this sheath factor more difficult to obtain. Previously published studies also differ with regards to the presence of FlaA proteins in *fcpA*$^{-}$ samples: Western blots of *L. interrogans fcpA*$^{-}$ flagella showed a reduction of both FlaA1 and FlaA2 expression, by ~26% and ~58% respectively (Wunder, Figueira et al. 2016), while Western blots of purified *L. biflexa fcpA*$^{-}$ flagella by a different group showed a large reduction in FlaA2 and an absence of FlaA1 and FlaB4 (Sasaki, Kawamoto et al. 2018).

The discrepancy between the various flagellar preparations may reflect the tenuous nature of the sheath-core connection that is observed in our structure, with less than a dozen total contacts between either FlaAP or FlaA2 and the FlaB core. Our current wild-type models indicate a potential interaction between the FcpA and FlaAP sheath proteins (Gibson, Trajtenberg et al. 2020) and likely between FcpA and the sheath protein present on the opposite inner curvature density lobe, making the *fcpA*$^{-}$ mutants (which would lack any FlaA-FcpA interactions) inherently less stable than wild-type filaments. This may contribute to the discrepancies in sheath composition seen across different flagella purifications, highlighting the importance in validating mutant flagellar composition through mass spectrometry and through comparison with the most stable wild-type structure.

The wild-type filament likely contains additional unknown flagellar proteins

The identification of FlaAP as a flagellar protein was surprising, as it was not identified in any previous study. However, its lower abundance in wild-type *L. interrogans* relative to other sheath proteins (with ~1500 copies of FlaAP per cell,

compared to ~8000, ~4500, ~4000, and ~3500 copies of FcpA, FlaA1, FcpB, and FlaA2 respectively) (Malmstrom, Beck et al. 2009), as well as a predicted molecular weight (~43 kDa) similar to that of other flagellar proteins, could explain its absence from previous work. FlaAP was first identified as a candidate for our unexplained helical density due to our mass spectrometry analysis, as FlaAP was both abundant in the *fcpA*$^{-}$ sample and found to crosslink to FlaA2 in wild-type samples.

However, other unidentified proteins were also found to crosslink to flagellar sheath proteins: a peptide corresponding to LEPBI_I2297 has four crosslinks with FlaA2 and three crosslinks with FcpA, and a peptide corresponding to LEPBI_I3081 has two crosslinks with FcpA and one crosslink with FcpB. As LEPBI_I3081 crosslinks to the Fcp sheath proteins and has homology to FcpB (based on an AlphaFold prediction), it may be located on the outer curvature face of the filament. LEPBI_I2297, on the other hand, may be located along the inner curvature, as a link between the FlaA inner curvature sheath and the Fcp outer curvature one; it is predicted (via AlphaFold) to form a jellyroll fold similar to that seen in both FlaA1 and FlaA2. While LEPBI_I3081 was present in our recent flagella preps, only a small quantity was found, with 30 copies total (compared to the 623 total copies of FlaAP); however, as it crosslinks with FcpA, this protein may be less likely to be found in these mutant *fcpA*$^{-}$ samples. LEPBI_I2297 was not present in any *fcpA*$^{-}$ replicate, as would be expected from a protein that crosslinks to both FcpA and FcpB and that could require those sheath factors in order to bind. The combination of mass spectrometry and structural analysis of the filaments from the *fcpA*$^{-}$ mutant proved crucial for the identification of FlaAP; we are hopeful that a similar procedure with freshly purified wild-type filaments could help to identify these other

unknown sheath factors. However, it is possible that these unknown factors may be loosely attached and could easily fall off the filament during purification, hindering their identification.

Remaining questions

We utilized a straighter *fcpA*$^-$ mutant flagellar filament to determine the structure of FlaA2 and surprisingly identified a novel flagellar sheath protein, FlaAP. However, our structure did not show FlaA1, either in complex with the other FlaA sheath proteins or by itself. It is likely that this sheath factor is present in our sample (as suggested by both the Western blots and mass spectrometry analysis), though it may be more flexible and more likely to 'fall off' the core, hindering the reconstruction. The size of the remaining inner-curvature density in the wild-type structure suggests that FlaA1 also likely has an unknown, closely associated sheath factor. A wild-type structure resolved to a high enough resolution could potentially allow for determination of the FlaA1 structure, and cross-link mass spectrometry could help to identify any additional sheath proteins.

A higher-resolution wild-type structure could also provide crucial insight into the flagellar core. At the current resolution, the specific isoforms and locations of glycosylation cannot be determined; however, a near-atomic resolution structure could show if the protofilaments underlying the coiling proteins have a different composition or glycosylation pattern than those underlying the FlaA factors. Additionally, a high-resolution core structure could show whether the tightly coiled form has additional 'R' protofilaments, as predicted by Calladine (Calladine, Luisi et al. 2013). While an improved wild-type structure can also provide further insight into the sheath-core

connectivity of FcpA, we were able to address this through the utilization of a *flaA2*$^{-}$ mutant, as described in Chapter 3.

Conclusion

The structure of a flagellar filament from an *fcpA*$^{-}$ mutant revealed in this chapter is the first near-atomic resolution structure of a spirochete flagellum. Our FlaA2 structure provides the first details of this conserved spirochete protein, showing interactions with the glycosylated FlaB core. Through mass spectrometry, we also identified a novel sheath protein (FlaAP) in this structure; together, FlaA2 and FlaAP form a lattice along the inner curvature face of the filament. Our near-atomic model of the FlaB core not only identifies FlaB4 as the dominant isoform in our structure, but also reveals a 'seam' underlying the FlaA2/FlaAP sheath binding that is suggestive of the R/L 'polymorphic switching' model of *Salmonella*. While this mutant structure revealed important insights into the structure of spirochete flagella, future work will be required to characterize FlaA1 and to fully identify all sheath proteins of the *Leptospira* wild-type flagellum.

Methods

Strains and culturing of *Leptospira*

Wild-type and *fcpA*$^{-}$ strains (Wunder, Slamti et al. 2018) of *Leptospira biflexa* serovar Patoc strain Patoc I (Paris) were grown in EMJH media at 30°C as previously described (Wunder, Figueira et al. 2016, Wunder, Slamti et al. 2018).

Flagellar purification

Flagella from *fcpA*$^{-}$ mutants were purified as previously described (Miller, Miller et al. 2016). Briefly, 500 mL of the *L. biflexa fcpA*$^{-}$ culture was centrifuged at 8,000xg for 15 min, before being washed with cold PBS and re-centrifuged. The pellet was then resuspended in 30 mL Tris buffer (150 mM Tris (hydroxymethyl amino methane, pH 6.8, with 0.9% sodium chloride), centrifuged as before, and the pellet was resuspended in 15 mL of the Tris buffer and stirred at 4°C for 10 min before the addition of 1.5 mL of 20% Triton X-100. The sample was stirred at room temperature for 1 h, centrifuged at 15,000xg for 45 min, and resuspended in 15 mL Tris buffer. 1000 units of mutanolysin were added to the sample, which was stirred at room temperature for 1 h and overnight at 4°C. The sample was then centrifuged at 8,000xg for 30 min, where the supernatant was kept, to which 2.2 mL of ammonium sulfate was added (giving a final ammonium sulfate concentration of 12.6%). The sample was stirred at 4°C for 20 min, and centrifuged at 120,000xg for 2 h. The resulting pellet was resuspended in water, and the centrifugation was repeated. This final pellet was resuspended in 600 μL water, and the resultant flagella were analyzed with SDS-page, Western blotting, and negative staining to ensure proper protein composition and filament form.

The wild-type *L. biflexa* flagella, as well as the *L. biflexa fcpA*$^{-}$ samples used in the mass spectrometry analysis, were purified using the same method as in Chapter 1 (Wunder, Figueira et al. 2016).

Western blot analysis

Western blots were performed as previously described (Wunder, Figueira et al. 2016, Wunder, Slamti et al. 2018). Briefly, SDS gels of both lysates and purified flagella were run, and Western blots were performed with rabbit anti-FcpA (1:1000), anti-FcpB (1:2000), anti-FlaA1 (1:1000), anti-FlaA2 (1:1000), and anti-FlaB (1:500); each antibody was raised against the respective purified recombinant protein (Cocalico Biological, Inc.). Detection was performed using an HRP-conjugated goat anti-rabbit secondary antibody (1:100,000) (Invitrogen), and images were acquired with a Bio-rad ChemiDoc.

Cryo-EM sample preparation

3-4 μL of purified flagella samples from the *L. biflexa fcpA*$^{-}$ mutant were applied to Quantifoil 1.2/1.3 Cu 300 mesh grids (Ted Pella, Inc., Redding, CA) and 2.5 μL of *L. biflexa* wild-type purified flagella were applied to Quantifoil 1.2/1.3 Cu 300 mesh grids (Ted Pella, Inc., Redding, CA). All grids were plasma discharged in a Gatan Model 950 Solarus Advanced Plasma System with H2/O2 for either 30 seconds (*fcpA*$^{-}$) or for 20 seconds (wild-type). The grids were incubated for 1 minute, and then plunge frozen with a Vitrobot Mark IV, with a blotting time of 6 seconds and a blotting force of 2, all at 18 - 22°C and 100% humidity.

Data collection

Initial movies of filaments from the *fcpA*$^-$ sample were collected through the program SerialEM on the 200kV Thermo Scientific Glacios containing a K2 detector. Movies were collected with a pixel size of 0.896 Å.

Wild-type and additional *fcpA*$^-$ movies were collected on the 300kV Titan Krios microscope, containing a K3 detector, through the program SerialEM. *fcpA*$^-$ movies were collected with super-resolution pixel size of 0.534 Å and a magnification of 81kx. A defocus between -1.5 μm to -3.2 μm was used for the *fcpA*$^-$ samples. For the wild-type sample, a magnification of 130kx was used, with a pixel size of 1.05 Å and a defocus of -2.5 to -3 μm. A total dose of 60 $e^-/Å^2$ was used for *fcpA*$^-$, and a total dose of 54 $e^-/Å^2$ was used for the wild-type, both collected over 40 frames. 11906 *fcpA*$^-$ movies were collected; for 4994 of those movies, image shift was used to take movies at four holes per stage position. 718 wild-type movies were collected; 197 at 0° tilt and 521 at 45° tilt. For all movies, only one image was taken per hole.

RELION reconstruction of the filaments from the *fcpA*$^-$ mutants

Initial Glacios reconstruction

802 movies collected on the Glacios were used to generate an initial helically symmetric model. Movies were motion corrected using MotionCor2 (Zheng, Palovcak et al. 2017), and the CTF was corrected with Gctf (Zhang 2016) in Relion 3.0 (Zivanov, Nakane et al. 2018). Filaments were manually selected using EMAN (Ludtke, Baldwin et al. 1999), resulting in 74,369 image segments. 2D classification was performed, and

26,552 image segments corresponding to filaments without a visible sheath were selected for further reconstruction; these classes were selected, as we desired an initial model of just the core of the flagellum. Helical 3D refinement was performed on these selected image segments, using a 30Å low-pass filter of a *Bacillus subtilis* flagellar filament (EMDB-8852) as a reference (Wang, Burrage et al. 2017). 0.5° local searches were used, and the following helical parameters were applied: 11 asymmetrical units, an initial helical rise of 4.72 Å with a 0.2 Å search between 4.42 Å and 5.02 Å, an initial helical twist of 65.3° with a 0.2° search between 62.3° and 68.3°, a central Z length of 30%, a range factor of local averaging of 2, a psi angular search range of 10°, a tilt angular search range of 15°, an outer tube diameter of 280Å, an initial angular sampling of 0.9°, and using fixed tilt-prior angles.

Initial Krios reconstruction

A total of 11906 movies were collected on the Krios. The initial 4461 movies were also processed with Relion 3.0 (Zivanov, Nakane et al. 2018). MotionCor2 (Zheng, Palovcak et al. 2017) and Gctf (Zhang 2016) were used for motion correction and CTF correction, respectively. crYOLO (Wagner and Raunser 2020) was used for filament selection; this program was trained with a subset of filaments selected from 20 random micrographs with EMAN2 (Tang, Peng et al. 2007), and was run with a box size of 320 pixels and a box distance of 56 pixels. This procedure yielded 807,689 image segments. 2D classification was performed to remove mis-selected particles, yielding 803,924 image segments, and outliers and discontinuities were removed from these filament trajectories using the *pf_smooth* program (Debs, Cha et al. 2020) with the following

filament parameters: a rise per subunit of 4.72Å, a twist per subunit of 65.356°, 11 protofilaments, a window size of 7, a fit order of 2, a minimum filament length of 10, a direction tolerance of 20, and a phi, theta, psi, distance, and twist tolerance of 10. Four iterations of smoothing were performed. After smoothing, 430,755 image segments remained (53.3% of the crYOLO-selected image segments).

3D helical refinement was performed on these remaining image segments, using a 10 Å low-pass filtered map of the FlaB core structure from the Glacios data as a reference, and using the same helical parameters as the Glacios refinement. To account for the asymmetry of the filament, the resulting star file was then expanded 11-fold with relion_particle_symmetry_expand (utilizing the helix function, with 11 protofilaments, a twist of 65.4°, and a rise of 4.72 Å), resulting in 4,738,305 image segments.

The asymmetry and/or heterogeneity present in the sample required additional image analysis steps (Mentes, Huehn et al. 2018) to obtain meaningful 3D reconstructions. Particle subtraction was used for focused 3D classification targeting the sheath. The mask used for the subtraction was cylindrical, 90Å in diameter and along the entire length of the reconstructed filament. The mask was placed to capture a small amount of the FlaB core, in addition to any potential sheath density. 3D classification was first performed on a random subset of ~70,000 image segments from the subtracted particles, using a 15 Å filtered copy of the helical refinement as a reference, the mask used in the subtraction, 5 classes, a regularization parameter of 100, no image alignment, and no helical reconstruction. The classification converged by 27 iterations, and resulted in a sheathed class (9,942 particles, or 14.2% of asymmetric subunits), three bare classes (totaling 48,371 particles, or 69.0% of asymmetric subunits), and a class possibly

corresponding to a part of the sheath that was cut off by the mask (11,823 particles, or 16.9% of asymmetric subunits). These classes were then used as a seed for the subsequent 3D classification on the full ~4 million particle dataset; the same parameters for classification were used. This run converged after 16 iterations, and resulted in one decorated sheath class (416,720 particles, or 8.8% of asymmetric subunits), three bare classes (totaling 4,011,067 particles, or 84.7% of asymmetric subunits), and a final class that may represent part of the sheath that was cut off by the mask (310,518 particles, or 6.6% of asymmetric subunits).

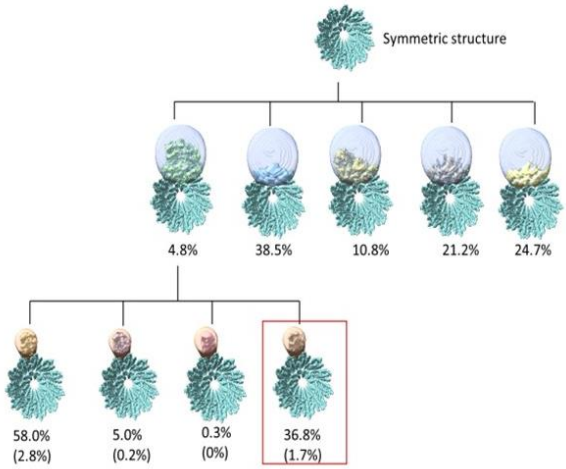

Figure 25. Focused classification was used for sheath refinement in the *fcpA*$^{-}$ mutant filaments. The initial particles were expanded 11-fold (due to the symmetric nature of the filament core). A circular mask was created that would capture sheath densities of the expected sizes (based on the wild-type tomography structure). An initial round of 3D classification resulted in five distinct classes, of which only one had a promising structure. An additional round of classification was performed on this class, with a smaller circular mask focusing on the beta-sheet density. The best-resolved structure from this classification scheme, representing ~2% of the initial expanded particles, was used for all of the analysis in this section.

Final Krios reconstruction

The above micrographs were then reprocessed, together with an additional 7445 micrographs, using Relion3.1 (Zivanov, Nakane et al. 2020). Motion correction, defocus estimation, and particle selection were performed as before, resulting in 1,632,891 total image segments. Initial efforts at helical processing failed to obtain converged alignment parameters for up to half of these image segments, as reflected by numerous discontinuities in the alignments of individual filaments. We therefore generated an initial alignment using one round of Relion refinement (Class3D) with a fine-grained, exhaustive search, with the following parameters: a regularization parameter of 4, an angular sampling interval of 1.8°, an offset search range of 20 pixels with a 1 pixel search step, with no local angular searches performed, and using helical reconstruction with a tube outer diameter of 170 Å, a tilt search range of 15°, a psi search range of 20°, with fixed tilt-priors, and without applying helical symmetry. The reference for this step was the FlaA2 decorated class from the original 4461-micrograph analysis. We discovered that using an 80 Å high-pass filtered reference improved the ability to track continuously along filaments, possibly by de-emphasizing sheath features that could interfere with the alignment. After this classification, *pf_smooth* was applied using the same parameters as before, resulting in 1,538,897 image segments. Mis-selected particles were then identified by 2D classification, resulting in ~1000 image segments being discarded, and Relion 3D helical refinement was performed on the remaining 1,538,130 image segments, using the same parameters as the initial Krios reconstruction. Particle expansion (yielding 16,919,430 image segments) and subtraction were performed as before. Asymmetric 3D classification was then performed, seeded by the classes identified in the 4461-

micrograph analysis, with convergence achieved after 19 iterations. This resulted in a decorated sheath class (816,888 particles, or 4.8% of asymmetric subunits), three bare classes (14,276,994 particles for 84.4% total of asymmetric subunits), and a final class that may represent part of the sheath cut-off by the mask (1,825,548 particles, or 10.8% of asymmetric subunits). An additional round of focused classification was performed on the sheathed class, utilizing a smaller mask focused on the FlaA2 density. This converged after 38 iterations, resulting in one class with strong density (300,616 particles, corresponding to 36.8% of the particles in the second classification and 1.8% of the overall particles), one class with moderate density (40,595 particles, corresponding to 5% of the particles in the second classification and 0.2% of particles overall), and two classes with poorly defined density (475,677 particles corresponding to 58.3% of the second classification and 2.8% of all the particles) (Figure 25).

Reconstruction of wild-type filaments

Gain referencing was performed in IMOD (Kremer, Mastronarde et al. 1996) using the program 'clip'. IMOD newstack was used to bin the particles to a pixel size of 2.19 Å. All further processing and refinement steps were carried out in cryoSPARC v.3.3.1 (Punjani, Rubinstein et al. 2017). Motion correction was performed with the Patch Correction tool (Rubinstein and Brubaker 2015), followed by multi-frame patch CTF estimation, both utilizing default parameters. Particles were then selected with the filament tracer tool in template-free mode, using a 300 Å diameter and a separation distance of ~52 Å. 40,061 particles were extracted, with a box size of 192 pixels. Per-particle local motion correction was applied, and 2D classification was used to remove

any misselected particles; 36,638 particles remained. Helical refinement was then performed, using an initial mask based on the tomographic wild-type structure (Gibson, Trajtenberg et al. 2020), with automasking enabled, and using initial helical parameters of a 0° twist, a 52 Å shift, and a maximum initial tilt search of ±20°. The resultant structure had a reported resolution of 5.44 Å. Local per-particle CTF correction was performed (using default parameters), and helical refinement was repeated; this structure gave a reported resolution of 4.5 Å. Global CTF refinement (which includes beam tilt, spherical aberration, and higher order trefoil and tetrafoil terms) and local CTF refinement were repeated once, followed by another round of helical refinement. The resulting structure reported a resolution of 4.28 Å, roughly the Nyquist limit of the sample. Heterogeneous refinement was used, utilizing two classes; the class with the most complete sheath contained 26,788 particles. Additional helical and non-uniform refinement rounds further improved the high-frequency signal.

Mass spectrometry

Flagella purification for LC-MS/MS

L. biflexa fcpA^{-} mutant filaments were purified for mass spectrometry using the method described above for the wild-type samples. Protein concentration of the extracts were determined by SDS-PAGE and densitometry analysis using the LMW-SDS Marker Kit (GE Healthcare) as a standard.

Nano LC-MS/MS analysis and protein identification

This procedure was provided by Fabiana San Martin. Flagellar extracts were run on 12% acrylamide SDS polyacrylamide gels and processed as previously described (Rossello, Lima et al. 2017). Briefly, 20 µg of proteins of each replicate were run until samples entered 1 cm into the resolving SDS-PAGE. After slicing, each band was destained and cysteine alkylation was performed in-gel by successive incubation with 10 mM dithiothreitol for 1 h at 56 °C and then 55 mM iodoacetamide at room temperature for 45 minutes. In-gel proteolytic digestion was performed overnight at 37°C using sequencing-grade trypsin (Promega). The resulting peptides were extracted at room temperature with 50% acetonitrile/0.1% trifluoroacetic acid. Peptides were desalted using C18 microcolumns (ZipTip® C18, Millipore), vacuum dried and resuspended in 20 µL of 0.1% formic acid.

Three biological replicates of flagella purifications were analyzed using a nano-HPLC (UltiMate 3000, Thermo) coupled to a Q-Orbitrap mass spectrometer (Q Exactive Plus, Thermo). Tryptic peptides (5 µg) were separated into a 75 µm × 50 cm, PepMapTM RSLC C18 analytical column (2 µm particle size, 100 Å, Thermo) at a flow rate of 200 nL/min using a 90 minutes gradient (from 1% to 35% of acetonitrile in 0.1% formic acid). Mass analysis was performed in a data-dependent mode (full scan followed by MS/MS of the top 12 m/z in each segment) using a dynamic exclusion list.

PatternLab for Proteomics (Version V, http://www.patternlabforproteomics.org/) was used for protein identifications (Carvalho, Lima et al. 2016). Briefly, raw data were searched against a target decoy database including *Leptospira biflexa* serovar

Patoc_UP000001847 sequences downloaded from Uniprot (December, 2021) and 127 most common mass spectrometry contaminants. Search parameters were set as follows: enzyme: trypsin; enzyme specificity: full specific; oxidation of methionine as variable modification and carbamidomethylation as fixed modification; 35 ppm of tolerance from the measured precursor m/z. Peptide spectrum matches were filtered using the Search Engine Processor (SEPro) using the following parameters: acceptable FDR: 1% at the protein level; a minimum of two peptides per protein and 10 ppm precursor mass tolerance. Patternlab for proteomics´s Venn Diagram module was used to identify proteins present in all replicates.

Model building

AlphaFold2 (Jumper, Evans et al. 2021) was used to generate structural predictions for the *L. biflexa* proteins FlaB1, FlaB2, FlaB3, FlaB4, FlaA2, and FlaA1. For the *fcpA*$^{-}$ core region, each FlaB AlphaFold model was manually fit into the density in ChimeraX (Pettersen, Goddard et al. 2021), and then Isolde (Croll 2018) was used to individually improve the fit of each isoform. Due to the glycan fingerprint overlaying the model (as described in the results), we identified FlaB4 as the presumptive isoform present in the sample, and used Isolde to individually fit the FlaB4 monomer into each of the 11 protofilaments.

The AlphaFold FlaA2 model was manually aligned to the sheath density in ChimeraX, and then Isolde was used to fit the sequence into the density. Bulky side chains were fit into lobes in the corresponding density, helping to confirm the sequence registration and helping to confirm the identity of the density as FlaA2. An AlphaFold2

structure of FlaA1 was also generated (Jumper, Evans et al. 2021), though the model did not fit well into the density.

Structures of uncharacterized proteins identified in abundance in our sample through mass spectrometry were also predicted with AlphaFold2. One protein, LEPBI_I0551, was predicted to form an alpha-helical bundle, similar to the observed helical sheath density. The model was manually aligned into our density with ChimeraX, and then Isolde (Croll 2018) was used to fit the sequence into the density.

Docking into the wild-type density

The program Situs (Wriggers 2012, Kovacs, Galkin et al. 2018) was employed to dock our FlaA2 and FlaAP models into the wild-type density. Fittings were performed with a model consisting of one FlaA2 monomer and one FlaAP monomer, and only the inner-curvature density (that remained unidentified in the wild-type structure) was used for the docking. The docking was performed with the program colores, using the ~12 Å inner core wild-type map with an applied Laplacian filter, an angular sampling of 10°, and an anisotropy factor of 4. Manual docking of these sheath factors in the wild-type density was consistent with the Situs fittings.

FlaA sequence alignment

Alignment of spirochete FlaA protein sequences was performed with Clustal Omega (Sievers, Wilm et al. 2011).

Chapter III: Loss of FlaA sheath factors allows FcpA and FcpB to bind symmetrically to the FlaB core

Contribution Statement: Dr. Charles Sindelar helped with the analysis of the final *flaA2*$^{-}$ filaments. This work is included in a bioRxiv preprint (doi: https://doi.org/10.1101/2022.03.03.482903).

Introduction

Spirochetes, including *Leptospira*, have a unique, corkscrew-like motility that is driven by the rotation of their periplasmic flagella. Each flagellum is attached to a flagellar motor via a hook region; this assembly is composed of many different proteins. While the loss of some of these proteins effects the motility of the bacteria, the loss of other flagellar-associated proteins seems to have little to no effect (Lambert, Picardeau et al. 2012, Wunder, Slamti et al. 2018). Wild-type *Leptospira* flagellar filaments are tightly coiled when purified (Holt 1978), and consist of at least nine proteins: four isoforms of FlaB (FlaB1-4) (Li, Motaleb et al. 2000), two isoforms of the spirochete sheath factor FlaA (FlaA1 and FlaA2) (Lambert, Picardeau et al. 2012), and three *Leptospira*-specific sheath factors (FcpA, FcpB, and FlaAP) (Wunder, Figueira et al. 2016, Wunder, Slamti et al. 2018). While FcpA and FcpB are not located in an operon and have little sequence

identity (Wunder, Slamti et al. 2018), FlaA1 and FlaA2 form an operon and are 28% identical (Lambert, Picardeau et al. 2012).

FlaA forms a sheath in all spirochete flagella (Charon, Greenberg et al. 1992, Li, Motaleb et al. 2000). Loss of FlaA in *B. burgdorferi*, caused by a harsh flagellar purification method, does not affect the helical parameters of the purified filaments, with similar helical diameters and helical pitches in both filaments lacking and containing FlaA (Motaleb, Sal et al. 2004). However, purified flagella from *flaA*$^{-}$ mutants in *B. hyodysenteriae* have a smaller filament diameter compared with wild-type flagella (~14 nm and ~21 nm, respectively) and remain motile and pathogenic (Kennedy, Rosey et al. 1997, Li, Corum et al. 2000, Li, Wolgemuth et al. 2008).

Leptospira has two FlaA isoforms, FlaA1 and FlaA2; while we were able to solve the structure of FlaA2 and place it into the inner curvature of the wild-type filament density (see Chapter 2), the structure of FlaA1 and its location and function in the wild-type structure remains unknown. Loss of only FlaA1 (*flaA1*$^{-}$ mutant) results in a flagellar morphology that is similar to wild-type; the filaments are still coiled, with no significant loss in filament diameter (Lambert, Picardeau et al. 2012). However, loss of both FlaA1 and FlaA2 (*flaA2*$^{-}$ mutant) results in purified filaments that appear to be much straighter than wild-type filaments (Lambert, Picardeau et al. 2012). As was seen in the *fcpA*$^{-}$ mutants (Wunder, Figueira et al. 2016), these *flaA2*$^{-}$ bacteria are also non-virulent (Lambert, Picardeau et al. 2012); however, unlike in the *fcpA*$^{-}$ sample, these *flaA2*$^{-}$ mutants do not have any significant loss of filament diameter (Lambert, Picardeau et al. 2012). Here I present the near-atomic structure of a flagellar filament from an *L.*

interrogans flaA2^{-} mutant, revealing a symmetric arrangement of the FcpA and FcpB sheath factors around a glycosylated FlaB1 core.

Results

Purified flagella from the *flaA2*$^{-}$ mutant have a straight, not curved, morphology

The purified flagellar filaments from the *L. interrogans flaA2*$^{-}$ sample were analyzed with Western blots, indicating that both FlaA1 and FlaA2 are absent, while FcpA and FcpB are present, along with a lower amount of FlaB (Figure 26). (The reason for this FlaB discrepancy in the *flaA2*$^{-}$ sample is not currently understood). While these mutant flagella appear straighter than the tightly coiled wild-type form, the filaments are highly heterogeneous in nature, with filaments of varying diameters and curvatures (Figure 27). A majority (~83%) of the imaged filaments were thin in nature (~12 nm in

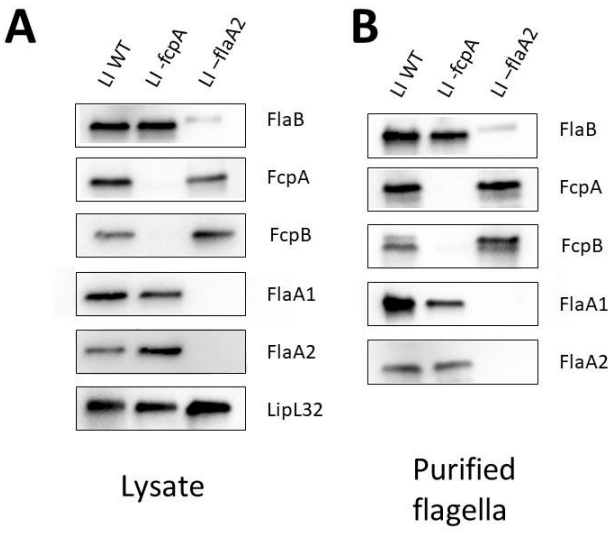

Figure 26. Western blots of *L. interrogans* samples indicates both coiling proteins are present in the *flaA2*$^{-}$ mutant. **A,** Western blot of lysate fractions of a wild-type sample (left lane), as well as an *fcpA*$^{-}$ sample (middle lane) and a *flaA2*$^{-}$ sample (right lane). All three samples are from *L. interrogans*. The samples were blotted against FlaB, FcpA, FcpB, FlaA1, FlaA2, and LipL32 (as a control). **B,** As in A, but with the purified flagellar samples. As LipL32 is an outer membrane protein, it was not included in the purified blots.

diameter), likely reflecting a FlaB core without any sheath proteins present, as has been observed in previous flagellar samples (Lambert, Picardeau et al. 2012, Gibson, Trajtenberg et al. 2020). Two distinct types of sheathed filaments were apparent in our sample: curved filaments (~20 nm in diameter), representing ~2% of the particles, and nearly straight filaments (~24 nm in diameter), corresponding to ~15% of the sample.

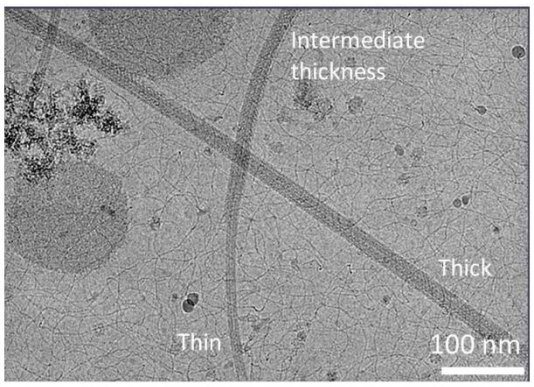

Figure 27. Filaments from *flaA2*$^{-}$ mutants exhibit heterogeneity in filament widths and curvatures. ~15% of the filaments are 'thick' and straight (~24 nm in diameter), 83% are 'thin' and curved (~12 nm, corresponding to just a FlaB core), and 2% have an intermediate thickness and curvature.

FcpA and FcpB assemble symmetrically around the entire *flaA2*$^{-}$ mutant filament

The straight, thick ~24 nm filaments were analyzed structurally, as it was assumed that they contain a complete FcpA and FcpB sheath. An initial 11-fold symmetric refinement of these particles, based upon the known symmetry of the FlaB core, resulted in a high-resolution structure with visible sheath density; when the same refinement was performed on the filaments from *fcpA*$^{-}$ mutants (see Methods, Chapter 2), the resulting symmetric structure did not have any strong visible sheath density. When asymmetric focused classification was performed on the thick *flaA2*$^{-}$ particles, utilizing a

mask that would capture only the sheath region of the filament, no obvious differences in the resultant classes was observed. This contrasts from the *fcpA*$^{-}$ reconstruction, where the same sheath-based asymmetric focused classification revealed distinctly different 'bare' and 'sheathed' classes. As this focused classification of filaments from the *flaA2*$^{-}$ sample did not reveal any stark differences in filament composition, we began to investigate whether these FcpA and FcpB-sheathed flagella may be symmetric, rather than asymmetric, in nature.

The 11-fold symmetric structure was reconstructed with cryoSPARC (Punjani, Rubinstein et al. 2017) and achieved an overall resolution of 2.9 Å (see Methods). Not only did the core achieve high resolution, but the sheath region reached high resolution as well, revealing two distinct sheath layers: a layer closest to the core that appeared to have the helical characteristic of FcpA, and an outer layer consisting of beta-sheets and similar to the structure of FcpB (Figure 28A). This relationship between the two sheath proteins is consistent with their dockings into the wild-type structure, with FcpA located adjacent to the core and FcpB located on the periphery of the filament (Gibson, Trajtenberg et al. 2020). FcpA and FcpB were modeled into the density with Isolde (Croll 2018), utilizing the crystal structures and AlphaFold models for initial alignments (see Methods) (Jumper, Evans et al. 2021). The resultant symmetric structure is drastically different from the asymmetric structures of filaments from the *fcpA*$^{-}$ and wild-type samples, both of which showed sheath proteins localized to just one side of the filament.

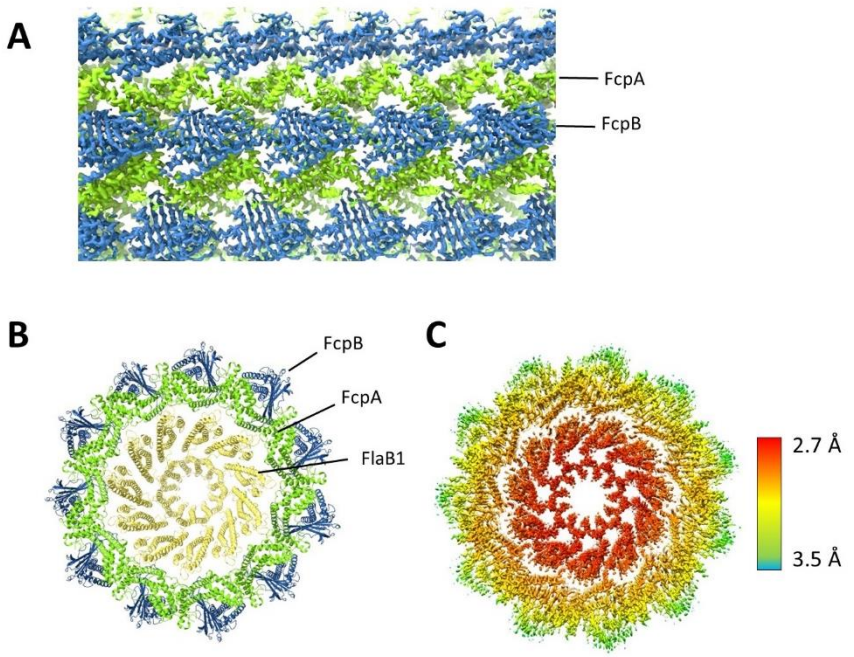

Figure 28. Lattice of FcpA and FcpB in the reconstructed *flaA2*$^{-}$ mutant flagellar filament symmetrically encloses the FlaB core. **A,** FcpA and FcpB lattice shows the characteristic helical and beta-sheet folds, respectively. **B,** Modeled structures, showing the 11-fold symmetry and arrangement of FlaB1, FcpA, and FcpB. FcpA forms a layer enclosing the FlaB core, with FcpB forming the outer most layer. **C,** Local resolution of the filament structure from the *flaA2*$^{-}$ sample, showing near-atomic resolution was reached in the core and sheath regions.

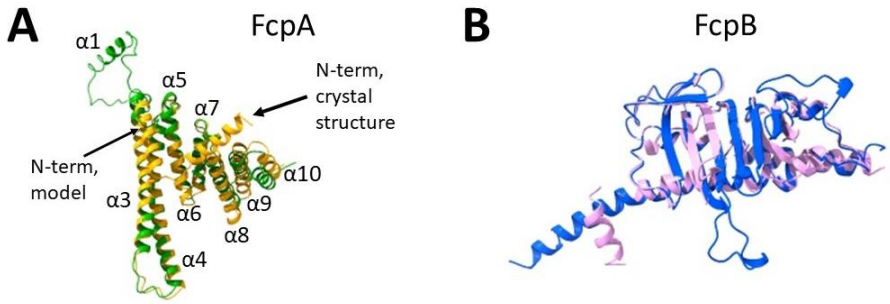

Figure 29. Comparison of Fcp crystal structures to the modeled filament reveals domain swapping and flexing in the FcpA/FcpB lattice. **A,** Comparison of the FcpA crystal structure (orange) and our model (green) reveal a large displacement in the N-terminal helices, as well as flexing between the helical bundles. **B,** Comparison between the FcpB crystal structure (pink) and our model (blue), showing better resolution of loops, and a displaced N-terminal helix.

Each FcpA monomer forms a 'Y'-shaped structure composed of 10 helices, with two long main helices forming one arm (α3,4), and a bundle of five shorter helices forming the other arm (α6-10) (Figure 29A). Each FcpA stretches across two FlaB protofilaments (i.e., protofilaments #1 and #2); the extended nature of this FcpA protein results in close association with six adjacent FcpA molecules, described in further detail below. Our FcpA model matches closely with the previously determined FcpA crystal structure (San Martin, Mechaly et al. 2017, Gibson, Trajtenberg et al. 2020), except in the N-terminal region (residues 44-102) (Figure 29A).

FcpB was modeled into the outer most density lobes of the *flaA2*$^{-}$ mutant flagellar filament, occupying all the remaining sheath density not accounted for by FcpA. Our model is similar to the crystal structure (Gibson, Trajtenberg et al. 2020), revealing a seven-strand beta-sheet overlaying two long helices (Figure 29B); each FcpB can interact with laterally adjacent FcpB monomers, but is unable to contact FcpB molecules forming the neighboring FcpB rows (described below). Each FcpB monomer lies above the FcpA lattice, with extensive interactions between the two coiling factors.

Sheath proteins associate with the FlaB1 core through glycosylation

Near-atomic resolution (~2.7 Å) was achieved in the core region of the filament density from the *flaA2*$^{-}$ sample, revealing the expected 11-fold coiled-coil pattern associated with the FlaB core (Figure 28B,C); in addition, five large, globular densities were seen in each FlaB density, likely corresponding to glycosylation. Each of the five glycans seen in the *flaA2*$^{-}$ mutant (Ser115, Ser126, Ser148, Thr167, and Ser186) were

also observed in the filament structure from the *fcpA*$^{-}$ mutant; however, two glycans (the FlaB4-specific Thr137 and Ser182) were only seen in our *fcpA*$^{-}$ reconstruction and are not present in the *flaA2*$^{-}$ structure.

While the glycan fingerprint of the *fcpA*$^{-}$ mutant filaments allowed for unambiguous FlaB isoform identification, the glycan pattern of the *flaA2*$^{-}$ mutant filaments remained ambiguous with regards to the specific FlaB identity, as the glycan densities would underlie Ser/Thr residues in three of the four FlaB isoforms (excluding only FlaB3). Instead, density of the loop at residues 176-183, as well as the absence of expected side-chain densities at certain FlaB2/FlaB3/FlaB4 residues, identified FlaB1 as the majority isoform in the filament from the *flaA2*$^{-}$ mutant. As in the *fcpA*$^{-}$ structure, it is probable that there may be a mixture of FlaB isoforms, though the above features and the previous *L. interrogans* mass spectrometry studies (Malmstrom, Beck et al. 2009) support a FlaB1-dominant core.

In the *flaA2*$^{-}$ structure, contact is made between three FcpA residues and three core-associated glycans, with one of the glycans located along one protofilament and the other two glycans located on the adjacent FlaB protofilament (Figure 30A). Two of these glycan contacts are with tryptophan residues, likely reflecting pi-stacking with the sugar (Samanta and Chakrabarti 2001). Surprisingly, it appears that a positively charged residue of FcpB interacts with a core glycan; this loop (corresponding to residues 111-124) is the only location where FcpB seems to directly interact with the core (Figure 30C), as otherwise FcpB only contacts the FcpA lattice.

In addition to these glycan interactions, both FcpA and FcpB also contact the FlaB1 core through protein-protein interactions. There is one protein-protein contact

between FcpB and FlaB1, occurring in the same extended loop region as the FcpB-glycan contact (Figure 30D). There are five contacts between FcpA and FlaB1, involving three

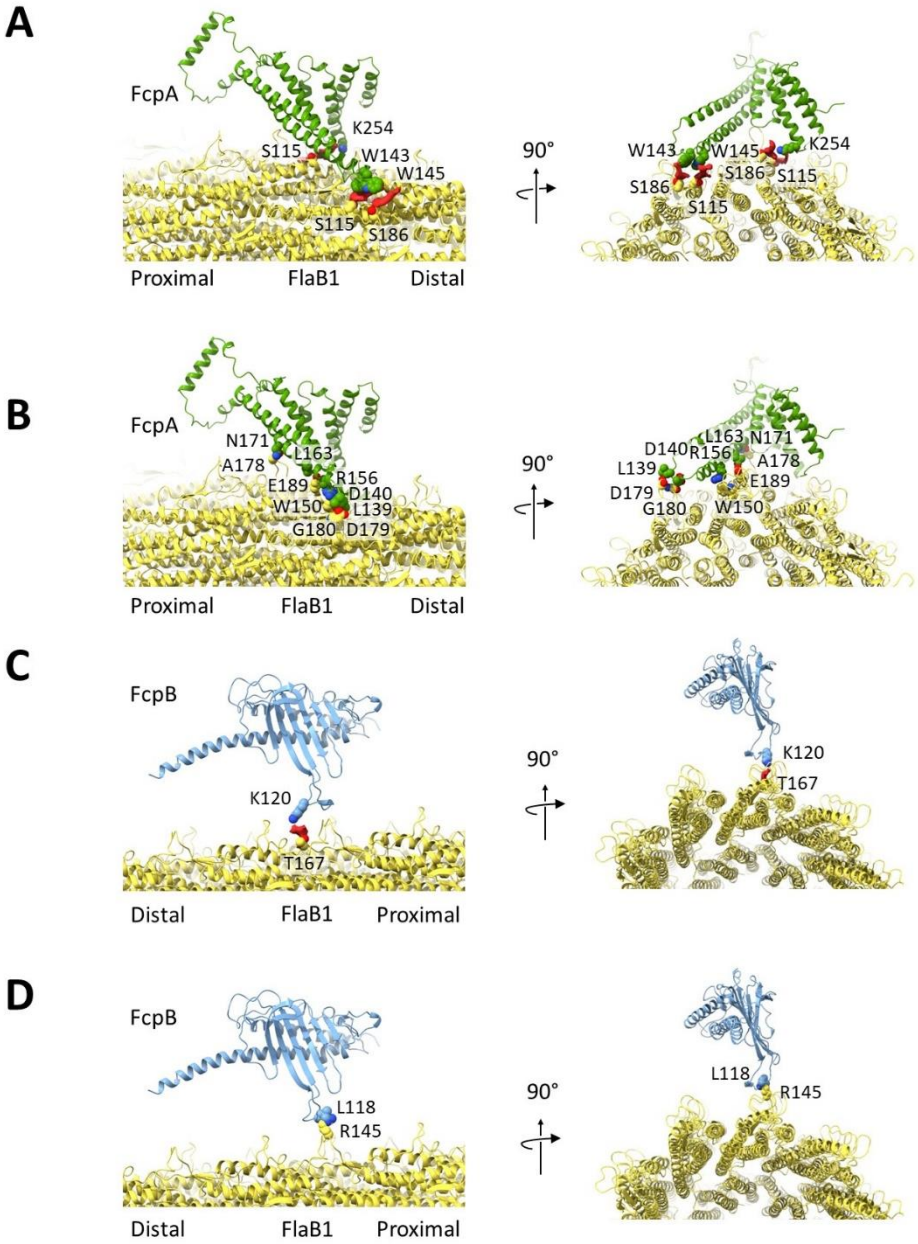

Figure 30. Core-sheath interactions with the Fcp proteins involve glycans and few protein-protein contacts. **A,** Glycans at Ser115 and Ser186 of FlaB1 contact three FcpA residues (Trp 143, Trp 145, and Lys254). Cross-sectional view is shown on the right, highlighting how these interactions bridge multiple FlaB1 protofilaments. **B,** Few protein-protein contacts between FlaB1 and FcpA are present, and mostly involve hydrophobic interactions. **C,** One loop of FcpB interacts with the underlying core; Lys120 contacts the glycan at Thr167. **D,** Lys118 on that same loop contacts Arg145 of FlaB1.

FlaB1 monomers on two FlaB1 protofilaments (Figure 30C); as was seen with the FlaA/FlaB4 interactions, these contacts are sparse and are mostly hydrophobic in nature.

FcpA and FcpB form a helical lattice enclosing the FlaB1 core

In the filament from the *flaA2*$^{-}$ mutant, FcpA assembles into 11 rows around the FlaB1 core. The extended, 'Y'-shaped structure of each FcpA monomer allows it to contact two adjacent FlaB1 protofilaments (Figure 31A), as well as six neighboring FcpAs: the two lateral FcpA monomers within the same FcpA row, as well as two longitudinally adjacent FcpAs in each of the two neighboring FcpA rows (Figure 31A).

Extensive lateral interactions occur within each FcpA row. Residues 65-96 of one FcpA monomer are located on two parallel helices connected by a loop (α1-2), with both α1 and α2 oriented perpendicular to the long helix (α3), and with α1 positioned ~15 Å proximally from the start of α3, apart from the main α3-10 bundle (Figure 32). This α1-2 segment interacts with residues on seven of the ten helices of the proximal FcpA: α3, 4, 6, 7, 8, 9, and 10. Significant contacts occur with the helix-loop-helix region spanning residues 225-240 (α7-8), including an especially prominent interaction between Trp78 (on the α1-2 loop) and this α7-8 segment (Figure 31E). Notably, this modeled N-terminal region of FcpA is starkly different from the crystal structure, as in the crystal structure α2 is positioned parallel to α3, rather than perpendicular (Figure 29A). While α1 in both the model and crystal structure maintains the same perpendicular orientation with α3, the modeled and crystal segments are positioned ~52 Å apart along the α3 axis: α1 in the crystal segment packs near the α9-10 loop, while α1 in our modeled structure is

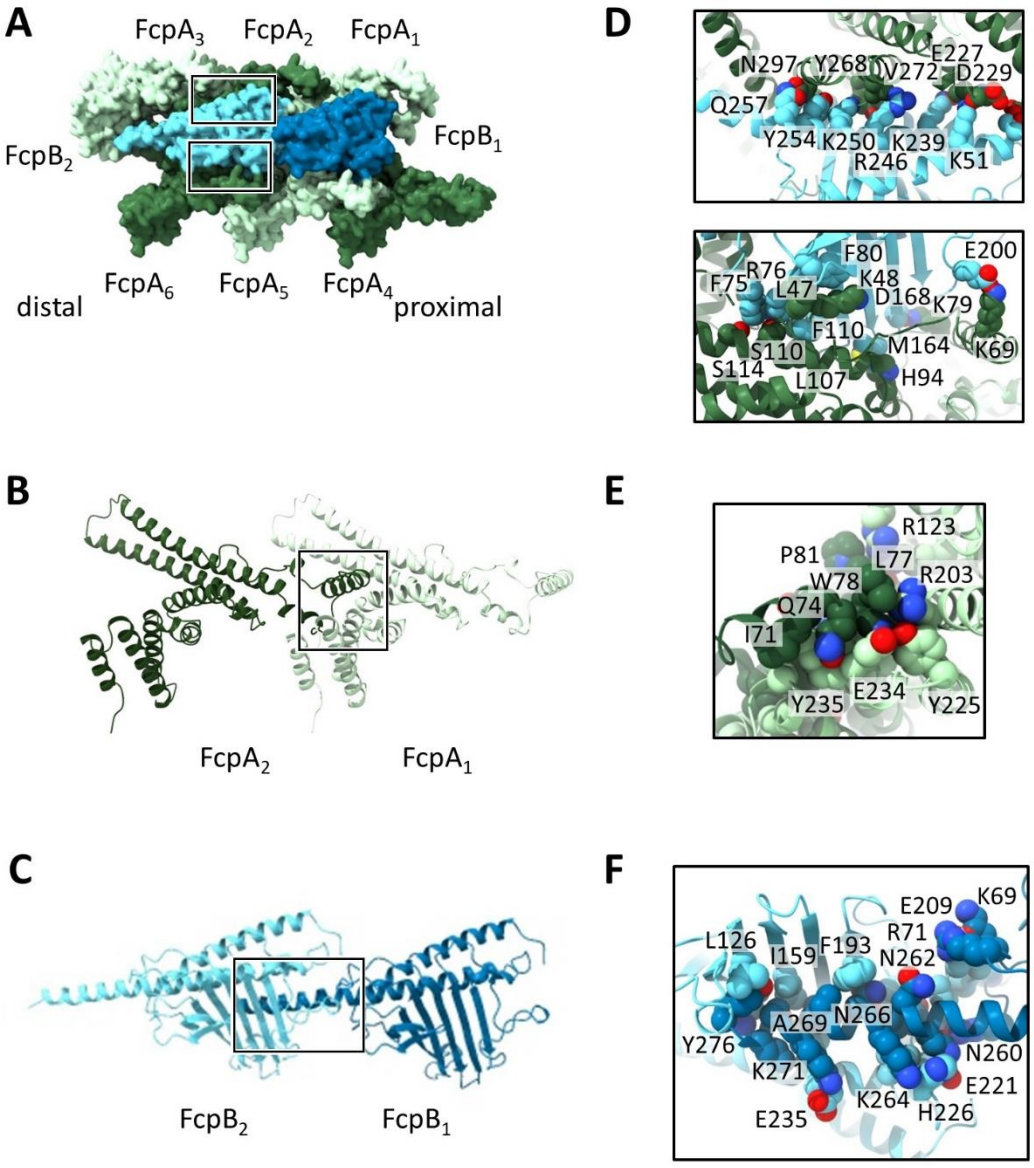

Figure 31. The FcpA/FcpB lattice is characterized by extensive longitudinal interactions. **A,** FcpB overlies a lattice of FcpA, with each FcpB contacting 5 FcpA monomers. Contacts with two lattices are shown in the boxes; top box corresponding to the top panel in D, and the bottom box corresponding to the bottom D panel. **B,** Extensive lateral FcpA interactions. The extended helix of one FcpA monomer contacts the proximal monomer. **C,** Lateral FcpB interactions occur through the elongated helix. **D,** FcpA and FcpB interactions. Top panel, contacts between $FcpB_2$ and $FcpA_2$. Bottom panel, contacts between $FcpB_2$ and $FcpA_6$. **E,** Close-up view of the axial FcpA interactions, including hydrophobic interactions with Trp78 of $FcpA_2$ and a helix of $FcpA_1$. **F,** Lateral interactions of FcpB involve contacts between the beta-sheet of one monomer and the helix of the proximal FcpB.

α1 helix, allowing for these extensive lateral interactions to occur between adjacent FcpA monomers (Figure 32), and likely contributing to the formation of the stable FcpA rows.

Additional contacts are formed between an FcpA monomer and two FcpAs located on each longitudinally adjacent row. Residues on the α3-4 loop of the central FcpA ($FcpA_5$ in Figure 31A) contacts residues on α4, 6, 7, 8, 9, and the α8-9 loop of a longitudinally adjacent FcpA ($FcpA_2$). The C-terminus of the next proximal FcpA in this adjacent row ($FcpA_1$) contacts α3 of the central FcpA ($FcpA_5$). Together, each FcpA monomer contacts six adjacent monomers: one extensive contact with the domain-swapped region of each lateral FcpA monomer, as well as with two separate longitudinal FcpA monomers on each adjacent FcpA row. This network creates a strong FcpA lattice around the entire FlaB core of the *flaA2*$^-$ mutant flagellar filament.

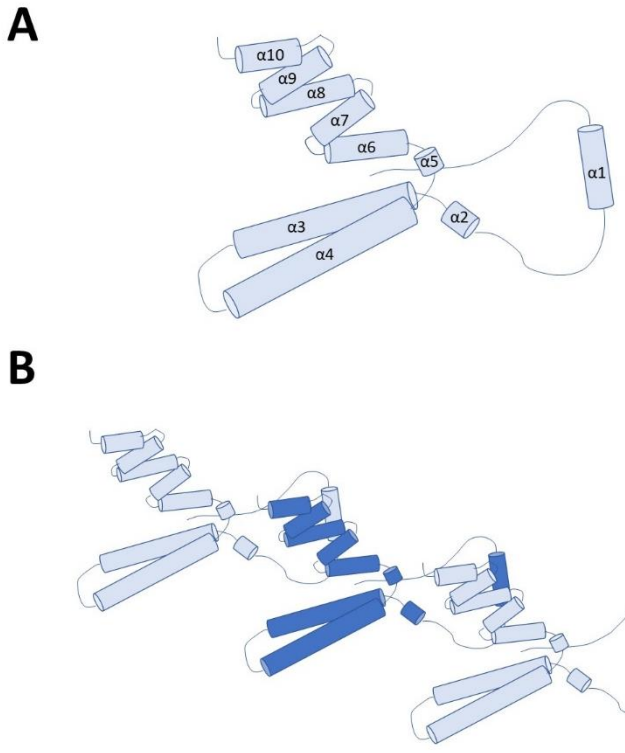

Figure 32. Domain swapping of helix α1 helps to form rows of FcpA. **A,** Diagram of FcpA, with labeled helices. **B,** Three adjacent FcpA monomers within a row. The central monomer is colored a darker blue, highlighting how α1 interacts with the helical bundle of the proximal FcpA monomer.

Each FcpB molecule can only associate with the laterally adjacent FcpB monomers, and cannot contact the FcpBs forming the longitudinally adjacent rows. The C-terminal helix of one FcpB ($FcpB_1$) contacts residues on three of the beta sheets of the distal FcpB monomer ($FcpB_2$), as well as with residues on the proximal part of the C-terminal helix of $FcpB_2$ (Figure 31C). Contacts are also made between extended loops of both adjacent FcpB molecules. These contacts are extensive, aiding in the formation of these FcpB rows that overlay the FcpA lattice.

There are extensive contacts made between FcpB and FcpA, with each FcpB contacting five different FcpA monomers from two different FcpA rows (Figure 31A). Contacts between $FcpB_2$ and three of the adjacent FcpAs ($FcpA_1$, $FcpA_3$, and $FcpA_5$) involve only minor contacts, while interactions with $FcpA_2$ and $FcpA_6$ are more extensive (Figure 31A). During the initial asymmetric classification of the straight filaments from the *flaA2*$^-$ mutant, all classes contained strong FcpB density, with no class solely containing the FcpA lattice. The strong interactions between FcpB and both underlying FcpA rows likely contributes to this stability of FcpB within the filament, aiding in the formation of the complete FcpA/FcpB helical lattice.

Discussion

FlaA2 templates flagellar asymmetry

The symmetric nature of the FcpA and FcpB sheath in the *flaA2*$^-$ structure was surprising, as these proteins were located solely on the outer curvature of the wild-type filament (Gibson, Trajtenberg et al. 2020), and as our *fcpA*$^-$ reconstruction was also

asymmetric in nature. Both the *fcpA*$^{-}$ and *flaA2*$^{-}$ mutant filaments are lacking two sheath proteins- *fcpA*$^{-}$ is lacking the two outer curvature Fcp sheath proteins, while *flaA2*$^{-}$ is missing the inner curvature FlaA1 and FlaA2 sheath proteins- and yet the resulting structures are vastly different in the arrangement of the remaining proteins. The remaining FlaA2 and FlaAP sheath proteins of the *fcpA*$^{-}$ mutant filament remain localized only to the inner curvature of the FlaB core, with no density in any other region of the filament. However, the same is not true with the Fcp sheath proteins.

When only FlaA1 is missing (the *flaA1*$^{-}$ mutant), the filaments are still tightly coiled, and the cells themselves remain highly motile and pathogenic (Lambert, Picardeau et al. 2012). The strong preferred orientation of these coiled filaments has prevented the determination of a high-resolution structure of these flagella. However, when both FlaA1 and FlaA2 are absent (as in the *flaA2*$^{-}$ flagellar mutants), the Fcp proteins are found in the entire sheath region, occupying the spaces where both FlaA1 and FlaA2 bind to the core (Figure 28B).

We hypothesize, therefore, that FlaA2 is able to ‘template’ the binding of the sheath proteins; the presence of FlaA2, bound strictly to the filament inner curvature, prevents the Fcp proteins from binding in that region. However, when FlaA2 is absent, FcpA is able to occupy the inner curvature space, forming a flagellum symmetrically decorated with FcpA (Figure 33). Since FcpB binds strongly to FcpA (while only weakly to the core itself), it likely binds to any available binding sites along the FcpA lattice, whether the lattice is symmetric or still asymmetric in nature.

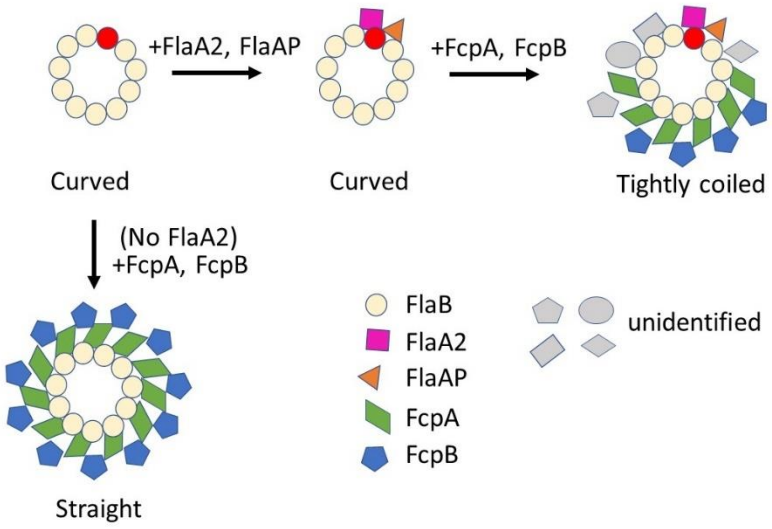

Figure 33. Model for joint activation of *Leptospira* flagellar filament function by multiple sheath proteins. The curved FlaB core (yellow) is formed first; when FlaA2 (pink) and FlaAP (orange) are present, they bind to the shifted protofilament (red) along the inner curvature of the FlaB core, preventing any additional proteins from binding in that area. The FcpA (green)/FcpB (blue) lattice can then form on the outer curvature side of the filament; this allows the filament to adopt a tightly coiled form. Several additional proteins, colored gray, appear to bind to the sheath of the wild-type filament. However, if FlaA2 and FlaAP are not present (as in the *flaA2*$^{-}$ mutant), the FcpA/FcpB lattice can form around the entire FlaB core, forming a symmetric and straight flagellar filament.

Structure of flagella in pathogenic versus saprophytic *Leptospira*

Purified flagellar filaments from the saprophyte *L. biflexa* were used for the reconstructions of wild-type, *fcpB*$^{-}$, and *fcpA*$^{-}$ samples, while the pathogenic *L. interrogans* was used to reconstruct filaments from the *flaA2*$^{-}$ mutant. While *L. biflexa* was primarily used due to its faster growth in culture (Picardeau, Bulach et al. 2008), certain flagellar mutations (such as the *flaA2*$^{-}$ mutant) were only present in *L. interrogans* strains, requiring assumptions on the similarities of the filaments.

All of the identified flagellar proteins (FlaB1-4, FlaA1, FlaA2, and FlaAP) are present in all *Leptospira* species, with degrees of identity between *L. biflexa* and *L. interrogans* ranging from 53% (FcpB) to 92% (FlaB4) (Gibson, Trajtenberg et al. 2020). The two candidates for the unidentified sheath densities (based on *L. biflexa* crosslink mass spectrometry results from Fabiana San Martin), LEPBI_I2297 and LEPBI_I3081, are also found across *Leptospira*, with slightly lower sequence identity with the orthologous *L. interrogans* (45% and 30% identity, respectively). This suggests that the overall structure of the filament is likely conserved in *Leptospira*, as all the expected protein components are conserved.

In our *L.biflexa fcpA*$^{-}$ structure, the glycan fingerprint suggested that the dominant FlaB isoform in the filament was FlaB4 (see Chapter 2). The mass spectrometry of the purified *L. biflexa fcpA*$^{-}$ filaments (performed by F. San Martin) showed similar levels of FlaB1 and FlaB4 in our samples (1111 copies versus 1007 copies across three replicates) (Table 2)- significantly different than the distribution that had previously been reported in wild-type *L. interrogans* (~12,000 copies/cell of FlaB1 versus ~3,500 copies/cell of FlaB4) (Malmstrom, Beck et al. 2009). Mass spectrometry on purified *L. biflexa* wild-type filaments (performed by Dr. Michael Lynch, in the lab of Dr. Brian Crane) also revealed similar levels of FlaB1, FlaB2, and FlaB4 in the sample, with 6-7 peptides corresponding to each; the most abundant protein in this sample was FcpA, with 10 peptides present. Unfortunately, the flagellar concentrations were low, resulting in the low scores, and only one replicate was completed (as the relative abundance of the flagellar proteins was an incidental result from their studies). However, these preliminary findings suggests that there may be a difference in relative FlaB isoform composition

between *L. biflexa* and *L. interrogans*, as in our reconstruction of the symmetric filaments from the *L. interrogans flaA2*$^{-}$ mutants, the dominant core isoform appears to be FlaB1. Further mass spectrometry analysis on both *L. biflexa* and *L. interrogans* wild-type and mutant purified flagella may help to further clarify the relative abundance of the FlaB isoforms in these filaments.

FlaAP and unidentified sheath proteins in the *flaA2*$^{-}$ mutant filament

Prior to the discovery of FlaAP (see Chapter 2), it was assumed that the only flagellar proteins that would be present in the *flaA2*$^{-}$ filaments, which lack both FlaA1 and FlaA2, would be the FlaB core and the FcpA and FcpB sheath proteins. FlaAP is not found in the same operon as FlaA1 and FlaA2, and therefore should not be affected by the mutation in that region of the genome; therefore, it is plausible that *flaAP* is expressed or that the FlaAP protein is present in the lysate. However, we do not see any evidence of FlaAP in our *flaA2*$^{-}$ structure, as all of the density is fully accounted for by FlaB1, FcpA, FcpB, and the core-associated glycans, suggesting that FlaAP is not incorporated into the straight, fully sheathed *flaA2*$^{-}$ flagella. As FlaAP is closely associated with FlaA2, the loss of this interacting partner could destabilize FlaAP, as it has very few contacts with the core: just four charged residues interacting with the core-associated glycans (in contrast to the eight FlaA2 residues that contact glycans), and a handful of weak protein-protein contacts. However, FcpA and FcpB also have a similar number of core-sheath contacts, and it is unclear how these Fcp sheath proteins could dislodge any bound FlaAP. Without an antibody against FlaAP, we cannot determine whether the protein is

present in the lysate or flagella preparations, though future RT-PCR studies may help to determine the levels of *flaAP* gene expression in this *flaA2*$^{-}$ strain.

Furthermore, our latest single particle structure of a coiled wild-type filament revealed additional sheath densities not accounted for by any of the known flagellar proteins. It is likely, therefore, that the *Leptospira* flagellar sheath consists of more than the five known components; as described in Chapter 2, potential candidates from crosslinking mass spectrometry experiments include LEPBI_I2297, which is predicted to form a jellyroll fold and crosslinks FcpA and FlaA2, and LEPBI_I3081, which is predicted to form an FcpB-like fold and crosslinks to both FcpA and FcpB. In order to form the symmetric FcpA/FcpB helical lattice, these other unidentified sheath proteins would also need to be displaced. It is especially curious how the FcpB-homolog would be displaced, as it presumably would bind FcpA in a similar manner.

Potential cooperative binding of the Fcp sheath

The curved, ~20 nm filaments that were present in these micrographs have not yet been reconstructed, as they represented a small minority of the sample (~2% of the picked particles). A structure of these filaments, which due to their curvature would be asymmetric in nature, could provide a better understanding of the assembly of the fully symmetric filaments. These curved filaments likely have only a partial sheath and may provide insight into how the filaments are symmetrized. If this partial sheath contains only the expected six FcpA/four FcpB that is found in the wild-type (but no inner curvature densities), then this would represent a filament that has not yet symmetrized but

has already displaced any additional sheath factors. However, if FlaAP density was observed along the inner curvature of this partially sheathed filament, along with FcpA and FcpB densities along the outer curvature, then this would suggest that FlaAP is indeed present in the sample, and that it must be dislodged before symmetrization of the Fcp proteins can occur.

The binding of the FcpA and FcpB sheath proteins may be cooperative, with the binding of each additional FcpA row (beyond the six in the wild-type) altering the structure to make subsequent binding more likely. It is probable that these subsequent FcpA sheath proteins would bind adjacent to the present FcpA protofilament, interacting with the two monomers along the neighboring protofilament. Therefore, this partial sheath in the curved ~20 nm filament may be heterogeneous, showing evidence of an expanding Fcp sheath. Each additional Fcp protein would alter the curvature of the filament; as FlaAP is located solely along the inner curvature of the filament, it is conceivable that this curvature change would disrupt the weak FlaAP/FlaB interactions, causing FlaAP to fall off and providing space for the Fcp proteins to continue to cooperatively bind around the entire filament. However, it is possible that FlaAP may not be assembled into these filaments at all, in which case this cooperative binding would not need to dislodge this inner curvature sheath protein before forming the full 11-fold symmetry.

As FcpB mainly makes contact with the underlying FcpA lattice (with only one protein-protein contact with the core and one interaction with a core glycan), it is probable that FcpA symmetrically assembles around the core before FcpB binds to any

available binding sites on the FcpA lattice. As FcpB monomers from adjacent rows do not come into contact with each other, each row likely forms independently of the others.

Speculation into FlaB core assembly

The FlaB core is the first part of the flagellar filament to be formed; each FlaB monomer travels through the central channel, assembling onto the distal end of the growing flagellar filament (Yonekura, Maki-Yonekura et al. 2002, Yonekura, Maki-Yonekura et al. 2003). However, it is unknown how the four *Leptospira* FlaB isoforms are assembled: are the isoforms localized to different regions of the filament?

The glycan fingerprint of our *fcpA*$^{-}$ structure seemed to suggest a FlaB4-dominant filament, with FlaB4-specific interactions present between the core and both FlaA2 and FlaAP sheath structures. However, mass spectrometry on the purified filaments identified all four isoforms present in our *fcpA*$^{-}$ sample, including FlaB1 at a similar level as FlaB4, suggesting that there is likely a mixed population that cannot be resolved by cryo-EM. The core of our *flaA2*$^{-}$ sample appears to be composed of FlaB1, though none of the interactions between FcpA and the core glycans are isoform-specific and only a few of the protein-protein contacts show a strong preference for FlaB1.

One can conceive of a FlaB core where FlaB4 proteins are assembled in the inner curvature of the growing filament and FlaB1 monomers assembled into the outer curvature of the filament. In the wild-type structure, this would partition FlaA2 and FlaAP along the inner curvature, and FcpA along the outer curvature. In the *flaA2*$^{-}$ mutant there is no FlaA2/FlaAP present along the FlaB4 inner curvature, and FcpA (with little preference for FlaB1 over FlaB4) can then bind along this inner curvature as well,

forming the fully sheathed, symmetric filaments. However, this is just speculation, and does not account for how FlaB4 appeared to populate the entire core of the *fcpA*$^{-}$ filament and does not account for the locations of the FlaB2 and FlaB3 isoforms. A high-resolution structure of a wild-type filament could provide insight into whether the different protofilaments of the core are composed of different FlaB isoforms, though the high levels of identity between the four isoforms (ranging from 57-72% in *L. interrogans*, for example) (Holzapfel, Bonhomme et al. 2020) may make this challenging if the glycosylation pattern is not distinct in all regions.

Conclusion

The *flaA2*$^{-}$ flagellar filament is symmetric in nature, with an FcpA/FcpB lattice that stretches around all protofilaments of the FlaB1 core. There is an extensive network of contacts between neighboring FcpA monomers, both within a row and with molecules on adjacent rows; in contrast, the connections between FcpA and the core are relatively weak in nature, involving just a handful of glycan contacts and protein-protein interactions. This structure highlights the importance of FlaA2 and its ability to 'template' sheath binding: when FlaA2 and FlaAP are present (as in the wild-type or the *fcpA*$^{-}$ mutant), these proteins localize to the inner curvature, recognizing a 'seam' that breaks the helical symmetry. FcpA and FcpB localize strictly to the outer curvature regions of the wild-type filament, with the FlaA sheath factors occupying the inner curvature. However, in the absence of both FlaA1 and FlaA2, the FcpA/FcpB lattice is able to expand around the entire filament, forming a straight, symmetric flagellum.

Methods

Purification of flagella

L. interrogans flaA2^{-} strains (Lambert, Picardeau et al. 2012) were grown in EMJH media at 30°C, and the flagella were purified as previously described in Chapter 2 (Miller, Miller et al. 2016).

Western blots

Western blots were performed as described in Chapter 2, with an additional rabbit anti-LipL32 antibody (1:1000) (Cocalico Biological Inc.) as a control.

Cryo-EM sample preparation

3-4 μL of the final flagella preparation from *flaA2*$^{-}$ mutants was added to Quantifoil R1.2/1.3 Cu 200 mesh grids (Ted Pella, Inc, Redding, CA) that had been plasma discharged in a Model 950, Solarus Advanced Plasma System with H_2/O_2 for 30 seconds (GATAN). The grids underwent a 30 second incubation time before vitrification with a Vitrobot Mark IV (FEI Company, Eindhoven, Netherlands); a blot time of 6-7.5 s, a blot offset of -2 mm, and humidity of 100% were used.

Data Collection

Grids were screened in a 200 kV Glacios (Thermofisher) for ice thickness and filament concentration. Selected grids were then imaged with a 300 kV Titan Krios, equipped with a K3 detector (GATAN, Inc, Pleasanton, CA), using a magnification of 81kx (for a pixel size of 1.068 Å), a total dose of 60 $e^-/Å^2$, and a defocus range of -1.5 – -2.6 μm. SerialEM (Mastronarde 2005) was used to acquire 2465 movies (40 frames per movie), with only one movie taken per hole.

Initial data processing

flaA2$^-$ movies were initially processed with Relion3.1 (Zivanov, Nakane et al. 2020), utilizing MotionCor for motion correction (Zheng, Palovcak et al. 2017) and gctf for CTF correction (Zhang 2016). crYOLO (Wagner, Lusnig et al. 2020) was used for filament selection; the program was trained using filaments manually selected from 20 micrographs with EMAN2 (Tang, Peng et al. 2007). All flagellar filaments (including those skinny and thick, straight and curved) were selected; 251,215 particles were selected from these 9670 filaments. Straight, thick filaments (~24 nm in diameter) were manually selected for further analysis from the resultant crYOLO selections (corresponding to 36,151 particles). These filaments were extracted in Relion with a box size of 384, and an initial 3D volume was generated using 11-fold symmetry (with a twist of 65.3° and a rise of 4.72 Å), using a low-passed volume (20 Å) of the core FlaB structure (from the initial *fcpA*$^-$ analysis) as an initial reference.

cryoSPARC reconstruction

All subsequent refinement steps were carried out in cryoSPARC (Punjani, Rubinstein et al. 2017). 2D classification was performed to ensure that the majority of the filaments were thicker and straight; no particles were discarded. Helical refinement was then performed, using a low-passed volume of the resultant Relion structure as an initial reference. For all steps, 11-fold symmetry, with a twist of 65.3° and a rise of 4.72 Å, was used. Global CTF refinement and local CTF refinement were used, before another round of helical refinement, utilizing the same parameters as before. The final reconstruction reported a resolution of 2.9 Å.

AlphaFold model building

AlphaFold models of *L. interrogans* FlaB1, FlaB2, FlaB3, FlaB4, FcpA, and FcpB were generated (Jumper, Evans et al. 2021). As with the *L. biflexa* samples (see Chapter 2), each FlaB isoform was individually fit into the density with Isolde (Croll 2018); in this structure, the glycan fingerprint suggested a FlaB1 population. As this flagellum is symmetric, the fit FlaB1 monomer was placed into each FlaB protofilament without additional Isolde refinements.

FcpA and FcpB have been crystallized in *L. biflexa* and *L. interrogans*, respectively (San Martin, Mechaly et al. 2017, Gibson, Trajtenberg et al. 2020). AlphaFold-predicted structures of these proteins in *L. interrogans* closely resembled the crystal structures, and were used as an initial fit into the density. Isolde was then used to further refine the models (Croll 2018).

Future directions

The work presented here naturally leads to many additional questions about the *Leptospira* flagellar filaments. Our refined single-particle wild-type structure highlights how recent advancements in reconstruction software can overcome the issue of filament preferential orientation, an issue which has complicated our studies for years, and can allow for a high-resolution structure to be achieved without the limitations imposed by tomography. An issue seen in all the samples in this work is stability of the sheath layer, with loss of the flagellar sheath observed across preparations. Studies by Fabiana San Martin into the ion-binding ability of the FlaA sheath factors suggests that a slightly modified purification protocol may yield more intact filaments; she graciously provided some samples, and we are hopeful that we may obtain a wild-type structure at a high enough resolution to be able to identify the additional sheath factors, and to potentially even observe any core-sheath contacts in greater detail in their normal, tightly coiled environment. Ideally we would also be able to determine the structure of FlaA1 and see where it assembles in the wild-type filament, and would be able to uniquely identify the FlaB isoform located at each protofilament of the core.

As the finding of FlaAP in our sheath was surprising, it leads one to wonder the impact that it has on the flagellum. Dr. Elsio Wunder recently received a culture of *L. interrogans* that is mutated in the *flaAP* region (LIC_12274); investigation of the motility of these bacteria and purification of their flagella would be interesting future studies. Furthermore, if an antibody was developed to FlaAP, one could then see if it is present in the various mutants (namely the *flaA2*$^{-}$ sample, where it appears to be displaced completely by the FcpA/FcpB lattice).

It would also be interesting to explore how the expression levels of flagellar genes are impacted in the various flagellar mutants. Of particular interest would be the *flaB* genes, to see if the expression levels mirror the isoform findings from our purified flagellar filaments. A full picture of the expression of all of the flagellar genes in both *L. biflexa* and *L. interrogans* wild-type and mutant strains could provide important insight into flagellar composition and how impacts to the flagellar network are affected in these different samples.

There is still so much to learn about *Leptospira* flagellar filaments, as well as filaments from the rest of the spirochete family. I am hopeful that continued work structurally, biochemically, and genetically will help to better understand these important protein complexes and their crucial roles in bacterial motility and virulence.

Agulleiro, J. I. and J. J. Fernandez (2011). "Fast tomographic reconstruction on multicore computers." Bioinformatics **27**(4): 582-583.
Alm, R. A., P. Guerry and T. J. Trust (1993). "Significance of duplicated flagellin genes in Campylobacter." J Mol Biol **230**(2): 359-363.
Andersen-Nissen, E., K. D. Smith, K. L. Strobe, S. L. Barrett, B. T. Cookson, S. M. Logan and A. Aderem (2005). "Evasion of Toll-like receptor 5 by flagellated bacteria." Proc Natl Acad Sci U S A **102**(26): 9247-9252.
Ardissone, S., N. Kint and P. H. Viollier (2020). "Specificity in glycosylation of multiple flagellins by the modular and cell cycle regulated glycosyltransferase FlmG." Elife **9**.
Asakura, H., Y. Churin, B. Bauer, J. P. Boettcher, S. Bartfeld, N. Hashii, N. Kawasaki, H. J. Mollenkopf, P. R. Jungblut, V. Brinkmann and T. F. Meyer (2010). "Helicobacter pylori HP0518 affects flagellin glycosylation to alter bacterial motility." Mol Microbiol **78**(5): 1130-1144.
Beatson, S. A., T. Minamino and M. J. Pallen (2006). "Variation in bacterial flagellins: from sequence to structure." Trends Microbiol **14**(4): 151-155.
Beeby, M., D. A. Ribardo, C. A. Brennan, E. G. Ruby, G. J. Jensen and D. R. Hendrixson (2016). "Diverse high-torque bacterial flagellar motors assemble wider stator rings using a conserved protein scaffold." Proc Natl Acad Sci U S A **113**(13): E1917-1926.
Berg, H. C. (1975). "Chemotaxis in bacteria." Annu Rev Biophys Bioeng **4**(00): 119-136.
Berg, H. C., D. B. Bromley and N. W. Charon (1978). Leptospiral Motility. Relations between structure and function in the prokaryotic cell. In 28th Symposium of the Society for General Microbiology. R. Y. Stanier, H. J. Rogers and J. B. Ward. Cambridge, Cambridge University Press**:** 285–294.
Berg, H. C. and D. A. Brown (1972). "Chemotaxis in Escherichia coli analysed by three-dimensional tracking." Nature **239**(5374): 500-504.
Bharat, T. A., C. J. Russo, J. Lowe, L. A. Passmore and S. H. Scheres (2015). "Advances in Single-Particle Electron Cryomicroscopy Structure Determination applied to Sub-tomogram Averaging." Structure **23**(9): 1743-1753.
Blum, T. B., S. Filippidou, M. Fatton, P. Junier and J. P. Abrahams (2019). "The wild-type flagellar filament of the Firmicute Kurthia at 2.8 A resolution in vivo." Sci Rep **9**(1): 14948.
Bromley, D. B. and N. W. Charon (1979). "Axial filament involvement in the motility of Leptospira interrogans." J Bacteriol **137**(3): 1406-1412.
Burnham, P. M., W. P. Kolar and D. R. Hendrixson (2020). "A Polar Flagellar Transcriptional Program Mediated by Diverse Two-Component Signal Transduction Systems and Basal Flagellar Proteins Is Broadly Conserved in Polar Flagellates." mBio **11**(2).
Calladine, C. R. (1975). "Construction of bacterial flagella." Nature **255**(5504): 121-124.
Calladine, C. R., B. F. Luisi and J. V. Pratap (2013). "A "mechanistic" explanation of the multiple helical forms adopted by bacterial flagellar filaments." J Mol Biol **425**(5): 914-928.
Carroll, B. L. and J. Liu (2020). "Structural Conservation and Adaptation of the Bacterial Flagella Motor." Biomolecules **10**(11).
Carvalho, P. C., D. B. Lima, F. V. Leprevost, M. D. Santos, J. S. Fischer, P. F. Aquino, J. J. Moresco, J. R. Yates, 3rd and V. C. Barbosa (2016). "Integrated analysis of shotgun proteomic data with PatternLab for proteomics 4.0." Nat Protoc **11**(1): 102-117.
Chang, Y., B. L. Carroll and J. Liu (2021). "Structural basis of bacterial flagellar motor rotation and switching." Trends Microbiol **29**(11): 1024-1033.
Chang, Y., H. Xu, M. A. Motaleb and J. Liu (2021). "Characterization of the Flagellar Collar Reveals Structural Plasticity Essential for Spirochete Motility." mBio **12**(6): e0249421.

Charon, N. W., A. Cockburn, C. Li, J. Liu, K. A. Miller, M. R. Miller, M. A. Motaleb and C. W. Wolgemuth (2012). "The unique paradigm of spirochete motility and chemotaxis." Annu Rev Microbiol **66**: 349-370.
Charon, N. W., S. F. Goldstein, S. M. Block, K. Curci, J. D. Ruby, J. A. Kreiling and R. J. Limberger (1992). "Morphology and dynamics of protruding spirochete periplasmic flagella." J Bacteriol **174**(3): 832-840.
Charon, N. W., E. P. Greenberg, M. B. Koopman and R. J. Limberger (1992). "Spirochete chemotaxis, motility, and the structure of the spirochetal periplasmic flagella." Res Microbiol **143**(6): 597-603.
Chen, S., M. Beeby, G. E. Murphy, J. R. Leadbetter, D. R. Hendrixson, A. Briegel, Z. Li, J. Shi, E. I. Tocheva, A. Muller, M. J. Dobro and G. J. Jensen (2011). "Structural diversity of bacterial flagellar motors." EMBO J **30**(14): 2972-2981.
Chu, J., J. Liu and T. R. Hoover (2020). "Phylogenetic Distribution, Ultrastructure, and Function of Bacterial Flagellar Sheaths." Biomolecules **10**(3).
Cope, J., J. Heumann and A. Hoenger (2011). "Cryo-electron tomography for structural characterization of macromolecular complexes." Curr Protoc Protein Sci **Chapter 17**: Unit17 13.
Costa, F., J. E. Hagan, J. Calcagno, M. Kane, P. Torgerson, M. S. Martinez-Silveira, C. Stein, B. Abela-Ridder and A. I. Ko (2015). "Global Morbidity and Mortality of Leptospirosis: A Systematic Review." PLoS Negl Trop Dis **9**(9): e0003898.
Courtney, C. R., L. M. Cozy and D. B. Kearns (2012). "Molecular characterization of the flagellar hook in Bacillus subtilis." J Bacteriol **194**(17): 4619-4629.
Crenshaw, H. C., C. N. Ciampaglio and M. McHenry (2000). "Analysis of the three-dimensional trajectories of organisms: estimates of velocity, curvature and torsion from positional information." J Exp Biol **203**(Pt 6): 961-982.
Croll, T. I. (2018). "ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps." Acta Crystallogr D Struct Biol **74**(Pt 6): 519-530.
Debs, G. E., M. Cha, X. Liu, A. R. Huehn and C. V. Sindelar (2020). "Dynamic and asymmetric fluctuations in the microtubule wall captured by high-resolution cryoelectron microscopy." Proc Natl Acad Sci U S A **117**(29): 16976-16984.
Duda, V. I., N. E. Suzina, V. N. Polivtseva and A. M. Boronin (2012). "[Ultramicrobacteria: Formation of the concept and contribution of ultramicrobacteria to biology]." Mikrobiologiia **81**(4): 415-427.
Eaton, K. A., S. Suerbaum, C. Josenhans and S. Krakowka (1996). "Colonization of gnotobiotic piglets by Helicobacter pylori deficient in two flagellin genes." Infect Immun **64**(7): 2445-2448.
Echazarreta, M. A. and K. E. Klose (2019). "Vibrio Flagellar Synthesis." Front Cell Infect Microbiol **9**: 131.
Edmondson, D. G., B. Hu and S. J. Norris (2018). "Long-Term In Vitro Culture of the Syphilis Spirochete Treponema pallidum subsp. pallidum." mBio **9**(3).
Ely, B., T. W. Ely, W. B. Crymes, Jr. and S. A. Minnich (2000). "A family of six flagellin genes contributes to the Caulobacter crescentus flagellar filament." J Bacteriol **182**(17): 5001-5004.
Evangelista, K. V. and J. Coburn (2010). "Leptospira as an emerging pathogen: a review of its biology, pathogenesis and host immune responses." Future Microbiol **5**(9): 1413-1425.
Ewing, C. P., E. Andreishcheva and P. Guerry (2009). "Functional characterization of flagellin glycosylation in Campylobacter jejuni 81-176." J Bacteriol **191**(22): 7086-7093.
Fontana, C., A. Lambert, N. Benaroudj, D. Gasparini, O. Gorgette, N. Cachet, N. Bomchil and M. Picardeau (2016). "Analysis of a Spontaneous Non-Motile and Avirulent Mutant Shows That FliM Is Required for Full Endoflagella Assembly in Leptospira interrogans." PLoS One **11**(4): e0152916.

Fouts, D. E., M. A. Matthias, H. Adhikarla, B. Adler, L. Amorim-Santos, D. E. Berg, D. Bulach, A. Buschiazzo, Y. F. Chang, R. L. Galloway, D. A. Haake, D. H. Haft, R. Hartskeerl, A. I. Ko, P. N. Levett, J. Matsunaga, A. E. Mechaly, J. M. Monk, A. L. Nascimento, K. E. Nelson, B. Palsson, S. J. Peacock, M. Picardeau, J. N. Ricaldi, J. Thaipandungpanit, E. A. Wunder, Jr., X. F. Yang, J. J. Zhang and J. M. Vinetz (2016). "What Makes a Bacterial Species Pathogenic?:Comparative Genomic Analysis of the Genus Leptospira." PLoS Negl Trop Dis **10**(2): e0004403.
Fujii, M., S. Shibata and S. Aizawa (2008). "Polar, peritrichous, and lateral flagella belong to three distinguishable flagellar families." J Mol Biol **379**(2): 273-283.
Ge, Y., C. Li, L. Corum, C. A. Slaughter and N. W. Charon (1998). "Structure and expression of the FlaA periplasmic flagellar protein of Borrelia burgdorferi." J Bacteriol **180**(9): 2418-2425.
Gibson, K. H., F. Trajtenberg, E. A. Wunder, M. R. Brady, F. San Martin, A. Mechaly, Z. Shang, J. Liu, M. Picardeau, A. Ko, A. Buschiazzo and C. V. Sindelar (2020). "An asymmetric sheath controls flagellar supercoiling and motility in the leptospira spirochete." Elife **9**.
Goldstein, S. F. and N. W. Charon (1988). "Motility of the spirochete Leptospira." Cell Motil Cytoskeleton **9**(2): 101-110.
Goon, S., J. F. Kelly, S. M. Logan, C. P. Ewing and P. Guerry (2003). "Pseudaminic acid, the major modification on Campylobacter flagellin, is synthesized via the Cj1293 gene." Mol Microbiol **50**(2): 659-671.
Guerry, P., R. A. Alm, M. E. Power, S. M. Logan and T. J. Trust (1991). "Role of two flagellin genes in Campylobacter motility." J Bacteriol **173**(15): 4757-4764.
Harshey, R. M. (2003). "Bacterial motility on a surface: many ways to a common goal." Annu Rev Microbiol **57**: 249-273.
Himes, B. A. and P. Zhang (2018). "emClarity: software for high-resolution cryo-electron tomography and subtomogram averaging." Nat Methods **15**(11): 955-961.
Holt, S. C. (1978). "Anatomy and chemistry of spirochetes." Microbiol Rev **42**(1): 114-160.
Holzapfel, M., D. Bonhomme, J. Cagliero, F. Vernel-Pauillac, M. Fanton d'Andon, S. Bortolussi, L. Fiette, C. Goarant, E. A. Wunder, Jr., M. Picardeau, A. I. Ko, D. Werling, M. Matsui, I. G. Boneca and C. Werts (2020). "Escape of TLR5 Recognition by Leptospira spp.: A Rationale for Atypical Endoflagella." Front Immunol **11**: 2007.
Horstmann, J. A., M. Lunelli, H. Cazzola, J. Heidemann, C. Kuhne, P. Steffen, S. Szefs, C. Rossi, R. K. Lokareddy, C. Wang, L. Lemaire, K. T. Hughes, C. Uetrecht, H. Schluter, G. A. Grassl, T. E. B. Stradal, Y. Rossez, M. Kolbe and M. Erhardt (2020). "Methylation of Salmonella Typhimurium flagella promotes bacterial adhesion and host cell invasion." Nat Commun **11**(1): 2013.
Huehn, A., W. Cao, W. A. Elam, X. Liu, E. M. De La Cruz and C. V. Sindelar (2018). "The actin filament twist changes abruptly at boundaries between bare and cofilin-decorated segments." J Biol Chem **293**(15): 5377-5383.
Human Microbiome Project, C. (2012). "Structure, function and diversity of the healthy human microbiome." Nature **486**(7402): 207-214.
Inoue, T., C. S. Barker, H. Matsunami, S. I. Aizawa and F. A. Samatey (2018). "The FlaG regulator is involved in length control of the polar flagella of Campylobacter jejuni." Microbiology (Reading) **164**(5): 740-750.
Izard, J., C. E. Hsieh, R. J. Limberger, C. A. Mannella and M. Marko (2008). "Native cellular architecture of Treponema denticola revealed by cryo-electron tomography." J Struct Biol **163**(1): 10-17.
Izard, J., C. Renken, C. E. Hsieh, D. C. Desrosiers, S. Dunham-Ems, C. La Vake, L. L. Gebhardt, R. J. Limberger, D. L. Cox, M. Marko and J. D. Radolf (2009). "Cryo-electron tomography elucidates the molecular architecture of Treponema pallidum, the syphilis spirochete." J Bacteriol **191**(24): 7566-7580.

Josenhans, C., A. Labigne and S. Suerbaum (1995). "Comparative ultrastructural and functional studies of Helicobacter pylori and Helicobacter mustelae flagellin mutants: both flagellin subunits, FlaA and FlaB, are necessary for full motility in Helicobacter species." J Bacteriol **177**(11): 3010-3020.
Jumper, J., R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Zidek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli and D. Hassabis (2021). "Highly accurate protein structure prediction with AlphaFold." Nature.
Kamiya, R. and S. Asakura (1976). "Helical transformations of Salmonella flagella in vitro." J Mol Biol **106**(1): 167-186.
Kan, W. and C. W. Wolgemuth (2007). "The shape and dynamics of the Leptospiraceae." Biophys J **93**(1): 54-61.
Kato, T., F. Makino, T. Miyata, P. Horvath and K. Namba (2019). "Structure of the native supercoiled flagellar hook as a universal joint." Nat Commun **10**(1): 5295.
Kawamoto, A., T. Miyata, F. Makino, M. Kinoshita, T. Minamino, K. Imada, T. Kato and K. Namba (2021). "Native flagellar MS ring is formed by 34 subunits with 23-fold and 11-fold subsymmetries." Nat Commun **12**(1): 4223.
Kennedy, M. J., E. L. Rosey and R. J. Yancey, Jr. (1997). "Characterization of flaA- and flaB-mutants of Serpulina hyodysenteriae: both flagellin subunits, FlaA and FlaB, are necessary for full motility and intestinal colonization." FEMS Microbiol Lett **153**(1): 119-128.
Klose, K. E. and J. J. Mekalanos (1998). "Differential regulation of multiple flagellins in Vibrio cholerae." J Bacteriol **180**(2): 303-316.
Ko, A. I., C. Goarant and M. Picardeau (2009). "Leptospira: the dawn of the molecular genetics era for an emerging zoonotic pathogen." Nat Rev Microbiol **7**(10): 736-747.
Kovacs, J. A., V. E. Galkin and W. Wriggers (2018). "Accurate flexible refinement of atomic models against medium-resolution cryo-EM maps using damped dynamics." BMC Struct Biol **18**(1): 12.
Kremer, J. R., D. N. Mastronarde and J. R. McIntosh (1996). "Computer visualization of three-dimensional image data using IMOD." J Struct Biol **116**(1): 71-76.
Kreutzberger, M. A. B., C. Ewing, F. Poly, F. Wang and E. H. Egelman (2020). "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**(29): 16985-16991.
Kubanov, A., A. Runina and D. Deryabin (2017). "Novel Treponema pallidum Recombinant Antigens for Syphilis Diagnostics: Current Status and Future Prospects." Biomed Res Int **2017**: 1436080.
Kudryashev, M., M. Cyrklaff, W. Baumeister, M. M. Simon, R. Wallich and F. Frischknecht (2009). "Comparative cryo-electron tomography of pathogenic Lyme disease spirochetes." Mol Microbiol **71**(6): 1415-1434.
Kurniyati, K., J. F. Kelly, E. Vinogradov, A. Robotham, Y. Tu, J. Wang, J. Liu, S. M. Logan and C. Li (2017). "A novel glycan modifies the flagellar filament proteins of the oral bacterium Treponema denticola." Mol Microbiol **103**(1): 67-85.
Kurniyati, K., J. Liu, J. R. Zhang, Y. Min and C. Li (2019). "A pleiotropic role of FlaG in regulating the cell morphogenesis and flagellar homeostasis at the cell poles of Treponema denticola." Cell Microbiol **21**(2): e12886.

Lambert, A., M. Picardeau, D. A. Haake, R. W. Sermswan, A. Srikram, B. Adler and G. A. Murray (2012). "FlaA proteins in Leptospira interrogans are essential for motility and virulence but are not required for formation of the flagellum sheath." Infect Immun **80**(6): 2019-2025.
Leifson, E. (1960). "Atlas of bacterial flagellation."
Lele, P. P., T. Roland, A. Shrivastava, Y. Chen and H. C. Berg (2016). "The flagellar motor of Caulobacter crescentus generates more torque when a cell swims backward." Nat Phys **12**(2): 175-178.
Li, C., L. Corum, D. Morgan, E. L. Rosey, T. B. Stanton and N. W. Charon (2000). "The spirochete FlaA periplasmic flagellar sheath protein impacts flagellar helicity." J Bacteriol **182**(23): 6698-6706.
Li, C., A. Motaleb, M. Sal, S. F. Goldstein and N. W. Charon (2000). "Spirochete periplasmic flagella and motility." J Mol Microbiol Biotechnol **2**(4): 345-354.
Li, C., C. W. Wolgemuth, M. Marko, D. G. Morgan and N. W. Charon (2008). "Genetic analysis of spirochete flagellin proteins and their involvement in motility, filament assembly, and flagellar morphology." J Bacteriol **190**(16): 5607-5615.
Li, H., J. Ruby, N. Charon and H. Kuramitsu (1996). "Gene inactivation in the oral spirochete Treponema denticola: construction of an flgE mutant." J Bacteriol **178**(12): 3664-3667.
Li, Z., F. Dumas, D. Dubreuil and M. Jacques (1993). "A species-specific periplasmic flagellar protein of Serpulina (Treponema) hyodysenteriae." J Bacteriol **175**(24): 8000-8007.
Lis, L. and I. F. Connerton (2016). "The Minor Flagellin of Campylobacter jejuni (FlaB) Confers Defensive Properties against Bacteriophage Infection." Front Microbiol **7**: 1908.
Liu, J., J. K. Howell, S. D. Bradley, Y. Zheng, Z. H. Zhou and S. J. Norris (2010). "Cellular architecture of Treponema pallidum: novel flagellum, periplasmic cone, and cell envelope as revealed by cryo electron tomography." J Mol Biol **403**(4): 546-561.
Liu, J., T. Lin, D. J. Botkin, E. McCrum, H. Winkler and S. J. Norris (2009). "Intact flagellar motor of Borrelia burgdorferi revealed by cryo-electron tomography: evidence for stator ring curvature and rotor/C-ring assembly flexion." J Bacteriol **191**(16): 5026-5036.
Logan, S. M. (2006). "Flagellar glycosylation - a new component of the motility repertoire?" Microbiology (Reading) **152**(Pt 5): 1249-1262.
Lozupone, C. A. and R. Knight (2007). "Global patterns in bacterial diversity." Proc Natl Acad Sci U S A **104**(27): 11436-11440.
Ludtke, S. J., P. R. Baldwin and W. Chiu (1999). "EMAN: semiautomated software for high-resolution single-particle reconstructions." J Struct Biol **128**(1): 82-97.
Lynch, M. J., M. Miller, M. James, S. Zhang, K. Zhang, C. Li, N. W. Charon and B. R. Crane (2019). "Structure and chemistry of lysinoalanine crosslinking in the spirochaete flagella hook." Nat Chem Biol.
Macnab, R. M. (2003). "How bacteria assemble flagella." Annu Rev Microbiol **57**: 77-100.
Maki-Yonekura, S., K. Yonekura and K. Namba (2010). "Conformational change of flagellin for polymorphic supercoiling of the flagellar filament." Nat Struct Mol Biol **17**(4): 417-422.
Malmstrom, J., M. Beck, A. Schmidt, V. Lange, E. W. Deutsch and R. Aebersold (2009). "Proteome-wide cellular protein concentrations of the human pathogen Leptospira interrogans." Nature **460**(7256): 762-765.
Mandadapu, K. K., J. A. Nirody, R. M. Berry and G. Oster (2015). "Mechanics of torque generation in the bacterial flagellar motor." Proc Natl Acad Sci U S A **112**(32): E4381-4389.
Mastronarde, D. N. (2005). "Automated electron microscope tomography using robust prediction of specimen movements." J Struct Biol **152**(1): 36-51.
Mastronarde, D. N. and S. R. Held (2017). "Automated tilt series alignment and tomographic reconstruction in IMOD." J Struct Biol **197**(2): 102-113.

Matsunami, H., C. S. Barker, Y. H. Yoon, M. Wolf and F. A. Samatey (2016). "Complete structure of the bacterial flagellar hook reveals extensive set of stabilizing interactions." Nat Commun **7**: 13425.
McGill, M. A., D. G. Edmondson, J. A. Carroll, R. G. Cook, R. S. Orkiszewski and S. J. Norris (2010). "Characterization and serologic analysis of the Treponema pallidum proteome." Infect Immun **78**(6): 2631-2643.
Mentes, A., A. Huehn, X. Liu, A. Zwolak, R. Dominguez, H. Shuman, E. M. Ostap and C. V. Sindelar (2018). "High-resolution cryo-EM structures of actin-bound myosin states reveal the mechanism of myosin force sensing." Proc Natl Acad Sci U S A **115**(6): 1292-1297.
Miller, M. R., K. A. Miller, J. Bian, M. E. James, S. Zhang, M. J. Lynch, P. S. Callery, J. M. Hettick, A. Cockburn, J. Liu, C. Li, B. R. Crane and N. W. Charon (2016). "Spirochaete flagella hook proteins self-catalyse a lysinoalanine covalent crosslink for motility." Nat Microbiol **1**(10): 16134.
Minamino, T. and K. Imada (2015). "The bacterial flagellar motor and its structural diversity." Trends Microbiol **23**(5): 267-274.
Montemayor, E. J., N. T. Ploscariu, J. C. Sanchez, D. Parrell, R. S. Dillard, C. W. Shebelut, Z. Ke, R. C. Guerrero-Ferreira and E. R. Wright (2021). "Flagellar Structures from the Bacterium Caulobacter crescentus and Implications for Phage varphi CbK Predation of Multiflagellin Bacteria." J Bacteriol **203**(5).
Moon, K. H., X. Zhao, A. Manne, J. Wang, Z. Yu, J. Liu and M. A. Motaleb (2016). "Spirochetes flagellar collar protein FlbB has astounding effects in orientation of periplasmic flagella, bacterial shape, motility, and assembly of motors in Borrelia burgdorferi." Mol Microbiol **102**(2): 336-348.
Motaleb, M. A., L. Corum, J. L. Bono, A. F. Elias, P. Rosa, D. S. Samuels and N. W. Charon (2000). "Borrelia burgdorferi periplasmic flagella have both skeletal and motility functions." Proc Natl Acad Sci U S A **97**(20): 10899-10904.
Motaleb, M. A., M. S. Sal and N. W. Charon (2004). "The decrease in FlaA observed in a flaB mutant of Borrelia burgdorferi occurs posttranscriptionally." J Bacteriol **186**(12): 3703-3711.
Murphy, G. E., J. R. Leadbetter and G. J. Jensen (2006). "In situ structure of the complete Treponema primitia flagellar motor." Nature **442**(7106): 1062-1064.
Nakamura, S. (2020). "Spirochete Flagella and Motility." Biomolecules **10**(4).
Nakamura, S. and T. Minamino (2019). "Flagella-Driven Motility of Bacteria." Biomolecules **9**(7).
Ng, H. M., N. Slakeski, C. A. Butler, P. D. Veith, Y. Y. Chen, S. W. Liu, B. Hoffmann, S. G. Dashper and E. C. Reynolds (2019). "The Role of Treponema denticola Motility in Synergistic Biofilm Formation With Porphyromonas gingivalis." Front Cell Infect Microbiol **9**: 432.
Nogales, E. and S. H. Scheres (2015). "Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity." Mol Cell **58**(4): 677-689.
Norris, S. J. (1993). "Polypeptides of Treponema pallidum: progress toward understanding their structural, functional, and immunologic roles. Treponema Pallidum Polypeptide Research Group." Microbiol Rev **57**(3): 750-779.
Norris, S. J., N. W. Charon, R. G. Cook, M. D. Fuentes and R. J. Limberger (1988). "Antigenic relatedness and N-terminal sequence homology define two classes of periplasmic flagellar proteins of Treponema pallidum subsp. pallidum and Treponema phagedenis." J Bacteriol **170**(9): 4072-4082.
Norris, S. J., B. J. Paster, A. Moter and U. B. Göbel (2006). The Genus Treponema. The Prokaryotes: Volume 7: Proteobacteria: Delta, Epsilon Subclass. M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer and E. Stackebrandt. New York, NY, Springer New York**:** 211-234.
Pallen, M. J., C. W. Penn and R. R. Chaudhuri (2005). "Bacterial flagellar diversity in the post-genomic era." Trends Microbiol **13**(4): 143-149.

Paster, B. J. (2010). Phylum XV. Spirochaetes Garrity and Holt 2001. Bergey's Manual® of Systematic Bacteriology: Volume Four The Bacteroidetes, Spirochaetes, Tenericutes (Mollicutes), Acidobacteria, Fibrobacteres, Fusobacteria, Dictyoglomi, Gemmatimonadetes, Lentisphaerae, Verrucomicrobia, Chlamydiae, and Planctomycetes. N. R. Krieg, J. T. Staley, D. R. Brown et al. New York, NY, Springer New York**:** 471-566.
Pettersen, E. F., T. D. Goddard, C. C. Huang, E. C. Meng, G. S. Couch, T. I. Croll, J. H. Morris and T. E. Ferrin (2021). "UCSF ChimeraX: Structure visualization for researchers, educators, and developers." Protein Sci **30**(1): 70-82.
Picardeau, M. (2017). "Virulence of the zoonotic agent of leptospirosis: still terra incognita?" Nat Rev Microbiol **15**(5): 297-307.
Picardeau, M., A. Brenot and I. Saint Girons (2001). "First evidence for gene replacement in Leptospira spp. Inactivation of L. biflexa flaB results in non-motile mutants deficient in endoflagella." Mol Microbiol **40**(1): 189-199.
Picardeau, M., D. M. Bulach, C. Bouchier, R. L. Zuerner, N. Zidane, P. J. Wilson, S. Creno, E. S. Kuczek, S. Bommezzadri, J. C. Davis, A. McGrath, M. J. Johnson, C. Boursaux-Eude, T. Seemann, Z. Rouy, R. L. Coppel, J. I. Rood, A. Lajus, J. K. Davies, C. Medigue and B. Adler (2008). "Genome sequence of the saprophyte Leptospira biflexa provides insights into the evolution of Leptospira and the pathogenesis of leptospirosis." PLoS One **3**(2): e1607.
Punjani, A., J. L. Rubinstein, D. J. Fleet and M. A. Brubaker (2017). "cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination." Nat Methods **14**(3): 290-296.
Raddi, G., D. R. Morado, J. Yan, D. A. Haake, X. F. Yang and J. Liu (2012). "Three-dimensional structures of pathogenic and saprophytic Leptospira species revealed by cryo-electron tomography." J Bacteriol **194**(6): 1299-1306.
Romeis, E., L. Tantalo, N. Lieberman, Q. Phung, A. Greninger and L. Giacani (2021). "Genetic engineering of Treponema pallidum subsp. pallidum, the Syphilis Spirochete." PLoS Pathog **17**(7): e1009612.
Rossello, J., A. Lima, M. Gil, J. Rodriguez Duarte, A. Correa, P. C. Carvalho, A. Kierbel and R. Duran (2017). "The EAL-domain protein FcsR regulates flagella, chemotaxis and type III secretion system in Pseudomonas aeruginosa by a phosphodiesterase independent mechanism." Sci Rep **7**(1): 10281.
Roux, V., K. El Karkouri, J. C. Lagier, C. Robert and D. Raoult (2012). "Non-contiguous finished genome sequence and description of Kurthia massiliensis sp. nov." Stand Genomic Sci **7**(2): 221-232.
Rubinstein, J. L. and M. A. Brubaker (2015). "Alignment of cryo-EM movies of individual particles by optimization of image translations." J Struct Biol **192**(2): 188-195.
Ruby, J. D., H. Li, H. Kuramitsu, S. J. Norris, S. F. Goldstein, K. F. Buttle and N. W. Charon (1997). "Relationship of Treponema denticola periplasmic flagella to irregular cell morphology." J Bacteriol **179**(5): 1628-1635.
Salah Ud-Din, A. I. M. and A. Roujeinikova (2018). "Flagellin glycosylation with pseudaminic acid in Campylobacter and Helicobacter: prospects for development of novel therapeutics." Cell Mol Life Sci **75**(7): 1163-1178.
Samanta, U. and P. Chakrabarti (2001). "Assessing the role of tryptophan residues in the binding site." Protein Eng **14**(1): 7-15.
San Martin, F., A. E. Mechaly, N. Larrieux, E. A. Wunder, Jr., A. I. Ko, M. Picardeau, F. Trajtenberg and A. Buschiazzo (2017). "Crystallization of FcpA from Leptospira, a novel flagellar protein that is essential for pathogenesis." Acta Crystallogr F Struct Biol Commun **73**(Pt 3): 123-129.

Sarkar, M. K., K. Paul and D. Blair (2010). "Chemotaxis signaling protein CheY binds to the rotor protein FliN to control the direction of flagellar rotation in Escherichia coli." Proc Natl Acad Sci U S A **107**(20): 9370-9375.
Sasaki, Y., A. Kawamoto, H. Tahara, K. Kasuga, R. Sato, M. Ohnishi, S. Nakamura and N. Koizumi (2018). "Leptospiral flagellar sheath protein FcpA interacts with FlaA2 and FlaB1 in Leptospira biflexa." PLoS One **13**(4): e0194923.
Scheres, S. H. (2012). "RELION: implementation of a Bayesian approach to cryo-EM structure determination." J Struct Biol **180**(3): 519-530.
Schirm, M., E. C. Soo, A. J. Aubry, J. Austin, P. Thibault and S. M. Logan (2003). "Structural, genetic and functional characterization of the flagellin glycosylation process in Helicobacter pylori." Mol Microbiol **48**(6): 1579-1592.
Schulz, H. N., T. Brinkhoff, T. G. Ferdelman, M. H. Marine, A. Teske and B. B. Jorgensen (1999). "Dense populations of a giant sulfur bacterium in Namibian shelf sediments." Science **284**(5413): 493-495.
Sievers, F., A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Soding, J. D. Thompson and D. G. Higgins (2011). "Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega." Mol Syst Biol **7**: 539.
Sourjik, V. and N. S. Wingreen (2012). "Responding to chemical gradients: bacterial chemotaxis." Curr Opin Cell Biol **24**(2): 262-268.
Spiro, R. G. (2002). "Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds." Glycobiology **12**(4): 43R-56R.
Sterba, J., M. Vancova, N. Rudenko, M. Golovchenko, T. L. Tremblay, J. F. Kelly, C. R. MacKenzie, S. M. Logan and L. Grubhoffer (2008). "Flagellin and outer surface proteins from Borrelia burgdorferi are not glycosylated." J Bacteriol **190**(7): 2619-2623.
Tan, J., X. Zhang, X. Wang, C. Xu, S. Chang, H. Wu, T. Wang, H. Liang, H. Gao, Y. Zhou and Y. Zhu (2021). "Structural basis of assembly and torque transmission of the bacterial flagellar motor." Cell **184**(10): 2665-2679 e2619.
Tang, G., L. Peng, P. R. Baldwin, D. S. Mann, W. Jiang, I. Rees and S. J. Ludtke (2007). "EMAN2: an extensible image processing suite for electron microscopy." J Struct Biol **157**(1): 38-46.
Thibault, P., S. M. Logan, J. F. Kelly, J. R. Brisson, C. P. Ewing, T. J. Trust and P. Guerry (2001). "Identification of the carbohydrate moieties and glycosylation motifs in Campylobacter jejuni flagellin." J Biol Chem **276**(37): 34862-34870.
Trueba, G. A., C. A. Bolin and R. L. Zuerner (1992). "Characterization of the periplasmic flagellum proteins of Leptospira interrogans." J Bacteriol **174**(14): 4761-4768.
Turk, M. and W. Baumeister (2020). "The promise and the challenges of cryo-electron tomography." FEBS Lett **594**(20): 3243-3261.
Wagner, T., L. Lusnig, S. Pospich, M. Stabrin, F. Schonfeld and S. Raunser (2020). "Two particle-picking procedures for filamentous proteins: SPHIRE-crYOLO filament mode and SPHIRE-STRIPER." Acta Crystallogr D Struct Biol **76**(Pt 7): 613-620.
Wagner, T. and S. Raunser (2020). "The evolution of SPHIRE-crYOLO particle picking and its application in automated cryo-EM processing workflows." Commun Biol **3**(1): 61.
Wang, F., A. M. Burrage, S. Postel, R. E. Clark, A. Orlova, E. J. Sundberg, D. B. Kearns and E. H. Egelman (2017). "A structural model of flagellar filament switching across multiple bacterial species." Nat Commun **8**(1): 960.
Wang, W., Z. Jiang, M. Westermann and L. Ping (2012). "Three mutations in Escherichia coli that generate transformable functional flagella." J Bacteriol **194**(21): 5856-5863.
Webre, D. J., P. M. Wolanin and J. B. Stock (2003). "Bacterial chemotaxis." Curr Biol **13**(2): R47-49.

Wolgemuth, C. W. (2015). "Flagellar motility of the pathogenic spirochetes." Semin Cell Dev Biol **46**: 104-112.
Wriggers, W. (2012). "Conventions and workflows for using Situs." Acta Crystallogr D Biol Crystallogr **68**(Pt 4): 344-351.
Wunder, E. A., H. Adhikarla, C. Hamond, K. A. Owers Bonner, L. Liang, C. B. Rodrigues, V. Bisht, J. E. Nally, D. P. Alt, M. G. Reis, P. J. Diggle, P. L. Felgner and A. Ko (2021). "A live attenuated-vaccine model confers cross-protective immunity against different species of the Leptospira genus." Elife **10**.
Wunder, E. A., Jr., C. P. Figueira, N. Benaroudj, B. Hu, B. A. Tong, F. Trajtenberg, J. Liu, M. G. Reis, N. W. Charon, A. Buschiazzo, M. Picardeau and A. I. Ko (2016). "A novel flagellar sheath protein, FcpA, determines filament coiling, translational motility and virulence for the Leptospira spirochete." Mol Microbiol **101**(3): 457-470.
Wunder, E. A., Jr., L. Slamti, D. N. Suwondo, K. H. Gibson, Z. Shang, C. V. Sindelar, F. Trajtenberg, A. Buschiazzo, A. I. Ko and M. Picardeau (2018). "FcpB Is a Surface Filament Protein of the Endoflagellum Required for the Motility of the Spirochete Leptospira." Front Cell Infect Microbiol **8**: 130.
Wyss, C. (1998). "Flagellins, but not endoflagellar sheath proteins, of Treponema pallidum and of pathogen-related oral spirochetes are glycosylated." Infect Immun **66**(12): 5751-5754.
Xie, Y., M. Xu, Y. Xiao, Z. Liu, C. Jiang, X. Kuang, C. Wang, H. Wu, J. Peng, C. Li, Y. Wang, H. Liu, B. Liu, X. Zhang, F. Zhao, T. Zeng, S. Liu and Y. Wu (2017). "Treponema pallidum flagellin FlaA2 induces IL-6 secretion in THP-1 cells via the Toll-like receptor 2 signaling pathway." Mol Immunol **81**: 42-51.
Xu, H., J. He, J. Liu and M. A. Motaleb (2020). "BB0326 is responsible for the formation of periplasmic flagellar collar and assembly of the stator complex in Borrelia burgdorferi." Mol Microbiol **113**(2): 418-429.
Yamaguchi, T., S. Toma, N. Terahara, T. Miyata, M. Ashihara, T. Minamino, K. Namba and T. Kato (2020). "Structural and Functional Comparison of Salmonella Flagellar Filaments Composed of FljB and FliC." Biomolecules **10**(2).
Yamashita, I., K. Hasegawa, H. Suzuki, F. Vonderviszt, Y. Mimori-Kiyosue and K. Namba (1998). "Structure and switching of bacterial flagellar filaments studied by X-ray fiber diffraction." Nat Struct Biol **5**(2): 125-132.
Yang, K. Y., C. Y. Kao, M. S. Su, S. Wang, Y. L. Chen, S. T. Hu, J. W. Chen, C. H. Teng, P. J. Tsai and J. J. Wu (2021). "Glycosyltransferase Jhp0106 (PseE) contributes to flagellin maturation in Helicobacter pylori." Helicobacter **26**(2): e12787.
Yonekura, K., S. Maki-Yonekura and K. Namba (2002). "Growth mechanism of the bacterial flagellar filament." Res Microbiol **153**(4): 191-197.
Yonekura, K., S. Maki-Yonekura and K. Namba (2003). "Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy." Nature **424**(6949): 643-650.
Yoon, S. I., O. Kurnasov, V. Natarajan, M. Hong, A. V. Gudkov, A. L. Osterman and I. A. Wilson (2012). "Structural basis of TLR5-flagellin recognition and signaling." Science **335**(6070): 859-864.
Zhang, K. (2016). "Gctf: Real-time CTF determination and correction." J Struct Biol **193**(1): 1-12.
Zhao, X., K. Zhang, T. Boquoi, B. Hu, M. A. Motaleb, K. A. Miller, M. E. James, N. W. Charon, M. D. Manson, S. J. Norris, C. Li and J. Liu (2013). "Cryoelectron tomography reveals the sequential assembly of bacterial flagella in Borrelia burgdorferi." Proc Natl Acad Sci U S A **110**(35): 14390-14395.

Zheng, S. Q., E. Palovcak, J. P. Armache, K. A. Verba, Y. Cheng and D. A. Agard (2017). "MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy." Nat Methods **14**(4): 331-332.
Zivanov, J., T. Nakane, B. O. Forsberg, D. Kimanius, W. J. Hagen, E. Lindahl and S. H. Scheres (2018). "New tools for automated high-resolution cryo-EM structure determination in RELION-3." Elife **7**.
Zivanov, J., T. Nakane and S. H. W. Scheres (2020). "Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1." IUCrJ **7**(Pt 2): 253-267.