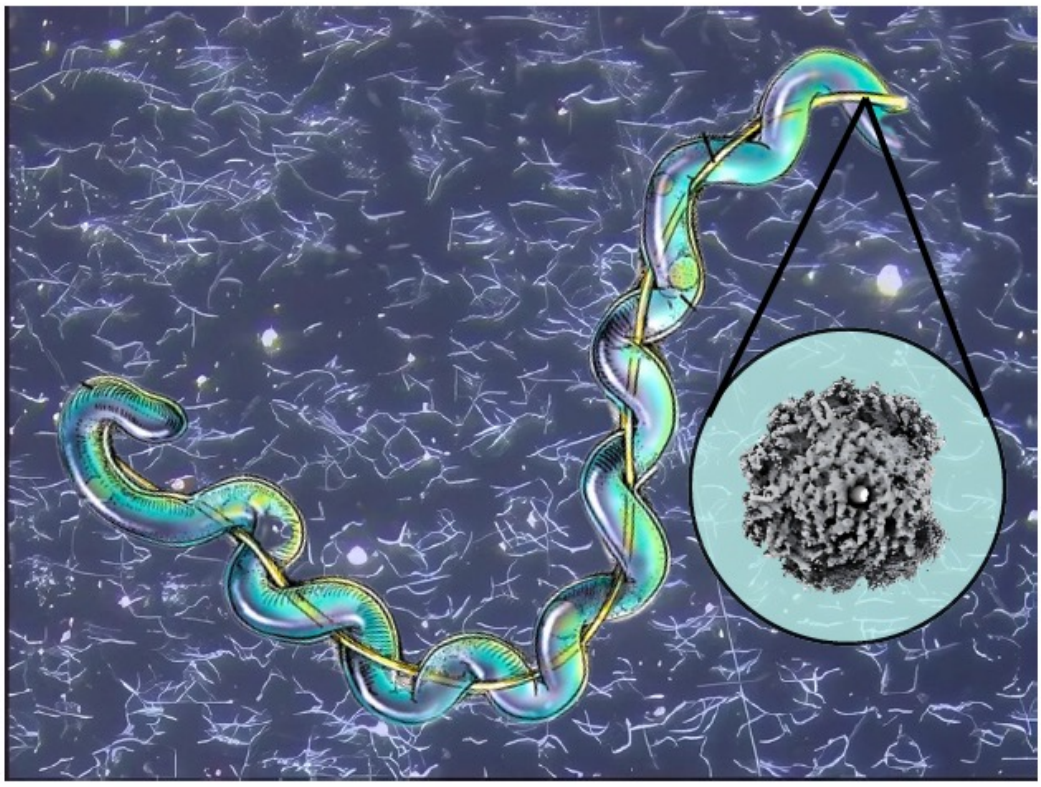
Universidad de la República
Facultad de Ciencias
PEDECIBA Biología

Aproximación estructural e integrativa al estudio del endoflagelo de Leptospira



Lic. Fabiana San Martín

Orientador: Dr. Alejandro Buschiazzo
Coorientador: Dr. Felipe Trajtenberg

Laboratorio de Microbiología Molecular y Estructural
Instituto Pasteur de Montevideo

Tribunal
Dr. Raúl Platero
Dra. Andrea Villarino
Dr. Ari Zeida

AGRADECIMIENTOS

Quisiera expresar mi más profundo agradecimiento a todas las personas que contribuyeron al desarrollo de esta tesis. En primer lugar, al Dr. Alejandro Buschiazzo, por darme la oportunidad de trabajar con él y por brindarme su apoyo a lo largo de todos estos años. Recordaré muchos momentos compartidos, pero con especial cariño aquella semana en la que trabajamos codo a codo en la mesada. También quiero agradecer enormemente a mi cotutor, el Dr. Felipe Trajtenberg, por acompañarme en mis primeros pasos en el laboratorio e impulsarme a ganar independencia. Su paciencia, motivación y experiencia fueron fundamentales en este camino.

Extiendo mi gratitud a todos nuestros colaboradores, en especial a la Dra. Rosario Durán, por los valiosos intercambios científicos y a todo su grupo, por dedicar su tiempo y esfuerzo a la recopilación y análisis de datos que formaron parte de esta tesis. Agradezco también al Dr. Charles Sindelar y a la Dra. Megan Brady, cuya investigación fue esencial para el desarrollo de nuestro modelo. Asimismo, quiero expresar mi agradecimiento a los Dres. Julia Chamot y Mathieu Picardeau, por su generosidad al recibirme en sus respectivos laboratorios. En particular, gracias a los Dres. Lucienne Noguchichian, Martial Rey y Lenka Fule, cuyo trabajo fue crucial para el avance de esta investigación.

A mis compañeros de laboratorio, les agradezco de todo corazón. Me siento profundamente afortunada por haber contado con su aliento y compañía diaria a lo largo de estos años. Por último, quiero agradecer a todas las personas que me acompañaron a lo largo de esta travesía: a mis amigos, mi esposo y nuestras familias, por celebrar conmigo los momentos de eureka y por sostenerme en los días de frustración. Su apoyo ha sido invaluable. Un agradecimiento especial a mi madre y a Virginia, mis más grandes escuchas fuera del mundo científico.

Finalmente agradezco a PEDICIBA, la ANII y la ANR por la financiación recibida.

INDICE

RESUMEN 6

I – INTRODUCCION

1.1 – Motilidad bacteriana y aparato flagelar 8

1.1.1 – Estructura general del flagelo bacteriano 9

1.1.2 – Control de la motilidad y quimiotaxis 10

1.1.3 – El filamento flagelar: estructura y composición 11

1.1.4 – Diversidad de los sistemas flagelares bacterianos 13

1.2 – Técnicas de biología estructural para el estudio del flagelo 16

1.2.1 – Cristalografía de rayos X 16

1.2.2 – Crío-microscopía electrónica de partículas individuales 17

1.2.3 – Crío-tomografía electrónica 18

1.2.4 – Integración de metodologías 19

1.3 – *Leptospira* como modelo de estudio:
características distintivas dentro de las espiroquetas 19

1.3.1 – Las espiroquetas: un filo bacteriano con morfología y motilidad únicas 19

1.3.2 – Singularidades de la estructura del flagelo en Espiroquetas 21

1.3.3 – El género *Leptospira*: características morfológicas y motilidad 24

1.3.4 – Leptospirosis: aspectos epidemiológicos y clínicos 26

1.3.5 – El endoflagelo de *Leptospira*: qué conocíamos hasta ahora? 28

II – OBJETIVOS Y ESTRATEGIA EXPERIMENTAL

2.1 – Objetivo General 31

2.2 – Objetivos Específicos de Maestría 31

2.3 – Objetivos Específicos de Doctorado 31

2.4 – Aproximación experimental 32

III – RESULTADOS

3.1 – Análisis filogenético de proteínas flagelares de Espiroquetas 33

3.2 – Caracterización estructural de componentes del filamento flagelar de *Leptospira*.37

3.2.1 – Modelado de FlaB y del núcleo del filamento flagelar 38

3.2.2 – Estructura tridimensional de las "Flagellar coiling proteins" o proteínas

inductoras del superenrrollamiento flagelar 44
3.2.3 – Estudios estructurales de FlaA1 y FlaA2 de *Leptospira* 48
3.2.3.1 – Expresión recombinante de FlaA1 y FlaA2 48
3.2.3.2 – Coexpresión de FlaA 51
3.2.3.3 – Expresión high-throughput de distintos mutantes de FlaA1 y FlaA2 ...51
3.2.3.4 – Escalado de la producción recombinante de FlaA1 y FlaA2 55
3.2.3.5 – Análisis de estabilidad de FlaA1 con nano-fluorimetría diferencial de barrido 58
3.2.3.6 – Optimización de la purificación de FlaA1 recombinante 59
3.2.3.7 – Ensayos de cristalización de FlaA1 61
3.2.3.8 – Difracción y estructura cristalográfica de FlaA1 62

3.3 – Interacción entre proteínas flagelares e identificación de nuevos componentes del filamento 65
3.3.1 – Interacción *in vitro* entre proteínas flagelares conocidas: FcpA y FcpB 65
3.3.2 – Identificación de nuevas proteínas flagelares utilizando espectrometría de masas 67
3.3.3 – Entrecruzamiento químico covalente de proteínas flagelares acoplado a espectrometría de masas (XL-MS) 70

3.4 – Enfoque integrativo para construir el modelo 3D del filamento flagelar nativo de *Leptospira* 75
3.4.1 – El filamento flagelar de *Leptospira* es asimétrico 75
3.4.2 – En el filamento flagelar de *Leptospira* se identifican nuevos componentes ..77
3.4.3 – Modelo final completo del filamento flagelar nativo de *L. biflexa* 84
3.4.4 – Estructura inicial del filamento flagelar salvaje de *L. interrogans* 97

IV – DISCUSIÓN Y PERSPECTIVAS

4.1 – Enfoque metodológico integrativo para modelar el filamento de *Leptospira* 100
4.2 – Arquitectura del filamento de *Leptospira* 101
4.3 – Núcleo del filamento 102
4.4 – Vaina del filamento 104
4.5 – Mecanismos de motilidad y origen de la asimetría 107
4.6 – Comparación con otros filamentos de espiroquetas 108

V – CONCLUSION 110

VI – MATERIALES Y METODOS

6.1 – Análisis filogenético de proteínas flagelares de Espiroquetas 112
6.2 – Análisis de características proteicas y predicción estructural 113
6.3 – Clonados y construcción de vectores 113
6.4 – Clonado *high-throughput* de isoformas de FlaA 117
6.5 – Mutagénesis sitio-dirigidas 118
6.6 – Cuantificación de ADN 118
6.7 – Secuenciación y análisis de secuencias de ADN 119
6.8 – Preparación de células competentes y transformación 119
6.9 – Expresión y purificación general de proteínas 120
6.10 – Expresión y purificación específica de proteínas individuales 121
6.11 – Expresión y purificación *high – throughput* de isoformas de FlaA 121
6.12 – Escalado de producción de isoformas de FlaA 125
6.13 – Análisis de proteínas por electroforesis (SDS-PAGE) 126
6.14 – Cuantificación de proteínas 126
6.15 – Análisis de estabilidad con nano-fluorimetría diferencial de barrido 126
6.16 – Cristalización de proteínas 127
6.17 – Difracción de rayos X y resolución de estructuras 129
6.18 – Gel filtración analítica 129
6.19 – Modelado de monómero y filamentos de FlaB 129
6.20 – Cultivos de *Leptospira* 130
6.21 – Purificación de filamentos flagelares de *L. biflexa* y *L. interrogans* 130
6.22 – Ensayos de reconstitución *in vitro* de filamentos flagelares de L. biflexa 131
6.23 – Entrecruzamiento químico de proteínas con NNP9 132
6.24 – Análisis por espectrometría de masas de péptidos entrecruzados 134
6.25 – Espectrometría de masas comparativa (MS/MS) 135
6.26 – Modelado de componentes del filamento con AlphaFold2 136
6.27 – Construcción de modelos 3D completos de filamentos flagelares salvajes de *L. biflexa* y *L. interrogans* 136

VII – BIBLIOGRAFIA 138

VIII – ANEXOS

ANEXO 1 – Tabla de Espectrometría de masas comparativa 149
ANEXO 2 – Tabla de entrecruzamientos químicos con NNP9 155
ANEXO 3 – Articulos científicos publicados durante esta tesis 162

RESUMEN

Esta tesis se centra en *Leptospira*, el agente causal de la Leptospirosis, una zoonosis de distribución global que afecta aproximadamente a un millón de personas cada año. *Leptospira* es un modelo particularmente interesante para el estudio de la motilidad bacteriana, la cual depende de endoflagelos confinados dentro del espacio periplasmico.

A diferencia de otros sistemas flagelares, cuyos filamentos suelen estar formados por un homopolimero de flagelina, los filamentos de *Leptospira* presentan una composicion inusualmente compleja. En el filo de las espiroquetas existen hasta cuatro isoformas de flagelina, denominadas FlaB, que conforman el núcleo del filamento. Este núcleo está recubierto por las proteinas de vaina, conocidas como FlaA, comunes a todas las espiroquetas. Asimismo *Leptospira* tiene un tipo de componente adicional en sus filamentos: las Fcps. Bautizadas "*Flagellar coiling proteins*" por ser responsables del pronunciado superenrrollamiento espontáneo de sus filamentos purificados *in vitro*, FcpA y FcpB son únicas del género *Leptospira* y esenciales para su virulencia.

Con el objetivo de avanzar en la comprensión la estructura flagelar y las bases moleculares de la motilidad de *Leptospira*, en esta tesis nos propusimos generar un modelo molecular de su filamento flagelar. Para ello realizamos un análisis bioinformático exhaustivo, integrando toda la información de las proteínas flagelares conocidas de espiroquetas. Trabajando con los componentes anotados del filamento modelamos por homología las estructuras 3D de las isoformas de FlaB identificando diferencias estructurales entre ellas a nivel de superficie. En pos de resolver la primer estructura cristalográfica experimental de una proteína FlaA, llevamos a cabo de forma paralela múltiples estrategias, incluyendo la coexpresión de ambas isoformas con otras proteínas de vaina y con diversas proteínas de fusión, logrando obtener varias especies de FlaA1 en forma recombinante, cristalizando y finalmente resolviendo FlaA1 de *L. borgpetersenii*. Utilizando Cristalografía de Rayos X también mejoramos la resolución obtenida para la estructura 3D de FcpA de *L. biflexa* previamente resuelta en nuestro laboratorio junto a FcpB de *L. interrogans*.

Para identificar nuevas proteínas del filamento llevamos a cabo proteómica comparativa y entrecruzamiento químico acoplado a espectrometría de masas de filamentos purificados de *L. biflexa*. Aquellas proteínas diferencialmente disminuidas o ausentes en muestras mutantes de proteínas conocidas, y/o proteínas entrecruzadas con ellas, fueron consideradas posibles candidatos a proteínas flagelares. En una aproximación estructural integrativa, sus modelos AlphaFold fueron ensamblados y refinados *in-silico,* junto a los componentes ya conocidos del filamento sobre mapas de Crio-electromicroscopia obtenidos por el grupo de nuestro colaborador, el Dr. Charles Sindelar. Construimos así un modelo completo de filamento flagelar

de *L. biflexa* que explica muy bien el mapa (FSC de 0,79) y es validado por las restricciones espaciales derivadas del entrecruzamiento químico.

El modelo final del filamento revela una arquitectura fascinante e inesperadamente compleja, compuesta por catorce identidades proteicas diferentes: FlaB1-4, FlaA2, FcpA, FcpB, y siete nuevos componentes identificados en esta tesis. Este heteropolímero mantiene características conservadas con otros filamentos bacterianos, como el núcleo helicoidalmente simétrico de proteínas homólogas a flagelina, pero se encuentra recubierto por una vaina radialmente asimétrica y con un surco lateral. Sorprendentemente, FlaA1 no aparece en el modelo final, por lo que resolvimos la estructura del filamento *wt* de *L. interrogans,* encontrando dímeros de FlaA1 posicionados sobre este surco lateral y una vaina completa que recubre todo el núcleo central de FlaBs. Nuestros resultados permiten proponer un modelo parcial para el ensamblaje secuencial del endoflagelo y el origen de su asimetría. Este proyecto abre ahora las puertas a demostrar estas hipótesis y terminar de comprender la natación de espiroquetas, que ciertamente desafían los paradigmas de locomoción bacteriana.

I – INTRODUCCION

1.1 – Motilidad bacteriana y aparato flagelar

La motilidad juega un papel determinante en la supervivencia de las especies permitiendo a los organismos vivos obtener nutrientes, escapar de toxinas o depredadores y reproducirse[1]. Esta característica definitoria de la vida fue lo que llevó a Van Leeuwenhoek en el siglo XVII a identificar por primera vez a las bacterias, "qué se movían con un propósito aparente", como organismos vivos[2]. Las bacterias han desarrollado diversos mecanismos de desplazamiento adaptados a sus necesidades ecológicas específicas. De los 18 modos de motilidad bacteriana descritos[3], se pueden destacar: (i) el "twitching" o tironeo, mediado por pili tipo IV que se extienden, adhieren a superficies y se retraen; (ii) el "gliding" o deslizamiento, que permite el movimiento sobre superficies sin apéndices externos visibles; (iii) el "swarming" o enjambre, un movimiento colectivo multicelular sobre superficies; y (iv) la natación o "swimming", impulsada por la rotación de flagelos. Este último es el el mecanismo de traslación más extendido en bacterias[4,5] estando presente en más del 70% de las especies bacterianas[6].

La motilidad flagelar es crucial para el crecimiento, supervivencia e infectividad de muchas bacterias[5,7,8]. Sin embargo, dada su diminuta escala, para ellas moverse en un medio acuoso es comparable a nadar en alquitrán[9]. Para superar este desafío han evolucionado el flagelo, un mecanismo de traslación excepcionalmente eficiente que permite a las bacterias alcanzar velocidades de entre 30 y 150 µm/s[10], lo que representa aproximadamente 100 longitudes corporales por segundo. Para contextualizar, esta capacidad equivaldría a una persona de 1,7 m nadando a más de 600 km/h[10].

El flagelo bacteriano es posiblemente uno de los orgánulos más fascinantes de la naturaleza [11], por su intrincada estructura molecular y exquisita funcionalidad que, además de proporcionar motilidad, media diversas funciones biológicas adicionales. La traslación dirigida es particularmente relevante para que los patógenos alcancen su sitio primario de infección, convirtiendo a los flagelos en importantes factores de virulencia. Durante los procesos infecciosos, los flagelos también participan en: (i) adhesión a células y tejidos del huésped, (ii) secreción de factores de virulencia y (iii) inmunomodulación y evasión de la respuesta inmune[1,10,12]. Además, la rotación flagelar constituye una forma eficiente de liberar vesículas de membrana externa[7], y su participación en la formación de biopelículas está bien documentada[9,12,13].

1.1.1 – Estructura general del flagelo bacteriano

La estructura y función del flagelo han sido extensamente estudiadas en bacterias modelo como *Escherichia coli* y *Salmonella enterica*[2,14,15]. El flagelo bacteriano es una de las nanomáquinas biológicas de mayor complejidad que se hayan caracterizado a la fecha. Su estructura básica consta de tres componentes principales (Figura 1):

1. **El cuerpo basal**: un complejo proteico de aproximadamente 8 MDa y 45 nm de diámetro[2,12] que contiene el motor rotatorio bidireccional –capaz de rotar en el sentido y en contra de las agujas del reloj–, varias estructuras de anclaje a membrana y su propio sistema de secreción flagelar[14,16]. Este complejo atraviesa tanto la membrana citoplasmática como la externa en bacterias Gram-negativas o didermas.

2. **El gancho**: una estructura curva que actúa como cardán universal, transmitiendo eficientemente la rotación del motor al filamento, independientemente de la orientación axial relativa entre ambos[17].

3. **El filamento**: una estructura muy larga que actúa como hélice[5,11,14], generando el empuje necesario para asegurar la natación de la célula bacteriana.

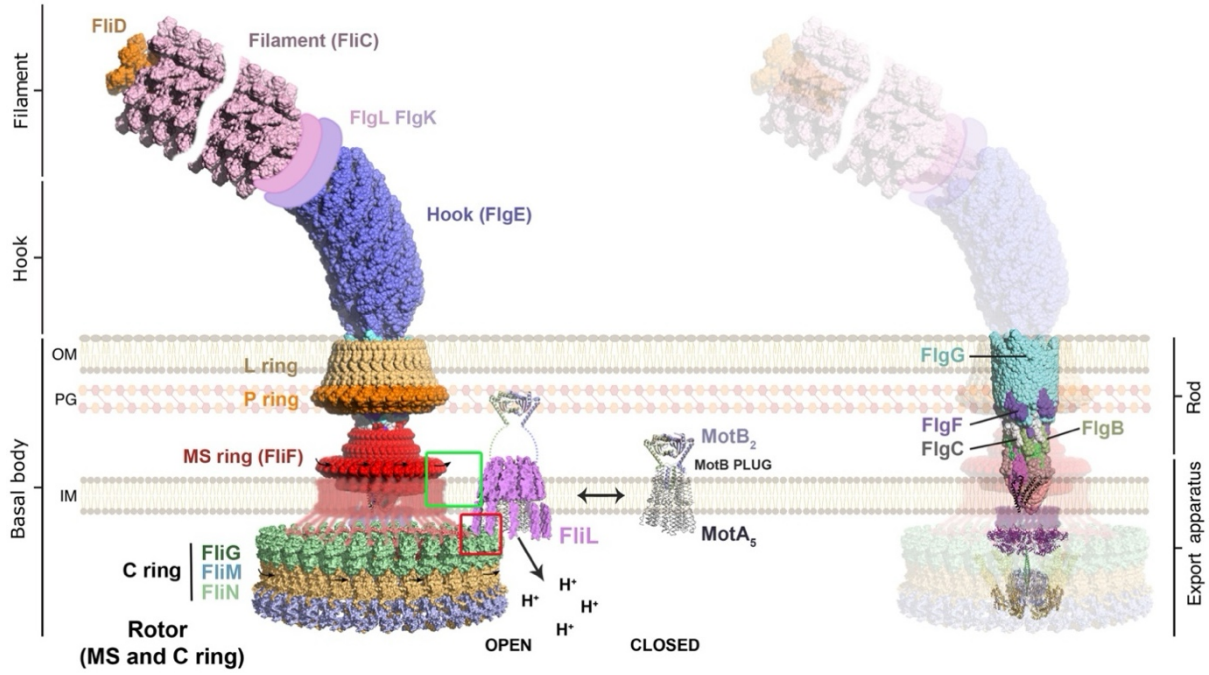


Figura 1: Arquitectura molecular del flagelo bacteriano. Representación completa del flagelo bacteriano construida a escala utilizando subconjuntos de estructuras disponibles a alta resolución para distintas especies bacterianas. El panel izquierdo muestra las superficies moleculares sólidas para la mayor parte del apéndice flagelar. El panel derecho es idéntico, excepto que los complejos proteicos más externos se muestran semitransparentes, descubriendo la composición interna del aparato de exportación y la varilla flagelar. Tomada de Partridge et al., 2024[18] que a su vez fue adaptada de Mondino et al.,2022[10].

El ensamblaje flagelar comienza con la formación del anillo MS en la membrana citoplasmática y continúa tanto hacia el interior como hacia el exterior de la célula[15]. Hacia el interior, se ensamblan el anillo C y el aparato de exportación flagelar, un sistema especializado de secreción de tipo III, que transporta a través del canal interno los componentes de la varilla, el gancho y el filamento[19]. Hacia el exterior, en las bacterias Gram-negativas, el proceso incluye la construcción de los anillos P y L, que atraviesan la pared de peptidoglicano y la membrana externa, respectivamente[9].

Todo motor rotatorio mecánico tiene un componente fijo, denominado estator, que ancla el dispositivo a un sistema de referencia estático; y un componente móvil, o rotor, que gira en relación al sistema fijo de referencia. El motor flagelar es uno de los motores biológicos más pequeños conocidos[13]; es capaz de rotar a alta velocidad en ambas direcciones, y lo hace con una eficiencia energética cercana al 100%[20]. En el flagelo, el rotor se encuentra representado por el anillo C[21], donde residuos clave en las subunidades FliG interactúan con los estatores, generando el torque que se transmite a través del eje, al gancho y el filamento[19]. Alrededor de los anillos MS y C se ensamblan lateralmente un cierto número de estatores[15], complejos proteicos individuales de dos proteínas, capaces de convertir la diferencia de potencial electroquímico a través de la membrana celular en el trabajo mecánico del rotor. Dicha diferencia de potencial deriva en general del gradiente de protones, transportados selectivamente a través de los estatores, aunque algunas especies pueden utilizar otros cationes como el Na^+.

La síntesis de un flagelo requiere de unas 30 proteínas estructurales, siendo otras 30 necesarias para su autoensamblaje y regulación[16,17,22]. Esto representa una inversión energética considerable para una bacteria, lo que refleja la importancia evolutiva de la motilidad[2].

1.1.2 – Control de la motilidad y quimiotaxis

Todas las bacterias móviles controlan de forma autónoma su locomoción, desplazándose hacia entornos favorables o alejándose de estímulos nocivos[14,15,21]. Este comportamiento, conocido como quimiotaxis, depende de sistemas quimiosensores que envían señales y se adaptan, dotando a las bacteria de una memoria temporal y una sensibilidad extraordinaria frente a gradientes químicos[5,23]. Debido a su pequeño tamaño, las bacterias son incapaces de detectar gradientes espaciales de manera directa[2]. En su lugar han desarrollado una estrategia temporal: detectan la concentración de efectores y la comparan con la registrada aproximadamente 2 segundos antes, cambiando regularmente de dirección para monitorear estos cambios de gradiente[2,14], y así nadar a favor (atractantes) o en contra (repelentes) de dicho gradiente.

Las vías de transducción de señales quimiotácticas se integran para controlar, mediante la fosforilación del regulador CheY (CheY-P), la frecuencia con la que el motor flagelar cambia de dirección[14,24,25]. El efecto final de este cambio sobre la motilidad depende del patrón de flagelación de cada especie, es decir, del número y localización de los flagelos[23,26]. Las especies bacterianas pueden ser monotricas (un solo flagelo polar o subpolar), lofótricas (múltiples flagelos en un mismo polo), anfítricas (un flagelo en cada polo) o perítricas (múltiples flagelos distribuidos por toda la superficie celular)[27,28]. El modo de motilidad más extensamente estudiado ha sido el de las bacterias perítricas como *E. coli* y *Salmonella*. La rotación de sus motores en sentido antihorario (visto desde fuera de la célula) permite que los filamentos flagelares se agrupen en fase formando un haz superhelicoidal (a modo de "cabellera") que impulsa a la bacteria en línea recta, en un movimiento denominado carrera ("*run"*)[24]. Cuando los motores giran en dirección horaria, la fuerza de torsión provoca que los flagelos se separen y se muevan independientemente unos de otros, causando que la célula gire ("*tumble*") prácticamente en su posición[24,25]. La alternancia entre carreras y giros permite a las bacterias realizar un paseo tridimensional totalmente aleatorio, que puede estar más o menos sesgado por los gradientes ambientales[5,15].

Actualmente se conocen diversos patrones de motilidad, además del modo "*run-tumble*", que pueden constar de dos o incluso de tres fases[29]. Con diferentes arquitecturas flagelares se logran movimientos como "*run-desynchronize*", "*run-pause*", "*run-reverse*" y "*run-reverse-flick*"[28,30]. Más recientemente, se ha descubierto que bajo ciertas condiciones, bacterias con distintas arquitecturas flagelares pueden realizar una maniobra denominada "*wrapping*" (envolvimiento flagelar), en la cual el flagelo se enrolla alrededor del cuerpo celular durante la locomoción[28].

1.1.3 – El filamento flagelar: estructura y composición

La subestructura más grande del flagelo es el filamento, un elemento largo y delgado que actúa como la hélice propulsora de un barco en la locomoción bacteriana[13,15]. En enterobacterias como *Salmonella* y *E. coli*, el filamento es el producto de la homo-polimerización de una única especie proteica denominada flagelina (FliC en *Salmonella*). Sus aproximadamente 30.000 protómeros[15] se autoensamblan en un proceso altamente dinámico en el que las subunidades viajan desde el citoplasma a través del canal central (~2nm de diámetro) del aparato flagelar hasta el extremo distal en crecimiento. Allí, son plegadas e incorporadas a la estructura con la ayuda del complejo homopentamérico de tapa FliD[8,12,31].

El filamento se organiza siguiendo una simetría helicoidal (~5,5Å de elevación y ~64,5° de vuelta), con típicamente 11 subunidades de flagelina por cada 2 vueltas de hélice. Esta arquitectura da lugar a un ensamblaje supramolecular que se puede visualizar como 11

protofilamentos empaquetados lateralmente, que corren a lo largo del eje mayor central hueco de la hélice[13,32]. La estructura atómica completa del filamento, resuelta mediante cryo-EM[33], revela que las subunidades de flagelina tienen forma de boomerang y se disponen en dirección radial: sus extremos N- y C-terminales forman el dominio D0 localizado en el canal central, mientras que el resto de la cadena polipeptídica se extiende hacia la superficie formando los dominios D1, D2 y D3 que se organizan como anillos concéntricos (Figura 2).

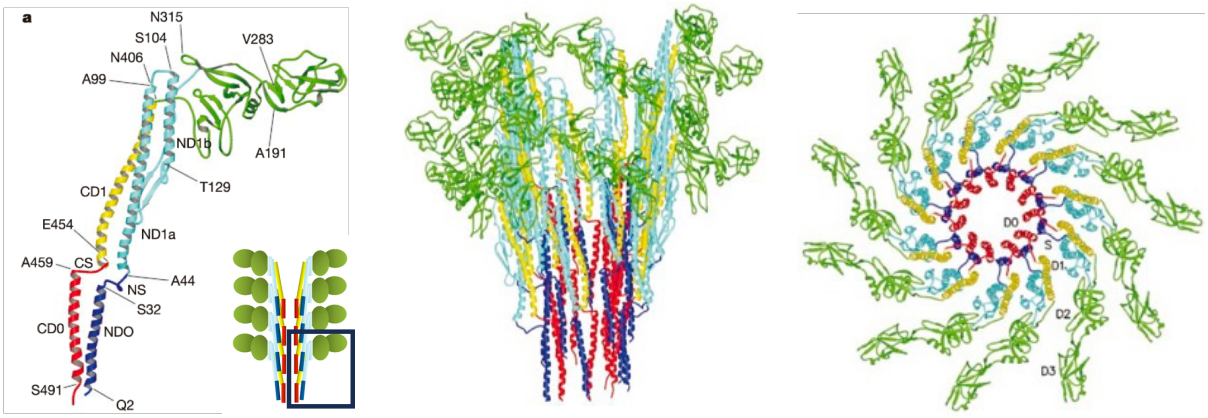


Figura 2: Estructura de la flagelina de *S. enterica* . A la izquierda diagrama estereoscópico de la cadena principal de C_α de la flagelina de *S. enterica*, coloreada con los residuos 1-44 en azul; 44-179 en cian; 179-406 en verde; 406-454 en amarillo y 454-494 en rojo. Al medio vista lateral del modelo del filamento en diagrama de cinta. A la derecha vista frontal desde el extremo distal del filamento, se muestran 11 subunidades de flagelina. Figuras tomadas de Yonekura et al., 2003[33].

Los dominios D0 y D1, compuestos por hélices alfa, constituyen un núcleo densamente empaquetado que ocupa aproximadamente la mitad del diámetro total del filamento. Las interacciones que se establecen en este núcleo son principalmente hidrofóbicas, lo que contribuye a su alta estabilidad. Por el contrario, los dominios D2 y D3, compuestos principalmente por hojas beta, se proyectan hacia el exterior y establecen pocos contactos intermoleculares, contribuyendo sólo parcialmente a la estabilidad general de la estructura[12].

Una propiedad fundamental de los filamentos flagelares es su capacidad para adoptar diferentes configuraciones helicoidales, un fenómeno conocido como polimorfismo. Estas diferentes formas helicoidales son esenciales para que las bacterias puedan cambiar de dirección durante la quimiotaxis. En *Salmonella*, cuando el motor rota en sentido antihorario, el filamento adopta una conformación denominada "normal", con una superhelicidad de~2,55 µm de paso y un diámetro de ~0,60 µm. Bajo distintas condiciones, el filamento puede cambiar a otras formas: "enrollada", "semi-enrollada", "rizada I" y "rizada II", cada una con parámetros superhelicoidales característicos[22,34].

1.1.4 – Diversidad de los sistemas flagelares bacterianos

Aunque los flagelos comparten una arquitectura general común, la visualización por cryo-ET de motores flagelares intactos en diversas especies bacterianas (Figura 3) ha revelado una notable diversidad estructural[35]. Este análisis comparativo demuestra que, si bien existe una estructura ancestral conservada, el flagelo se ha adaptado evolutivamente para funcionar óptimamente en distintos contextos filogenéticos y ecológicos[10].

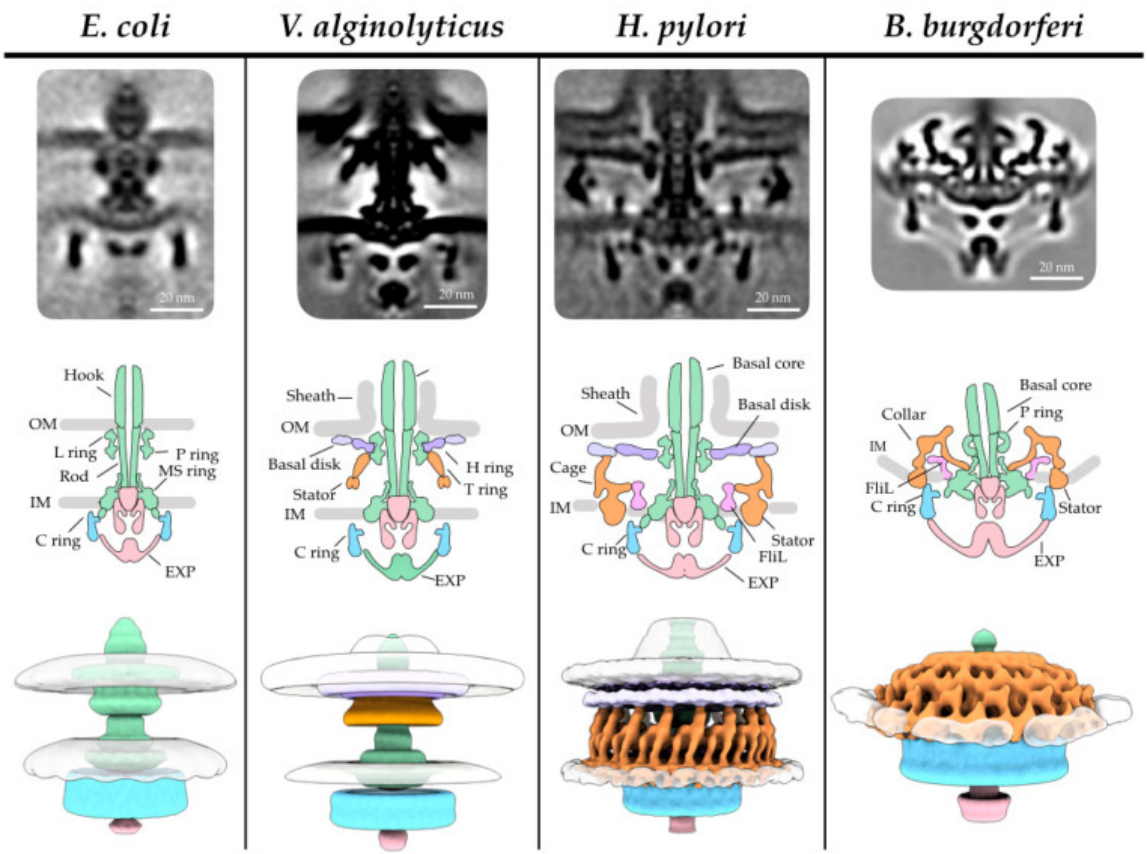


Figura 3: Comparación de la estructura de los motores flagelares de distintas especies. Para cada especie se representan, de arriba a abajo, el promediado de la clase proveniente de la CryoET depositada para el motor, el modelo gráfico extraído del mismo y una reconstrucción 3D del mapa. *E. coli*, a la izquierda, posee el motor más simple que resulta en un flagelo funcional (EMDB 5311). *Vibrio spp*. ha desarrollado anillos adicionales que aumentan la velocidad de rotación. *H. pylori* (EMDB 8459), que representa a las ε-proteobacterias, y *B. burgdorferi* (EMDB 0534), que representa a las espiroquetas, desarrollaron por separado estructuras que estabilizan los estatores y aumentan el diámetro del rotor, lo que genera un mayor torque. Figura tomada de Carrol & Liu.,2020[14].

Entre los ejemplos estructuralmente más divergentes encontramos los flagelos envainados presentes en algunas bacterias Gram-negativas con flagelación polar o lofótrica, como en varias especies de *Vibrio*, *Helicobacter* y *Brucella*. Aunque sus flagelos se extienden al espacio extracelular, están completamente rodeados por una vaina lipídica continua de membrana externa[12,14], cuya composición y función exactas siguen siendo objeto de estudio[13,36].

Una tendencia evolutiva particularmente relevante es el desarrollo de adaptaciones estructurales que aumentan el torque de los motores flagelares. El torque es una magnitud vectorial que mide la fuerza que causa o intenta causar la rotación en un objeto alrededor de un eje. Matemáticamente se define como el producto vectorial de la fuerza aplicada para el giro rotacional por la distancia del eje al punto donde se genera esta fuerza (interfaz rotor-estator)[10]. Un mayor torque no sólo permite alcanzar velocidades más altas, sino que también facilita la formación de enjambres[37] y el desplazamiento a través de medios viscosos[10].

Para lograr un mayor torque, algunas bacterias han adoptado diversas estrategias:

1. **Sistemas de flagelación duales**: Especies de géneros como *Aeromonas*, *Azospirillum* y *Vibrio* pueden cambiar condicionalmente su patrón de flagelación, típicamente de polar a perítrico o lateral, permitiéndoles generar una fuerza mayor[13,38].

2. **Estatores especializados**: Mientras que la mayoría de bacterias utilizan estatores impulsados por un gradiente de protones (PMF), algunas especies han desarrollado estatores que funcionan con iones Na^+[11,16]. Los motores impulsados por sodio pueden alcanzar aproximadamente 1300 revoluciones por segundo, en comparación con las 300 r.p.s. de los motores protónicos[2]. Algunas bacterias, como *Shewanella oneidensis* y *Bacillus subtilis*, poseen dos conjuntos de estatores que utilizan selectivamente distintos cationes según las condiciones ambientales[2,3,23,24].

3. **Proteína FliL**: Al menos 10 filos bacterianos utilizan la proteína FliL para aumentar el torque del motor[18]. Esta proteína periplásmica se ensambla como un anillo decamérico alrededor de los estatores, facilitando su reclutamiento y retención sobre el motor[37], además de modular el flujo iónico que pasa a través del estator.

4. **Rotores de mayor diámetro**: En bacterias adaptadas a medios viscosos, los motores flagelares presentan rotores de mayor diámetro y mayor simetría rotacional[14,23,39]. Este incremento en el diámetro no sólo aumenta significativamente el torque generado aumentando la distancia al eje de rotación, sino que permite incorporar más estatores cada uno de los cuales contribuye con fuerza rotacional adicional[10,16,39]. Por ejemplo, *Salmonella* incorpora 11 estatores en un rotor de 22 nm de radio, mientras que la espiroqueta *Borrelia burgdorferi* incorpora 16 estatores en un rotor de 30 nm de radio[3,16].

5. **Estructuras de anclaje adicionales**: Para estabilizar estos anillos de mayor simetría, diversas bacterias han desarrollado estructuras específicas que actúan como andamiajes, facilitando el reclutamiento y permanencia de los estatores[14,39]. En γ-

proteobacterias como *Vibrio* y *Shewanella*, estas estructuras incluyen los anillos H y T; en ε-proteobacterias como *Helicobacter* y *Campylobacter*, se observan complejos de "jaula" formados por tres discos; mientras que las espiroquetas han evolucionado una estructura alternativa denominada "collar P" (Figura 3)[14,29,39].

Los valores de torque varían considerablemente entre especies bacterianas, reflejando sus distintos estilos de vida: desde 350 pN·nm en *Caulobacter crescentus*, pasando por 2000 pN·nm en *Salmonella* y *E. coli*, hasta valores superiores a 3600 pN·nm en especies adaptadas a medios viscosos como *Helicobacter pylori*. Las espiroquetas generan los torques más elevados reportados, superando los 4000 pN·nm[10,16].

A nivel del gancho, distintos filos bacterianos también presentan variaciones adaptativas que soportan regímenes de torque más o menos elevados[10]. Los análisis *in vivo* de la rotación del gancho han demostrado que su rigidez en la flexión aumenta en respuesta a la carga del motor[40]. Las γ-proteobacterias como *V. alginolyticus*, aprovechan esta propiedad para girar rápidamente[10]. Cuando la carga del gancho supera un umbral determinado, la natación genera una compresión homogénea que provoca su flexión[41]. En las ε-proteobacterias el gancho tiene un diámetro mayor y presenta dominios de superficie adicionales, D3 y D4, en las subunidades de la proteína constituyente del gancho FlgE[42]. Además, un segmento desordenado que conecta los dominios D0 y D1 de FlgE, común a todas las bacterias[43], se estructura en *C. jejuni* como una horquilla β más larga de 17 aminoácidos, denominada "estiramiento L". Estos contactos adicionales dentro y entre los protofilamentos hacen que los ganchos sean más rígidos y resistentes en comparación con los de bacterias que usan menor torque[10,42]. Por su parte, en las espiroquetas, FlgE forma un complejo de alto peso molecular inusualmente estable. Estudios recientes explicaron que las bandas observables en SDS-PAGE[44] corresponden a monómeros de FlgE químicamente entrecruzados. Durante un proceso de autocatálisis, el residuo Lys165 en el dominio D1 se entrecruza con la Cys178 en el dominio D2 del protómero adyacente, formando un aducto lisino-alanina[43]. Esta reacción, observada en varias especies de espiroquetas[10], no es necesaria para el ensamblaje flagelar, pero sí para la motilidad celular y la infección[43]. La polimerización covalente del gancho proporciona una gran estabilidad mecánica ante sus altas demandas de torque[3,45].

En el filamento, la diversidad filogenética es aún más pronunciada. Aproximadamente el 45% de las bacterias codifican más de un gen de flagelina[46], con ejemplos extremos como *Magnetococcus*, que posee hasta 15 genes en su genoma[47]. En la mayoría de las especies, una célula generalmente expresa una o dos flagelinas simultáneamente, aunque algunas, como *Rhizobium leguminosarum*, pueden expresar hasta siete[12].

En especies como *Salmonella*, la expresión de sus dos flagelinas (FliC y FljB) está regulada mediante un proceso denominado variación de fase[48], donde FliC muestra ventaja selectiva durante la invasión, mientras que FljB se favorece en condiciones de alto torque. En bacterias como *Helicobacter pylori* y *Campylobacter jejuni*, múltiples flagelinas se ensamblan simultáneamente en el mismo filamento, pero con distribuciones espaciales específicas[12].

A nivel proteico, los dominios D0 y D1 están altamente conservados y contienen los determinantes moleculares que guian la polimerización en una estructura de 11 protofilamentos que se conserva sin excepciones[13,49,50]. Por el contrario, la región media de las secuencias polipeptídicas de las flagelina son hipervariables haciendo que los dominios externos D2 y D3 sean significativamente diversos entre especies[51]. Esta variabilidad en la superficie del filamento se refleja en pesos moleculares que oscilan entre 26 y 115 kDa [52]. Algunas bacterias presentan flagelinas con dominios adicionales, como la inserción de metaloproteasas descubierta por Eckhard et al. [53,54], o carecen completamente de dominios externos, como en *B. subtilis* [32].

Las modificaciones postraduccionales de flagelinas parecen ser comunes en muchas especies bacterianas [12]. Por ejemplo, en Campylobacter se ha descubierto que múltiples residuos de serina y treonina expuestos en la superficie se O-glicosilan con derivados del ácido pseudoamínico [55], siendo esta modificación esencial para la biosíntesis flagelar. Diversas O- y N-glicosilaciones han sido descritas en *Pseudomonas* [56], *Shewanella* [57], *Listeria* [58], *Clostridium* [59], así como en espiroquetas [60,61]. Asimismo, en bacterias como Salmonella se han decrito metilaciones de la flagelina que aumentan significativamente su hidrofobicidad de superficie, mejorando la adhesión a la célula huésped.

1.2 – Técnicas de biología estructural para el estudio del flagelo

La comprensión actual de la estructura y función del flagelo bacteriano es resultado de la aplicación complementaria de diversas técnicas de biología estructural. Cada aproximación metodológica proporciona información específica y presenta limitaciones propias, por lo que la integración de múltiples técnicas ha sido fundamental para desarrollar modelos completos de este complejo supramolecular, y ha sido el abordaje seguido en este trabajo de Tesis.

1.2.1 – Cristalografía de rayos X

Desde la determinación de la estructura de la mioglobina en 1957, la cristalografía de rayos X ha revolucionado la biología estructural, convirtiéndose en el método principal para resolver estructuras atómicas de macromoléculas [62]. Esta técnica se basa en la dispersión de rayos X por los electrones de los átomos organizados periódicamente en un cristal, generando patrones de difracción característicos [63]. En un experimento cristalográfico típico, los

cristales crioprotegidos son iluminados con rayos X de una única longitud de onda, y se rotan secuencialmente mientras se van registrando los patrones de difracción para cada arco de rotación. Las “reflexiones” (puntos de difracción) observadas se caracterizan por su amplitud y fase, aunque experimentalmente sólo pueden medirse directamente las amplitudes, directamente proporcionales a las intensidades [64]. Para resolver el "problema de las fases", se emplean distintas estrategias, siendo el reemplazo molecular (MR) la más utilizada [64,65]. Esta técnica aprovecha las fases de una estructura similar previamente determinada como punto de partida. Con amplitudes y fases para cada factor de estructura, se puede calcular un primer mapa de densidad electrónica, lo que permite construir (interpretar) un modelo atómico de la proteína. Mediante ciclos iterativos de refinamiento y reconstrucción interactiva del modelo, este se ajusta para predecir lo más certeramente posible el patrón de difracción medido [63]. Al cabo de este proceso de refinamiento, el mapa de densidad electrónica va mejorando, a medida que las fases se van corrigiendo para converger con las de la estructura cristalina problema.

La cristalografía de rayos X ha sido crucial para el estudio del flagelo bacteriano, habiendo resuelto aproximadamente 100 estructuras de sus componentes [14]. Estas determinaciones estructurales han revelado aspectos clave de su funcionamiento [14] y han permitido el diseño racional de mutantes para diversos componentes flagelares [52]. Sin embargo, la cristalografía presenta limitaciones importantes para el estudio de complejos macromoleculares grandes o flexibles, como el flagelo completo, requiriendo su complementación con otras técnicas.

1.2.2 – Crío-microscopía electrónica de partículas individuales

Desde 2012, la crío-microscopía electrónica de partículas individuales (“*cryo-EM*”) ha ganado protagonismo en la determinación estructural de proteínas y complejos macromoleculares [64,66,67]. Esta técnica implica vitrificar la muestra de proteína (típicamente por congelación brusca a temperaturas criogénicas: < -150°C) en un soporte sólido (rejilla) que contiene pequeños orificios donde el hielo vítreo pueda formar una fina película conteniendo las partículas individuales de proteína en distintas orientaciones [67]. Dicha grilla es luego introducida en el tubo de alto vacío de un microscopio electrónico de transmisión, siempre manteniendo la muestra bajo temperaturas criogénicas. Las micrografías resultantes representan proyecciones bidimensionales de objetos tridimensionales, en las que se analizan miles de partículas purificadas [14]. Dado que las moléculas en solución se orientan aleatoriamente, las imágenes contienen proyecciones de la proteína en diferentes ángulos [64]. Estas partículas son clasificadas, y se promedian todas aquellas que están en la misma orientación (clases), lo que mejora sustancialmente la relación señal-ruido. Mediante algoritmos de reconstrucción 3D, se genera un mapa tridimensional de volumen que posteriormente permite construir un modelo atómico. A diferencia de la cristalografía, la cryo-

EM no requiere la formación de cristales, lo que representa una ventaja fundamental para el estudio de complejos macromoleculares que resultan difíciles o imposibles de cristalizar.

La cryo-EM ha proporcionado información valiosa sobre numerosos componentes flagelares. Aproximadamente 50 mapas de cryo-EM de estructuras flagelares han sido depositados en bases de datos [5,12,14], revelando interacciones fundamentales entre subunidades flagelares. Por ejemplo, esta técnica ha permitido determinar la estructura del filamento completo de *Salmonella* [33], el complejo de la tapa del filamento FliD [68], y el anillo MS [69]. En los últimos años, gracias a los avances en detectores de electrones, software de procesamiento de imágenes y algoritmos de reconstrucción, la cryo-EM ha experimentado lo que se conoce como "revolución de la resolución", permitiendo obtener estructuras con resoluciones cercanas o equivalentes a las obtenidas por cristalografía de rayos X [66,67].

1.2.3 – Críо-tomografía electrónica

La críо-tomografía electrónica ("*cryo-ET*") representa otra técnica fundamental para el estudio de complejos macromoleculares en su contexto celular. Al igual que la cryo-EM, se basa en la microscopía electrónica de transmisión, pero en este caso los objetos tridimensionales se reconstruyen a partir de una serie de imágenes del mismo especimen tomadas a diferentes ángulos de inclinación [70]. El tomograma resultante (que es una representación 3D del objeto iluminado) puede segmentarse en regiones (o partículas) idénticas si estas aparecen en la célula, resultando así en la generación de una cantidad de sub-tomogramas. Dichos sub-tomogramas son luego clasificados (determinando sus orientaciones), y pueden promediarse (*subtomogram averaging* STA en inglés), mejorando enormemente el contraste y resolución finales. Una ventaja fundamental de la cryo-ET por STA es que permite visualizar macromoleculas en su entorno celular nativo sin fijación, deshidratación, seccionamiento o artefactos de tinción [14,71]. Sin embargo, presenta limitaciones en cuanto a resolución en comparación con la cristalografía y la cryo-EM, y requiere que las estructuras estén cerca de la superficie o que las células y/o muestras sean suficientemente delgadas [2].

La cryo-ET ha sido particularmente valiosa para el estudio del flagelo bacteriano, permitiendo visualizar *in situ* los flagelos de más de 20 especies bacterianas [14], incluidos representantes de casi todas las espiroquetas [71]. Esta técnica ha revelado la arquitectura completa del motor flagelar en diversas especies, incluyendo *Borrelia burgdorferi* [72,73], *Treponema primitia* [35], *Campylobacter jejuni* [39] y *Leptospira interrogans* [70]. Mediante cryo-ET se han podido identificar adaptaciones estructurales específicas que diferencian los motores flagelares de distintas especies, como el collar en espiroquetas [73] o los discos basales en γ-proteobacterias [39]. Además, ha permitido observar cambios conformacionales asociados con la rotación del motor [29] y estudiar mejor el proceso de ensamblaje flagelar [72].

1.2.4 – Integración de metodologías

Ninguna técnica de biología estructural por sí sola puede proporcionar una comprensión completa de estructuras tan complejas como el flagelo bacteriano. La integración de datos provenientes de diferentes metodologías ha sido fundamental para desarrollar modelos holísticos de este orgánulo. La cristalografía de rayos X proporciona información atómica detallada de componentes individuales, la cryo-EM permite visualizar complejos macromoleculares purificados, y la cryo-ET revela la organización de estas estructuras en su contexto celular nativo. La combinación de estas técnicas ha permitido, por ejemplo, ajustar estructuras cristalográficas o modelos obtenidos por cryo-EM en mapas de cryo-ET de menor resolución, generando modelos pseudo-atómicos de flagelos completos [5,14].

Adicionalmente, técnicas bioquímicas y de espectroscopía de masas han contribuido a identificar la composición proteica de los filamentos flagelares y las interacciones entre sus componentes [74–76]. La mutagénesis dirigida, combinada con estudios estructurales y funcionales, ha permitido validar modelos y establecer correlaciones estructura-función [52,77]. Y a su vez la aparación de métodos potentes de predicción de estructura 3D de macromoléculas a partir de su secuencia de aminoácidos [78,79] también está revolucionando la Biología Estructural moderna. Utilizando algoritmos de inteligencia artificial, y gracias a la compilación masiva de datos estructurales experimentales de estructuras proteicas, estas aproximaciones puramente computacionales resultan hoy una herramienta fundamental en enfoques integrativos. En mi Tesis utilicé extensamente AlphaFold2 para predecir estructuras de proteínas nuevas (anotadas como hipotéticas en las bases genómicas) que íbamos identificando por espectrometría de masas, para poder ajustarlas dentro de los mapas de cryoEM. Las estructuras cristalográficas que habíamos resuelto de algunos componentes del filamento también fueron fundamentales. La utilización complementaria de espectrometría de masas con y sin entrecruzamiento químico, y a partir de filamentos de cepas salvajes *vs* mutantes KO de componentes específicos, fue en suma lo que permitió la interpretación atomística de los mapas de cryo-EM.

1.3 – *Leptospira* como modelo de estudio: características distintivas dentro de las espiroquetas

1.3.1 – Las espiroquetas: un filo bacteriano con morfología y motilidad únicas

Las espiroquetas constituyen un grupo antiguo y monofilético de bacterias con considerable relevancia médica, ya que incluyen los agentes etiológicos de diversas enfermedades en humanos y animales [5,43,61]. Entre las patologías causadas por espiroquetas se encuentran la leptospirosis (*Leptospira*), la enfermedad de Lyme (*Borrelia burgdorferi*), fiebres recurrentes (otras especies de *Borrelia*), sífilis (*Treponema pallidum*), enfermedades periodontales (*T.*

denticola), gastrointestinales (*B. pilosicoli*) y disentería porcina (*Brachyspira hyodysenteriae*) [43,71,80].

Además de su importancia clínica, las espiroquetas se distinguen por su morfología y motilidad características [5,14,81,82]. Sus cuerpos celulares largos y finos presentan forma helicoidal o espiralada, y su natación se asemeja al movimiento de un sacacorchos o una serpentina plana [3,61,83]. Aunque poseen una doble membrana similar a la de las bacterias Gram-negativas [83,84], las espiroquetas difieren considerablemente de éstas debido a su divergencia evolutiva temprana [17]. En particular, su membrana externa se asocia más laxamente al cuerpo celular, y su pared de peptidoglicano es más delgada y está muy próxima a la membrana interna [71,83].

La característica más distintiva de las espiroquetas es la localización de sus flagelos. A diferencia de otras bacterias, cuyos flagelos protruyen hacia el espacio extracelular, los flagelos de las espiroquetas permanecen confinados en el espacio periplásmico, entre la pared de peptidoglicano y la membrana externa [5,11,81]. Estos flagelos periplásmicos (PFs), también denominados endoflagelos [3,61], se insertan en la membrana interna con su cuerpo basal subterminalmente anclado en los polos de la célula, y extienden sus filamentos de forma axial hacia el extremo opuesto [85,86], enrollándose alrededor del cuerpo protoplasmático en sentido dextrógiro [81].

Diferentes especies de espiroquetas varían considerablemente en cuanto al tamaño, la cantidad y la longitud de los FPs que emergen desde cada polo celular, pudiendo superponerse o no en la región central [45,85–87]. A pesar de estas diferencias, los flagelos periplásmicos mantienen similitudes estructurales y funcionales con los exoflagelos típicos [10]. La arquitectura flagelar básica se conserva [72,86,88], y aunque su ubicación es inusual, los endoflagelos preservan su capacidad de rotación y su papel fundamental para la motilidad por natación [45,82,85]. Esto ha sido demostrado mediante diversos estudios con mutantes aflagelados, que no sólo pierden la motilidad sino también su virulencia, subrayando la importancia de la motilidad flagelar en la patogenicidad de las espiroquetas [3,72,87,89–91].

Una diferencia notable entre los endo- y exo-flagelos, es la función citoesquelética que los primeros desempeñan en algunas especies [14,57,72,80,92]. En varias espiroquetas, como *B. burgdorferi* y *T. denticola*, los mutantes aflagelados presentan una morfología marcadamente diferente a las células silvestres [86,89,93]. En algunos casos, sólo se afecta la curvatura de los extremos celulares, mientras que en otros se pierde completamente la forma espiralada característica [3,81], lo que sugiere interacciones complejas entre el cilindro celular y los endoflagelos. No obstante, el filo también contiene una minoría de especies, como *Spirochaeta aurantia* y *Treponema primitia*, en las que la rotación flagelar no parece deformar

el cilindro celular [93,94]. En estas especies la morfología helicoidal dependería exclusivamente del citoesqueleto y la pared celular.

Los mecanismos que gobiernan la locomoción de las espiroquetas difieren fundamentalmente de los de bacterias con flagelos externos[84,95]. Mientras los exoflagelos funcionana como helices propulsoras cuya interacción directa con el medio circundante genera el empuje [3,11,92], en las espiroquetas la rotación de los endoflagelos deforma dinámicamente el cilindro celular, propulsando a la bacteria mediante la ondulación del cuerpo entero [3,5,86]. Para que este movimiento sea efectivo, los motores flagelares de ambos polos celulares deben rotar de manera coordinada: en direcciones opuestas durante el movimiento de traslación y en el mismo sentido durante los cambios de dirección (denominados "flexiones" en espiroquetas) [5,11,85].

Una ventaja adaptativa singular de las espiroquetas es su excepcional rendimiento en medios viscosos. Mientras que la mayoría de las bacterias ven reducida su movilidad al aumentar la viscosidad del medio, muchas espiroquetas incrementan su velocidad de natación en estas condiciones [82,86]. Esta propiedad les permite desplazarse eficientemente a través de matrices extracelulares y tejidos, facilitando su diseminación durante los procesos infecciosos [96] lo cual es una característica clínica típica en infecciones por espiroquetas, la rapidez en la diseminación y establecimiento de la infección.

1.3.2 – Singularidades de la estructura del flagelo en Espiroquetas

Los endoflagelos periplásmicos de las espiroquetas comparten los principales componentes de los flagelos externos (cuerpo basal, gancho y filamento), pero presentan adaptaciones específicas que les permiten funcionar óptimamente en su ubicación intracelular y generar el elevado torque requerido para deformar el cilindro celular [11,45].

El cuerpo basal de los flagelos de espiroquetas muestra modificaciones estructurales significativas en comparación con bacterias modelo como *E. coli* y *Salmonella* (Figura 4). La espiroqueta mejor caracterizada estructuralmente es *B. burgdorferi*, cuyos motores flagelares han sido estudiados extensivamente mediante cryo-ET [5,11,14,73]. A diferencia de los complejos de estatores, altamente dinámicos observados en otras bacterias (uniéndose y disociándose del rotor), *B. burgdorferi* posee un anillo con 16 estatores ensamblados de manera simétrica y muy estable [11,14]. Estos estatores se disponen alrededor de un rotor de mayor tamaño: mientras que en *Salmonella* el anillo C tiene una simetría de 34 y un diámetro de 45 nm, en *B. burgdorferi* la simetría es de 46 y el diámetro de aproximadamente 57 nm [11,45].

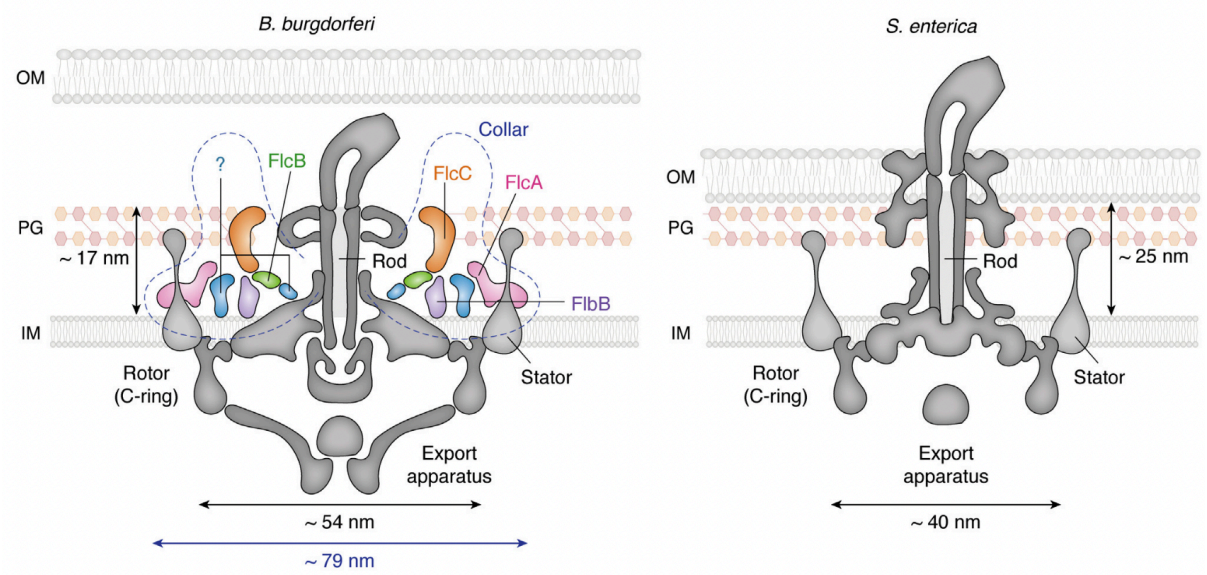


Figura 4: Comparación del cuerpo basal del endo- y exo-flagelos: Ilustraciones esquemáticas comparativas dibujadas a escala en base a los datos disponibles de crio-ET in situ. Panel izquierdo: representación completa del cuerpo basal del endoflagelo de *B. burgdorferi*, la espiroqueta mejor estudiada. Panel derecho: representación comparativa completa del cuerpo basal de *S. entérica* como modelo de bacteria exoflagelada. Nótese en el panel izquierdo la presencia del collar, un gran complejo proteico de 79nm de ancho exclusivo de los endoflagelos espiroquetales. El motor, el anillo estator y el anillo MS son comparativamente más grandes en las espiroquetas, mientras que su varilla es más corta (17nm vs 25nm en S. enterica). Tomado de Mondino et al.,2022[10].

Además, los motores de espiroquetas han desarrollado una estructura completamente nueva en la porción periplásmica del cuerpo basal, denominada "collar P" [10]. Este complejo multiproteico presenta una morfología similar a una turbina, con un diámetro de 79 nm y una altura de 20 nm [14,73]. Su arquitectura modular y flexible le permite adaptarse a la superficie curva de la célula [73]. El collar funciona como andamiaje para el reclutamiento y estabilización de los estatores, de forma análoga pero estructuralmente distinta a los discos descritos en γ- y ε-proteobacterias [39].

Otra característica distintiva de las espiroquetas es la presencia de entrecruzamientos covalentes en los ganchos flagelares que, como se mencionó antes, proporcionan estabilidad mecánica adicional para soportar las elevadas fuerzas de torsión generadas durante la rotación flagelar[10,93].

Los filamentos flagelares de las espiroquetas también difieren considerablemente de los de otras bacterias [61,87,89,97]. Mientras que la mayoría de los flagelos bacterianos consisten en homopolímeros de una única flagelina [36,77,86], los filamentos periplásmicos son heteropolímeros compuestos por múltiples proteínas [82,97]. Estos componentes se clasificaban clásicamente en dos tipos: proteínas de clase A (FlaA) y de clase B (FlaB) [36,61,82,85].

Las proteínas FlaA son específicas de espiroquetas y están presentes en una o dos isoformas según la especie [10,87,89] mostrando limitada similitud de secuencia pero clara homología. Incluyen un péptido señal N-terminal, lo que sugiere que su secreción al espacio periplásmico ocurre a través de la vía Sec[84,87,89]. En contraste, las proteínas FlaB carecen de esta secuencia escindible y se exportan al periplasma mediante el sistema de secreción flagelar[61,89,96]. Las FlaB muestran alta similitud inmunológica y de secuencia con las flagelinas convencionales, particularmente en las regiones N- y C-terminales correspondientes a los dominios D0 y D1, lo que sugiere que cumplirian una funcion similar en la polimerización del filamento flagelar [44,85,87,96].

El modelo hasta ahora más aceptado para la organización del filamento flagelar de espiroquetas proponía que las proteínas FlaB polimerizan formando un núcleo central, similar al filamento de los exoflagelos convencionales, mientras que las proteínas FlaA formarían una vaina externa recubriendo parcial o totalmente al núcleo de FlaBs[10,44,98]. Diversas evidencias experimentales respaldan este modelo incluyendo estudios de inmunomicroscopía electrónica[82] y análisis de mutantes[36,61,82]. En especies como *B. hyodysenteriae* y *T. denticola*, las FlaA determinan el diámetro del filamento, que varía entre 13 y 25 nm, con anticuerpos anti-flaA (y no anti-FlaB) marcando la superficie de los filamentos más gruesos. Asimismo, en *B. hyodysenteriae* se ha observado una vaina hueca de FlaA en ausencia de FlaB, cuya función se postula podría ser aumentar la rigidez y la fuerza del filamento[36,87]. Sin embargo, en *B. burgdorferi* y *L. interrogans*, los mutantes FlaA mantienen diámetros similares a los de las cepas silvestres[36,89], dejando claro que más invesigación es necesaria para dilucidar estos escenarios.

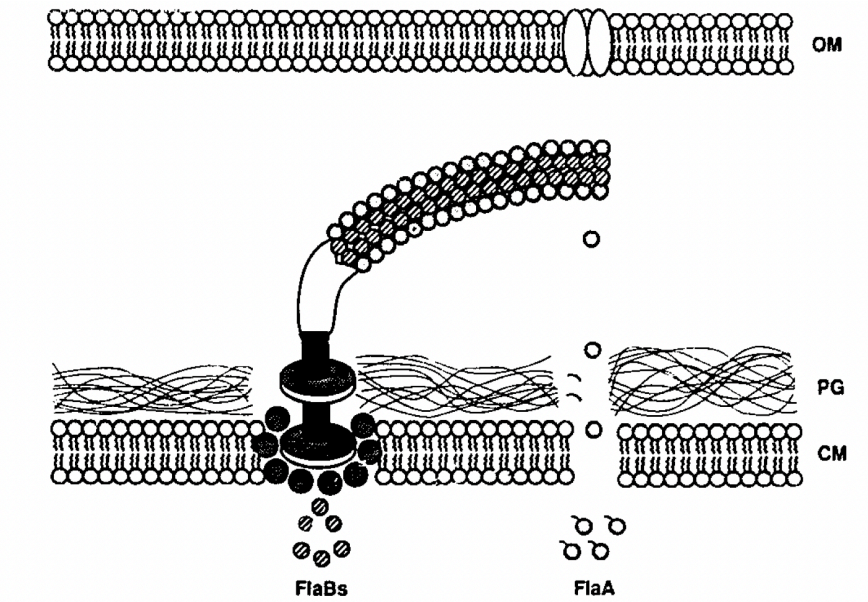


Figura 5: Modelo del endoflagelo envainado de espiroquetas: Dependiendo de la especie los PF constan de una o dos proteínas FlaA y de dos a cuatro proteínas FlaB diferentes[96], todas codificadas por genes monocistrónicos

con una regulación compleja[11,99]. Según el modelo más aceptado las proteínas FlaB son exportadas al periplasma por el aparato de secreción flagelar y polimerizan en 11 protofilamentos formando el núcleo central helicoidal del filamento flagelar. Las proteínas FlaA serían exportadas por la vía de secreción dependiente de Sec y se unirían a las proteínas FlaB formando una vaina que reviste el núcleo central del filamento. Figura tomada de Charon et al.,1992 [85]. OM= membrana externa, PG= capa de peptidoglicano, CM=membrana citoplasmática.

1.3.3 – El género *Leptospira*: características morfológicas y motilidad

Leptospira constituye un género de espiroquetas que comprende especies saprófitas de vida libre y patógenas causantes de leptospirosis[17,90,100]. Las leptospiras son bacterias delgadas, con forma helicoidal dextrógira de paso corto[3,101,102]. Tienen un diámetro de 0,1 a 0,2 µm y una longitud de 6 a 20 µm[86,94], siendo generalmente más cortas y enrolladas las especies patógenas.

La envoltura externa de *Leptospira* se asemeja más a la de bacterias Gram-negativas convencionales que la de otras espiroquetas, con abundantes proteínas intrínsecas en membrana externa[83]. Mediante cryo-ET se ha observado asimismo una estructura de membrana externa típica y relativamente ancha, con tres capas bien definidas correspondientes a las hemicapas interna y externa de la membrana y al lipopolisacárido (LPS) expuesto al medio extracelular[70,74,103]. Este LPS desempeña un papel fundamental en generar la enorme variedad de cepas de *Leptospira*. Dichas variedades, conocidas también como serovares –dado que generan respuestas serológicas bien diferentes cuando se inyectan en animales– correlacionan fuertemente con características fenotípicas importantes para las cepas patógenas, incluyendo su nivel de virulencia, preferencia de hospedero, o incluso distribuciones epidemiológicas. El lipopolisacárido de cada serovar tiene una región antigénica muy variable, y en general forma una capa más espesa en especies patógenas como *L. interrogans* (~9 nm) que en las saprófitas como *L. biflexa* (~6 nm)[70].

Otra característica estructural distintiva observada mediante cryo-ET es la presencia de una estructura en forma de "gorra o casquete" en los polos celulares, cuya función exacta se desconoce[70]. En estos polos se inserta un único filamento periplásmico que se enrolla alrededor del cilindro celular en dirección al centro de la célula[75,77,90]. A diferencia de otras espiroquetas, este endoflagelo es demasiado corto (Figura 6) para superponerse con el proveniente del extremo opuesto[3,17,102]. *Leptospira* es de este modo un género ideal para estudiar el aparato locomotor de Espiroquetas, pues exhiben la configuración celular más simple.

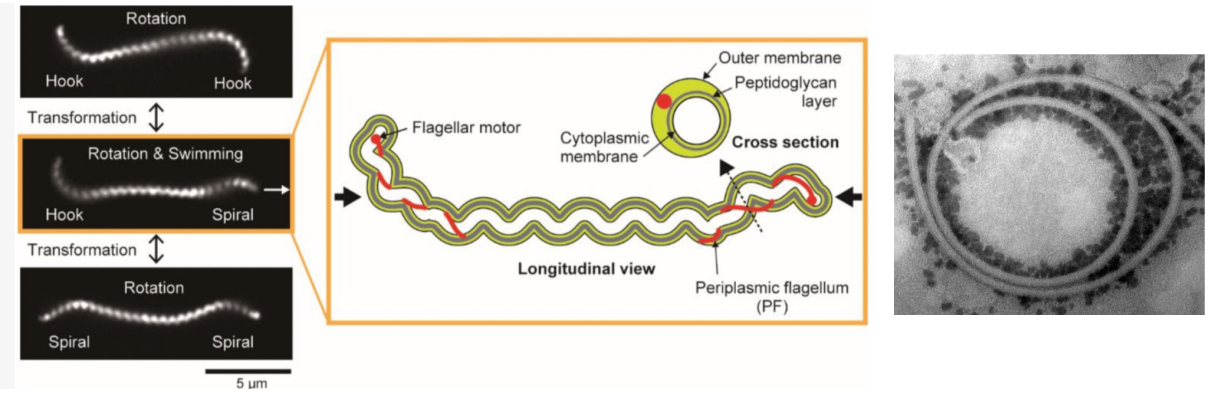


Figura 6: Singularidades de los endoflagelos de *Leptospira*. A la izquierda micrografias de campo oscuro mostrando las morfologías de los extremos celulares en cada modo de motilidad. En una célula en traslación el extremo anterior siempre se curva en forma de espiral (S) y su extremo posteiror adopta forma de gancho (H). La transformación a una morfología simétrica (S-S ó H-H) provoca que la céula interrumpa el nado pero siga rotando para cambiar de dirección. Al medio representación esquemática de *Leptospira* y su nado asimétrico en vista longitudinal y corte transversal. Los endoflagelos cortos que nacen de cada extremo celular envolviendose alrededor del cuerpo celular se muestran en rojo y la región periplásmica en verde. Las flechas negras indican la dirección de giro del extremo celular si se observan los polos celulares desde fuera de la célula. Figuras tomadas de Nakamura., 2022[104] A la derecha microscopia electrónica de transmisión (MET) de flagelos purificados de *L. interrogans*, mostrando su superenrrollamiento característico.

Una propiedad particularmente característica de los flagelos de *Leptospira* es que, al ser purificados *in vitro*, en lugar de formar ondas, se enrollan en espirales planas similares a resortes. Esta configuración superenrollada parece ser fundamental para su función motriz[105]. La motilidad de *Leptospira* presenta características únicas[101,102]. Por microscopía de campo oscuro se puede correlacionar la morfología celular con la dirección del movimiento[77]: el extremo anterior de una célula en traslación siempre presenta forma de espiral (S, *spiral*), mientras que el extremo posterior adopta forma de gancho (H, *hook*)[92,101] (Figura 6). La transformación a una morfología celular simétrica (S-S o H-H) anticipa la interrupción transitoria del nado, seguida de un cambio de dirección[71,86,100]. Es interesante notar que cuando las células están paralizadas, fijas o muertas, ambos extremos adoptan forma de gancho[86,90,92].

El mecanismo de natación de *Leptospira* fue propuesto inicialmente por Howard Berg y colaboradores, tras analizar la motilidad de *L. illini[106,107]*. Según este modelo, la rotación del endoflagelo anterior hace que ese extremo gire generando una onda espiral que se propaga hacia atrás e impulsa a la célula[89,92,101]. En medios viscosos, se añade un segundo modo de motilidad en el cual la célula gira alrededor del endoflagelo. Ambos movimientos ocurren de manera simultanea: el endoflagelo rota en una dirección y el cilindro celular en la dirección opuesta[81,101], produciendo un desplazamiento similar al de un tornillo[77,81]. El modelo predice que, dado que los cuerpos celulares son helicoidales dextrógiros, los extremos en espiral se generan cuando los motores giran en sentido antihorario (CCW, *counterclockwise*), mientras que los extremos en gancho se asocian con motores girando en sentido horario (CW, *clockwise*)[101,102,108]. Las diferentes configuraciones móviles (S-H, H-

H y S-S)[90] se explican mediante distintas combinaciones rotacionales: en la configuración asimétrica traslativa (S-H), los endoflagelos giran en direcciones opuestas, mientras que en las configuraciones simétricas (S-S y H-H), ambos motores rotan en el mismo sentido[90].

Numerosas evidencias experimentales respaldan este modelo[89,100], aunque aún se desonocen las bases moleculares que conecten la mecánica celular con la arquitectura flagelar. Estudios con células fijadas a superficies mediante partículas de látex recubiertas con anticuerpos han demostrado que el extremo en espiral efectivamente gira CCW[82,101]. Análisis más recientes utilizando microscopía de campo oscuro unilateral han correlacionado las velocidades de natación de células individuales con las velocidades de rotación de sus diferentes regiones en medios de distinta viscosidad[100]. Estos estudios revelan que tanto el extremo en espiral como el cilindro protoplasmático contribuyen significativamente a la propulsión celular, mientras que el extremo en gancho parece tener un papel más relacionado con el equilibrio y la estabilidad del movimiento[90,100].

A diferencia de lo observado en otras espiroquetas, la morfología helicoidal de *Leptospira* no depende de sus endoflagelos[70,80]. Tanto los mutantes con flagelos desenrollados como las células con membrana externa eliminada mantienen su forma helicoidal, aunque con extremos casi rectos[17,86]. Estos resultados indican que: (i) la helicidad del cuerpo celular está determinada por el citoesqueleto y/o la pared celular, y no por los flagelos; (ii) la mera presencia de los endoflagelos interactuando con el cilindro celular deforma los extremos formando ganchos[92,108]; y (iii) la rotación flagelar es necesaria para generar la morfología característicamente asimétrica (espiral/gancho) de las células móviles[70,80,86].

1.3.4 – Leptospirosis: aspectos epidemiológicos y clínicos

La leptospirosis es una zoonosis de distribución mundial que ha sido objeto de estudio desde que Adolf Weil realizó la primera descripción clínica moderna de la enfermedad en 1886[109]. Aunque la principal especie patógena en humanos es *L. interrogans[110]*, se han identificado más de 60 especies del género *Leptospira*, de las cuales aproximadamente 20 pueden causar enfermedad[109]. Estas bacterias se clasifican tradicionalmente en tres grupos (patógenas, intermedias y saprófitas) según su potencial patogénico[111–113], aunque recientemente se ha revisado esa clasificaciónen base a un aumento significativo de datos genómicos, lo que permitió extender la filogenia a dos clados –correspondientes a las antiguas especies patógenas intermedias, y a las saprófitas– cada uno sub-dividido en dos subclados[114].

La transmisión de la leptospirosis requiere una circulación enzoótica continua del patógeno entre el medio ambiente y reservorios animales[115]. Casi todos los mamíferos[111] y algunos reptiles, anfibios[116] pueden actuar como portadores. Algunos de estos animales incluidos los humanos, actúan como huéspedes accidentales que manifiestan una infección

aguda[111,115]; mientras que otros, particularmente roedores en los medios urbanos[116], desarrollan infecciones crónicas asintomáticas, volviendose reservorios que eliminan continuamente leptospiras viables a través de la orina[115,117]. A diferencia de lo que ocurre con otras espiroquetas, *Leptospira* puede sobrevivir y replicarse en el ambiente durante períodos prolongados, especialmente en suelos húmedos y agua con bajos nutrientes[118,119].

La infección humana ocurre por contacto directo con fluidos corporales de animales infectados o indirectamente a través del agua y suelos contaminados[113,120,121]. Las leptospiras ingresan al organismo a través de heridas, abrasiones cutáneas y mucosas[115,120] y se diseminan rápidamente por vía hematógena, invadiendo prácticamente todos los órganos y tejidos[122,123]. Clínicamente, la leptospirosis presenta un espectro de manifestaciones extremadamente variable. Aproximadamente el 90% de los casos son leves y autolimitados[116], con sintomatología inespecífica (fiebre, cefalea, mialgias) que frecuentemente conduce a diagnósticos erróneos[113]. Sin embargo, entre el 5-15% de los pacientes desarrollan formas graves como el síndrome de Weil (caracterizado por ictericia, insuficiencia renal y hepática, hemorragia y colapso vascular) con tasas de letalidad del 5-50%[124], o el síndrome de hemorragia pulmonar asociado a leptospirosis (LPHS), cuya letalidad supera el 50-70%[113,115,124].

A nivel global, la leptospirosis es la zoonosis más extendida causando aproximadamente un millón de casos[123,125–128] y 60.000 muertes anuales[113,125], aunque estas cifras probablemente subestiman su impacto real debido al subdiagnóstico y a deficiencias en los sistemas de notificación[113,120,128]. La incidencia es particularmente elevada en regiones tropicales y subtropicales con condiciones climáticas y socioeconómicas favorables para su transmisión[129]. Estas cifras de morbi-mortalidad ubican a la leptospirosis como una zoonosis de mayor gravedad que el dengue severo o las formas graves de leishmaniasis. Dada su extensión geográfica (no hay países libres de leptospirosis), y su asociación a aspectos ambientales que corrleacionan con el cambio global climático (como las inundaciones), la leptospirosis es una enfermdad infecciosa desatendida que se pronostica aumentará su prevalencia, y para la que no contamos con vacunas eficaces.

En Uruguay, aunque la incidencia de casos humanos es relativamente baja (~15 casos cada 100.000 hab, datos MSP), aproximadamente el 20% del ganado bovino presenta evidencia de infección[130,131]. La leptospirosis en bovinos (y en otros animales de producción como los ovinos) es causal de abortos, generando un impacto económico significativo en la industria ganadera. Además, considerando que estos animales pueden eliminar grandes cantidades de leptospiras infecciosas en la orina (~6 x 10^8 bacterias diarias por animal)[132], el riesgo de salud pública es también muy importante[123,130,133].

1.3.5 – El endoflagelo de *Leptospira*: qué conocíamos hasta ahora?

Los filamentos flagelares de *Leptospira* presentan una complejidad estructural y composición proteica notablemente más diversa que los de otras espiroquetas[74,77,103]. Al igual que en otros miembros del filo, estos filamentos comprenden un núcleo interno de aproximadamente 16 nm de diámetro, rodeado por una vaina de 21 nm, aunque esta última se pierde en algunos segmentos de los filamentos nativos[74,77,103].

Leptospira presenta la mayor cantidad de isoformas de proteínas flagelares descritas en espiroquetas, todas las cuales se expresan simultáneamente[3,134]. Los análisis de filamentos purificados mediante SDS-PAGE realizados desde finales de la década de 1960 han identificado al menos 6-7 bandas proteicas prominentes[74,75], lo que sugiere una composición más compleja que la originalmente propuesta para otras espiroquetas.

El endoflagelo de *Leptospira* contiene cuatro isoformas de proteínas FlaB (denominadas FlaB1, FlaB2, FlaB3 y FlaB4) que comparten alto grado de conservación entre sí (aproximadamente 70% de identidad) y con los dominios D0 y D1 de flagelinas de otras bacterias[77]. Estas cuatro isoformas se expresan en distintas proporciones[134], y su longitud similar (281-285 aminoácidos) sugiere que podrían tener funciones redundantes o intercambiables[77]. Sin embargo, se desconoce si estas proteínas se distribuyen homogéneamente a lo largo del filamento o se concentran en dominios específicos[82,87].

Las proteínas FlaA, por su parte, son los componentes adicionales característicos y únicos de las vainas del filamento de espiroquetas. Su estructura es desconocida y de gran interés para entender su función. Mientras que la mayoría de las bacterias del filo tienen una única isoforma, *Leptospira* tiene dos. FlaA1 y FlaA2 muestran tan solo un 28% de identidad entre sí y difieren considerablemente en longitud (315 y 239 aminoácidos, respectivamente)[77]. Esta divergencia sugiere que desempeñan funciones diferentes en el filamento. Los caracterización de mutantes realizados por Picardeau et al. en *L. interrogans*[77] han demostrado que FlaA2 tiene un papel más importante que FlaA1 en la motilidad bacteriana. Los mutantes *flaA1*$^{-}$ de *L. interrogans* presentan motilidad reducida pero conservan su virulencia y la morfología normal de sus filamentos. En contraste, los mutantes *flaA2*$^{-}$ no expresan FlaA1 (posiblemente debido a efectos polares o porque la incorporación estable de FlaA1 en el filamento requiere la presencia de FlaA2) y los dobles mutantes *flaA1*$^{-}$*/flaA2*$^{-}$ pierden completamente la motilidad traslacional, aunque mantienen la rotacional[77].

En 2016, nuestro grupo de investigación, en colaboración con los equipos dirigidos por los Dres. Albert Ko, Mathieu Picardeau y Charles Sindelar, reportó el descubrimiento de una nueva proteína en el filamento flagelar de *L. interrogans*, denominada FcpA (Flagellar coiling protein A)[105]. Esta proteína fue identificada a partir de un aislado clínico con motilidad deficiente, en el cual se observó la inserción de una timidina que introducía un cambio en el

marco de lectura truncando el producto proteico. FcpA resultó ser específica de *Leptospira*, sin ortólogos en otras espiroquetas[105,135] y constituye la decimotercer proteína más abundante del proteoma de *L. interrogans*[134].

Los estudios fenotípicos revelaron que los mutantes *fcpA*$^-$ pierden el superenrollamiento característico de los filamentos flagelares purificados[105]. Aunque estos mutantes aún ensamblan flagelos, pierden completamente la motilidad traslacional y su virulencia en modelos animales[105,136]. Microscópicamente, las células mantienen cierto movimiento rotacional pero pierden la morfología asimétrica característica de sus extremos celulares. Los filamentos purificados de cepas *fcpA*$^-$ presentan un diámetro reducido (aproximadamente 16 nm) en comparación con los de cepas salvajes (21 nm)[105], lo que sugiere que FcpA forma parte integral de la vaina flagelar.

Posteriormente, mediante rastreo de mutantes aleatorios, se identificó una segunda proteína flagelar exclusiva de *Leptospira*, denominada FcpB (Flagellar coiling protein B)[137]. Los mutantes knockout *fcpB*$^-$ presentan un fenotipo similar, aunque menos severo que el de las cepas *fcpA*$^-$: pérdida parcial del superenrollamiento de los filamentos y de la deformación característica de los extremos celulares, pero conservan cierta capacidad de motilidad traslacional[137]. El análisis de las trayectorias de natación de leptospiras individuales revela que estos mutantes pierden específicamente la subpoblación celular que nada a máxima velocidad[137].

Es importante destacar que, aunque FcpA y FcpB determinan fenotipos similares, carecen de homología de secuencia significativa entre sí y con otras proteínas conocidas. Ambas son exclusivas de *Leptospira* y poseen péptidos señal N-terminales que sugieren su translocación al periplasma a través de la vía Sec, lo que apoya su potencial localización en la vaina flagelar. Esta hipótesis ha sido corroborada mediante experimentos de inmunomicroscopía electrónica, donde tanto los anticuerpos anti-FcpA como anti-FcpB marcan específicamente la superficie del filamento flagelar, particularmente en el lado convexo de la curvatura[105,137].

Estudios adicionales realizados por Sasaki et al., 2018[76] mediante ensayos bacterianos de doble híbrido y *pull-down* han demostrado que FcpA interactúa directamente con FlaA2 y FlaB1, proporcionando evidencia bioquímica de su participación en la estructura de la vaina flagelar. En conjunto, estos hallazgos sugieren que el endoflagelo de *Leptospira* es más complejo que el propuesto originalmente para otras espiroquetas. Este modelo incluiría un núcleo central formado por las cuatro isoformas de FlaB, rodeado por una vaina compuesta por FlaA1, FlaA2, FcpA y FcpB. La organización precisa de estos componentes dentro de la estructura filamentar, sus interacciones moleculares y sus contribuciones específicas a las propiedades funcionales del flagelo constituyen aspectos fundamentales que esta tesis

pretende dilucidar. La comprensión detallada de la estructura del endoflagelo de *Leptospira* permitiría abordar preguntas fundamentales sobre su biología, como ser: ¿Por qué las espiroquetas, y en particular *Leptospira* han desarrollado una composición flagelar más compleja? ¿Qué interacciones moleculares específicas determinan el superenrollamiento característico del flagelo? ¿Cómo contribuyen las distintas proteínas flagelares a la motilidad bacteriana y, consecuentemente, a la virulencia? ¿De qué manera interactúan estos apéndices con el cuerpo celular para generar un movimiento traslacional tan eficiente?

Responder estas incógnitas permitiría avanzar no solo en el conocimiento de la biología básica de estas bacterias, sino también en el desarrollo de nuevas estrategias terapéuticas y preventivas para combatirlas.

II – OBJETIVOS Y ESTRATEGIA EXPERIMENTAL

2.1 – Objetivo General:

El objetivo central de este proyecto fue comprender con detalle molecular la constitución del filamento flagelar de *Leptospira* para entender su rol funcional en la locomoción de Espiroquetas.

Esta tesis constituye una Maestría con pasaje a Doctorado, por lo que a continuación se resumen los objetivos específicos de cada etapa.

2.2 – Objetivos Específicos de Maestría:

1 - Expresión, purificación y determinación de la estructura 3D de proteínas flagelares de especies saprófitas y patógenas de *Leptospira* (FcpA, FcpB, FlaA1 y FlaA2) por Cristalografia de Rayos X.

2 - Estudios estructurales bioinformáticos de los componentes más conservados del filamento de *Leptospira*. Modelado de las 4 isoformas de FlaB de *L. interrogans*, y estudio de posibles superficies de interacción con el resto de los componentes flagelares.

3 - Evaluación de la interacción entre distintos componentes del filamento flagelar de *Leptospira*.

2.3 – Objetivos Específicos de Doctorado:

4 - Determinar la estructura tridimensional de las isoformas de FlaA de *Leptospira*. (Continuación del objetivo 1 de Maestría)

5 - Determinar el proteoma de filamentos flagelares de *L. biflexa*.

6 - Determinar contactos intermoleculares del filamento flagelar de *Leptospira* por entrecruzamiento químico y espectrometría de masas.

7 - Reconstituir el filamento flagelar de *Leptospira in vitro* para dilucidar interacciones intermoleculares del ensamblaje.

8 - Generar un modelo integrativo de la arquitectura del filamento flagelar de *Leptospira*.

2.4 – Aproximación experimental

Empleamos una estrategia experimental multidisciplinaria que combinó:

a) La obtención de las estructuras 3D de los componentes individuales ya conocidos de *Leptospira*, es decir: FlaBs, FlaAs, FcpA y FcpB (objetivos específicos 1, 2 y 4). Si bien trabaje con todos los componentes previamente anotados, en particular mi contribución aquí fue el modelado por homología de las cuatro isoformas de FlaB de *L. interrogans*, la mejora en la resolución de la estructura 3D de FcpA de *L. biflexa* y la resolución de la estructura cristalográfica de FlaA1 *L. borgpetersenii.*

b) Análisis bioinformático exhaustivo de genes/proteínas flagelares en Espiroquetas: búsqueda de homólogos lejanos, alineamiento múltiple de secuencias con curado manual. En colaboración con Gregorio Iraola, estos alineamientos permitieron realizar un análisis filogenético e identificación de genes específicos de *Leptospira*.

c) La identificación de nuevas proteínas flagelares previamente no anotadas mediante el análisis proteómico de filamentos flagelares aislados, tanto nativos como sometidos a entrecruzamiento químico (objetivos especificos 5 y 6).

d) La identificación de interacciones entre proteínas flagelares, analizando la interacción entre los componentes FcpA y FcpB (objetivos específicos 3 y 7) y en una aproximación más global empleando entrecruzamiento químico covalente acoplado a espectrometría de masas (XL-MS) para analizar las interacciones en proteoma del filamento entero (objetivo especifico 6).

e) La interpretación de mapas de cryo-EM obtenidos por nuestros colaboradores (Laboratorio de Charles Sindelar en la Univ de Yale). Utilizando un enfoque integrativo analizamos la estructura del filamento flagelar completo, ubicando y ajustando las estructuras cristalográficas, así como otras predichas con algoritmos de inteligencia artificial a partir de las secuencias, para lograr un modelo completo de los filamentos de cepas salvajes de *L. biflexa* y de *L. interrogans* (objetivo especifico 8). En particular mi contribución aquí fue en las etapas finales de la generación de los modelos, identificando y posicionando FlaA1 y los nuevos componentes identificados, así como refinando y validando el modelo final 3D.

III – RESULTADOS

3.1 – Análisis filogenético de proteínas flagelares de Espiroquetas

Un primer análisis bioinformático minucioso de las secuencias de proteínas flagelares anotadas de *Leptospira*, fue útil para informarnos sobre la variabilidad de secuencia (global y por dominios) en vistas a construcciones recombinantes y posterior estudio estructural de componentes individuales. Este abordaje, en colaboración con el Laboratorio de Genómica Microbiana del Instituto Pasteur de Montevideo, se fue enriqueciendo y derivó en un estudio genómico comparativo exhaustivo[138]. El mismo permitió contextualizar nuestros hallazgos estructurales (ver más abajo) en el marco de la evolución del sistema endoflagelar en Espiroquetas, incluyendo el conocimiento actual sobre otros componentes flagelares del Phylum.

Aplicando herramientas bioinformáticas modernas, identificamos secuencias homólogas de 59 componentes flagelares conocidos, en 864 genomas representativos, correspondientes a 115 especies de Espiroquetas distribuidas en 9 géneros. Estas secuencias fueron alineadas utilizando el algoritmo MUSCLE[139], generando alineamientos múltiples para cada proteína. Los alineamientos fueron posteriormente curados de forma manual, eliminando regiones con baja cobertura, corrigiendo errores de alineamiento en regiones conservadas y verificando la correcta asignación de dominios funcionales mediante análisis de conservación de residuos catalíticos y estructurales.

A partir de estos alineamientos curados, calculamos un mapa de calor (*heat map*) (Figura 3.1) que ilustra la presencia/ausencia y el grado de similitud de las diferentes proteínas flagelares entre las especies analizadas.

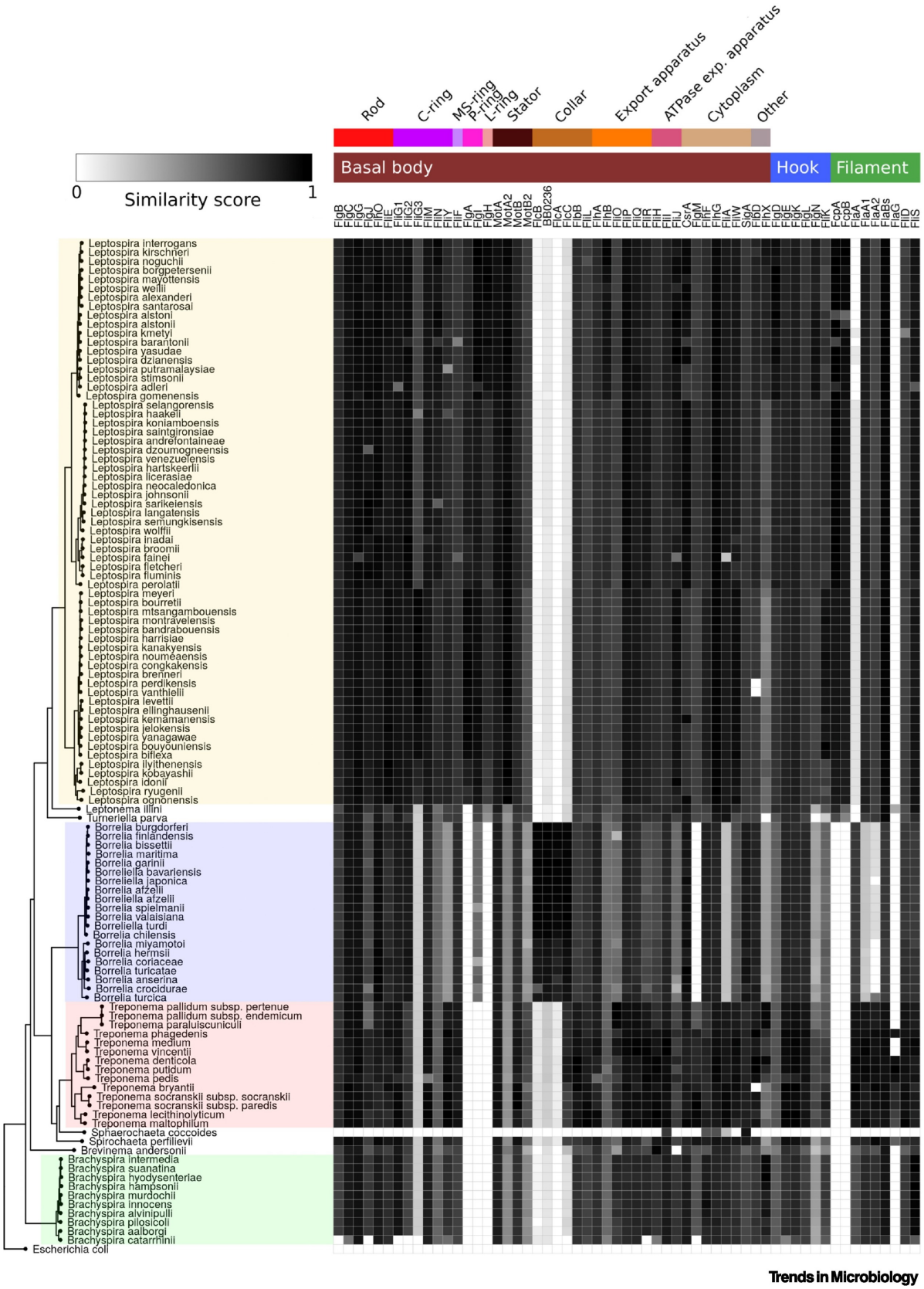


Figura 7: Mapa de calor mostrando la conservación de proteínas flagelares en el Phylum Espiroquetas. Las columnas corresponden a genes codificantes de proteínas flagelares y las filas representan diferentes especies del phylum. La intensidad del color indica el grado de similitud de secuencias, desde ausencia (blanco) hasta alta conservación (negro). Para su construcción, realizamos alineamientos de secuencias aminoacídicas de 59 proteínas flagelares conocidas, utilizando 971 genomas disponibles del Phylum. Los valores de bitscore

resultantes fueron normalizados en escala de 0 a 1 respecto al valor máximo obtenido para cada proteína. Para contextualizar evolutivamente, se construyó un árbol filogenético de 864 especies de Espiroquetas y una de *E. coli* como grupo externo. El árbol fue colapsado a nivel de especies, resultando en 115 representantes de los géneros *Leptospira* (n=63), *Leptonema* (n=1), *Turneriella* (n=1), *Borrelia* (n=21), *Treponema* (n=15), *Sphaerochaeta* (n=1), *Spirochaeta* (n=1), *Brevinema* (n=1) y *Brachyspira* (n=10). Para las especies colapsadas, se muestra el valor medio normalizado de bitscore para cada proteína.

El patrón filogenético resultante concuerda notablemente con árboles filogenéticos establecidos para Espiroquetas [140,141], validando nuestro enfoque. Aproximadamente el 66% de las proteínas flagelares están conservadas en todo el phylum, aunque también identificamos componentes específicos de género que revelan particularidades evolutivas y funcionales.

Entre los hallazgos más relevantes del análisis comparativo, destacan:

a) Componentes del aparato de exportación y cuerpo basal: El aparato de exportación en el cuerpo basal está altamente conservado tanto en Espiroquetas como en otras bacterias flageladas. Sin embargo, identificamos que las Espiroquetas poseen un componente adicional exclusivo (FlhX), ausente únicamente en *Borrelia burgdorferi*. FlhX presenta clara homología con el dominio C-terminal de FlhB, una proteína involucrada en el reconocimiento de sustratos y el cambio de especificidad durante el ensamblaje flagelar, sugiriendo una función similar para FlhX en Espiroquetas. Mediante crio-electrotomografía se han descripto diferencias estructurales significativas en *Leptospira*[70], donde el canal transmembrana es considerablemente más grande, evidenciando adaptaciones específicas de este género.

b) Estructuras anulares del cuerpo basal: Las Espiroquetas presentan diferencias significativas en la composición de sus anillos flagelares. *Leptospira*, *Turneriella* y *Leptonema* poseen los cuatro anillos presentes en bacterias Gram-negativas: MS (asociada a membrana plasmática), P (peptidoglicano), L (membrana externa) y C (citoplásmico). En contraste, *Borrelia* spp. carece de FlgH (formadora del anillo L), mientras que *Treponema*, *Brevinema*, *Brachyspira* y *Spirochaeta* carecen también de FlgI (componente del anillo P). *Leptospira* destaca por su complejidad, siendo la única Espiroqueta que posee la chaperona periplásmica FlgA (necesaria para el ensamblaje del anillo P), una proteína FlgI más extensa, y un anillo C más complejo compuesto por FlgG, FliM y ambas proteínas FliN y FliY. La presencia simultánea de FliN y FliY es atípica, observada únicamente en *Leptospira*, *Helicobacter* y *Campylobacter*, donde posiblemente desempeñen funciones complementarias no redundantes[142].

c) Estatores y sistemas de generación de torque: Las Espiroquetas contienen hasta 16 estatores (heterooligómeros de MotA y MotB) ensamblados alrededor del rotor, comparado con los 12 habituales en otras bacterias[143]. Este mayor número correlaciona con un diámetro motor ampliado[95] y la generación de mayor torque [3], adaptaciones

fundamentales para la motilidad endoflagelar. Mientras la mayoría de bacterias exoflageladas poseen una única isoforma de FliG (proteína que interactúa con los estatores para generar torque), las Espiroquetas tienen dos parálogos, y *Leptospira* se distingue con tres parálogos de FliG y dos de cada proteína Mot, evidenciando adaptaciones específicas de su sistema de generación de fuerza motriz.

d) Estructura del collar: El collar, estructura característica de las Espiroquetas, se ancla a la membrana citoplasmática y al anillo MS, interactuando con estatores y rotor mediante FliL. Su composición varía considerablemente entre géneros. De los cinco componentes identificados en *B. burgdorferi* (FlbB, BB0236, FlcA, FlcB y FlcC), únicamente FlbB está presente en todas las Espiroquetas. FlcA y FlcB se encuentran solo en *Borrelia*; BB0236 en *Borrelia* y *Treponema*; y FlcC en *Leptospira*, *Borrelia* y *Treponema*. Esta variabilidad sugiere diferencias estructurales significativas en el collar entre los distintos géneros de Espiroquetas.

e) Eje del cuerpo basal y gancho flagelar: El eje del cuerpo basal en Espiroquetas consta de seis proteínas conservadas: FliE, FlgB, FlgC, FlgF (denominada FlhO en Espiroquetas), FlgG y FlgJ. Notablemente, FlgJ carece del dominio C-terminal muramidasa presente en otras bacterias, sugiriendo que esta función de remodelación del peptidoglicano está codificada en otros genes. En cuanto al gancho, las proteínas FlgD (capuchón), FlgE (componente principal) y FliK (reguladora de longitud) están conservadas. Sin embargo, las Espiroquetas presentan la particularidad del auto-entrecruzamiento covalente de las subunidades de FlgE[43,144], y en *B. burgdorferi* esta proteína también participa en la regulación postranscripcional de FlaA y FlaB[145]. *Leptospira* muestra una sección basal del gancho más ancha (aproximadamente 21 nm de diámetro) que otras Espiroquetas como *B. burgdorferi* (aproximadamente 16 nm)[70], y su FlgD posee una secuencia extendida con un extremo C-terminal no conservado.

f) Componentes del filamento flagelar: Las Espiroquetas conservan los componentes básicos del filamento presentes en bacterias modelo: FlgK y FlgL (unión gancho-filamento); FliD (capuchón distal); y, luego múltiples subunidades de flagelina polimerizadas (formando el filamento propiamente dicho). Sin embargo, las flagelinas de Espiroquetas carecen de los dominios D2 y D3 menos conservados de FliC de *Salmonella* [33], están glicosiladas, y presentes en número variable de isoformas según el género: *Borrelia* spp. tiene una sola especie de FlaB, *Treponema* spp. y *Brachyspira hyodysenteriae* tienen 3, y *Leptospira* spp. posee 4 isoformas FlaB1, FlaB2, FlaB3 y FlaB4[17]. Rodeando el núcleo de FlaB se encuentra una vaina compuesta por FlaA: una isoforma única en *Borrelia* y *Brachyspira*, y dos en *Leptospira* y *Treponema* [44]. Nuestro análisis reveló que *Brevinema*, *Brachyspira*, *Treponema* y *Spirochaeta* poseen una isoforma adicional de FlaA, lo cual también es visible en *Leptospira*: un tercer gen parálogo a FlaA1 y FlaA2, que no había sido anotado previamente. Adicionalmente, nuestro grupo y colaboradores han descubierto los componentes FcpA y FcpB en *Leptospira*[105,137], también presentes en *Turneriella* y *Leptonema*. Por su parte,

Treponema spp. es la única Espiroqueta con un homólogo de FlaG, que cumple una función diferente a sus parálogos en *C. jejuni*, *P. aeruginosa* y *P. fluorescens*, afectando la motilidad celular y los niveles de FlaA y FlaB[146].

g) Regulación de la biogénesis flagelar: Los genes de clase I y la mayoría de los genes de clase II están presentes en todas las Espiroquetas analizadas[16,97]. Sorprendentemente, *B. burgdorferi* posee un único factor sigma (sigma70/SigA/RpoD) para transcribir los genes flagelares de clase III[45], mientras que las demás Espiroquetas expresan también el factor sigma28 (FliA/RpoF) y su antisigma FlgM[97]. Las Espiroquetas presentan una regulación postranscripcional más rica, con proteínas como FliW y CsrA que en *Leptospira* y *B. burgdorferi* regulan la expresión de FlaB[147,148]. Asimismo, FlhF y FlhG participan en la regulación transcripcional y postranscripcional en *L. biflexa* y *B. burgdorferi*, respectivamente[149,150], controlando la distribución espacial y el número de flagelos. En cuanto a chaperonas, FliS parece ser la única conservada en Espiroquetas, posiblemente actuando como chaperona general con amplia especificidad de sustratos. *Leptospira* es el único género que contiene la chaperona FlgN, presente en otras bacterias.

Este análisis comparativo nos permitió, además, identificar y corregir inconsistencias en la anotación automática de genes flagelares. Detectamos casos de asignación errónea, como FlgC y FlhO frecuentemente anotadas como FlgG; varias FlgL anotadas como FlaB o como proteínas hipotéticas; y confusiones entre proteínas con cierto grado de similitud (FlhX con FlhB, FliY con FliN). Particularmente importante para nuestro proyecto fue corregir la anotación de las variantes parálogas de FlaA y FlaB en distintas especies de Espiroquetas, donde en algunos casos se habían asignado nombres idénticos a proteínas no ortólogas[138]. Esta estandarización de la nomenclatura resulta indispensable para facilitar comparaciones entre especies, realizar búsquedas precisas en bases de datos y establecer una comunicación efectiva entre grupos de investigación.

La particular complejidad del sistema flagelar de *Leptospira*, evidenciada por la presencia de componentes exclusivos o adaptaciones específicas en prácticamente todos los niveles estructurales del flagelo (cuerpo basal, gancho y filamento), justifica el enfoque molecular detallado que adoptamos en este trabajo. Este análisis bioinformático representó no solo un primer resultado fundamental, sino también una herramienta esencial para interpretar los hallazgos estructurales subsiguientes en un contexto evolutivo más amplio.

3.2 - Caracterización estructural de componentes del filamento flagelar de *Leptospira*

Tras el análisis filogenético y la identificación precisa de los componentes flagelares en *Leptospira*, procedimos a la caracterización estructural de proteínas clave del filamento. Esta sección describe los resultados obtenidos para FcpA, FcpB, FlaA1 y FlaA2, componentes

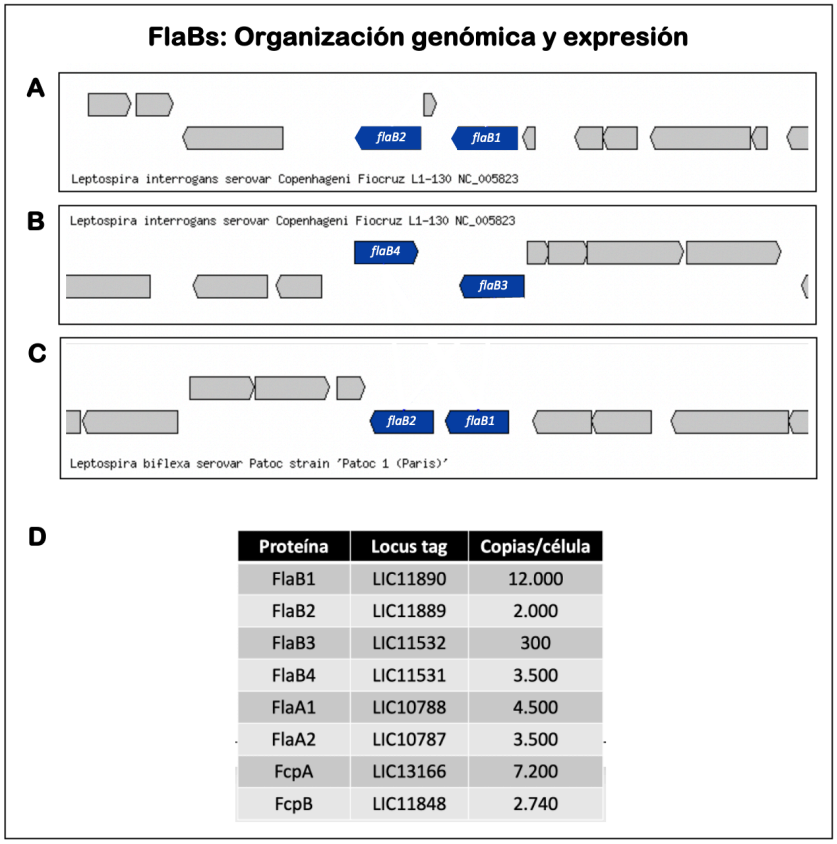
principales de la vaina del filamento, cuya caracterización tridimensional resultaba fundamental para comprender la arquitectura del filamento.

3.2.1 – Modelado de FlaB y del núcleo del filamento flagelar

Las proteínas FlaB son ortólogas a la flagelina que conforma el filamento flagelar en bacterias modelo como *Salmonella* y *Bacillus*. Al carecer de péptido señal para la vía clásica de secreción, son exportadas al periplasma a través del aparato de secreción flagelar homopolimerizando *in situ* durante el crecimiento del filamento. Esta homología con la flagelina sustenta que FlaB constituye la especie proteica del núcleo de polimerización del filamento flagelar en Espiroquetas, consistente también con observaciones reportadas de que algunos mutantes knockout de *flaB* parecen perder completamente su filamento flagelar*[80]*.

Leptospira posee cuatro isoformas de FlaB (FlaB1-4) con identidad de secuencia en el rango de 45-72%. Las isoformas FlaB1 y FlaB4 presentan la mayor similitud de secuencia y parecen compartir un ancestro común, mientras que FlaB3 es filogenéticamente la más distante con homología menor al 50% con las otras isoformas. En cuanto a su organización genética, *flaB1* y *flaB2* se encuentran en un mismo locus separados por aproximadamente 450 nucleótidos en *L. interrogans* y 200 nucleótidos en *L. biflexa*, sugiriendo un evento de duplicación génica y regulación coordinada. Las isoformas FlaB3 y FlaB4 son codificadas por genes ubicados en loci génicos contiguos en *L. interrogans*, aunque en sitios distintos del genoma en *L. biflexa*, indicando cierta variabilidad entre especies saprófitas y patógenas.

Las cuantificaciones absolutas de proteína en *L. interrogans* (Malmström et al., 2009) revelan que FlaB1 es la isoforma más expresada, seguida de FlaB4, FlaB2 y FlaB3, esta última con solo ~300 copias por célula (Figura 8).



Proteína	Locus tag	Copias/célula
FlaB1	LIC11890	12.000
FlaB2	LIC11889	2.000
FlaB3	LIC11532	300
FlaB4	LIC11531	3.500
FlaA1	LIC10788	4.500
FlaA2	LIC10787	3.500
FcpA	LIC13166	7.200
FcpB	LIC11848	2.740

Figura 8: Organización genómica y expresión de FlaB. A) Locus génico de *flaB1* y *flaB2* de *L. interrogans* serovar Copenhageni. **B)** Locus génico de *flaB3* y *flaB4* de *L. interrogans*. **C)** Locus génico de *flaB1* y *flaB2* de *L. biflexa* serovar Patoc. **D)** Cuantificación absoluta del número de copias por célula para las proteínas del filamento flagelar de *L. interrogans* (datos de Malmström et al., 2009[134]).

Esta expresión diferencial y los diferentes grados de homología plantean interrogantes sobre la redundancia estructural y funcional entre las distintas isoformas. Para investigar si las diferencias de secuencia de FlaB se traducen en diferencias estructurales que generen propiedades de interacción y autoensamblaje diferenciales, realizamos modelado molecular de las cuatro isoformas de FlaB de *L. interrogans*. Dada que la tendencia natural de estas proteínas tipo flagelina a polimerizar formando filamentos muy largos, dificulta el estudio estructural del monómero en solución, adoptamos una aproximación bioinformática para predecir la estructura tridimensional utilizando modelización por homología.

El alineamiento de secuencias de las cuatro isoformas de FlaB con la flagelina FliC de *S. enterica* reveló alta conservación en los dominios D0 y D1, mientras que los dominios D2 y D3 están ausentes (Figura 9). Los alineamientos de mejor score y cobertura se obtuvieron con la estructura PDB ID 1UCU de FliC de *S. enterica*, resuelta por crio-microscopía electrónica a 4 Å de resolución.

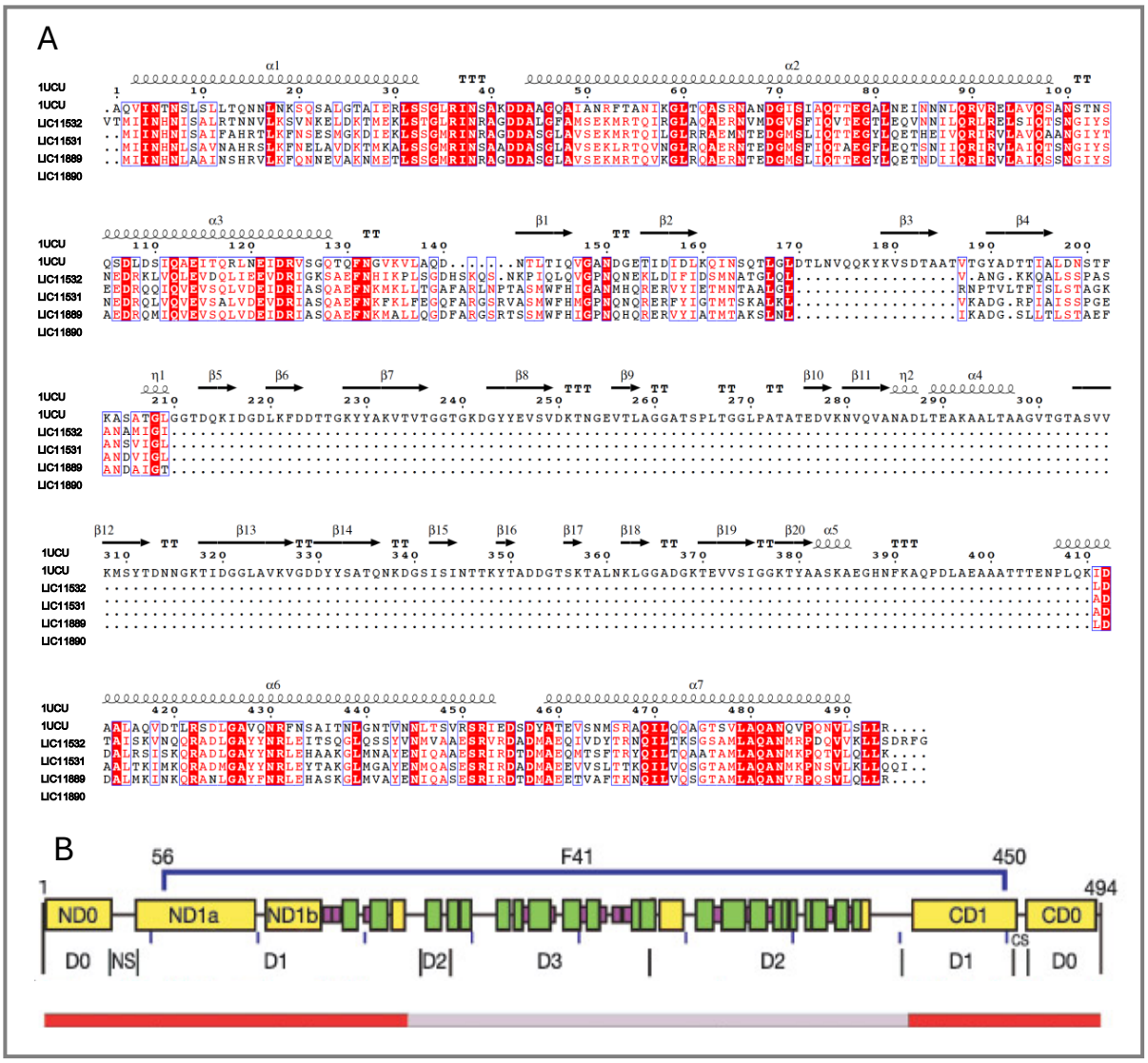


Figura 9: Análisis comparativo de flagelinas. A) Alineamiento múltiple de secuencias de la flagelina FliC de *S. enterica* (1UCU) con FlaB1 (gen LIC11890), FlaB2 (LIC11889), FlaB3 (LIC11532) y FlaB4 (LIC11531) de *L. interrogans*. Los elementos de estructura secundaria de FliC están anotados arriba de las secuencias, numerando las hélices α y hebras β secuencialmente. Los aminoácidos totalmente conservados están marcados con fondo rojo, mientras que aquellos que conservan propiedades similares se marcan con caracteres rojos. Los segmentos de mayor conservación se delinearon dentro de cajas. **B)** Posición y arreglo de elementos estructurales en la secuencia de FliC de *S. enterica* (tomado de Yonekura et al., 2003 [33]). Las regiones que conforman los cuatro dominios D0 a D3 están indicadas. La estructura de la versión truncada F41 utilizada para resolver la estructura cristalográfica está marcada. Los rectángulos indican estructuras secundarias: amarillo para hélices α, verde para hebras β, y violeta para vueltas β. Abajo se muestran las regiones de secuencias aminoacídicas bien conservadas (rojo) y variables (violeta). Comparando ambos paneles, nótese la ausencia de los dominios D2 y D3 en las cuatro isoformas de FlaB de *L. interrogans*.

Utilizando el programa Rosetta, generamos 600 modelos por homología para cada proteína FlaB de *Leptospira* basados en los alineamientos pareados con FliC. Dado que los dominios D0 y D1 compartidos con FliC son justamente los que ésta involucra para autoensamblarse formando el homo-polímero helicoidal del filamento, asumimos que FlaB generará ensamblajes con simetría helicoidal análoga. Inicialmente utilizamos los parámetros de

simetría helicoidal de la estructura del filamento ensamblado de *S. enterica* [33]. Los modelos de las cuatro isoformas de FlaB revelan que los dominios D0 y D1 comparten conformaciones similares, mientras que se aprecian mayores diferencias en los bucles que quedarán expuestos al solvente en el ensamblaje flagelar, coherente con la menor similitud de secuencia en esas regiones (Figura 10). Los modelos presentan estructuras generalmente similares con patrones algo diferentes de hidrofobicidad. Considerando los valores de energía calculados para cada filamento ensamblado (entre -700 y -950 REU [*Rosetta Energy Unit*]), es altamente probable que el ensamblaje de FlaB constituya el núcleo del filamento flagelar.

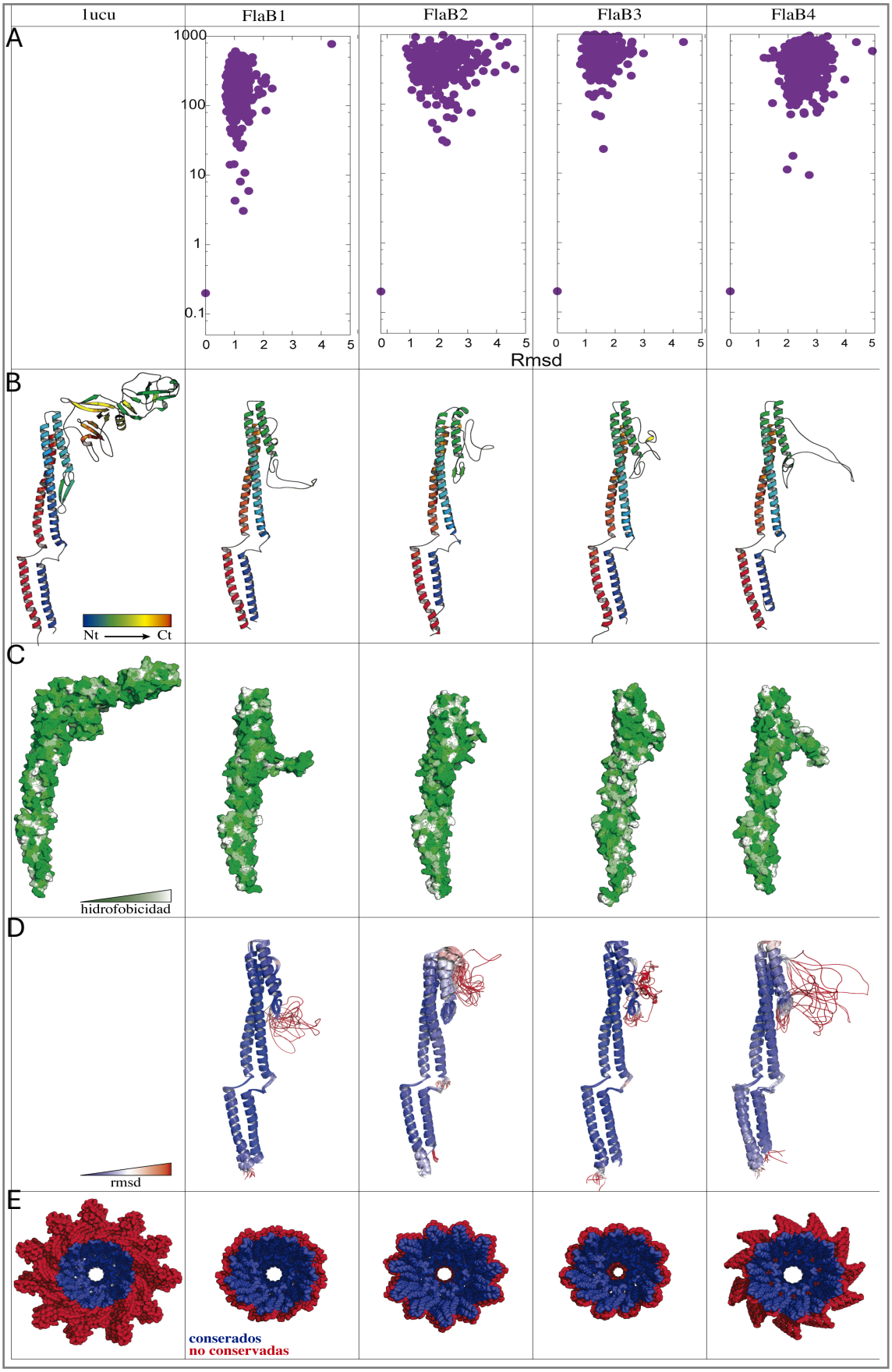


Figura 10: Estructura tridimensional de FlaB de *L. interrogans*. En la primera columna se muestran los modelos experimentales de la flagelina FliC de *S. enterica* y en las columnas 2-5 los modelos calculados in silico correspondientes a las cuatro isoformas de FlaB. La primera fila **A)** muestra gráficos donde cada modelo

generado está representado por un punto. Las desviaciones cuadráticas medias (RMSD) se calcularon para cada modelo en comparación con el de mejor score de energía. En la fila **B)** se muestran los modelos de menor *score* energético, considerados de mejor calidad, correspondientes a un protómero calculado en contexto de simetría helicoidal. Las estructuras están coloreadas desde el extremo N-terminal al C-terminal con rampa de azul a rojo. En la fila **C)** se mapeó en la superficie de cada protómero la hidrofobicidad calculada de los residuos. En **D)**, los 10 mejores modelos del muestreo de partida se muestran alineados entre sí utilizando el modelo de mejor *score* como referencia. Las regiones con mayor variabilidad, detectadas calculando el RMSD residuo a residuo, se muestran en rojo y sugieren mayor flexibilidad en esa región de las moléculas. En la fila **E)** se muestran cortes axiales para cada modelo, coloreados según su nivel de conservación de secuencia comparando cada FlaB vs FliC. Las regiones más internas, en azul, se encuentran altamente conservadas entre todas las flagelinas, mientras que los residuos expuestos en la superficie, en rojo, presentan la mayor variabilidad.

El diámetro de los filamentos homopoliméricos de FlaB generados varía entre 145 y 192 Å, valores muy próximos a los reportados experimentalmente en imágenes de crio-tomografía electrónica de filamentos purificados de una cepa *knock-out fcpA*$^{-}$ de *L. interrogans*, que probablemente presenten un núcleo desnudo de FlaB (Figura 11).

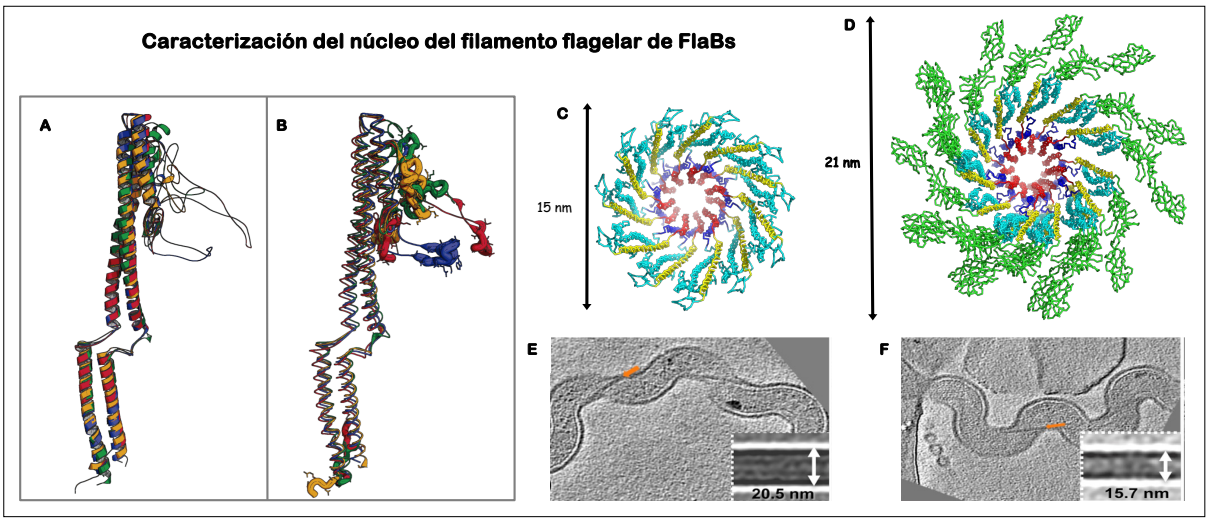


Figura 11: Diferencias entre las cuatro isoformas de FlaB. A) Los cuatro modelos superpuestos de las isoformas de FlaB de *L. interrogans* se muestran como cartoon, cada una con un color diferente (FlaB1: azul, FlaB2: verde, FlaB3: naranja, FlaB4: rojo). Se nota la gran similitud estructural general, con la mayor diversidad concentrada en la región correspondiente al bucle entre las hélices α2 y α3 del dominio D1. **B)** Los mismos modelos superpuestos de manera de subrayar la similitud de secuencia, con las regiones de mayor variabilidad representadas por el grosor del cartoon. **C)** Modelo del filamento flagelar compuesto por homopolímeros de FlaB1 de *L. interrogans* modelados por homología con la flagelina de *S. enterica*. **D)** Modelo experimental del filamento flagelar de *S. enterica*, que presenta un diámetro de 21 nm. **E-F)** Crio-electrotomografía de cepas Fiocruz L1-130 *fcpA*$^{-}$ y Fiocruz L1-130 *fcpA*$^{-/+}$. Las flechas rojas indican la localización de los filamentos periplásmicos. Los insertos muestran los mapas promediados de los filamentos de ambas cepas, mostrando sus diámetros (figuras tomadas de Wunder et al., 2016 [105]).

Dada la alta conservación estructural y de secuencia en las regiones que guían la polimerización en el núcleo interno de los filamentos, postulamos que las cuatro isoformas de FlaB tienen la capacidad de autoensamblarse, solas o en combinación. La diferencia

estructural más notable de FlaB comparada con la flagelina de *Salmonella* es la ausencia de los dominios D2 y D3, que en *Salmonella* quedan expuestos en la superficie flagelar generando filamentos con diámetro aproximado de 21 nm. *Leptospira* carece de estos dominios, sin embargo el diámetro reportado para los filamentos de cepa salvaje es similar (~20 nm), mientras que el de los filamentos de la cepa *knock-out fcpA*$^{-}$, es ~15 nm (Figura 11). Postulamos que en *Leptospira*, el volumen correspondiente a los dominios D2/D3 es reemplazado por las proteínas que desempeñan el rol de vaina, probablemente incluyendo FcpA, FcpB, FlaA1 y FlaA2. La mayor diversidad estructural entre las distintas isoformas de FlaB se concentra entre las hélices α2 y α3 del dominio D1, en los bucles flagelares expuestos al solvente, que coinciden con las regiones de menor similitud de secuencia en las cuatro isoformas. Estos sitios presentan diferencias en conformación, hidrofobicidad y flexibilidad, por lo que podrían mediar distintas capacidades de interacción de las cuatro isoformas con otros componentes flagelares.

3.2.2 – Estructura tridimensional de las "*Flagellar coiling proteins*" o proteínas inductoras del superenrrollamiento flagelar

Las proteínas FcpA y FcpB son componentes del filamento flagelar presentes únicamente en el género *Leptospira*. Estas proteínas desempeñan un rol relevante en la motilidad traslacional y patogenicidad[105]. Dada su importancia biológica y papel estructural en los endoflagelos de *Leptospira*, nos propusimos resolver su estructura tridimensional por cristalografía de rayos X.

FcpA es uno de los componentes más abundantes de *L. interrogans*[134] y se encuentran ortólogos en todas las especies de *Leptospira*, con identidad de secuencia de 95-100% en patógenas, 88% en intermedias y 76-79% en saprófitas. Trabajamos con FcpA de una especie patógena (*L. interrogans*) y de una saprófita (*L. biflexa*). La estructura de esta última había sido previamente resuelta en nuestro laboratorio, pero nos propusimos mejorar su resolución y el rendimiento con el que se obtiene la proteína soluble. Cambiando la cepa de expresión y las condiciones de purificación, la expresión recombinante de FcpA de *L. biflexa* fue optimizada obteniendo rendimientos típicos de 30 mg de proteína por litro de cultivo con alta pureza (Figura 12). La cromatografía de exclusión molecular mostró que FcpA se separa isocráticamente como un dímero, consistente con el peso molecular estimado. Con esta muestra se realizó un rastreo de 384 condiciones de cristalización, observándose cristales en 57 condiciones. Se seleccionaron siete condiciones basándose en el tamaño de los cristales, calidad del hábito de crecimiento e intensidad de luz polarizada desviada por los mismos.

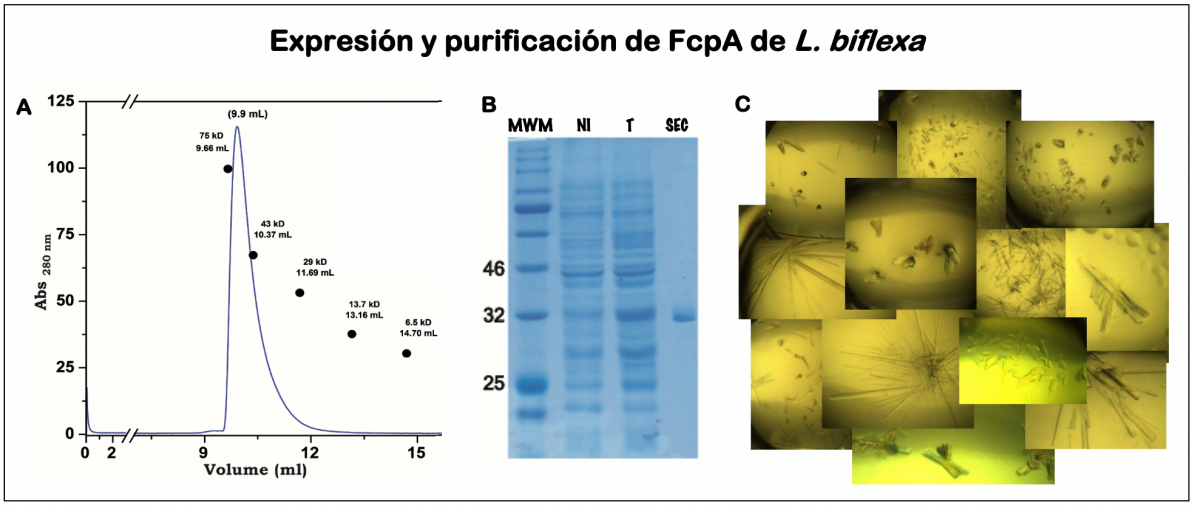


Figura 12: Expresión y purificación de FcpA de *L. biflexa*. A. La cromatografía de exclusión por tamaño (SEC) es el último paso en la purificación de FcpA recombinante de *L. biflexa*. En solución, esta proteína tiene un tamaño aparente que es coherente con una estructura cuaternaria dimérica (33kDa es el peso estimado para él monómero). **B.** SDS-PAGE mostrando la purificación de bFcpA, cuya sobre-expresión en la cepa Rosetta-gami 2 (DE3) de *E. coli* fue inducida a 30ºC por 4hs con IPTG. 1. Marcadores de peso molecular (MPM), 2. Extracto total de la cepa transformada no inducida, 3. Extracto total de la cepa inducida, 4. bFcpA luego de la purificación por SEC. **C.** Cristales de bFcpA obtenidos luego del ensayo de cristalización.

Los cristales fueron analizados mediante difracción de rayos X, resultando en la resolución de FcpA en tres formas cristalinas en distintos grupos de espacio (Figura 13). Las dos formas cristalinas monoclínicas fueron obtenidas a partir de cristales optimizados en gota colgante, mientras que la variante hexagonal fue obtenida directamente del rastreo inicial y sorpresivamente difractó mejor alcanzando resoluciones de 2 Å. El grupo de espacio hexagonal P622 es muy poco frecuente en macromoléculas biológicas (0.086%). Las estructuras fueron resueltas por reemplazo molecular y depositadas en la PDB bajo los códigos 6NQW, 6NQX y 6NQY.

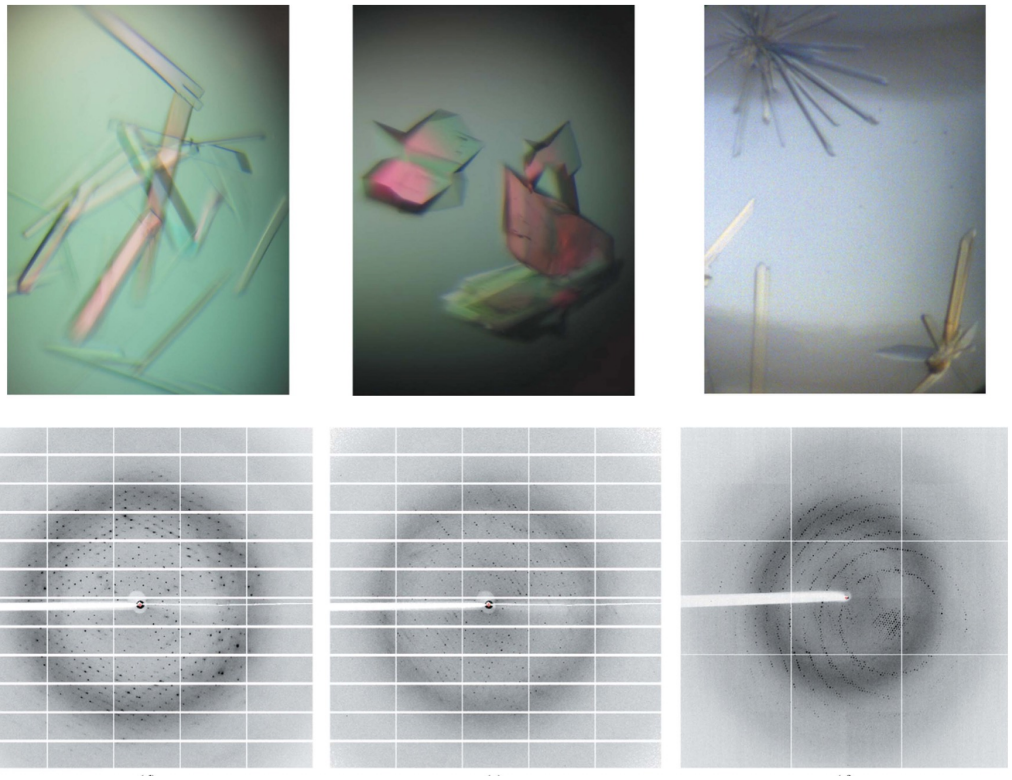
Resolución de la estructura de bFcpA

A

Crystallization.

Crystal form	1 (centred monoclinic)	2 (primitive monoclinic)	3 (hexagonal)
Method	Hanging-drop vapour diffusion	Hanging-drop vapour diffusion	Sitting-drop vapour diffusion
Plate type	VDX 24-well	VDX 24-well	96-well CrystalQuick
Temperature (K)	293	293	293
Protein concentration (mg ml^{-1})	20	20	20
Buffer composition of protein solution	20 m*M* Tris–HCl pH 8.0, 300 m*M* NaCl	20 m*M* Tris–HCl pH 8.0, 300 m*M* NaCl	20 m*M* HEPES pH 7.0, 200 m*M* NaCl
Composition of reservoir solution	20–25%(*v*/*v*) PEG 400, 100 m*M* Tris–HCl pH 7.5–8.5	3–5%(*w*/*v*) PEG 4000, 100 m*M* Tris–HCl pH 7.5–8.0	12%(*w*/*v*) PEG 8000, 0.1 *M* Na/K phosphate pH 6.2, 0.2 *M* NaCl
Volume and ratio of drop	4 µl (1:1)	4 µl (1:1)	0.6 µl (1:1)
Volume of reservoir (ml)	1	1	0.1

B



C

Crystal form	Form 1: centred monoclinic	Form 2: primitive monoclinic	Form 3: hexagonal
Diffraction source	PROXIMA 1, SOLEIL	PROXIMA 1, SOLEIL	PROXIMA 2A, SOLEIL
Wavelength (Å)	0.97857	0.97857	0.97910
Temperature (K)	100	100	100
Detector	Hybrid photon counting (PILATUS 6M, Dectris)	Hybrid photon counting (PILATUS 6M, Dectris)	Area detector (Q315r CCD, ADSC)
Crystal-to-detector distance (mm)	484.81	485.34	274.35
Rotation range per image (°)	0.2	0.2	0.5
Total rotation range (°)	180	180	100
Exposure time per image (s)	0.2	0.2	0.5
Space group	*C*2	$P2_1$	*P*622
a, *b*, *c* (Å)	82.3, 99.6, 106.7	85.5, 96.5, 121.1	132.7, 132.7, 67.6
α, β, γ (°)	90, 91.95, 90	90, 105.2, 90	90, 90, 120
Mosaicity (°)	0.24	0.15	0.20
Resolution range (Å)	45.12–2.50 (2.60–2.50)	48.23–2.95 (3.07–2.95)	43.76–2.00 (2.11–2.00)
Total No. of reflections	99658 (10749)	137503 (15749)	286178 (40385)
No. of unique reflections	29300 (3278)	39975 (4485)	24246 (3460)
Completeness (%)	98.4 (97.9)	99.5 (99.7)	99.9 (99.6)
Mulitplicity	3.4 (3.3)	3.4 (3.5)	11.8 (11.7)
$\langle I/\sigma(I)\rangle$	18.2 (2.3)	17.2 (2.0)	17.5 (2.5)
R_{meas} (%)	4 (57)	6 (76)	11 (119)
Overall *B* factor from Wilson plot (Å^2)	61.3	75.6	30.5

Figura 13: Resolución de la estructura de FcpA de *L. biflexa*. A. Tabla mostrando las condiciones de cristalización de las tres formas cristalinas analizadas para FcpA de *L. biflexa* obtenida en forma recombinante y concentrada a 20mg/mL. **B.** Cristales analizados de bFcpA y abajo imágenes de difracción de rayos X representativas para cada uno de ellos. A la izquierda para la forma centrada monoclínica; al medio para la forma primitiva monoclinica; y a la derecha para la forma hexagonal. **C.** Tabla mostrando la colección de datos por difracción de rayos X y el procesamiento de las tres formas cristalinas. Los valores entre paréntesis son para la capa de resolución más alta. Figuras tomadas de San Martin et al., 2017[136].

La estructura de FcpA de *L. biflexa* representa un plegamiento completamente nuevo sin homología estructural conocida en la base de datos PDB (Figura 3.8). FcpA está organizada como una proteína todo-α, con diez α-hélices agrupadas en dos subdominios. La hélice α4, la más larga de la estructura, conduce a la dimerización de la proteína por interacciones hidrofóbicas con la hélice α3 de la molécula adyacente y puentes de hidrógeno entre las dos α4. Esta superficie de dimerización esconde 960 Å², se evidencia en los tres empaquetamientos cristalinos obtenidos y es consistente con el dímero observado en solución durante la purificación.

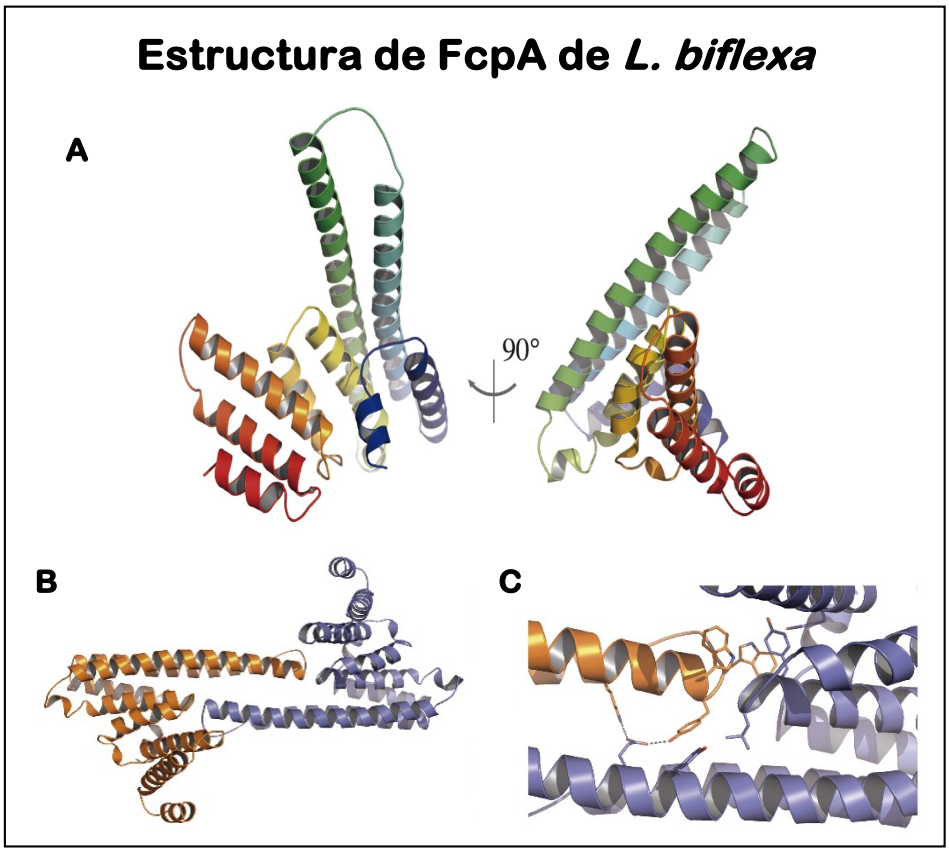


Figura 14: Estructura cristalográfica de FcpA de *L. biflexa*. A) Estructura de FcpA en dos vistas ortogonales, coloreada con rampa de colores del azul (extremo N-terminal) hasta el rojo (C-terminal). Su estructura representa un plegamiento nuevo y está compuesta únicamente por α-hélices dispuestas en dos subdominios. **B)** Dímero de FcpA observado en las tres formas cristalinas. Cada monómero está coloreado en un color diferente y se observa el estrecho contacto entre la hélice α-4 más larga de cada FcpA con las hélices α-3 y α-4 de la subunidad vecina, que esconde 960 Å² de superficie. **C)** Imagen aumentada de la interacción mostrada en B, resaltando los residuos que participan en la dimerización.

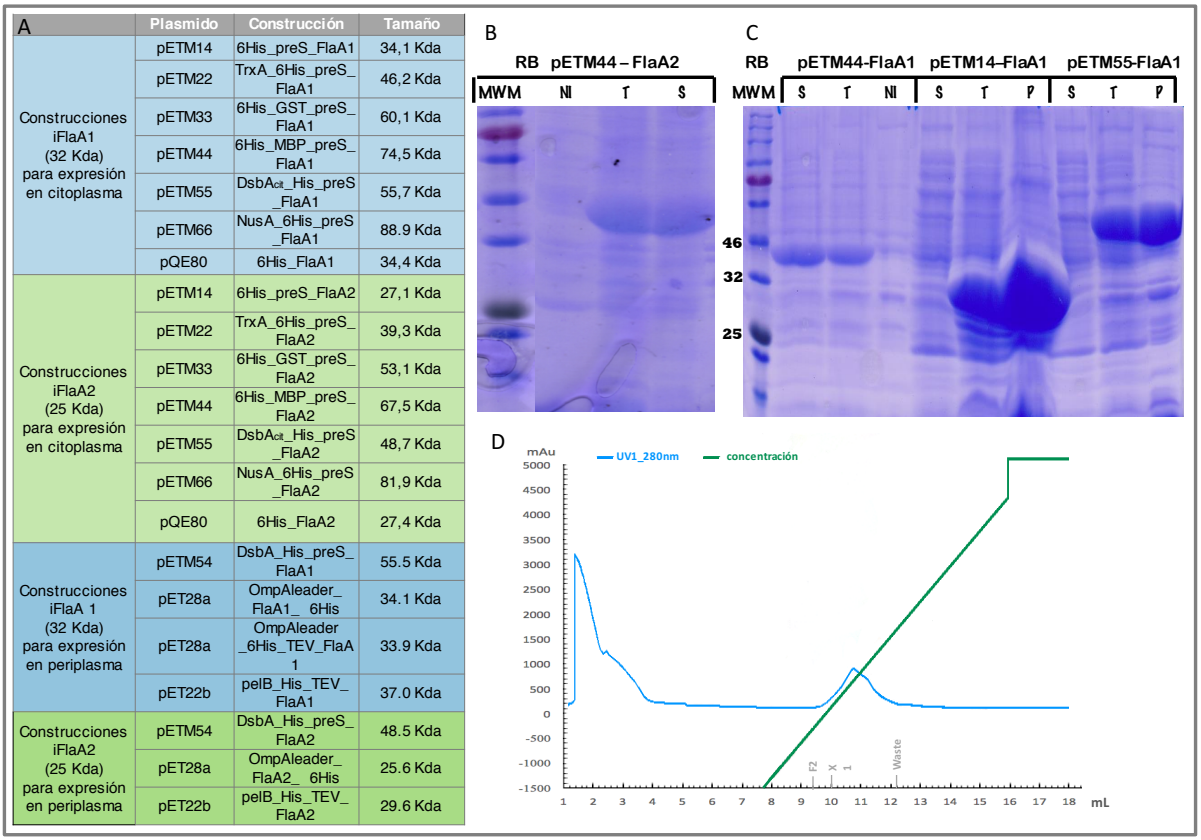
Para FcpA de *L. interrogans*, aunque se habían obtenido previamente cristales, estos resultaron ser pobres difractores obteniéndose señal solo hasta ~10 Å. Se produjeron mutantes N-terminal truncados basados en ensayos de proteólisis limitada, pero no se obtuvieron cristales. Considerando que FcpA de *L. interrogans* comparte 77% de identidad con la de *L. biflexa* sin gaps predichos, diseñamos un mutante que facilitara la cristalización. Una diferencia importante es la presencia de dos cisteínas (Cys221 y Cys241) que podrían formar un puente disulfuro solo en FcpA de *L. interrogans*. Estas cisteínas podrían estar reducidas debido a la expresión recombinante en el citoplasma de *E. coli*, siendo una posible fuente de heterogeneidad que afectaría la calidad cristalina. Se diseñó y clonó la variante FcpA_cysless que incluye las sustituciones Cys221Thr y Cys241Ala, equivalentes a los residuos presentes en *L. biflexa*. Esta variante fue sobre-expresada y purificada, estableciendo un equilibrio monómero-dímero en solución. Por cromatografía de exclusión se separaron ambas formas que fueron utilizadas independientemente para rastreo de cristalización. Se probaron 384 condiciones diferentes con ambas formas monomérica y dimérica, incluyendo técnicas de microsembrado con semillas de cristales de FcpA de *L. biflexa*. Sin embargo, no se lograron obtener cristales de FcpA de *L. interrogans*.

En paralelo, durante mi Tesis, en el laboratorio se logró resolver la estructura de FcpB de *L. interrogans* a 2.6 Å de resolución. FcpB es una proteína α+β con una gran hoja β de 8 hebras expuesta al solvente e interactuando por una de sus caras con dos largas hélices α. Este plegamiento tiene cierta homología con la proteína humana p32. La estructura de FcpB fue depositada en la wwPDB bajo el código 6NQZ.

3.2.3 – Estudios estructurales de FlaA1 y FlaA2 de *Leptospira*

3.2.3.1 – Expresión recombinante de FlaA1 y FlaA2

Las proteínas FlaA son componentes esenciales y exclusivos de los endoflagelos de Espiroquetas. Se desconocen sus estructuras tridimensionales y el rol que cumplen en los filamentos, por lo que nos propusimos resolverlas por cristalografía. Como estrategia inicial, las secuencias codificantes de FlaA1 y FlaA2 de *L. interrogans* fueron subclonadas en diez vectores de expresión de *E. coli* dirigidos a producir distintas proteínas de fusión en el citoplasma (Figura 15). Los insertos fueron truncados para no contener los primeros 33 residuos N-terminales de FlaA1 y los primeros 22 de FlaA2, evitando las secuencias señal de exportación al periplasma.



A	Plasmido	Construcción	Tamaño
Construcciones iFlaA1 (32 Kda) para expresión en citoplasma	pETM14	6His_preS_FlaA1	34,1 Kda
	pETM22	TrxA_6His_preS_FlaA1	46,2 Kda
	pETM33	6His_GST_preS_FlaA1	60,1 Kda
	pETM44	6His_MBP_preS_FlaA1	74,5 Kda
	pETM55	DsbA$_{cit}$_His_preS_FlaA1	55,7 Kda
	pETM66	NusA_6His_preS_FlaA1	88.9 Kda
	pQE80	6His_FlaA1	34,4 Kda
Construcciones iFlaA2 (25 Kda) para expresión en citoplasma	pETM14	6His_preS_FlaA2	27,1 Kda
	pETM22	TrxA_6His_preS_FlaA2	39,3 Kda
	pETM33	6His_GST_preS_FlaA2	53,1 Kda
	pETM44	6His_MBP_preS_FlaA2	67,5 Kda
	pETM55	DsbA$_{cit}$_His_preS_FlaA2	48,7 Kda
	pETM66	NusA_6His_preS_FlaA2	81,9 Kda
	pQE80	6His_FlaA2	27,4 Kda
Construcciones iFlaA 1 (32 Kda) para expresión en periplasma	pETM54	DsbA_His_preS_FlaA1	55.5 Kda
	pET28a	OmpAleader_FlaA1_ 6His	34.1 Kda
	pET28a	OmpAleader_6His_TEV_FlaA1	33.9 Kda
	pET22b	pelB_His_TEV_FlaA1	37.0 Kda
Construcciones iFlaA2 (25 Kda) para expresión en periplasma	pETM54	DsbA_His_preS_FlaA2	48.5 Kda
	pET28a	OmpAleader_FlaA2_ 6His	25.6 Kda
	pET22b	pelB_His_TEV_FlaA2	29.6 Kda

Figura 15: Ensayos iniciales de expresión y purificación de isoformas de FlaA de *Leptospira*. A. Tabla mostrando las construcciones utilizadas para los ensayos de expresión de FlaA1 y FlaA2 de *L. interrogans*. **B y C.** SDS-PAGE mostrando algunos de los ensayos de expresión a pequeña escala en la cepa Rosetta Blue de *E. coli*. NI: Fracciones totales de las cepas transformadas, no inducidas; T: Fracciones totales de las cepas inducidas; S: Fracciones solubles purificadas; P: Fracciones insolubles. La expresión a partir de los plásmidos pETM44-FlaA1 y pETM44-FlaA2, en todas las cepas es soluble, pero con un tamaño aparente de 43kDa correspondiente al *partner* de expresión soluble, MBP. **D.** La expresión a partir de pETM66-FlaA2 es soluble en todas las cepas. La proteína recombinante NusA-His-FlaA2 fue escalada a 1Lt en la cepa de *E. coli* BL21 (DE3) y purificada por IMAC como se muestra en la figura, pero precipita al cortar el tag.

La expresión en algunas cepas resultó tóxica para el crecimiento bacteriano, mientras que la mayoría fueron capaces de crecer y expresar las proteínas, aunque en forma insoluble. En el caso de los vectores pETM44-FlaA1 y pETM44-FlaA2, se observó proteína recombinante en la fracción soluble, pero su tamaño correspondía al de His-MBP sin la proteína de interés, probablemente producto de proteólisis *in vivo*. Para la expresión con pETM66-FlaA2, que en todas las cepas produjo la proteína recombinante de fusión esperada NusA-His-FlaA2, se decidió producirla a mayor escala y purificarla por IMAC. Sin embargo, la digestión con proteasa TEV para eliminar la proteína de fusión provocó la precipitación de FlaA2, indicando posibles problemas de plegamiento.

Dadas las dificultades para conseguir expresión soluble, se generaron nuevas construcciones para expresión recombinante en el espacio periplasmático, considerando que las FlaA maduras son proteínas periplasmáticas en *Leptospira*. Se subclonaron las secuencias en vectores pET22b, pET28a y pETM54, generando siete construcciones diferentes (Figura 16).

Los ensayos de expresión a pequeña escala en medio 2YT con inducción a 20°C fueron auspiciosos en tres construcciones: DsbA_His_pS_FlaA1, DsbA_His_pS_FlaA2 y OmpA leader_His_FlaA1, mostrando sobre-expresión tanto en BL21 como en C41.

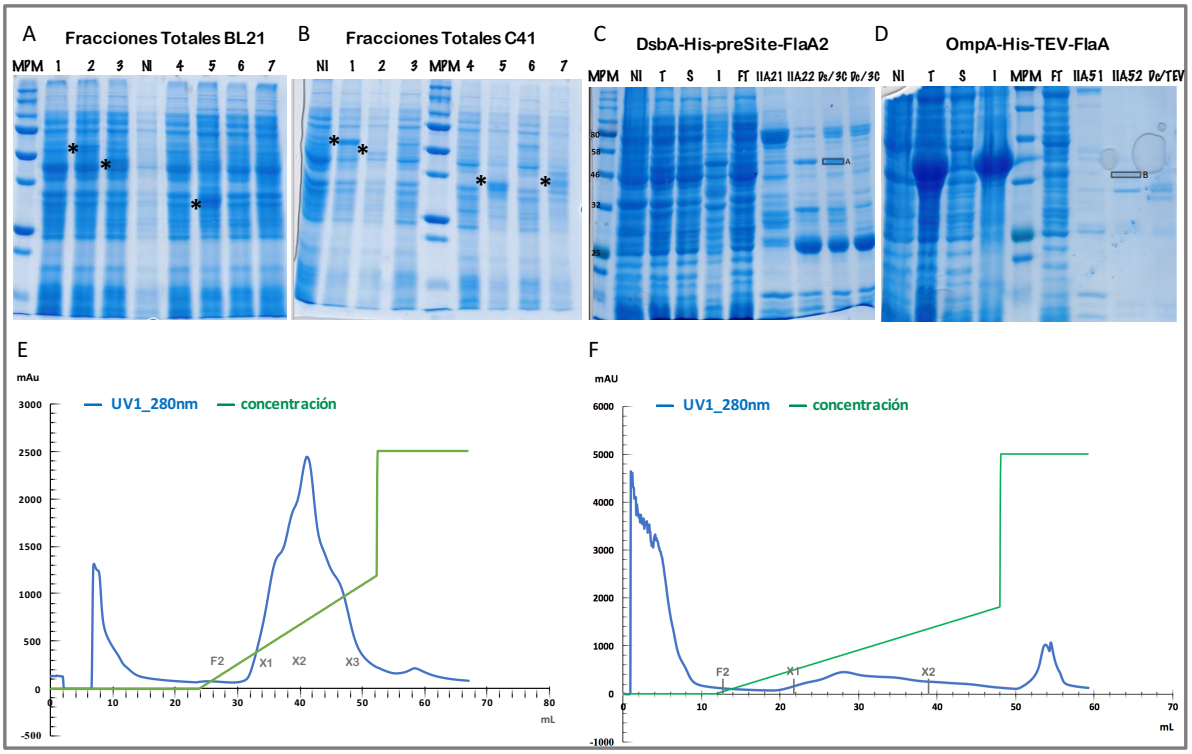


Figura 16: Expresión en periplasma de FlaA1 y FlaA2 de *L. Interrogans*. En los dos paneles superiores de la izquierda se muestran los SDS-PAGE correspondientes a los ensayos de expresión a pequeña escala en las cepas **A**: BL21(DE3) y **B**: C41 a partir de las construcciones: 1: pETM54-FlaA1; 2: pETM54-FlaA2; 3: pET28a-FlaA2-6His; 4: pET28a-FlaA1-6His; 5: pET28a-6His-FlaA1; 6: pET22b-FlaA2 y 7: pET22b-FlaA1. * marcan las proteínas expresadas en la fracción total. Los paneles **C y D** muestran las purificaciones de los cultivos escalados en la cepa C41 para las construcciones 2 y 5 respectivamente. NI: Fracciones totales de C41 transformadas, no inducidas; T: Fracciones totales de las cepas inducidas; S: Fracciones solubles purificadas; I: Fracciones insolubles; FT: *flow-through* de la IMAC; ; IIA21: muestra de la fracción X1 de la IMAC en el panel E; IIA22: muestra de la fracción X2 de la misma IMAC; Ds/3C: Fracción del pico dialisado luego de la IMAC; Dc/3C; Fracción del mismo pico dialisado con la proteasa 3C; IIA51 primer pico de elución de la IMAC mostrada en el panel F; IIA52 segundo pico de elución de la IMAC; Dc/TEV: Fracción luego de la diálisis con TEV. Las bandas con recuadros A y B de los paneles C y D, fueron identificadas por espectrometría de masas y corresponden a FlaA2 (25KDa) y FlaA1 (32KDa) respectivamente.

La expresión a mayor escala indicó que efectivamente había cierta cantidad de proteína recombinante soluble, confirmada como FlaA1 y FlaA2 por espectrometría de masas. Sin embargo, considerando el bajo rendimiento y la necesidad de agregar varios pasos adicionales de purificación, se decidió no continuar optimizando esta purificación.

3.2.3.2 – Coexpresión de FlaA

Para intentar resolver las dificultades anteriores, hipotetizamos que las interacciones específicas entre interactores flagelares fisiológicamente relevantes podrían promover el correcto plegamiento de las isoformas de FlaA. Esta hipótesis tiene sustento en resultados previos que sugerían interacción directa entre FcpA y FcpB de *L. biflexa*, y el reporte sobre la interacción directa entre FlaA2 y FcpA de *L. biflexa* (Sasaki et al., 2018). Se realizaron extensos esfuerzos en este sentido, pues los complejos serían de doble interés, permitiendo analizar interacciones específicas potencialmente relevantes para entender el ensamblado supramolecular del filamento nativo. Se clonaron las secuencias codificantes de FcpA y FlaA2 de *L. biflexa* en el plásmido pColaDuet y las secuencias de FlaA1 y FcpB en el plásmido compatible pETDuet. Estos plásmidos fueron co-transformados en células Rosetta-gami2 de *E. coli* y co-expresados a gran escala para obtener las cuatro proteínas con diferentes etiquetas de afinidad. Por SDS-PAGE y espectrometría de masas pudimos concluir que las proteínas FcpA, FlaA1 y FlaA2 de *L. biflexa,* efectivamente se co-expresan en estas condiciones. Sin embargo, los productos resultanes terminaron siempre siendo o bien agregados grandes insolubles, o con rendimientos extremadamente bajos.

3.2.3.3 – Expresión *high-throughput* de distintos mutantes de FlaA1 y FlaA2

Integrando información detallada de los alineamientos múltiples (ver sección 3.1), se implementaron tres estrategias diferentes para producir las isoformas de FlaA en forma soluble (Figura 17). La primera fue probar la expresión recombinante de FlaA1 y FlaA2 de tres especies distintas de *Leptospira* (ortólogos), notando que para las dos isoformas de FlaA hay dos clusters distintos: un grupo propio de especies patógenas y otro de saprófitas. Se seleccionaron *L. biflexa* como especie saprófita, y *L. interrogans* y *L. borgpetersenii* como patógenas.

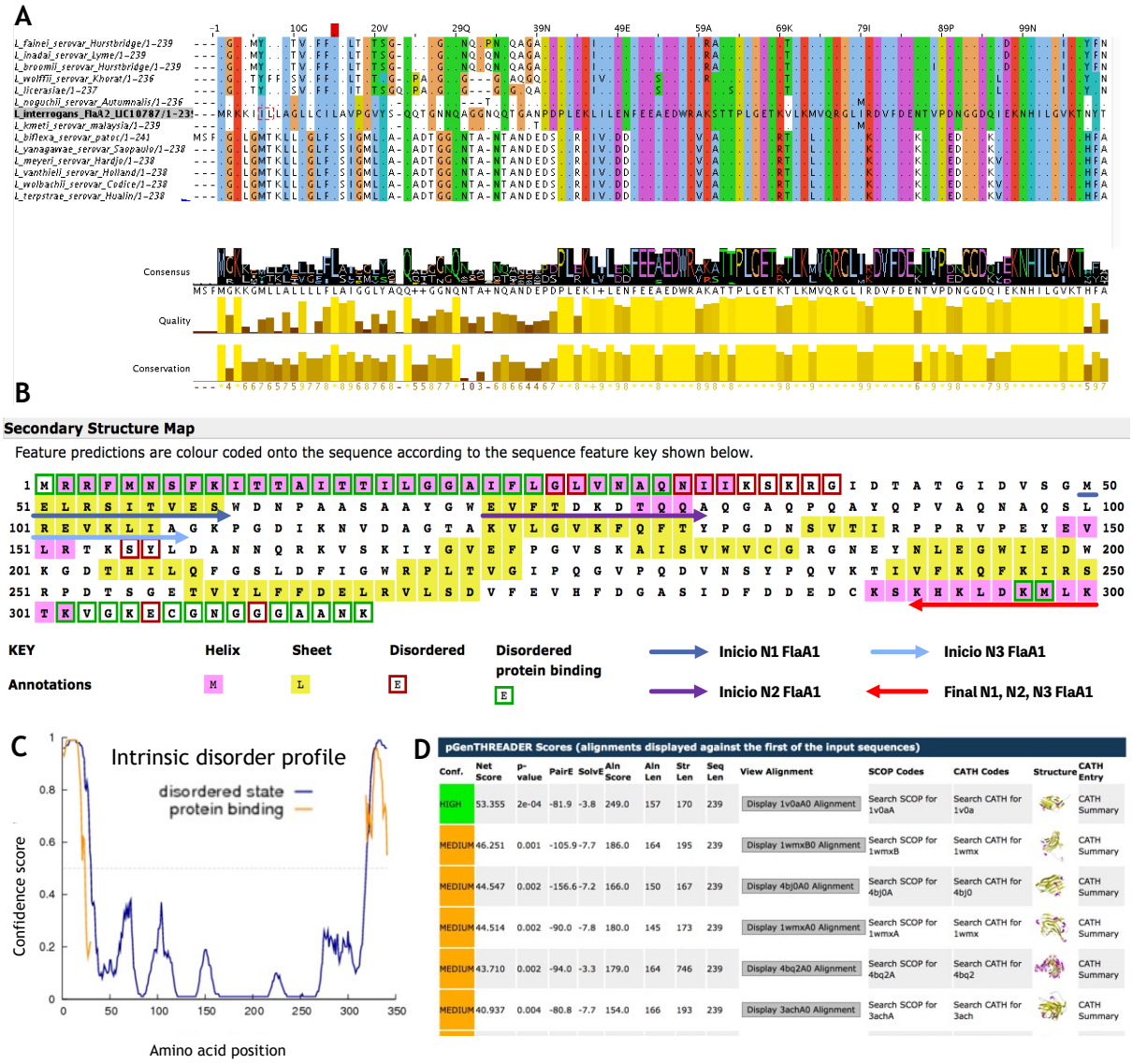


Figura 17: Análisis de secuencias de isoformas de FlaA y diseño de construcciones. A. Alineamientos múltiples de secuencia ClustalW[151] de distintas especies de *Leptospira* para la isoforma FlaA2, analizados con Jalview[152]. El código de colores ZAPPO muestra fácilmente la conservación de secuencia. En la primera fila se puede observar el número de residuo, y el foco se hace en la región N-terminal, donde hay mayor divergencia. **B.** Gráfico de secuencia resultante de la predicción de estructuras secundarias con PSIPRED[153]. Se muestra en particular la predicción para FlaA1 de *L. borgpetersenii* (locus_tag="LBL_2375") y las secuencias cubiertas por los cebadores diseñados. **C.** Perfil de desorden intrínseco predicho por PSIPRED para la secuencia de FlaA2 de *L. borgpetersenii* (locus_tag="LBL_2376"). **D.** Predicción de homólogos estructurales de FlaA2 de *L. borgpetersenii*, utilizando el servidor PHYRE2[154]. Todas las secuencias que se identifican son de módulos de unión a carbohidratos. La de mejor *score* corresponde a la estructurta PDB 1v0a (modulo de unión a carbohidratos de familia 11 de *Clostridium thermocellum*.

Como segunda estrategia, considerando que una de las regiones de mayor divergencia de secuencia se encuentra en el extremo N-terminal y que las predicciones de estructura secundaria sugieren que los extremos N- y C-terminales estarían desordenados, se generaron tres mutantes N- y/o C-terminal truncados para cada ortólogo, diseñados para evitar la disrupción de estructuras secundarias definidas.

Como tercera estrategia complementaria para aumentar el espacio de búsqueda, las nueve secuencias codificantes de cada isoforma se clonaron en el Vector Suite diseñado por Agustín Correa[155], una batería de vectores que permite el clonado paralelo utilizando un único par de cebadores. Se utilizaron los nueve vectores disponibles para expresar las secuencias sin fusión y fusionadas a ocho proteínas diferentes: MBP, SUMO, TrX, PelB, OsmY, CelDnc, DsbC y GST.

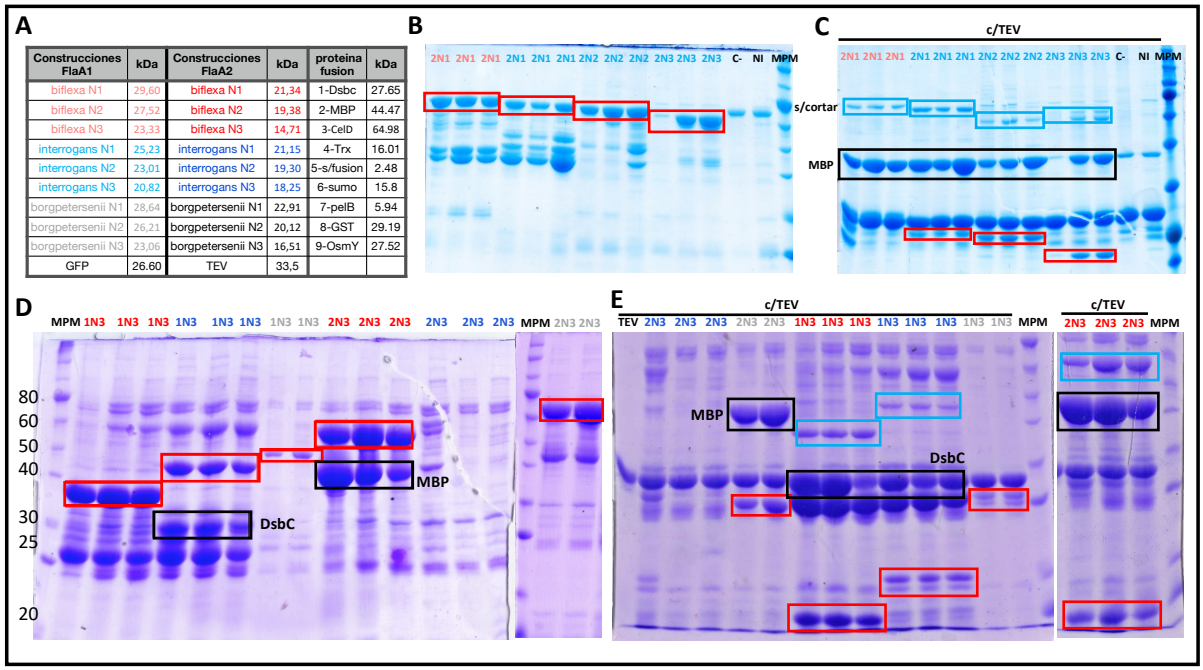
Se lograron clonar un total de 61 construcciones de FlaA1 (22 de *L. interrogans*, 23 de *L. biflexa* y 16 de *L. borgpetersenii*) y 54 construcciones de FlaA2 (15 de *L. interrogans*, 16 de *L. biflexa* y 23 de *L. borgpetersenii*), representando un 85% de las construcciones totales (Tabla 1). Para la expresión heteróloga se utilizó la cepa Lemo21 de *E. coli*, diseñada para expresar blancos difíciles variando finamente los niveles de inducción con distintas cantidades de ramnosa.

	1 DsbC	2 MBP	3 CelD	4 Trx	5 S/fusión	6 sumo	7 pelB	8 GST	9 OsmY
bFlaA1 N1	✓ -	✓ ✓	✓ -	✓ ✗	✓ ✓	✓ ✗	✓ ✓	✓ ✗	✗ ✗
bFlaA1 N2	✓ ✗	✓ ✗	✓ -	✗ ✗	✓ ✗	✓ ✗	✓ ✗	✓ ✗	✓ -
bFlaA1 N3	✓ ✗	✓ ✗	✓ ✗	✓ ✗	✓ ✗	✓ -	✓ ✗	✗ ✗	✗ ✗
iFlaA1 N1	✗ ✗	✓ ✓	✗ ✗	✓ ✗	✗ ✗	✓ -	✓ -	✓ ✗	✓ -
iFlaA1 N2	✗ ✗	✓ ✓	✓ ✗	✓ -	✓ ✗	✓ ✗	✓ -	✗ ✗	✓ ✓
iFlaA1 N3	✓ ✓	✓ ✓	✓ ✗	✓ ✗	✓ -	✓ -	✓ ✓	✓ -	✓ ✓
boFlaA1 N1	✓ -	✓ ✓	✗ ✗	✓ ✓	✓ ✓	✓ -	✓ ✓	✓ ✗	✓ -
boFlaA1 N2	✗ ✗	✗ ✗	✗ ✗	✗ ✗	✗ ✗	✗ ✗	✗ ✗	✗ ✗	✗ ✗
boFlaA1 N3	✓ ✓	✓ ✓	✓ -	✓ ✓	✓ ✓	✗ ✗	✓ ✗	✓ ✗	✓ -
bFlaA2 N1	✓ ✓	✓ ✓	✓ -	✓ -	✓ -	✓ -	✓ -	✗ ✗	✓ -
bFlaA2 N2	✗ ✗	✗ ✗	✗ ✗	✗ ✗	✗ ✗	✗ ✗	✗ ✗	✗ ✗	✗ ✗
bFlaA2 N3	✓ ✓	✓ ✓	✓ -	✓ -	✓ ✗	✓ ✗	✓ ✗	✗ ✗	✓ -
iFlaA2 N1	✓ ✓	✓ ✓	✓ -	✓ -	✓ -	✓ -	✓ -	✓ ✗	✗ ✗
iFlaA2 N2	✗ ✗	✗ ✗	✗ ✗	✗ ✗	✗ ✗	✗ ✗	✗ ✗	✗ ✗	✗ ✗
iFlaA2 N3	✓ ✓	✓ -	✓ -	✗ ✗	✗ ✗	✓ -	✓ -	✓ ✗	✓ -
boFlaA2 N1	✓ ✓	✓ ✓	✓ ✗	✗ ✗	✓ -	✓ -	✓ -	✓ ✗	✓ ✗
boFlaA2 N2	✓ ✗	✓ ✓	✓ -	✗ ✗	✓ -	✓ ✗	✓ -	✓ ✗	✓ -
boFlaA2 N3	✓ ✗	✓ ✗	✓ ✗	✗ ✗	✓ ✗	✓ ✗	✓ ✗	✗ ✗	✓ ✗

Tabla 1: Construcciones de FlaA para clonado y expresión *high-throughput*. Arriba, en colores más claros, se muestran las construcciones para la isoforma FlaA1, abajo en colores más oscuros las de FlaA2. Para cada isoforma se utilizaron los genes de 3 especies: *L biflexa* (señaladas en rojos), *L. interrogans* (señaladas en azules) y *L. borgpetersenii* (señaladas en negro). Para cada isoforma de cada especie se diseñaron 3 construcciones distintas (N1, N2 y N3 según se indican en el panel de la izquierda), que se insertaron en 9 vectores (señalados arriba del 1 al 9, según la proteína de fusión que expresan). En cada celda de la tabla, el símbolo a la izquierda se usa para representar los clonados, un ✓ (tick) para las construcciones clonadas y confirmadas, y una ✗ para

los no clonados o los que no se intentaron clonar. Mientras que el simbolo a la derecha representa el resultado de la expresión a pequeña escala, (✓ para aquellas proteínas de fusión que parecen expresarse según los geles SDS-PAGE, un - para aquellas que no se expresan y una ✗ para los que no se intentaron expresar).

Se pusieron a punto aproximaciones de tipo *high-throughput* para transformar los 115 vectores obtenidos y llevar a cabo ensayos de expresión a pequeña escala, probando simultáneamente tres concentraciones de ramnosa (0, 500 µM y 1 mM) para cada construcción (Figura 18). Para FlaA1, en total se expresaron 36 construcciones, de las cuales pareció haber expresión soluble en 18: 3 de *L. biflexa*, 7 de *L. interrogans* y 8 de *L. borgpetersenii*. Para FlaA2, de 54 construcciones obtenidas, se ensayaron las primeras 36, constatándose expresión soluble en 10: 4 de *L. biflexa*, 3 de *L. interrogans* y 3 de *L. borgpetersenii*.



Construcciones FlaA1	kDa	Construcciones FlaA2	kDa	proteina fusion	kDa
biflexa N1	29,60	biflexa N1	21,34	1-Dsbc	27.65
biflexa N2	27,52	biflexa N2	19,38	2-MBP	44.47
biflexa N3	23,33	biflexa N3	14,71	3-CelD	64.98
interrogans N1	25,23	interrogans N1	21,15	4-Trx	16.01
interrogans N2	23,01	interrogans N2	19,30	5-s/fusion	2.48
interrogans N3	20,82	interrogans N3	18,25	6-sumo	15.8
borgpetersenii N1	28,64	borgpetersenii N1	22,91	7-pelB	5.94
borgpetersenii N2	26,21	borgpetersenii N2	20,12	8-GST	29.19
borgpetersenii N3	23,06	borgpetersenii N3	16,51	9-OsmY	27.52
GFP	26.60	TEV	33,5		

Figura 18: Resultados de expresión HT a pequeña escala de ambas isoformas de FlaA. A. Tabla de tamaños esperados para cada construcción y sus compañeros de fusión; sumando ambos se deduce el tamaño de la proteína recombinante esperada. Para cada construcción se realizaron ensayos de expresión a pequeña escala y las fracciones solubles fueron purificadas por IMAC. **B-E.** Geles SDS-PAGE 12% con las eluciones provenientes de IMAC. En los geles C. y E. las muestras fueron posteriormente incubadas con la proteasa TEV. Para identificar las muestras sembradas se utilizó el siguiente código: Las construcciones con FlaA1 y FlaA2 de las distintas especies de *Leptospira* se muestran con el mismo código de color en la Tabla A. y en los geles B-E: FlaA1 de *L. biflexa* en rosado y FlA2 de la misma especie en rojo, FlaA1 de *L.interrogans* en celeste y FlaA2 en azul, y FlaA1 *L. borgpeterseni* en gris y FlaA2 en negro. Cada muestra sembrada se identifica con el número de la proteína de fusión seguido de la construcción de FlaA1 o FlaA2. En los casos en que aparecen tres carriles con las mismas referencias, se trata de las tres concentraciones de Ramnosa testeadas (0, 500 y 1000mM en este orden); y si aparecen dos, son las de menor concentración en el mismo orden. ie. en B. el carril 1 corresponde a la fracción soluble purificada de la proteína de fusión MBP_ FlaA1 de *L.biflexa*, sin incubación con Ramnosa. El MPM utilizado fue NEB #P7703 10-250kDa, con los tamaños de las bandas mostradas en D. NI: muestra purificada a partir de una cepa transformada pero no inducida, y C-: muestra purificada a partir de una cepa transformada

con el vector parental con GFP. Asimismo en los geles, se marcan con recuadros rojos las proteínas de tamaño esperado, con recuadros negros las bandas con tamaño esperado para las proteínas de fusión solas y en recuadros azules, las proteínas que no fueron cortadas por TEV.

Entre los resultados destacados, las fusiones con DsbC para todas las construcciones analizadas de FlaA2 de las tres especies parecen expresarse abundantemente, y luego del corte con TEV se observan bandas nuevas del tamaño esperado para los distintos mutantes de FlaA2. Para FlaA1, solo se obtienen proteínas recombinantes del tamaño esperado para las versiones más cortas (DsbC_FlaA1N3) de *L. interrogans* y *L. borgpetersenii*. Las fusiones con MBP mostraron que para FlaA1 todas las proteínas analizadas se expresan bien, obteniéndose los tamaños esperados después del corte, sin importar el largo de la construcción.

3.2.3.4 – Escalado de la producción recombinante de FlaA1 y FlaA2

A partir de los resultados de expresión a pequeña escala de FlaA2, se decidieron escalar a 4 litros de cultivo los transformantes de tres construcciones: DsbC_FlaA2_N1 de *L. biflexa*, MBP_FlaA2_N1 de *L. interrogans* y DsbC_FlaA2_N1 de *L. borgpetersenii*. Estas proteínas recombinantes fueron purificadas por pasos sucesivos de cromatografía IMAC1, corte con TEV en diálisis, IMAC2 y SEC (Figura 19). Las proteínas recombinantes esperadas parecen salir principalmente en la cola del pico de la primera IMAC, consistente con proteínas que interactúan más específicamente con la columna. En el flowthrough de la segunda IMAC aparecen bandas del tamaño esperado para FlaA2 sola, aunque en el caso de *L. interrogans* la banda es más tenue. En todas las purificaciones, la proteína recombinante está acompañada de bandas de mayor peso molecular, reflejándose en el paso final de purificación por cromatografía de exclusión molecular donde aparecen al menos tres picos discretos.

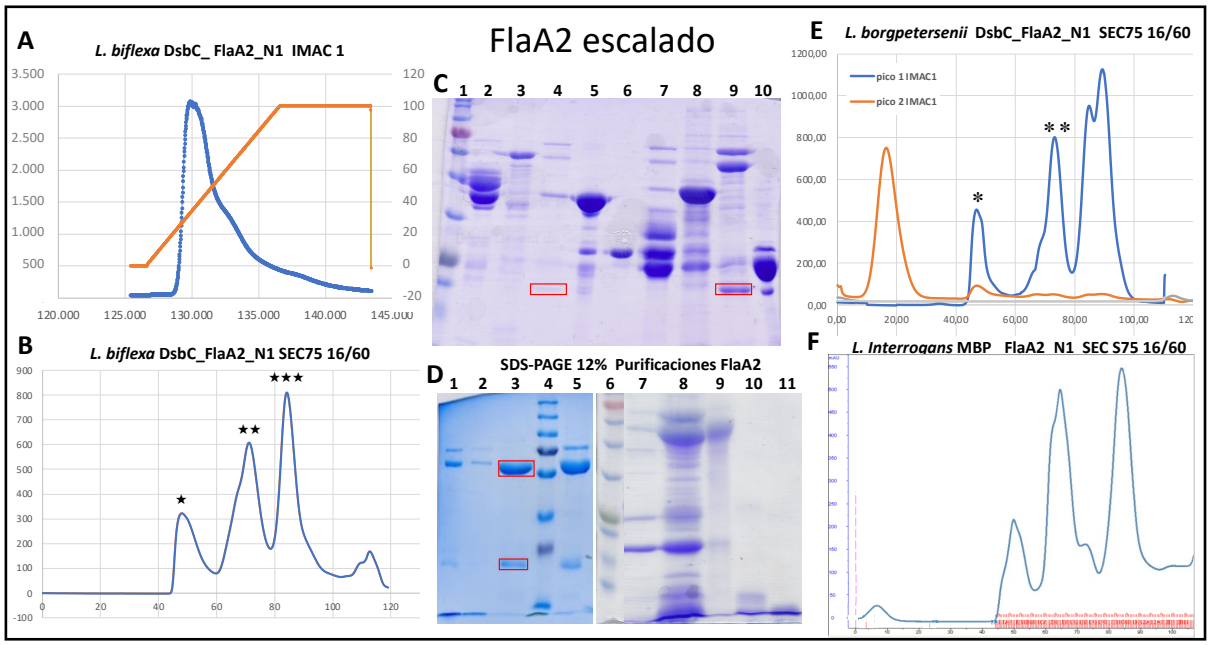


Figura 19: Purificación de FlaA2 recombinante: La sobreexpresión de DsbC_FlaA2_N1 de *L. biflexa* en cepas *Lemo21* de *E.coli* fue inducida con IPTG a 20ºC por 18hs en 4Lts de TB. La fraccion soluble fue purificada por IMAC (columna 5mL); en **A.** se muestra la elución resultante. El eluato fue colectado, dializado *on* con TEV y purificado en una segunda IMAC. El flowthrough se concentró e inyectó en una columna SEC75 16/60 cuyo cromatograma se muestra en **B. C.** SDS-PAGE 12%. Carril 1: MPM Biolabs #P7711S; Carriles 2-4: purificación del mutante N1 de FlaA2 de *L. interrogans* fusionada a MBP. Carril 2: primer fracción de elución de la IMAC 1. Carril 3: cola del pico de elución de la IMAC1. Carril 4: Flowthrough de la IMAC 2. Carril 5: Elución de la segunda IMAC. Carril 6: proteasa TEV sola Carriles 7-10: purificación de del mutante N1 de FlaA2 de *L. biflexa* fusionada a DsbC. Carril 7: primer fracción de elución de IMAC 1 mostrado en **A.** Carril 8: Cola del mismo pico. Carril 9: Flowthrough de la IMAC 2. Carril 10: Fracción eluida de la 2da IMAC **D.** SDS-PAGE 12% mostrando las purificaciones de FlaA2 a partir de de DcbC_FlaA2_N1 de *L. borgpetersenii* (carriles 1-5) y DcbC_FlaA2_N1 de *L. biflexa* (carriles7-11) Carriles 1 y 2: Flowthrough de la IMAC 2; Carriles 3 y 5: picos de SEC75 mostrados en **E.** con los símbolos (✻) y (✻✻) respectivamente, Carriles 4 y 6: MPM Biolabs #P7711S; Carril 7: Flowthrough IMAC 2 de otra purificación de DsbC_FlaA2_N1 de *L. biflexa,* Carril 8: idem 7 con la muestra más concentrada, Carriles 9-11: picos de SEC75 mostrados en **B.** con los símbolos (★), (★ ★) y (★ ★ ★) respectivamente. **E.** Cromatograma SEC 75 16/60 de DcbC_FlaA2_N1 de *L. borgpetersenii.* En la IMAC 1 eluyen 2 picos más discretos que para los casos anteriores, por lo que se separan en los siguientes pasos de purificación. En esta figura se muestran las eluciones SEC superpuestas para ambos purificados por separado. **F.** Cromatografía SEC 75 16/60 de la purificación de MBP_FlaA2_N1 de *L. interrogans* mostrada en **C.** Las dos bandas mostradas en recuadros rojos en **C.** y la de mayor PM mostrada en **D.** fueron identificadas por espectrometría de masas y corresponden a FlaA2, la restante es una chaperona.

El análisis por espectrometría de masas reveló que FlaA2 se purifica en forma soluble pero formando agregados, y el hecho de que copurifique con una chaperona es señal de plegamiento incorrecto. Cromatogramas similares se observaron para el resto de las purificaciones de FlaA2, sugiriendo una situación similar en todos los casos.

Para FlaA1, se adoptó una estrategia diferente escalando a 200 mL las 11 construcciones más prometedoras para comparar rendimientos. Basándose en los resultados, se decidió escalar a 4 litros y purificar las construcciones que expresan: FlaA1_N1 de *L. biflexa* sin fusión, DsbC_FlaA1_N3 y MBP_FlaA1_N1 ambas de *L. interrogans*, posteriormente sumando FlaA1_N1 de *L. borgpetersenii* (Figura 20). Para las primeras tres construcciones, si bien se

obtienen solubles y con buen rendimiento, prácticamente toda la proteína sale en el volumen de exclusión de la cromatografía, formando agregados solubles. Por DLS se confirmó la formación de agregados de aproximadamente 600 kDa que no son reversibles variando el pH o la fuerza iónica. No obstante, las muestras parecen ser monodispersas con índice de polidispersión menor a 0.2 y contienen únicamente dos poblaciones de tamaño definido, indicativo de que se encuentran bien plegadas.

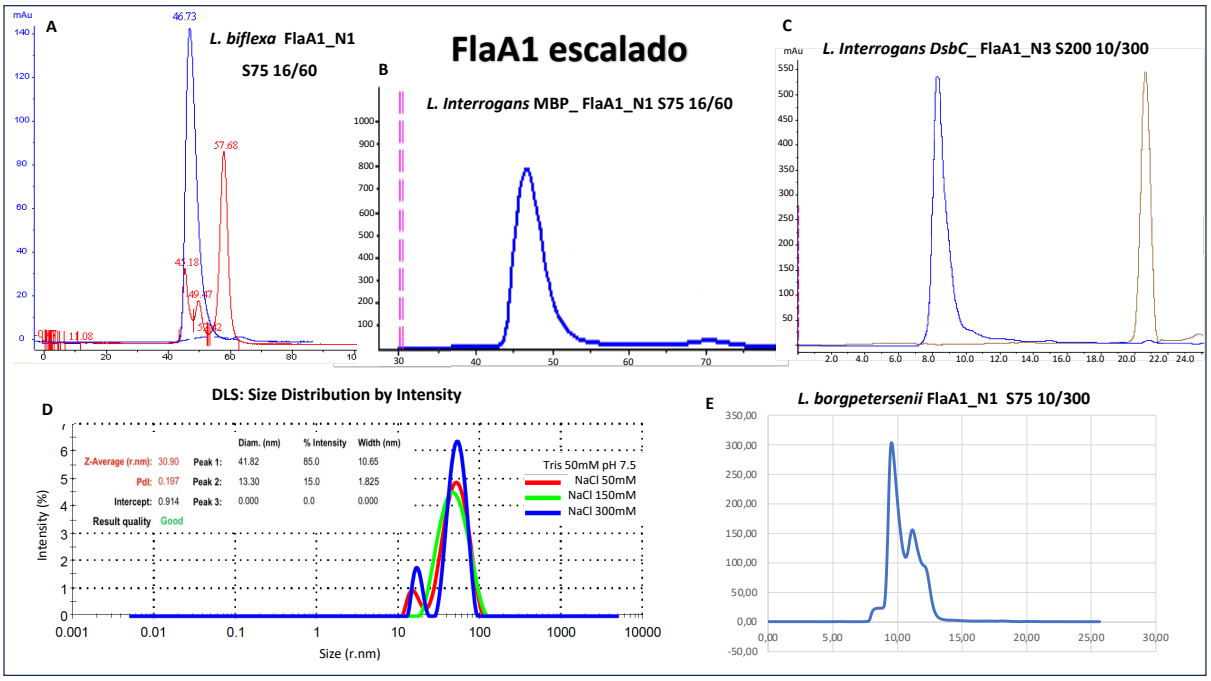


Figura 20: Expresión y purificación de FlaA1 recombinante a mayor escala. Células de *E. coli* Lemo21 transformadas con el plásmido que expresa la proteína de fusión que se indica fueron cultivadas en 4Lts de medio TB e inducidas con IPTG a 20ºC por 18hs. Las fracciones solubles resultantes fueron purificadas por cromatografía de afinidad IMAC, colectadas y dializadas *on* con TEV. Al día siguiente se realizó una segunda IMAC colectando el flowthrough que fue concentrado e inyectado en Cromatografías de exclusión moleular. **A.** Cromatograma de SEC 75 16/60 para la purificación recombinante de FlaA1_N1 de *L. biflexa*, con la curva de calibración de la columna superpuesta para mostrar que la proteína forma agregados que eluyen en el volumen de exclusión de la columna. **B.** Cromatograma de SEC 75 16/60 para la purificación la proteína de fusión MBP_FlaA1_N1 de *L. interrogans*. **C.** Cromatograma de SEC: S200 10/300, de la proteína de fusión DsbC_FlaA1_N3 de *L. interrogans.* Esta SEC analítica fue realizada inyectando la muestra colectada a partir de la SEC75 16/60, y como forma de comprobar la formación de agregados solubles de gran tamaño que eluyen en el volumen de exclusión tanto de una SEC75 como de una S200. **D.** Tabla de datos y distribución de Intensidad vs. Tamaño obtenido del análisis por dispersión dinámica de luz visible (DLS) de FlaA1_N1 de *L. biflexa* mostrando una muestra poco polidispersa con 2 poblaciones distintas y homogeneas, que no se ven notoriamente alteradas ante cambios en la fuerza iónica. **E.** Cromatograma de SEC 75 10/300 para la purificación recombinante de FlaA1_N1 de *L. borgpetersenii*. Esta cromatografía analítica se obtuvo al inyectar una muestra colectada el día anterior como el pico no excluído de menor tamaño de la SEC75 26/60, mostrando que se establece un equilibrio en la formación de estos agregados.

Para FlaA1_N1 de *L. borgpetersenii*, si bien también se observan dos poblaciones en la purificación por SEC, la proporción de proteína no excluida de la columna es mayor. Esta fracción pudo ser colectada por separado y reanalizada al día siguiente por SEC analítica,

observándose que se establece un equilibrio con FlaA1 en distintos estados de agregación. La población más grande en tamaño correspondería a los agregados solubles, que se vuelven a formar y se observan como un pequeño hombro al inicio del cromatograma, y las de tamaños más pequeños podrían corresponder al dímero y monómero.

3.2.3.5 – Análisis de estabilidad de FlaA1 con nano-fluorimetría diferencial de barrido

Para optimizar las purificaciones preparativas de FlaA1 recombinante, promoviendo su solubilidad y evitando la formación de agregados, realizamos un rastreo de condiciones que aumentaran la termoestabilidad para el mutante N1 de FlaA1 (Δ49-Δ301-316) de *L. borgpetersenii*, que tiene la mayor proporción de monómero soluble obtenido. La nano-fluorimetría diferencial de barrido permite medir la fluorescencia intrínseca de las proteínas dependiente de sus residuos aromáticos, principalmente triptófanos. Cuando las proteínas se desnaturalizan, estos residuos que normalmente se pliegan dentro del núcleo hidrofóbico quedan expuestos, aumentando el cociente 350/330 nm de fluorescencia emitida.

Comenzamos comparando la estabilidad de FlaA1 en presencia de concentraciones crecientes de $CaCl_2$ (Figura 21). La fundamentación se basó en predicciones estructurales que sugerían un plegamiento propio de los módulos de unión a carbohidratos (dominios CBM), donde varias familias unen cationes de calcio contribuyendo el metal a estabilizar la estructura y función. Efectivamente, la temperatura de fusión presentó un corrimiento de 19°C cuando la muestra contenía 10 mM $CaCl_2$ comparada con la misma muestra en ausencia de calcio. Este gran aumento en la estabilidad se mantiene al seguir aumentando la concentración de $CaCl_2$, concluyendo que el catión calcio es un ligando necesario para el correcto plegamiento de FlaA1. En una siguiente etapa, se evaluó la termoestabilidad de la proteína con $CaCl_2$, en solcuiones amortiguadoras de diferente pH y fuerza iónica. Los resultados indican que la proteína es más estable en los amortiguadores HEPES pH 7 y Tris pH 8, y la fuerza iónica óptima se sitúa en torno a 150 mM NaCl. También se probaron $LiSO_4$, LiCl, N-acetil-glucosamina y ácido N-acetil-murámico, con y sin calcio. A excepción del ácido N-acetil-murámico, se observan diferencias de estabilidad únicamente cuando se comparan las soluciones con y sin $CaCl_2$. Los efectos observados con ácido N-acetil-murámico resultaron ser los mismos que los obtenidos para la proteína TEV utilizada como control negativo, concluyendo que los cambios en fluorescencia son provocados por altas concentraciones del compuesto sobre cualquier proteína.

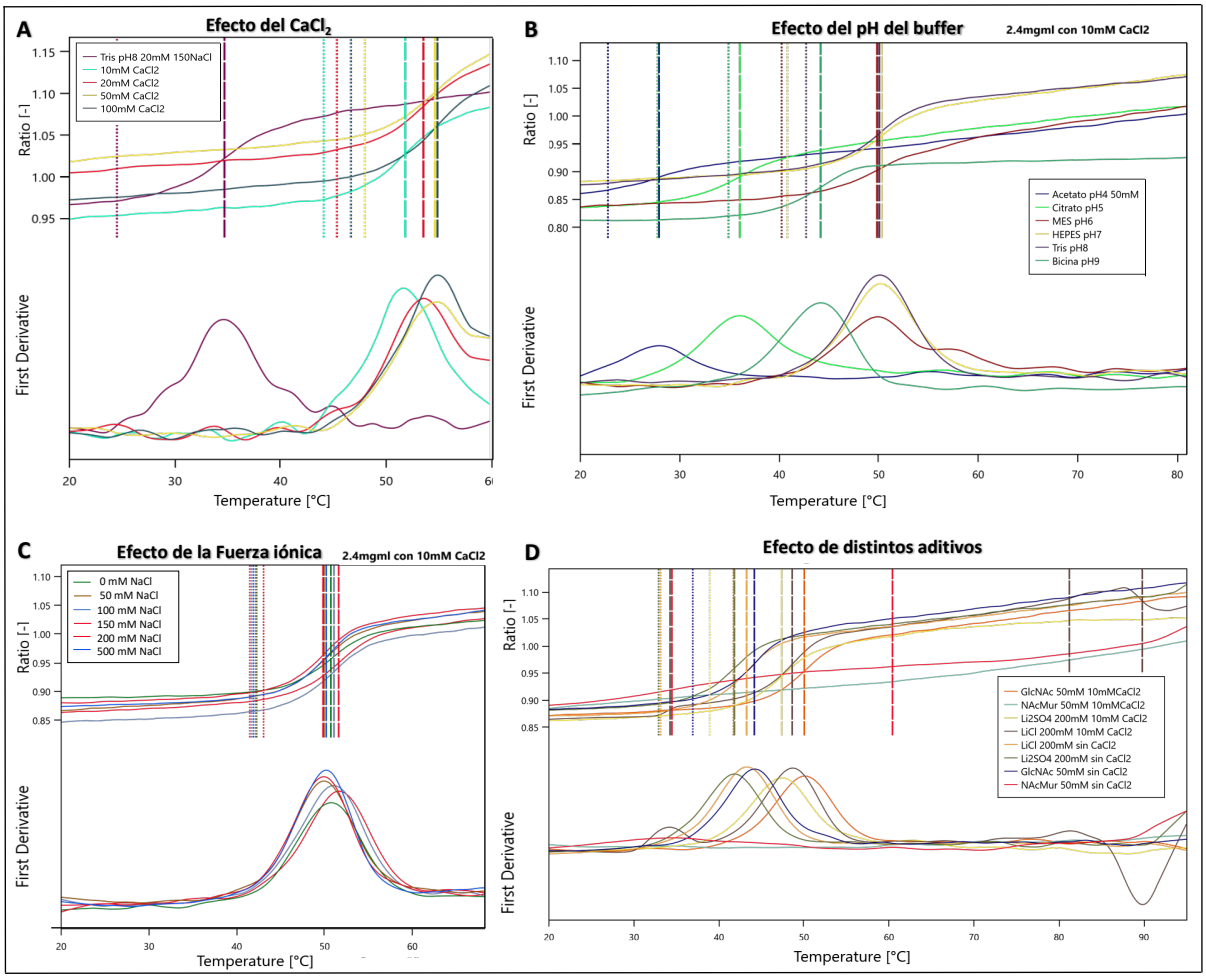


Figura 21: Análisis de estabilidad del mutante N1 de FlaA1 de *L. borgpetersenii* por nanoDSF. A. Efecto de concentraciones crecientes de $CaCl_2$ en la termoestabilidad de la proteína. **B.** Efecto de distintas soluciones amortiguadores con diferente pH en la estabilidad de FlaA1. **C.** Rastreo de las mejores condiciones de fuerza iónica (a distintas concentraciones de NaCl). **D.** Rastreo de posibles ligandos de FlaA1. Se analiza en particular el efecto de distintas sales de litio así como de azúcares presentes en el peptidoglicano: ácido N-acetil murámico (NAcMur), y N-acetil glucosamina (GlcNAc), con y sin $CaCl_2$.

3.2.3.6 – Optimización de la purificación de FlaA1 recombinante

Habiendo concluido que el calcio es un posible ligando de FlaA1, se agregó $CaCl_2$ a lo largo de toda la purificación, desde el inicio del cultivo. Los protocolos fueron optimizados agregando 10 mM $CaCl_2$. El principal cambio observado fue el aumento de solubilidad de la proteína, disminuyendo su precipitación entre pasos sucesivos de purificación. Este efecto acompañó cambios en el estado de agregación de la proteína finalmente obtenida. Si bien siguen apareciendo dos poblaciones mayoritarias (Figura 22b), ambos picos se encuentran en el rango de tamaños 20-60 kDa, no observándose agregados supramoleculares, y comenzaron a distinguirse mejor los picos correspondientes al monómero y dímero. El análisis por espectrometría de masas de la banda principal identifica únicamente la proteína esperada.

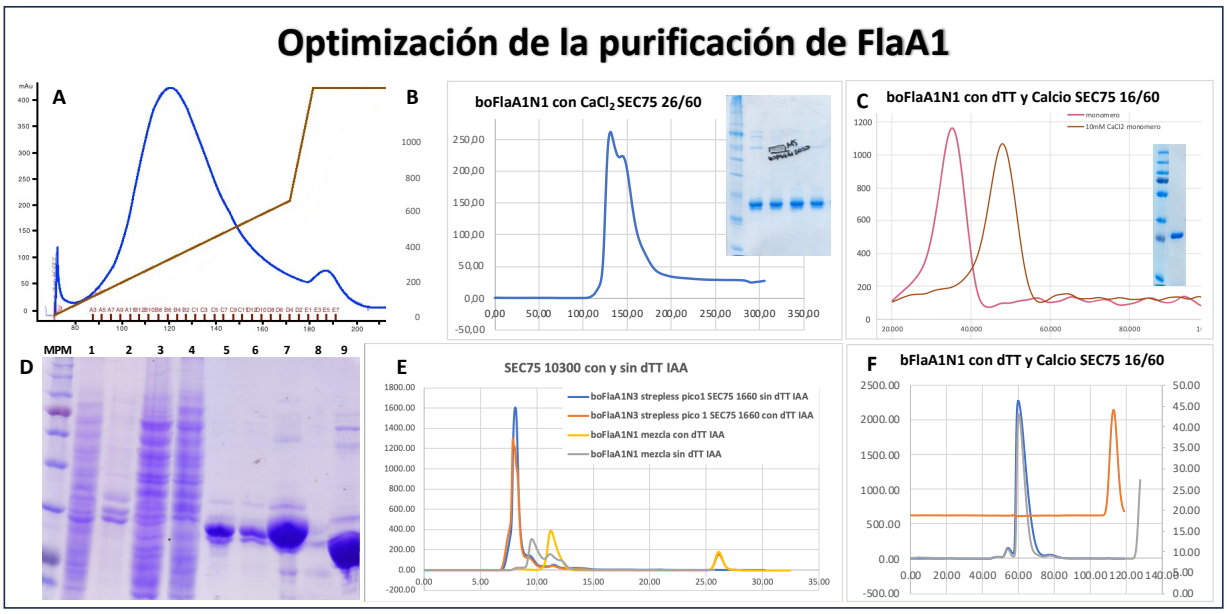


Figura 22: Optimización de la purificación de FlaA1 recombinante. A. Cromatograma de la IMAC (5mL), el primer paso de purificación de la proteína FlaA1_N1 de *L. borgpetersenii* (siempre con 10mM $CaCl_2$). **B.** Cromatograma del último paso de purificación de la misma proteína por SEC75 26/60. En la esquina dererecha puede observarse el gel SDS-PAGE 12% asociado a esta purificación. Carril 1: MPM Biolabs #P7711S prestained; Carriles 2-5: Fracciones sucesivas colectadas de los picos en el cromatograma SEC. Puede observarse la aparición bandas de mucho mayor peso molecular que el correspondiente al monómero de boFlaA1N1 (28,6 KDa). La banda indicada en el carril 3 fue cortada y analizada por espectrometría de masas confirmando la identidad de esta proteína. **C.** Idem B. pero comparando las purificaciónes en presencia de DTT por SEC 75 16/60 con y sin calcio. A la derecha, SDS-PAGE 12% asociado a la purificación con DTT y $CaCl_2$. carril 1: MPM Biolabs #P7711S prestained; carril 2: muestra obtenida del único pico de la SEC 75 16/60. **D.** SDS-PAGE 12% siguiendo los primeros pasos de la pruificación de boFlaA1N. carril 1: MPM, 1. Fracción total del cultivo inducido, 2. Fracción insoluble del cultivo inducido, 3. Fracción soluble del mismo cultivo, 4. *Flowthrough* con las proteínas que no quedan retenidas en la primer IMAC, 5. Fracción B2 tomada del cromatograma en A., 6. Fracción D8 tomada del cromatograma en A. en la cola del pico eluído, 7. muestra de la fracción total eluída finalmente colectada de esta IMAC, 8. *Flowthrough* de la segunda IMAC luego del corte con TEV, 9. elución de la segunda IMAC. **E.** Cromatografías analíticas, SEC 75 10/300 comparando las construcciones boFlaA1_N1 y _N3, con y sin tratamiento con DTT e iodoacetamida; **F.** Cromatografía SEC75 16/60 mostrando el paso final de purificación de FlaA1_N1 de *L. biflexa* con el método de purificación optimizado.

El análisis por SDS-PAGE reveló varias banas minoritarias de mayor peso molecular al esperado para el monomero (28,64 kDa) o dimero de FlaA1N1 de *L. borgpetersenii,* por lo que se consideró la posible formación de puentes disulfuro. El agregado de DTT y posterior incubación con iodoacetamida eliminaron por completo dichas bandas. Concluimos que uno o más de los tres residuos de cisteína (Cys185, Cys289 y Cys307) ubicados cerca del extremo C-terminal participan en la formación de puentes disulfuro, los cuales conducen a la pérdida de solubilidad. Se optimizó el protocolo de purificación incluyendo β-mercaptoetanol en los primeros pasos e incubando con DTT más iodoacetamida antes de la cromatografía de exclusión molecular. Esto provoca la reducción irreversible de las cisteínas, resultando en un único pico correspondiente al monómero en la SEC. Se obtiene entonces esta proteína en escala preparativa, soluble y homogénea para cristalización.

Para obtener versiones proteicas alternativas que aumentaran las probabilidades de cristalización, se expresaron y purificaron mediante los protocolos análogos las versiones FlaA1_N3 de *L. borgpetersenii* y FlaA1_N1 de *L. biflexa*. El efecto de la reducción permanente de las cisteínas observado para la versión N1, donde toda la proteína pasa al estado monomérico, no se mantiene para FlaA1_N3 donde prácticamente toda la proteína eluye en el volumen de exclusión como agregados solubles. Sin embargo, el protocolo optimizado sí funciona para aumentar la solubilidad y disminuir la formación de agregados solubles en FlaA1_N1 de *L. biflexa*.

3.2.3.7 – Ensayos de cristalización de FlaA1

Se rastrearon condiciones de cristalogénesis con 32 placas de cristalización con la proteína FlaA1_N1 de *L. borgpetersenii*, analizando 96 condiciones en cada placa y ensayando distintas concentraciones de proteína. Solo se obtuvieron fibras, cristales de $Ca(OH)_2$ y sales. Tampoco se obtuvieron cristales en las 12 placas de rastreo realizadas con FlaA1_N1 de *L. biflexa*. Para obtener cristales, fue necesario eliminar la etiqueta Strep fusionada en el extremo C-terminal, como se ha reportado en la literatura para otras proteínas de fusión[156]. Se diseñó un cebador que codifica un codón stop al final de la secuencia codificante. La versión N1 sin etiqueta fue expresada y purificada a homogeneidad según el protocolo optimizado, y se realizaron 16 placas rastreando un total de 1,536 condiciones de cristalización.

Se obtuvieron cristales de proteína en seis condiciones diferentes (Figura 23), de las cuales una se descartó por su apariencia y formación tardía (Figura 23E) y otra no se pudo reproducir (Figura 23D). Las dos condiciones más prometedoras fueron: 0.1M Hepes pH 7, 10% m/v PEG 6000 y 1.0M LiCl, donde inicialmente se obtuvieron agujas muy finas que se optimizaron variando pH, concentración de PEG y LiCl, resultando en barritas birrefringentes y 0.1M cacodilato de sodio pH 6.5, 1.0M sulfato de amonio, donde se obtuvieron cristales cuadrados que aparecían entre los 15-20 días, optimizados mediante variación de pH, concentración de $(NH_4)_2SO_4$ y técnicas de micro-sembrado para disminuir el tiempo de cristalización y obtener cristales más grandes.

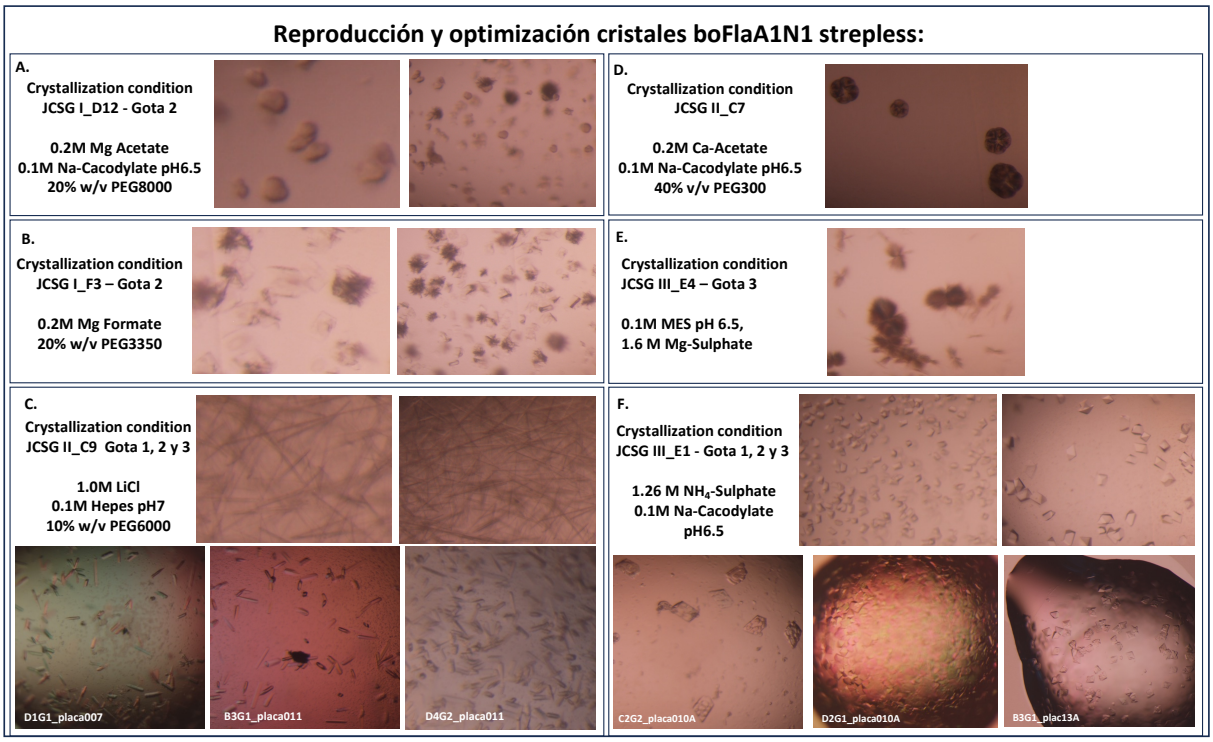


Figura 23: Cristalización de boFlaAN1_strepless. Este rastreo de condiciones se realizó con la versión strepless de la proteína purificada con DTT/iodoacetamida y calcio. La proteína fue concentrada a ~23 mg/mL y su buffer final es 20mM Tris pH8, 150mM NaCl sin aditivos en la gota 1; con 10mM citrato de sodio en la gota 2; y con 10mM N-Acetil Glucosamina en la gota 3. Se muestran los ensayos de cristalización con los Kits comerciales JCSG I, II, III y IV donde se obtuvieron *hits* en seis condiciones diferentes. En cada caso se muestran a la izquierda las condiciones en que se obtuvieron; en medio los primeros cristales observados; y a la derecha los cristales que pudieron ser reproducidos o mejorados. **A.** Esférulas sin birrefringencia en las tres gotas del ensayo de rastreo (con y sin aditivos) que pudieron ser reproducidas pero no optimizadas. **B.** Cristales con forma de barras/agujas que se obtienen en las tres gotas y tienen poca birrefringencia. Pudieron ser reproducidas. **C.** Agujas birrefringentes son obtenidas en las tres gotas de esta condición, se observa mucho menos cantidad de cristales en la gota 3 y son más chicas en la gota 1 que en la 2. Pudieron ser reproducidas y optimizadas, obteniendo los cristales en forma de barritas que se muestran abajo (proteína a ~32,7 mg/mL en reservorio 0.1M Hepes pH7.0, 4% m/v PEG 6000, 1.5M LiCl (izquierda); ~32,7 mg/mL en reservorio 0.1M Hepes pH7.0, 3% m/v, PEG 6000, 2.0M LiCl (medio); ~32,5 mg/mL con 2mM $CaCl_2$ en reservorio 0.1M Hepes pH7.0, 5% m/v, PEG 6000, 2.5M LiCl (derecha)). **D.** Esférulas birrefringentes aparecen en las gotas 1 y 2 de esta condición, no pudieron ser reproducidas. **E.** Agujas formando racimos que aparecen en la tres gotas de forma tardía, luego del mes del ensayo. No se intentaron reproducir. **F.** Cuadraditos inicialmente bastante planos salen en las tres gotas del ensayo también de forma tardía, inicialmente a los 15-20 días. Pudieron ser reproducidos y optimizados, como se muestra en las figuras de abajo (proteína a 50 mg/mL en reservorio 0.1M cacodilato pH6.5, 1.0M $(NH_4)_2S0_4$ (izquierda); proteína a 32,5 mg/mL en reservorios: 0.1M cacodilato pH 7.0, $(NH_4)_2S0_4$ 1M (medio) y 0.1M cacodilato pH 7.5, $(NH_4)_2S0_4$ 1M (derecha).

3.2.3.8 – Difracción y estructura cristalográfica de FlaA1

Los cristales del mutante N1 de FlaA1 (Δ49-Δ301-316) de *L. borgpetersenii* fueron analizados mediante difracción de rayos X. Las barritas obtenidas en 0.1M Hepes pH 7, 10% m/v PEG 6000 y 1.0M LiCl difractan con ~3 Å de resolución. La estructura, sin embargo fue resuelta a partir de los cristales cuadrados optimizados en la condición 0.1M cacodilato de sodio pH 6.5, 1.0M sulfato de amonio. Los datos finales fueron colectados en el sincrotrón Diamond Light

Source (Oxford) y las imagenes de difracción procesadas con autoProc[157], revelando que los cristales crecieron en el grupo de espacio ortorrómbico $P2_12_12_1$ (Figura 24A).
Al momento de resolver la estructura se estaba produciendo la revolución en la predicción de estructuras 3D de proteínas a partir de la secuencia. Si bien AlphaFold aún no había sido lanzado, la predicción de FlaA1 con RoseTTAfold[158] fue suficiente para obtener un modelo inicial con el cual realizar reemplazo molecular. Se encontraron tres moléculas de FlaA1 por unidad asimétrica (Figura 24C). El modelo atómico se generó utilizando Phaser[159] y se refinó utilizando Phenix.refine[160], iterando ciclos de refinamiento y reconstrucción en Coot 0.9[161]. La estructura final resuelta fue depositada en la Protein Data Bank con PDB ID 9CYN.

Resolución de la estructura de FlaA1 de *L. borgpetersenii*

A Estadísticas de colecta y procesamiento de datos de difracción de rayos X de FlaA1

Spacegroup			P212121
Cell parameters			81.417 105.776 108.734 90.0 90.0 90.0
Wavelength [A]			0.97951
	Overall	**Inner Shell**	**Outer Shell**
Low resolution limit	75.819	75.819	1.848
High resolution limit	1.816	4.930	1.816
Rmerge (all I+ & I-)	0.106	0.042	5.769
Rmeas (all I+ & I-)	0.110	0.044	5.987
Rpim (all I+ & I-)	0.030	0.012	1.596
Total number of observations	1173515	58174	59257
Total number unique	85228	4538	4238
Mean(I)/sd(I)	13.1	47.5	0.5
Completeness	100.0	100.0	100.0
Multiplicity	13.8	12.8	14.0
CC(1/2)	0.998	0.997	0.375

B Estadísticas de refinamiento del modelo cristalográfico de FlaA1 de *L. borgpetersenii.*

Resolution range	75.82 - 1.816 (1.84 - 1.82)**
Unique reflections used in refinement	85179 (2786)
Reflections used for R-free	4230 (146)
R-work	0.1842 (0.4334)
R-free	0.2151 (0.4366)
Number of non-hydrogen atoms	5528
macromolecules	5073
ligands	38
solvent	417
Protein residues	643
RMS (bonds)	0.017
RMS (angles)	1.32
Ramachandran favored (%)	96.85
Ramachandran allowed (%)	3.15
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	1.13
Clashscore	2.92
Average B-factor	50.15
macromolecules	49.59
ligands	70.54
solvent	55.09

C

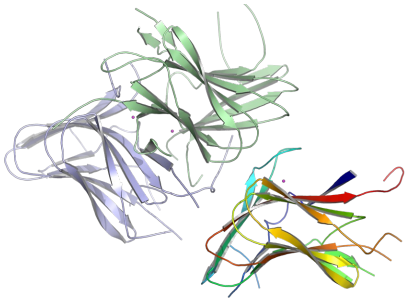


Figura 24: Resolución de la estructura de FlaA1 de *L. borgpetersenii* por difracción de rayos X. A. Tabla mostrando las estadísticas de la colecta y procesamiento de datos de difracción de rayos X para FlaA1 (Δ49-Δ301-316) de *L. borgpetersenii.* **B.** Tabla con las estadisticas de refinamiento final del modelo cristalográfico de FlaA1 depositado en la PDB con id 9CYN. ** Las estadísticas para la capa de mayor resolución se muestran entre paréntesis. **C.** La unidad asimétrica de los cristales resueltos tienen tres copias de FlaA que se muestran en verde, violeta y la tercera coloreada en rampa de colores de azul (extremo N-terminal) a rojo (C-terminal).

La estructura cristalográfica final de FlaA1 de *L. borgpetersenii* se resolvió a 1.82 Å de resolución. Esta estructura representa la primera determinación experimental de la

arquitectura tridimensional de una proteína FlaA, proporcionando información estructural fundamental para comprender la organización del filamento flagelar de *Leptospira*.

Los mapas de densidad electrónica refinados para la forma $P2_12_12_1$ fueron claros y permitieron un rastreo seguro desde el inicio del constructo en Met50 N-terminal, hasta los aminoácidos 278 (cadena A), 281 (cadena B) o 286 (cadena C), con los ~30 aminoácidos de los segmentos C-terminales y algunos bucles desordenados indetectables en los mapas de densidad electrónica. La estuctura de FlaA1 se pliega formando un sándwich β cóncavo (Figura 25A), con una lámina antiparalela (hebras β2-β6-β13-β10-β11) del lado cóncavo y dos láminas antiparalelas del lado convexo formando en conjunto una superficie triangular de tres lados. Las dos hojas que forman el lado convexo del sándwich son de cuatro hebras (β1-β15-β9-β12 y β3-β4-β5-β14) y la que forma el lado cóncavo de cinco. Uniendo las dos laminas antiparalelas del lado convexo, encontramos un átomo de Calcio(Figura 25B). Este catión se encuentra en una esfera de coordinación octaédrica constituida por dos átomos de oxígeno de la cadena principal de los residuos Ser59 y Lys122, dos átomos de oxigeno de la cadena lateral de los grupos carboxilato de Asp61 y Asp264, y los oxígenos de dos moléculas de agua. La posición del metal Ca^{2+} une la punta de las largas hebras β5-β14 (a través de Lys122 y Asp264), con el bucle que conecta las hebras β1 y β2 (a través de Ser59 y Asp61) manteniendolas en posición en la porción bisagra de la cara convexa del sándwich β. Esta participación del calcio en el mantenimiento de la estructura 3D de FalA1 explica por qué para aumentar la solubilidad de las isoformas recombinantes de FlaA fue necesario agregar de Cloruro de Calcio durante la purificación.

El plegamiento de FlaA1 es característico de módulos de unión a carbohidratos(CBM) dependiente de Ca^{2+}, validando las predicciones bioinformáticas iniciales que sugirieron su similitud. Aunque estructuralmente homóloga a los dominios CBM, FlaA1 es una proteína de dominio único por lo que no forma parte de una entidad mayor con actividad enzimática activa en carbohidratos como otros CBM. Pertenece a la extensa clase general de lectinas y basandonos en ello anticipamos que unirá específicamente carbohidratos en el contexto del filamento flagelar. La superposicion estructural de FlaA1 con CBMs resueltos en complejo con ligandos de carbohidratos afines nos permitió identificar los sitios más probables a través de los cuales FlaA1 podría unirse a ligandos de azúcar: en la cara cóncava del sándwich β (Figura 25D) y/o en su superficie β3-β4-β5-β14 (Figura 25C).

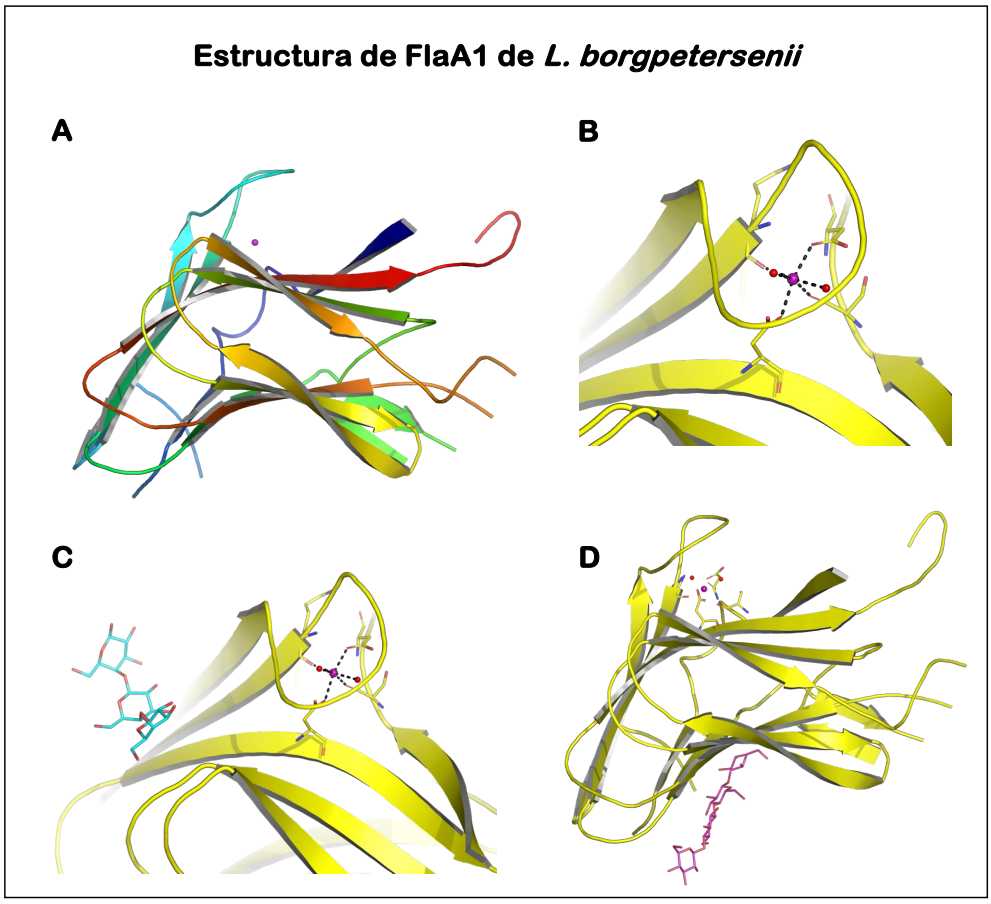


Figura 25: Estructura resuelta de FlaA1 de *L. borgpetersenii*. A. Estructura de FlaA1 (Δ49-Δ301-316) de *L. borgpetersenii,* coloreada en una rampa de colores de azul (extremo N-terminal) a rojo (C-terminal). Su plegamiento contiene tres hojas β que se pliegan formando un sándwich β cóncavo. **B.** Región aumentada en la porción bisagra de la cara convexa del sándwich β mostrando la coordinación de la proteína con el cation calcio. **C.** Región aumentada mostrando un posible sitio de unión a carbohidrato, sobre la superficie de la hoja β3-β4-β5-β14. **D.** Estructura de FlaA1 mostrando el otro posible sitio de unión a carbohidratos en la hoja β del lado cóncavo del sándwich.

3.3 – Interacción entre proteínas flagelares e identificación de nuevos componentes del filamento

Con el fin de dilucidar la arquitectura del filamento de *Leptospira*, nos propusimos caracterizar las interacciones que establecen sus componentes conocidos (FlaB, FlaA1, FlaA2, FcpA y FcpB), tanto entre ellos como eventualmente con otras proteínas aún no identificadas que pudieran constituir nuevos componentes flagelares. Para ello se llevaron a cabo varias estrategias complementarias que se describen a continuación.

3.3.1 – Interacción *in vitro* entre proteínas flagelares conocidas: FcpA y FcpB

Comenzamos evaluando la interacción entre FcpA y FcpB. Nos enfocamos en estas dos proteínas no solo porque son proteínas flagelares únicas del género *Leptospira*, sino porque la ausencia de cualquiera de ellas resulta en fenotipos de motilidad traslacional reducida y

pérdida de la morfología en gancho o espiralada típicamente observada en los extremos de las células de *Leptospira*. Las cepas knockout *fcpA*- presentan el fenotipo más drástico, con pérdida completa de motilidad traslacional y virulencia, así como menor diámetro de sus filamentos flagelares.

Previamente se había observado por microscopía electrónica que los anticuerpos conjugados a oro contra ambas proteínas colocalizan en la superficie de los filamentos superenrollados purificados[105]. Los filamentos *fcpA*$^-$ no solo pierden la banda de 36 kDa esperada para FcpA –detectable por SDS-PAGE y por Western blot–, sino que también desaparece una banda por debajo de FlaB1 que corresponde a FcpB. Esto no sucede con la cepa knockout *fcpB*-, cuyos filamentos pierden únicamente FcpB y siguen presentando FcpA en cantidades similares a las de la cepa salvaje[137]. Estas observaciones nos permitieron hipotetizar que FcpA y FcpB interactúan directa o indirectamente entre sí, y muy posiblemente que FcpA es reclutada al ensamblaje supramolecular primero durante la biogénesis del filamento flagelar *in vivo*.

Evaluamos la posibilidad de una interacción directa entre FcpA y FcpB utilizando cromatografía analítica de exclusión por tamaño, comparando los perfiles cromatográficos de FcpA y FcpB recombinantes solas versus la mezcla entre ambas (Figura 26). Los resultados no permitieron detectar una interacción directa entre ambas proteínas *in vitro* en las condiciones del ensayo. No podemos excluir que la constante de afinidad *in vitro* con las dos proteínas libres en solución no sea muy elevada, y que por ello no podamos detectar el complejo por cromatografía de exclusión por tamaño (por ejemplo, si la constante de disociación fuera del orden de ~50-100 µM o mayor).

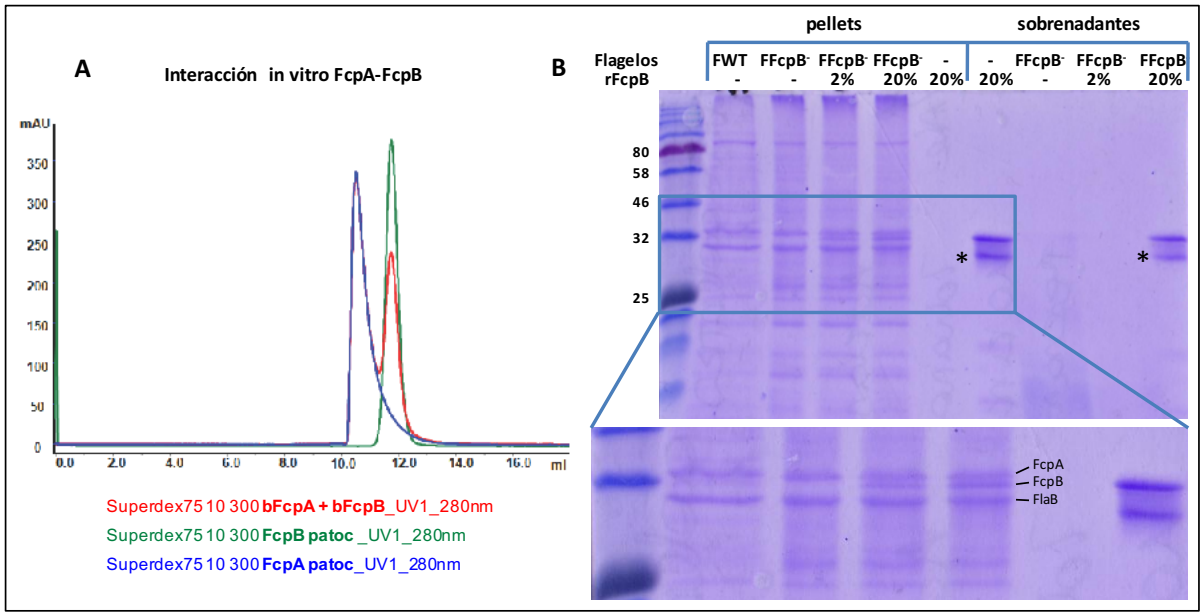


Figura 26: Ensayos de interacción entre componentes del filamento de *L. biflexa*. A) Ensayo de interacción *in vitro* entre bFcpA y bFcpB por cromatografía de exclusión por tamaño (SEC). **B)** Ensayo de asociación por *pull-*

down de filamentos flagelares purificados de la cepa *knockout fcpB-* de *L. biflexa* con bFcpB recombinante (rFcpB). Luego de la incubación, los filamentos reconstituidos fueron separados de la proteína rFcpB soluble por ultracentrifugación. En los carriles 3-6 se muestran los precipitados de esta centrifugación, en los carriles 7-10 los sobrenadantes concentrados por TCA. Según las medidas de proteómica cuantitativa, FcpB representa un 2% de las proteínas totales del filamento. Las muestras con 2% rFcpB contienen suficiente bFcpB para la unión (1X), mientras que las muestras con 20% rFcpB exceden 10 veces la cantidad normalmente presente (10X). Como control se incluye una muestra sin filamentos donde la proteína aparece solo en el sobrenadante. El control negativo sin bFcpB muestra la misma cantidad de filamento purificado. Para comparación se incluyen filamentos purificados de la cepa salvaje de *L. biflexa*. * indica una banda inespecífica que aparece solo en sobrenadantes que contienen rFcpB.

Considerando que la incorporación de FcpA en el filamento no requiere de FcpB, quizás la afinidad entre ambas aumenta significativamente cuando FcpA se encuentra polimerizada en el flagelo. Para probar esto analizamos la asociación de FcpB recombinante libre a filamentos flagelares purificados a partir de la cepa mutante *fcpB*$^{-}$ de *L. biflexa*, que verificamos contiene FcpA. Cuando comparamos los precipitados de filamentos provenientes de la cepa salvaje con el knockout de *fcpB*$^{-}$, observamos que se pierde una banda del tamaño esperado para FcpB, ubicada entre las bandas correspondientes a FcpA y a las FlaB (Figura 26B). Esta banda se recupera al agregar la proteína recombinante FcpB de la misma cepa, tanto en la cantidad que esperamos esté presente en el filamento como 10 veces en exceso.

Agregar FcpB recombinante en exceso no aumenta la intensidad de la banda, es decir, la proporción de FcpB en el filamento se mantiene, pero sí se observa el exceso en el sobrenadante que contiene las proteínas no unidas al filamento. Esto no ocurre cuando FcpB se agrega solo en cantidad suficiente para unirse al filamento; toda la proteína recombinante agregada al inicio precipita con el filamento, mostrando alta afinidad por el mismo. Nuestros resultados preliminares indican que FcpB sí interacciona con FcpA, pero solo en el contexto del filamento ensamblado.

3.3.2 – Identificación de nuevas proteínas flagelares utilizando espectrometría de masas

En paralelo, para tener un panorama global de las proteínas que componen el filamento, recurrimos a técnicas de espectrometría de masas. Siguiendo una estrategia *top-down*, se cultivó *L. biflexa* salvaje hasta fase logarítmica y sus filamentos flagelares fueron purificados según el protocolo establecido. Los filamentos fueron digeridos con tripsina y los péptidos resultantes purificados y analizados por LC-MS/MS en la Unidad de Bioquímica y Proteómica Analíticas (UByPA, Instituto Pasteur Montevideo).

En el proteoma de las preparaciones flagelares encontramos todas las proteínas esperadas: las cuatro isoformas de FlaB, las dos de FlaA, FcpA y FcpB, junto a 600-1300 proteínas adicionales. Estos resultados comprueban lo observado por SDS-PAGE: nuestras

purificaciones flagelares son enriquecimientos periplásmicos que contienen un gran número de proteínas del periplasma además de componentes flagelares. Gran parte de este sub-proteoma incluye proteínas que no han sido identificadas aún.

Para discriminar proteínas del filamento flagelar de otros componentes periplásmicos, ensayamos distintos protocolos de purificación para optimizar la pureza de las preparaciones. Se agregó un paso de ultracentrifugación en gradiente de CsCl, sin embargo, la cantidad de proteína total recuperada disminuyó drásticamente sin observarse grandes cambios en el perifil de proteínas por SDS-PAGE. Se probó sustituir el paso de concentración por PEG con sulfato de amonio en su lugar. Si bien no logramos aumentar la pureza de las preparaciones, nuestros colaboradores pudieron comprobar por microscopía electrónica de transmisión que usando el protocolo estándar los filamentos estaban presentes e intactos, observándose incluso diferencias ultraestructurales entre los distintos mutantes *knockout* de las proteínas flagelares conocidas.

Para identificar nuevos componentes flagelares, decidimos entonces realizar análisis comparativos de los proteomas obtenidos por cromatografía líquida acoplada a espectrometría de masas, analizando filamentos purificados de la cepa salvaje *L. biflexa* versus mutantes *knockout* de proteínas del filamento ya conocidas. Filamentos purificados de cepas salvaje, *fcpA*- y *flaA2*- fueron cuantificados por densitometría en gel SDS-PAGE, para luego sembrar cantidades equivalentes en geles SDS-PAGE 12%, recuperarlos en gel y digerirlos con tripsina. El análisis por triplicado de los péptidos resultantes por LC-MS/MS permitió identificar aquellas proteínas exclusivamente encontradas en filamentos salvajes y no en las mutantes, así como proteínas específicas de los mutantes *fcpA*$^-$ y *flaA2*$^-$. Los filamentos de la cepa *flaA2*$^-$ tienen un mayor número de proteínas que aquellos de *fcpA*$^-$, y ambos mutantes contienen a su vez más proteínas que las purificaciones provenientes de la cepa salvaje (725). Si bien alguna de estas proteínas podría estar verdaderamente aumentada en los mutantes, indicando por ejemplo un reemplazo de función, es más probable que las purificaciones de la cepa salvaje tengan un mayor nivel de pureza. En este escenario, las 24 proteínas presentes exclusivamente en salvaje versus *fcpA*$^-$ y las 8 presentes en salvaje versus *flaA2*$^-$ tienen altas probabilidades de corresponder a proteínas flagelares.

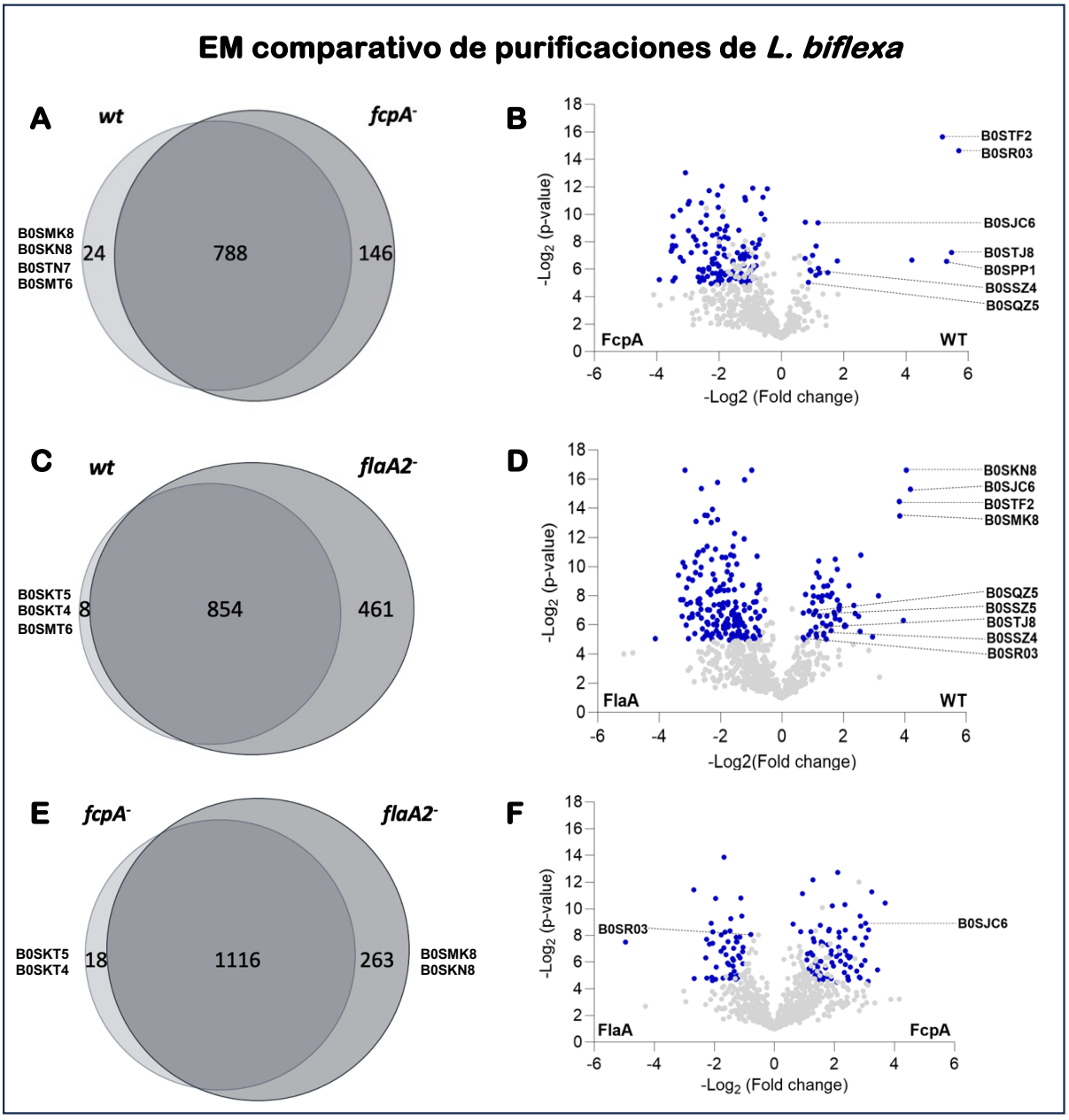


Figura 27: Espectrometría de masas comparativa entre purificaciones flagelares de cepas salvaje, *fcpA*- y *flaA2*- de *L. biflexa*. A) Proteínas identificadas exclusivamente en salvaje vs *fcpA*-. **B)** Gráfico de volcán mostrando las proteínas diferencialmente expresadas entre salvaje vs *fcpA*-. **C-D)** Análisis semejantes comparando salvaje vs *flaA2*-. **E-F)** Análisis comparando *fcpA*- vs *flaA2*-. Para construir los diagramas de Venn se utilizó el módulo estadístico de PatternLab for Proteomics, identificando proteínas detectadas de forma única en cada condición biológica utilizando un valor de corte de significancia estadística <0.05. Se utilizó el módulo TFold para identificar proteínas encontradas en ambas condiciones pero que muestran diferencia estadística significativa (BH q-value <0.05). Cada proteína presente en al menos 4 de las 6 réplicas totales está representada por un punto en los gráficos de volcán. Los puntos azules corresponden a proteínas que satisfacen todos los filtros estadísticos y se consideran diferencialmente expresadas. Las proteínas de interés que se expresan diferencialmente y fueron confirmadas como parte del filamento flagelar se señalan con sus códigos de referencia Uniprot.

Aquellas proteínas cuya abundancia disminuye considerablemente en ausencia de las proteínas de vaina FcpA y FlaA2 fueron evaluadas con el módulo TFold del software PatternLab for Proteomics[162]. Entre las proteínas que se expresan diferencialmente en

salvaje versus las cepas mutantes, evaluamos primero como candidatas aquellas con mayor fold change (Figura 27), hipotetizando que dichos componentes podrían corresponder con mayor probabilidad a proteínas del filamento o interactores de las proteínas FcpA o FlaA2. Como se muestra en la Figura 27, la mayor parte de ellos terminaron efectivamente siendo identificados como nuevos componentes del filamento.

De entre las 8 proteínas presentes en la cepa salvaje y ausentes en *flaA2*$^{-}$ se encuentra FlaA2 (como era de esperar), pero también FlaA1 y con menor señal B0SMT6, un componente del filamento que no interactúa en forma directa con FlaA2. Entre las 24 proteínas presentes exclusivamente en salvaje y no en *fcpA*-, 4 de los 6 primeros hits con mayor diferencia de señal entre ambas muestras terminaron siendo auténticas proteínas flagelares que nunca se habían anotado como tal. Sorprendentemente, FcpA y FcpB no están en esta lista, lo que sugiere que estas proteínas disminuyen en cantidad pero no desaparecen completamente en las preparaciones mutantes, posiblemente debido a contaminación cruzada o degradación parcial.

3.3.3 – Entrecruzamiento químico covalente de proteínas flagelares acoplado a espectrometría de masas (XL-MS)

Esta estrategia fue planificada por su complementariedad con las aproximaciones anteriores: apunta a definir qué componentes están cerca uno de otro y a través de qué regiones específicas. De este modo, permite validar o complementar la información estructural derivada de estudios cristalográficos de proteínas individuales, así como servir de guía para posicionar componentes en mapas de crio-microscopía electrónica. Nuestras primeras pruebas de entrecruzamiento químico fueron llevadas a cabo con el agente comercial BS^2G-d_0, también conocido como bis-sulfosuccinimidilo glutarato, utilizando proteínas recombinantes FcpA y FcpB, pero la eficiencia de entrecruzamiento fue muy baja, por lo que nos inclinamos por el agente de entrecruzamiento NNP9[163,164], desarrollado por el grupo de nuestra colaboradora, la Dra. Julia Chamot-Rooke (Laboratorio de Espectrometría de Masa para Biología/LEMB, Institut Pasteur, Paris). Por su química *click*, consideramos que el NNP9 mejoraría el rendimiento global en la captura y análisis de péptidos entrecruzados.

El compuesto consta de un anillo fenilo con tres brazos espaciadores, dos de los cuales contienen en sus extremos grupos N-succinimidil-carbamato. Estos grupos permiten el entrecruzamiento de aminas primarias, de modo semejante a sus contrapartes NHS-éster, pero son menos propensos a la hidrólisis rápida en agua, aumentando la eficiencia del entrecruzamiento. En ambos casos quedan unidas de forma covalente e irreversible las aminas primarias del extremo N-terminal de las proteínas o de sus cadenas laterales de lisina que estén en contacto próximo en el espacio. El tercer brazo de NNP9 contiene un grupo azido terminal, que permite establecer uniones covalentes con grupos alcalinos unidos por

enlaces fotoclivables a matrices sólidas de agarosa[165]. De este modo, luego de la reacción con NNP9, los péptidos entrecruzados se pueden unir a una matriz sólida que permite purificarlos y enriquecerlos por el protocolo eFASP (*enhanced Filter Aided Sample Preparation*), aumentando significativamente el rendimiento.

El largo del brazo espaciador del NNP9 es de 9 Å, pero se ha demostrado que pueden unir residuos que estén a una distancia de entre 9 y 27 Å[165], ya sea dentro de una misma proteína o en proteínas que interactúan. Esto convierte al entrecruzamiento químico acoplado a espectrometría de masas en una poderosa herramienta para descifrar la estructura de complejos multiproteicos[164,166,167].

Durante una pasantía semestral en el LEMB, utilizamos el compuesto NNP9 para entrecruzar las proteínas de filamentos flagelares purificados a partir de *L. biflexa* y *L. interrogans*. Luego del entrecruzamiento, las proteínas fueron digeridas, los péptidos entrecruzados purificados y analizados por LC-MS/MS.

Los primeros resultados del interactoma de *L. biflexa* salvaje mostraron péptidos entrecruzados principalmente en la superficie del filamento, pero la intensidad de sus señales en espectrometría de masas era relativamente baja. El hecho de que observáramos entrecruzamientos entre FlaB nos permitió verificar que el agente NNP9 es suficientemente pequeño para penetrar la estructura del filamento, aunque lo mismo no ocurría con la tripsina. Dado que los filamentos precipitaban y solo lográbamos su digestión parcial, inicialmente fue difícil establecer las mejores condiciones de digestión y entrecruzamiento.
Para incrementar la señal, se variaron las concentraciones de NNP9 y se probaron diferentes estrategias para optimizar la digestión de los filamentos purificados, incluyendo el agregado de distintas concentraciones de urea y detergentes como Rapigest, SDS y Sarkosyl, e incluso la gelificación en tubo. Realizamos ocho intentos con mejoras limitadas, tanto en la intensidad de los espectros como en el número de péptidos entrecruzados.

Intentamos mejorar la purificación removiendo el ADN con benzonasa para mejorar el pipeteado de las muestras y evitar que el ADN pudiera tapar los filtros durante los pasos de lavado. También notamos que el protocolo inicial de purificación en el paso final involucraba resuspender los flagelos en agua miliQ, lo que quizás dificultaba la detección de péptidos entrecruzados. Con estos cambios incluidos, entrecruzamos 100 µg de purificaciones de filamentos con 150 nmoles de NNP9. Si bien este cociente de NNP9/proteína total es tres veces mayor al estándar normalmente utilizado para XL-MS, razonamos que el agente de entrecruzamiento estaría siendo secuestrado por una gran cantidad de otras proteínas dado que la purificación de filamentos es solo parcial.

Estos ajustes llevaron a obtener un primer juego de datos completo de entrecruzamiento. Aunque la intensidad de los péptidos entrecruzados siguió siendo relativamente baja, el cromatograma se vio poblado con un número mucho mayor de especies altamente cargadas. Detectamos 469 entrecruzamientos en un triplicado, muchos de ellos perteneciendo a las proteínas flagelares esperadas. Enfocándonos en las proteínas flagelares, la mayor parte de los entrecruzamientos fueron encontrados en proteínas del núcleo del filamento (FlaB) y entre las proteínas de vaina FcpA y FcpB.

Se realizaron otros experimentos adicionales, escalando el contenido total de proteínas para incrementar la señal y determinar si estas interacciones se reproducían. A lo largo del proyecto hubo mejoras sustanciales en el análisis de los datos de espectrometría de masas. Inicialmente era necesario buscar los entrecruzamientos en la base de datos y validarlos de forma manual. Con las innovaciones, sobre el final del proceso de modelado los nuevos algoritmos de análisis automático permitieron una validación automática y sin sesgo humano. Asimismo con las sucesivas innovaciones del proyecto, fuimos capaces de buscar proteínas únicas de interés e identificar todas aquellas proteínas con las que interaccionan.

Para evitar sesgos al comparar diferentes juegos de datos analizados de forma diferente, decidimos analizar todos los juegos de datos que dieron señales suficientemente intensas con la nueva versión del software MS Studio, fijando como punto de corte un FDR (*False Discovery Rate*) de 1%. Las purificaciones flagelares contienen al menos 1,327 proteínas distintas (Figura 28), la mayoría de las cuales no son componentes del filamento, confirmando así la justificación de triplicar la cantidad de NNP9 utilizado.

A **Entrecruzamiento químico de flagelos de *L. biflexa***

Labeled type	Trial 1	Trial 2	Trial 3	Trial 4	Combined data
Dead-end	3733	4207	3810	2691	8551
Cross-linked	469	334	531	429	1308
Loop-linked	1987	1404	1921	1562	3923
Total labeled proteins	841	708	823	627	1327

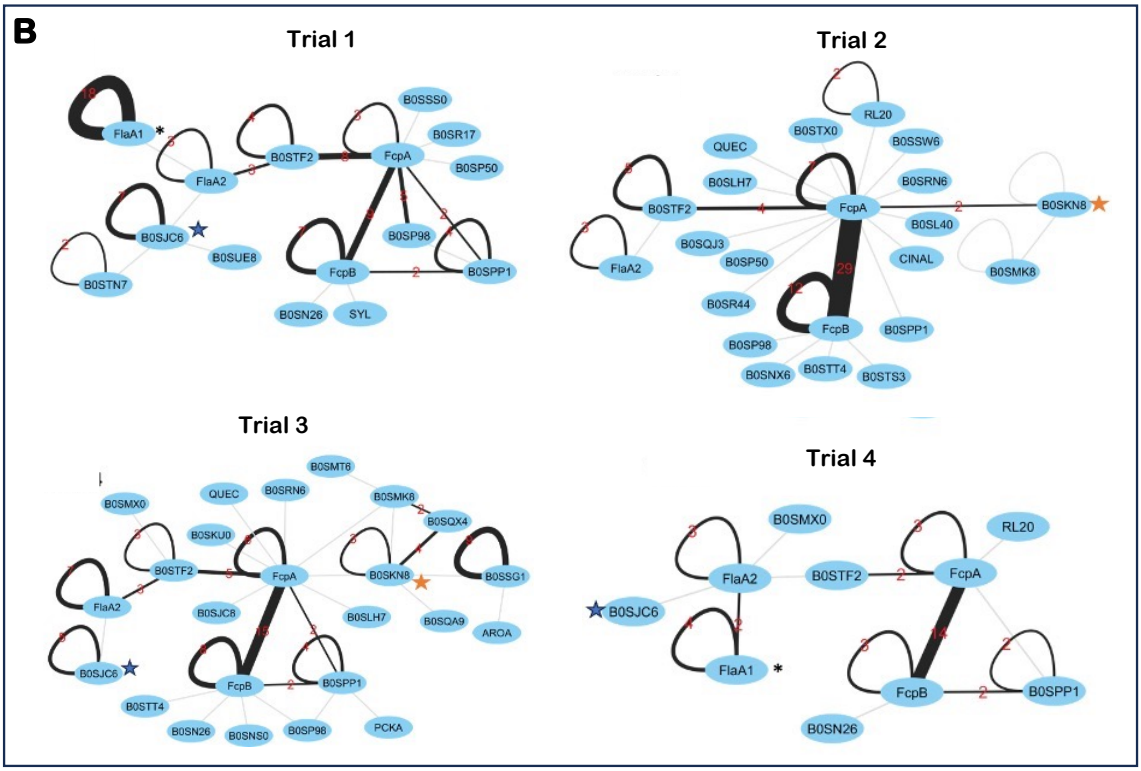
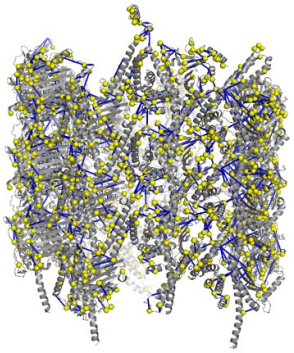


Figura 28: Ensayos de entrecruzamiento de *L. biflexa* salvaje. A) Tabla resumiendo el número de péptidos marcados con el agente NNP9 en cada ensayo. El ensayo/*trial* 1 consiste en un duplicado, mientras que los *trials* 2, 3 y 4 son triplicados, totalizando 11 corridas en los datos unificados. Las columnas muestran los distintos tipos de marcaje con NNP9: "*Dead ends*" refiere a residuos donde se unió el agente NNP9 solo en un extremo sin entrecruzar péptidos; "*Cross links*" refiere a péptidos entrecruzados de proteínas distintas o dentro de un mismo monómero espaciados por 10 o más residuos aminoacídicos; "*Loop-links*" refiere a entrecruzamientos intra-proteína espaciados por menos de 10 aminoácidos. Las proteínas totales marcadas representan la cantidad de proteínas distintas a las que se unió el NNP9, totalizando 1,327 en todas las muestras combinadas. **B)** Redes de interacción proteína-proteína (PPI) de la proteína de vaina FcpA incluyendo todos los ensayos de entrecruzamiento. De las proteínas que componen el filamento, FcpA establece la red más grande de entrecruzamientos, con interacciones tanto inter- como intra-moleculares.

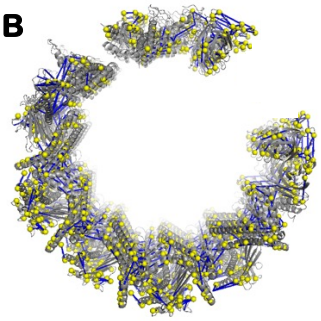
En todos los juegos de datos, cuando FcpA se encuentra en el centro, la red de interacciones es más grande (Figura 28B). Esto se debe en parte a que estas proteínas de vaina se encuentran expuestas al solvente y tienen una gran superficie disponible para que el agente de entrecruzamiento reaccione, y por otra parte, debido a que FcpA es una de las proteínas más abundantes del filamento. En las cuatro redes, la proteína que establece un mayor número de entrecruzamientos con FcpA es FcpB.

Resumen del número de entrecruzamientos y péptidos flagelares marcados por el agente NNP9

A



B



C

	Dead ends	Loop - XLs	Intra - XLs	Inter - XLs
FlaB1	13	0	8	17
FlaB2	9	1	5	10
FlaB3	3	0	1	0
FlaB4	12	6	3	12
FlaA1	19	18	4	10
FlaA2	32	11	8	6
FcpA	40	23	17	74
FcpB	33	23	5	47
B0SJC6	22	18	14	4
B0SMT6	0	0	0	1
B0SKN8	14	15	9	22
B0SMK8	18	5	2	15
B0SPP1	17	3	5	9
B0STF2	25	14	8	10
B0STN7	16	9	1	1

Figura 29: A) Vista de los marcajes químicos mediados por el agente NNP9 en la vaina del filamento flagelar de *L. biflexa* salvaje. Los entrecruzamientos químicos se muestran en azul mientras que los dead-ends se ven en amarillo. **B)** Vista en corte axial de A. **C)** Tabla resumiendo el número de entrecruzamientos y péptidos marcados por agente de entrecruzamiento NNP9 para cada uno de los componentes detectados en filamentos purificados de *L. biflexa* salvaje.

Con respecto a las otras proteínas previamente conocidas, tanto FlaA1 como FlaA2 están presentes en las muestras (Figura 29). Para FlaA1 se detectan 19 *dead-ends*, 22 entrecruzamientos intra-moleculares y 9 *inter-links* consigo misma (incluyendo un homo-entrecruzamiento que se establece entre dos residuos de Lys_{113} de dos monómeros de FlaA1). La única otra proteína con la que entrecruza FlaA1 es FlaA2, entre las que se detecta un único entrecruzamiento. Para FlaA2 se detectan 32 *dead-ends*, 19 entrecruzamientos intra-moleculares y 5 *inter-links*: 1 consigo misma, 1 con FlaA1 y 3 con proteínas no identificadas.
Para los cuatro ensayos, una de estas proteínas previamente desconocidas (Uniprot B0STF2) conecta a FlaA2 con FcpA, por lo que probablemente esta proteína constituya un componente flagelar en sí mismo. El otro caso similar es el de la proteína con código B0SPP1, que entrecruza tanto con FcpA como con FcpB. Además, otras 70 proteínas desconocidas aparecen interactuando con al menos un componente flagelar. Para discriminar cuáles son buenos candidatos a ser verdaderamente proteínas del filamento, fue necesario combinar

estos resultados con los datos comparativos de espectrometría de masas. Integrando estas metodologías con los mapas de crio-microscopía electrónica (ver más abajo), pudimos identificar y modelar 5 de estas proteínas previamente anotadas como hipóteticas, formando parte fehacientemente del filamento flagelar de *L. biflexa* salvaje. Únicamente 5 de los entrecruzamientos finales no son consistentes con el modelo tridimensional final del filamento ensamblado.

3.4 – Enfoque integrativo para construir el modelo 3D del filamento flagelar nativo de *Leptospira*

3.4.1 – El filamento flagelar de *Leptospira* es asimétrico

A comienzos del 2020, nuestro grupo publicó un primer modelo atomístico del filamento flagelar de *Leptospira*[168]. Este trabajo combinó la información estructural de algunos componentes individuales resueltos por cristalografía de rayos-X, ajustados sobre mapas de crio-tomografía electrónica (crio-ET) del filamento flagelar completo de *L. biflexa* salvaje (Figura 30). El mapa de crio-ET fue obtenido a ~12-14 Å de resolución usando la aproximación de promediado de subtomogramas, por el grupo de nuestro colaborador Charles Sindelar (Universidad de Yale, USA).

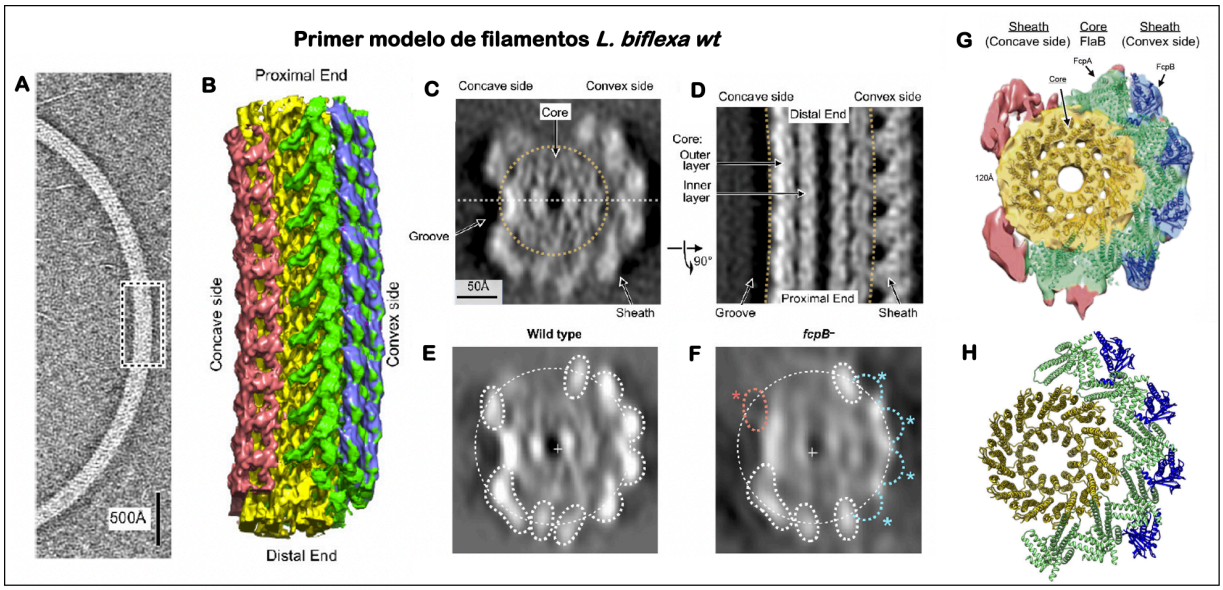


Figura 30: Modelo 3D atomístico del filamento de *L. biflexa* salvaje, figura adaptada a partir del artículo "An asymmetric sheath controls flagellar supercoiling and motility in the leptospira spirochete" Gibson et al.,2020[168] **A.** Imagen representativa de uno de los filamentos purificados vitrificado, visto por crio-microscopía electrónica de transmisión; estas muestras fueron observadas por tomografía en el mismo microscopio, tomando múltiples imágenes de cada filamento mientras se giraba la platina. **B.** Volumen final por crio-tomografía electrónica (crio-ET) del filamento flagelar. La resolución fue mejorada lo máximo posible utilizando la técnica de promediado de subtomogramas, identificando regiones repetitivas de la supraestructura.

Nótese la vaina asimétrica con diferentes componentes identificados en distintos colores (verde, azul y rosado), rodeando al núcleo central (amarillo) helicoidalmente simétrico. **C-D.** Corte transversal y longitudinal respectivamente de los mapas de crio-ET. Las regiones con asimetría se destacan con flechas, incluyendo un surco del lado cóncavo, en la curvatura interna del filamento. La línea punteada en C, es el sitio de corte axial desde donde se observa el panel D. **E.** Corte transversal del mapa de crio-ET de la cepa salvaje, filtrado a 18 Å de resolución (para mejorar la comparación con el panel F). **F.** Corte transversal del mapa de crio-ET de la cepa *fcpB*- (que fue obtenido a 18 Å de resolución); con círculos punteados en azul se subrayan las diferencias con el mapa del panel E, correspondientes a cuatro densidades del lado convexo que coinciden con la ubicación ajustada de los protómeros de FcpB en la cepa salvaje. Una porción de densidad adicional del lado cóncavo (* marcado en rojo) no pudo ser asignada en el modelo, pero podría corresponder a FlaA1 y/o FlaA2. **G.** Corte transversal como en el panel B, aquí con el modelo atómico del núcleo y vaina del filamento ajustados dentro del mapa crio-ET. El núcleo se muestra en amarillo y contiene 11 protofilamentos de FlaB1, mientras que en la vaina, del lado convexo, encontramos seis protofilamentos de FcpA (en azul) y, sobre ellos en una capa más externa, cuatro protofilamentos de FcpB (en verde). Nótese densidad en el lado cóncavo del apéndice (coloreada en rojo), que debido a la limitada resolución del mapa crio-ET, no se podía aún interpretar en términos moleculares. **H.** Mismo modelo que en el panel G, pero aquí ilustrado sin el mapa superpuesto, para mayor claridad.

La construcción del modelo inicial se realizó bajo la hipótesis de la interacción FcpA-FcpB, modelando heterodímeros de estas proteínas utilizando: i) el modelo por homología de FcpB de *L. biflexa* utilizando como molde la estructura experimental de *L. interrogans* (introducida en la sección 3.2.2); y ii) la estructura cristalográfica de FcpA de *L. biflexa* (sección 3.2.1). Los heterodímeros candidato con menor energía fueron seleccionados para realizar una búsqueda exhaustiva de ajuste sobre el mapa de crio-ET del filamento. La correcta ubicación de FcpB fue además confirmada por comparación con crio-tomogramas obtenidos con la cepa *fcpB*- de *L. biflexa* (E y F).

En la construcción del núcleo interno de flagelinas (FlaB) se asumió que la organización helicoidal propia de los flagelos bacterianos, con 11 protofilamentos de flagelina se mantenía en *Leptospira*. Un modelo por homología de FlaB1 (Uniprot B0SSZ5, ortóloga a la isoforma más abundante de FlaB en *L. interrogans*), calculado usando como templado a la flagelina FliC de *Salmonella enterica* (modelos descritos en la sección 3.1.2) fue posicionado de forma manual sobre el mapa de crio-ET. Utilizando luego el suite de programas Rosetta, la posición de cada protómero de FlaB1 fue refinada por un proceso iterativo: ciclos de ajuste en cuerpo rígido, seguidos de minimización de energía de la cadena principal y cadenas laterales.

Como puede verse en la Figura 30, el modelo atómico interpretado revela una estructura organizada en dos capas. La capa interna consiste en el núcleo de proteínas FlaB donde, como se esperaba, ajustan bien 11 protofilamentos de FlaB empaquetados lado a lado a lo largo del eje largo de la hélice. Este núcleo central es muy similar al observado para otros filamentos flagelares, que en la mayoría de las bacterias exoflageladas conforman el filamento en su totalidad. En *Leptospira* sin embargo, el polímero esta recubierto por una vaina altamente asimétrica (con distintas densidades proteicas en distintos lados del apéndice). La superhélice del filamento muestra una curvatura pronunciada con un lado cóncavo y otro convexo. En el lado convexo de la vaina se ubican las dos proteínas Fcp (FcpA y FcpB), también dispuestas en

hileras longitudinales de protofilamentos a lo largo del eje central. Las FcpA se ubican en una capa más interna de cinco protofilamentos que interactúan en forma directa con el núcleo central de FlaBs. Los protofilamentos de FcpB por su parte, se disponen en el surco externo entre dos de FcpA contiguos, conformando una segunda capa más externa alejada del núcleo central. Del lado cóncavo del filamento, en el lado opuesto de la vaina, se observa un surco lateral sin densidad electrónica, y hacia cada uno de sus flancos, dos grandes porciones de densidad (en rojo) que no podíamos asignar aún a especies moleculares conocidas, pero que muestran ser divergentes en estructura con respecto a las Fcp. Estas densidades podrían corresponder a las isoformas de FlaA1 y FlaA2 (para las que, a esta altura de avance del proyecto, no disponíamos de estructura 3D conocida). Dado el volumen de la densidad podían de hecho corresponder a dímeros de FlaA, y/o proteínas adicionales aún no anotadas.

Este primer modelo atómico del filamento flagelar de *Leptospira* representó un importante hito, descubriendo las posiciones de componentes flagelares previamente conocidos dentro del mapa reconstruido por crio-tomografía electrónica. Una vez realizados los análisis de proteómica comparativa y *cross-linking* de filamentos purificados, pudimos interrogar este modelo inicial de baja resolución verificando su compatibilidad con la red de interacciones proteína-proteína. Los 29 entrecruzamientos entre FcpA y FcpB (Figura 29) coinciden y confirman el modelo propuesto. Lo mismo sucede con 6 de los 7 inter-entrecruzamientos entre monómeros distintos de FcpA y los 6 entre FcpBs dentro de un mismo protofilamento. FcpA establece 40 entrecruzamientos intra-moleculares, 23 de los cuales son loop-links que se establecen entre residuos a una distancia menor a 10 aminoácidos. De los 17 intra-links restantes hay 3 que no se encuentran modelados en la estructura final propuesta. Por su parte FcpB establece 28 intra-links, 23 de los cuales son loop-links y todos son coherentes con la estructura tridimensional de la proteína. Este número de interacciones y el hecho de que prácticamente todas coincidan con las estructuras cristalográficas y el modelo final propuesto del ensamblaje valida además la metodología XL-MS, a pesar de la baja intensidad de sus espectros. La espectrometría de masas también logra distinguir entre las cuatro isoformas homólogas de FlaB, y las 70 marcaciones diferentes de NNP9 que involucran interacciones entre FlaB coinciden con nuestro modelo tridimensional actual del filamento.

3.4.2 – En el filamento flagelar de *Leptospira* se identifican nuevos componentes

El siguiente paso adelante en la interpretación de la arquitectura molecular del filamento de *Leptospira* fue publicado en 2022 en bioRxiv[169]. En él se presentan las estructuras de los filamentos salvajes de *L. biflexa* (aún a una resolución relativamente limitada de ~6.5 Å), y de los mutantes *fcpA*- de *L. biflexa* (3.8 Å) y *flaA2*- de *L. interrogans* (2.9 Å), significando estas últimas dos estructuras de filamentos mutantes un avance enorme en términos de la resolución a la que los mapas podían resolverse.

El adelanto mayor en este sentido fue conseguido en el Laboratorio del Dr C. Sindelar, donde por vez primera se pudo aplicar análisis de partícula única para reconstruir los volúmenes por crio-microscopía electrónica (crio-EM). Debido a que los filamentos tienen una geometría superenrollada y adoptan una orientación preferencial, alcanzar estas resoluciones por crio-EM no fue una tarea trivial para nuestros colaboradores; requirió la aplicación de métodos innovadores de procesamiento de imágenes específicos para estructuras con curvatura como Emclarity[170].

La tarea fue aún más desafiante considerando la heterogeneidad en la composición de los filamentos. De hecho, observando preparaciones de filamentos de la cepa mutante *fcpA*-, se terminaron resolviendo dos tipos de poblaciones estructurales netamente diferentes que estaban mezcladas en las muestras: una con filamentos más gruesos que la otra. La población de filamentos finos, que se observa en la Figura 31, tiene aproximadamente 120 Å de diámetro, y consiste únicamente del núcleo interno de FlaB. La resolución local de 3.4 Å que alcanzó la reconstrucción 3D de este mapa nos permitió trazar las cadenas principales de las proteínas FlaB, muchas cadenas laterales, y explicar la mayor parte del mapa.

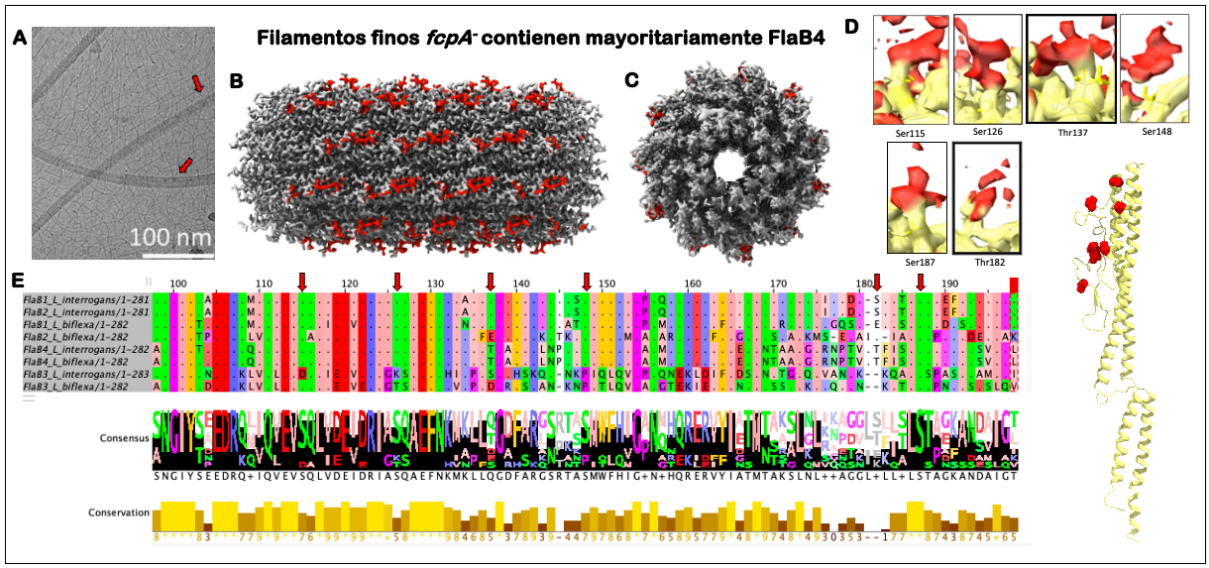


Figura 31: Los filamentos finos de *Leptospira biflexa fcpA*- están compuestos principalmente de FlaB4. A. Micrografía mostrando la heterogeneidad en el grosor y la curvatura de los filamentos flagelares de *fcpA*-. Con flechas rojas se marcan los filamentos gruesos y finos. **B.** Volumen final promediado de los filamentos flagelares finos de *fcpA*- en vista lateral. Las densidades correspondientes a los carbohidratos se muestran en rojo. **C.** Corte axial de B. **D.** Vista aumentada de los residuos glicosilados de FlaB. Los marcados con recuadro más grueso, treoninas 137 y 182, son exclusivos de la isoforma FlaB4. La posición de estos residuos en FlaB se muestra abajo. **E.** Alineamiento de las cuatro isoformas de FlaB de *L. interrogans* y *L. biflexa* en la región donde se identifican los carbohidratos. Se muestran con flechas rojas los residuos que se observan en D. Figura adaptada de Brady et al.2022[169].

En la superficie del filamento encontramos pequeñas densidades (en rojo en las Figuras 31B, C y D) no explicadas por ninguna de las cuatro isoformas de FlaB. Dado que los residuos de FlaB que se ubican en estos sitios corresponden a serinas o treoninas, concluimos que se trata de *O*-glicosilaciones de superficie. Si bien la identidad de estos carbohidratos permanece hasta el momento desconocida, los sitios de glicosilación nos permiten discriminar entre isoformas de FlaB. En particular, de las siete *O*-glicosilaciones que se muestran en la Figura 31D, cinco están modificando residuos conservados en todas las isoformas de FlaB; mientras que las últimas dos, correspondientes a las Thr137 y Thr182, nos permiten determinar que se trata de FlaB4, la única isoforma que tiene treoninas en dichas posiciones (ver alineamiento en Figura 31E).

Por otra parte, de los filamentos gruesos, la clase mayoritaria (~53% de los segmentos considerados) pudo ser resuelta a 3.8 Å (con resoluciones locales en el rango de 3.4 - 4.4 Å): se trata de filamentos curvos, pero con menor curvatura que la observada para la cepa salvaje (Figura 32A). Su estructura contiene, además del núcleo interno de FlaB, una densidad ubicada en la vaina, de tamaño consistente con la presencia de dos proteínas que se asocian entre sí y al núcleo de FlaB en el lado cóncavo del filamento (Figuras 32B y C).

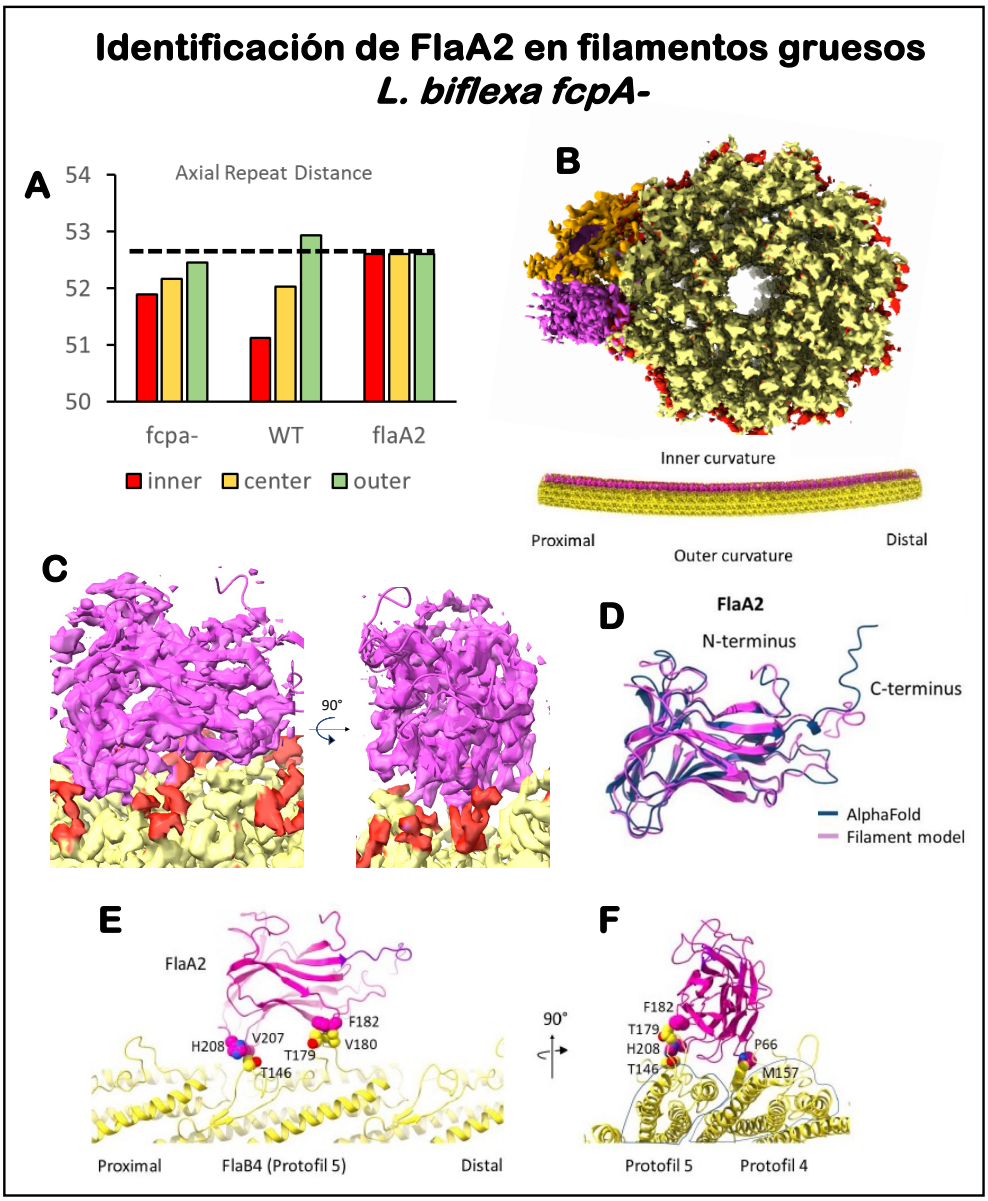


Figura 32: Identificación de FlaA2 en los filamentos gruesos de *L. biflexa fcpA*-. A. Distancias de repetición axial de protómeros consecutivos de FlaB, comparando filamentos salvajes vs mutantes *fcpA*- y *flaA2*-, medidas por separado sobre la cara interna (en rojo), externa (verde) de las superhélices. Su promedio (en amarillo) se considera la distancia al centro de la hélice, sobre la cual se puede medir el paso. A mayor diferencia entre los valores de estas tres posiciones, mayor es la curvatura del filamento. El filamento salvaje es por tanto el que presenta mayor curvatura, le sigue *fcpA*-, mientras que los filamentos de *flaA2*- son completamente derechos. **B.** Reconstrucción 3D final en corte transversal (arriba) y longitudinal (abajo) de los filamentos gruesos de la cepa *fcpA*-. La estructura contiene dos proteínas de vaina que interactúan entre sí (se muestran en fucsia y naranja). **C.** Vista aumentada de la proteína que se muestra en fucsia en el panel B, con el modelo AlphaFold2 de FlaA2 de *L. biflexa* superpuesto en el mismo color. Se muestra la interacción con FlaB (en amarillo) mediada por glicosilaciones (en rojo). **D.** Modelo AlphaFold (en azul) superpuesto a la estructura final refinada de FlaA2 (en fucsia). **E.** Interacciones proteína-carbohidrato entre FlaA2 y FlaB: El presunto glicano en Ser148 (FlaB4) interactúa con Arg145 y His208 (FlaA2), y el glicano en Thr137 (FlaB4) interactúa con Arg163 y Arg166 (FlaA2). **F.** Interacciones proteínas FlaA2-FlaB: His208 y Val207 (FlaA2) interactúan con Ser146 (FlaB4), y Phe182 (FlaA2) interactúa con Thr179 y Val180 (FlaB4). Figura adaptada de Brady et al.,2022[169].

Considerando que en el modelo previo obtenido por crio-ET con filamentos de la cepa salvaje, sobre esta cara interna cóncava del filamento se ubicaban densidades que postulábamos podrían corresponder a las isoformas de FlaA, y considerando además que los filamentos *fcpA*$^-$ carecen de las otras dos proteínas de vaina mayoritarias conocidas, FcpA y FcpB, intentamos ajustar la estructura cristalográfica de FlaA1 y el modelo AlphaFold2 de FlaA2 en esta densidad de vaina de la cepa *fcpA*$^-$. Resultó claro que una de las porciones de dicha densidad correspondía con FlaA2 de *L. biflexa*, lo que permitió refinar el modelo atómico de FlaA2 *in situ* sobre este mapa de crio-EM.

En la Figura 32D se comparan los modelos inicial y final de FlaA2; la predicción es muy buena, aunque se observan diferencias en los bucles de FlaA2 y principalmente en el extremo C-terminal, donde FlaA2 forma un "tentáculo" que se proyecta hacia afuera de la estructura y contacta a la subunidad FlaA2 adyacente en el protofilamento permitiendo la polimerización. Por otra parte, como se observa en las Figuras 32E y F, la interacción de FlaA2 con el núcleo central de FlaBs, está mediada por dos sitios de contacto proteína-proteína y por dos interacciones que FlaA2 establece con los residuos glicosilados de FlaB: la Ser148 conservada en todas las isoformas de FlaB, y la Thr137, que es exclusiva de FlaB4. Dichas interacciones mediadas por azúcares se localizan en FlaA2 precisamente en una de las regiones de unión a carbohidratos que habíamos pronosticado para FlaA1 al analizar su estructura cristalográfica (ver sección 3.2.3).

La segunda porción de densidad crio-EM sin identificar en los filamentos gruesos de *fcpA*$^-$, no coincidía con ninguna proteína conocida, y la resolución local de esa zona no era lo suficientemente alta (4.4 Å) como para trazar una cadena principal directamente en base a la densidad crio-EM. Es precisamente en este contexto que nuestra aproximación integrativa, y más específicamente los resultados descritos en las secciones 3.3.2 y 3.3.3 cobran especial relevancia. La proteína no caracterizada más abundante en los análisis proteómicos de filamentos de la cepa *fcpA*-, así como la cuarta más abundante en los provenientes de la cepa salvaje, corresponde a una proteína con código Uniprot B0SJC6.

A partir de la secuencia de B0SJC6, su estructura 3D fue predicha con AlphaFold2 –con alta fiabilidad en la mayor parte de su extensión. Su tamaño 3D y topología de estructuras secundarias –compuesta únicamente por α-hélices– coincidieron perfectamente con la densidad crio-EM. En la Figura 33B se ilustra un primer posicionamiento aproximado de esta nueva proteína sobre la densidad misteriosa, revelando que el ajuste es bastante bueno. Reanalizando nuestros datos obtenidos por XL-MS, B0SJC6 arroja efectivamente un entrecruzamiento entre las cadenas laterales de los residuos K282 de B0SJC6 y K129 de FlaA2 (Figura 33D), que aunque tenía un *score* relativamente bajo, se repite en forma consistente en cinco muestras diferentes de *L. biflexa*.

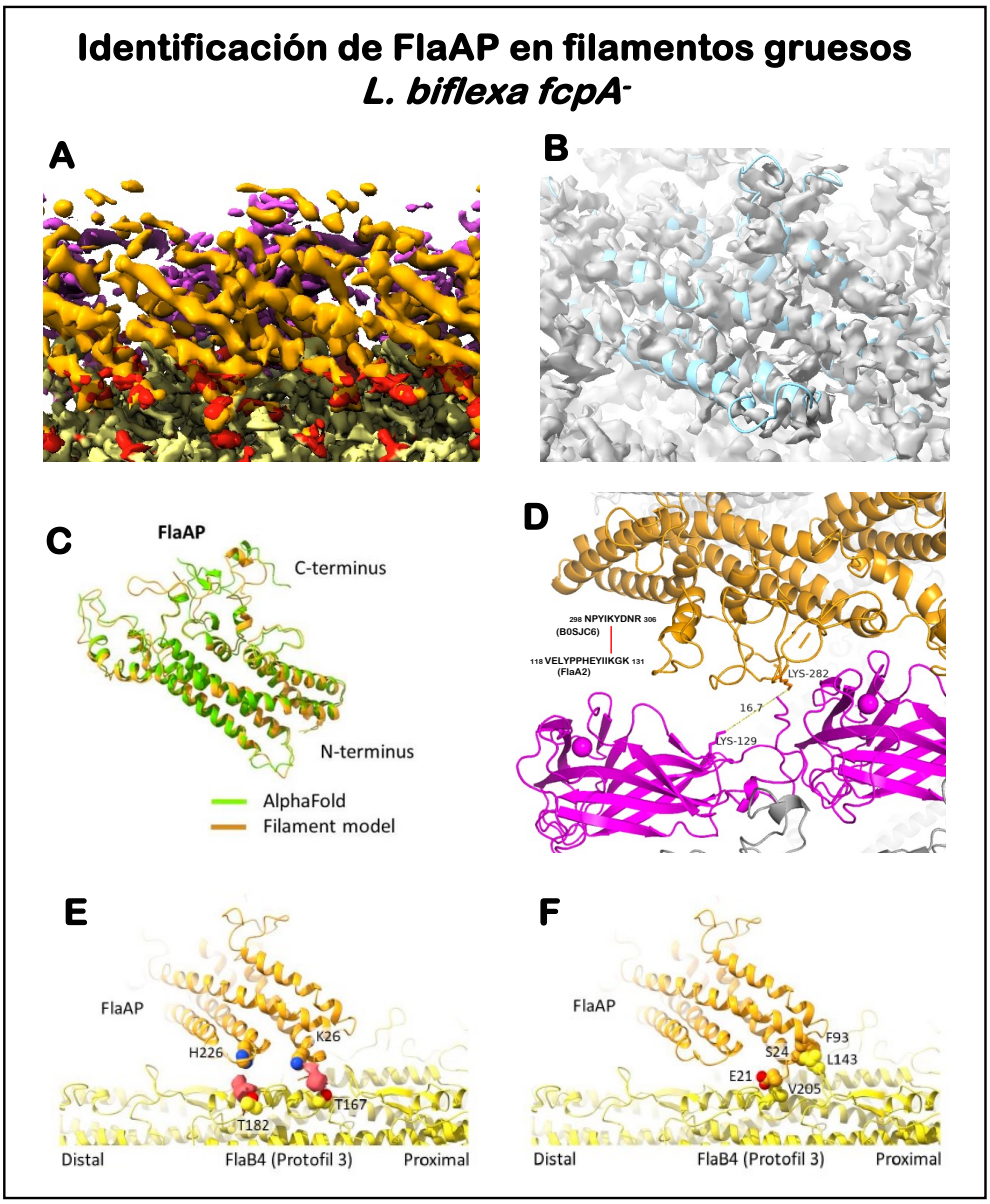


Figura 33: Identificación de FlaAP en los filamentos gruesos de *L. biflexa fcpA*-. A. Vista aumentada de la Figura 32B, enfocando la densidad crio-EM no identificada, que revela una estructura secundaria compuesta principalmente por α-hélices. **B.** Superposición inicial del modelo predicho por AlphaFold2 de la proteína B0SJC6, ajustada manualmente dentro de la densidad crio-EM no identificada. **C.** La estructura final refinada de FlaAP (marrón) tiene muy buena coincidencia con el modelo predicho con AlphaFold2 inicial (en verde). **D.** Posición del entrecruzamiento detectado por XL-MS entre el residuo K282 de FlaAP y K129 de FlaA2 (a la izquierda se muestran los péptidos entrecruzados por NNP9). **E.** Las interacciones proteína-carbohidrato que asocian FlaAP con FlaB se establecen entre la His226 de FlaAP y el presunto glicano unido a la Thr182 de FlaB4, y entre la Lys126 de FlaAP y el glicano unido a la Thr167 de FlaB4. **F.** Interacciones proteína-proteína entre FlaAP y FlaB4: la Ala172 (FlaB4) interacciona con el Glu21 (FlaAP), y la Leu143 (FlaB4) interacciona con las Phe93 y Ser24 (FlaAP). Figura adaptada de Brady et al.,2022[169].

En suma, pudimos identificar a B0SJC6 como un nuevo componente del filamento flagelar, que es característico y resulta ser exclusivo de *Leptospira*. Se lo nombra inicialmente FlaAP (*FlaA-associated protein*) por su interacción directa con FlaA2, y al igual que FlaA2, FlaAP

interactúa también directamente con el núcleo interno de FlaB mediante dos contactos proteína-proteína y dos interacciones proteína-carbohidrato (Figuras 33E y F).

En los filamentos gruesos de la cepa *fcpA*-, la unión conjunta de FlaAP y FlaA2 al núcleo interno del filamento crea una estructura superhelicoidal, en la que se observa curvatura. En contraste, la cepa mutante KO de *flaA2*- de *L. interrogans*, que es también viable aunque no móvil, tiene filamentos completamente derechos y simétricos (Figura 34A). La reconstrucción de estos filamentos reveló una estructura sorprendente: el núcleo de flagelinas FlaB se encuentra completamente rodeado por una vaina simétrica de FcpA+FcpB (Figura 34A). Al igual que ocurre en el lado convexo de la estructura *wt*, la vaina consta de una primer capa de proteínas FcpA rodeada por otra capa externa de FcpBs, pero en este caso la estructura cubre la totalidad de la superficie del filamento de un modo radialmente simétrico.

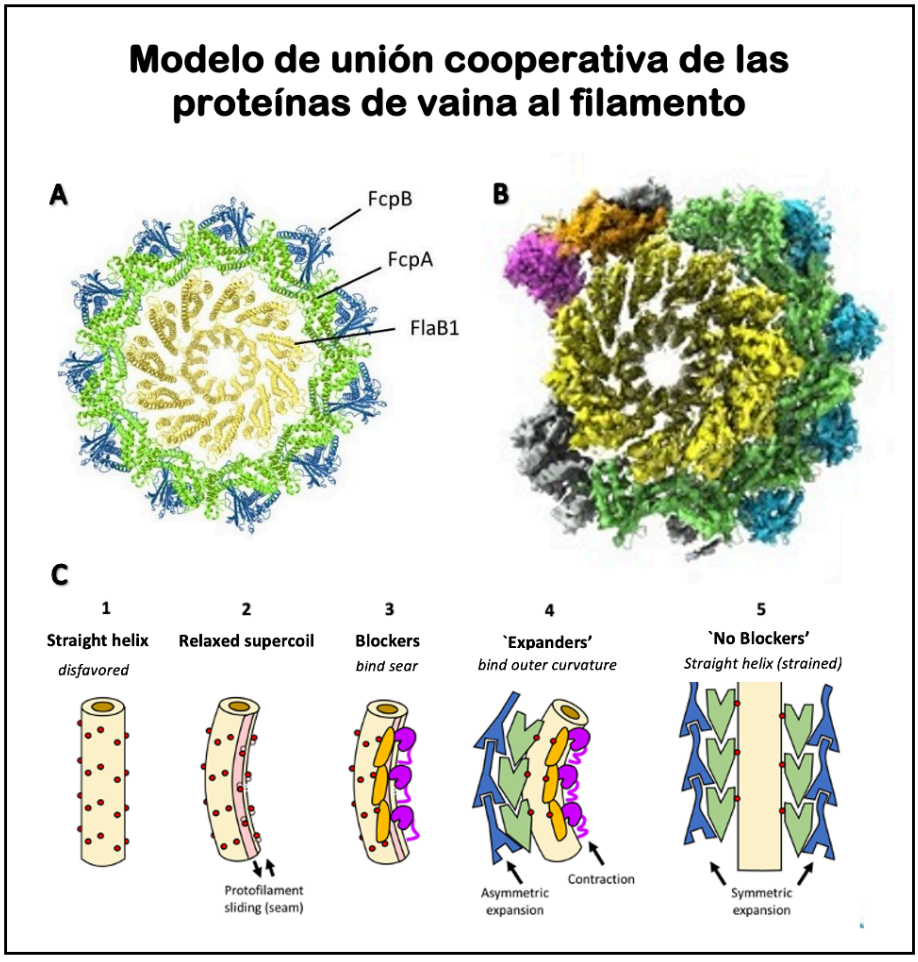


Figura 34: Modelo de unión cooperativa de las proteínas de vaina en el filamento de *L. biflexa* salvaje. A. Vista axial del modelo del filamento *flaA2*- de *L. interrogans*. El trazado de los mapas de crio-EM muestra una estructura simétrica con el núcleo central de FlaBs (en amarillo) rodeado por una vaina de proteínas FcpA (en verde) y FcpB (en azul). **B.** Modelo esquemático de unión de las proteínas de vaina. (1) El homopolímero de FlaB es una hélice simétrica derecha, aunque esta conformación está desfavorecida y se relaja mediante el "sliding" de protofilamentos (2). Esto genera una costura sobre la cual se unen las proteínas "bloqueadoras" FlaA2 y FlaAP (3) creando una estructura asimétrica sobre la cual pueden unirse primero FcpA y luego FcpB (4). Estas proteínas

funcionan como expansoras estirando la cara convexa del filamento al unirse y aumentando la curvatura del filamento. (5) En ausencia de las proteínas bloqueadoras, las Fcps son libres de unirse en todas las superficies del núcleo de FlaBs, lo que explica la estructura simétrica del mutante *flaA2*-. Figuras tomadas de "*Structural basis of flagellar filament asymmetry and supercoil templating by Leptospira spirochete sheath proteins" Brady et al.,2022[169].*

3.4.3 – Modelo final completo del filamento flagelar nativo de *L. biflexa*

Pudiendo ahora integrar la estructura experimental de FlaA1 de *L. borgpetersenii*, y develando la importancia que juega el calcio para su correcto plegamiento (sección 3.2.3), optimizamos variantes del protocolo de purificación de filamentos de *L. biflexa*, para purificarlos con y sin calcio (utilizando el quelante *o*-fenantrolina para secuestrar el metal divalente). Distintas preparaciones de estos filamentos fueron analizadas en simultáneo por espectrometría de masas y por crio-EM. Si bien no se observaron diferencias apreciables en la composición o estructura de los filamentos purificados con y sin Ca^{2+}, la resolución de los mapas de crio-EM por análisis de partícula única de filamentos nativos de *L. biflexa* salvaje mejoró notablemente.

Ya habiamos observado que hay diferencias entre las purificaciones realizadas en los distintos laboratorios, pero aquí además descubrimos de forma accidental que cuando los filamentos purificados pasan un tiempo a -80°C y luego se descongelan, tienden a quebrarse de forma axial generando fragmentos más cortos. Dichos fragmentos presentan una curvatura de superenrollamiento más homogénea en comparación con filamentos largos, lo que redundó en mejoras sustanciales durante el promediado de imágenes de partículas y por ende en la resolución final obtenida en las reconstrucciones 3D. Así se logró obtener mapas finales crio-EM de *L. biflexa* salvaje (en presencia de calcio) a 3.6 Å de resolución.

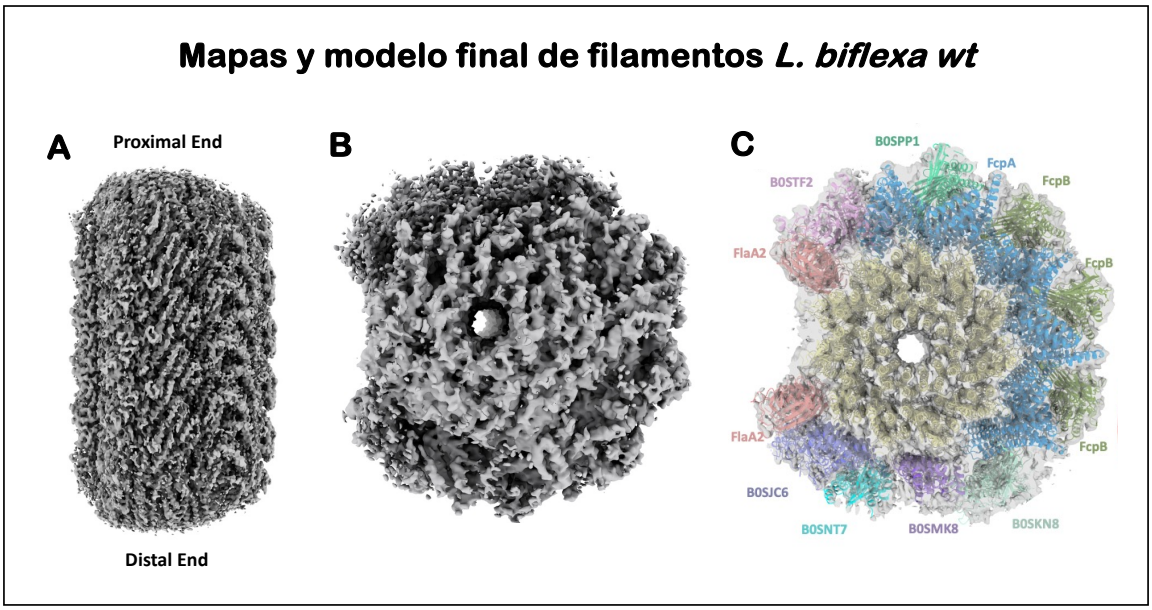


Figura 35: Mapas y modelo final del filamento salvaje de *L. biflexa*. A. Vista longitudinal del mapa de crio-EM del filamento salvaje de *L. biflexa* obtenido a 3.6 Å de resolución **B.** Corte axial de A. **C.** Modelo final del filamento de *L. biflexa* con todos los componentes fiteados sobre el mapa en B.

El posicionamiento y ajuste de proteínas de vaina previamente identificadas (FcpA, FcpB, FlaA2 y FlaAP) dentro de sus respectivas densidades fue mucho más simple y evidente, por la mejor resolución del mapa. Comenzamos modelando las porciones de densidad faltantes con FlaA1, la última proteína flagelar previamente conocida. Para ello, a partir de la estructura cristalográfica de FlaA1 de *L. borgpetersenii*, modelamos por homología FlaA1 de *L. biflexa* y la fiteamos en la única porción de densidad con forma similar; del otro lado del surco donde se encuentra FlaA2.

Para identificar los componentes del filamento restantes, integramos los resultados de los estudios XL-MS (ver sección 3.3.3), así como también los análisis de proteómica comparativa *bottom up* (ver sección 3.3.2), que revelaban la presencia y la abundancia relativa de ciertas proteínas clave. La Figura 36A muestra la red final de interacciones entre las proteínas de vaina halladas por XL-MS, y la Tabla 2 resume los datos MS para cada uno de los componentes de vaina, que se describirán a continuación en esta sección.

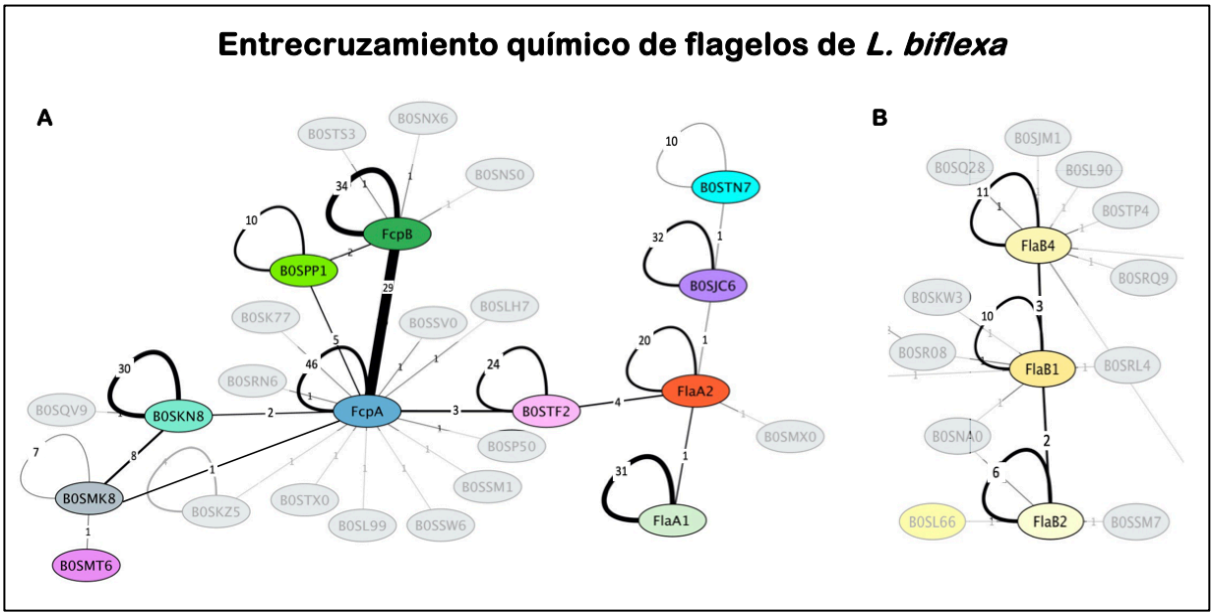


Figura 36: Redes finales de interacciones proteína-proteína del filamento de *L. biflexa* salvaje obtenidas por XL-MS. A. Red final de interacciones proteína-proteína (PPI) de las proteínas de vaina del filamento de *L. biflexa* salvaje, obtenidas a partir de los ensayos de entrecruzamiento con el agente NNP9 (sección 3.3.3). En ella se detectan todas las proteínas que conforman la vaina de nuestro modelo final y al menos un inter-entrecruzamiento valida cada interacción proteína-proteína, a excepción de las interacciones entre B0STN7-B0SMT6 y B0STN7-B0SMK8. **B.** Red final de interacciones proteína-proteína (PPI) en el núcleo del filamento flagelar de *L. biflexa* salvaje, que no tiene nodo en común o interacción alguna detectada con la vaina del filamento. De las cuatro isoformas de FlaB se detectan tres, con FlaB3 ausente en la red. El nodo en común entre las tres isoformas restantes es FlaB1 para la cual se detectan interacciones con FlaB4 y FlaB2.

Resumen de resultados de Espectrometría de masas		
Proteína	Cepas fcpA-	Cepas flaA2-
FcpA	↓ 52 veces	↓ 2 veces
FcpB	↓ 44 veces	↓ 2 veces
FlaA1	↓ 1.9 veces	Ausente
FlaA2	↓ 1.8 veces	Ausente
B0SJC6	↓ 2.3 veces	↓ 18 veces
B0SPP1	↓ 40 veces	↓ 9 veces
B0STF2	↓ 36 veces	↓ 14 veces
B0SMT6	Ausente	Ausente
B0SMK8	Ausente	↓ 14 veces
B0SKN8	Ausente	↓ 17 veces
B0STN7	Ausente	-

Tabla 2: Resumen de resultados de espectrometría de masas comparativa entre cepas *L. biflexa* salvaje y los mutantes *knockout fcpA-* y *flaA2-* según resultados obtenidos en la sección 3.3.2. Para cada uno de los componentes de vaina se observa la presencia y abundancia relativa en las cepas mutantes con respecto a la cepa salvaje.

La proteína con código Uniprot B0SPP1 fue el primer nuevo componente que logramos identificar y ajustar en mapa. B0SPP1 es la segunda proteína en cuanto al número de entrecruzamientos que establece con FcpA, siguiendo al interactor previamente conocido y presente en varias copias que es FcpB. B0SPP1 es además una proteína cuyos niveles disminuyen unas 40 veces cuando se comparan cepas *fcpA- vs* salvaje por espectrometría de masas (Figura 36 y Tabla 2). AlphaFold2 permitió predecir la estructura 3D de B0SPP1 con alta fiabilidad, lo que nos sorprendió pues resultó ser claramente homóloga a FcpB: aun cuando comparten tan solo un ~7% de identidad de secuencia, el rmsd entre ambas estructuras es de 2.3 Å.

Para discernir si (y dónde) B0SPP1 forma parte del filamento en el mapa crio-EM, y dada su semejanza con FcpB, la posicionamos sistemáticamente sobre todas las ubicaciones de FcpB y densidades adicionales, constatando que el único caso en que B0SPP1 ajusta mejor que FcpB al mapa, es donde terminamos posicionándola en el modelo final (Figura 37). A su vez, cinco inter-entrecruzamientos entre FcpA y B0SPP1 validan la presencia y posición de B0SPP1 en el filamento (Figuras 36 y 37A). Entre las diferencias estructurales con FcpB, que permiten diferenciarlas inequívocamente, B0SPP1 tiene una hebra menos en su hoja β y un extremo N-terminal sustancialmente más largo. Dicha extensión se posiciona entre dos monómeros de FcpA, y efectivamente genera un entrecruzamiento entre la Lys130 de FcpB y la Lys23 de B0SPP1 que detectamos (en la Figura 37B se dibuja desde la Ser26, que es el primer residuo de B0SPP1 que se observa nítidamente en el mapa).

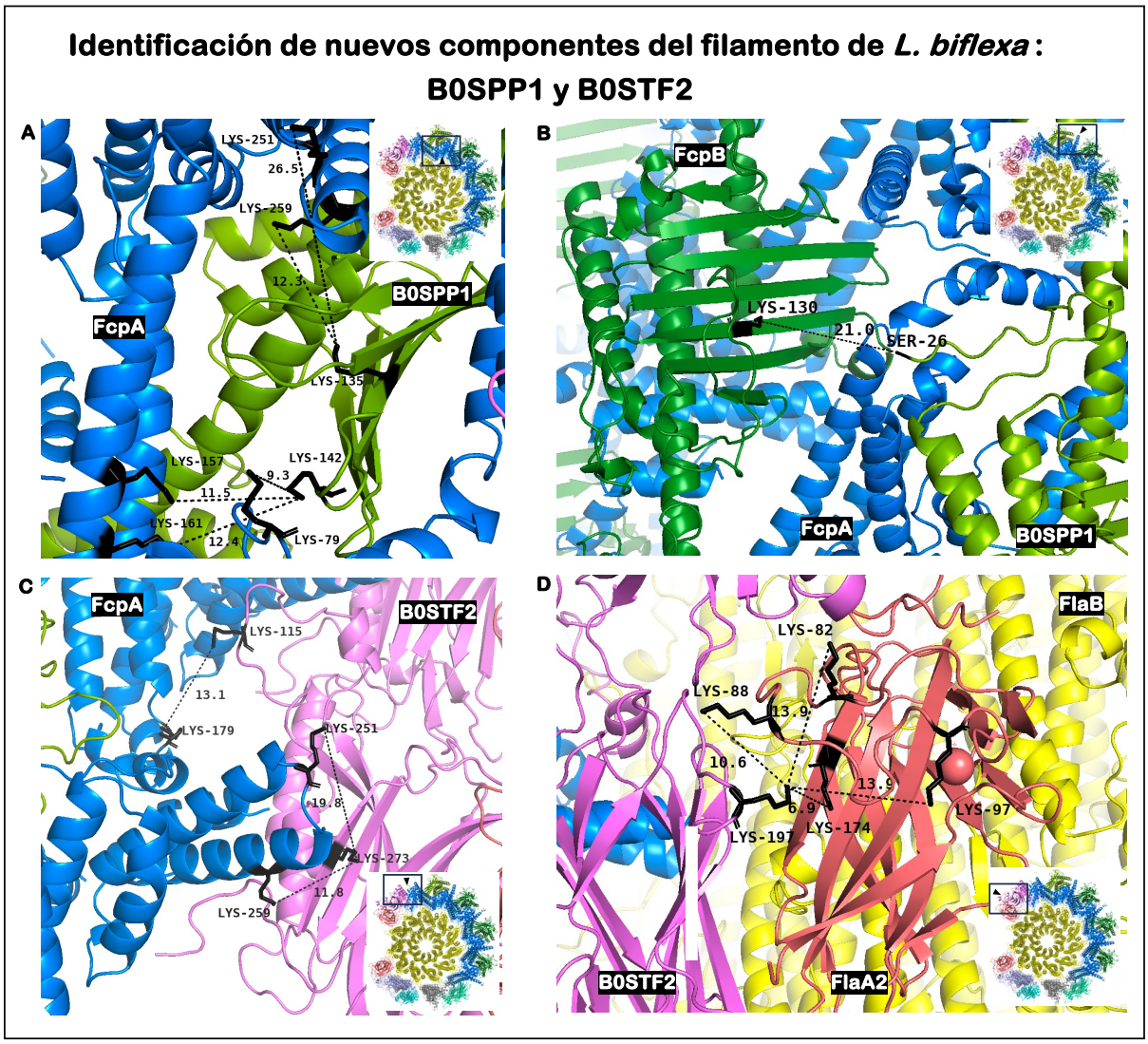


Figura 37: Identificación de B0SPP1 y B0STF2 en el filamento flagelar de *L. biflexa* salvaje. A. Se observan cinco inter-entrecruzamientos entre B0SPP1 y FcpA. El residuo K135 de B0SPP1 entrecruza con los residuos K251 y K259 de FcpA (arriba), mientras que la K142 de B0SPP1 interacciona covalentemente con los residuos K157, K161 y K79 de FcpA (abajo). **B.** El residuo K130 de FcpB entrecruza con K23 en el extremo N-terminal de B0SPP1. Nótese que esta interacción se anota aquí sobre la serina S26, razón es que éste es el primer residuo visible en el mapa crio-EM de manera fiable; sin embargo, por la topología del modelo refinado, la distancia del entrecruzamiento con la K23 será menor a 21 Å, lo cual es factible. **C.** Entrecruzamientos químicos que se establecen entre FcpA y B0STF2. Las lisinas 179, 251 y 259 en FcpA se entrecruzan respectivamente con las 115, 273 y 273 en B0STF2. **D.** La K197 de FlaA2 establece entrecruzamientos con los residuos K82, K88, K97 y K174 de B0STF2, todos a una distancia concordante con el rango del agente de entrecruzamiento NNP9. El código de colores de las proteínas sigue el del modelo: en amarillo, FlaB; en azul, FcpA; en verde claro, B0SPP1; y en verde inglés, FcpB. En rosado se identifica a la proteína B0STF2, y en rojo a FlaA2. En cada panel se ilustra abajo a la derecha una vista general del corte transversal del filamento completo, encuadrando la región magnificada del panel, y con una flecha, marcando la vista desde donde se observa.

El siguiente candidato que analizamos en detalle fue B0STF2. No sólo se entrecruza con FcpA (mediante tres *inter-XL* que se observan en la Figura 37C), sino que también establece cuatro inter-entrecruzamientos con FlaA2, conectando a ambas proteínas en la red (Figura 37D). Por

MS, los niveles de B0STF2 se ven disminuidos ~36 veces en ausencia de FcpA, y ~14 veces en ausencia de FlaA2. Hay una región en la densidad crio-EM con muy buena resolución local, bordeando directamente a FcpA y FlaA2, que permitió un ajuste fiable de B0STF2 en el modelo.

Del lado diametralmente opuesto a B0STF2, la vaina muestra una gran porción de densidad que no podíamos interpretar y cuya resolución local es menor que en el resto de la vaina, por lo que su modelado resultó más desafiante. A modo anecdótico, el primer componente de esta densidad multiproteica, B0SMT6, fue correctamente posicionado sobre el mapa de forma manual durante una reunión por zoom con mi tutor Alejandro Buschiazzo. La fortuna estuvo de nuestro lado si consideramos que B0SMT6 era uno de múltiples candidatos a proteínas de vaina del filamento, y que en simultáneo ninguna de las proteínas adyacentes había sido identificada, por lo que no conocíamos los bordes donde terminaba la proteína en la densidad electrónica. En aquel momento no teníamos ningún entrecruzamiento vinculando a B0SMT6 con el filamento, sin embargo por espectrometría de masas esta proteína se encuentra ausente tanto en los mutantes *flaA2*- como *fcpA*-.

A la izquierda de B0SMT6 se encuentra una densidad electrónica que fue interesante de modelar. Inicialmente, e incluso en modelos atómicos previamente publicados, esta densidad se había interpretado como un protómero adicional de FcpB. Sin embargo, ahora con mejor resolución, notamos que la densidad crio-EM no coincidía bien con FcpB, principalmente en sus regiones de superficie (Figura 38A). A su vez, ahora contábamos con datos complementarios de XL-MS. La proteína B0SKN8 parecía ser un mejor candidato en esta posición, pues interactúa con FcpA por dos inter-entrecruzamientos diferentes (no así con B0SMT6). Adicionalmente, y de forma recurrente en todas las réplicas de MS, B0SKN8 está ausente tanto en los mutantes *flaA2*- como *fcpA*-.

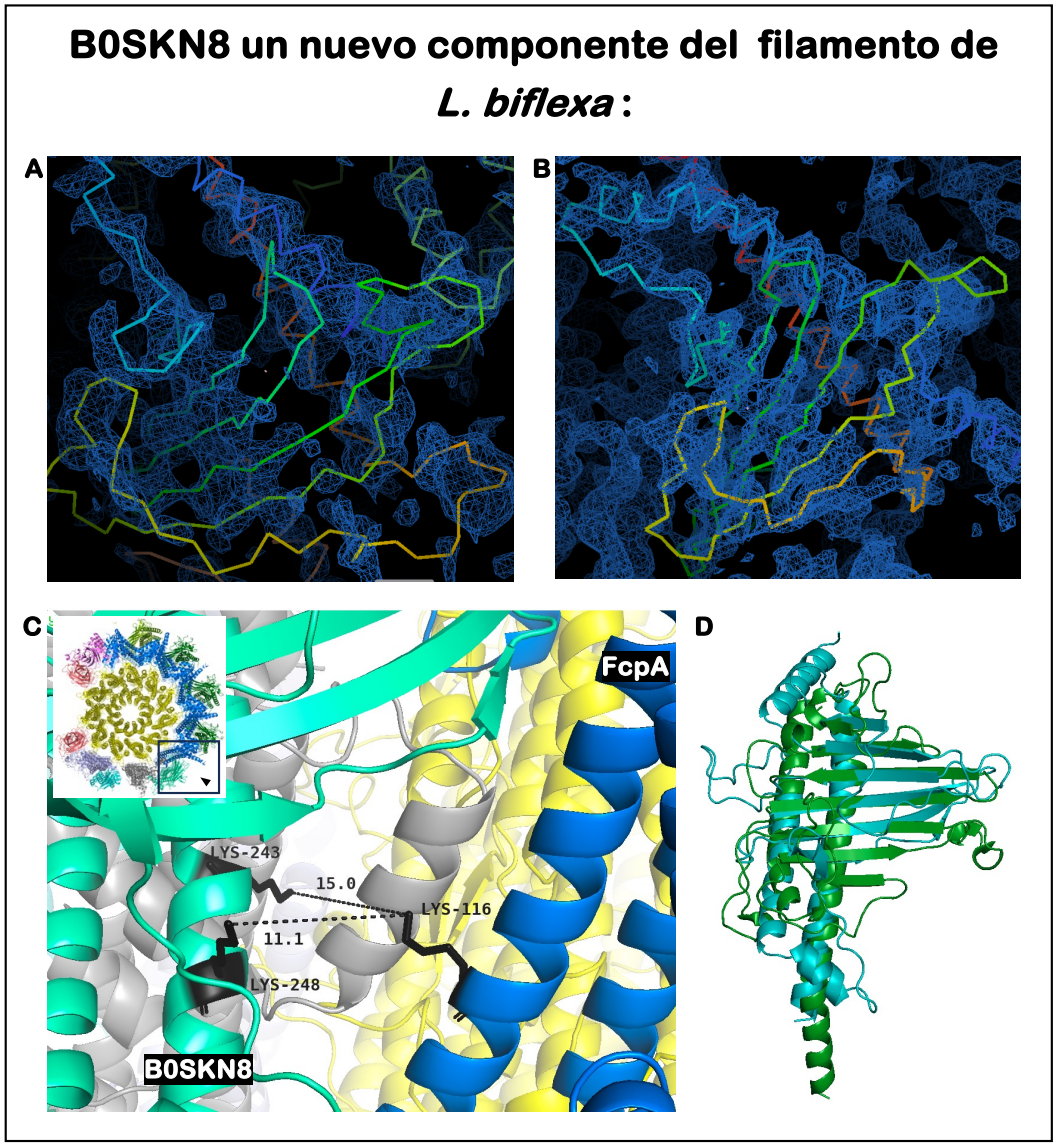


Figura 38: B0SKN8 un nuevo componente del filamento flagelar de *L. biflexa* salvaje. A. Estructura de FcpB coloreada con una rampa de azul (extremo N-terminal) a rojo (C-terminal), alineada sobre la densidad crio-EM. **B.** Modelo 3D inicial de B0SKN8 predicho con AlphaFold2 (también en una rampa de azul a rojo), sobre la misma densidad que en el panel A. Su posicionamiento y refinamiento automático inicial es mucho mejor que el de FcpB. **C.** Vista magnificada, mostrando los dos entrecruzamientos entre FcpA (en azul) y B0SKN8 (en verde agua). El residuo K116 de FcpA interactúa con los residuos K243 y K248 de B0SKN8. **D.** Modelos finales superpuestos de B0SKN8 (en verde agua) y FcpB (en verde inglés). Ambas comparten ~14.5% de identidad de secuencia y un rmsd final de 2.2 Å.

La estructura 3D de B0SKN8 predicha por AlphaFold2 reveló además una arquitectura semejante a la de FcpB (aun cuando su superposición estructural no era inicialmente buena [rmsd = 7.5 Å] y con una identidad de secuencia también baja ~14.5%). Luego del refinamiento automático y manual sobre el mapa crio-EM (Figura 38B), B0SKN8 fue rearreglada ajustando mejor en la densidad, y dicho modelo refinado ahora pasó a superponerse con un rmsd de 2.2 Å comparada con FcpB (Figura 38D) confirmando la validez de nuestras observaciones.

Aun así, observamos que permanecía una densidad adicional entre ella y su vecina B0SMT6, demasiado grande para ser explicada como un posible carbohidrato, aunque en principio muy pequeña para ser una proteína entera (Figura 39A). Este contratiempo pudo ser resuelto al encontrar a B0SMK8. El descubrimiento de que varias proteínas de vaina forman "familias estructurales" aun exhibiendo bajos niveles de identidad de secuencia, nos condujo eventualmente a alinear todos los candidatos a proteínas flagelares entre sí, y de este modo logramos descubrir a B0SMK8, una proteína que alineaba estructuralmente muy bien con B0SMT6 (rmsd 1.2 Å e identidad de secuencia de ~22.3%, Figura 39C).

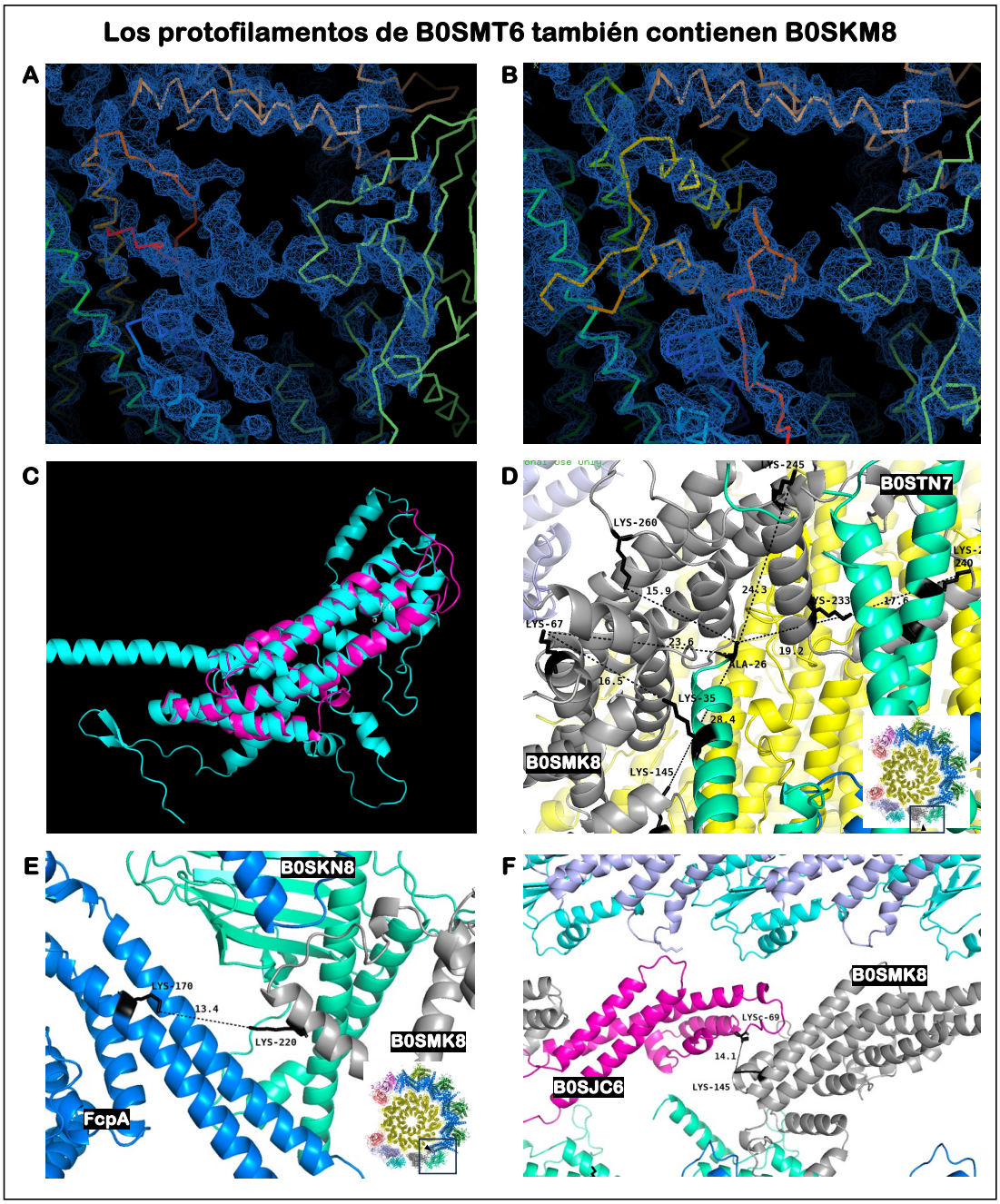


Figura 39: Los protofilamentos de B0SMT6 en el filamento flagelar de *L. biflexa* salvaje, parecen también contener B0SMK8, probablemente de una forma alternada. A. Se observa el modelo superpuesto a la densidad

electrónica en la región de contacto entre B0SMT6 (en una rampa de colores de azul a rojo, marcando la dirección N- a C-terminal) y B0SKN8 (en verde, a la derecha). Entre ambas proteínas queda una densidad electrónica no explicada por ninguno de los dos modelos refinados. **B.** Modelo superpuesto sobre la densidad crio-EM en la región de contacto entre B0SMK8 (en una rampa de colores de azul a rojo) y B0SKN8 (en verde a la derecha) que explica mejor la densidad electrónica en esta región. **C.** Modelos superpuestos de B0SMT6 (en rosado) y B0SMK8 (en cian) que exhiben ~22.3% de identidad de secuencia, y un rmsd de 1.2 Å. **D.** Ocho entrecruzamientos covalentes entre B0SMK8 (en gris) y B0SKN8 (en verde agua) son mediados por el agente NNP9. El residuo de K22 de B0SKN8 interactúa con los residuos K67, K145, K233, K245 y K260 de B0SMK8, el residuo K35 de B0SKN8 interactúa con el residuo K67 de B0SMK8, y los residuos K240 y K243 de B0SKN8 interactúan con el residuo K233 de B0SMK8. **E.** El agente NNP9 media un único inter-entrecruzamiento entre el residuo K220 de B0SMK8 (en gris) y K170 de FcpA (en azul). **F.** Sorprendentemente, se observa un inter entrecruzamiento entre el residuo K69 de B0SMT6 (en rosado) y K145 de B0SMK8 (en gris) que, a diferencia de todos los demás entrecruzamientos parece estarse produciendo dentro de un mismo protofilamento y no entre distintos protofilamentos de las proteínas de vaina.

En ese momento, al no presentar ningún entrecruzamiento con proteínas flagelares conocidas, B0SMK8 aparecía únicamente en los resultados de MS, pero de forma reproduciblemente ausente en las cepas *flaA*- y *fcpA*-. Si bien B0SMK8 es más grande, y explica por ende mejor el mapa de densidad que B0SMT6 (Figura 39B), al buscar esta proteína específica en nuestra base de datos de entrecruzamiento encontramos algo sorprendente: B0SMK8 interacciona efectivamente con FcpA (mediante un entrecruzamiento), con B0SKN8 (mediante ocho inter-entrecruzamientos) y, sorprendentemente, también con B0SMT6 (mediante un único inter-entrecruzamiento) haciendo de puente entre ellas en la red de interactores (Figura 39F).

Esto indicaría que, de hecho, B0SKM8 y B0SMT6 polimerizan entre ellas dentro de un mismo protofilamento. Teniendo este modelo hipotético como escenario, la distancia predicha entre la Lys145 de B0SMK8 y la Lys69 de B0SMT6 que se entrecruzan es de ~14 Å. Asimismo, la formación de un protofilamento que contenga ambas proteínas explicaría por qué el mapa de densidad electrónica se ve un poco peor en aquellas regiones donde las estructuras de ambas proteínas difieren; al haber copias divergentes de estructuras dentro de una misma región del mapa promediado, este se ve con peor resolución local.

Como se muestra en la Figura 40, nos quedaba entonces una única densidad por interpretar, la más pequeña de las densidades hasta el momento descritas. De todos los candidatos de proteínas flagelares que hemos podido listar, el más pequeño era B0STN7, una proteína exclusiva de la cepa salvaje en comparación con la *fcpA*- (aunque también presente en *flaA2*-) que, de forma consistente en todos los experimentos realizados, acompañaba la desaparición de las proteínas ya descritas B0SMK8 y B0SKN8. Sin embargo, la estructura 3D de B0STN7 predicha con AlphaFold2, no coincidía con la forma de la densidad "semi-globular" visible en el mapa crio-EM (Figuras 40A y B).

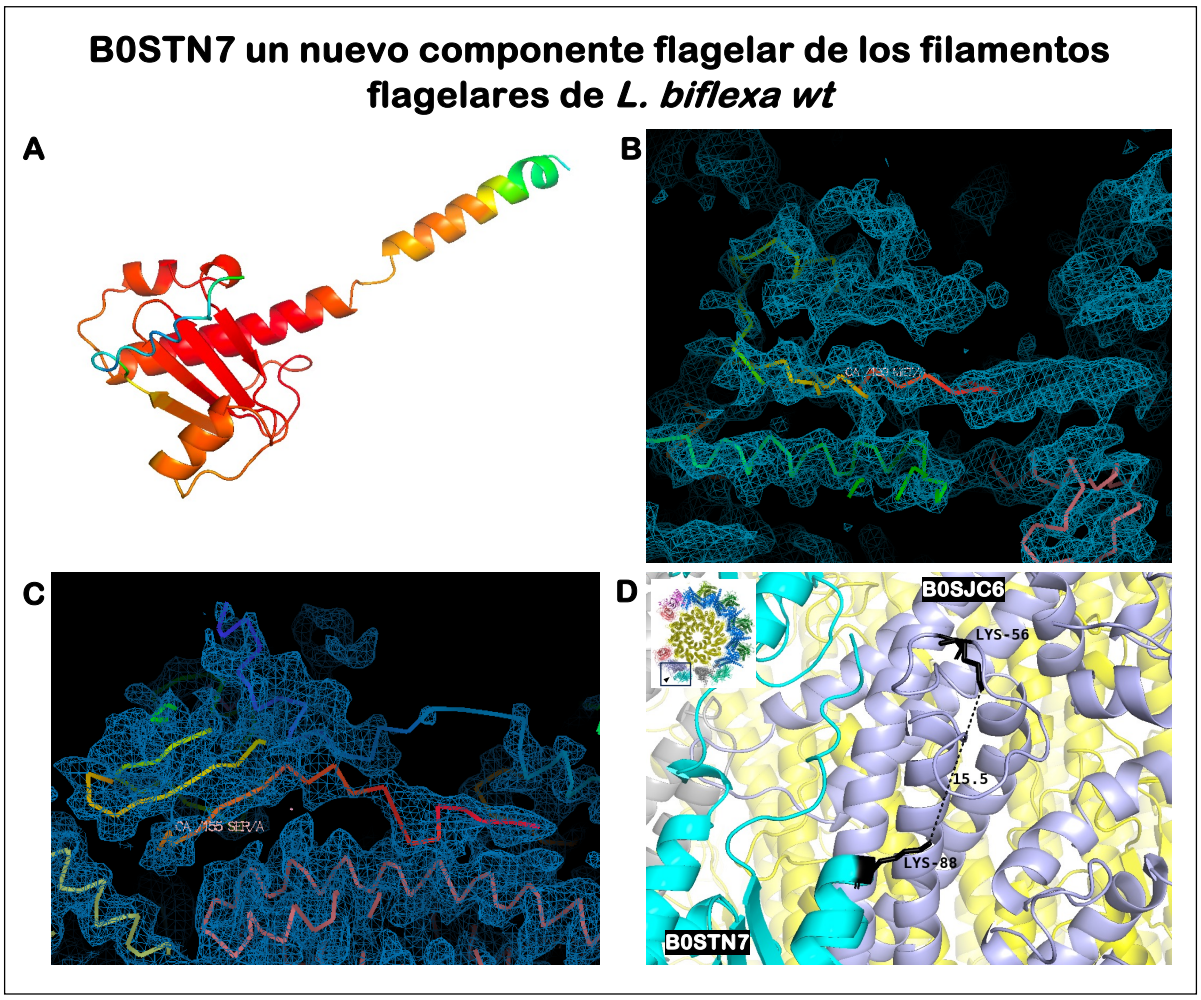


Figura 40: B0STN7, otra nueva proteína flagelar de *L. biflexa* salvaje, el componente más pequeño del filamento. A. Modelo atómico de B0STN7 predicho con AlphaFold2, coloreado según sus valores pLDDT, con una rampa del rojo (mejor calidad de predicción) al azul (menor calidad). **B.** Primer posicionamiento de B0STN7 en el mapa crio-EM. Aproximadamente la mitad de la densidad electrónica que se ve en la figura no se explicaba inicialmente con el modelo predicho, el cual se ilustra aquí coloreado con una rampa desde el azul (extremo N-terminal) al rojo (C-terminal). **C.** Reconstrucción manual/interactiva de los extremos N- y C-terminales de B0STN7 dentro de la densidad crio-EM. Nótese que el extremo N-terminal de una proteína se despliega para contactar y plegarse sobre el monómero adyacente en el protofilamento. **D.** El agente NNP9 media un único inter-entrecruzamiento entre el residuo K88 de B0STN7 (en cian) y K56 de B0SJC6 (en violeta).

Fue la dilucidación de la red final de interactores, obtenida por XL-MS, la que nos llevó a dilucidar esta aparente discordancia. El resultado clave fue haber detectado un entrecruzamiento entre B0STN7 y B0SJC6 (previamente descrita como FlaAP) en un experimento de XL-MS. El refinamiento inicial automático e interactivo del modelo AlphaFold2 de B0STN7 se muestra en la Figura 40B. El extremo C-terminal (en rojo en la figura) se pliega hacia atrás en el modelo predicho, por lo que fue necesario reconstruirlo *de novo* en forma manual/interactiva para poder ajustar su estructura dentro del mapa crio-EM. La construcción del extremo N-terminal por su parte resultó más desafiante, debido a que proviene del monómero adyacente y no se ve densidad crio-EM en la conexión entre ambas.

El entrecruzamiento entre los residuos de Lys de B0STN7 y B0SJC6 tiene una distancia de ~15.5 Å, lo cual valida el posicionamiento relativo entre ambos componentes.

Con la red final de entrecruzamientos completa, también notamos que es FlaA2 y no FlaA1 quien se encuentra entrecruzada con B0STF2. Esto nos llevó a una inspección cuidadosa del mapa crio-EM en la región de los protómeros de FlaA, encontrando que sí, FlaA1 tiene de hecho dos bucles no explicados por el mapa: uno entre los residuos K71 y L102 que se encuentra en la superficie expuesto al solvente, y más importante aún, otro entre los residuos P147 y G174 que si se modelara coincidiría en espacio con B0STF2. Al superponer y comparar ambas estructuras sobre la densidad electrónica se observa además que la coincidencia de FlaA2 con el modelo es en general mucho mayor (Figura 41). Por lo tanto donde inicialmente fue modelada FlaA1, se sitúa otra copia de FlaA2.

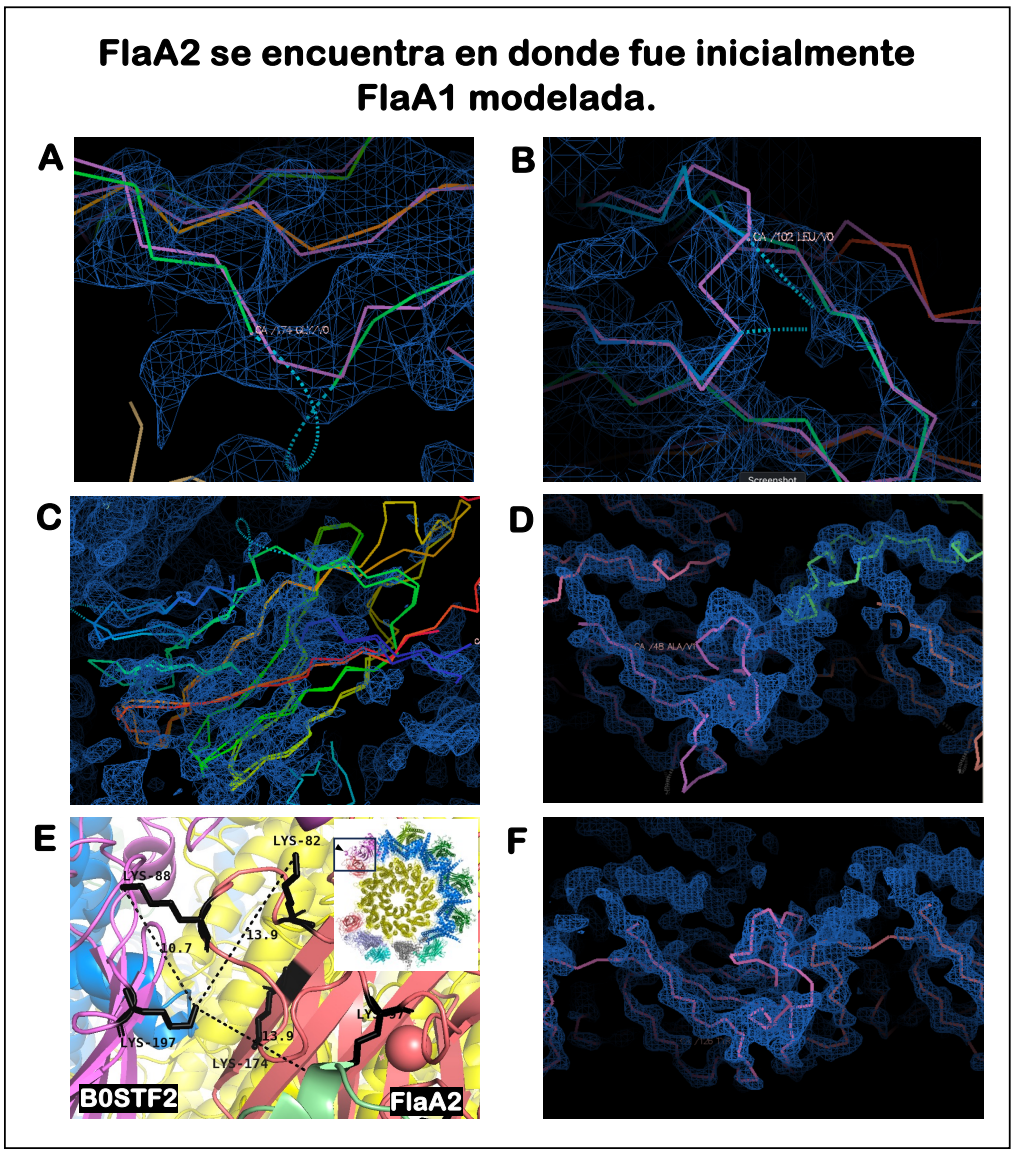


Figura 41: En los filamentos flagelares de *L. biflexa* salvaje FlaA2 se encuentra en donde FlaA1 fue inicialmente modelada. **A.** Estructuras de FlaA1 (en verde) y FlaA2 (en violeta) de *L. biflexa* sobre la densidad electrónica en el bucle entre P147 y G174. **B.** Ídem A en el bucle entre K71 y L102 que se encuentra en la superficie expuesto al solvente. **C.** Estructuras de FlaA1 y FlaA2 superpuestas sobre la densidad electrónica. Ambas isoformas de FlaA comparten un 29.5% de identidad de secuencia y alinean con un rmsd de 1.48 Å. **D.** Estructuras de FlaA1 de *L. biflexa* alineada sobre la densidad electrónica. **E.** Cuatro entrecruzamientos químicos entre B0STF2 (en rosado) y FlaA2 (en rojo) son mediados por el agente NNP9, validando el posicionamiento de FlaA2 en esta región del modelo. El residuo de K197 de B0STF2 se entrecruza con los residuos K82, K88, K97 y K174 de FlaA2. **F.** Ídem D. pero con el modelo de FlaA2 en vez de FlaA1. En esta región el modelo de FlaA2 explica mucho mejor la densidad electrónica.

Dada la similitud estructural sustancial entre ambas isoformas (rmsd de 1.5 Å) y razonando que FlaA1 debía de estar presente en los filamentos salvajes, se justifica nuestra suposición inicial errónea. Fue para nosotros intrigante que FlaA1, una de las pocas proteínas que sí se esperaba encontrar en el filamento, no formara parte del modelo final. Se sabe que FlaA1 es una proteína del filamento y que está presente en las preparaciones flagelares siendo detectada por espectrometría de masas *bottom up*. Nuestro análisis de entrecruzamiento indica que estaría interactuando con FlaA2, por lo que es posible que FlaA1 esté en el surco que se ve entre las dos copias de FlaA2 y que se pierda en una alta proporción de los filamentos durante la purificación. Esto explicaría por qué una pequeña proporción de las interacciones transitorias pueden ser capturadas por XL-MS y por qué se observa un gran hueco en la vaina del filamento que por lo demás es cilíndrica, como todos los filamentos flagelares observados hasta el momento. Esta hipótesis está apoyada además por los resultados de entrecruzamiento de FlaA1 consigo misma, ya que para FlaA1 observamos un homo-inter-entrecruzamiento entre sus residuos de Lys88 lo que indicaría que FlaA1 dimeriza.

Nuestro modelo final del filamento nativo de *L. biflexa* salvaje (manuscrito en preparación para ser enviado próximamente a publicación) (Figura 42), al momento contiene 146 cadenas pertenecientes a 10 proteínas diferentes. Una de ellas corresponde a la isoforma dominante de FlaB que componen el núcleo del filamento (Uniprot B0SSZ5, - anotada en el genoma de *L. biflexa* como FlaB2, pero FlaB1 a los efectos de esta tesis- y ortóloga de FlaB4 en *L. interrogans* donde se demostró es la isoforma más abundante (Malmström et al, 2009)), mientras que los 9 restantes envainan al núcleo interno, quedando expuestas en la superficie del apéndice en mayor o menor grado. A esto debemos agregar que (i) B0SMK8 parecería estar alternándose con B0SMT6 (no tenemos suficiente resolución en el mapa para definir protofilamentos alternando en el modelo atomístico); (ii) que FlaA1 probablemente se encuentre en los filamentos nativos intactos, “llenando” el surco lateral en la cara cóncava; y, que (iii) las otras tres isoformas de FlaB probablemente también se encuentren dispersas (homogénea, o heterogéneamente) dado que se detectan en el proteoma. Considerando dichas notas, el número total de proteínas del filamento parecería ser de al menos 15 proteínas diferentes: 4 en el núcleo y 11 conformando la vaina.

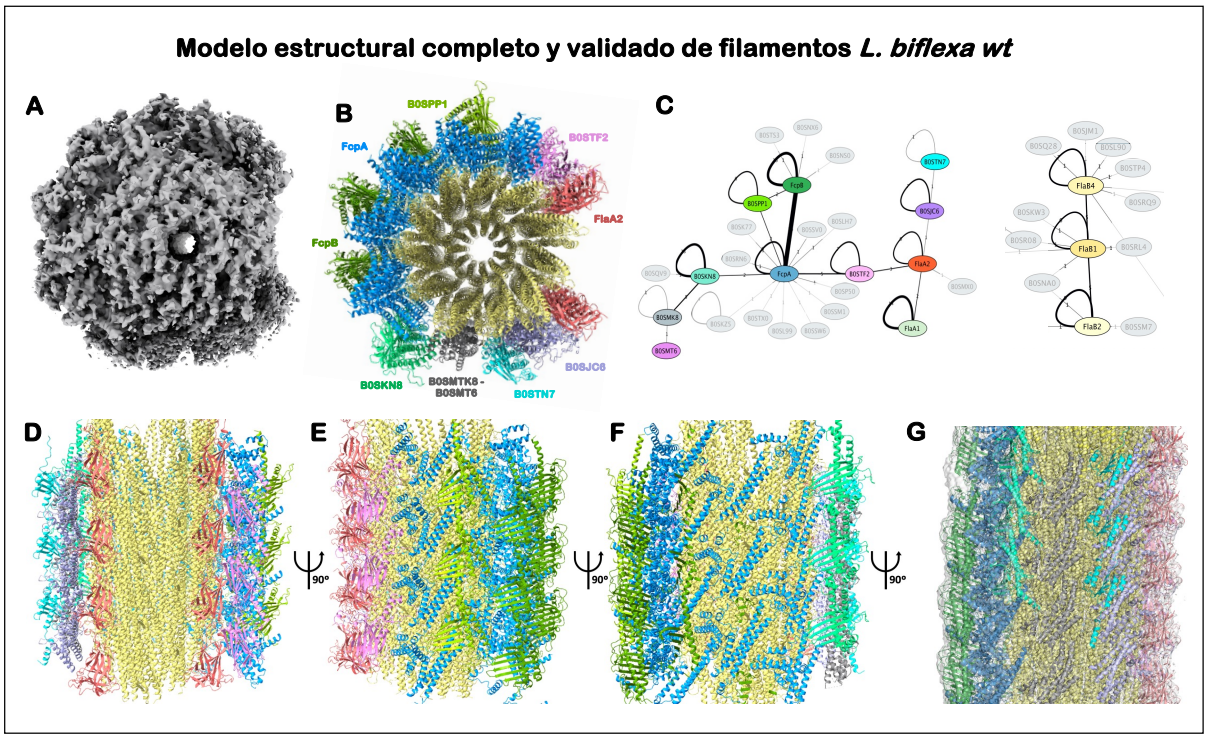


Figura 42: Modelo estructural completo y validado del filamento flagelar de *L. biflexa* salvaje. A. Vista en corte transversal del mapa obtenido de crio-microscopía electrónica de los filamentos de *Leptospira biflexa* salvaje a 3.7 Å de resolución. **B.** Vista en corte transversal del modelo molecular del filamento generado en esta tesis. La arquitectura del filamento consta de un núcleo central de proteínas FlaB (en amarillo) organizadas en 11 protofilamentos formando una estructura simétrica cilíndrica con un canal hueco de ~20 Å de diámetro en el centro. Envainando al núcleo, encontramos una red cuasi-concéntrica de proteínas que recubre el núcleo central, pero que se ve interrumpida por un surco lateral entre los protofilamentos de FlaA2. Esta vaina consta así de protofilamentos a lo largo del eje mayor del filamento, que incluyen 10 componentes proteicos distintos, y que se muestran en diferentes colores: en rojo dos protofilamentos de FlaA2 (uno a cada lado del surco lateral), en rosado un protofilamento de la proteína B0STF2, en azul cinco protofilamentos de FcpA, en verde claro un protofilamento de B0SPP1, en verde más oscuro tres protofilamentos de FcpB, en verde agua un protofilamento de B0SKN8, en gris el único protofilamento que está compuesto por dos componentes alternantes distintos: B0SMT6 y B0SMK8, en cian un protofilamento de B0STN7 y en violeta un protofilamento de FlaAP (cuyo código Uniprot es B0SJC6). **C.** La red de entrecruzamientos químicos con el agente NNP9 obtenida por XL-MS valida el modelo final obtenido. **D.** Vista lateral del modelo en B. Vista desde el surco lateral. **E.** Ídem al panel D., con vista desde B0SPP1. **F.** Ídem al panel D., con vista frontal opuesta. **G.** Ídem al panel D., con vista frontal en los filamentos de B0SMK8-B0SMT6, en esta perspectiva sólo se observa la primera de ambas proteínas y se superpone el modelo atómico sobre la densidad crio-EM para subrayar su buen ajuste.

Los parámetros que miden la calidad del modelo son buenos; con valores de desviación globales (rmsd) de ángulos y distancias de enlace de, respectivamente 1.7° y 0.008 Å con respecto a valores ideales. Se observa una buena concordancia entre el mapa derivado de nuestro modelo y el de la densidad observada por microscopía (*Fourier Shell Correlation* promedio de 0.78), sobre todo considerando que se trata de un modelo muy grande. Además, este mapa queda validado por los resultados de entrecruzamiento químico; todos los interactores esperados del modelo (a excepción sólo de B0STN7-B0SMT6/B0SMK8), efectivamente se encuentran en la red final de XL-MS que se muestra en la Figura 36.

Un aspecto importante que también destaca de la red final de interacciones es que el núcleo del filamento flagelar y la vaina no tienen ningún nodo en común. De todos los experimentos XL-MS llevados a cabo, un único entrecruzamiento entre el núcleo flagelar (FlaB1) y los componentes que lo envainan (B0SJC6) fue detectado en el *trial* 1, y su *score* era de los más bajos aceptados estadísticamente; de hecho, el mismo se encuentra entre los pocos que no coinciden con el modelo final. La falta de entrecruzamientos que validen estas interacciones puede deberse a dos motivos:

(i) La primera razón podría ser que el agente NNP9 sea incapaz de cruzar la vaina para alcanzar la zona de interacción con el núcleo; y que los entrecruzamientos observados entre proteínas del núcleo provengan de aquellos segmentos del filamento que por microscopía se han observado desprovistos de vaina (ver sección 3.4.2).

(ii) La segunda razón, y a nuestro entender la más probable, podría ser que prácticamente todas las interacciones entre las isoformas de FlaB y las proteínas de la vaina, involucran carbohidratos. En la sección 3.4.2 se describió como estos glicanos se encuentran modificando residuos de serina y treonina en FlaB4, en general interaccionando con residuos positivamente cargados (incluidas argininas y lisinas) de FlaA2, FlaAP y muy probablemente otras proteínas de vaina, que constituyen exactamente los mismos residuos reactivos que tiene el agente NNP9 para actuar. Los carbohidratos pueden de esta forma ocupar los sitios de unión del agente de entrecruzamiento en ambos interactores. Aun más, incluso si algún entrecruzamiento estuviera ocurriendo en esa región, la masa extra de los glicanos sin identificar y sin considerar sobre esos péptidos, estaría obstruyendo su detección por espectrometría de masas.

Es de destacar que por XL-MS se detectan 37 entrecruzamientos que involucran interacciones en y entre FlaBs. La espectrometría logra distinguir entre las cuatro isoformas homólogas y de hecho sólo tres de las cuatro están presentes en el mapa de la Figura 36B. La ausencia de FlaB3 en este mapa no implica su ausencia en la muestra, de hecho se detectan marcaciones "*dead end*" con el agente NNP9 para FlaB3 en cuatro de los 11 ensayos (Figura 29C), observándose en total un único "*loop link*" y tres "*dead ends*" distintos. Estos resultados son consistentes con el hecho de que FlaB3 es la isoforma menos abundante en el núcleo interno del filamento y está presente únicamente en cantidades muy bajas[134].

De las tres isoformas que sí muestran entrecruzamientos, notamos que no todas interactúan entre sí; FlaB1 interacciona tanto con FlaB2 (dos inter-entrecruzamientos) como con FlaB4 (tres inter-entrecruzamientos), pero en ninguno de los 11 juegos de datos se ve un entrecruzamiento que conecte estas últimas dos isoformas de FlaB entre sí. Es posible que FlaB2 y FlaB4 efectivamente interactúen, y que simplemente no hayan sido capturadas por el agente de entrecruzamiento. Sin embargo, considerando que FlaB4 es la isoforma mayoritaria

en los filamentos de *L. biflexa* salvaje, y dado que el entrecruzamiento típico y muy abundante entre el extremo N-terminal y la K17 conservada en todas las isoformas de FlaB se observa entre monómeros FlaB4 y en sus interacciones con FlaB1 pero no con FlaB2, resulta más probable que haya un arreglo específico de isoformas: que FlaB4 y FlaB2 estén físicamente separadas en el núcleo, por protofilamentos de FlaB1.

3.4.4 – Estructura inicial del filamento flagelar salvaje de *L. interrogans*

Finalmente, para investigar si podíamos ubicar a FlaA1 como parte del filamento flagelar nativo en *Leptospira*, se purificaron a partir de la cepa salvaje de otra especie, *L. interrogans*. El laboratorio del Dr C Sindelar (Univ de Yale) logró obtener reconstrucciones 3D de dichos filamentos por crio-EM utilizando un abordaje de análisis de partículas únicas, con un protocolo similar al utilizado para resolver la estructura de los filamentos de *L. biflexa*. Como se observa en la Figuras 43B y E, si bien la resolución final alcanzada en este mapa (4.8 Å) es más pobre que la obtenida para *L. biflexa*, la especie patógena muestra una vaina que recubre completamente al núcleo central de FlaB, y particularmente no hay un surco lateral en el lado cóncavo.

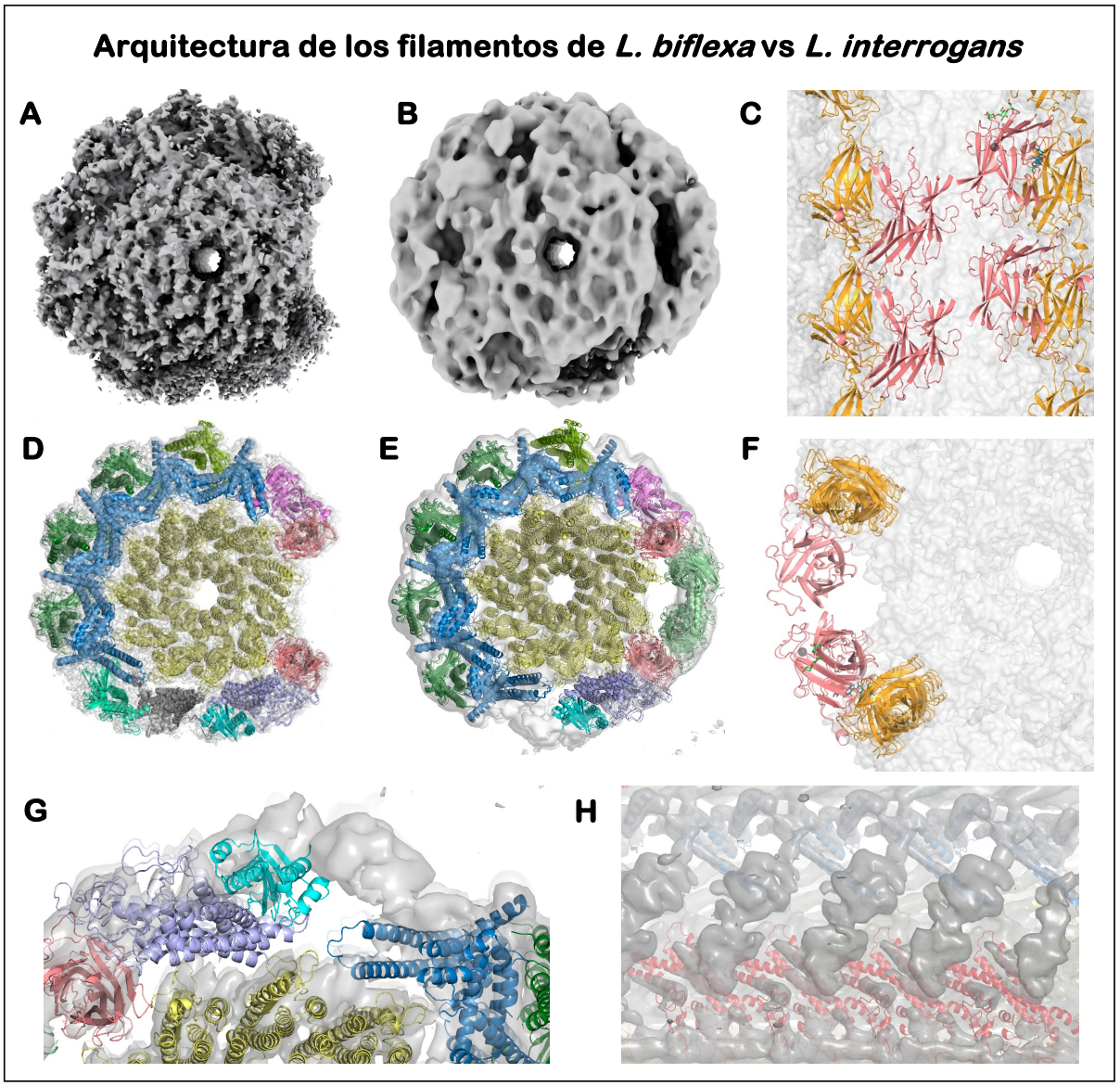


Figura 43: Arquitectura de los filamentos de *L. biflexa* vs *L. interrogans*. A. Vista en corte transversal del mapa crio-EM de filamentos de *Leptospira biflexa* salvaje a 3.6 Å de resolución. **B.** Vista en corte transversal del mapa crio-EM de filamentos de *Leptospira interrogans* salvaje a ~4.8 Å de resolución. **C.** Vista lateral del mapa de crio-EM de filamentos de *L. interrogans* mostrando la interacción de FlaA1 (rosado) con FlaA2 (anaranjado). **D.** Ídem al panel A. con el modelo molecular superpuesto de *L. biflexa* salvaje que no contiene FlaA1, resultando en la aparición de un surco lateral en la vaina. Se muestran los protofilamentos de FlaB en amarillo, de FlaA2 en rojo, de B0STF2 en rosado, de FcpA en azul, de B0SPP1 en verde claro, de FcpB en verde más oscuro, de B0SKN8 en verde agua, de B0SMT6 - B0SMK8 en gris, de B0STN7 en cian y en violeta de FlaAP. **E.** Ídem al panel B. con el modelo molecular superpuesto de *L. interrogans* salvaje. Estos filamentos contienen dos protofilamentos de FlaA1 (en verde claro), completando el surco lateral de la vaina. Además, se observan otras diferencias con respecto a los filamentos de *L. biflexa*: los de *L. interrogans* poseen seis protofilamentos de FcpA en lugar de cinco, cuatro protofilamentos de FcpB en vez de tres, y al menos una porción pequeña de densidad, aún no identificada que está ausente en los otros filamentos analizados. Por otro lado, los filamentos de la especie patógena carecen de protofilamentos de B0SKN8, B0SMT6-B0SMK8. **F.** Vista en corte transversal del mapa crio-EM de filamentos de *L. biflexa*, con FlaA1 de *L. biflexa* modelada y superpuesta sobre la isoforma de *L. interrogans*. **G.** Imagen magnificada del panel E. enfocando la región con el componente no identificado en filamentos de *L. interrogans* salvaje. **H.** Vista lateral del panel G, en la región de la densidad crio-EM no identificada y B0STN7 (no mostrada).

Si bien la interpretación de este mapa en términos de un modelo atómico completo está en curso, podemos afirmar con confianza que efectivamente, un dímero de FlaA1 se ubica "llenando" el surco lateral de *L. biflexa*, formando de hecho dos protofilamentos que corren en forma antiparalela. Estos dos protofilamentos de FlaA1 sirven así de puente entre los dos protofilamentos de FlaA2, cerrando completamente la vaina alrededor del núcleo de FlaB (Figura 43C). Al ser antiparalelos, las superficies de interacción de cada protómero de FlaA1 con su vecino FlaA2 son diferentes, ya que los dos protofilamentos de FlaA2 exponen caras opuestas de la proteína hacia el "surco lateral" donde se ubica el dímero de FlaA1. Por otro lado, se observa un espacio entre FlaA1 y el núcleo interno de FlaBs, distinto al poro central por donde se exportan las subunidades desplegadas de FlaB para polimerizar. Si bien la presencia de este segundo hueco resulta un tanto inesperada, no parece ser un artefacto de la técnica de crio-EM, ni una cuestión de resolución; de hecho si quisiéramos desplazar los monómeros de FlaA1 hacia el centro, para que hagan contacto con el núcleo de FlaBs, no podría entrar un dímero sin generar colisiones entre proteínas vecinas.

Otro aspecto sorprendente que descubrimos durante el análisis y modelado de los filamentos de *L. interrogans*, es que la composición y estructura de su vaina son distintas a la de *L. biflexa*. En particular la región que contiene B0SKN8, B0SMK8 y B0SMT6 en *L. biflexa* se sustituye por una copia adicional de FcpA y FcpB y por una proteína nueva aún no identificada (Figuras 43E, G y H). Dado que la resolución del mapa crio-EM de *L. interrogans* salvaje es pobre en esta región, aún resta identificar este nuevo componente flagelar (Figura 43H) para completar el modelo.

IV– DISCUSIÓN Y PERSPECTIVAS

La motilidad flagelar única de las espiroquetas les permite nadar eficientemente a través de medios viscosos, colonizando nichos inalcanzables para otras bacterias y explicando su éxito como patógenos[89]. Esta ventaja selectiva se debe a sus flagelos periplásmicos que se enrollan alrededor del cuerpo celular, provocando la rotación de la célula sin aparatos externos[171]. Además de facilitar el movimiento, los endoflagelos tienen funciones citoesqueléticas que influyen en la morfología espiral característica de espiroquetas[172]. Estas particularidades funcionales están acompañadas de una estructura flagelar más compleja, evolutivamente adaptada para producir mayor torque[10]. Aunque estudios de crio-ET han permitido avanzar en el entendimiento de su arquitectura flagelar[71,173], aún se desconocen los detalles estructurales completos de sus componentes menos conservados como el filamento flagelar. A pesar de compartir ciertas similitudes con los filamentos de flagelos típicos de enterobacterias, los PF de espiroquetas presentan una vaina proteica cuya función y estructura aún no han sido dilucidadas[82].

4.1 – Enfoque metodológico integrativo para modelar el filamento de *Leptospira*

Para modelar la arquitectura del filamento flagelar de *Leptospira* empleamos una estrategia integrativa que combinó varios enfoques y tecnologías. Las estructuras de los componentes conocidos fueron obtenidas por aproximaciones "bottom-up" que incluyeron el modelado por homología de las isoformas de FlaB y la resolución cristalográfica de FlaA1, FcpA y FcpB. A medida que la resolución de los mapas de criomicroscopia electrónica aumentó, se hizo evidente una complejidad estructural mayor a la esperada, por lo que fue esencial integrar estrategias "top-down" analizando filamentos flagelares completos.

El análisis de la composición diferencial por espectrometría de masas de filamentos purificados de cepas salvajes y mutantes, combinado con información del entrecruzamiento químico acoplado a espectrometría de masas (XL-MS), permitió identificar nuevos componentes flagelares y dilucidar sus interacciones. Estas metodologías complementarias resultaron en una sinergia poderosa para sortear los desafíos planteados por la baja pureza de nuestras preparaciones flagelares y la complejidad estrucutral del filamento. Si bien desconociamos dónde estaban los límites entre una proteína y otra en el mapa de crio-EM y fue necesario evaluar aprox 30 candidatos flagelares diferentes, el surgimiento del algoritmo AlphaFold en el momento preciso permitió la obtención de modelos estructurales para cada candidato. El ensamblado in silico manual, facilitado por herramientas bioinformáticas, nos permitió ubicar cada pieza del "rompecabezas" estructural para obtener un modelo final completo y detallado de la estructura molecular del filamento *de L. biflexa*.

Los parámetros de calidad de este modelo son sólidos: los ángulos y distancias de enlace se desvían respectivamente 1.710º y 0.013Å de los parámetros ideales. La densidad electrónica experimental, mejorada a una resolución de 3.8 Å, muestra buena concordancia (FSC promedio de 0.79) con el mapa derivado de nuestro modelo. Los interactores esperados , salvo B0STN7-B0SMT6/B0SMK8, efectivamente están presentes en la red final de XL-MS y sus distancias de entrecruzamiento mapeadas sobre el modelo final coinciden en su mayoría con las esperadas para el agente NNP9, validando nuestro enfoque integrativo para modelar esta estructura compleja.

4.2 – Arquitectura del filamento de *Leptospira*

El modelo final del filamento revela una arquitectura fascinante e inesperadamente compleja, con 15 componentes diferentes: FlaB1-4, FlaA1-2, FcpA, FcpB, FlaAP (B0SJC6), B0SMT6, B0SKM8, B0STN7, B0STF2, B0SKN8 y B0SPP1; los últimos siete fueron identificados en esta tesis. Este heteropolímero mantiene características conservadas con otros filamentos bacterianos, como el núcleo helicoidalmente simétrico formado por proteínas homólogas a flagelina, que polimerizan formando 11 protofilamentos alrededor de un eje hueco central. Como en otras espiroquetas, la superficie de este polímero está recubierta por una vaina que contiene FlaA, pero en *Leptospira* esta estructura contiene no uno sino once componentes distintos, convirtiéndolo en el filamento con arquitectura más compleja del filo espiroquetal y del reino bacteriano en general.

Una característica notable es la organización asimétrica radial de la vaina proteica. Observando el filamento en corte axial, la cara convexa del superenrollamiento presenta múltiples copias de FcpA y FcpB organizadas en capas concéntricas, con FcpB en la superficie[169]. En el lado cóncavo se ubican las proteínas FlaA: dos copias de FlaA1 diametralmente opuestas a FcpA y flanqueadas por una copia de FlaA2 a cada lado. Al observar los filamentos en corte longitudinal, cada componente de la vaina se ensambla en protofilamentos específicos, de modo que la mayoría son homopolímeros, con la excepción del filamento compuesto por B0SMK8 y B0SMT6 cuya disposición precisa aún se desconoce.
Nuestro modelo explica casi todas las observaciones previas sobre los filamentos de *Leptospira*, desde las múltiples bandas en los SDS-PAGE de filamentos purificados[74,75] y las cantidades relativas de proteínas flagelares cuantificadas por Malmström et al.[134], hasta los resultados de inmunomarcado con anticuerpos anti-FcpA, anti-FcpB (que solo se unen al lado convexo de los filamentos de cepas *wt*) y anti-FlaB (que solo marcan la superficie de endoflagelos provenientes de cepas *fcpA*$^{-}$) [105,168]. Asimismo, el modelo es consistente con los resultados de ensamblaje secuencial (donde FcpA se recluta antes que FcpB) y con los diámetros medidos para los filamentos de cepas salvajes (21.5nm) y los distintos mutantes FcpA, FcpB (13-16nm)[105,168] y flaA2-. Sin embargo, nuestros resultados difieren de lo propuesto por Sasaki et al.[76], quienes sugieren que en *L. biflexa* FcpA interacciona con la

región C-terminal de FlaA2 y la región N-terminal de FlaB1 cuando, según nuestro modelo, FlaA2 y FcpA están separadas por al menos una proteína de distancia (B0STF2) y la región N-terminal de todas las isoformas de FlaB se encuentra en la región más central, alejada de la vaina.

4.3 – Núcleo del filamento

La característica estructural más conservada de las espiroquetas es su núcleo interno con simetría helicoidal, compuesto de proteínas homólogas a flagelina. Nuestros alineamientos y modelos muestran que las isoformas de FlaB están ampliamente conservadas con los dominios D0 y D1 de FliC de S. enterica. Estos dominios permiten la polimerización alrededor de un cilindro hueco formando una superhélice de 11 protofilamentos, universalmente conservada en todos los flagelos bacterianos[12,34]. Se ha postulado que debido al impedimento estérico que genera, este número de protofilamentos es ideal para facilitar el deslizamiento entre ellos, alcanzando un estado de energía mínimo naturalmente superenrollado que permite su función[25,34,169].

Por otra parte, a diferencia de la flagelina convencional, las isoformas de FlaB de espiroquetas carecen completamente de los dominios hipervariables D2 y D3. Sin embargo, lejos de simplificar la estructura, los endoflagelos reemplazan estos dominios de superficie por polipéptidos adicionales. La codificación de la vaina en proteínas separadas permite mayor variabilidad estructural, añadiendo un nivel de complejidad adicional sobre una base conservada de auto-polimerización, que incrementa mínimamente el diámetro flagelar pero añade complejidad funcional[174].

Exceptuando *B. burgdorferi*[17,36], la mayoría de las espiroquetas tienen múltiples isoformas de FlaB, generalmente tres[87,97]. Aún no se comprende del todo el rol funcional de los distintos homólogos estructurales de FlaB ni su organización específica dentro del filamento flagelar[96]. En especies como *T. denticola*[36] y *B. hyodysenteriae*[82,87,97], los mutantes de estas isoformas conservan endoflagelos con helicidad característica pero pierden parcialmente su motilidad traslacional, con algunas isoformas jugando roles desiguales en la determinación de la velocidad de natación[36,87,97]. Estos hallazgos sugieren que las distintas isoformas son parcialmente intercambiables, pero no absolutamente redundantes en función. Además, los distintos géneros de espiroquetas evidencian sus propios mecanismos reguladores, de forma tal que la eliminación de una determinada isoforma de FlaB puede ser compensada (*en T. denticola*) o no (*en B. hyodisenteriae*) por un aumento de las restantes. Dado que todos los homólogos de FlaB contribuyen a la motilidad pero ninguna es esencial, se postula que las distintas isoformas de FlaB no residen en dominios separados del PF sino que están distribuidas de forma dispersa a lo largo del filamento[36].

Leptospira expresa constitutivamente cuatro isoformas de flagelina que en *L. interrogans* se encuentran en el siguiente orden de abundancia: FlaB1(12.000 copias/célula), FlaB4(3.500), FlaB2(2.000) y FlaB3(300)[134]. Los mutantes flaB1^{-} en *L. interrogans* compensan la eliminación de esta isoforma mayoritaria sobreexpresando FlaB4, produciendo filamentos más rígidos y frágiles que pierden capacidad infectiva y motilidad en medios semisólidos. En *L. biflexa*, el fenotipo del mutante *flaB1*$^{-}$ es diferente; las isoformas FlaB2 y FlaB3 continúan expresándose y la espiroqueta mantiene su capacidad de progresar en medios viscosos. Estas diferencias entre cepas pueden explicarse por nuestros resultados, ya que tanto en el modelo final de *L. biflexa wt*, como en nuestro análisis de EM, FlaB4 (y no FlaB1) parece ser la isoforma mayoritaria en *L. biflexa*. Además, y en contraste con lo propuesto por Sasaki et al.[76] en cuanto a que en *L. biflexa wt* solo se expresan las isoformas FlaB1 y FlaB2; nuestros análisis de espectrometría de masas y entrecruzamiento químico permiten detectar y distinguir a las cuatro isoformas de FlaB en los filamentos purificados. Si bien FlaB3 no establece entrecruzamientos químicos con las demás, creemos que esto puede deberse a su abundancia relativa sustancialmente más baja en comparación con las demás isoformas.

Para explorar el rol diferencial de las cuatro isoformas de FlaB, caracterizamos sus diferencias estructurales mediante modelado in silico. Nuestros modelos homo-poliméricos resultaron confiables e identificaron una región con mayor variabilidad de secuencia y estructura que da lugar a diferencias sutiles en tamaño, hidrofobicidad y flexibilidad entre las distintas isoformas. Estas regiones se mapean a un motivo desestructurado formado por las secuencias a cada lado de la tercer alfa hélice de FlaB, que se localiza en la superficie del núcleo del filamento en los puntos de interacción con las proteínas de vaina. La resolución alcanzada en los mapas de cryoEM nos permitió validar estas predicciones de manera independiente, confirmando que las principales diferencias en la estructura primaria de las cuatro isoformas de FlaB se encuentran expuestas en la superficie del ensamblaje, y de hecho se amplifican con la glicosilación de las isoformas de FlaB en esta región.

Las glicosilaciones son un recurso más común de lo que inicialmente se pensaba en bacterias[96,175]. En particular las O-glicosilaciones son muy poco frecuentes en procariotas pero se encuentran en las flagelinas de las bacterias con flagelos más complejos como *C. jejuni* y *H. pylori*[12,176]. En espiroquetas, se ha descrito la glicosilación de FlaA de *B. burgdorferi, B. hyodysenteriae y S. aurantia*, que también glicosila dos de sus tres isoformas de FlaB[61,96,175,177]. El único glicano identificado en espiroquetas es el ácido noulosónico monoacetilado (Non) con una configuración de ácido pseudoamínico que glicosila las serinas y treoninas del dominio D1 de las isoformas de FlaB de *T. denticola*[61]. Estos glicanos de estructura inusual son críticos para la motilidad de la espiroqueta[61].

Mientras que en otras bacterias se ha postulado que las glicosilaciones de superficie en el filamento permitirian el escape de los patógenos al reconocimiento del Sistema Inmune

Innato del húesped[42], en las espiroquetas las glicosilaciones quedan internas en la estructura del filamento, mediando las interacciones que permiten el ensamblaje de la vaina. Estas glicosilaciones nos permiten entender las bajas intensidades observadas en los péptidos entrecruzados entre las FlaB y las proteínas de vaina ya que se glicosilan justamente los sitios de entrecruzamiento putativos del agente NNP9. Más relevante aún, estas glicosilaciones, que en el modelo final aparecen como pequeñas densidades electrónicas no asignadas, nos permiten diferenciar entre las distintas isoformas mayoritarias de FlaB en base a la presencia o ausencia de ciertos residuos glicosilados. En particular, de las siete O-glicosilaciones de FlaB en nuestro modelo de *L. biflexa wt*, 5 están conservadas en todas las FlaBs, mientras que las otras dos se encuentran localizadas en los residuos de Thr137 y Thr182 que solo son Treoninas en FlaB4. El carbohidrato en Thr137 es clave en la interacción con FlaA2, mientras que Thr182 lo es en la interacción con B0SJC6.

A partir de los resultados obtenidos, nuestra hipótesis es que las isoformas de FlaB pueden rescatarse parcialmente entre sí, de modo que ninguna isoforma es estrictamente esencial para la formación de endoflagelos y para mantener un nivel basal de motilidad. Sin embargo, para lograr motilidad completa se requiere la presencia de todas las isoformas, y los hallazgos sugieren que, al menos en *Leptospira*, las isoformas de FlaB no son funcionalmente intercambiables, como se ha planteado en otras espiroquetas[36,87]. Cada isoforma podría modular la estructura de la vaina a su alrededor al reclutar preferencialmente determinadas proteínas. Como parte de una respuesta adaptativa al ambiente, las bacterias podrían ajustar las proporciones relativas de isoformas FlaB y, consecuentemente, el tipo de vaina que producen.

Nuestra hipótesis, aunque implica un mecanismo más complejo debido a la presencia de la vaina, tiene cierta analogía con lo que ocurre en los flagelos de otras especies con múltiples isoformas de flagelina. En estos casos, la inactivación de genes específicos de flagelina permite aún la formación de filamentos móviles, aunque con pérdida de propiedades o ventajas adaptativas en entornos específicos[48,178].

4.4 – Vaina del filamento

Las proteínas FlaA son características y están universalmente presentes en todas las espiroquetas. En la mayoría de estas bacterias se ha postulado que FlaA forma una vaina continua alrededor del núcleo flagelar[45]. Resolver la primera estructura de una FlaA fue uno de nuestros objetivos principales. Tras superar retos considerables de expresión y solubilidad mediante técnicas de clonado, expresión y purificación high-throughput, logramos obtener FlaA1 de *L. borgpetersenii* y FlaA2 de *L. biflexa* en forma soluble.

La estructura tridimensional de FlaA1 revela un plegamiento conservado de la superfamilia de unión a carbohidratos, con tres hojas beta, una de ellas antiparalela a las otras dos. El ion calcio resulta fundamental para su correcto plegamiento, ya que se coordina con dos de estas tres hojas beta manteniéndolas en posición. La reducción permanente de las cisteínas conservadas nos permitió evitar agregados, confirmando que estas cisteínas están dispuestas a una distancia capaz de formar puentes disulfuro.

En cuanto a su función, FlaA parece tener configuraciones estructurales variadas en diferentes espiroquetas. En *B. hyodysenteriae*[82,97] y *T. denticola*[36], los mutantes deficientes en FlaA presentan disminución en motilidad y alteraciones morfológicas en sus endoflagelos. A pesar de que sus endoflagelos conservan la capacidad de rotar y mantienen una forma helicoidal levógira, muestran una morfología alterada, con un diámetro y un paso de hélice notablemente menores en comparación con las cepas salvajes[3]. En *Borrelia*, FlaA no forma una vaina completa, sino que se localiza principalmente en la base proximal del filamento[3,45,98]. En *Leptospira*, la ausencia de proteínas FlaA no afecta el diámetro flagelar[17] porque los mutantes *flaA2*$^{-}$ reemplazan la vaina nativa por una envoltura simétrica compuesta de FcpA y FcpB. Al igual que ocurre en *B. hyodysenteriae* y *T. denticola,* estos mutantes *flaA2*$^{-}$, que también carecen de FlaA1, pierden su motilidad translacional y superenrrollamiento característicos. Aunque el mecanismo que media estos fenotipos parece ser al menos en parte diferencial en *Leptospira*, involucrando también a las proteínas Fcp; creemos que al ser una proteína conservada en espiroquetas debería mantener un mismo rol funcional en todo el filo.

Además de la función propuesta que involucra un aumento en la rigidez de la vaina mediado por FlaA[97], estudios recientes en *Salmonella* han identificado a FlhE, una proteína estructuralmente similar a FlaA, que previene la formación de flagelos periplásmicos[91]. Los mutantes *flhE*$^{-}$ ensamblan filamentos en el periplasma capaces de girar y curvar la morfología celular, pero la tensión sobre la pared celular compromete la viabilidad. La prevención del ensamblaje o la reducción de la longitud del filamento flagelar restaura el fenotipo salvaje[91]. El fenotipo de estos mutantes, y la existencia de otros con filamentos flagelares ectópicos en el periplasma [179] igualmente inviables, sugieren que la capacidad distintiva de las espiroquetas radica en su habilidad para enrollar los filamentos periplásmicos alrededor del protoplasma sin comprometer la integridad celular. Dada la similitud estructural entre FlhE y FlaA[91], es posible que sus pliegues permitan interacciones con un mismo componente del periplasma bacteriano.

Las proteínas FcpA y FcpB, exclusivas de la familia *Leptospiraceae*, son componentes principales de su vaina flagelar. Nuestro laboratorio ha cristalizado y resuelto las estructuras de FcpA de *L. biflexa* y FcpB de *L. interrogans*, revelando plegamientos únicos sin ortólogos en otras espiroquetas ni similitud estructural con otros componentes flagelares

conocidos[168]. Aunque FcpA es capaz de polimerizar, su alta solubilidad (~1 mM) sugiere que requiere una plataforma de ensamblaje proporcionada por el núcleo flagelar. Asimismo, FcpA parece actuar como plataforma para el ensamblaje secuencial y ordenado de FcpB.

Estas proteínas desempeñan un papel central en el enrollamiento del filamento, localizándose exclusivamente en el lado convexo y actuando como "expansores de curvatura"[168]. En los mutantes flaA2^{-}, FcpA y FcpB forman una vaina radialmente simétrica que produce filamentos sin curvatura, demostrando que su disposición asimétrica es esencial para su función. Aunque aún no comprendemos completamente el mecanismo subyacente, la curvatura inducida por las Fcps es crucial para la motilidad de *Leptospira*, ya que filamentos con menor curvatura no pueden producir el empuje necesario para que las bacterias se trasladen, deformen sus extremos celulares o colonizen a sus huéspedes.

El análisis de similitud estructural y conservación de secuencias (Figura 44) permitió agrupar las proteínas de vaina en familias: i) FlaA1 y FlaA2, proteínas de unión a carbohidratos dependientes de calcio; ii) la familia de FcpB, que incluye B0SPP1 y B0SKN8; y iii) B0SMT6, B0SJC6 y B0SMK8, grandes proteínas con al menos 6 α-hélices. Componentes como B0STF2, FcpA y B0STN7 tienen estructuras únicas sin homólogos en el filamento.

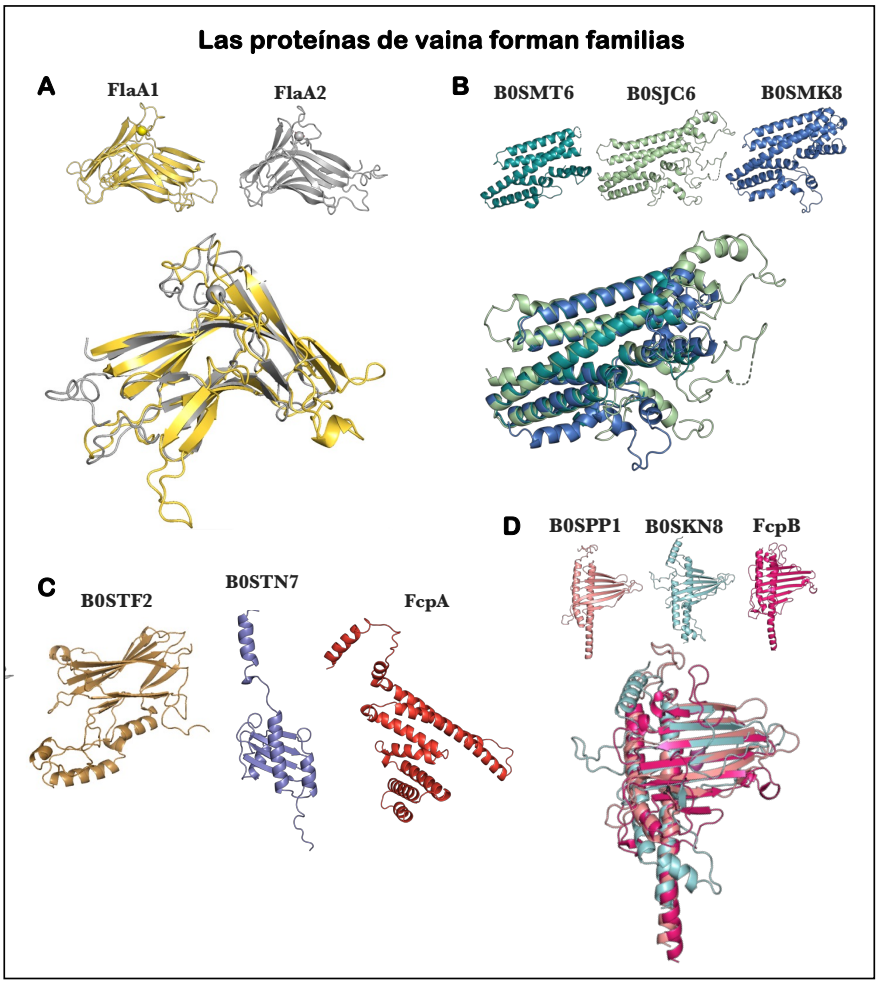


Figura 44: Las proteínas de vaina se agrupan en familias. Se muestran los alineamientos estructurales de todos los carbonos α de las proteínas de vaina de *L. biflexa* y su agrupación en familias. **A.** La familia de las FlaAs (FlaA1 y FlaA2) comparten un ~30% de identidad de secuencia y sus estructuras alineadas tienen un rmsd de ~1.5Å. **B.** Las proteínas de vaina B0SMT6, B0SJC6 y B0SMK8 conforman otra familia de proteínas de vaina. Tomando como referencia a B0SMK8, B0SMT6 tiene un rmsd de ~1.2Å y ~22% de identidad, mientras que B0SJC6 tiene ~1.5Å y ~17.7%. **C.** Las proteínas FcpA y B0STN7, al igual que B0STF2, son cada una singular, no conformando familias con otros componentes. **D.** La tercer familia está conformada por las proteínas FcpB, B0SPP1 y B0SKN8, con rmsd de ~2.3Å y ~7% de identidad para B0SPP1, y rmsd de ~2.2Å y ~14.5 % de identidad para B0SKN8, tomando como referencia a FcpB.

4.5 – Mecanismos de motilidad y origen de la asimetría

La vaina asimétrica de *Leptospira* plantea preguntas fundamentales sobre cómo se genera esta asimetría y cuál es su función en la motilidad. La motilidad de las espiroquetas, y en particular el mecanismo mediante el cual el torque generado por el motor se transmite al filamento flagelar para inducir su movimiento ondulatorio y deformar el cuerpo celular, es un misterio de larga data. Un modelo matemático reciente en *B. burgdorferi*[45,105] sugiere que el filamento y su vaina rotan libremente alrededor del periplasma, sin contacto con el cuerpo celular y el peptidoglicano. Sin embargo, considerando el espacio confinado y la capacidad de los filamentos para deformar la envoltura celular, parece probable alguna interacción entre filamentos y estructuras celulares.

En el caso de *Leptospira*, la asimetría de la vaina introduce una complejidad adicional: la hipótesis de rotación libre se torna menos viable, ya que las Flagellar Coiling Proteins (Fcps), necesarias para mantener la curvatura del filamento, deberían estar constantemente posicionadas en el lado convexo para inducir la curvatura hacia el cuerpo celular y no hacia el exterior. Una hipótesis alternativa sería que la vaina y el núcleo flagelar no roten de manera conjunta, sino que la vaina permanezca en una posición fija en contacto con el cuerpo celular mientras el núcleo rota. Algunas imágenes de microscopía electrónica sugieren interacción entre el flagelo y el cuerpo celular[74], y Wunder et al. han planteado que FcpA podría ser la interfaz entre el endoflagelo y el peptidoglicano (PG), cuya pérdida en los mutantes *fcpA*$^{-}$ podría impedir el enganche correcto y comprometer el movimiento rotacional del extremo celular[105].

Nuestros resultados permiten proponer un modelo para el ensamblaje secuencial del endoflagelo y el origen de su asimetría. El proceso comenzaría con la secreción flagelar de las isoformas de FlaB, que son los únicos componentes sin una secuencia señal de exportación en el N-terminal[87,97]. Este proceso sería conservado y similar al autoensamblaje de la flagelina en otras bacterias. Sin embargo, en *Leptospira*, es probable que existan mecanismos reguladores adicionales que controlen los niveles de las cuatro isoformas de FlaB secretadas, dejando segmentos con una mayor afinidad por ciertas proteínas de vaina. Aunque distintos homólogos y parálogos de FlaB presentan superficies diferenciales de unión a otras proteínas, consideramos que este mecanismo no es suficiente, por sí mismo, para generar la notable

asimetría observada en la vaina. Postulamos entonces, la existencia de un mecanismo adicional que aprovecha el deslizamiento de los protofilamentos entre sí que se ha descrito en otras bacterias introduce perturbaciones no helicoidales en el filamento[50,180], para encontrar una asimetría temporal en el núcleo flagelar y propagarla, volviéndola permanente y más marcada[169]. Nuestro modelo, publicado en Brady et al.[169], propone que FlaA2 podría ser el candidato principal que se une en la primeras articulaciones asimétricas del filamento flagelar. Esta unión inicial de FlaA2 bloquearía el acceso de otras proteínas de vaina (como FcpA y FcpB) a la cara cóncava del filamento y proporcionaria una superficie de unión diferencial y asimétrica para el resto de las proteínas de vaina. Es posible que su ensamblaje reclute ciertos interactores como B0STN7 y FcpA que se encuentran adyacentes FlaA2 y FlaAP en la vaina flagelar. Aunque al momento no podemos concluir quién se ensambla primero, lo cierto es que en esta etapa se secreta y ensambla FcpA antes que FcpB y probablemente que B0SKN8 y B0SPP1, que también tienen sus plataformas de ensamblaje sobre FcpA.

Finalmente, FlaA1 sería la última proteína en reclutarse para formar la vaina del filamento. Proponemos que la vaina incompleta con un surco que se observa en los filamentos *L. biflexa*, representa de hecho el último estado intermedio en el ensamblaje de la vaina. Si bien esto en principio podría reflejar una diferencia inherente a las especies saprófitas *vs* patógenas, el hecho es que sí encontramos FlaA1 en el proteoma flagelar de *L. biflexa*, que además se entrecruza con FlaA2 e incluso dimeriza en un homo-inter-entrecruzamiento de FlaA1. Estas observaciones concuerdan con el modelo de *L. interrogans* que postula dos protofilamentos de FlaA1 antiparalelos y flanqueados a ambos lados por FlaA2. Además, nuestro resultados de espectrometría de masas indican que las cantidades relativas de FlaA1 varían en las distintas preparaciones flagelares, posiblemente debido a su expresión diferencial bajo distintas condiciones, a su interacción con un componente periplásmico de alta afinidad y/o a una asociación más laxa con el filamento. Nuestro modelo del filamento presenta un segundo canal hueco en el centro, lo cual podría explicar la unión inestable de FlaA1 al mismo. Aunque no descartamos la ausencia de un último componente flagelar en esta posición, FlaA1 parece asociarse a FlaA2 sin interactuar directamente con el núcleo flagelar de FlaB.

4.6 – Comparación con otros filamentos de espiroquetas

Nuestro modelo cuestiona lo conocido sobre filamentos flagelares de otras espiroquetas. ¿Realmente poseen una vaina completamente simétrica compuesta únicamente por FlaA? En los filamentos de *B. hyodysenteriae* copurifican de manera reproducible proteínas con masas moleculares de 39, 35 y 29 kDa que no reaccionan con antisueros FlaA o FlaB[82]. Hallazgos similares se han descrito para *T. pallidum, T. phagedenis y T. denticola*[80,100], y en esta última especie se ha identificado FlaG, una proteína minoritaria asociada al filamento[146]. Sin embargo, a diferencia de *Leptospira*, estas espiroquetas presentan un proteoma aparentemente más simple y muestran reducción evidente en diámetro al eliminar

FlaA[36,71,87]. En *B. hyodysenteriae*, la crioET muestra una estructura de vaina simétrica y reticular[71,87], más sencilla que la de *Leptospira*.

Otra gran interrogante es ¿Por qué *Leptospira* necesita filamentos más complejos? Encontramos tres diferencias estructurales que lo distinguen: i) tiene el menor paso de hélice en su cuerpo celular, requiriendo filamentos más curvos; ii) logra con solo dos flagelos cortos no solapantes la misma velocidad que otras espiroquetas con hasta 11 filamentos por polo; y iii) presenta la mayor rigidez relativa de endoflagelos frente al cuerpo celular[93]. Estos factores podrían explicar su estrategia alternativa con filamentos estructuralmente más complejos que generan mayor torque[3], permitiéndole alcanzar las mayores velocidades de traslación observadas en espiroquetas tanto en medios líquidos[3] como en geles viscosos.

La morfología de las espiroquetas parece estar determinada por dos factores. El primero la capa de peptidoglicano y las proteínas del citoesqueleto celular, que juegan un rol determinante en la forma, haciendo que algunas espiroquetas como *B. burgdorferi* tengan forma naturalmente cilíndrica y otras como *Leptospira* tengan forma helicoidal. El segundo los flagelos periplásmicos, cuya función se superpone a la del citoesqueleto, terminando de darle forma a la célula[100]. Esta morfología dependiente de endoflagelos parece ser consecuencia de la interacción mecánica entre estos dos cuerpos elásticos de diferente rigidez[181]. En *Leptospira*, sus endoflagelos cortos solo determinan la forma espiralada o de gancho de los extremos celulares[70].

V– CONCLUSIÓN

Hace más de un siglo, Dobell (1912) destacó que "el movimiento de las espiroquetas está rodeado de misterio". Un siglo después, aunque se han logrado avances, muchas preguntas intrigantes sobre la motilidad de las espiroquetas siguen sin respuesta[45]. Con sus modos de motilidad tan distintos a los de flagelos bacterianos canónicos, las espiroquetas constituyen un sistema biológico notable para estudiar la interrelación de elasticidad, morfología y autopropulsión[86] dependientes de sus endoflagelos que rotan en el periplasma.

Utilizando enfoques estructurales integrativos, dilucidamos la estructura del filamento periplásmico de *Leptospira*, generando un modelo molecular exacto que revela una arquitectura increíblemente compleja con 15 componentes diferentes. Este filamento, con un núcleo de 11 protofilamentos de FlaB rodeado por una vaina asimétrica, queda validado por la calidad de sus parámetros estructurales, las distancias esperadas entre sus interactores y su capacidad para explicar diversos fenotipos experimentales.

Nuestro trabajo ha permitido responder varias de las preguntas fundamentales planteadas inicialmente. Si bien el filamento mantiene características importantes y conservadas con los filamentos de las bacterias modelo y otras espiroquetas, a diferencia de cualquier otro filamento jamás descrito presenta una estructura altamente asimétrica que plantea un cambio en el paradigma aceptado en la motilidad bacteriana. El reclutamiento temprano de las proteínas FlaA2 en el lado cóncavo del filamento generaría esta asimétria, dejando a las proteínas FcpA y FcpB del lado convexo del filamento para generar la curvatura necesaria para el superenrrollamiento.

Respecto a la contribución de distintas proteínas flagelares a la motilidad y virulencia, nuestros estudios de mutantes revelan que, mientras las isoformas de FlaB pueden compensarse parcialmente entre sí, la eliminación de proteínas de vaina como FcpA, FcpB o FlaA2 provoca pérdidas significativas de motilidad traslacional y capacidad infectiva. Esto demuestra que cada componente cumple funciones específicas no completamente redundantes.

Finalmente, hemos propuesto que la complejidad excepcional del filamento de *Leptospira*, comparada con otras espiroquetas, responde a tres características únicas del género: su menor paso de hélice celular que requiere filamentos más curvos; la necesidad de lograr alta velocidad con solo dos flagelos cortos; y la mayor rigidez relativa de sus endoflagelos frente al cuerpo celular. Esta configuración estructural única permite a *Leptospira* generar el mayor torque entre las espiroquetas y alcanzar velocidades de traslación superiores tanto en medios líquidos como viscosos.

Este proyecto abre las puertas a explorar en detalle la estructura de los glicanos flagelares y el rol de las diferentes proteínas de vaina, avanzando hacia una comprensión integral de la biología de *Leptospira* como modelo de espiroquetas. El conocimiento adquirido sobre estos aspectos fundamentales podría aplicarse eventualmente al tratamiento y prevención de enfermedades causadas por espiroquetas patógenas[3], e incluso inspirar el diseño de microrrobots de propulsión autónoma con alta eficiencia[12].

VI – MATERIALES Y METODOS

6.1 – Análisis filogenético de proteínas flagelares de Espiroquetas

Las secuencias genómicas y aminoacídicas fueron descargadas de las bases de datos NCBI GenBank y RefSeq[182] y Uniprot[183], salvo cuando se indica lo contrario. Para el análisis genómico comparativo completo de los sistemas endoflagelares en Espiroquetas, se descargaron de la base de datos PATRIC[184] todos los genomas disponibles pertenecientes al Phylum Espiroquetas (n = 971). Se removieron aquellos genomas no asignados a especies conocidas o correspondientes a cepas no cultivadas, resultando en 864 genomas de alta calidad. Los 864 genomas fueron reanotados automáticamente utilizando Prokka[185] para asegurar consistencia en la anotación génica. Para la visualización de los genomas y la distribución de los genes flagelares de *Leptospira* se empleó el software Genoscope[186].

Se realizaron búsquedas de las 59 proteínas flagelares conocidas en los genomas procesados, manteniendo por conveniencia una única isoforma de FlaB para el análisis comparativo. Se calcularon modelos ocultos de Markov (HMMs, por Hidden Markov Models) para cada proteína flagelar utilizando hmmbuild v3.3.2[187]. Los hits positivos se reportaron utilizando un umbral de e-value menor a $1x10^{-10}$ para asegurar la especificidad de las identificaciones. Los *scores* resultantes fueron escalados de 0 a 1 mediante normalización por el *score* más alto obtenido para cada proteína flagelar específica. Esta normalización permitió la comparación directa entre diferentes familias de proteínas con rangos de similitud variables.

Para el análisis comparativo y alineamiento múltiple de secuencias, tanto genómicas como proteicas, se utilizó el programa Jalview[152], calculando los alineamientos mediante el algoritmo ClustalW[151]. Los alineamientos resultantes del pipeline automático fueron visualizados y curados manualmente en Jalview, proceso que incluyó la eliminación de secuencias repetidas, corrección de genomas mal anotados, remoción de regiones con cobertura inferior al 30% de las secuencias, y verificación de la correcta asignación de dominios funcionales. El proceso de curación manual también involucró la corrección de errores de alineamiento en regiones altamente conservadas mediante inspección de perfiles de conservación y la identificación de dominios funcionales utilizando análisis con bases de datos especializadas.

Para contextualizar evolutivamente los resultados, se construyó un árbol filogenético de las 864 Espiroquetas con el programa gtdb-tk v3.0[188], utilizando un genoma de *Escherichia coli* como grupo externo. El árbol se basó en un alineamiento concatenado de 120 genes marcadores universalmente presentes en bacterias, proporcionando una base filogenética robusta para el análisis comparativo. El árbol resultante fue colapsado a nivel de especies, resultando en 115 especies representativas de los géneros *Leptospira* (n = 63), *Leptonema* (n

= 1), *Turneriella* (n = 1), *Borrelia* (n = 21), *Treponema* (n = 15), *Sphaerochaeta* (n = 1), *Spirochaeta* (n = 1), *Brevinema* (n = 1) y *Brachyspira* (n = 10), más el genoma externo de *E. coli*.

Se realizaron análisis de conservación de secuencias para identificar posiciones altamente conservadas en cada familia de proteínas. La significancia estadística de la conservación se evaluó mediante métodos estadísticos apropiados, y la clasificación funcional de proteínas se realizó mediante análisis de dominios con bases de datos especializadas. La revisión y corrección de anotaciones génicas se llevó a cabo mediante análisis filogenético de cada familia de proteínas, construcción de árboles de genes específicos y comparación con anotaciones de genomas de referencia curados manualmente. Los resultados se visualizaron utilizando ggtree[188] para la representación del árbol filogenético y el mapa de calor correspondiente. El mapa de calor se generó integrando los datos de conservación de proteínas flagelares con la información filogenética, permitiendo la identificación de patrones evolutivos y la distribución de componentes específicos entre los diferentes géneros de Espiroquetas.

Todos los archivos, alineamientos de secuencia y scripts utilizados para realizar los análisis están disponibles públicamente en el repositorio FigShare (https://doi.org/10.6084/m9.figshare.19686282), permitiendo la reproducción completa de los resultados presentados en este estudio.

6.2 – Análisis de características proteícas y predicción estructural

Para analizar las características fisicoquímicas de las proteínas identificadas, incluyendo punto isoeléctrico, masas moleculares esperadas, coeficiente de extinción y contenido en cisteínas, se utilizó el programa ProtParam[189]. La predicción de péptidos señal de secreción al periplasma se realizó mediante SignalP[190]. Para la predicción de estructuras secundarias de las proteínas, particularmente en el caso de las isoformas de FlaA, se empleó el programa PSIPRED[153]. La búsqueda de homólogos estructurales se realizó utilizando los servidores Phyre2 (Protein Homology/analogy Recognition Engine V2.0)[154] y DALI[191]. La predicción de estructuras 3D de proteínas no anotadas fue realizada con AlphaFold2[78] con parámetros por defecto en búsqueda de homólogos y generación de alineamientos múlitples de secuencia. Se consideraron fiables aquellas predicciones (o regiones) con índice pLDDT ≥ 0.75.

6.3 – Clonados y construcción de vectores

Todas las construcciones de esta tesis se realizaron mediante la técnica de Restriction-Free Cloning[192], que no utiliza enzimas de restricción y permite insertar la secuencia blanco en cualquier posición deseada del vector. Los oligonucleótidos específicos fueron diseñados

como cebadores considerando parámetros básicos estándar: contenido GC aproximado de 50%, temperatura de fusión entre 55°C y 65°C, extremos 3' terminando en G o C. Los cebadores se diseñaron con la herramienta en línea RF-cloning [192], y se verificaron con Primer3[193]. La lista completa con los cebadores comprados en IDT (Integrated DNA Technologies), así como los vectores parentales y los obtenidos por nosotros, pueden verse en la Tabla 3.

Construcción	Plásmido parental	Resistencia	Diseño de la construcción	Primer Fw	Primer Rev	Tamaño del megaprime	Tamaño del plásmido
pQE80-bFcpA		ampicilina					5577 pb
pQE80-bFcpB		ampicilina					5559 pb
pQE80-iFcpA		ampicilina					5625 pb
pQE80-iFcpB		ampicilina					5550 pb
pQE80-iFcpA cysless	pQE80_. iFcpA	ampicilina		5'_CGATTTACAAAATCTTCCTCAGACCATCCCCATTCTCGAGCTC_3´	5´_GATCATGTTTTCTTCGAAAGCGTAGGCACTTGCTAGATACTTATGAG_3´	127 pb	5625 pb
pQE80-iFlaA1							
pETM14-FlaA1	pETM14	kanamicina	6His_preS_FlaA1	5'_CTGGAAGTTCTGTTCCAGGGGCCCGGTTCACAAAACATCATCAAAGGC_3'	5'_TGAGATCCGGCTGCTAACAAAGCCGCAGGTCGACGCCAAGATCC_3'	939 pb	6157 pb
pETM22-FlaA1	pETM22	kanamicina	trxA_6His_preS_FlaA1			939 pb	6500 pb
pETM33-FlaA1	pETM33	kanamicina	6His_GST_preS_FlaA1			939 pb	6826 pb
pETM44-FlaA1	pETM44	kanamicina	6His_MBP_preS_FlaA1			939 pb	7261 pb
pETM55-FlaA1	pETM55	kanamicina	$DsbA_{cit}$_6His_preS_FlaA1			939 pb	6745 pb
pETM66-FlaA1	pETM66	kanamicina	NusA_6His_preS_FlaA1			939 pb	7648 pb
pQE80-iFlaA2		ampicilina					
pETM14-FlaA2	pETM14	kanamicina	6His_preS_FlaA2	5'_CTGGAAGTTCTGTTCCAGGGGCCCTCAGGTCAGCAGACCGGAAATAATCAG_3'	5'_TGAGATCCGGCTGCTAACAAAGCCGCAGGTCGACGCCAAGATCC_3'	732 pb	5950 pb
pETM22-Fla2	pETM22	kanamicina	trxA_6His_preS_FlaA2			732 pb	6293 pb
pETM33-FlaA2	pETM33	kanamicina	6His_GST_preS_FlaA2			732 pb	6619 pb
pETM44-FlaA2	pETM44	kanamicina	6His_MBP_preS_FlaA2			732 pb	7054 pb
pETM55-FlaA2	pETM55	kanamicina	$DsbA_{cit}$_6His_preS_FlaA2			732 pb	5950 pb
pETM66-FlaA2	pETM66	kanamicina	NusA_6His_preS_FlaA2			732 pb	7441 pb
pETM54-FlaA1	pETM54	kanamicina	DsbA_His_preSite_FlaA1	5'_CTGGAAGTTCTGTTCCAGGGGCCCGGTTCACAAAACATCATCAAAGGC_3'	5'_TGAGATCCGGCTGCTAACAAAGCCGCAGGTCGACGCCAAGATCC_3'	939 pb	6797 pb
pET28a-FlaA1_His	pET28a	kanamicina	OmpAleader_FlaA1_His	5'_CGCTACCGTGGCCCAGGCGCATCACCATCACCATCACGGATCAGGATCGG_3'	5'_GCCATTCGATTTTCTGAGCCTCGTTTGCCTGCAGCACCGCCGCCGCCAC_3'	900 pb	6182 pb
pET28a_His_FlaA1	pET28a	kanamicina	OmpAleader_His_TEV_FlaA1	5'_CGCTACCGTGGCCCAGGCGCATCACCATCACCATCACGGATCAGGATCGG_3'	5'_ACAGGAGAGTCTCAGAGATTATTTGCCTGCAGCACCGCCGCCGCCAC_3'	958 pb	6586 pb
pET22b-FlaA1	pET22b	ampicilina	pelB_His_TEV_FlaA1	5'_CTCGAGCACCACCACCACCACCACCATCACCATCACCATCACGGATCAGG_3'	5'_CAAGACCCGTTTAGAGGCCCCTTATTTGCCTGCAGCACCGCCGCCGCC_3'	966 pb	6333 pb
pETM54-FlaA2	pETM54	kanamicina	DsbA_His_preSite_FlaA2	5'_CTGGAAGTTCTGTTCCAGGGGCCCTCAGGTCAGCAGACCGGAAATAATCAG_3'	5'_TGAGATCCGGCTGCTAACAAAGCCGCAGGTCGACGCCAAGATCC_3'	732 pb	6590 pb
pET28a-FlaA2_His	pET28a	kanamicina	OmpAleader_FlaA2_His	5'_CGCTACCGTGGCCCAGGCGCAGCAGACCGGAAATAATCAGGCTGGTGG_3'	5'_GCCGGATCTCAGTGGTGGTGGTGGTGGTGCCAATTATCTTTAATTTCCGATCCAGG_3'	699 pb	5941 pb
pET22b-FlaA2	pET22b	ampicilina	pelB_His_TEV_FlaA2	5`_CTCGAGCACCACCACCACCACCACCATCACGGATCAGGATCGG_3'	5'_CAAGACCCGTTTAGAGGCCCCTTACCAATTATCTTTAATTTCCGATCCAGGATACTTCGC_3'	750 pb	6117 pb
pCola_bFcpA	pCola	kanamicina	His_bFcpA	5´_GCAGCAGCCATCACCATCATCACCACAGCGGATCAGGATCGGAAAACCTGTATTTTCAGG_3´	5'_CGGCCGTGTACAATACGATTACTTTCTGTTCGACTTAAGCATTATTGTGCTACCGAAATAACCTCATCTCGG_3'	910 pb	4535 pb
pCola_bFlaA2	pCola	kanamicina	FlaA2_strep	5'_ACATATGGCAGATCTCAATTGGATATCGGCCGGCGATACTGGCGGACAAAATACAGCTAACAC_3'	5'_CGTTCAAATTTCGCAGCAGCGGTTTCTTTACCAGACCAATTGTCCTTAATTTCAGAACCAGGGT_3'	711 pb	4337 pb
pCola_bFcpA_bFlaA2	pCola_FlaA2	kanamicina	His_bFcpA y FlaA2_strep				5177 pb
pETDuet_bFlaA1	pETDuet	ampicilina	His_FlaA1	5'_CATGGGCAGCAGCCATCACCATCATCACCACGGTTCAGGTTCACAATACAAAGTCTCCAACGGC_3'	5'_TCGATTATGCGGCCGTGTACAATACGATTACTTTCTGTTCGACCTATTTGTTTTGGGAAGATCCGTTGCTTC_3'	998 pb	6240 pb
pETDuet_bFcpB	pETDuet	ampicilina	bFcpB_strep	5'_GAAGGAGATATACATATGGCAGATCTCAATTGGATATCGACTGGTAGCTCGGAAACTGGC_3'	5'_CGCAGCAGCGGTTTCTTTACCAGGTATTTCAAAGAACCTTTCAATGCTTCTACGTTC_3'	836 pb	6090 pb
pETDuet_bFlaA1_bFcpB	pETDuet_FlaA1	ampicilina	His_FlaA1 y bFcpB_strep				7014 pb

Tabla 3: Construcciones diseñadas para el clonado de vectores. Se muestran todas las construcciones generadas, incluyendo cebadores utilizados, tamaño esperado del megaprimer, vectores parentales y generados, así como las proteínas producidas a partir de sus secuencias codificantes. Las construcciones se organizan por colores: rojo para Fcp, azul para FlaA1, verde para FlaA2 y amarillo para coexpresión de proteínas de *L. biflexa*.

El proceso involucra dos PCR consecutivas (Tabla 4). En la primera PCR se amplificaron las secuencias codificantes blanco provenientes de ADN genómico de *Leptospira* o de otros vectores previamente clonados. Para verificar si la PCR funcionó y recuperar los "megaprimers", se corrieron geles de agarosa 1% en solución TAE (Tris 40 mM, ácido acético 1%, EDTA 2 mM), teñidos con bromuro de etidio (0.5 µg/mL) y visualizados en transiluminador ultravioleta. Las bandas del tamaño adecuado fueron cortadas del gel y purificadas por el kit QIAquick Gel Extraction (Qiagen) según protocolo del fabricante.

A

PCR 1 mix		PCR 2 mix	
Reactivo	**Concentración /volumen final**	**Reactivo**	**Tiempo**
5X Phusion buffer HF	1X – 10uL	5X Phusion buffer HF	1X – 5uL
10 mM dNTPs	200uM – 1uL	10 mM dNTPs	200uM – 0.5uL
10uM primer For	1,25 uL	Megaprimer	50-100 ng
10uM primer Rev	1,25 uL		
ADN molde Lepto	50-250 ng	Vector	0.5 ng
ADN Polimerasa Phusion 20U/uL	10U- 0.5uL	ADN Polimerasa Phusion 20U/uL	5U- 0.25uL
H_2O	csp 50uL	4ºC	csp 25uL

B

Termociclado PCR 1		Termociclado PCR 2	
Temperatura	**Tiempo**	**Temperatura**	**Tiempo**
98ºC	1 min	98ºC	1 min
98ºC	10 seg	98ºC	10 seg
Tm primer más bajo + 3ºC	10-30 seg	60ºC-65ºC	40-60 seg
72ºC	15-30 seg	72ºC	2.5 - 3 min
72ºC	7 min	72ºC	7 min
4ºC	10 min	4ºC	10 min
30 ciclos		30 - 35 ciclos	

Tabla 4: Reactivos y condiciones de termociclado para RF-cloning. A) Reactivos utilizados en cada PCR. Los volúmenes de ADN molde de *Leptospira* (PCR1) y de megaprimer y vector (PCR2) son variables según concentraciones determinadas. La polimerasa Phusion utilizada en PCR2 fue High Fidelity comercial (Biolabs), mientras que para PCR1 se utilizaron preparaciones propias cuando estaban disponibles. Se utilizó Buffer HF en todos los casos por su menor propensión a errores. **B)** Condiciones de termociclado para ambas PCR. Los ciclos de amplificación se marcan en gris, con pasos sucesivos de desnaturalización, annealing (temperatura 3°C mayor a la Tm del primer más bajo) y extensión (72°C). Las temperaturas de hibridación y tiempos difieren entre PCR1 y PCR2 considerando que la extensión de ADN a amplificar es mayor en PCR2.

Los megaprimers purificados contienen la secuencia codificante de interés flanqueada por cebadores diseñados para hibridar en los sitios deseados para inserción en el vector. En la segunda PCR, estos plásmidos se extienden para obtener los vectores deseados. El producto de esta segunda amplificación fue tratado con la enzima DpnI (20U, incubando 1 hora a 37°C e inactivando posteriormente a 80°C), que corta únicamente el ADN del plásmido parental metilado de *E. coli*.

Se transformaron células quimiocompetentes de *E. coli* DH5α o TOP10F' y se plaquearon 100 µL de reacción. Para confirmar los vectores previo a secuenciación, en la mayoría de los casos se llevó a cabo confirmación mediante digestión con enzimas de restricción o realizando el mismo protocolo que para la PCR1, confirmando que las colonias transformadas contenían el inserto de interés por visualización del megaprimer.

6.4 – Clonado *high-throughput* de isoformas de FlaA

Para clonar las isoformas de FlaA1 y FlaA2 de *L. biflexa*, *L. borgpetersenii* y *L. interrogans* en el vector suite, se empleó un método que permitiera utilizar un único par de cebadores para introducir paralelamente cada secuencia codificante en múltiples vectores[194]. Para cada especie de interés se diseñaron tres cebadores forward y un cebador reverso (Tabla 5). Los primers forward fueron diseñados para unirse en 5' al sitio de corte con TEV, hibridando con todos los vectores de la batería, mientras que en 3' hibridan con el genoma de *Leptospira* en las secuencias que codifican para las tres primeras estructuras secundarias predichas como hélices α u hojas β. De esta forma se obtienen tres mutantes N-terminal truncados de cada isoforma de FlaA. Los cebadores reversos fueron diseñados para que su reverso complementario hibride con las secuencias que codifican para las últimas estructuras secundarias codificadas por los genes de *Leptospira* y con la etiqueta *strep-tag* y el codón *stop* de los vectores. Se seleccionaron las baterías de vectores con promotores T7 y la proteína GFP clonada en marco de lectura con las proteínas de fusión.

Construcción del plásmido parental	Tamaño del plásmido	Promotor	Resistencia
His_TEV_DsbC_ GFP_strep	5419 pb	T7	ampicilina
His_TEV_MBP_ GFP_strep	5877 pb	T7	ampicilina
His_TEV_CelD_ GFP_strep	6403 pb	T7	ampicilina
His_TEV_Trx_ GFP_strep	5798 pb	T7	ampicilina
His_TEV_ GFP_strep	4759 pb	T7	ampicilina
His_TEV_sumo_ GFP_strep	5074 pb	T7	ampicilina
His_TEV_pelB_ GFP_strep	5519 pb	T7	ampicilina
His_TEV_GST_ GFP_strep	6160 pb	T7	ampicilina
His_TEV_OsmY_ GFP_strep	6109 pb	T7	ampicilina

Nombre del inserto	Proteína que se traduce	Primer Fw	Tm (ºC)	Primer Rev	Tm (ºC)	Tamaño megaprim
boFlaA1ΔN1	FlaA1 Δ1-49 _Δ303-316 *L. borgpetersenii*	5'_GGTAGCGGATCGGAAAACCTGTATTTTCAGGGATCCGGTTCAATGGAACTTAGATCCATTACTGTGGAATC_3'	61	5'_GGTGGCTCCAGCTGCCGGATCCTTTCGTCTTAAGCATTTTATCTAACTTATGTTTGCTTTTGC_3' Para la versión strepless: 5'_GGCTGCAGGTCGACGCCAAGATCCTTATTTCGTCTTAAGCATTTTATCTAACTTATGTTTGCTTTTGC_3'	61	823 pb
boFlaA1ΔN2	FlaA1 Δ1-71_Δ303-316 *L. borgpetersenii*	5'_GGTAGCGGATCGGAAAACCTGTATTTTCAGGGATCCGGTTCAGAAGTTTTTACCGATAAGGATACTCAACAAGC_3'	61			757 pb
boFlaA1ΔN3	FlaA1 Δ1-100_Δ303-316 *L. borgpetersenii*	5'_GGTAGCGGATCGGAAAACCTGTATTTTCAGGGATCCGGTTCACGTGAAGTGAAACTGATCGCC_3'	61			670 pb
iFlaA1ΔN1	FlaA1 Δ1-49_Δ273-317 *L. interrogans*	5'_GGATCGGAAAACCTGTATTTTCAGGGATCCGGTTCACTGGAACTCAGATCTATCACGGTGGAATC_3'	61	5'_GGGTGGCTCCAGCTGCCGGATCCTACATCGGACAATACTCTCAATTCGTCAAAGA_3'	61	718 pb
iFlaA1ΔN2	FlaA1 Δ1-69_Δ273-317 *L. interrogans*	5'_GGATCGGAAAACCTGTATTTTCAGGGATCCGGTTCAGGTTGGGAAGTTACTACCGATAAAGATACACAAC_3'	62			668 pb
iFlaA1ΔN3	FlaA1 Δ1-89_Δ273-317 *L. interrogans*	5'_GGATCGGAAAACCTGTATTTTCAGGGATCCGGTTCATATCAAGCGACTGCCCAAAACGCT_3'	62			608 pb
bFlaA1ΔN1	FlaA1 Δ1-45_Δ307-328 *L. biflexa*	5'_GGATCGGAAAACCTGTATTTTCAGGGATCCGGTTCACTCAAAGCCATTACCATTGAATCTTGGGACA_3'	62	5'_GGGTGGCTCCAGCTGCCGGATCCCCAGTTATCCTTAATTTCTGATCCAGGATACTTC_3'	61	842 pb
bFlaA1ΔN2	FlaA1 Δ1-63_Δ307-328 *L. biflexa*	5'_GGATCGGAAAACCTGTATTTTCAGGGATCCGGTTCAGGTTGGGAAGTGTTCACTGACAAAGA_3'	61			788 pb
bFlaA1ΔN3	FlaA1 Δ1-101_Δ307-328 *L. biflexa*	5'_GGATCGGAAAACCTGTATTTTCAGGGATCCGGTTCACTTGTGAGTAGAGAAGTGAAACTGGTTCCAG_3'	62			675 pb
boFlaA2ΔN1	FlaA2 _Δ1-43 *L. borgpetersenii*	5'_GGATCGGAAAACCTGTATTTTCAGGGATCCGGTTCAGAAAAACTCATTCTCGAGAATTTCGAGGAGG_3'	61	5'_GGGTGGCTCCAGCTGCCGGATCCCCAGTTATCCTTAATTTCTGATCCAGGATACTTCG_3'	61	650 pb
boFlaA2ΔN2	FlaA2 _Δ1-67 *L. borgpetersenii*	5'_GGATCGGAAAACCTGTATTTTCAGGGATCCGGTTCAACCAAAGTTCTTAAGATGGTTCAAAGAGGTTTG_3'	61			578 pb
boFlaA2ΔN3	FlaA2 _Δ1-99 *L. borgpetersenii*	5'_GGATCGGAAAACCTGTATTTTCAGGGATCCGGTTCACATATTCTCGGAGTAAAAACCAATTACACCGTC_3'	61			482 pb
iFlaA2ΔN1	FlaA2 Δ1-42 _Δ221-239 *L. interrogans*	5'_GGTAGCGGATCGGAAAACCTGTATTTTCAGGGATCCGGTTCACTTGAAAAGCTCATACTAGAGAATTTCGAAGAG_3'	63	5'_GGTGGCTCCAGCTGCCGGATCCGGCACGAACTTCCAAATCGTCC_3'	60	607 pb
iFlaA2ΔN2	FlaA2 Δ1-57 _Δ221-239 *L. interrogans*	5'_GGTAGCGGATCGGAAAACCTGTATTTTCAGGGATCCGGTTCAAGAGCGAAGTCAACAACTCCTC_3'	63			562 pb
iFlaA2ΔN3	FlaA2 Δ1-67 _Δ221-239 *L. interrogans*	5'_GGTAGCGGATCGGAAAACCTGTATTTTCAGGGATCCGGTTCAACCAAGGTTCTCAAAATGGTCCA_3'	63			532 pb
bFlaA2ΔN1	FlaA2 Δ1-43 _Δ226-241 *L. biflexa*	5'_GGTAGCGGATCGGAAAACCTGTATTTTCAGGGATCCGGTTCACCTCTAAGAAAAATCGTTTTAGATGATTTTGAAGAGGC_3'	64	5'_GGTGGCTCCAGCTGCCGGATCCTGTTCTCACTTGTAAATCATCTACATAAAAGTAGAAGTC_3'	60	610 pb
bFlaA2ΔN2	FlaA2 Δ1-59 _Δ226-241 *L. biflexa*	5'_GGTAGCGGATCGGAAAACCTGTATTTTCAGGGATCCGGTTCAAGAGTCAAAGCAACCACTCCAC_3'	63			562 pb
bFlaA2ΔN3	FlaA2 Δ1-101 _Δ226-241 *L. biflexa*	5'_GGTAGCGGATCGGAAAACCTGTATTTTCAGGGATCCGGTTCACATATTTTAGGAGTCAAAACACATTTTGCTGC_3`	63			436 pb

Tabla 5: Construcciones diseñadas para clonado high-throughput de isoformas de FlaA. En gradientes de azul se muestran las construcciones para FlaA1 y en naranja las de FlaA2. Los colores más claros corresponden a *L.*

borgpetersenii, los más oscuros a *L. biflexa* y el intermedio a *L. interrogans*. Cada secuencia codificante fue subclonada en el juego de vectores "*vector suite*" mostrado en amarillo mediante RF-cloning. Se incluyen cebadores utilizados, temperaturas de fusión (Tm) y tamaño esperado del megaprimer.

El protocolo de RF cloning fue análogo al previamente descripto, con modificaciones para facilitar el clonado *high-throughput*. Las reacciones de PCR1 se escalaron aproximadamente cinco veces para producir megaprimer suficiente para toda la batería de vectores. Para PCR2 se preparó una master mix con H_2O, Buffer HF, dNTPs y polimerasa Phusion de alta fidelidad, que se dividió en varios tubos para añadir los megaprimers correspondientes y se alicuotaron en *blisters* de PCR. Los productos de PCR se transformaron en células *E. coli* TOP10F' utilizando bolitas de vidrio para facilitar el plaqueo. Como método confirmatorio, y para manejar una gran cantidad de plásmidos en paralelo, se utilizó el método *Colony PCR*. Por cada colonia a analizar se preparó un tubo de pre-cultivo (3 mL LB con antibiótico correspondiente) y uno de PCR (80 µL agua estéril). Cada colonia fue picada con tip, agitada en el tubo de PCR sin tocar las paredes y transferida al tubo de precultivo. Los precultivos se incubaron a 37°C toda la noche, mientras que los tubos de PCR se incubaron 5 minutos a 95-98°C para liberar el ADN. Posteriormente, 5 µL del agua de cada tubo se utilizaron para inocular un nuevo tubo de PCR con 20 µL de master mix alicuotado.

6.5 – Mutagénesis sitio-dirigidas

Las mutagénesis puntuales sitio-dirigidas se llevaron a cabo con RF cloning utilizando el mismo protocolo descripto antes. Los cebadores fueron diseñados para contener las mutaciones en un punto intermedio de su secuencia manteniendo la capacidad de hibridización. Para la mutagénesis de FcpA de *L. interrogans* para obtener la versión "cysless" en el plásmido pQE80-iFcpA, cada primer se diseñó alrededor de uno de los dos residuos de cisteína a mutar, sustituyendo Cys221 por treonina (C221T) y Cys241 por alanina (C241A).
Para la versión *strepless* de FlaA1 de *L. borgpetersenii*, se eliminó la etiqueta *streptag* fusionada al extremo C-terminal para facilitar cristalización. Se añadió un codón *stop* (TTA) al final de la secuencia codificante de FlaA1 y se eliminaron las secuencias posteriores río-abajo que codifican para la etiqueta *streptag*, haciendo que la secuencia 5' del cebador hibridice con el plásmido en una región 3’ en relación a las etiquetas.

6.6 – Cuantificación de ADN

La presencia de ADN plasmídico, su concentracón y su calidad, fueron determinadas utilizando un equipo NanoDrop (Thermo Scientific o DeNovix DS-11). La concentración de los ADN plasmidicos se estimó por determinación de su absorbancia (1 UA_{260} correspondiendo aproximadamente a 50µg/mL). Asímismo se verificó la calidad del ADN observando el

espectro de absorción y verificando que la relación de Abs 260nm/280nm estuviera en el rango 1.8-2.0.

6.7 – Secuenciación y análisis de las secuencias de ADN

Todas las construcciones generadas fueron confirmadas en última instancia por secuenciación, ya sea en el servicio de la Unidad de Biología molecular del Institut Pasteur de Montevideo o enviándolas a Macrogen. Para la secuenciación se utilizaron cebadores específicos de los vectores, y/o los cebadores del clonado. Para el análisis de las secuencias se construyeron con SnapGene Viewer software[195], los vectores con las secuencias esperadas del clonado, y se utilizó el programa Bioedit[196] para confirmar los plásmidos. En los análisis de similitud de secuencias se utiizaron los algoritmos de BLAST disponibles en la base de datos GeneBank.

6.8 – Preparación de células competentes y transformación

Células quimiocompetentes: Se partió de un precultivo fresco crecido a 37°C, inoculando 1/100 en 50 mL de medio SOC o LB estéril con antibiótico correspondiente. Se incubó a 37°C con 220 rpm por 2-3 horas hasta DO_{600} = 0.7. Las células se cosecharon centrifugando a 3,000g y 4°C durante 15 minutos.

Se resuspendió el pellet en ½ volumen de $CaCl_2$ 0.1M estéril, se incubó en baño de agua 30 minutos para permeabilizar membranas y se centrifugó a 3,000g y 4°C durante 15 minutos. El pellet se resuspendió en 1/40 volúmenes de $CaCl_2$ 0.1M y glicerol 10% frío. Las células competentes se fraccionaron en alícuotas de 50 µL y se guardaron a -80°C.

Células electrocompetentes: Se resuspendió el pellet en 1 volumen de agua estéril fría en baño de hielo. Se centrifugaron las células a 3,000g y 4°C durante 15 minutos y se resuspendió con ½ volumen. Se repitieron los pasos resuspendiendo en ¼ volumen de agua estéril y luego en 1/8 volumen de glicerol 10% frío. Se centrifugó una vez más resuspendiendo finalmente en 1 mL de glicerol 10% frío. Las células se fraccionaron y guardaron a -80°C.

Transformación química: Se incubaron 30-50 µL de células quimiocompetentes con 0.5-2 µL de ADN plasmídico en hielo (4°C) durante media a 1 hora. Los viales fueron incubados a 37°C por 45 segundos, inmediatamente sumergidos en hielo y se agregó 1 mL de medio LB o SOC estéril. Las células se incubaron 1 hora a 37°C con agitación vigorosa y 100 µL fueron sembrados en placas LB agar con antibiótico de resistencia correspondiente.

Transformación por electroporación: Se incubaron 30-50 µL de células electrocompetentes con 0.5-1.5 µL de ADN plasmídico en hielo durante 30 minutos y se transformaron mediante

pulso eléctrico de 2.5 voltios por 8 ms. Se agregó inmediatamente 1 mL de medio LB o SOC estéril y las bacterias fueron incubadas a 37°C 1 hora con agitación vigorosa antes de sembrar 100 µL en placas selectivas.

6.9 – Expresión y purificación general de proteínas

La estrategia general de expresión involucra transformación de células competentes con los vectores de interés, selección de colonias transformantes en placas LB-agar con antibiótico correspondiente, y cultivo en medios apropiados inoculados con 1/100 del pre-cultivo. Los cultivos se incubaron a 37°C y 220 rpm hasta fase exponencial media (DO_{600} entre 0.6 y 1.7 dependiendo del medio). La sobre-expresión se indujo con 1 mM IPTG por tiempos y temperaturas variables según cada proteína.

Para cosechar las bacterias, los cultivos se centrifugaron a 5,000g por 20 minutos a 4°C. Los precipitados celulares se resuspendieron en Buffer de lisis (50 mM Tris-HCl pH 8, 500 mM NaCl con inhibidores de proteasas *EDTA-free cocktail*, Roche). La lisis celular se realizó con lisozima (1 mg/mL) y por congelamiento a -80°C seguido de descongelado. Los lisados fueron incubados con DNAsa (100 µg/mL) y sonicados con Branson Digital sonifier (4 ciclos de 30 segundos, amplitud 30%, 1 minuto entre ciclos) o tratados con homogeneizador Emulsiflex (2-3 ciclos a 4°C). La fracción total fue centrifugada a 9,000g por 45 minutos a 4°C para separar fracción soluble de insoluble.

Todos los vectores fueron elegidos para expresar etiqueta de poli-histidinas (6xHisTag), generalmente fusionada en el extremo N-terminal, para purificación mediante cromatografía de afinidad a níquel inmovilizado (IMAC). El sobrenadante (fracción soluble) fue filtrado, llevado a la composición del buffer de lavado (generalmente con imidazol 20-40 mM) y cargado en columna HisTrap (GE Healthcare) de 1 o 5 mL pre-equilibrada con buffer A (50 mM Tris-HCl pH 8, 0.5M NaCl, 20 mM imidazol). La columna fue conectada a un equipo AKTA Xpress o Purifier, durante el lavado con Buffer A. Las eluciones se realizaron con gradientes de imidazol 0-100% buffer B (50 mM Tris-HCl pH 8, 0.5 M NaCl, 500 mM imidazol), siguiendo la purificación por Abs_{280}.

En construcciones con etiqueta de poli-histidinas seguida de secuencia de corte para proteasas TEV o PreScission, las proteínas fueron incubadas con la proteasa correspondiente (1 mg de proteasa por cada 40 mg de proteína eluída) y dializadas durante la noche contra un mínimo de 10 volúmenes de buffer (50 mM Tris-HCl pH 8, 0.5 M NaCl) para eliminar el imidazol. Las muestras se sometieron a una segunda separación por IMAC para remover la proteasa y purificar la proteína blanco. El *flowthrough* con la proteína de interés (sin His-tag) fue colectado y concentrado por ultrafiltración (Vivaspin, Sartorius).

El paso final de purificación consistió en cromatografía de exclusión molecular usando columnas Superdex S75 o S200 HiLoad 16/60 o 26/60 (GE Healthcare) según volumen de muestra a inyectar y el tamaño esperado de la proteína blanco. Todas las cromatografías se realizaron en equipos AKTA con columnas pre-equilibradas con Buffer SEC indicado en cada caso. Cada pico de elución fue colectado, concentrado por ultrafiltración y alicuotado para almacenamiento a -80°C.

6.10 – Expresión y purificación específica de proteínas individuales

FcpA y FcpB de *L. biflexa* y FcpB de *L. interrogans*: Transformación de células quimiocompetentes Rosetta-gami 2 (DE3) de *E. coli* con plásmidos pQE80-bFcpA, pQE80-bFcpB o pQE80-iFcpB. Cultivo a 37°C en 4 L de medio 2YT suplementado con $MgSO_4$ (2 mM), cloranfenicol (30 mg/L) y ampicilina (100 mg/L). Inducción por 4 horas a 30°C. Buffer SEC: 20 mM Hepes pH 7, 200 mM NaCl.

FcpA de *L. interrogans* con y sin cisteínas: Transformación de células electrocompetentes TOP10F' de *E. coli* con plásmido pQE80-iFcpA o pQE80-iFcpA cysless. Cultivo a 37°C en 4 L de medio TB con $MgSO_4$ (2 mM) y ampicilina (100 mg/L). A DO_{600} = 0.9, se indujo con 1 mM IPTG por 4 horas a 30°C. Buffer SEC: 20 mM Hepes pH 7, 100 mM NaCl. El cromatograma mostró mezcla de monómero y dímero, que fueron colectados y concentrados independientemente.

Isoformas de FlaA en citoplasma: Los plásmidos pETM conteniendo la región codificante de FlaA1 o FlaA2 de *L. interrogans* (sin secuencia señal) en marco con distintas proteínas de fusión fueron transformados en cepas Rosetta-Blue, BL21(DE3) o C43. Cultivo a 37°C en 10 mL LB hasta DO_{600} = 0.8, inducción por 20 horas a 20°C o 4 horas a 30°C. Para escalado de FlaA2 con vector pETM66-FlaA2, se utilizó cepa BL21 en 1 L de cultivo con inducción a 20°C por 20 horas.

Isoformas de FlaA en periplasma: Los vectores pETM54, pET28a y pET22b se transformaron en cepas BL21(DE3) y C41, cultivando en 50 mL de medio 2YT hasta DO_{600} = 0.5-0.8 con inducción a 20°C por 18 horas. Las células fueron recuperadas y resuspendidas en buffer TES para extracción periplásmica según protocolo de Pardon et al., 2014. La fracción periplásmica se purificó por IMAC en beads usando Ni-Sepharose. Para escalado, se utilizaron 3 L de cultivo con columna HisTrap de 5 mL.

6.11 – Expresión y purificación *high – throughput* de isoformas de FlaA

Las construcciones (Tabla 5) fueron transformadas en células quimiocompetentes Lemo21 de *E. coli* que permiten ajustar la expresión mediante ramnosa[197].

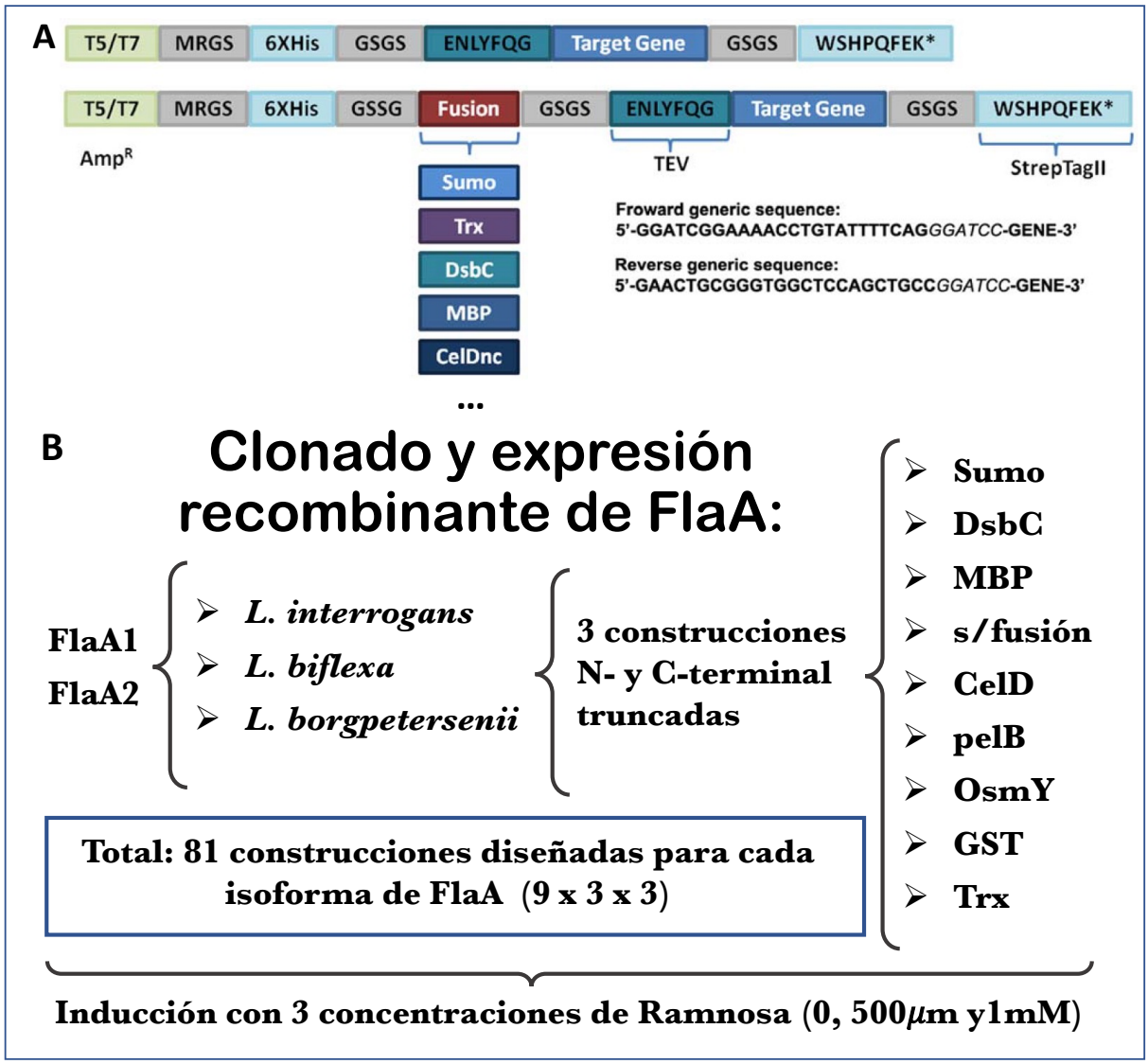


Figura 45: Estrategia de clonado *high-throughput* para expresar las dos isoformas de FlaA. A. Representación esquemática del vector suite. Se eligieron los vectores con el promotor T7, que tienen un gen de resistencia a ampicilina para la selección, y la secuencia codificante genérica que se muestra. Este gen comienza con la secuencia MRGS y una etiqueta de 6 hisidinas en el extremo N-terminal, seguidas de las proteínas de fusión separadas por secuencais GSGS. Se utilizaron 9 proteínas de fusión diferentes: Sumo, Trx, DsbC, MBP, CelDnC, y las no mostradas: PelB, OsmY, GST y Trx. Entre la proteína de fusión y la secuencia codificante para GFP (no mostrada), que luego será reemplazada por el gen de interés, hay una secuencia de corte con la proteasa TEV (Tobacco Etch Virus). Por último, se muestran también las secuencias genericas para la hibridación de los cebadores a los vectores. **B.** Construcciones clonadas y expresadas: para cada isoforma y cada una de las tres especies se generaron tres construcciones N y C-terminal truncadas. Para la isoforma FlaA1 de *L. biflexa* las construcciones*:* bFlaA1N1(Δ1-45_Δ307-328), bFlaA1N2(Δ1-63_Δ307-328) y bFlaA1N3(Δ1-101_Δ307-328). Para FlaA2 de *L. biflexa*: bFlaA2N1(Δ1-43 _Δ226-241), bFlaA2N2(Δ1-59 _Δ226-241) y bFlaA2N3(Δ1-101 _Δ226-241). Para la isoforma FlaA1 de *L. interrogans* las construcciones*:* iFlaA1N1(Δ1-49_Δ273-317), iFlaA1N2(Δ1-69_Δ273-317) y iFlaA1N3(Δ1-89_Δ273-317). Para FlaA2 de *L. interrogans*: iFlaA2N1(Δ1-42 _Δ221-239), iFlaA2N2(Δ1-57 _Δ221-239) y iFlaA2N3 (Δ1-67 _Δ221-239). Para la isoforma FlaA1 de *L. borgpetersenii* las construcciones*:* boFlaA1N1(Δ1-49_Δ303-316), boFlaA1N2(Δ1-71_Δ303-316) y boFlaA1N3(Δ1-100_Δ303-316). Para FlaA2 de *L. borgpetersenii*: boFlaA2N1(Δ43), boFlaA2N2(Δ67) y boFlaA2N3(Δ99). Cada una de estas secuencias codificantes fue subclonada en los 9 plásmidos del vector suite diseñando en total 81 construcciones para cada isoforma de FlaA. Cada construcción fue utilizada para evaluar ensayos de expresión bajo 3 concentraciones diferentes de Ramnosa, por lo que en total se pueden probar 243 condiciones para cada isoforma.

Los ensayos de expresión high-throughput se realizaron usando pipetas multicanal y placas de cultivo de 24 pocillos cubiertas con cinta permeable al aire (AirPore tape, Qiagen) (Figura 46). Se iniciaron precultivos de transformantes en medio LB, se incubaron a 37°C y al día siguiente se transfirieron a placas nuevas de 24 pocillos cultivando 4 mL/pocillo de medio TB. Para asegurar crecimiento parejo fue necesario agregar un paso intermedio generando precultivos nuevos antes de inducir la sobre-expresión con cultivos frescos al tercer día. Los cultivos se realizaron con mix de medio TB, sales TB (1X), glicerol (2 mL/L), cloranfenicol (34 µg/mL), ampicilina (100 µg/mL) y $MgSO_4$ (2 mM). Cada columna de placa se llevó a la concentración final apropiada de ramnosa (0 mM, 500 µM o 1 mM). Las placas se incubaron a 37°C con agitación a 220 rpm. A DO_{600} aproximadamente 1.6 se indujeron con 100 µL de medio TB completo con ampicilina y 40 mM IPTG (inducción final 1 mM). Las placas se incubaron a 20°C, 220 rpm por 17 horas.

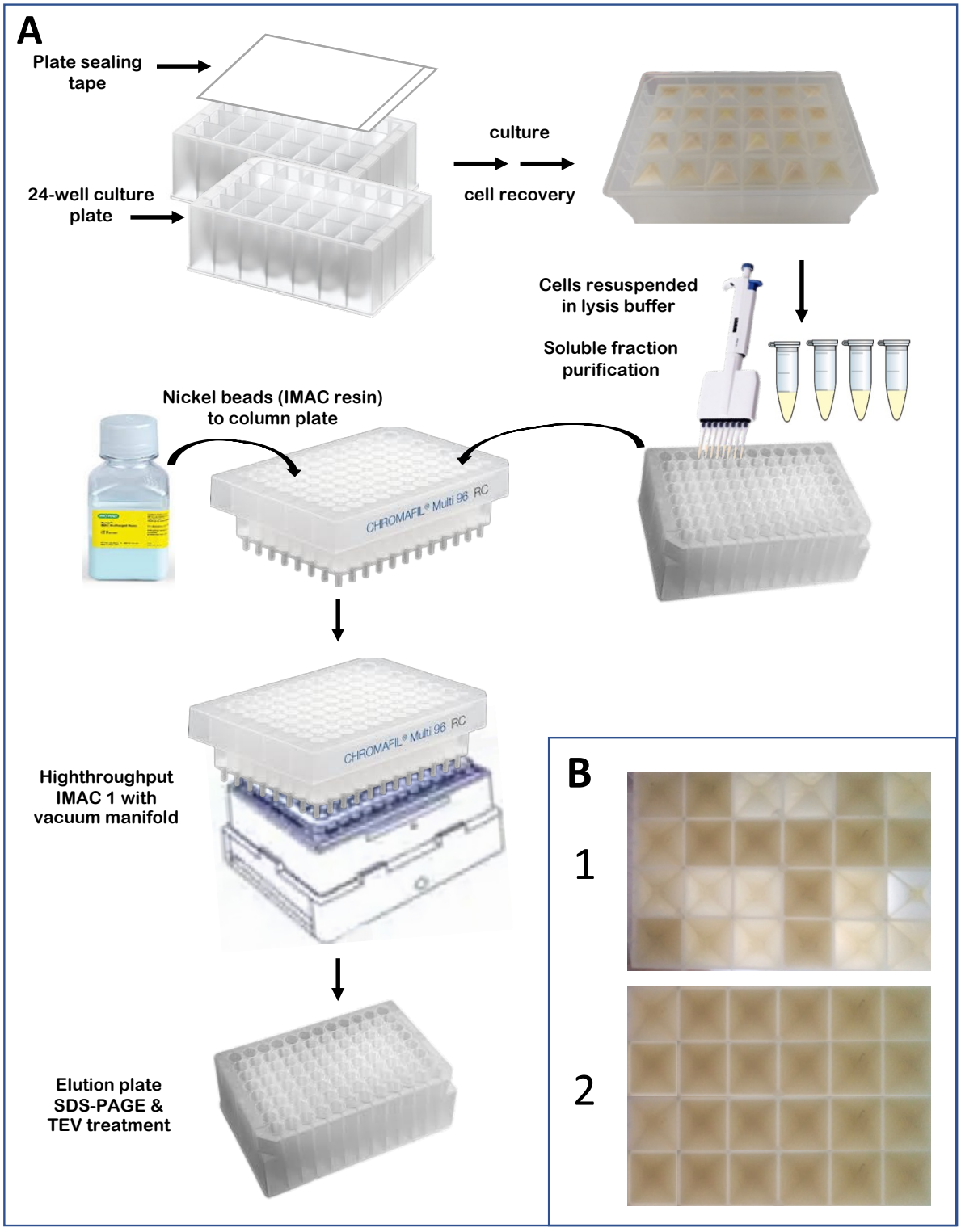


Figura 46: Estrategia de expresión y purificación *high-throughput* para expresar las dos isoformas de FlaA. A. Los precultivos y cultivos se realizan en placas de 24 pocillos con fondo redondo, cubiertos de cinta permeable al aire (AirPore). Los pellet celulares son recuperados por centrifugación y resuspendidos en buffer de lisis. Se agrega DNAsa y se purifica la fracción soluble según se describe en el texto. Estas fracciones solubles son transferidas a placas de 96 pocillos y llevadas a la composición del buffer A. Se toman alícuotas para SDS-PAGE y el resto se transfiere a placas de microcolumnas precargadas con resina IMAC y pre-equilibradas con buffer A. Se realizan todas las IMAC en simultaneo con la ayuda de un dispositivo de vacío para placas. Se guarda el FT de las columnas en una placa y se lleva a cabo el lavado cargando con micropipeta multicanal dos veces con buffer A y descartando el contenido eluído de las placas. Se eluye con 100% buffer B y la elución se colecta en otra placa de 96 pocillos. Se guardan fracciones para SDS-PAGE y el resto de la muestra se transfiere con pipetas multicanal

a blisters de PCR para su tratamiento posterior con TEV. Las fracciones resultantes también se guardan para correr por SDS-PAGE. **B**. 1- placa de 24 pocillos resultante de cultivar las células transformadas con medio TB luego de un paso de precultivo, se ven en algunos pocillos mucha mayor densidad celular. 2- placa de 24 pocillos resultantes de cultivar las mismas células transformadas, pero ahora luego de 2 pasos de precultivo consecutivos antes del cultivo para inducir.

Los pellets celulares se recuperaron por centrifugación en placa a 5,000g por 20 minutos a 4°C. Se descartaron transformantes con precipitados fluorescentes (expresión de GFP). Los pellets restantes se resuspendieron en 1 mL de Buffer de lisis con lisozima. La lisis se completó por congelamiento a -80°C. Los lisados fueron descongelados a 30°C y tratados con DNAsa (10 µg/mL) y $MgSO_4$ (10 µM) por 1.5 horas. Las fracciones totales se transfirieron a tubos eppendorf y se centrifugaron a 9,000g por 30 minutos a 4°C. Los sobrenadantes con proteína soluble se transfirieron a placas Deep well de 96 pocillos y se llevaron a composición de buffer A de IMAC. Se preparó resina de IMAC (Ni Sepharose 6FF) que se lavó con buffer A. La resina se cargó en placas con fracción soluble e incubó 20 minutos a 20°C. Las muestras se transfirieron a placas de columnas con filtro y se filtró con bomba de vacío por presión positiva. Las columnas se lavaron dos veces con 1 mL buffer A y se eluyó después de incubar 2 minutos con 150 µL buffer B. Las eluciones se trataron con proteasa TEV guardando muestras para SDS-PAGE en cada paso.

6.12 – Escalado de producción de isoformas de FlaA

Transformación de células quimiocompetentes *E. coli* Lemo21 con plásmidos de interés. Cultivo a 37°C en 4 L de medio TB con sales TB (10X), $MgSO_4$ (2 mM), glicerol (2 mL/L), ampicilina (100 µg/L), cloranfenicol (34 µg/L) y concentración especificada de ramnosa. A DO_{600} = 1.5, se cambió temperatura a 20°C por 20 minutos y se indujo con 1 mM IPTG por 18 horas a 20°C. Buffer SEC: 20 mM Tris pH 7, 150 mM NaCl.

Durante la optimización de purificación de FlaA1 N49 de *L. borgpetersenii* y FlaA1 de *L. biflexa* se realizaron las siguientes modificaciones:

Primera etapa: Agregado de 1 mM $CaCl_2$ en medio de cultivo inicial. Agregado de $CaCl_2$ en Buffer de lisis (composición final: 50 mM Tris pH 8, 0.5M NaCl, 20 mM imidazol, inhibidor de proteasa EDTA-free cocktail, 10 mM $CaCl_2$, lisozima 1/100). Agregado de $CaCl_2$ a concentración final 10 mM en muestra eluída de primera IMAC, buffer de diálisis y buffer SEC.

Segunda etapa: Agregado de β-mercaptoetanol en Buffer de lisis (composición final: 50 mM Tris pH 8, 0.5M NaCl, 20 mM imidazol, inhibidor de proteasa *EDTA-free cocktail*, 10 mM $CaCl_2$, 20 mM β-mercaptoetanol, lisozima 1/100). Después de IMAC 2 se agregó iodoacetamida 5 veces en exceso, se incubó 1 hora y se realizó diálisis durante la noche en el mismo buffer usado para SEC para eliminar el exceso.

6.13 – Análisis de proteínas por electroforesis (SDS-PAGE)

En todas las purificaciones de proteínas la pureza y el tamaño de las mismas fue verificado por SDS-PAGE. Lo mismo para el análisis y seguimiento de la purificación de filamentos flagelares de *Leptospira* (ver más abajo). En todos estos casos se realizó electroforesis en geles separadores de poliacrilamida 12% con SDS (con geles concentradores 5%), de acuerdo a los protocolos estándar. Las muestras se mezclaron con buffer de carga 5X (10% SDS, 0.5M DTT, 50% glicerol, 0.31M Tris-HCl y 0.05% Azul de Bromofenol, 7% β-mercaptoetanol). En los casos puntuales donde se uso otro buffer de muestra (ie. sin β -mercaptoetanol) esto se detalla. Las muestras se calentaron a 100ºC durante 5 minutos. Los geles se corrieron en cámaras verticales de electroforesis (Hoefer o Biorad). Se utilizó buffer de corrida 1X (25mM Tris pH 8.3, 200mM Glicina, 0.1% SDS) a 25 mA por gel. Para la estimación del tamaño molecular se utilizaron marcadores de peso molecular comerciales indicados en cada caso. Los geles se tiñeron con una solución de azul de Coomassie (0.025% azul de Coomasie, 10% ácido acético, 90% etanol) y se decoloraron en una solución de decoloración (10% ácido acético, 50% etanol)

6.14 – Cuantificación de proteínas

Para cuantificar proteínas recombinantes, las cantidades fueron estimadas por la Abs_{280} con los equipos AKTA, que permiten además seguir las purificaciones. Para mediciones más exactas de la cantidad total de proteína purificada se utilizó el espectrofotómetro Cary 50 Bio (Varían) que también permite medir absorbancia a 280nm, y hacer espectros de absorción para chequear la calidad de las muestras. Para estimar las cantidades de proteínas puras, y las presentes en purificaciones de filamentos flagelares para ensayos posteriores de espectrometría de masas, se utilizó densitometría de geles SDS-PAGE 12%, donde las muestras fueron corridas junto al estandar kit LMW-SDS Marker (GE Healthcare) que permite luego construir una curva de calibración. Sabiendo la cantidad de proteína sembrada se puede medir la intensidad de las banda por densitometría. Finalmente para algunas purificaciones de filamentos flagelares, la cantidad de proteína total se midio por Bradford[198] construyendo una curva estándar de albúmina sérica bovina.

6.15 – Análisis de estabilidad con nano-fluorimetría diferencial de barrido

Utilizando el equipo Prometheus NT.48 (Nanotemper), se realizaron curvas de temperatura de 20 a 60-110°C con incrementos de 1°C/min y potencia de LED del 30%. Las muestras del mutante N1 de FlaA1 de *L. borgpetersenii* en distintos buffers fueron cargadas en capilares que se sellaron para experimentos con rampas mayores a 95°C.

Se utilizó proteína a concentración inicial de 2.4 mg/mL diluyéndola en el buffer seleccionado unas 20 veces hasta concentración final de 0.12 mg/mL. La proteína TEV utilizada como control negativo se ensayó a concentración final de 0.29 mg/mL, diluyendo proteína a 7.3 mg/mL unas 25 veces en los buffers seleccionados.

6.16 – Cristalización de proteínas

Para cristalizar las proteínas se comenzó realizando rastreo robótico de condiciones de cristalización por difusión de vapor al equilibrio en gota sentada, utilizando placas de 96 pocillos CrystalQuick (Greiner). Se utilizaron matrices dispersas de condiciones comercialmente disponibles: JCSG Core I-IV Suites (Qiagen), PEGs Suite (Qiagen), pH clear (Qiagen), $AmSO_4$ suites (Qiagen), Crystal Screen y Crystal Screen II (Hampton Research).

Las gotas se prepararon mezclando 300 nL de proteína con 300 nL de cada solución de cristalización, dispensados con el robot Honeybee963® (Isogen Life Science) disponible en el laboratorio. En el reservorio de cada pocillo se colocaron 100 µL de solución de cristalización. Los cristales de mejor apariencia fueron eventualmente optimizados manualmente, utilizando un *setup* de gota colgante, con gotas de 4 µL (2 µL proteína + 2 µL reservorio) dispensadas manualmente en placas VDX de 24 pocillos (Hampton Research). Los reservorios contenían 1 mL de soluciones de cristalización preparadas manualmente, o bien preparadas utilizando un robot Alchemist (Rigaku). Las placas se dejaron reposar en sala con temperatura controlada y se revisaron periódicamente hasta aparición de cristales.

Proteína	Concentración	Buffer	Kit de cristalización
bFcpA	27,5 mg/mL	20 mM Hepes pH 7, 200 mM NaCl	JCSG Core Suites I al IV PEG Suite
iFcpA_cysless forma monomérica	51 mg/mL	20 mM Hepes pH 7, 200 mM NaCl	JCSG Core Suites I al IV
iFcpA_cysless forma monomérica	51 mg/mL	20 mM Hepes pH 7, 200 mM NaCl	PEG Suite / PHclear $AmSO_4$ Suite (70%) Crystal Screen (85%)
bFcpB	31 mg/mL	20mM Tris pH 8, 300mM NaCl	JCSG Core Suites I al IV
bFcpB	68 mg/mL	20 mM Hepes pH 7, 200 mM NaCl	PEG Suite / PHclear $AmSO_4$ Suite (70%) Crystal Screen (85%)
boFlaA2N1_streptag pico 1 SEC	8,3 mg/mL	Tris 20mM pH8, 150mM NaCl	PEG Suite / PHclear $AmSO_4$ Suite (70%) Crystal Screen (85%)
boFlaA2N1_streptag pico 2 SEC	12,8 mg/mL	Tris 20mM pH8, 150mM NaCl	PEG Suite / PHclear $AmSO_4$ Suite (70%) Crystal Screen (85%)
boFlaA1N1_streptag pico 1 SEC	4,2 mg/mL	Tris 20mM pH8, 150mM NaCl	JCSG Core Suites I al IV
boFlaA1N1_streptag pico 1 SEC	4,2 mg/mL	Tris 20mM pH8, 150mM NaCl 10mM $CaCl_2$	JCSG Core Suites I al IV
boFlaA1N1_streptag pico 2 SEC	8,1 mg/mL	Tris 20mM pH8, 150mM NaCl 10mM $CaCl_2$	JCSG Core Suites I al IV
boFlaA1N1_streptag purificada con $CaCl_2$	25,0 mg/mL	Tris 20mM pH8, 150mM NaCl 10mM $CaCl_2$	PEG Suite / PHclear $AmSO_4$ Suite (70%) Crystal Screen (85%)
boFlaA1N1_streptag purificada con $CaCl_2$ (quelado al final)	20,0 mg/mL	Tris 20mM pH8, 150mM NaCl	PEG Suite / PHclear $AmSO_4$ Suite (70%) Crystal Screen (85%)
bFlaA1N1_streptag purificada con $CaCl_2$	14,3 mg/ mL	Tris 20mM pH8, 150mM NaCl	JCSG Core Suites I al IV
bFlaA1N1_streptag purificada con $CaCl_2$	14,3 mg/ mL	Tris 20mM pH8, 150mM NaCl 10mM NaCitrate	JCSG Core Suites I al IV
bFlaA1N1_streptag purificada con $CaCl_2$	14,3 mg/ mL	Tris 20mM pH8, 150mM NaCl 10mM N-Acetil Glucosamine	JCSG Core Suites I al IV
boFlaA1N_strepless purificada con $CaCl_2$	22,9 mg/mL	Tris 20mM pH8, 150mM NaCl	JCSG Core Suites I al IV
boFlaA1N_strepless purificada con $CaCl_2$	22,9 mg/mL	Tris 20mM pH8, 150mM NaCl 10mM NaCitrate	JCSG Core Suites I al IV
boFlaA1N_strepless purificada con $CaCl_2$	22,9 mg/mL	Tris 20mM pH8, 150mM NaCl 10mM N-Acetil Glucosamine	JCSG Core Suites I al IV
boFlaA1N_strepless purificada con $CaCl_2$	22,9 mg/mL	Tris 20mM pH8, 150mM NaCl 10mM Acido N-Acetil Muramico	PEG Suite / PHclear $AmSO_4$ Suite (70%) Crystal Screen (85%)

Tabla 6: Condiciones utilizadas para rastreos de condiciones de cristalización. En todos los casos se utilizó formato de gota sentada mezclando 300 nL de proteína indicada con igual cantidad del reservorio correspondiente. El dispensado se realizó con robot Honeybee.
Para difracción de rayos X, los cristales fueron colectados en loops de nylon y congelados en nitrógeno líquido después de incubación rápida con solución crio-protectora compuesta de solución de cristalización y 25% (v/v) glicerol.

6.17 – Difracción de rayos X y resolución de estructuras

Los cristales se analizaron mediante difracción de rayos X en condiciones criogénicas utilizando el generador microfuente de rayos X de ánodo rotatorio de cobre Micromax007-HF (Rigaku) disponible en la plataforma de cristalografía de proteinas del IPMontevideo, equipado de un detector de área Mar345 (Mar Research). Para las proteínas FcpA y FcpB de *L. biflexa*, los cristales con mejor difracción fueron llevados al sincrotrón Soleil (París), mientras que para FlaA1 de *L. borgpetersenii* se enviaron muestras al sincrotrón Diamond (Oxfordshire).

Los datos de difracción se procesaron con diversos paquetes de programas: XDS, CCP4, PHENIX, ShelX. El procesamiento de imágenes se realizó principalmente con iMosflm[199] y Aimless[200] o bien con autoPROC. La resolución de estructuras de FcpA y FcpB se realizó por reemplazo molecular a partir de estructuras previamente resueltas usando programa Phaser. Para la resolución de la estructura de FlaA1 se empleó una estrategia combinada de derivatización con yodo y reemplazo molecular. Inicialmente, se utilizó un modelo de FlaA1 generado con Rosetta como plantilla para el reemplazo molecular. Posteriormente, las fases se mejoraron mediante datos de dispersión anómala a una sola longitud de onda (SAD), obtenidos a partir de cristales derivatizados con yoduro.

6.18 – Gel filtración analítica

La interacción *in vitro* entre proteínas recombinantes bFcpA y bFcpB se evaluó por cromatografía analítica de exclusión por tamaño. Se prepararon en buffer 20 mM Hepes pH 7.5, 100 mM NaCl, diluciones 50 µM de ambas proteínas por separado, así como mezcla equimolar de ambas. Las muestras fueron incubadas durante 1 hora a temperatura ambiente y centrifugadas a 16,000g por 10 minutos a 4°C previo a inyección. Para cada muestra se analizaron dos corridas cargando en loop de igual volumen 100 µL de muestras que fueron inyectadas en columna Superdex S75 10/300 pre-equilibrada en buffer 50 mL.

6.19 – Modelado de monómero y filamentos de FlaB

Para construir modelos por homología de cada una de las cuatro isoformas de FlaB de *L. interrogans*, se realizó búsqueda de proteínas homólogas con estructura conocida utilizando el servidor HHpred[201]. Utilizando el software Rosetta suite[202], a partir de los

alineamientos y de la estructura pdb 1UCU (correspondiente al filamento de *S. enterica*), se construyeron modelos estructurales de filamentos generados por múltiples copias de una misma FlaB. Basándose en la simetría helicoidal del filamento de *S. enterica* se restringieron los grados de libertad en la ubicación de cada protómero[202]. Para cada filamento de FlaB se generaron 600 modelos independientes utilizando RosettaCM[203]. Para el análisis posterior, tanto de las propiedades estructurales de cada proteína como del filamento que generan, se seleccionó el ensamble de los 10 modelos con menor REU (*Rosetta Energy Unit*). Se utilizó PyMol para alinearlos estructuralmente, y luego comparar y analizar las propiedades de hidrofobicidad, así como para generar las figuras.

6.20 – Cultivos de *Leptospira*

Para cultivar las leptospiras se utilizó medio líquido Ellinghausen-McCullough-Johnson-Harris (EMJH) base (Difco) suplementado con 4 mg/L $ZnSO_4 \bullet 7H_2O$, 15 mg/L $MgCl_2 \bullet 6H_2O$, 15 mg/L $CaCl_2 \bullet 6H_2O$, 100 mg/L de piruvato de sodio, 0.2 mg/L de vitamina B12, 0.05 g/L $FeSO_4 \bullet 7H_2O$, 0.4 g/L de glicerol y 1.25 g/L de Tween80. El medio se llevó a pH 7.6, fue filtrado y autoclavado antes de controlar su esterilidad.

Los precultivos de *Leptospira* a partir de stocks de *L. biflexa* salvaje o mutantes congelados en DMSO 0.025% se iniciaron descongelando rápidamente e inoculando 1 mL en 36 mL de medio de cultivo EMJH base. Los cultivos líquidos se realizaron en medio EMJH inoculando 1/100 de un cultivo anterior. Todos los cultivos se crecieron a 30°C sin agitación, y se midió densidad óptica a 420 nm (DO_{420}) en un espectrofotómetro tomando 1 mL de cultivo cada 24 horas. La observación de las bacterias se realizó con un microscopio de campo oscuro (Unidad Mixta Pasteur-INIA). Se realizaron curvas de crecimiento de las distintas cepas para estimar los tiempos de incubación.

6.21 – Purificación de filamentos flagelares de *L. biflexa* y *L. interrogans*

Para obtener filamentos flagelares, se cultivaron la cepa salvaje y los mutantes knockout *fcpA*- y *fcpB*- de *L. biflexa* y *L. interrogans* en 250 mL (500 mL o 1000 mL, escalando proporcionalmente todas las cantidades) de medio EMJH suplementado. La purificación se basó en el protocolo descrito anteriormente por Wunder et al.[105]con pequeñas modificaciones.

En todos los casos a una DO_{600} = 0.2-0.4, los cultivos fueron centrifugados 15 minutos a 8,000g y 4°C, y los precipitados lavados con PBS, re-centrifugados y finalmente resuspendidos en 15 mL de buffer 0.5M sacarosa, 0.15 M Tris, pH 8.0. Las células se agitaron primero 10 minutos en hielo y luego 30 minutos a temperatura ambiente con 1% Triton X-100 (concentración final). A continuación, se añadieron gota a gota 150 µL de lisozima (10 mg/mL), agitando 5

minutos a 4°C y luego se agregó lentamente 1.5 mL de 20 mM EDTA pH 8.0, agitando la solución 2 horas a temperatura ambiente.

Este paso fue seguido de dos ciclos de agitación a temperatura ambiente por 5 minutos; el primero luego de añadir 300 µL 0.1 M $MgSO_4$, y el segundo agregando 300 µL 0.1 M EDTA pH 8. Finalmente, los filamentos flagelares fueron separados de los restos celulares centrifugando a 17,000g por 15 minutos a 4°C. El sobrenadante fue precipitado con 2 mL de una solución 20% PEG 8000, 1M NaCl. Luego de incubar 30 minutos en hielo, los filamentos fueron recuperados centrifugando 30 minutos a 27,000g y 4°C, y resuspendidos en 4.5 mL H_2O. En el último paso de la purificación, los flagelos fueron ultracentrifugados a 80,000g por 45 minutos y resuspendidos en 100 µL H_2O con 2 gotas de 1% azida de sodio en PBS.

Durante los experimentos de espectrometría de masas y XL-MS este protocolo fue optimizado agregando:

- Benzonasa o DNAsa a una concentración final de 0.005 mg/mL y RNAsa a una concentración final de 0.01 mg/mL después de la incubación con lisozima y previo a la adición del EDTA.
- Resuspensión final de los filamentos en buffer PBS o Tris y no en agua.

Para experimentos de purificación de filamentos con y sin calcio se modificó este protocolo de dos formas alternativas:

Protocolo 1: Luego del agregado de lisozima la muestra se incubó durante 2.5 horas, y luego toda la noche a 4°C con agitación lenta para ayudar al disgregado de los filamentos. Al día siguiente se continuó el protocolo agregando 20 mM final de calcio a la mitad de la muestra y 10 mM *o*-fenantrolina en DMSO a la otra mitad.

Protocolo 2: Desde la primera resuspensión del pellet en buffer Tris-Sacarosa, se mantuvo la concentración de $CaCl_2$ en 20 mM, agregando luego de cada adición de EDTA 20 mM $CaCl_2$. En el último paso la mitad de la muestra fue resuspendida en buffer con 20 mM $CaCl_2$ y la otra mitad en buffer sin calcio y con 10 mM *o*-fenantrolina.

6.22 – Ensayos de reconstitución *in vitro* de filamentos flagelares de *L. biflexa*

Para analizar la interacción de las proteínas en el contexto del filamento flagelar, se evaluó la incorporación de la proteína recombinante bFcpB en filamentos purificados a partir de la cepa *fcpB*- de *L. biflexa*. Para ello, 420 µL de flagelos provenientes de la mezcla de tres purificaciones independientes se llevaron al mismo buffer en que se purificó bFcpB (20 mM Hepes pH 7.5, 100 mM NaCl) agregando 8.4 µL 5M NaCl, 4.2 µL 1M Hepes ácido y 4.2 µL 1 M Hepes base.

A continuación, 140 µL de estos filamentos (concentración final de proteínas totales = 69.3 mg/mL, medidos por la técnica de Bradford) fueron incubados durante 0.5 horas a temperatura ambiente con: (a) 135 µL bFcpB (30 mg/mL), que representan un 20% de la proteína total en la mezcla; (b) 7.75 µL bFcpB (30 mg/mL), que representan un 2% de la cantidad total de proteína en la mezcla; o (c) 135 µL 20 mM Hepes pH 7.5, 100 mM NaCl.

Transcurrido el tiempo de incubación, las muestras (a), (b) y (c), así como una cuarta muestra (d) 135 µL de bFcpB (30 mg/mL) sin filamentos, fueron lavadas dos veces con 4.5 mL de buffer y recuperadas por ultracentrifugación a 80,000g por 45 minutos. Los sobrenadantes de ambas centrifugaciones para cada muestra fueron mezclados y concentrados con ácido tricloroacético. Para evaluar la reconstitución de los filamentos se sembraron 10 µL de cada muestra junto a 10 µL de una muestra de filamentos purificados a partir de la cepa salvaje de *L. biflexa*, y una calle con marcadores comerciales de pesos moleculares.

6.23 – Entrecruzamiento químico de proteínas con NNP9

Los filamentos purificados fueron entrecruzados inicialmente con 5 nmoles del agente NNP9 cada 10 µg de proteína total en la muestra (Figura 47). La reacción se incubó 1 hora a 4°C y luego se frenó con 100 mM bicarbonato de amonio. El producto de la reacción fue transferido a un filtro molecular (Amicon, con corte de 10 kDa) previamente pasivado con 10% Tritón o Tween. Las muestras se lavaron 3 veces con 400 µL de buffer de digestión (50 mM bicarbonato de amonio) centrifugando en cada paso 1 minuto a 10,000g.

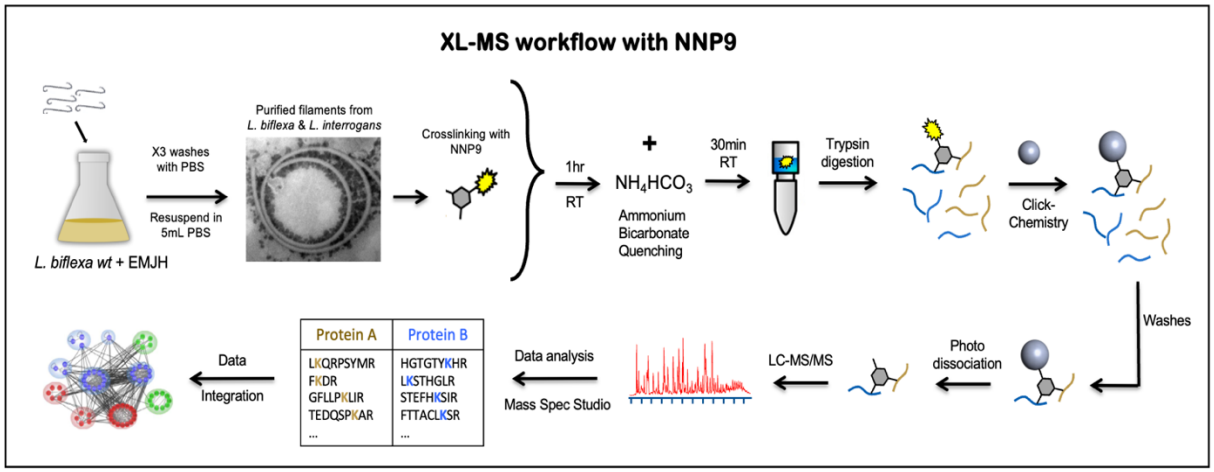


Figura 47: Esquema de trabajo para el entrecruzamiento con NNP9. Cepas de *Leptospira* fueron cultivadas en medio EMJH. Sus precipitados lavados y resuspendidos en buffer PBS1x. Sus filamentos fueron purificados como se detalla en la sección, fueron incubados con el agente de entrecruzamiento NNP9 por 1hr a temperatura ambiente (T_{amb}). La reacción fue terminada incubada media hora a T_{amb} con bicarbonato de amonio. Las proteínas entrecruzadas fueron digeridas con tripsina durante toda la noche, y los péptidos trípticos purificados con bolitas sólidas de agarosa unidas a grupos alcalinos. Esta purificación se basa en la unión covalente del grupo azido del

NNP9, a los grupos alcalinos de la matriz. Esta última unión entre la matriz y los grupos alcalinos es fotoclivable por lo que después del lavado de la muestra en por ultrafiltración (Centricon, 10kDa punto de corte). Los péptidos entrecruzados purificados fueron disociados de la matriz sólida por exposición a luz UV, y recuperados por centrifugación. El sobrenadante fue sometido a LC-MS-MS y los péptidos analizados por el software Mass Spec Studio. La integración de los datos permite mapear las distancias de nuestro complejo.

Se añadió 1 µg de tripsina por cada 50 µg de proteína y se incubó la reacción durante la noche a 37°C. Al día siguiente se recuperaron los péptidos trípticos por centrifugación a 10,000g. Se lavó el filtro con buffer de digestión y se sonicó la muestra por 15 minutos para liberar los péptidos unidos a la membrana del dispositivo Centricon. Los péptidos recuperados fueron pasados a un Eppendorf (Lowbind) y se midió y comparó la cantidad de péptidos recuperados por espectrofotometría.

Las muestras fueron incubadas con 1 µL de cuentas de agarosa alcalinas fotoclivables cada 10 nmoles de NNP9 inicialmente utilizado. La matriz fue lavada 3 veces con agua miliQ, incubando a temperatura ambiente, agitando durante 5 minutos y recuperando el sobrenadante en cada ciclo. Los péptidos entrecruzados fueron así purificados de sus pares no unidos a la matriz sólida (protocolo eFASP, *enhanced Filter Aided Sample Preparation*). Para remover el sobrenadante sin pipetear las cuentas de agarosa, se utilizaron tips de 200 µL conectados a 20 µL geLoader tips.

Luego de los lavados, la matriz con los péptidos entrecruzados fue incubada con sulfato de cobre en un cociente molar 5:1 con NNP9, con THPTA (Tris-benziltriazolil-metil-amina) en un cociente molar 25:1, y posteriormente con ascorbato de sodio en un cociente molar 50:1, todas incubaciones de 30 minutos a temperatura ambiente. Estos reactivos permiten y aceleran la cicloadición posterior que da lugar a la ruptura del enlace covalente con radiación UV.

Luego de los lavados posteriores, con o sin reducción y alquilación de cisteínas, los péptidos fueron transferidos a placas de 96 pocillos previamente pasivadas donde se sembró 50 µL de matriz por pocillo y se incubó 10 minutos a temperatura ambiente con lámpara UV (365 nm) y agitación. La reacción de cicloadición hace que la agarosa se torne amarillenta y precipite al fondo de la placa, por lo que los péptidos entrecruzados fueron recuperados del sobrenadante, liofilizados y conservados a -20°C. Para su análisis por espectrometría de masas los péptidos fueron descongelados y resuspendidos en 0.1% ácido fórmico en agua.

Para facilitar la solubilización y desnaturalización de las proteínas en la muestra, ayudando a promover la digestión tríptica, se ensayaron las siguientes condiciones previo a la incubación con tripsina:

- Incubación con SDS a concentración final de 0.1% por 30 minutos a temperatura ambiente, con agitación.

- Incubación con urea a distintas concentraciones (2-8M) a temperatura ambiente por 1-4 horas.
- Incubación con 2-5% N-lauroilsarcosina (Sarkosil) en buffer 20 mM Tris pH 8, 1 mM EDTA a 37°C por 45 minutos.
- Incubación con Rapigest concentrado en 200 µL, hirviendo o no la muestra por 45 minutos.
- Incubación con LysC en urea 8M por 4 horas.
- Digestión en gel según protocolo Müller et al., 2019[204] que involucra el atrapado de la muestra en una matriz 30% acrilamida/bis-acrilamida con 10% SDS, el cortado del gel en cuadrados de aproximadamente 2 mm^2, y la deshidratación previo a la incubación con tripsina y la extracción de los péptidos.

6.24 – Análisis por espectrometría de masas de péptidos entrecruzados

En todos los casos, para el análisis por espectrometría de masas, se inyectaron aproximadamente 6 µL de muestra a un flujo de 10 µL/min y con un gradiente en columna C-18 de 1 a 35% acetonitrilo en 0.1% ácido fórmico en 120 minutos, en un Nano-HPLC (Thermo Easy-nLC) acoplado a un espectrómetro de masas HF-Orbitrap o Fusion Lumos MS.

El análisis de masas fue realizado en modo dependiente de datos (escaneo completo en modo de polaridad positiva con una carga de +2, seguido de fragmentación HCD MS/MS en picos de intensidad mínima 4e4, evitando la fragmentación de aquellos péptidos con cargas +1, +2 y mayores a +8). Los péptidos fueron procesados automáticamente con el software Mass Spec Studio aceptando la presencia de al menos dos masas iónicas reporteras para confirmar la identificación de los péptidos. Los péptidos entrecruzados fueron inicialmente verificados de forma manual.

Se utilizó el software Cytoscape[205] para la visualización final de las redes de interacción flagelar. PyMOL (Schrödinger, LLC; https://pymol.org/) fue utilizado para inspeccionar de forma manual los entrecruzamientos a nivel estructural y el plugin PyXlinkViewer[206] de PyMOL para trazar y medir las distancias de todos los entrecruzamientos en el modelo final.

6.25 – Espectrometría de masas comparativa (MS/MS)

Cantidades similares de filamentos (aproximadamente 20 µg) previamente estimadas por densitometría en gel fueron corridas en SDS-PAGE 12% por aproximadamente 10-15 minutos. Una vez que corrieron 1 cm por el gel separador, la electroforesis fue frenada y las bandas cortadas. Las bandas fueron desteñidas con acetonitrilo 50%/bicarbonato de amonio 50% y

posteriormente las cisteínas proteicas fueron reducidas (con 10 mM DTT durante 1 hora a 56°C) y alquiladas (con 55 mM iodoacetamida 1 hora a temperatura ambiente).

Las muestras fueron secadas con acetonitrilo puro y digeridas en gel con tripsina (Promega), incubando durante la noche a 30°C. Al día siguiente los péptidos trípticos fueron extraídos del gel por 3 incubaciones sucesivas con 0.1% ácido trifluoroacético/50% acetonitrilo. Para el análisis por espectrometría de masas los péptidos fueron desalados en micro-columnas Zip-tip C18 (Millipore), sus eluatos secados al vacío con SpeedVac, y resuspendidos en 0.1% ácido fórmico.

En cada experimento de espectrometría de masas al menos tres réplicas biológicas fueron analizadas en nano-HPLC (UltiMate 3000, Thermo), acoplado a un espectrómetro de masas Q-Exactive Plus (Thermo). Aproximadamente 5 µg de péptidos trípticos fueron inyectados y separados en columnas analíticas C18 (75 µM×50 cm, PepMapTM RSLC, Thermo) a un flujo de 200 nL/min usando un gradiente de 1 a 35% de acetonitrilo en 0.1% ácido fórmico en 90 minutos. El análisis de masas fue realizado en modo dependiente de datos (escaneo completo seguido de MS/MS de los 12 primeros m/z en cada segmento) usando una lista de exclusión dinámica.

Para la identificación proteica se utilizó el PatternLab for Proteomics (Carvalho et al., 2016). Los datos crudos fueron analizados contra una base de datos señuelo que incluía las secuencias de *Leptospira biflexa* serovar Patoc_UP000001847 descargadas de Uniprot y los 127 contaminantes más comunes en espectrometría de masas. Los parámetros de búsqueda incluyeron la oxidación de metioninas como modificación variable, la carbamidometilación como modificación fija y 35 ppm de tolerancia para el precursor m/z.

Los espectros coincidentes fueron filtrados con SEPro (Search Engine Processor) usando un FDR de 1% a nivel de la proteína, un mínimo de dos péptidos identificados por proteína y 10 ppm de tolerancia para la masa del precursor. Para comparar visualmente las proteínas presentes en todas las réplicas se utilizó el módulo Venn Diagram del PatternLab for Proteomics.

6.26 – Modelado de componentes del filamento con AlphaFold2

Durante el transcurso de este trabajo de doctorado, a mediados del año 2021, se abrió el uso público del software bioinformático AlphaFold2 (AF2)[78,79] El desarrollo de esta tecnología permitió la predicción con alta fidelidad de la estructura tridimensional de muchas proteínas, a partir únicamente del input de sus secuencias aminoacídicas.

Mediante el uso de AF2 se volvieron a predecir las estructuras 3D de las cuatro isoformas de FlaB (dando resultados muy similares a los obtenidos antes por modelización por homología), así como la estructura de FlaA2. Además, se predijeron las estructuras de todos los candidatos a proteínas flagelares de vaina obtenidos a partir de los datos de espectrometría de masas y XL-MS, habiendo removido previamente su secuencia señal de exportación, predicha por SignalP.

AF2 es capaz de evaluar la calidad de sus predicciones analizando la cantidad y cobertura de secuencias y brinda un valor pLDDT (*predicted local distance difference test*) promedio –global para toda la proteína– y por residuo. El pLDDT está normalizado en una escala de 0 a 100 que a medida que aumenta indica una mayor fiabilidad del modelo predicho (considerándose relativamente fiable por encima de 70, y altamente fiable >90).

6.27 – Construcción de modelos 3D completos de filamentos flagelares salvajes de *L. biflexa* y *L. interrogans*

Para la construcción de estos modelos se utilizaron las estructuras tridimensionales de protómeros, tanto experimentales (FcpA, FcpB, FlaA1) como estructuras predichas con AF2 (FlaB, FlaA2, y las nuevas proteínas descubiertas en este trabajo: en *L. biflexa* codigos Uniprot B0SPP1, B0SKN8, B0STN7, B0SMK8, B0STF2 y B0SJC6; que corresponden en *L. interrogans* respectivamente a las especies Q72NA0, Q72MH5, Q75FY6, Q72SV9, Q72QN8 y Q72SU9 respectivamente).

Los mapas cryoEM fueron reconstruidos por nuestro colaborador el Dr Charles Sindelar (Universidad de Yale, USA) utilizando técnica de análisis de partícula única. Los volúmenes fueron reconstruidos a partir de micrografías de filamentos purificados de cepas salvajes de *L. biflexa* y *L. interrogans*. Brevemente, las muestras fueron aplicadas en grillas Quantifoil, sometidas a descarga de plasma y congeladas por inmersión usando un Vitrobot Mark IV. Las micrografías fueron recolectadas en microscopios Glacios (200kV, Thermo) y Titan Krios (300kV, Thermo) equipados con detectores K2 y K3 (Gatan) respectivamente, utilizando el programa SerialEM. El procesamiento de imágenes y refinamiento helicoidal se realizó con el programa cryoSPARC v.3.3.1 [207], incluyendo corrección de movimiento, estimación de CTF, trazado de filamentos, clasificación 2D, y refinamiento helicoidal iterativo con correcciones locales y globales de CTF, alcanzando resoluciones globales de 3.6Å (*L. biflexa*) y 4.8Å (*L. interrogans*).

El ajuste en espacio real sobre los mapas cryoEM y refinamiento inicial fue realizado con Coot (versiones 0.96 y 1.1), y posteriormente los modelos fueron refinados alternando ciclos de refinamiento en espacio real con Phenix.real_space_refine, y en espacio recíproco con

Servalcat (versión 0.2.85 en CCPEM). Iterando con los ciclos de refinamiento, el modelo se reconstruyó/corrigió refinando interactivamente en espacio real con Coot. Todos los análisis estructurales, y la preparación de figuras fueron realizados con PyMOL, Coot y ChimeraX.

El procedimiento seguido para construir y refinar los modelos atómicos fue similar tanto para los filamentos de *L. biflexa* como de *L. interrogans*.

Inicialmente, se realizó el ajuste de cuerpo rígido de los modelos de partida en los mapas cryoEM utilizando Chimera v1.15 [208]. El refinamiento adicional se llevó a cabo en Coot v0.9.8.95 y v1.1.15 [161], calculando primero restricciones adicionales intra-cadena (esfera de 5 Å), seguido por refinamiento en espacio real de cada cadena utilizando restricciones Geman-McClure (parámetro α 0.01). Se adoptó un enfoque conservador para todas las estructuras, eliminando cadenas laterales, residuos o bucles cortos en regiones donde la densidad cryoEM no era confiable (densidad muy débil o indetectable a niveles de contorno sensibles). La validación del modelo se realizó continuamente dentro de Coot v1.1, verificando especialmente valores atípicos de Ramachandran y otros parámetros geométricos, corrigiendo mediante refinamiento esférico con rotámeros "*backrub*". Se calcularon restricciones de estructura secundaria a partir de las estructuras cristalográficas o modelos AlphaFold2 iniciales, utilizando phenix.secondary_structure_restraints implementado en Phenix v1.20.1-4487[209]. Estas restricciones adicionales se incluyeron en los primeros ciclos de refinamiento en espacio real con Phenix, utilizando también restricciones de Ramachandran[210]. Posteriormente, se utilizó Servalcat v0.4.72[211] para refinar adicionalmente los modelos atómicos en espacio recíproco contra los medio-mapas ("*half-maps*") cryoEM finales, sin post-procesamiento ninguno ("*sharpening*", etc). Durante el refinamiento con Servalcat se emplearon restricciones estereoquímicas/geométricas externas, basadas en las estructuras iniciales ProSMART v0.860; [212]. La validación final se realizó con MolProbity v4.02-528[213]Todos los análisis estructurales, y la preparación de figuras fueron realizados con PyMOL, Coot y ChimeraX.

VII - BIBLIOGRAFIA

1. Miyata, M.; Robinson, R.C.; Uyeda, T.Q.P.; Fukumori, Y.; Fukushima, S.-I.; Haruta, S.; Homma, M.; Inaba, K.; Ito, M.; Kaito, C.; et al. Tree of Motility - A Proposed History of Motility Systems in the Tree of Life. *Genes Cells* **2020**, *25*, 6–21, doi:10.1111/gtc.12737.
2. Armitage, J.P. Microbial Primer: The Bacterial Flagellum – How Bacteria Swim. *Microbiology (Reading)* **2024**, *170*, 001406, doi:10.1099/mic.0.001406.
3. Nakamura, S. Spirochete Flagella and Motility. *Biomolecules* **2020**, *10*, 550, doi:10.3390/biom10040550.
4. Arkhipov, A.; Freddolino, P.L.; Imada, K.; Namba, K.; Schulten, K. Coarse-Grained Molecular Dynamics Simulations of a Rotating Bacterial Flagellum. *Biophys J* **2006**, *91*, 4589–4597, doi:10.1529/biophysj.106.093443.
5. Zhao, X.; Norris, S.J.; Liu, J. Molecular Architecture of the Bacterial Flagellar Motor in Cells. *Biochemistry* **2014**, *53*, 4323–4333, doi:10.1021/bi500059y.
6. Shibata, S.; Matsunami, H.; Aizawa, S.-I.; Wolf, M. Torque Transmission Mechanism of the Curved Bacterial Flagellar Hook Revealed by Cryo-EM. *Nat Struct Mol Biol* **2019**, *26*, 941–945, doi:10.1038/s41594-019-0301-3.
7. Horstmann, J.A.; Lunelli, M.; Cazzola, H.; Heidemann, J.; Kühne, C.; Steffen, P.; Szefs, S.; Rossi, C.; Lokareddy, R.K.; Wang, C.; et al. Methylation of Salmonella Typhimurium Flagella Promotes Bacterial Adhesion and Host Cell Invasion. *Nat Commun* **2020**, *11*, 2013, doi:10.1038/s41467-020-15738-3.
8. Nedeljković, M.; Postel, S.; Pierce, B.G.; Sundberg, E.J. Molecular Determinants of Filament Capping Proteins Required for the Formation of Functional Flagella in Gram-Negative Bacteria. *Biomolecules* **2021**, *11*, 1397, doi:10.3390/biom11101397.
9. Kaplan, M.; Ghosal, D.; Subramanian, P.; Oikonomou, C.M.; Kjaer, A.; Pirbadian, S.; Ortega, D.R.; Briegel, A.; El-Naggar, M.Y.; Jensen, G.J. The Presence and Absence of Periplasmic Rings in Bacterial Flagellar Motors Correlates with Stator Type. *Elife* **2019**, *8*, e43487, doi:10.7554/eLife.43487.
10. Mondino, S.; Martin, F.S.; Buschiazzo, A. 3D Cryo-EM Imaging of Bacterial Flagella: Novel Structural and Mechanistic Insights into Cell Motility. *Journal of Biological Chemistry* **2022**, *298*, doi:10.1016/j.jbc.2022.102105.
11. Chang, Y.; Moon, K.H.; Zhao, X.; Norris, S.J.; Motaleb, M.A.; Liu, J. Structural Insights into Flagellar Stator–Rotor Interactions. *eLife* **2019**, *8*, e48979, doi:10.7554/eLife.48979.
12. Nedeljković, M.; Sastre, D.E.; Sundberg, E.J. Bacterial Flagellar Filament: A Supramolecular Multifunctional Nanostructure. *Int J Mol Sci* **2021**, *22*, 7521, doi:10.3390/ijms22147521.
13. Altegoer, F.; Schuhmacher, J.; Pausch, P.; Bange, G. From Molecular Evolution to Biobricks and Synthetic Modules: A Lesson by the Bacterial Flagellum. *Biotechnol Genet Eng Rev* **2014**, *30*, 49–64, doi:10.1080/02648725.2014.921500.
14. Carroll, B.L.; Liu, J. Structural Conservation and Adaptation of the Bacterial Flagella Motor. *Biomolecules* **2020**, *10*, 1492, doi:10.3390/biom10111492.
15. Yonekura, K.; Maki-Yonekura, S.; Namba, K. Growth Mechanism of the Bacterial Flagellar Filament. *Res Microbiol* **2002**, *153*, 191–197, doi:10.1016/s0923-2508(02)01308-6.

16. Minamino, T.; Kinoshita, M. Structure, Assembly, and Function of Flagella Responsible for Bacterial Locomotion. *EcoSal Plus* **2023**, *11*, eesp-0011-2023, doi:10.1128/ecosalplus.esp-0011-2023.
17. Picardeau, M.; Brenot, A.; Saint Girons, I. First Evidence for Gene Replacement in Leptospira Spp. Inactivation of L. Biflexa flaB Results in Non-Motile Mutants Deficient in Endoflagella. *Mol Microbiol* **2001**, *40*, 189–199, doi:10.1046/j.1365-2958.2001.02374.x.
18. Partridge, J.D.; Harshey, R.M. Flagellar Protein FliL: A Many-Splendored Thing. *Mol Microbiol* **2024**, doi:10.1111/mmi.15301.
19. Evans, L.D.B.; Hughes, C.; Fraser, G.M. Building a Flagellum Outside the Bacterial Cell. *Trends Microbiol* **2014**, *22*, 566–572, doi:10.1016/j.tim.2014.05.009.
20. Yamaguchi, T.; Makino, F.; Miyata, T.; Minamino, T.; Kato, T.; Namba, K. Structure of the Molecular Bushing of the Bacterial Flagellar Motor. *Nat Commun* **2021**, *12*, 4469, doi:10.1038/s41467-021-24715-3.
21. Terashima, H.; Hori, K.; Ihara, K.; Homma, M.; Kojima, S. Mutations in the Stator Protein PomA Affect Switching of Rotational Direction in Bacterial Flagellar Motor. *Sci Rep* **2022**, *12*, 2979, doi:10.1038/s41598-022-06947-5.
22. Kato, T.; Makino, F.; Miyata, T.; Horváth, P.; Namba, K. Structure of the Native Supercoiled Flagellar Hook as a Universal Joint. *Nat Commun* **2019**, *10*, 5295, doi:10.1038/s41467-019-13252-9.
23. Armitage, J. Twists and Turns: 40 Years of Investigating How and Why Bacteria Swim. *Microbiology (Reading, England)* **2024**, *170*, doi:10.1099/mic.0.001432.
24. Manson, M.D. Dynamic Motors for Bacterial Flagella. *Proceedings of the National Academy of Sciences* **2010**, *107*, 11151–11152, doi:10.1073/pnas.1006365107.
25. Gambelli, L.; Isupov, M.N.; Conners, R.; McLaren, M.; Bellack, A.; Gold, V.; Rachel, R.; Daum, B. An Archaellum Filament Composed of Two Alternating Subunits. *Nat Commun* **2022**, *13*, 710, doi:10.1038/s41467-022-28337-1.
26. Kreutzberger, M.A.B.; Ewing, C.; Poly, F.; Wang, F.; Egelman, E.H. Atomic Structure of the Campylobacter Jejuni Flagellar Filament Reveals How ε Proteobacteria Escaped Toll-like Receptor 5 Surveillance. *Proc Natl Acad Sci U S A* **2020**, *117*, 16985–16991, doi:10.1073/pnas.2010996117.
27. Gao, B.; Lara-Tejero, M.; Lefebre, M.; Goodman, A.L.; Galán, J.E. Novel Components of the Flagellar System in Epsilonproteobacteria. *mBio* **2014**, *5*, e01349-01314, doi:10.1128/mBio.01349-14.
28. Grognot, M.; Taute, K.M. More than Propellers: How Flagella Shape Bacterial Motility Behaviors. *Curr Opin Microbiol* **2021**, *61*, 73–81, doi:10.1016/j.mib.2021.02.005.
29. Carroll, B.L.; Nishikino, T.; Guo, W.; Zhu, S.; Kojima, S.; Homma, M.; Liu, J. The Flagellar Motor of Vibrio Alginolyticus Undergoes Major Structural Remodeling during Rotational Switching. *Elife* **2020**, *9*, e61446, doi:10.7554/eLife.61446.
30. Xie, L.; Altindal, T.; Chattopadhyay, S.; Wu, X.-L. From the Cover: Bacterial Flagellum as a Propeller and as a Rudder for Efficient Chemotaxis. *Proc Natl Acad Sci U S A* **2011**, *108*, 2246–2251, doi:10.1073/pnas.1011953108.
31. Namba, K.; Yamashita, I.; Vonderviszt, F. Structure of the Core and Central Channel of Bacterial Flagella. *Nature* **1989**, *342*, 648–654, doi:10.1038/342648a0.
32. Samatey, F.A.; Imada, K.; Nagashima, S.; Vonderviszt, F.; Kumasaka, T.; Yamamoto, M.; Namba, K. Structure of the Bacterial Flagellar Protofilament and Implications for a Switch for Supercoiling. *Nature* **2001**, *410*, 331–337, doi:10.1038/35066504.

33. Yonekura, K.; Maki-Yonekura, S.; Namba, K. Complete Atomic Model of the Bacterial Flagellar Filament by Electron Cryomicroscopy. *Nature* **2003**, *424*, 643–650, doi:10.1038/nature01830.
34. Kreutzberger, M.A.B.; Sonani, R.R.; Liu, J.; Chatterjee, S.; Wang, F.; Sebastian, A.L.; Biswas, P.; Ewing, C.; Zheng, W.; Poly, F.; et al. Convergent Evolution in the Supercoiling of Prokaryotic Flagellar Filaments. *Cell* **2022**, *185*, 3487-3500.e14, doi:10.1016/j.cell.2022.08.009.
35. Chen, S.; Beeby, M.; Murphy, G.E.; Leadbetter, J.R.; Hendrixson, D.R.; Briegel, A.; Li, Z.; Shi, J.; Tocheva, E.I.; Müller, A.; et al. Structural Diversity of Bacterial Flagellar Motors. *EMBO J* **2011**, *30*, 2972–2981, doi:10.1038/emboj.2011.186.
36. Kurniyati, K.; Chang, Y.; Liu, J.; Li, C. Transcriptional and Functional Characterizations of Multiple Flagellin Genes in Spirochetes. *Mol Microbiol* **2022**, *118*, 175–190, doi:10.1111/mmi.14959.
37. Partridge, J.D.; Dufour, Y.; Hwang, Y.; Harshey, R.M. Flagellar Motor Remodeling during Swarming Requires FliL. *Mol Microbiol* **2023**, *120*, 670–683, doi:10.1111/mmi.15148.
38. McCarter, L.L. Dual Flagellar Systems Enable Motility under Different Circumstances. *J Mol Microbiol Biotechnol* **2004**, *7*, 18–29, doi:10.1159/000077866.
39. Beeby, M.; Ribardo, D.A.; Brennan, C.A.; Ruby, E.G.; Jensen, G.J.; Hendrixson, D.R. Diverse High-Torque Bacterial Flagellar Motors Assemble Wider Stator Rings Using a Conserved Protein Scaffold. *Proc Natl Acad Sci U S A* **2016**, *113*, E1917-1926, doi:10.1073/pnas.1518952113.
40. Nord, A.L.; Biquet-Bisquert, A.; Abkarian, M.; Pigaglio, T.; Seduk, F.; Magalon, A.; Pedaci, F. Dynamic Stiffening of the Flagellar Hook. *Nat Commun* **2022**, *13*, 2925, doi:10.1038/s41467-022-30295-7.
41. Son, K.; Guasto, J.S.; Stocker, R. Bacteria Can Exploit a Flagellar Buckling Instability to Change Direction. *Nature Phys* **2013**, *9*, 494–498, doi:10.1038/nphys2676.
42. Matsunami, H.; Barker, C.S.; Yoon, Y.-H.; Wolf, M.; Samatey, F.A. Complete Structure of the Bacterial Flagellar Hook Reveals Extensive Set of Stabilizing Interactions. *Nat Commun* **2016**, *7*, 13425, doi:10.1038/ncomms13425.
43. Miller, M.R.; Miller, K.A.; Bian, J.; James, M.E.; Zhang, S.; Lynch, M.; Callery, P.S.; Hettick, J.M.; Cockburn, A.; Liu, J.; et al. Spirochete Flagella Hook Protein Self-Catalyze a Lysinoalanine Covalent Cross-Link for Motility. *Nat Microbiol* **2016**, *1*, 16134, doi:10.1038/nmicrobiol.2016.134.
44. Charon, N.W.; Goldstein, S.F. Genetics of Motility and Chemotaxis of a Fascinating Group of Bacteria: The Spirochetes. *Annu Rev Genet* **2002**, *36*, 47–73, doi:10.1146/annurev.genet.36.041602.134359.
45. Charon, N.W.; Cockburn, A.; Li, C.; Liu, J.; Miller, K.A.; Miller, M.R.; Motaleb, Md.; Wolgemuth, C.W. The Unique Paradigm of Spirochete Motility and Chemotaxis. *Annu Rev Microbiol* **2012**, *66*, 349–370, doi:10.1146/annurev-micro-092611-150145.
46. Kühn, M.J.; Schmidt, F.K.; Farthing, N.E.; Rossmann, F.M.; Helm, B.; Wilson, L.G.; Eckhardt, B.; Thormann, K.M. Spatial Arrangement of Several Flagellins within Bacterial Flagella Improves Motility in Different Environments. *Nat Commun* **2018**, *9*, 5369, doi:10.1038/s41467-018-07802-w.
47. Faulds-Pain, A.; Birchall, C.; Aldridge, C.; Smith, W.D.; Grimaldi, G.; Nakamura, S.; Miyata, T.; Gray, J.; Li, G.; Tang, J.X.; et al. Flagellin Redundancy in Caulobacter Crescentus and Its Implications for Flagellar Filament Assembly. *Journal of Bacteriology* **2011**, *193*, 2695–2707, doi:10.1128/jb.01172-10.

48. Yamaguchi, T.; Toma, S.; Terahara, N.; Miyata, T.; Ashihara, M.; Minamino, T.; Namba, K.; Kato, T. Structural and Functional Comparison of Salmonella Flagellar Filaments Composed of FljB and FliC. *Biomolecules* **2020**, *10*, 246, doi:10.3390/biom10020246.
49. Beatson, S.A.; Minamino, T.; Pallen, M.J. Variation in Bacterial Flagellins: From Sequence to Structure. *Trends Microbiol* **2006**, *14*, 151–155, doi:10.1016/j.tim.2006.02.008.
50. Wang, F.; Burrage, A.M.; Postel, S.; Clark, R.E.; Orlova, A.; Sundberg, E.J.; Kearns, D.B.; Egelman, E.H. A Structural Model of Flagellar Filament Switching across Multiple Bacterial Species. *Nat Commun* **2017**, *8*, 960, doi:10.1038/s41467-017-01075-5.
51. Hu, D.; Reeves, P.R. The Remarkable Dual-Level Diversity of Prokaryotic Flagellins. *mSystems* **2020**, *5*, e00705-19, doi:10.1128/mSystems.00705-19.
52. Fujii, T.; Kato, T.; Hiraoka, K.D.; Miyata, T.; Minamino, T.; Chevance, F.F.V.; Hughes, K.T.; Namba, K. Identical Folds Used for Distinct Mechanical Functions of the Bacterial Flagellar Rod and Hook. *Nat Commun* **2017**, *8*, 14276, doi:10.1038/ncomms14276.
53. Eckhard, U.; Bandukwala, H.; Mansfield, M.J.; Marino, G.; Cheng, J.; Wallace, I.; Holyoak, T.; Charles, T.C.; Austin, J.; Overall, C.M.; et al. Discovery of a Proteolytic Flagellin Family in Diverse Bacterial Phyla That Assembles Enzymatically Active Flagella. *Nat Commun* **2017**, *8*, 521, doi:10.1038/s41467-017-00599-0.
54. Eckhard, U.; Blöchl, C.; Jenkins, B.G.L.; Mansfield, M.J.; Huber, C.G.; Doxey, A.C.; Brandstetter, H. Identification and Characterization of the Proteolytic Flagellin from the Common Freshwater Bacterium Hylemonella Gracilis. *Sci Rep* **2020**, *10*, 19052, doi:10.1038/s41598-020-76010-8.
55. Thibault, P.; Logan, S.M.; Kelly, J.F.; Brisson, J.-R.; Ewing, C.P.; Trust, T.J.; Guerry, P. Identification of the Carbohydrate Moieties and Glycosylation Motifs in Campylobacter Jejuni Flagellin *. *Journal of Biological Chemistry* **2001**, *276*, 34862–34870, doi:10.1074/jbc.M104529200.
56. Khemiri, A.; Naudin, B.; Franck, X.; Song, P.C.T.; Jouenne, T.; Cosette, P. N-Glycosidase Treatment with 18O Labeling and de Novo Sequencing Argues for Flagellin FliC Glycopolymorphism in Pseudomonas Aeruginosa. *Anal Bioanal Chem* **2013**, *405*, 9835–9842, doi:10.1007/s00216-013-7424-x.
57. Sun, L.; Jin, M.; Ding, W.; Yuan, J.; Kelly, J.; Gao, H. Posttranslational Modification of Flagellin FlaB in Shewanella Oneidensis. *J Bacteriol* **2013**, *195*, 2550–2561, doi:10.1128/JB.00015-13.
58. Schirm, M.; Kalmokoff, M.; Aubry, A.; Thibault, P.; Sandoz, M.; Logan, S.M. Flagellin from Listeria Monocytogenes Is Glycosylated with Beta-O-Linked N-Acetylglucosamine. *J Bacteriol* **2004**, *186*, 6721–6727, doi:10.1128/JB.186.20.6721-6727.2004.
59. Twine, S.M.; Reid, C.W.; Aubry, A.; McMullin, D.R.; Fulton, K.M.; Austin, J.; Logan, S.M. Motility and Flagellar Glycosylation in Clostridium Difficile. *J Bacteriol* **2009**, *191*, 7050–7062, doi:10.1128/JB.00861-09.
60. Arora, S.K.; Bangera, M.; Lory, S.; Ramphal, R. A Genomic Island in Pseudomonas Aeruginosa Carries the Determinants of Flagellin Glycosylation. *Proc Natl Acad Sci U S A* **2001**, *98*, 9342–9347, doi:10.1073/pnas.161249198.
61. Kurniyati, K.; Kelly, J.F.; Vinogradov, E.; Robotham, A.; Tu, Y.; Wang, J.; Liu, J.; Logan, S.M.; Li, C. A Novel Glycan Modifies the Flagellar Filament Proteins of the Oral Bacterium Treponema Denticola. *Mol Microbiol* **2017**, *103*, 67–85, doi:10.1111/mmi.13544.

62. Shi, Y. A Glimpse of Structural Biology through X-Ray Crystallography. *Cell* **2014**, *159*, 995–1014, doi:10.1016/j.cell.2014.10.051.
63. Wlodawer, A.; Minor, W.; Dauter, Z.; Jaskolski, M. Protein Crystallography for Non-Crystallographers, or How to Get the Best (but Not More) from Published Macromolecular Structures. *FEBS J* **2008**, *275*, 1–21, doi:10.1111/j.1742-4658.2007.06178.x.
64. Masrati, G.; Landau, M.; Ben-Tal, N.; Lupas, A.; Kosloff, M.; Kosinski, J. Integrative Structural Biology in the Era of Accurate Structure Prediction. *J Mol Biol* **2021**, *433*, 167127, doi:10.1016/j.jmb.2021.167127.
65. Forster, M.J. Molecular Modelling in Structural Biology. *Micron* **2002**, *33*, 365–384, doi:10.1016/S0968-4328(01)00035-X.
66. Assaiya, A.; Burada, A.P.; Dhingra, S.; Kumar, J. An Overview of the Recent Advances in Cryo-Electron Microscopy for Life Sciences. *Emerg Top Life Sci* **2021**, *5*, 151–168, doi:10.1042/ETLS20200295.
67. White, J.B.R.; Maskell, D.P.; Howe, A.; Harrow, M.; Clare, D.K.; Siebert, C.A.; Hesketh, E.L.; Thompson, R.F. Single Particle Cryo-Electron Microscopy: From Sample to Structure. *J Vis Exp* **2021**, doi:10.3791/62415.
68. Al-Otaibi, N.S.; Taylor, A.J.; Farrell, D.P.; Tzokov, S.B.; DiMaio, F.; Kelly, D.J.; Bergeron, J.R.C. The Cryo-EM Structure of the Bacterial Flagellum Cap Complex Suggests a Molecular Mechanism for Filament Elongation. *Nat Commun* **2020**, *11*, 3210, doi:10.1038/s41467-020-16981-4.
69. Kawamoto, A.; Miyata, T.; Makino, F.; Kinoshita, M.; Minamino, T.; Imada, K.; Kato, T.; Namba, K. Native Flagellar MS Ring Is Formed by 34 Subunits with 23-Fold and 11-Fold Subsymmetries. *Nat Commun* **2021**, *12*, 4223, doi:10.1038/s41467-021-24507-9.
70. Raddi, G.; Morado, D.R.; Yan, J.; Haake, D.A.; Yang, X.F.; Liu, J. Three-Dimensional Structures of Pathogenic and Saprophytic Leptospira Species Revealed by Cryo-Electron Tomography. *J Bacteriol* **2012**, *194*, 1299–1306, doi:10.1128/JB.06474-11.
71. Liu, J.; Howell, J.K.; Bradley, S.D.; Zheng, Y.; Zhou, Z.H.; Norris, S.J. Cellular Architecture of *Treponema Pallidum*: Novel Flagellum, Periplasmic Cone, and Cell Envelope as Revealed by Cryo Electron Tomography. *Journal of Molecular Biology* **2010**, *403*, 546–561, doi:10.1016/j.jmb.2010.09.020.
72. Zhao, X.; Zhang, K.; Boquoi, T.; Hu, B.; Motaleb, M.A.; Miller, K.A.; James, M.E.; Charon, N.W.; Manson, M.D.; Norris, S.J.; et al. Cryoelectron Tomography Reveals the Sequential Assembly of Bacterial Flagella in Borrelia Burgdorferi. *Proc Natl Acad Sci U S A* **2013**, *110*, 14390–14395, doi:10.1073/pnas.1308306110.
73. Chang, Y.; Xu, H.; Motaleb, M.A.; Liu, J. Characterization of the Flagellar Collar Reveals Structural Plasticity Essential for Spirochete Motility. *mBio* **2021**, *12*, e0249421, doi:10.1128/mBio.02494-21.
74. Nauman, R.K.; Holt, S.C.; Cox, C.D. Purification, Ultrastructure, and Composition of Axial Filaments from Leptospira. *Journal of Bacteriology* **1969**, *98*, 264–280, doi:10.1128/jb.98.1.264-280.1969.
75. Xiao, J.G.; Dai, B.M. [Purification and SDS-page analysis of axial filaments from leptospires]. *Hua Xi Yi Ke Da Xue Xue Bao* **1989**, *20*, 1–4.
76. Sasaki, Y.; Kawamoto, A.; Tahara, H.; Kasuga, K.; Sato, R.; Ohnishi, M.; Nakamura, S.; Koizumi, N. Leptospiral Flagellar Sheath Protein FcpA Interacts with FlaA2 and FlaB1 in Leptospira Biflexa. *PLoS One* **2018**, *13*, e0194923, doi:10.1371/journal.pone.0194923.

77. Lambert, A.; Picardeau, M.; Haake, D.A.; Sermswan, R.W.; Srikram, A.; Adler, B.; Murray, G.A. FlaA Proteins in Leptospira Interrogans Are Essential for Motility and Virulence but Are Not Required for Formation of the Flagellum Sheath. *Infect Immun* **2012**, *80*, 2019–2025, doi:10.1128/IAI.00131-12.
78. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596*, 583–589, doi:10.1038/s41586-021-03819-2.
79. Mirdita, M.; Schütze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M. ColabFold: Making Protein Folding Accessible to All. *Nat Methods* **2022**, *19*, 679–682, doi:10.1038/s41592-022-01488-1.
80. Motaleb, M.A.; Corum, L.; Bono, J.L.; Elias, A.F.; Rosa, P.; Samuels, D.S.; Charon, N.W. Borrelia Burgdorferi Periplasmic Flagella Have Both Skeletal and Motility Functions. *Proc Natl Acad Sci U S A* **2000**, *97*, 10899–10904, doi:10.1073/pnas.200221797.
81. Goldstein, S.F.; Buttle, K.F.; Charon, N.W. Structural Analysis of the Leptospiraceae and Borrelia Burgdorferi by High-Voltage Electron Microscopy. *J Bacteriol* **1996**, *178*, 6539–6545, doi:10.1128/jb.178.22.6539-6545.1996.
82. Li, C.; Corum, L.; Morgan, D.; Rosey, E.L.; Stanton, T.B.; Charon, N.W. The Spirochete FlaA Periplasmic Flagellar Sheath Protein Impacts Flagellar Helicity. *J Bacteriol* **2000**, *182*, 6698–6706, doi:10.1128/JB.182.23.6698-6706.2000.
83. Zückert, W.R. Protein Secretion in Spirochetes. *Microbiology Spectrum* **2019**, *7*, 10.1128/microbiolspec.psib-0026–2019, doi:10.1128/microbiolspec.psib-0026-2019.
84. Trueba, G.A.; Bolin, C.A.; Zuerner, R.L. Characterization of the Periplasmic Flagellum Proteins of Leptospira Interrogans. *J Bacteriol* **1992**, *174*, 4761–4768.
85. Charon, N.W.; Greenberg, E.P.; Koopman, M.B.; Limberger, R.J. Spirochete Chemotaxis, Motility, and the Structure of the Spirochetal Periplasmic Flagella. *Res Microbiol* **1992**, *143*, 597–603, doi:10.1016/0923-2508(92)90117-7.
86. Wolgemuth, C.W.; Charon, N.W.; Goldstein, S.F.; Goldstein, R.E. The Flagellar Cytoskeleton of the Spirochetes. *Microb Physiol* **2006**, *11*, 221–227, doi:10.1159/000094056.
87. Li, C.; Wolgemuth, C.W.; Marko, M.; Morgan, D.G.; Charon, N.W. Genetic Analysis of Spirochete Flagellin Proteins and Their Involvement in Motility, Filament Assembly, and Flagellar Morphology. *J Bacteriol* **2008**, *190*, 5607–5615, doi:10.1128/JB.00319-08.
88. Xu, H.; Hu, B.; Flesher, D.A.; Liu, J.; Motaleb, M.A. BB0259 Encompasses a Peptidoglycan Lytic Enzyme Function for Proper Assembly of Periplasmic Flagella in Borrelia Burgdorferi. *Front. Microbiol.* **2021**, *12*, doi:10.3389/fmicb.2021.692707.
89. Li, C.; Motaleb, A.; Sal, M.; Goldstein, S.F.; Charon, N.W. Spirochete Periplasmic Flagella and Motility. *J Mol Microbiol Biotechnol* **2000**, *2*, 345–354.
90. Takabe, K.; Kawamoto, A.; Tahara, H.; Kudo, S.; Nakamura, S. Implications of Coordinated Cell-Body Rotations for Leptospira Motility. *Biochem Biophys Res Commun* **2017**, *491*, 1040–1046, doi:10.1016/j.bbrc.2017.08.007.
91. Halte, M.; Andrianova, E.P.; Goosmann, C.; Chevance, F.F.V.; Hughes, K.T.; Zhulin, I.B.; Erhardt, M. FlhE Functions as a Chaperone to Prevent Formation of Periplasmic Flagella in Gram-Negative Bacteria. *bioRxiv* **2024**, 2024.03.11.584431, doi:10.1101/2024.03.11.584431.
92. Kan, W.; Wolgemuth, C.W. The Shape and Dynamics of the *Leptospiraceae*. *Biophysical Journal* **2007**, *93*, 54–61, doi:10.1529/biophysj.106.103143.

93. Dombrowski, C.; Kan, W.; Motaleb, Md.A.; Charon, N.W.; Goldstein, R.E.; Wolgemuth, C.W. The Elastic Basis for the Shape of Borrelia Burgdorferi. *Biophys J* **2009**, *96*, 4409–4417, doi:10.1016/j.bpj.2009.02.066.
94. Slamti, L.; de Pedro, M.A.; Guichet, E.; Picardeau, M. Deciphering Morphological Determinants of the Helix-Shaped Leptospira ▿. *J Bacteriol* **2011**, *193*, 6266–6275, doi:10.1128/JB.05695-11.
95. Qin, Z.; Tu, J.; Lin, T.; Norris, S.J.; Li, C.; Motaleb, M.A.; Liu, J. Cryo-Electron Tomography of Periplasmic Flagella in Borrelia Burgdorferi Reveals a Distinct Cytoplasmic ATPase Complex. *PLoS Biol* **2018**, *16*, e3000050, doi:10.1371/journal.pbio.3000050.
96. Ge, Y.; Li, C.; Corum, L.; Slaughter, C.A.; Charon, N.W. Structure and Expression of the FlaA Periplasmic Flagellar Protein of Borrelia Burgdorferi. *J Bacteriol* **1998**, *180*, 2418–2425, doi:10.1128/JB.180.9.2418-2425.1998.
97. Li, C.; Sal, M.; Marko, M.; Charon, N.W. Differential Regulation of the Multiple Flagellins in Spirochetes. *J Bacteriol* **2010**, *192*, 2596–2603, doi:10.1128/JB.01502-09.
98. Motaleb, M.A.; Sal, M.S.; Charon, N.W. The Decrease in FlaA Observed in a flaB Mutant of Borrelia Burgdorferi Occurs Posttranscriptionally. *J Bacteriol* **2004**, *186*, 3703–3711, doi:10.1128/JB.186.12.3703-3711.2004.
99. Mongodin, E.F.; Casjens, S.R.; Bruno, J.F.; Xu, Y.; Drabek, E.F.; Riley, D.R.; Cantarel, B.L.; Pagan, P.E.; Hernandez, Y.A.; Vargas, L.C.; et al. Inter- and Intra-Specific Pan-Genomes of Borrelia Burgdorferi Sensu Lato: Genome Stability and Adaptive Radiation. *BMC Genomics* **2013**, *14*, 693, doi:10.1186/1471-2164-14-693.
100. Nakamura, S.; Leshansky, A.; Magariyama, Y.; Namba, K.; Kudo, S. Direct Measurement of Helical Cell Motion of the Spirochete *Leptospira*. *Biophysical Journal* **2014**, *106*, 47–54, doi:10.1016/j.bpj.2013.11.1118.
101. Charon, N.W.; Daughtry, G.R.; McCuskey, R.S.; Franz, G.N. Microcinematographic Analysis of Tethered Leptospira Illini. *J Bacteriol* **1984**, *160*, 1067–1073, doi:10.1128/jb.160.3.1067-1073.1984.
102. Goldstein, S.F.; Charon, N.W. Multiple-Exposure Photographic Analysis of a Motile Spirochete. *Proc Natl Acad Sci U S A* **1990**, *87*, 4895–4899, doi:10.1073/pnas.87.13.4895.
103. Bulach, D.M.; Seemann, T.; Zuerner, R.L.; Adler, B. The Organization of Leptospira at a Genomic Level. In *Bacterial Genomes and Infectious Diseases*; Chan, V.L., Sherman, P.M., Bourke, B., Eds.; Humana Press: Totowa, NJ, 2006; pp. 109–123 ISBN 978-1-58829-496-8.
104. Nakamura, S. Motility of the Zoonotic Spirochete Leptospira: Insight into Association with Pathogenicity. *Int J Mol Sci* **2022**, *23*, 1859, doi:10.3390/ijms23031859.
105. Wunder, E.A.; Figueira, C.P.; Benaroudj, N.; Hu, B.; Tong, B.A.; Trajtenberg, F.; Liu, J.; Reis, M.G.; Charon, N.W.; Buschiazzo, A.; et al. A Novel Flagellar Sheath Protein, FcpA, Determines Filament Coiling, Translational Motility and Virulence for the Leptospira Spirochete. *Mol Microbiol* **2016**, *101*, 457–470, doi:10.1111/mmi.13403.
106. Berg, H.C.; Turner, L. Movement of Microorganisms in Viscous Environments. *Nature* **1979**, *278*, 349–351, doi:10.1038/278349a0.
107. Goldstein, S.F.; Charon, N.W. Motility of the Spirochete Leptospira. *Cell Motil Cytoskeleton* **1988**, *9*, 101–110, doi:10.1002/cm.970090202.
108. Bromley, D.B.; Charon, N.W. Axial Filament Involvement in the Motility of Leptospira Interrogans. *J Bacteriol* **1979**, *137*, 1406–1412, doi:10.1128/jb.137.3.1406-1412.1979.

109. de Oliveira, N.R.; Santos, F.D.S.; dos Santos, V.A.C.; Maia, M.A.C.; Oliveira, T.L.; Dellagostin, O.A. Challenges and Strategies for Developing Recombinant Vaccines against Leptospirosis: Role of Expression Platforms and Adjuvants in Achieving Protective Efficacy. *Pathogens* **2023**, *12*, 787, doi:10.3390/pathogens12060787.
110. Xue, F.; Dong, H.; Wu, J.; Wu, Z.; Hu, W.; Sun, A.; Troxell, B.; Yang, X.F.; Yan, J. Transcriptional Responses of Leptospira Interrogans to Host Innate Immunity: Significant Changes in Metabolism, Oxygen Tolerance, and Outer Membrane. *PLOS Neglected Tropical Diseases* **2010**, *4*, e857, doi:10.1371/journal.pntd.0000857.
111. Adler, B.; de la Peña Moctezuma, A. Leptospira and Leptospirosis. *Vet Microbiol* **2010**, *140*, 287–296, doi:10.1016/j.vetmic.2009.03.012.
112. Haake, D.A.; Matsunaga, J. Leptospira: A Spirochete with a Hybrid Outer Membrane. *Mol Microbiol* **2010**, *77*, 805–814, doi:10.1111/j.1365-2958.2010.07262.x.
113. Dellagostin, O.A.; Grassmann, A.A.; Rizzi, C.; Schuch, R.A.; Jorge, S.; Oliveira, T.L.; McBride, A.J.A.; Hartwig, D.D. Reverse Vaccinology: An Approach for Identifying Leptospiral Vaccine Candidates. *Int J Mol Sci* **2017**, *18*, 158, doi:10.3390/ijms18010158.
114. Vincent, A.T.; Schiettekatte, O.; Goarant, C.; Neela, V.K.; Bernet, E.; Thibeaux, R.; Ismail, N.; Khalid, M.K.N.M.; Amran, F.; Masuzawa, T.; et al. Revisiting the Taxonomy and Evolution of Pathogenicity of the Genus Leptospira through the Prism of Genomics. *PLOS Neglected Tropical Diseases* **2019**, *13*, e0007270, doi:10.1371/journal.pntd.0007270.
115. Ko, A.I.; Goarant, C.; Picardeau, M. Leptospira: The Dawn of the Molecular Genetics Era for an Emerging Zoonotic Pathogen. *Nat Rev Microbiol* **2009**, *7*, 736–747, doi:10.1038/nrmicro2208.
116. Plank, R.; Dean, D. Overview of the Epidemiology, Microbiology, and Pathogenesis of Leptospira Spp. in Humans. *Microbes Infect* **2000**, *2*, 1265–1276, doi:10.1016/s1286-4579(00)01280-6.
117. Cullen, P.A.; Haake, D.A.; Adler, B. Outer Membrane Proteins of Pathogenic Spirochetes. *FEMS Microbiol Rev* **2004**, *28*, 291–318, doi:10.1016/j.femsre.2003.10.004.
118. Xue, F.; Zhao, X.; Yang, Y.; Zhao, J.; Yang, Y.; Cao, Y.; Hong, C.; Liu, Y.; Sun, L.; Huang, M.; et al. Responses of Murine and Human Macrophages to Leptospiral Infection: A Study Using Comparative Array Analysis. *PLoS Negl Trop Dis* **2013**, *7*, e2477, doi:10.1371/journal.pntd.0002477.
119. Yanagihara, Y.; Villanueva, S.Y.A.M.; Nomura, N.; Ohno, M.; Sekiya, T.; Handabile, C.; Shingai, M.; Higashi, H.; Yoshida, S.; Masuzawa, T.; et al. Leptospira Is an Environmental Bacterium That Grows in Waterlogged Soil. *Microbiology Spectrum* **2022**, *10*, e02157-21, doi:10.1128/spectrum.02157-21.
120. Eshghi, A.; Cullen, P.A.; Cowen, L.; Zuerner, R.L.; Cameron, C.E. Global Proteome Analysis of Leptospira Interrogans. *J Proteome Res* **2009**, *8*, 4564–4578, doi:10.1021/pr9004597.
121. Lo, M.; Cordwell, S.J.; Bulach, D.M.; Adler, B. Comparative Transcriptional and Translational Analysis of Leptospiral Outer Membrane Protein Expression in Response to Temperature. *PLoS Negl Trop Dis* **2009**, *3*, e560, doi:10.1371/journal.pntd.0000560.
122. Leptospirosis humana: Guía para el diagnóstico, vigilancia y control - OPS/OMS | Organización Panamericana de la Salud Available online: https://www.paho.org/es/documentos/leptospirosis-humana-guia-para-diagnostico-vigilancia-control (accessed on 3 August 2024).

123. Fraga, T.R.; Carvalho, E.; Isaac, L.; Barbosa, A.S. *Leptospira* and Leptospirosis. In *Molecular Medical Microbiology (Third Edition)*; Tang, Y.-W., Hindiyeh, M.Y., Liu, D., Sails, A., Spearman, P., Zhang, J.-R., Eds.; Academic Press, 2024; pp. 1849–1871 ISBN 978-0-12-818619-0.
124. Barazzone, G.C.; Teixeira, A.F.; Azevedo, B.O.P.; Damiano, D.K.; Oliveira, M.P.; Nascimento, A.L.T.O.; Lopes, A.P.Y. Revisiting the Development of Vaccines Against Pathogenic Leptospira: Innovative Approaches, Present Challenges, and Future Perspectives. *Front. Immunol.* **2022**, *12*, doi:10.3389/fimmu.2021.760291.
125. Costa, F.; Hagan, J.E.; Calcagno, J.; Kane, M.; Torgerson, P.; Martinez-Silveira, M.S.; Stein, C.; Abela-Ridder, B.; Ko, A.I. Global Morbidity and Mortality of Leptospirosis: A Systematic Review. *PLoS Negl Trop Dis* **2015**, *9*, e0003898, doi:10.1371/journal.pntd.0003898.
126. Azevedo, I.R.; Amamura, T.A.; Isaac, L. Human Leptospirosis: In Search for a Better Vaccine. *Scandinavian Journal of Immunology* **2023**, *98*, e13316, doi:10.1111/sji.13316.
127. Barbosa, A.S.; Patarakul, K.; Isaac, L. Editorial: Leptospirosis: Pathogenesis, Clinical and Epidemiological Aspects. *Front. Cell. Infect. Microbiol.* **2023**, *13*, doi:10.3389/fcimb.2023.1210178.
128. Bradley, E.A.; Lockaby, G. Leptospirosis and the Environment: A Review and Future Directions. *Pathogens* **2023**, *12*, 1167, doi:10.3390/pathogens12091167.
129. Beauté, J.; Innocenti, F.; Aristodimou, A.; Špačková, M.; Eves, C.; Kerbo, N.; Rimhanen-Finne, R.; Picardeau, M.; Faber, M.; Dougas, G.; et al. Epidemiology of Reported Cases of Leptospirosis in the EU/EEA, 2010 to 2021. *Eurosurveillance* **2024**, *29*, 2300266, doi:10.2807/1560-7917.ES.2024.29.7.2300266.
130. Zarantonelli, L.; Suanes, A.; Meny, P.; Buroni, F.; Nieves, C.; Salaberry, X.; Briano, C.; Ashfield, N.; Da Silva Silveira, C.; Dutra, F.; et al. Isolation of Pathogenic Leptospira Strains from Naturally Infected Cattle in Uruguay Reveals High Serovar Diversity, and Uncovers a Relevant Risk for Human Leptospirosis. *PLoS Negl Trop Dis* **2018**, *12*, e0006694, doi:10.1371/journal.pntd.0006694.
131. Dorsch, M.A.; Francia, M.E.; Tana, L.R.; González, F.C.; Cabrera, A.; Calleros, L.; Sanguinetti, M.; Barcellos, M.; Zarantonelli, L.; Ciuffo, C.; et al. Diagnostic Investigation of 100 Cases of Abortion in Sheep in Uruguay: 2015-2021. *Front Vet Sci* **2022**, *9*, 904786, doi:10.3389/fvets.2022.904786.
132. Samrot, A.V.; Sean, T.C.; Bhavya, K.S.; Sahithya, C.S.; Chan-drasekaran, S.; Palanisamy, R.; Robinson, E.R.; Subbiah, S.K.; Mok, P.L. Leptospiral Infection, Pathogenesis and Its Diagnosis—A Review. *Pathogens* **2021**, *10*, 145, doi:10.3390/pathogens10020145.
133. Nieves, C.; Ferrés, I.; Díaz-Viraqué, F.; Buschiazzo, A.; Zarantonelli, L.; Iraola, G. Draft Genome Sequences of 40 Pathogenic Leptospira Strains Isolated from Cattle in Uruguay. *Microbiol Resour Announc* **2019**, *8*, e00893-19, doi:10.1128/MRA.00893-19.
134. Malmström, J.; Beck, M.; Schmidt, A.; Lange, V.; Deutsch, E.W.; Aebersold, R. Proteome-Wide Cellular Protein Concentrations of the Human Pathogen Leptospira Interrogans. *Nature* **2009**, *460*, 762–765, doi:10.1038/nature08184.
135. Fouts, D.E.; Matthias, M.A.; Adhikarla, H.; Adler, B.; Amorim-Santos, L.; Berg, D.E.; Bulach, D.; Buschiazzo, A.; Chang, Y.-F.; Galloway, R.L.; et al. What Makes a Bacterial Species Pathogenic?:Comparative Genomic Analysis of the Genus Leptospira. *PLoS Negl Trop Dis* **2016**, *10*, e0004403, doi:10.1371/journal.pntd.0004403.

136. San Martin, F.; Mechaly, A.E.; Larrieux, N.; Wunder, E.A.; Ko, A.I.; Picardeau, M.; Trajtenberg, F.; Buschiazzo, A. Crystallization of FcpA from Leptospira, a Novel Flagellar Protein That Is Essential for Pathogenesis. *Acta Crystallogr F Struct Biol Commun* **2017**, *73*, 123–129, doi:10.1107/S2053230X17002096.
137. Wunder, E.A.; Slamti, L.; Suwondo, D.N.; Gibson, K.H.; Shang, Z.; Sindelar, C.V.; Trajtenberg, F.; Buschiazzo, A.; Ko, A.I.; Picardeau, M. FcpB Is a Surface Filament Protein of the Endoflagellum Required for the Motility of the Spirochete Leptospira. *Front Cell Infect Microbiol* **2018**, *8*, 130, doi:10.3389/fcimb.2018.00130.
138. San Martin, F.; Fule, L.; Iraola, G.; Buschiazzo, A.; Picardeau, M. Diving into the Complexity of the Spirochetal Endoflagellum. *Trends Microbiol* **2023**, *31*, 294–307, doi:10.1016/j.tim.2022.09.010.
139. Edgar, R.C. MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput. *Nucleic Acids Research* **2004**, *32*, 1792–1797, doi:10.1093/nar/gkh340.
140. Parte, A.; Krieg, N.R.; Ludwig, W.; Whitman, W.B.; Hedlund, B.P.; Paster, B.J.; Staley, J.T.; Ward, N.; Brown, D. *Bergey's Manual of Systematic Bacteriology: Volume 4: The Bacteroidetes, Spirochaetes, Tenericutes (Mollicutes), Acidobacteria, Fibrobacteres, Fusobacteria, Dictyoglomi, Gemmatimonadetes, Lentisphaerae, Verrucomicrobia, Chlamydiae, and Planctomycetes*; Springer Science & Business Media, 2011; ISBN 978-0-387-68572-4.
141. Gupta, R.S.; Mahmood, S.; Adeolu, M. A Phylogenomic and Molecular Signature Based Approach for Characterization of the Phylum Spirochaetes and Its Major Clades: Proposal for a Taxonomic Revision of the Phylum. *Front. Microbiol.* **2013**, *4*, doi:10.3389/fmicb.2013.00217.
142. Henderson, L.D.; Matthews-Palmer, T.R.S.; Gulbronson, C.J.; Ribardo, D.A.; Beeby, M.; Hendrixson, D.R. Diversification of Campylobacter Jejuni Flagellar C-Ring Composition Impacts Its Structure and Function in Motility, Flagellar Assembly, and Cellular Processes. *mBio* **2020**, *11*, e02286-19, doi:10.1128/mBio.02286-19.
143. Chang, Y.; Zhang, K.; Carroll, B.L.; Zhao, X.; Charon, N.W.; Norris, S.J.; Motaleb, M.A.; Li, C.; Liu, J. Molecular Mechanism for Rotational Switching of the Bacterial Flagellar Motor. *Nat Struct Mol Biol* **2020**, *27*, 1041–1047, doi:10.1038/s41594-020-0497-2.
144. Lynch, M.J.; Miller, M.; James, M.; Zhang, S.; Zhang, K.; Li, C.; Charon, N.W.; Crane, B.R. Structure and Chemistry of Lysinoalanine Crosslinking in the Spirochaete Flagella Hook. *Nat Chem Biol* **2019**, *15*, 959–965, doi:10.1038/s41589-019-0341-3.
145. Sal, M.S.; Li, C.; Motalab, M.A.; Shibata, S.; Aizawa, S.; Charon, N.W. Borrelia Burgdorferi Uniquely Regulates Its Motility Genes and Has an Intricate Flagellar Hook-Basal Body Structure. *J Bacteriol* **2008**, *190*, 1912–1921, doi:10.1128/JB.01421-07.
146. Kurniyati, K.; Liu, J.; Zhang, J.-R.; Min, Y.; Li, C. A Pleiotropic Role of FlaG in Regulating the Cell Morphogenesis and Flagellar Homeostasis at the Cell Poles of Treponema Denticola. *Cell Microbiol* **2019**, *21*, e12886, doi:10.1111/cmi.12886.
147. Sze, C.W.; Morado, D.R.; Liu, J.; Charon, N.W.; Xu, H.; Li, C. Carbon Storage Regulator A (CsrA(Bb)) Is a Repressor of Borrelia Burgdorferi Flagellin Protein FlaB. *Mol Microbiol* **2011**, *82*, 851–864, doi:10.1111/j.1365-2958.2011.07853.x.
148. Phoka, T.; Fule, L.; Da Fonseca, J.P.; Cokelaer, T.; Picardeau, M.; Patarakul, K. Investigating the Role of the Carbon Storage Regulator A (CsrA) in Leptospira Spp. *PLoS One* **2021**, *16*, e0260981, doi:10.1371/journal.pone.0260981.
149. Fule, L.; Halifa, R.; Fontana, C.; Sismeiro, O.; Legendre, R.; Varet, H.; Coppée, J.-Y.; Murray, G.L.; Adler, B.; Hendrixson, D.R.; et al. Role of the Major Determinant of Polar

Flagellation FlhG in the Endoflagella-Containing Spirochete Leptospira. *Mol Microbiol* **2021**, *116*, 1392–1406, doi:10.1111/mmi.14831.
150. Zhang, K.; He, J.; Catalano, C.; Guo, Y.; Liu, J.; Li, C. FlhF Regulates the Number and Configuration of Periplasmic Flagella in Borrelia Burgdorferi. *Mol Microbiol* **2020**, *113*, 1122–1139, doi:10.1111/mmi.14482.
151. Thompson, J.D.; Higgins, D.G.; Gibson, T.J. CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Research* **1994**, *22*, 4673–4680, doi:10.1093/nar/22.22.4673.
152. Troshin, P.V.; Procter, J.B.; Barton, G.J. Java Bioinformatics Analysis Web Services for Multiple Sequence Alignment--JABAWS:MSA. *Bioinformatics* **2011**, *27*, 2001–2002, doi:10.1093/bioinformatics/btr304.
153. McGuffin, L.J.; Bryson, K.; Jones, D.T. The PSIPRED Protein Structure Prediction Server. *Bioinformatics* **2000**, *16*, 404–405, doi:10.1093/bioinformatics/16.4.404.
154. Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J.E. The Phyre2 Web Portal for Protein Modeling, Prediction and Analysis. *Nat Protoc* **2015**, *10*, 845–858, doi:10.1038/nprot.2015.053.
155. Correa, A.; Ortega, C.; Obal, G.; Alzari, P.; Vincentelli, R.; Oppezzo, P.J. Generation of a Vector Suite for Protein Solubility Screening. *Front. Microbiol.* **2014**, *5*, doi:10.3389/fmicb.2014.00067.
156. Bucher, M.H.; Evdokimov, A.G.; Waugh, D.S. Differential Effects of Short Affinity Tags on the Crystallization of Pyrococcus Furiosus Maltodextrin-Binding Protein. *Acta Cryst D* **2002**, *58*, 392–397, doi:10.1107/S0907444901021187.
157. Vonrhein, C.; Flensburg, C.; Keller, P.; Sharff, A.; Smart, O.; Paciorek, W.; Womack, T.; Bricogne, G. Data Processing and Analysis with the autoPROC Toolbox. *Acta Crystallogr D Biol Crystallogr* **2011**, *67*, 293–302, doi:10.1107/S0907444911007773.
158. Baek, M.; DiMaio, F.; Anishchenko, I.; Dauparas, J.; Ovchinnikov, S.; Lee, G.R.; Wang, J.; Cong, Q.; Kinch, L.N.; Schaeffer, R.D.; et al. Accurate Prediction of Protein Structures and Interactions Using a Three-Track Neural Network. *Science* **2021**, *373*, 871–876, doi:10.1126/science.abj8754.
159. McCoy, A.J.; Grosse-Kunstleve, R.W.; Adams, P.D.; Winn, M.D.; Storoni, L.C.; Read, R.J. Phaser Crystallographic Software. *J Appl Crystallogr* **2007**, *40*, 658–674, doi:10.1107/S0021889807021206.
160. Afonine, P.V.; Grosse-Kunstleve, R.W.; Echols, N.; Headd, J.J.; Moriarty, N.W.; Mustyakimov, M.; Terwilliger, T.C.; Urzhumtsev, A.; Zwart, P.H.; Adams, P.D. Towards Automated Crystallographic Structure Refinement with Phenix.Refine. *Acta Crystallogr D Biol Crystallogr* **2012**, *68*, 352–367, doi:10.1107/S0907444912001308.
161. Emsley, P.; Lohkamp, B.; Scott, W.G.; Cowtan, K. Features and Development of Coot. *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 486–501, doi:10.1107/S0907444910007493.
162. Carvalho, P.C.; Lima, D.B.; Leprevost, F.V.; Santos, M.D.M.; Fischer, J.S.G.; Aquino, P.F.; Moresco, J.J.; Yates, J.R.; Barbosa, V.C. PatternLab for Proteomics 4.0: A One-Stop Shop for Analyzing Shotgun Proteomic Data. *Nat Protoc* **2016**, *11*, 102–117, doi:10.1038/nprot.2015.133.
163. Nury, C.; Redeker, V.; Dautrey, S.; Romieu, A.; van der Rest, G.; Renard, P.-Y.; Melki, R.; Chamot-Rooke, J. A Novel Bio-Orthogonal Cross-Linker for Improved Protein/Protein Interaction Analysis. *Anal Chem* **2015**, *87*, 1853–1860, doi:10.1021/ac503892c.

164. Rey, M.; Dupré, M.; Lopez-Neira, I.; Duchateau, M.; Chamot-Rooke, J. eXL-MS: An Enhanced Cross-Linking Mass Spectrometry Workflow To Study Protein Complexes. *Anal Chem* **2018**, *90*, 10707–10714, doi:10.1021/acs.analchem.8b00737.
165. Rey, M.; Dhenin, J.; Kong, Y.; Nouchikian, L.; Filella, I.; Duchateau, M.; Dupré, M.; Pellarin, R.; Duménil, G.; Chamot-Rooke, J. Advanced In Vivo Cross-Linking Mass Spectrometry Platform to Characterize Proteome-Wide Protein Interactions. *Anal Chem* **2021**, *93*, 4166–4174, doi:10.1021/acs.analchem.0c04430.
166. Jurėnas, D.; Rosa, L.T.; Rey, M.; Chamot-Rooke, J.; Fronzes, R.; Cascales, E. Mounting, Structure and Autocleavage of a Type VI Secretion-Associated Rhs Polymorphic Toxin. *Nat Commun* **2021**, *12*, 6998, doi:10.1038/s41467-021-27388-0.
167. Nouchikian, L.; Fernandez-Martinez, D.; Renard, P.-Y.; Sabot, C.; Duménil, G.; Rey, M.; Chamot-Rooke, J. Do Not Waste Time–Ensure Success in Your Cross-Linking Mass Spectrometry Experiments before You Begin. *Anal Chem* **2024**, *96*, 2506–2513, doi:10.1021/acs.analchem.3c04682.
168. Gibson, K.H.; Trajtenberg, F.; Wunder, E.A.; Brady, M.R.; San Martin, F.; Mechaly, A.; Shang, Z.; Liu, J.; Picardeau, M.; Ko, A.; et al. An Asymmetric Sheath Controls Flagellar Supercoiling and Motility in the Leptospira Spirochete. *eLife* **2020**, *9*, e53672, doi:10.7554/eLife.53672.
169. Brady, M.R.; Martin, F.S.; Debs, G.E.; Gibson, K.H.; Rodríguez, A.; Durán, R.; Wunder, E.A.; Ko, A.I.; Buschiazzo, A.; Sindelar, C.V. Structural Basis of Flagellar Filament Asymmetry and Supercoil Templating by Leptospira Spirochete Sheath Proteins 2022, 2022.03.03.482903.
170. Himes, B.A.; Zhang, P. emClarity: Software for High-Resolution Cryo-Electron Tomography and Subtomogram Averaging. *Nat Methods* **2018**, *15*, 955–961, doi:10.1038/s41592-018-0167-z.
171. Strnad, M.; Koizumi, N.; Nakamura, S.; Vancová, M.; Rego, R.O.M. It's Not All about Flagella - Sticky Invasion by Pathogenic Spirochetes. *Trends Parasitol* **2024**, *40*, 378–385, doi:10.1016/j.pt.2024.03.004.
172. Wolgemuth, C.W. Flagellar Motility of the Pathogenic Spirochetes. *Seminars in cell & developmental biology* **2015**, *46*, 104, doi:10.1016/j.semcdb.2015.10.015.
173. Brahamsha, B.; Greenberg, E.P. Cloning and Sequence Analysis of flaA, a Gene Encoding a Spirochaeta Aurantia Flagellar Filament Surface Antigen. *J Bacteriol* **1989**, *171*, 1692–1697, doi:10.1128/jb.171.3.1692-1697.1989.
174. Williams, A.W.; Yamaguchi, S.; Togashi, F.; Aizawa, S.I.; Kawagishi, I.; Macnab, R.M. Mutations in fliK and flhB Affecting Flagellar Hook and Filament Assembly in Salmonella Typhimurium. *Journal of Bacteriology* **1996**, *178*, 2960–2970, doi:10.1128/jb.178.10.2960-2970.1996.
175. Wyss, C. Flagellins, but Not Endoflagellar Sheath Proteins, of Treponema Pallidum and of Pathogen-Related Oral Spirochetes Are Glycosylated. *Infect Immun* **1998**, *66*, 5751–5754.
176. Chu, J.; Liu, J.; Hoover, T.R. Phylogenetic Distribution, Ultrastructure, and Function of Bacterial Flagellar Sheaths. *Biomolecules* **2020**, *10*, 363, doi:10.3390/biom10030363.
177. Sm, L. Flagellar Glycosylation - a New Component of the Motility Repertoire? *Microbiology (Reading, England)* **2006**, *152*, doi:10.1099/mic.0.28735-0.
178. Kim, S.Y.; Thanh, X.T.T.; Jeong, K.; Kim, S.B.; Pan, S.O.; Jung, C.H.; Hong, S.H.; Lee, S.E.; Rhee, J.H. Contribution of Six Flagellin Genes to the Flagellum Biogenesis of Vibrio

Vulnificus and In Vivo Invasion. *Infection and Immunity* **2014**, *82*, 29–42, doi:10.1128/iai.00654-13.
179. Zhu, S.; Nishikino, T.; Kojima, S.; Homma, M.; Liu, J. The Vibrio H-Ring Facilitates the Outer Membrane Penetration of the Polar Sheathed Flagellum. *Journal of Bacteriology* **2018**, *200*, e00387, doi:10.1128/JB.00387-18.
180. Kreutzberger, M.A.B.; Sobe, R.C.; Sauder, A.B.; Chatterjee, S.; Peña, A.; Wang, F.; Giron, J.A.; Kiessling, V.; Costa, T.R.D.; Conticello, V.P.; et al. Flagellin Outer Domain Dimerization Modulates Motility in Pathogenic and Soil Bacteria from Viscous Environments. *Nat Commun* **2022**, *13*, 1422, doi:10.1038/s41467-022-29069-y.
181. Sal, Melanie S Periplasmic Flagella of the Spirochetes Borrelia Burgdorferi and Brachyspira Hyodysenteriae Graduate Theses, Dissertations, and Problem Reports. 2280., 2005.
182. Sayers, E.W.; Beck, J.; Bolton, E.E.; Bourexis, D.; Brister, J.R.; Canese, K.; Comeau, D.C.; Funk, K.; Kim, S.; Klimke, W.; et al. Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **2021**, *49*, D10–D17, doi:10.1093/nar/gkaa892.
183. The UniProt Consortium UniProt: The Universal Protein Knowledgebase. *Nucleic Acids Res* **2017**, *45*, D158–D169, doi:10.1093/nar/gkw1099.
184. Davis, J.J.; Wattam, A.R.; Aziz, R.K.; Brettin, T.; Butler, R.; Butler, R.M.; Chlenski, P.; Conrad, N.; Dickerman, A.; Dietrich, E.M.; et al. The PATRIC Bioinformatics Resource Center: Expanding Data and Analysis Capabilities. *Nucleic Acids Res* **2020**, *48*, D606–D612, doi:10.1093/nar/gkz943.
185. Seemann, T. Prokka: Rapid Prokaryotic Genome Annotation. *Bioinformatics* **2014**, *30*, 2068–2069, doi:10.1093/bioinformatics/btu153.
186. Vallenet, D.; Calteau, A.; Dubois, M.; Amours, P.; Bazin, A.; Beuvin, M.; Burlot, L.; Bussell, X.; Fouteau, S.; Gautreau, G.; et al. MicroScope: An Integrated Platform for the Annotation and Exploration of Microbial Gene Functions through Genomic, Pangenomic and Metabolic Comparative Analysis. *Nucleic Acids Research* **2020**, *48*, D579–D589, doi:10.1093/nar/gkz926.
187. Eddy, S.R. Accelerated Profile HMM Searches. *PLoS Comput Biol* **2011**, *7*, e1002195, doi:10.1371/journal.pcbi.1002195.
188. Chaumeil, P.-A.; Mussig, A.J.; Hugenholtz, P.; Parks, D.H. GTDB-Tk: A Toolkit to Classify Genomes with the Genome Taxonomy Database. *Bioinformatics* **2020**, *36*, 1925–1927, doi:10.1093/bioinformatics/btz848.
189. Wilkins, M.R.; Gasteiger, E.; Bairoch, A.; Sanchez, J.C.; Williams, K.L.; Appel, R.D.; Hochstrasser, D.F. Protein Identification and Analysis Tools in the ExPASy Server. *Methods Mol Biol* **1999**, *112*, 531–552, doi:10.1385/1-59259-584-7:531.
190. Almagro Armenteros, J.J.; Tsirigos, K.D.; Sønderby, C.K.; Petersen, T.N.; Winther, O.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 5.0 Improves Signal Peptide Predictions Using Deep Neural Networks. *Nat Biotechnol* **2019**, *37*, 420–423, doi:10.1038/s41587-019-0036-z.
191. Holm, L. Dali Server: Structural Unification of Protein Families. *Nucleic Acids Res* **2022**, *50*, W210–W215, doi:10.1093/nar/gkac387.
192. van den Ent, F.; Löwe, J. RF Cloning: A Restriction-Free Method for Inserting Target Genes into Plasmids. *J Biochem Biophys Methods* **2006**, *67*, 67–74, doi:10.1016/j.jbbm.2005.12.008.

193. Rozen, S.; Skaletsky, H. Primer3 on the WWW for General Users and for Biologist Programmers. *Methods Mol Biol* **2000**, *132*, 365–386, doi:10.1385/1-59259-192-2:365.
194. Correa, A.; Oppezzo, P. Tuning Different Expression Parameters to Achieve Soluble Recombinant Proteins in E. Coli: Advantages of High-Throughput Screening. *Biotechnol J* **2011**, *6*, 715–730, doi:10.1002/biot.201100025.
195. Drummond, E.; Kavanagh, T.; Pires, G.; Marta-Ariza, M.; Kanshin, E.; Nayak, S.; Faustin, A.; Berdah, V.; Ueberheide, B.; Wisniewski, T. The Amyloid Plaque Proteome in Early Onset Alzheimer's Disease and Down Syndrome. *Acta Neuropathol Commun* **2022**, *10*, 53, doi:10.1186/s40478-022-01356-1.
196. Alzohairy, A. BioEdit: An Important Software for Molecular Biology. *GERF Bulletin of Biosciences* **2011**, *2*, 60–61.
197. Schlegel, S.; Löfblom, J.; Lee, C.; Hjelm, A.; Klepsch, M.; Strous, M.; Drew, D.; Slotboom, D.J.; de Gier, J.-W. Optimizing Membrane Protein Overexpression in the Escherichia Coli Strain Lemo21(DE3). *J Mol Biol* **2012**, *423*, 648–659, doi:10.1016/j.jmb.2012.07.019.
198. Bradford, M.M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Anal Biochem* **1976**, *72*, 248–254, doi:10.1016/0003-2697(76)90527-3.
199. Battye, T.G.G.; Kontogiannis, L.; Johnson, O.; Powell, H.R.; Leslie, A.G.W. iMOSFLM: A New Graphical Interface for Diffraction-Image Processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* **2011**, *67*, 271–281, doi:10.1107/S0907444910048675.
200. Evans, P.R.; Murshudov, G.N. How Good Are My Data and What Is the Resolution? *Acta Crystallogr D Biol Crystallogr* **2013**, *69*, 1204–1214, doi:10.1107/S0907444913000061.
201. Söding, J.; Biegert, A.; Lupas, A.N. The HHpred Interactive Server for Protein Homology Detection and Structure Prediction. *Nucleic Acids Res* **2005**, *33*, W244-248, doi:10.1093/nar/gki408.
202. Bender, B.J.; Cisneros, A.; Duran, A.M.; Finn, J.A.; Fu, D.; Lokits, A.D.; Mueller, B.K.; Sangha, A.K.; Sauer, M.F.; Sevy, A.M.; et al. Protocols for Molecular Modeling with Rosetta3 and RosettaScripts. *Biochemistry* **2016**, *55*, 4748–4763, doi:10.1021/acs.biochem.6b00444.
203. Song, Y.; DiMaio, F.; Wang, R.Y.-R.; Kim, D.; Miles, C.; Brunette, T.; Thompson, J.; Baker, D. High-Resolution Comparative Modeling with RosettaCM. *Structure* **2013**, *21*, 1735–1742, doi:10.1016/j.str.2013.08.005.
204. Muller, L.; Fornecker, L.; Cianferani, S.; Carapito, C. Tube-Gel: A Fast and Effective Sample Preparation Method for High-Throughput Quantitative Proteomics. In *Proteomics for Biomarker Discovery: Methods and Protocols*; Brun, V., Couté, Y., Eds.; Springer: New York, NY, 2019; pp. 123–127 ISBN 978-1-4939-9164-8.
205. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res* **2003**, *13*, 2498–2504, doi:10.1101/gr.1239303.
206. Schiffrin, B.; Radford, S.E.; Brockwell, D.J.; Calabrese, A.N. PyXlinkViewer: A Flexible Tool for Visualization of Protein Chemical Crosslinking Data within the PyMOL Molecular Graphics System. *Protein Sci* **2020**, *29*, 1851–1857, doi:10.1002/pro.3902.

207. Punjani, A.; Rubinstein, J.L.; Fleet, D.J.; Brubaker, M.A. cryoSPARC: Algorithms for Rapid Unsupervised Cryo-EM Structure Determination. *Nat Methods* **2017**, *14*, 290–296, doi:10.1038/nmeth.4169.
208. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera--a Visualization System for Exploratory Research and Analysis. *J Comput Chem* **2004**, *25*, 1605–1612, doi:10.1002/jcc.20084.
209. Liebschner, D.; Afonine, P.V.; Baker, M.L.; Bunkóczi, G.; Chen, V.B.; Croll, T.I.; Hintze, B.; Hung, L.W.; Jain, S.; McCoy, A.J.; et al. Macromolecular Structure Determination Using X-Rays, Neutrons and Electrons: Recent Developments in Phenix. *Acta Crystallogr D Struct Biol* **2019**, *75*, 861–877, doi:10.1107/S2059798319011471.
210. Afonine, P.V.; Poon, B.K.; Read, R.J.; Sobolev, O.V.; Terwilliger, T.C.; Urzhumtsev, A.; Adams, P.D. Real-Space Refinement in PHENIX for Cryo-EM and Crystallography. *Acta Crystallogr D Struct Biol* **2018**, *74*, 531–544, doi:10.1107/S2059798318006551.
211. Yamashita, K.; Palmer, C.M.; Burnley, T.; Murshudov, G.N. Cryo-EM Single-Particle Structure Refinement and Map Calculation Using Servalcat. *Acta Cryst D* **2021**, *77*, 1282–1291, doi:10.1107/S2059798321009475.
212. Nicholls, R.A.; Fischer, M.; McNicholas, S.; Murshudov, G.N. Conformation-Independent Structural Comparison of Macromolecules with ProSMART. *Acta Crystallogr D Biol Crystallogr* **2014**, *70*, 2487–2499, doi:10.1107/S1399004714016241.
213. Williams, C.J.; Headd, J.J.; Moriarty, N.W.; Prisant, M.G.; Videau, L.L.; Deis, L.N.; Verma, V.; Keedy, D.A.; Hintze, B.J.; Chen, V.B.; et al. MolProbity: More and Better Reference Data for Improved All-Atom Structure Validation. *Protein Sci* **2018**, *27*, 293–315, doi:10.1002/pro.3330.

VII - ANEXOS

ANEXO 1 – Tablas de Espectrometría de masas comparativa

Análisis SC de proteínas exclusivas en *wt* vs *fcpA*⁻					
Locus	Replica 1	Replica 2	Replica 3	Señal total	Descripción
tr|B0SMK8|B0SMK8_LEPBP	288	265	293	846	Uncharacterized protein OX=456481 GN=LEPBI_I1029 PE=4 SV=1
tr|B0SKN8|B0SKN8_LEPBP	207	194	176	577	Uncharacterized protein OX=456481 GN=LEPBI_I0662 PE=4 SV=1
tr|B0SNJ6|B0SNJ6_LEPBP	26	12	43	81	Uncharacterized protein OX=456481 GN=LEPBI_I1161 PE=4 SV=1
tr|B0STN7|B0STN7_LEPBP	62	7	9	78	Uncharacterized protein OX=456481 GN=LEPBI_II0033 PE=4 SV=1
tr|B0SSE3|B0SSE3_LEPBP	5	8	7	20	Putative TPR-repeat-protein OX=456481 GN=LEPBI_I1930 PE=4 SV=1
tr|B0SMT6|B0SMT6_LEPBP	4	2	4	10	Uncharacterized protein OX=456481 GN=LEPBI_I2732 PE=4 SV=1
tr|B0SL54|B0SL54_LEPBP	3	4	3	10	Uncharacterized protein OX=456481 GN=LEPBI_I0729 PE=4 SV=1
sp|B0SNJ9|HEMH_LEPBP	3	3	3	9	Ferrochelatase OX=456481 GN=hemH PE=3 SV=1
tr|B0SQ23|B0SQ23_LEPBP	3	3	3	9	Uncharacterized protein OX=456481 GN=LEPBI_I3109 PE=4 SV=1
tr|B0SS20|B0SS20_LEPBP	3	2	3	8	Uncharacterized protein OX=456481 GN=LEPBI_I1804 PE=4 SV=1
tr|B0SMH9|B0SMH9_LEPBP	2	4	2	8	Putative lipid A core-O-antigen ligase putative MP OX=456481 GN=LEPBI_I0999 PE=4 SV=1
tr|B0SR15|B0SR15_LEPBP	2	3	3	8	Uncharacterized protein OX=456481 GN=LEPBI_I1609 PE=4 SV=1
tr|B0SSZ1|B0SSZ1_LEPBP	3	2	3	8	Uncharacterized protein OX=456481 GN=LEPBI_I2129 PE=4 SV=1
sp|B0SMW9|RL31_LEPBP	1	3	3	7	50S ribosomal protein L31 OX=456481 GN=rpmE PE=3 SV=1
tr|B0SMW5|B0SMW5_LEPBP	3	2	2	7	Uncharacterized protein OX=456481 GN=LEPBI_I2761 PE=4 SV=1
tr|B0SLU0|B0SLU0_LEPBP	2	3	2	7	Uncharacterized protein OX=456481 GN=LEPBI_I0867 PE=4 SV=1
tr|B0SKY2|B0SKY2_LEPBP	1	4	2	7	Uncharacterized protein OX=456481 GN=LEPBI_I2385 PE=4 SV=1
tr|B0SNN0|B0SNN0_LEPBP	2	2	2	6	Enoyl-[acyl-carrier-protein] reductase [NADH] OX=456481 GN=LEPBI_I1195 PE=3 SV=1
tr|B0SPN3|B0SPN3_LEPBP	6	5		11	Cell shape protein MreC OX=456481 GN=mreC PE=3 SV=1
tr|B0SM39|B0SM39_LEPBP	4		3	7	Uncharacterized protein OX=456481 GN=LEPBI_I2588 PE=4 SV=1
tr|B0SSX5|B0SSX5_LEPBP	5	2		7	Putative RNA pseudouridine synthase OX=456481 GN=LEPBI_I2113 PE=4 SV=1
tr|B0SQF3|B0SQF3_LEPBP		4	2	6	Putative esterase putative signal peptide OX=456481 GN=LEPBI_I1504 PE=4 SV=1
tr|B0SU70|B0SU70_LEPBP	2	3		5	Ankyrin_rpt-contain_dom domain-containing protein OX=456481 GN=LEPBI_II0221 PE=4 SV=1
tr|B0SML0|B0SML0_LEPBP		2	3	5	Putative transcriptional regulator OX=456481 GN=LEPBI_I1031 PE=4 SV=1

Tabla 7: Proteínas detectadas de forma exclusiva en cepas *wt vs fcpA*⁻ (p<0,05) identificadas utilizando el módulo estadístico Venn diagram de PatternLab for Proteomics V. Se muestran las 24 proteínas exclusivas de la cepa *wt* frente a la *fcpA*⁻, pudiéndose detectar además 146 proteínas exclusivas de *fcpA*⁻ (no mostradas). Cuatro de las seis proteínas exclusivas de *wt* que presentan mayor señal total terminan siendo indentificadas como nuevos componentes flagelares y se muestran en amarillo en la tabla.

Análisis SC de proteínas aumentadas en wt vs fcpA⁻			
#Locus	Fold Change	pValue	Description
tr|B0SR03|B0SR03_LEPBP	51,94	0,00004	Uncharacterized protein OX=456481 GN=LEPBI_I1597 PE=1 SV=1
tr|B0STJ8|B0STJ8_LEPBP	44,14	0,00670	Uncharacterized protein OX=456481 GN=LEPBI_I0267 PE=1 SV=1
tr|B0SPP1|B0SPP1_LEPBP	39,42	0,01050	Uncharacterized protein OX=456481 GN=LEPBI_I3081 PE=4 SV=1
tr|B0STF2|B0STF2_LEPBP	35,92	0,00002	Uncharacterized protein OX=456481 GN=LEPBI_I2297 PE=4 SV=1
tr|B0SQV9|B0SQV9_LEPBP	18,29	0,00982	Uncharacterized protein OX=456481 GN=LEPBI_I3291 PE=4 SV=1
tr|B0SSW5|B0SSW5_LEPBP	3,47	0,01025	Uncharacterized protein OX=456481 GN=LEPBI_I2103 PE=4 SV=1
tr|B0SJG1|B0SJG1_LEPBP	2,81	0,01848	Putative serine-type endopeptidase OX=456481 GN=LEPBI_I0523 PE=4 SV=1
tr|B0SR80|B0SR80_LEPBP	2,35	0,01897	General secretion pathway protein C OX=456481 GN=gspC PE=4 SV=1
tr|B0SPD3|B0SPD3_LEPBP	2,29	0,01487	Uncharacterized protein OX=456481 GN=LEPBI_I1347 PE=4 SV=1
tr|B0SJC6|B0SJC6_LEPBP	2,26	0,00150	Uncharacterized protein OX=456481 GN=LEPBI_I0551 PE=4 SV=1
tr|B0SKP2|B0SKP2_LEPBP	2,17	0,00484	Uncharacterized protein OX=456481 GN=LEPBI_I0666 PE=4 SV=1
tr|B0SU19|B0SU19_LEPBP	2,16	0,02134	AB hydrolase-1 domain-containing protein OX=456481 GN=LEPBI_II0168 PE=4 SV=1
sp|B0SPA9|GLPK_LEPBP	2,00	0,00774	Glycerol kinase OX=456481 GN=glpK PE=3 SV=1
tr|B0SL60|B0SL60_LEPBP	1,94	0,01701	Diadenylate cyclase OX=456481 GN=dacA PE=3 SV=1
tr|B0SSZ4|B0SSZ4_LEPBP	1,90	0,01625	Flagellin OX=456481 GN=flaB3 PE=3 SV=1
tr|B0SQZ5|B0SQZ5_LEPBP	1,83	0,03023	Flagellin OX=456481 GN=flaB PE=3 SV=1
tr|B0SK68|B0SK68_LEPBP	1,70	0,00145	Uncharacterized protein OX=456481 GN=LEPBI_I0039 PE=4 SV=1
tr|B0SPX1|B0SPX1_LEPBP	1,70	0,00909	Uncharacterized protein OX=456481 GN=LEPBI_I1429 PE=4 SV=1

Tabla 8: Proteínas aumentadas en cepas *L. biflexa wt* con respecto a las cepas *fcpA*⁻. Para identificar las proteínas encontradas en ambas cepas pero que muestran una diferencia estadística en sus niveles (BH q-value < 0.05) se utilizó el módulo TFold de PatternLab for Proteomics V. En esta tabla se muestran los códigos uniprot que identifican las proteínas aumentadas en cepas *wt*, ordenadas según su ¨fold change¨ es decir el cociente relativo encontrado entre ambas condiciones y marcadas en amarillo aquellas que finalmente se confirman como componentes del filamento.

Análisis SC de proteínas exclusivas en *wt* vs *flaA2*⁻					
Locus	Replica 1	Replica 2	Replica 3	Señal total	Descripción
tr|B0SKT5|B0SKT5_LEPBP	283	135	199	617	Flagellar filament outer layer protein (Sheath protein) OX=456481 GN=flaA2 PE=4 SV=1
tr|B0SKT4|B0SKT4_LEPBP	101	12	75	188	Flagellar filament outer layer protein A OX=456481 GN=flaA1 PE=4 SV=1
tr|B0SNJ6|B0SNJ6_LEPBP	26	12	43	81	Uncharacterized protein OX=456481 GN=LEPBI_I1161 PE=4 SV=1
tr|B0SQV9|B0SQV9_LEPBP	20	20	21	61	Uncharacterized protein OX=456481 GN=LEPBI_I3291 PE=4 SV=1
tr|B0SQJ6|B0SQJ6_LEPBP	7	3	4	14	Putative short-chain dehydrogenase/reductase SDR OX=456481 GN=LEPBI_I1547 PE=4 SV=1
tr|B0SMT6|B0SMT6_LEPBP	4	2	4	10	Uncharacterized protein OX=456481 GN=LEPBI_I2732 PE=4 SV=1
tr|B0SPA5|B0SPA5_LEPBP		3	3	6	tRNA N6-adenosine threonylcarbamoyltransferase OX=456481 GN=tsaD PE=3 SV=1
tr|B0SU70|B0SU70_LEPBP	2	3		5	Ankyrin_rpt-contain_dom domain-containing protein OX=456481 GN=LEPBI_II0221 PE=4 SV=1

Tabla 9: Resultados del módulo estadístico Venn diagram del software PatternLab for Proteomics V utilizado para identificar las proteínas detectadas de forma exclusiva en cepas *L. biflexa flaA2*⁻ o *wt* (p<0,05). Esto nos permitió la identificación de 461 proteínas exclusivas de cepas *flaA2*⁻ y de las 8 proteínas exclusivas de *wt* que se muestran en esta tabla. La proteína de mayor señal total es FlaA2, seguida por FlaA1, otro componente previamente conocido del filamento que se encuentra ausente en los filamentos *flaA2*⁻ ya sea por una cuestión estructural del filamento o porque están formando parte de un mismo operón[77]. En amarillo se destacan las proteínas que efectivaemente constituyen nuevos componentes identificados en el filamento.

Análisis SC de proteínas aumentadas en *wt* vs *flaA2*⁻			
#Locus	Fold Change	pValue	Description
tr|B0SJC6|B0SJC6_LEPBP	18,15	0,000025	Uncharacterized protein OX=456481 GN=LEPBI_I0551 PE=4 SV=1
tr|B0SKN8|B0SKN8_LEPBP	16,58	0,000010	Uncharacterized protein OX=456481 GN=LEPBI_I0662 PE=4 SV=1
tr|B0SM18|B0SM18_LEPBP	15,49	0,012670	O-succinylhomoserine sulfhydrylase OX=456481 GN=metZ PE=3 SV=1
tr|B0SMK8|B0SMK8_LEPBP	14,27	0,000087	Uncharacterized protein OX=456481 GN=LEPBI_I1029 PE=4 SV=1
tr|B0STF2|B0STF2_LEPBP	14,17	0,000044	Uncharacterized protein OX=456481 GN=LEPBI_I2297 PE=4 SV=1
tr|B0SPP1|B0SPP1_LEPBP	8,84	0,003906	Uncharacterized protein OX=456481 GN=LEPBI_I3081 PE=4 SV=1
tr|B0SKF0|B0SKF0_LEPBP	7,74	0,027587	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase OX=456481 GN=lpxA PE=4 SV=1
tr|B0SUF9|B0SUF9_LEPBP	5,94	0,000565	Uncharacterized protein OX=456481 GN=LEPBI_p0029 PE=4 SV=1
sp|B0STJ3|SYL_LEPBP	5,82	0,021223	Leucine--tRNA ligase OX=456481 GN=leuS PE=3 SV=1
tr|B0SJA5|B0SJA5_LEPBP	5,63	0,010347	Polyphosphate kinase OX=456481 GN=ppk PE=3 SV=1
tr|B0STS1|B0STS1_LEPBP	5,22	0,009023	Glutamate synthase large chain OX=456481 GN=gltB PE=3 SV=1
tr|B0SK22|B0SK22_LEPBP	4,54	0,002413	3-oxoacyl-[acyl-carrier-protein] reductase OX=456481 GN=fabG4 PE=4 SV=1
tr|B0SMM0|B0SMM0_LEPBP	4,23	0,016370	Uncharacterized protein OX=456481 GN=LEPBI_I2666 PE=4 SV=1
tr|B0SQY0|B0SQY0_LEPBP	4,14	0,015717	Beta-lactamase domain-containing protein OX=456481 GN=LEPBI_I3312 PE=4 SV=1
tr|B0SKE6|B0SKE6_LEPBP	4,10	0,016727	Putative integral OM protein TolC, efflux pump component OX=456481 GN=LEPBI_I0360 PE=3 SV=1
tr|B0SLY4|B0SLY4_LEPBP	3,69	0,012243	Putative PLP-dependent aminotransferase OX=456481 GN=LEPBI_I0912 PE=4 SV=1
tr|B0SQM4|B0SQM4_LEPBP	3,64	0,006829	Putative alpha-glucosidase II OX=456481 GN=LEPBI_I3206 PE=3 SV=1
tr|B0SJX7|B0SJX7_LEPBP	3,64	0,007600	Uncharacterized protein OX=456481 GN=LEPBI_I0145 PE=4 SV=1
tr|B0SKQ9|B0SKQ9_LEPBP	3,33	0,000690	Uncharacterized protein OX=456481 GN=LEPBI_I0684 PE=4 SV=1
tr|B0SJ54|B0SJ54_LEPBP	3,33	0,011478	Uncharacterized protein OX=456481 GN=LEPBI_I0512 PE=4 SV=1
tr|B0SM40|B0SM40_LEPBP	3,31	0,009573	3-oxoacyl-[acyl-carrier-protein] reductase OX=456481 GN=fabG PE=3 SV=1
tr|B0SS97|B0SS97_LEPBP	3,10	0,003495	Iron(III) dicitrate TonB-dependent receptor putative signal peptide OX=456481 GN=fecA PE=3 SV=1
tr|B0STW0|B0STW0_LEPBP	3,05	0,001890	Uncharacterized protein OX=456481 GN=LEPBI_II0108 PE=4 SV=1
tr|B0SLY1|B0SLY1_LEPBP	3,04	0,020336	Putative periplasmic component of the Tol biopolymer transport system OX=456481 GN=LEPBI_I0909 PE=4 SV=1
tr|B0SK54|B0SK54_LEPBP	3,01	0,015142	3-oxoacyl-[acyl-carrier-protein] reductase OX=456481 GN=fabG3b PE=4 SV=1
tr|B0SPM8|B0SPM8_LEPBP	2,96	0,004060	Putative ankyrin-like protein OX=456481 GN=LEPBI_I3068 PE=4 SV=1
tr|B0STW1|B0STW1_LEPBP	2,80	0,002428	Uncharacterized protein OX=456481 GN=LEPBI_II0109 PE=4 SV=1
tr|B0SQ25|B0SQ25_LEPBP	2,78	0,003815	Uncharacterized protein OX=456481 GN=LEPBI_I3111 PE=4 SV=1
tr|B0SPD3|B0SPD3_LEPBP	2,76	0,010124	Uncharacterized protein OX=456481 GN=LEPBI_I1347 PE=4 SV=1
tr|B0SR80|B0SR80_LEPBP	2,73	0,016358	General secretion pathway protein C putative signal peptide OX=456481 GN=gspC PE=4 SV=1
tr|B0SMX3|B0SMX3_LEPBP	2,57	0,002503	Uncharacterized protein OX=456481 GN=LEPBI_I1037 PE=4 SV=1
tr|B0SKQ7|B0SKQ7_LEPBP	2,57	0,023329	Uncharacterized protein OX=456481 GN=LEPBI_I0682 PE=4 SV=1
tr|B0SQ77|B0SQ77_LEPBP	2,57	0,025274	Putative acylglycerol lipase OX=456481 GN=LEPBI_I3164 PE=4 SV=1
tr|B0SLN1|B0SLN1_LEPBP	2,49	0,003957	Uncharacterized protein OX=456481 GN=LEPBI_I2536 PE=4 SV=1
tr|B0SSW5|B0SSW5_LEPBP	2,35	0,010233	Uncharacterized protein OX=456481 GN=LEPBI_I2103 PE=4 SV=1
tr|B0SKP2|B0SKP2_LEPBP	2,31	0,001612	Uncharacterized protein OX=456481 GN=LEPBI_I0666 PE=4 SV=1
tr|B0SQQ8|B0SQQ8_LEPBP	2,29	0,000745	Putative TPR-repeat-containing protein putative signal peptide OX=456481 GN=LEPBI_I3240 PE=4 SV=1
sp|B0SJR2|ASSY_LEPBP	2,27	0,004132	Argininosuccinate synthase OX=456481 GN=argG PE=3 SV=1
tr|B0SQC6|B0SQC6_LEPBP	2,19	0,001318	Uncharacterized protein OX=456481 GN=LEPBI_I1477 PE=4 SV=1
tr|B0SKQ8|B0SKQ8_LEPBP	2,17	0,024085	Putative lipoprotein putative signal peptide OX=456481 GN=LEPBI_I0683 PE=4 SV=1
tr|B0STR9|B0STR9_LEPBP	2,15	0,008058	DUF5683 domain-containing protein OX=456481 GN=LEPBI_II0065 PE=4 SV=1
tr|B0STJ8|B0STJ8_LEPBP	2,14	0,017432	Uncharacterized protein OX=456481 GN=LEPBI_I0267 PE=1 SV=1
tr|B0SSZ5|B0SSZ5_LEPBP	2,04	0,009481	Flagellin OX=456481 GN=flaB2 PE=1 SV=1
tr|B0SS87|B0SS87_LEPBP	2,04	0,003942	Putative outer membrane protein, OmpA domain putative MP OX=456481 GN=LEPBI_I1873 PE=4 SV=1
tr|B0SSX9|B0SSX9_LEPBP	2,02	0,002592	Uncharacterized protein OX=456481 GN=LEPBI_I2117 PE=4 SV=1
sp|B0SSI8|RL11_LEPBP	2,00	0,005419	50S ribosomal protein L11 OX=456481 GN=rplK PE=3 SV=1
tr|B0SSZ4|B0SSZ4_LEPBP	1,92	0,021052	Flagellin OX=456481 GN=flaB3 PE=3 SV=1
tr|B0SQZ5|B0SQZ5_LEPBP	1,85	0,008047	Flagellin OX=456481 GN=flaB PE=3 SV=1
tr|B0SLP1|B0SLP1_LEPBP	1,71	0,003652	Carbamoyl-phosphate synthase small chain OX=456481 GN=carA PE=3 SV=1
tr|B0SMZ1|B0SMZ1_LEPBP	1,67	0,031831	Uncharacterized protein OX=456481 GN=LEPBI_I1055 PE=4 SV=1
tr|B0SR03|B0SR03_LEPBP	1,61	0,028380	Uncharacterized protein OX=456481 GN=LEPBI_I1597 PE=1 SV=1

Tabla 10: Proteínas aumentadas en cepas *L. biflexa wt* comparada con cepas *flaA2*⁻. Para identificar las proteínas

encontradas en ambas cepas pero que muestran una diferencia estadística en sus niveles (BH q-value < 0.05) se utilizó el módulo TFold de PatternLab for Proteomics V. En esta tabla se muestran los códigos uniprot que identifican las proteínas aumentadas en cepas wt, ordenadas según su ¨fold change¨ es decir el cociente relativo encontrado entre ambas condiciones y marcadas en amarillo aquellas que finalmente se confirman como componentes del filamento. Cinco nuevos componentes identificados en el filamento aparecen entre los candidatos que se encuentran con mayor diferencia en sus niveles, y otras 5 proteínas previamente conocidas del filamento también se expresan pero casi a los mismos niveles que en los mutantes *flaA2*$^{-}$.

Análisis SC de proteínas exclusivas en *flaA2*⁻ vs *fcpA*⁻					
Locus	**Replica 1**	**Replica 2**	**Replica 3**	**Señal total**	**Descripción**
tr|B0SK46|B0SK46_LEPBP	7	7	5	19	Uncharacterized protein GN=LEPBI_I0025 PE=4 SV=1
tr|B0SJX2|B0SJX2_LEPBP	4	3	7	14	Uncharacterized protein GN=LEPBI_I0134 PE=4 SV=1
tr|B0SJY6|B0SJY6_LEPBP	11	7	3	21	Uncharacterized protein GN=LEPBI_I0163 PE=4 SV=1
tr|B0SJD6|B0SJD6_LEPBP	17	14	19	50	Uncharacterized protein GN=LEPBI_I0311 PE=4 SV=1
tr|B0SJI3|B0SJI3_LEPBP	5	5	2	12	Uncharacterized protein GN=LEPBI_I0393 PE=4 SV=1
tr|B0SIW6|B0SIW6_LEPBP	6	7	3	16	Uncharacterized protein GN=LEPBI_I0417 PE=4 SV=1
tr|B0SJ45|B0SJ45_LEPBP	2	3	4	9	Uncharacterized protein GN=LEPBI_I0475 PE=4 SV=1
tr|B0SJ80|B0SJ80_LEPBP	2	2	14	18	Uncharacterized protein GN=LEPBI_I0492 PE=4 SV=1
tr|B0SKJ0|B0SKJ0_LEPBP	8	3	7	18	Uncharacterized protein GN=LEPBI_I0612 PE=4 SV=1
tr|B0SKM7|B0SKM7_LEPBP	5	2	1	8	Uncharacterized protein GN=LEPBI_I0651 PE=4 SV=1
tr|B0SKN8|B0SKN8_LEPBP	21	15	21	57	Uncharacterized protein GN=LEPBI_I0662 PE=4 SV=1
tr|B0SKR1|B0SKR1_LEPBP	5	4	9	18	Uncharacterized protein GN=LEPBI_I0686 PE=4 SV=1
tr|B0SKR9|B0SKR9_LEPBP	2	4		6	Uncharacterized protein GN=LEPBI_I0694 PE=4 SV=1
tr|B0SL54|B0SL54_LEPBP	3	3	5	11	Uncharacterized protein GN=LEPBI_I0729 PE=4 SV=1
tr|B0SL90|B0SL90_LEPBP	1	3	2	6	Uncharacterized protein GN=LEPBI_I0766 PE=4 SV=1
tr|B0SL92|B0SL92_LEPBP	2	4	2	8	Uncharacterized protein GN=LEPBI_I0768 PE=4 SV=1
tr|B0SLR6|B0SLR6_LEPBP	3		16	19	Uncharacterized protein GN=LEPBI_I0841 PE=4 SV=1
tr|B0SLU0|B0SLU0_LEPBP	7	3	4	14	Uncharacterized protein GN=LEPBI_I0867 PE=4 SV=1
tr|B0SLW4|B0SLW4_LEPBP	9	7	1	17	Uncharacterized protein GN=LEPBI_I0891 PE=4 SV=1
tr|B0SLY2|B0SLY2_LEPBP	5	5	3	13	Uncharacterized protein GN=LEPBI_I0910 PE=4 SV=1
tr|B0SME0|B0SME0_LEPBP	6	4	6	16	Uncharacterized protein GN=LEPBI_I0960 PE=4 SV=1
tr|B0SMF2|B0SMF2_LEPBP	5	1	6	12	Uncharacterized protein GN=LEPBI_I0972 PE=4 SV=1
tr|B0SMI4|B0SMI4_LEPBP	3	3		6	Uncharacterized protein GN=LEPBI_I1004 PE=4 SV=1
tr|B0SMI5|B0SMI5_LEPBP		3	3	6	Uncharacterized protein GN=LEPBI_I1005 PE=4 SV=1
tr|B0SMK1|B0SMK1_LEPBP	5	3	1	9	Uncharacterized protein GN=LEPBI_I1022 PE=4 SV=1
tr|B0SMK8|B0SMK8_LEPBP	22	23	51	96	Uncharacterized protein GN=LEPBI_I1029 PE=4 SV=1
tr|B0SMY1|B0SMY1_LEPBP	2	1	9	12	Uncharacterized protein GN=LEPBI_I1045 PE=4 SV=1
tr|B0SNJ0|B0SNJ0_LEPBP	29	20	17	66	Uncharacterized protein GN=LEPBI_I1155 PE=4 SV=1
tr|B0SNM7|B0SNM7_LEPBP	6	13		19	Uncharacterized protein GN=LEPBI_I1192 PE=4 SV=1
tr|B0SPV7|B0SPV7_LEPBP	3	3		6	Uncharacterized protein GN=LEPBI_I1414 PE=4 SV=1
tr|B0SQ00|B0SQ00_LEPBP	4	6		10	Uncharacterized protein GN=LEPBI_I1459 PE=4 SV=1
tr|B0SQK9|B0SQK9_LEPBP	1	6	26	33	Uncharacterized protein GN=LEPBI_I1560 PE=4 SV=1
tr|B0SQL7|B0SQL7_LEPBP	7	5	13	25	Uncharacterized protein GN=LEPBI_I1568 PE=4 SV=1
tr|B0SR15|B0SR15_LEPBP	21	9	15	45	Uncharacterized protein GN=LEPBI_I1609 PE=4 SV=1
tr|B0SR61|B0SR61_LEPBP	4	2		6	Uncharacterized protein GN=LEPBI_I1655 PE=4 SV=1
tr|B0SRJ7|B0SRJ7_LEPBP	4	2	6	12	Uncharacterized protein GN=LEPBI_I1685 PE=4 SV=1
tr|B0SRN1|B0SRN1_LEPBP	6	5	3	14	Uncharacterized protein GN=LEPBI_I1720 PE=4 SV=1
tr|B0SRQ7|B0SRQ7_LEPBP	4	4		8	Uncharacterized protein GN=LEPBI_I1746 PE=4 SV=1
tr|B0SS20|B0SS20_LEPBP	3	5	7	15	Uncharacterized protein GN=LEPBI_I1804 PE=4 SV=1
tr|B0SS28|B0SS28_LEPBP	2	2	2	6	Uncharacterized protein GN=LEPBI_I1812 PE=4 SV=1
tr|B0SS69|B0SS69_LEPBP	3	2	4	9	Uncharacterized protein GN=LEPBI_I1854 PE=4 SV=1
tr|B0SSA9|B0SSA9_LEPBP	8	8		16	Uncharacterized protein GN=LEPBI_I1896 PE=4 SV=1
tr|B0SSJ1|B0SSJ1_LEPBP	6	5	1	12	Uncharacterized protein GN=LEPBI_I1979 PE=4 SV=1
tr|B0SSK5|B0SSK5_LEPBP	4	4	2	10	Uncharacterized protein GN=LEPBI_I1993 PE=4 SV=1
tr|B0SSN3|B0SSN3_LEPBP	15	7		22	Uncharacterized protein GN=LEPBI_I2021 PE=4 SV=1
tr|B0SSS6|B0SSS6_LEPBP	4	6		10	Uncharacterized protein GN=LEPBI_I2064 PE=4 SV=1
tr|B0SSW7|B0SSW7_LEPBP	15	8	9	32	Uncharacterized protein GN=LEPBI_I2105 PE=4 SV=1
tr|B0SSZ1|B0SSZ1_LEPBP		3	4	7	Uncharacterized protein GN=LEPBI_I2129 PE=4 SV=1
tr|B0ST42|B0ST42_LEPBP	4	5		9	Uncharacterized protein GN=LEPBI_I2183 PE=4 SV=1
tr|B0STE8|B0STE8_LEPBP	4	4	9	17	Uncharacterized protein GN=LEPBI_I2293 PE=4 SV=1
tr|B0SKX4|B0SKX4_LEPBP	13	9		22	Uncharacterized protein GN=LEPBI_I2376 PE=4 SV=1
tr|B0SKZ3|B0SKZ3_LEPBP	9	3	4	16	Uncharacterized protein GN=LEPBI_I2396 PE=4 SV=1
tr|B0SKZ4|B0SKZ4_LEPBP	4	2	3	9	Uncharacterized protein GN=LEPBI_I2397 PE=4 SV=1
tr|B0SM22|B0SM22_LEPBP		5	7	12	Uncharacterized protein GN=LEPBI_I2571 PE=4 SV=1
tr|B0SM30|B0SM30_LEPBP	4	3	1	8	Uncharacterized protein GN=LEPBI_I2579 PE=4 SV=1
tr|B0SM39|B0SM39_LEPBP	14	14	26	54	Uncharacterized protein GN=LEPBI_I2588 PE=4 SV=1
tr|B0SML2|B0SML2_LEPBP	2	1	1	4	Uncharacterized protein GN=LEPBI_I2657 PE=4 SV=1
tr|B0SML3|B0SML3_LEPBP	5	2		7	Uncharacterized protein GN=LEPBI_I2658 PE=4 SV=1
tr|B0SMM4|B0SMM4_LEPBP		1	14	15	Uncharacterized protein GN=LEPBI_I2670 PE=4 SV=1
tr|B0SMW5|B0SMW5_LEPBP	3	8	2	13	Uncharacterized protein GN=LEPBI_I2761 PE=4 SV=1

tr|B0SNB8|B0SNB8_LEPBP	3	4		7	Uncharacterized protein GN=LEPBI_I2809 PE=4 SV=1
tr|B0SNC0|B0SNC0_LEPBP	5		5	10	Uncharacterized protein GN=LEPBI_I2812 PE=4 SV=1
tr|B0SNE7|B0SNE7_LEPBP	4	3		7	Uncharacterized protein GN=LEPBI_I2844 PE=4 SV=1
tr|B0SNH0|B0SNH0_LEPBP	5	4	5	14	Uncharacterized protein GN=LEPBI_I2868 PE=4 SV=1
tr|B0SNX8|B0SNX8_LEPBP	2	1	3	6	Uncharacterized protein GN=LEPBI_I2921 PE=4 SV=1
tr|B0SNY0|B0SNY0_LEPBP	5	4	4	13	Uncharacterized protein GN=LEPBI_I2923 PE=4 SV=1
tr|B0SQN4|B0SQN4_LEPBP	13	14	4	31	Uncharacterized protein GN=LEPBI_I3216 PE=4 SV=1
tr|B0SQP7|B0SQP7_LEPBP	6	4	2	12	Uncharacterized protein GN=LEPBI_I3229 PE=4 SV=1
tr|B0SQT5|B0SQT5_LEPBP	5	3	2	10	Uncharacterized protein GN=LEPBI_I3267 PE=4 SV=1
tr|B0SQV2|B0SQV2_LEPBP	9	5	9	23	Uncharacterized protein GN=LEPBI_I3284 PE=4 SV=1
tr|B0SQV4|B0SQV4_LEPBP	6	5	8	19	Uncharacterized protein GN=LEPBI_I3286 PE=4 SV=1
tr|B0SS00|B0SS00_LEPBP		6	17	23	Uncharacterized protein GN=LEPBI_I3476 PE=4 SV=1
tr|B0STN7|B0STN7_LEPBP	1	8		9	Uncharacterized protein GN=LEPBI_II0033 PE=4 SV=1
tr|B0STW4|B0STW4_LEPBP	4	6		10	Uncharacterized protein GN=LEPBI_II0112 PE=4 SV=1
tr|B0STX4|B0STX4_LEPBP	10	6	8	24	Uncharacterized protein GN=LEPBI_II0122 PE=4 SV=1
tr|B0SU67|B0SU67_LEPBP	3		4	7	Uncharacterized protein GN=LEPBI_II0218 PE=4 SV=1
tr|B0SU97|B0SU97_LEPBP	3	3	7	13	Uncharacterized protein GN=LEPBI_II0248 PE=4 SV=1

Tabla 11: Resultados del módulo estadístico Venn diagram del software PatternLab for Proteomics V utilizado para identificar las proteínas detectadas de forma exclusiva en cepas *L. biflexa flaA2*$^{-}$ vs *fcpA*$^{-}$ ($p<0,05$). Esto nos permitió la identificación de 461 proteínas exclusivas de cepas *flaA2*$^{-}$ de las cuales solo se muestran las proteínas no caracterizadas. Se destacan en amarillo aquellas que resultan ser flagelares. Estas proteínas acompañan más la perdida de FlaA2 del filamento que la de FcpA. Mientras que entre las 8 proteínas exclusivas de *fcpA*$^{-}$ que no se detectan en cepas *flaA2*$^{-}$*,* encontramos a FlaA1 y FlaA2 como únicas proteínas del filamento (resultado no mostrados).

Análisis SC de proteínas aumentadas en *flaA2*⁻ vs fcpA⁻			
Locus	**Fold Change**	**pValue**	**Description**
tr|B0SR03|B0SR03_LEPBP	-31,05	0,0055	Uncharacter+E86:E138ized protein GN=LEPBI_I1597 PE=1 SV=1
tr|B0SRM6|B0SRM6_LEPBP	-6,43	0,0004	Uncharacterized protein GN=LEPBI_I1715 PE=4 SV=1
tr|B0SPC0|B0SPC0_LEPBP	-6,38	0,0364	Uncharacterized protein GN=LEPBI_I1334 PE=4 SV=1
tr|B0SNK0|B0SNK0_LEPBP	-4,88	0,0125	Protoporphyrinogen oxidase (PPO) GN=hemG PE=4 SV=1
tr|B0STP2|B0STP2_LEPBP	-4,77	0,0047	Putative alkyldihydroxyacetonephosphate synthase GN=eapA PE=4 SV=1
tr|B0SSS4|B0SSS4_LEPBP	-4,75	0,0357	Lipid-A-disaccharide synthase GN=lpxB PE=4 SV=1
tr|B0SJP1|B0SJP1_LEPBP	-4,48	0,0061	Acyl-CoA dehydrogenase GN=acdA2 PE=3 SV=1
tr|B0SUD2|B0SUD2_LEPBP	-4,30	0,0021	Putative chromosome partitioning protein ParB GN=LEPBI_p0002 PE=3 SV=1
sp|B0SSI1|RS7_LEPBP	-4,23	0,0343	30S ribosomal protein S7 GN=rpsG PE=3 SV=1
tr|B0SPZ1|B0SPZ1_LEPBP	-4,18	0,0405	SPOR domain-containing protein GN=LEPBI_I1449 PE=4 SV=1
tr|B0SJG1|B0SJG1_LEPBP	-4,16	0,0059	Putative serine-type endopeptidase GN=LEPBI_I0523 PE=4 SV=1
tr|B0SML4|B0SML4_LEPBP	-4,13	0,0032	Uncharacterized protein GN=LEPBI_I2660 PE=4 SV=1
sp|B0SMC0|LPXK_LEPBP	-4,08	0,0384	Tetraacyldisaccharide 4'-kinase GN=lpxK PE=3 SV=1
tr|B0SM43|B0SM43_LEPBP	-3,89	0,0006	Uncharacterized protein GN=LEPBI_I2594 PE=4 SV=1
tr|B0SRR4|B0SRR4_LEPBP	-3,87	0,0380	UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine ligase GN=murF PE=3 SV=1
tr|B0SME8|B0SME8_LEPBP	-3,82	0,0201	Trigger factor GN=tig PE=3 SV=1
tr|B0STZ6|B0STZ6_LEPBP	-3,42	0,0089	Adenylate/guanylate cyclase GN=LEPBI_II0145 PE=4 SV=1
tr|B0STM5|B0STM5_LEPBP	-3,40	0,0038	NADH dehydrogenase GN=ndh PE=4 SV=1
tr|B0STV0|B0STV0_LEPBP	-3,27	0,0357	Rubrerythrin domain-containing protein GN=LEPBI_II0098 PE=4 SV=1
sp|B0SM57|TRPB_LEPBP	-3,21	0,0074	Tryptophan synthase beta chain GN=trpB PE=3 SV=1
tr|B0SJR6|B0SJR6_LEPBP	-3,20	0,0001	Putative tRNA (Uracil-5-)-methyltransferase GN=LEPBI_I0091 PE=3 SV=1
tr|B0STI6|B0STI6_LEPBP	-3,12	0,0033	OmpA-like domain-containing protein GN=LEPBI_I0255 PE=4 SV=1
tr|B0SLB3|B0SLB3_LEPBP	-3,05	0,0106	Signal peptidase I GN=lepB2 PE=3 SV=1
tr|B0SSM9|B0SSM9_LEPBP	-2,94	0,0164	ABC-type transport system, ATP binding protein GN=LEPBI_I2017 PE=3 SV=1
tr|B0SJF6|B0SJF6_LEPBP	-2,92	0,0054	Putative FAD-linked oxidase GN=LEPBI_I0538 PE=4 SV=1
tr|B0SKX3|B0SKX3_LEPBP	-2,84	0,0178	Hemolysin putative signal peptide GN=LEPBI_I2375 PE=4 SV=1
tr|B0SSB0|B0SSB0_LEPBP	-2,81	0,0362	Lysine--tRNA ligase GN=lysS PE=3 SV=1
tr|B0SJ29|B0SJ29_LEPBP	-2,78	0,0031	Mannose-1-phosphate guanylyltransferase [GDP] GN=manC PE=4 SV=1
tr|B0SSD4|B0SSD4_LEPBP	-2,74	0,0016	Putative diguanylate phosphodiesterase GN=LEPBI_I1921 PE=4 SV=1
sp|B0SPA9|GLPK_LEPBP	-2,73	0,0301	Glycerol kinase GN=glpK PE=3 SV=1
tr|B0SL67|B0SL67_LEPBP	-2,62	0,0130	Glutamate-1-semialdehyde 2,1-aminomutase GN=hemL2 PE=3 SV=1
tr|B0STL1|B0STL1_LEPBP	-2,60	0,0191	Uncharacterized protein GN=LEPBI_II0007 PE=4 SV=1
tr|B0SMF6|B0SMF6_LEPBP	-2,58	0,0142	Hydrolase_4 domain-containing protein GN=LEPBI_I0976 PE=4 SV=1
tr|B0SN68|B0SN68_LEPBP	-2,57	0,0076	Putative acyl-CoA dehydrogenase, GN=LEPBI_I1138 PE=3 SV=1
tr|B0SRR8|B0SRR8_LEPBP	-2,54	0,0354	UDP-N-acetylglucosamine transferase GN=murG PE=3 SV=1
tr|B0SMI0|B0SMI0_LEPBP	-2,54	0,0334	Glycosyltransferase GN=LEPBI_I1000 PE=4 SV=1
tr|B0SKU7|B0SKU7_LEPBP	-2,52	0,0250	Iron-sulfur cluster carrier protein GN=mrp PE=3 SV=1
tr|B0SR13|B0SR13_LEPBP	-2,43	0,0380	Peptidyl-prolyl cis-trans isomerase GN=LEPBI_I1607 PE=3 SV=1
tr|B0SM29|B0SM29_LEPBP	-2,42	0,0043	Uncharacterized protein GN=LEPBI_I2578 PE=4 SV=1
tr|B0SQ75|B0SQ75_LEPBP	-2,41	0,0332	Uncharacterized protein GN=LEPBI_I3162 PE=4 SV=1
tr|B0STF7|B0STF7_LEPBP	-2,40	0,0054	Acyl coenzyme A dehydrogenase GN=LEPBI_I2302 PE=3 SV=1
tr|B0SQL3|B0SQL3_LEPBP	-2,37	0,0054	30S ribosomal protein S21 GN=rpsU PE=3 SV=1
tr|B0SR55|B0SR55_LEPBP	-2,34	0,0312	ABC-type transport system, ATPase GN=LEPBI_I1649 PE=4 SV=1
tr|B0SNZ9|B0SNZ9_LEPBP	-2,31	0,0345	Aldehyde dehydrogenase GN=LEPBI_I2943 PE=3 SV=1
tr|B0SN63|B0SN63_LEPBP	-2,29	0,0111	Electron transfer flavoprotein alpha-subunit (ETFLS) GN=etfA PE=3 SV=1
tr|B0SP47|B0SP47_LEPBP	-2,22	0,0043	Acyl-CoA dehydrogenase family protein GN=LEPBI_I1259 PE=3 SV=1
sp|B0SLC8|ATPB_LEPBP	-2,21	0,0124	ATP synthase subunit beta GN=atpD PE=3 SV=1
tr|B0SPS9|B0SPS9_LEPBP	-2,17	0,0006	Poly(A) polymerase I GN=pcnB PE=3 SV=1
tr|B0STX2|B0STX2_LEPBP	-2,11	0,0014	Putative UDP-glucose 4-epimerase GN=LEPBI_II0120 PE=4 SV=1
tr|B0SQT0|B0SQT0_LEPBP	-2,08	0,0090	Uncharacterized protein GN=LEPBI_I3262 PE=4 SV=1
tr|B0STT2|B0STT2_LEPBP	-2,07	0,0169	Uncharacterized protein GN=LEPBI_II0078 PE=4 SV=1
tr|B0SMN8|B0SMN8_LEPBP	-2,06	0,0072	Cytochrome c oxidase subunit 2 GN=ctaC PE=3 SV=1
tr|B0SRD8|B0SRD8_LEPBP	-1,72	0,0037	Endolytic murein transglycosylase GN=mltG PE=3 SV=1

Tabla 12: Proteínas aumentadas en cepas *L. biflexa flaA2*⁻ comparada con cepas *fcpA*⁻. Para identificar las proteínas encontradas en ambas cepas pero que muestran una diferencia estadística en sus niveles (BH q-value < 0.05) se utilizó el módulo TFold de PatternLab for Proteomics V. Se identifica únicamente la proteína del filamento FcpB (código Uniprot B0SR03) que se expresa 31 veces menos en cepas *fcpA*⁻, que en cepas *flA2*⁻

ANEXO 2 – Tabla de entrecruzamientos químicos con NNP9

Muestras rep	Proteina 1	Peptido 1	Inicio	Fin	Sitio 1	Proteina 2	Peptido 2	Inicio	Fin	Sitio 2	Clasificación	Distancia	modelo	comentarios
3	B0SJC6	NFRGEDPEGKPSELPEGKEGFR	47	70	K56	B0STN7	DYSKESAESIKK	85	96	K88	inter XL	15.5	SI	
2	B0SJC6	NLEKHQKLQPSHR	215	227	K218	FlaB1	VLKFQNEEVSK	15	25	K17	inter XL	91.7	NO	
3	B0SJC6	TEASKTGASVEGQPVKIDGVTYDFTK	270	297	K274	B0SUE8	MEKTLMKQSEFSNLMR	1	16		n/c	NO	n/c	
1;3	B0SJC6	ELEPIVKNFRGEDPEGKPSELPEGK	40	66	K46	B0SJC6	VYELDSNNMVFMKYDQER	330	347	K342	intra XL	20.5	SI	
9	B0SJC6	ELEPIVKNFR	40	49	K46	B0SJC6	YLEAKR	74	79	K78	intra XL	14.1	SI	
3;5;9	B0SJC6	KALNVPPKPAQGSTTTPQQ	349	367	K349	B0SJC6	TEASKTGASVEGQPVK	270	287	K274	intra XL	17.2	SI	1
9	B0SJC6	ESVSSKSYQK	21	30	K26	B0SJC6	HQKLQPSHR	219	227	K221	intra XL	19.0	SI	
1;2	B0SJC6	IDGVTYDFTK	288	297	T292	B0SJC6	NPYIKYDNR	298	306	K302	intra XL	19.5	SI	
1;3;5;10;11;9	B0SJC6	VYELDSNNMVFMKYDQER	330	347	K342	B0SJC6	ELEPIVKNFR	40	49	K46	intra XL	23.3	SI	
3;4;5;10;11;9	B0SJC6	TEASKTGASVEGQPVKIDGVTYDFTK	270	297	K274					K286	intra XL	14.5	SI	
3	B0SJC6	NFRGEDPEGKPSELPEGKEGFRMKK	47	73	K56					K72	intra XL	12.9	SI	
2;3;4;5;10	B0SJC6	NFRGEDPEGKPSELPEGKEGFR	47	70	K56					K66	intra XL	5.2	SI	
1;3	B0SJC6	ELEPIVKNFRGEDPEGKPSELPEGK	40	66	K46					K56	intra XL	15.3	SI	
2;3;4;5;10	B0SJC6	NFRGEDPEGKPSELPEGKEGFR	47	70	K56					K66	intra XL	9.3	SI	
3;4;5;6;7;8;9	B0SJC6	GEDPEGKPSELPEGKEGFR	50	70	K56					K66	intra XL	9.3	SI	
1	B0SJC6	IDGVTYDFTKNPYIKYDNR	288	306	K302					S335	intra XL	23.8	SI	
1;2;3;5	B0SJC6	ALNVPPKPAQGSTTTPQQ	350	367	K356					T364	intra XL	NO	n/c	
3;9	B0SJC6	ALKVER	207	214	K212	B0SJC6	NLEKHQK	215	221	K218	loop XL	14.0	SI	
3	B0SJC6	TEASKTGASVEGQPVK	270	287	K274	B0SJC6	IDGVTYDFTK	288	297	T292	loop XL	10.6	SI	1
1;2;4;5;6;7;11;9	B0SJC6	TEASKTGASVEGQPVK	270	287	K274					S278	loop XL	NO	n/c	
1;2;3;4;5;10;9	B0SJC6	NFRGEDPEGKPSELPEGK	47	66	K56					S58	loop XL	13.2	SI	
1;2;3;4;5;6;7	B0SJC6	NLEKHQKLQPSHR	215	227	K218					K221	loop XL	9.1	SI	
5;6;7;11;9	B0SJC6	HQKLQPSHR	219	227	K221					S225	loop XL	11.7	SI	
1;5	B0SJC6	LKYPYDNYYLNNSQAK	254	269	K255					S266	loop XL	18.6	SI	
1	B0SJC6	IDGVTYDFTKNPYIKYDNR	288	306	K297					K302	loop XL	14.0	SI	
5	B0SJC6	VYELDSNNMVFMKYDQERKK	330	349	K342					K348	loop XL	10.7	SI	
10;11;9	B0SJC6	KKALNVPPKPAQGSTTTPQQ	348	367	K348					K349	loop XL	4.0	SI	
1;4;5	B0SJC6	ESVSSKSYQKR	21	31	K26					K30	loop XL	6.2	SI	
3;4;5	B0SJC6	EAHYSYNKYDIEK	178	190	S182					K185	loop XL	5.7	SI	
9	B0SJC6	NMELGYKHLGLAK	191	203	Y196					K197	loop XL	3.8	SI	
3	B0SJC6	AVNIYKLK	246	255	Y252					K253	loop XL	3.8	SI	
1	B0SJC6	NPYIKYDNR	298	306	Y300					K302	loop XL	7.3	SI	
1;3;5	B0SJC6	VYELDSNNMVFMKYDQER	330	347	S335					K342	loop XL	24.2	SI	
1;2;7	B0SJC6	TGASVEGQPVKIDGVTYDFTK	275	297	S278					K287	loop XL	NO	n/c	
5	B0SJC6	VYELDSNNMVFMKYDQERKK	330	349	S335					K348	loop XL	7.0	SI	
4;5	B0SJC6	HLGLAKEAR	198	206	K203						dead end	n/c	SI	
1;3;4;5;9	B0SJC6	ALKVER	207	214	K211						dead end	n/c	SI	
3;4;5;10	B0SJC6	NLEKHQK	215	221	K218						dead end	n/c	SI	
1;2;3;4;5;6;7;8;10;11;9	B0SJC6	ELEPIVKNFR	40	49	K46						dead end	n/c	SI	
1;2;3;4;5;10;11;9	B0SJC6	IDGVTYDFTKNPYIK	288	302	K297						dead end	n/c	SI	
3;4;5;6;10;11	B0SJC6	ESVSSKSYQK	21	30	K26						dead end	n/c	SI	
1;2	B0SJC6	ALLDISIEYGK	142	152	S147						dead end	n/c	SI	
1;2;5	B0SJC6	EAHYSYNK	178	185	S182						dead end	n/c	SI	
1;2	B0SJC6	TGASVEGQPVK	275	287	S280						dead end	n/c	SI	
1	B0SJC6	IYQEGLQYHFEGNFSSAFQR	80	99	S94						dead end	n/c	SI	
1;2	B0SJC6	RIELLTYLR	31	39	T36						dead end	n/c	SI	
1;2	B0SJC6	IDGVTYDFTK	288	297	T292						dead end	n/c	SI	
1;2	B0SJC6	MTEELSQLYILR	110	121	T111						dead end	n/c	SI	
1;2	B0SJC6	AVNIYK	246	253	Y252						dead end	n/c	SI	
3	B0SJC6	AVNIYKLK	246	255	K253						dead end	n/c	SI	
1	B0SJC6	NPYIKYDNR	298	306	K302						dead end	n/c	SI	
9	B0SJC6	NMELGYKHLGLAK	191	203	K197						dead end	n/c	SI	
5;6;7;11;9	B0SJC6	HQKLQPSHR	219	227	S225						dead end	n/c	SI	
1;5	B0SJC6	LKYPYDNYYLNNSQAK	254	269	S266						dead end	n/c	SI	
5	B0SJC6	VPEEYRVDHADVR	315	327	Y319						dead end	n/c	SI	
3;4;5;6;7;11;9	B0SJC6	GEDPEGKPSELPEGK	50	66	K56						dead end	n/c	SI	
1;3	B0SJC6	ELEPIVKNFRGEDPEGKPSELPEGK	40	66	K46						dead end	n/c	SI	
3	B0SKN8	ESQKQIDLQR	32	41	K35	B0SMK8	KNLDFDYFK	67	75	K67	inter XL	16.5	SI	
4	B0SKN8	EKPASATLK	21	29	K22	B0SMK8	NYEKVR	230	235	K233	inter XL	19.2	SI	3
9	B0SKN8	KIAKQTAFK	240	248	K240	B0SMK8	NYEKVR	230	235	K233	inter XL	17.6	SI	
2;10;9	B0SKN8	IAKQTAFK	241	248	K243	B0SMK8	NYEKVR	230	235	K233	inter XL	9.4	SI	
4;5;9	B0SKN8	EKPASATLK	21	29	K22	B0SMK8	SNDSKALLR	139	149	K145	inter XL	28.4	NO	2
1;2;3;9	B0SKN8	EKPASATLK	21	29	K22	B0SMK8	VRDTLINFIENKR	234	246	K245	inter XL	24.3	SI	2
10;11;9	B0SKN8	EKPASATLK	21	29	K22	B0SMK8	TVVPPGNPDAKPLDLLEQHDDNYGFITSKK	250	279	K260	inter XL	15.9	SI	2
3	B0SKN8	EKPASATLK	21	29	K22	B0SMK8	KNLDFDYFK	67	75	K67	inter XL	16.5	SI	2
3;5;10;11;9	B0SKN8	EKPASATLK	21	29	K22	B0SMK8	DTLINFIENKR	236	246	K245	inter XL	24.3	SI	
3;5	B0SKN8	SEGTEKKGEAR	166	178	K171	B0SQA9	GTVQADILK	763	771		n/c	NO	n/c	
1;2;3;4;5;10;9	B0SKN8	EKPASATLK	21	29	K22	B0SQV9	IIVSDKEK	210	217		n/c	NO	n/c	
2	B0SKN8	EKPASATLK	21	29	K22	B0SR40	LAYVAITR	591	598		n/c	NO	n/c	
2	B0SKN8	SVELDQKQMITK	251	262	K257	B0SSG1	AGIQDVLTK	117	125		n/c	NO	n/c	
3;4;5	B0SKN8	SEGTEKKGEAR	166	178	K174	B0SKN8	EKPASATLK	21	29	K22	inter XL	24.3	SI	2
3;4	B0SKN8	SEGTEKK	166	174	K171	B0SKN8	EKPASATLK	21	29	K22	inter XL	26.6	SI	2
3;4;5;10;11;9	B0SKN8	EKPASATLK	21	29	K22	B0SKN8	IAKQTAFK	241	248	K243	inter XL	21.2	SI	2
1;2;3;4;5;6;7;8;10;11;9	B0SKN8	SVELDQKQMITK	251	262	K257	B0SKN8	EKPASATLK	21	29	K22	inter XL	24.3	SI	2
3	B0SKN8	KLQDLGVDKTAAQYLAQTPQEHK	66	88	K66	B0SKN8	EKPASATLK	21	29	K22	inter XL	24.2	SI	2
1;2	B0SKN8	TAAQYLAQTPQEHKELYK	75	92	K88	B0SKN8	QTAFKLR	244	250	K248	inter XL	32	NO	
1;2;3;4;5;6;7;10;11;9	B0SKN8	KLQDLGVDK	66	74	K66	B0SKN8	EKPASATLK	21	29	K22	inter XL	22.5	SI	
2;3;4;5;10;9	B0SKN8	ESQKQIDLQR	32	41	K35	B0SKN8	EKPASATLK	21	29	K22	intra XL	12.1	SI	
3;10;11;9	B0SKN8	KNGFGDNEIDTLHANIIGNLR	42	62	K42	B0SKN8	EKPASATLK	21	29	K22	intra XL	21.3	SI	2
3;4	B0SKN8	SVELDQKQMITK	251	262	K257	B0SKN8	KLQDLGVDK	66	74	K66	intra XL	15.0	SI	
9	B0SKN8	KLQDLGVDKTAAQYLAQTPQEHK	66	88	K66	B0SKN8	KDKDGKPYLEIK	93	104	K93	intra XL	17.9	SI	
10;11;9	B0SKN8	LQDLGVDKTAAQYLAQTPQEHK	67	88	K74	B0SKN8	KDKDGKPYLEIK	93	104	K93	intra XL	18.0	SI	
7	B0SKN8	EKPASATLK	21	29	K22					T27	intra XL	NO	n/c	
2	B0SKN8	KNGFGDNEIDTLHANIIGNLRK	42	63	K42					T52	intra XL	21.8	SI	
1;2;3;5;10;9	B0SKN8	LQDLGVDKTAAQYLAQTPQEHK	67	88	K74					T75	intra XL	3.8	SI	
9	B0SKN8	KNGFGDNEIDTLHANIIGNLR	42	62	K42					T52	intra XL	25.6	SI	
3;5;10;11;9	B0SKN8	KNGFGDNEIDTLHANIIGNLR	42	62	K42	B0SKN8	ESQKQIDLQR	31	41	K35	loop XL	10.4	SI	
3;5;9	B0SKN8	KLQDLGVDKTAAQYLAQTPQEHK	66	88	K66					K74	loop XL	18.2	SI	
1;2;3;5;10;9	B0SKN8	LQDLGVDKTAAQYLAQTPQEHK	67	88	K74					T83	loop XL	19.8	SI	
1;2;3;4;5;10	B0SKN8	SVELDQKQMITK	251	262	S251					K257	loop XL	11.8	SI	
1	B0SKN8	LRSVELDQKQMITK	249	262	S251					K257	loop XL	11.8	SI	
1;2;6;10;11;9	B0SKN8	TAAQYLAQTPQEHKELYK	75	92	T75					K88	loop XL	12.5	SI	
1;2;3;4	B0SKN8	EKPASATLKER	21	31	K22					K29	loop XL	NO	n/c	
1;2	B0SKN8	QKHIMMSYK	227	235	K228					S233	loop XL	8.6	SI	
1;2;3;4;5;10;11;9	B0SKN8	KIAKQTAFK	240	248	K240					K243	loop XL	5.1	SI	
3;5	B0SKN8	IAKQTAFKLR	241	250	K243					K248	loop XL	8.6	SI	
3;4;5	B0SKN8	IKKLQDLGVDK	64	74	K65					K66	loop XL	3.8	SI	
3;5;9	B0SKN8	KLQDLGVDKTAAQYLAQTPQEHK	66	88	K66					K74	loop XL	13	SI	

1;2;3;4;5;10;11;9	B0SKN8	TAAQYLAQTPQEHKELYKK	75	93	K88					K92	loop XL	9.5	SI	
11;9	B0SKN8	KDKDGKPYLEIK	93	104	K93					K95	loop XL	6.8	SI	
5	B0SKN8	VDGTIHVKEMR	145	157	T150					K154	loop XL	12.3	SI	
1;2;3;5;7;10;11;9	B0SKN8	EDYSDLDKIILSFR	130	143	K137						dead end	n/c	SI	
2	B0SKN8	RVDGTIHVKEMR	144	157	K154						dead end	n/c	SI	
1;2	B0SKN8	ESQKQIDLQR	32	41	K35						dead end	n/c	SI	
1;2	B0SKN8	IAKQTAFK	241	248	K243						dead end	n/c	SI	
1;2;3;4;5;6;7;9	B0SKN8	SVELDQKQMITK	251	262	K257						dead end	n/c	SI	
2;5	B0SKN8	KLQDLGVDK	66	74	K66						dead end	n/c	SI	
1;2	B0SKN8	DGKPYLEIK	96	104	K98						dead end	n/c	SI	
1	B0SKN8	VDGTIHVK	145	154	T150						dead end	n/c	SI	
7	B0SKN8	EKPASATLK	21	29	K22						dead end	n/c	SI	
1;2;3;4	B0SKN8	EKPASATLKER	21	31	K29						dead end	n/c	SI	
5	B0SKN8	VDGTIHVKEMR	145	157	K154						dead end	n/c	SI	
1;2	B0SKN8	NGFGDNEIDTLHANIIGNLRK	43	63	T52						dead end	n/c	SI	
1;2	B0SKN8	NGFGDNEIDTLHANIIGNLR	43	62	T52						dead end	n/c	SI	
1;2	B0SKN8	TAAQYLAQTPQEHK	75	88	T75						dead end	n/c	SI	
3;4;5;9	B0SMK8	SNDSKALLR	139	149	K145	B0SMT6	FKSLHIQGK	68	76	K69	inter XL	14.1	SI	4
3;4;5	B0SMK8	SNEDLKEAR	215	223	K220	B0SK77	KIAK	69	72		n/c	NO	n/c	
3	B0SMK8	TVVPPGNPDAKPLDLLEQHDDNYGFITSKK	250	279	K260	B0SQX4	EKEIQGIK	11	18		n/c	NO	n/c	
3;4;5	B0SMK8	SNEDLKEAR	215	223	K220	B0SRN6	KLAK	446	449		n/c	NO	n/c	
5	B0SMK8	SNDSKALLR	139	149	K145	B0SQV9	VSYGLKDNYEFLR	33	45		n/c	NO	n/c	
1;2	B0SMK8	SNDSKALLR	139	149	K145	B0SMK8	YIEDAR	121	126	Y121	intra XL	21.9	SI	
10;11;9	B0SMK8	KLDLLLEAQIKETEGAR	279	297	K279					K291	intra XL	22.4	SI	
1;2	B0SMK8	LDLLLEAQIKETEGARR	280	298	K291					T293	loop XL	5.9	SI	
9	B0SMK8	KNLDFDYFKLLK	67	78	K67					K75	loop XL	12.0	SI	
1	B0SMK8	LLKAANQSDFSGNMWYLQSNYVYGFR	76	101	K78					S83	loop XL	8.5	SI	
9	B0SMK8	LTLYKEGILTLR	177	188	T178					K181	loop XL	5.1	SI	
1	B0SMK8	LDLLLEAQIKETEGAR	280	297	K291					T293	loop XL	12.2	SI	
1;2;10;11;9	B0SMK8	QAQGELKNIFEIVLQK	105	120	K111						dead end	n/c	SI	
9	B0SMK8	NIFEIVLQKYIEDAR	112	126	K120						dead end	n/c	SI	
1;3;10;11;9	B0SMK8	SNDSKALLR	139	149	K145						dead end	n/c	SI	
1	B0SMK8	TPDKPEYQYR	203	214	K208						dead end	n/c	SI	
2;4;5;9	B0SMK8	SNEDLKEAR	215	223	K220						dead end	n/c	SI	
10;9	B0SMK8	VRDTLINFIENKR	234	246	K245						dead end	n/c	SI	
9	B0SMK8	DTLINFIENKR	236	246	K245						dead end	n/c	SI	
1	B0SMK8	TVVPPGNPDAKPLDLLEQHDDNYGFITSK	250	278	K260						dead end	n/c	SI	
1	B0SMK8	FNGKAIYPEEK	307	319	K312						dead end	NO	n/c	
1;10;11;9	B0SMK8	NLDFDYFKLLK	68	78	K75						dead end	n/c	SI	
3;4;5;10;11;9	B0SMK8	SSEDLYTTGLNSSPHQYR	157	174	S157						dead end	n/c	SI	
2	B0SMK8	AANQSDFSGNMWYLQSNYVYGFR	79	101	S83						dead end	n/c	SI	
1;2	B0SMK8	EGILTLR	182	188	T186						dead end	n/c	SI	
10;11;9	B0SMK8	KLDLLLEAQIKETEGAR	279	297	K291						dead end	n/c	SI	
9	B0SMK8	KNLDFDYFKLLK	67	78	K75						dead end	n/c	SI	
9	B0SMK8	LTLYKEGILTLR	177	188	K181						dead end	n/c	SI	
1	B0SMK8	LLKAANQSDFSGNMWYLQSNYVYGFR	76	101	S83						dead end	n/c	SI	
1;2	B0SMK8	NLDFDYFK	68	75	Y73						dead end	n/c	SI	
10;9	B0SPP1	ADEKNESVSQSDATK	20	34	K23	B0SPF0	TINLDPKTEPEIYAAIRR	275	292		n/c	NO	n/c	
3;4;5;7;10;11;9	B0SPP1	ADEKNESVSQSDATK	20	34	K23	B0SPP1	VKFLSFEK	72	79	K73	inter XL	29.2	SI	2
3	B0SPP1	IGKEVDDYHAMINEK	57	71	K59	B0SPP1	ADEKNESVSQSDATK	20	34	K21	inter XL	26.2	SI	
9	B0SPP1	YIQAKGNLETK	211	221	K215	B0SPP1	FLSFEKK	74	80	K79	intra XL	12.3	SI	
3;4;5;11;9	B0SPP1	LNPDGSIALIPYKQQQK	35	53	K49	B0SPP1	ADEKNESVSQSDATK	20	34	K21	intra XL	11.6	SI	2
3;4;5;10;11;9	B0SPP1	ADEKNESVSQSDATK	20	34	K23	B0SPP1	NNTTGGVKGLVLSVR	153	167	K160	intra XL	26.0	SI	2
7	B0SPP1	IGKEVDDYHAMINEK	57	71	K59	B0SPP1	NNTTGGVKGLVLSVR	153	167	K160	intra XL	11.6	SI	
3;5;10;11;9	B0SPP1	ADEKNESVSQSDATKLNPDGSIALIPYNAK	20	49	K23					K34	intra XL	18.5	SI	2
3;4;5	B0SPP1	ADEKNESVSQSDATK	20	34	A20					S30	loop XL	NO	n/c	
1;2;3;10;9	B0SPP1	SLTEKGKDPVTTYR	136	149	K140					K142	loop XL	6.0	SI	
9	B0SPP1	ELKNNTTGGVKGLVLSVR	150	167	K152					K160	loop XL	10.9	SI	2
1;2	B0SPP1	GKDPVTTYR	141	149	K142						dead end	n/c	SI	
3;5	B0SPP1	ELKNNTTGGVK	150	160	K152						dead end	n/c	SI	
1;2;3;4;5;10;11;9	B0SPP1	NNTTGGVKGLVLSVR	153	167	K160						dead end	n/c	SI	
1;2	B0SPP1	ADEKNESVSQSDATKLNPDGSIALIPYNAK	20	49	K34						dead end	n/c	SI	
5	B0SPP1	TVTDDNTLNFELEKIQSPWER	168	190	K183						dead end	n/c	SI	
1;2;4	B0SPP1	YIQAKGNLETK	211	221	K215						dead end	n/c	SI	
1;2;3;5;9	B0SPP1	LNPDGSIALIPYKQQQK	35	53	K49						dead end	n/c	SI	
2	B0SPP1	VKFLSFEK	72	79	K73						dead end	n/c	SI	
2;9	B0SPP1	LKGGGGAEGGGFSLISFWSR	113	134	K114						dead end	n/c	SI	
1;2	B0SPP1	GGGGAEGGGFSLISFWSR	115	134	S125						dead end	n/c	SI	
1;2	B0SPP1	GLVLSVR	161	167	S165						dead end	n/c	SI	
1;2	B0SPP1	IQSPWER	184	190	S186						dead end	n/c	SI	
1;2	B0SPP1	LNPDGSIALIPYK	35	49	S40						dead end	n/c	SI	
3;4;5	B0SPP1	ADEKNESVSQSDATK	20	34	K23						dead end	n/c	SI	
3;5;10;11;9	B0SPP1	ADEKNESVSQSDATKLNPDGSIALIPYNAK	20	49	K34						dead end	n/c	SI	
1;2;3;10;9	B0SPP1	SLTEKGKDPVTTYR	136	149	K142						dead end	n/c	SI	
11;9	B0SPP1	EIHLPYEPSYVMHSR	94	108	S102						dead end	n/c	SI	
1;2;3;5;6;7;10;11;9	B0STF2	ALPFTLFGETGPSKEK	184	199	K197	FlaA2	KLTATVPGFIPQSTR	174	188	K174	inter XL	16.9	SI	
11;9	B0STF2	ALPFTLFGETGPSKEK	184	199	K197	FlaA2	GLIKDVFDEK	79	88	K82	inter XL	13.9	SI	
10;9	B0STF2	ALPFTLFGETGPSKEK	184	199	K197	FlaA2	TVPEDGGDKIEK	89	100	K97	inter XL	13.9	SI	
3;9	B0STF2	ALPFTLFGETGPSKEK	184	199	K197	FlaA2	DVFKTVPEDGGDKIEK	85	100	K88	inter XL	10.6	SI	
3;5;10;11;9	B0STF2	ALPFTLFGETGPSKEK	184	199	K197	B0STF2	QVIYNYVGQTAK	200	213	T209	inter XL	12.8	SI	
2	B0STF2	TLYFSNIMPTPFTAKEIWGEETYAK	124	148	S128	B0STF2	TLYFSNIMPTPFTAKEIWGEETYAK	124	148	K138	inter XL	12.9	SI	
1;2	B0STF2	EKQVIYNYVGQTAKK	198	214	K199	B0STF2	EKQVIYNYVGQTAKK	198	214	K213	inter XL	20.9	SI	
2	B0STF2	KPTNSLLFR	264	272	K264	B0STF2	VAILVFDTK	255	263	T262	intra XL	5.6	SI	
1;2	B0STF2	KFIDISDLK	273	281	K273	B0STF2	KPTNSLLFR	264	272	K264	intra XL	19.3	SI	
9	B0STF2	HADTGKGEFLDVQVEIESR	62	80	K67	B0STF2	QVHETLPKHTYTIYNNK	216	232	K223	intra XL	13.2	SI	
2	B0STF2	DNRDLYPLSMYDAR	37	50	S45	B0STF2	KFIDISDLK	273	281	K273	intra XL	9.8	SI	
1;2;3;5;10;11;9	B0STF2	KFIDISDLK	273	281	K273	B0STF2	LKNVTFVR	53	60	K54	intra XL	12.1	SI	
3;4;5;6;7;8;10;9	B0STF2	QESGNPQGTQVSAASSKDNR	20	39	Q20					K36	intra XL	NO	n/c	
1;3;5;9	B0STF2	NNAKALPFTLFGETGPSK	180	197	K183					S196	intra XL	25.2	SI	
1;2;3;4;5;6;7;10;11;9	B0STF2	EKQVIYNYVGQTAK	198	213	K199					T209	intra XL	23.1	SI	
2	B0STF2	HTYTIYNNK	224	232	T225	B0STF2	RQVHETLPK	215	223	T220	loop XL	10.9	SI	
5	B0STF2	KKEEMEKR	149	156	K149					K150	loop XL	3.8	SI	
3;5;9	B0STF2	ALPFTLFGETGPSKEKQVIYNYVGQTAK	184	213	K197					K199	loop XL	5.4	SI	
1;2;3;4;5;6;7;8;10;11;9	B0STF2	QVIYNYVGQTAKKR	200	215	K213					K214	loop XL	3.8	SI	
1;2	B0STF2	YKTTITSHHYTQYRPNFLSFNK	233	254	K234					S239	loop XL	16.3	SI	
1;2;3;10;11;9	B0STF2	VAILVFDTKKPTNSLLFR	255	272	K263					K264	loop XL	3.8	SI	
9	B0STF2	KFIDISDLK	273	281	K273					S278	loop XL	16.0	SI	
3;4;5;6;7	B0STF2	KADPEKR	115	123	K115					K120	loop XL	12.1	SI	
1;3;4;5;10;11;9	B0STF2	ALPFTLFGETGPSKEK	184	199	S196					K197	loop XL	3.8	SI	
1;2	B0STF2	QVIYNYVGQTAKK	200	214	T209					K213	loop XL	6.4	SI	
3;4;5;10;11;9	B0STF2	QVHETLPKHTYTIYNNK	216	232	T220					K223	loop XL	5.6	SI	

1	B0STF2	HTYTIYNNKYK	224	234	T225					K232	loop XL	23.3	SI	
1	B0STF2	QVIYNYVGQTAK	200	213	Y203					T209	loop XL	12.1	SI	
2	B0STF2	DNRDLYPLSMYDAR	37	50	Y42					S45	loop XL	7.0	SI	
11;9	B0STF2	EIWGEETYAKK	139	149	K148						dead end	n/c	SI	
1	B0STF2	RQVHETLPKHTYTIYNNK	215	232	K223						dead end	n/c	SI	
2;4;5	B0STF2	KPTNSLLFR	264	272	K264						dead end	n/c	SI	
1;2	B0STF2	LKNVTFVR	53	60	K54						dead end	n/c	SI	
1;2	B0STF2	ALPFTLFGETGPSK	184	197	S196						dead end	n/c	SI	
1;2	B0STF2	TTITSHHYTQYRPNFLSFNK	235	254	S239						dead end	n/c	SI	
1;2	B0STF2	FIDISDLK	274	281	S278						dead end	n/c	SI	
1;2	B0STF2	EIWGEETYAK	139	148	T145						dead end	n/c	SI	
9	B0STF2	QVHETLPK	216	223	T220						dead end	n/c	SI	
1;2;5	B0STF2	HTYTIYNNK	224	232	T225						dead end	n/c	SI	
1;2	B0STF2	VAILVFDTK	255	263	T262						dead end	n/c	SI	
3;4;5;6;7	B0STF2	KADPEKR	115	123	K120						dead end	n/c	SI	
2	B0STF2	TLYFSNIMPTPFTAKEIWGEETYAK	124	148	K138						dead end	n/c	SI	
1;3;4;5;10;11;9	B0STF2	ALPFTLFGETGPSKEK	184	199	K197						dead end	n/c	SI	
1	B0STF2	HTYTIYNNKYK	224	234	K232						dead end	n/c	SI	
3;4;5;10;11;9	B0STF2	QVHETLPKHTYTIYNNK	216	232	K223						dead end	n/c	SI	
1;2	B0STF2	YKTTITSHHYTQYRPNFLSFNK	233	254	S239						dead end	n/c	SI	
1;2;3;10;11;9	B0STF2	VAILVFDTKKPTNSLLFR	255	272	K264						dead end	n/c	SI	
9	B0STF2	KFIDISDLK	273	281	S278						dead end	n/c	SI	
5	B0STF2	KKEEMEKR	149	156	K150						dead end	n/c	SI	
1;3;5;9	B0STF2	NKALPFTLFGETGPSK	180	197	S196						dead end	n/c	SI	
1;2	B0STF2	EKQVIYNYVGQTAKK	198	214	K213						dead end	n/c	SI	
1;2;3;4;5;6;7;10;11;9	B0STF2	EKQVIYNYVGQTAK	198	213	T209						dead end	n/c	SI	
1;2;3;4;5;6;7;8;10;11;9	B0STF2	QVIYNYVGQTAKKR	200	215	T214						dead end	n/c	SI	
10;9	B0STF2	KKEEMEK	149	155	K149						dead end	n/c	SI	
3	B0STN7	AKLEYNLIKTLK	67	78	K68	B0STN7	ESAESIKK	89	96	K95	intra XL	18.6	SI	
3	B0STN7	KIDATNIQYER	96	106	K96	B0STN7	DYSKESAESIK	85	95	K88	loop XL	13.3	SI	
3;5;10;11;9	B0STN7	QSHLVFSGKDKFEITIPKP	157	175	K165					K167	loop XL	7.2	SI	
3;4;5;10;11;9	B0STN7	AKLEYNLIKTLKK	67	79	K68					K75	loop XL	10.5	SI	
2;3;4;5;6;7;10;11;9	B0STN7	LEYNLIKTLKK	69	79	K75					K78	loop XL	5.3	SI	
11;9	B0STN7	DYSKESAESIKK	85	96	K88					K95	loop XL	11.3	SI	
3;4;5	B0STN7	ESAESIKKIDATNIQYER	89	106	K95					K96	loop XL	3.8	SI	
5	B0STN7	KIDATNIQYER	96	106	K96					T100	loop XL	10.1	SI	
1;2;3;4;5;6;7;10;11;9	B0STN7	QDTKPQSTSTDVTQENHDAK	24	43	Q24					K27	loop XL	NO	n/c	
1	B0STN7	QSHLVFSGKDK	157	167	S158					K165	loop XL	19.3	SI	
11	B0STN7	DKFEITIPKP	166	175	K167						dead end	n/c	SI	
1;2	B0STN7	FEITIPKP	168	175	K174						dead end	n/c	SI	
9	B0STN7	EGQDKELFEYYYKTRPENLPPNK	44	66	K48						dead end	n/c	SI	
1;2	B0STN7	AKLEYNLIK	67	75	K68						dead end	n/c	SI	
1;2;11;9	B0STN7	LEYNLIKTLK	69	78	K75						dead end	n/c	SI	
2;4	B0STN7	DYSKESAESIK	85	95	K88						dead end	n/c	SI	
1;2;3;4;5	B0STN7	ESAESIKK	89	96	K95						dead end	n/c	SI	
1;2	B0STN7	YMCFVTFDVNPENYLQQSR	138	156	S155						dead end	n/c	SI	
1;2;3;4;5	B0STN7	QSHLVFSGK	157	165	S158						dead end	n/c	SI	
1	B0STN7	ESAESIK	89	95	S90						dead end	n/c	SI	
1;2	B0STN7	IDATNIQYER	97	106	T100						dead end	n/c	SI	
1;2	B0STN7	ELFEYYYK	49	56	Y53						dead end	n/c	SI	
5	B0STN7	KIDATNIQYER	96	106	T100						dead end	n/c	SI	
1	B0STN7	QSHLVFSGKDK	157	167	K165						dead end	n/c	SI	
1;2	B0STN7	LEYNLIK	69	75	Y71						dead end	n/c	SI	
1;2	B0STN7	EGQDKELFEYYYK	44	56	K48						dead end	n/c	SI	
1;9	FcpA	LALSNYQSGVNTIMKNR	102	118	K116	B0SKN8	IAKQTAFK	241	248	K243	inter XL	15.0	SI	
1;2	FcpA	LALSNYQSGVNTIMKNR	102	118	K116	B0SKN8	QTAFKLR	244	250	K248	inter XL	11.1	SI	
9	FcpA	QDALNYFSKAINHLIKNPDLR	162	184	K170	B0SMK8	SNEDLKEAR	215	223	K220	inter XL	13.4	SI	
3;10;9	FcpA	EKMK	156	159	K157	B0SPP1	GKDPVTTYR	141	149	K142	inter XL	11.5	SI	
1;2;3;4;5;6;7;11;9	FcpA	AKQDALNYFSK	160	170	K161	B0SPP1	GKDPVTTYR	141	149	K142	inter XL	12.4	SI	
1	FcpA	TKGPDDMLFKYR	250	261	K251	B0SPP1	KSLTEK	135	140	K135	inter XL	26.5	SI	
9	FcpA	GPDDMLFKYR	252	261	K259	B0SPP1	KSLTEK	135	140	K135	inter XL	12.3	SI	
3;5	FcpA	NPDFKGYNK	75	83	K79	B0SPP1	GKDPVTTYR	141	149	K142	inter XL	9.3	SI	
3;9	FcpA	AINHLIKNPDLR	171	184	K179	B0STF2	KADPEKR	115	123	K115	inter XL	13.1	SI	
1;2;3;4;5;10;11;9	FcpA	TKGPDDMLFKYR	250	261	K251	B0STF2	KFIDISDLK	273	281	K273	inter XL	19.8	SI	
1;2;3;5;6;7;8;10;11;9	FcpA	GPDDMLFKYR	252	261	K259	B0STF2	KFIDISDLK	273	281	K273	inter XL	11.8	SI	
11;9	FcpA	AKQDALNYFSKAINHLIKNPDLR	160	184	K161	FcpB	SSKNLGAK	120	127	K122	inter XL	24.9	SI	
11;9	FcpA	QDALNYFSKAINHLIKNPDLR	162	184	K170	FcpB	SSKNLGAK	120	127	K122	inter XL	11.7	SI	
1;2;7;10;11;9	FcpA	DQVLLKGELVPENDDAELTEDQKK	36	61	K43	FcpB	KMELFFEGSNDPDPR	130	146	K130	inter XL	21.9	SI	
1;2	FcpA	ERPEFKR	185	191	K190	FcpB	KQFYFK	235	240	K235	inter XL	15.8	SI	
1;2	FcpA	TKGPDDMLFKYR	250	261	K251	FcpB	NLGAKFKK	123	130	K127	inter XL	25.7	SI	
1;2;10;11;9	FcpA	TKGPDDMLFKYR	250	261	K251	FcpB	GSLKY	273	277	K276	inter XL	22.8	SI	
1;2;3;4;5;6;7;8;10;11;9	FcpA	GPDDMLFKYR	252	261	K259	FcpB	NLGAKFK	123	129	K127	inter XL	15.2	SI	
1;2	FcpA	GPDDMLFKYR	252	261	K259	FcpB	GSLKY	273	277	K276	inter XL	9.3	SI	
11;9	FcpA	YGKDSPEYKHIVNVINRVISVAQ	276	300	K278	FcpB	YILSNVENTKSNPIR	216	230	K225	inter XL	15.7	SI	
1;2;9	FcpA	DSPEYKHIVNVINRVISVAQ	279	300	K284	FcpB	YILSNVENTKSNPIR	216	230	K225	inter XL	12.4	SI	
1;2	FcpA	KEIMEQESLWKNPDFK	64	79	K64	FcpB	IFDYNDRDSNKHYK	252	265	K262	inter XL	22.0	SI	
1;2	FcpA	KEIMEQESLWKNPDFK	64	79	K64	FcpB	KNVEALK	266	272	K266	inter XL	23.0	SI	
1;2;3;4;5;6;7;8;10;11;9	FcpA	NPDFKGYNK	75	83	K79	FcpB	SSKNLGAK	120	127	K122	inter XL	16.2	SI	
1;2;6;7;8;10;9	FcpA	GYNKTFQELHQLSK	80	93	K83	FcpB	SSKNLGAK	120	127	K122	inter XL	16.0	SI	
1	FcpA	TFQELHQLSKTFANNQFR	84	101	K93	FcpB	KMELFFEGSNDPDPR	130	146	K130	inter XL	21.1	SI	
1;2;3;6;7;10;11;9	FcpA	TFQELHQLSKTFANNQFR	84	101	K93	FcpB	YILSNVENTKSNPIR	216	230	K225	inter XL	18.9	SI	
1;2;4;5;10;11	FcpA	KEEAEKK	126	132	K126	FcpB	KNVEALK	266	272	K266	inter XL	11.6	SI	
7	FcpA	KEEAEK	126	131	K126	FcpB	NLGAKFK	123	129	K127	inter XL	20.6	SI	
1;2;6;7;8;9	FcpA	DSPEYKHIVNVINR	279	292	K284	FcpB	YILSNVENTKSNPIR	216	230	K225	inter XL	18.3	SI	
3;5;7;9	FcpA	KEEAEK	126	131	K126	FcpB	IFDYNDRDSNKHYK	252	265	K262	inter XL	15.1	SI	
9	FcpA	EKMK	156	159	K157	FcpB	KNVEALK	266	272	K266	inter XL	21.7	SI	
1;2;3;4;6;10;11	FcpA	KEEAEK	126	131	K126	FcpB	KNVEALK	266	272	K266	inter XL	12.1	SI	
1	FcpA	KEEAEKK	126	132	K126	FcpB	NVEALKGSLKY	267	277	K272	inter XL	12.7	SI	
1	FcpA	KEEAEKK	126	132	K126	FcpB	NVEALKGSLK	267	276	K272	inter XL	12.7	SI	
7;10;11;9	FcpA	KEEAEK	126	131	K126	FcpB	NVEALKGSLK	267	276	K272	inter XL	12.7	SI	
1;2	FcpA	YGKDSPEYK	276	284	K278	FcpB	SMADTEKELDDNISEVNKR	40	58	K46	inter XL	27.1	SI	
1;2;7;11;9	FcpA	DQVLLKGELVPENDDAELTEDQK	36	60	K43	FcpB	KMELFFEGSNDPDPR	130	146	K130	inter XL	18.4	SI	
1;2;3;4;5;6;7;8;10;11;9	FcpA	TKGPDDMLFKYR	250	261	K259	FcpB	NLGAKFK	123	129	K127	inter XL	24.2	SI	
1;2	FcpA	GPDDMLFKYR	252	261	K259	FcpB	NLGAKFKK	123	130	K127	inter XL	10.8	SI	
1;2	FcpA	LALSNYQSGVNTIMKNR	102	118	K116	B0SK77	KIAK	69	72		n/c	NO	n/c	
2	FcpA	KEIMEQESLWKNPDFK	64	79	K64	B0SKZ5	SAVMK	37	41		n/c	NO	n/c	
7;8;9	FcpA	AKQDALNYFSK	160	170	K161	B0SKZ5	SAVMK	37	41		n/c	NO	n/c	
1;2	FcpA	TKGPDDMLFKYR	250	261	K251	B0SSV0	GKDYK	113	117		n/c	NO	n/c	
2	FcpA	YLASAYSFEENMIK	235	248	S238	B0SKR3	SLGEDHLK	7	14		n/c	NO	n/c	
1	FcpA	AINHLIKNPDLR	171	184	K179	B0SKT9	EKEKAEK	443	449		n/c	NO	n/c	
6	FcpA	EIMEQESLWKNPDFK	65	79	K74	B0SL99	VAESK	45	49		n/c	NO	n/c	
3;5	FcpA	KEEAEK	126	131	K126	B0SLF1	KYHK	209	212		n/c	NO	n/c	
3;4;5	FcpA	NPDFKGYNK	75	83	K79	B0SLH7	KEMK	84	87		n/c	NO	n/c	
2	FcpA	KEIMEQESLWKNPDFK	64	79	K64	B0SLH7	KEMK	84	87		n/c	NO	n/c	
9	FcpA	EIMEQESLWKNPDFK	65	79	K74	B0SLH7	KEMK	84	87		n/c	NO	n/c	
1	FcpA	TKGPDDMLFKYR	250	261	K251	B0SMH4	SLHYLK	879	884		n/c	NO	n/c	

1	FcpA	GPDDMLFKYR	252	261	K259	B0SNV6	EITKSK	280	285		n/c	NO	n/c	
1;2	FcpA	EYPAHKYLASAYSFEENMIK	229	248	K234	B0SP50	SLVDKYGK	297	304		n/c	NO	n/c	
3;11	FcpA	EIMEQESLWKNPDFK	65	79	K74	B0SP50	SLVDKYGK	297	304		n/c	NO	n/c	
3;4;5	FcpA	AKQDALNYFSK	160	170	K161	B0SP98	EKMK	305	308		n/c	NO	n/c	
1;2	FcpA	KTKGPDDMLFKYR	249	261	K249	B0SQJ3	VQKFK	78	82		n/c	NO	n/c	
5	FcpA	AINHLIKNPDLR	171	184	K179	B0SR17	AVGSPLNDVSK	238	248		n/c	NO	n/c	
1	FcpA	TKGPDDMLFKYR	250	261	K251	B0SR44	TKYR	35	38		n/c	NO	n/c	
1;2	FcpA	LALSNYQSGVNTIMKNR	102	118	K116	B0SSK7	KALK	450	453		n/c	NO	n/c	
2	FcpA	KEIMEQESLWKNPDFK	64	79	K64	B0SSM1	EKSVVYGK	139	146		n/c	NO	n/c	
2	FcpA	EYPAHKYLASAYSFEENMIK	229	248	K234	B0SSM1	EKSVVYGK	139	146		n/c	NO	n/c	
4	FcpA	AINHLIKNPDLR	171	184	K179	B0SSS0	AVSDDVK	412	418		n/c	NO	n/c	
2	FcpA	KEEAEKK	126	132	K126	B0SST3	KEKEALK	206	212		n/c	NO	n/c	
1	FcpA	NPDFKGYNK	75	83	K79	B0SSW6	DKQKSNK	562	568		n/c	NO	n/c	
1	FcpA	AINHLIKNPDLR	171	184	K179	B0STX0	SIEGLYFSDR	25	34		n/c	NO	n/c	
1;2	FcpA	TFQELHQLSK	84	93	S92	B0SL40	NKEPDLHSIQK	888	898		n/c	NO	n/c	
1;2;3;5;10;11;9	FcpA	AKQDALNYFSK	160	170	K161	FcpA	NPDFKGYNK	75	83	K79	inter XL	15.4	SI	
3;5;10;9	FcpA	AINHLIKNPDLR	171	184	K179	FcpA	AKQDALNYFSK	160	170	K166	inter XL	29.1	NO	
3;5;7	FcpA	AINHLIKNPDLR	171	184	K179	FcpA	KEEAEK	126	131	K126	inter XL	15.2	SI	
9	FcpA	YLASAYSFEENMIKK	235	249	K248	FcpA	GYNKTFQELHQLSK	80	93	K83	inter XL	17.5	SI	
1;2	FcpA	KEIMEQESLWKNPDFK	64	79	K64	FcpA	ATELKYGK	271	278	K275	inter XL	11.5	SI	
1;2;6;7;9	FcpA	TFQELHQLSKTFANNQFR	84	101	K93	FcpA	ATELKYGK	271	278	K275	inter XL	13.7	SI	
2	FcpA	TFANNQFR	94	101	T94	FcpA	ATELKYGK	271	278	K275	inter XL	16.9	SI	
1;2	FcpA	EYPAHKYLASAYSFEENMIKK	229	249	K234	FcpA	ATELKYGK	271	278	K275	intra XL	9.9	SI	
1;2	FcpA	ATELKYGK	271	278	K275	FcpA	KNVHLLR	264	270	K264	intra XL	16.6	SI	
10	FcpA	GYNKTFQELHQLSK	80	93	K83	FcpA	AINHLIKNPDLR	171	184	K179	intra XL	17.0	SI	
1;2;7	FcpA	EYPAHKYLASAYSFEENMIK	229	248	K234	FcpA	ATELKYGK	271	278	K275	intra XL	9.2	SI	
3;5;9	FcpA	DQVLLKGELVPENDDAELTEDQKKKKK	36	64	K43					K62	intra XL	NO	n/c	
5	FcpA	DQVLLKGELVPENDDAELTEDQKKKK	36	63	K43					K62	intra XL	NO	SI	
5;10;9	FcpA	DQVLLKGELVPENDDAELTEDQKKK	36	62	K43					K62	intra XL	n/c	SI	
1;2;3;5;6;8;9	FcpA	DQVLLKGELVPENDDAELTEDQK	36	60	K43					K62	intra XL	NO	n/c	
1	FcpA	KRLKWYWQK	132	142	K132					K137	intra XL	12.7	SI	
5;10;9	FcpA	DQVLLKGELVPENDDAELTEDQKKK	36	62	K43					K60	intra XL	NO	n/c	
1;2	FcpA	KEIMEQESLWKNPDFKGYNK	64	83	K64					K79	intra XL	24.1	SI	
1	FcpA	GYNKTFQELHQLSKTFANNQFR	80	101	K83					K93	intra XL	15.4	SI	
1;2;3;4;5;6;7;9	FcpA	GYNKTFQELHQLSK	80	93	K83					S92	intra XL	15.2	SI	
1	FcpA	YLASAYSFEENMIKK	235	249	S238					K248	intra XL	15.4	SI	
1;2;3;5;6;7;8;10;9	FcpA	TFQELHQLSKTFANNQFR	84	101	S92					K93	intra XL	3.8	SI	
3;5	FcpA	LALSNYQSGVNTIMKNR	102	118	S105					K116	intra XL	16.0	SI	
1	FcpA	DQVLLKGELVPENDDAELTEDQKK	36	61	K43					K60	intra XL	n/c	SI	
3;5;6;7;8;10;11;9	FcpA	AKQDALNYFSKAINHLIKNPDLR	160	184	K161					K170	loop XL	14.3	SI	
3;7	FcpA	AKQDALNYFSK	160	170	K161					S169	loop XL	12	SI	
1;2;3;5;6;7;8;10;11;9	FcpA	QDALNYFSKAINHLIKNPDLR	162	184	K170					K179	loop XL	16.6	SI	
1;2	FcpA	EYPAHKYLASAYSFEENMIK	229	248	K234					S238	loop XL	6.4	SI	
1;2	FcpA	KTKGPDDMLFK	249	259	K249					K251	loop XL	5.5	SI	
1;2;3;4;5;9	FcpA	TKGPDDMLFKYR	250	261	K251					K259	loop XL	14.6	SI	
1;3	FcpA	GPDDMLFKYRYKK	252	264	K259					K263	loop XL	6.2	SI	
1	FcpA	YKKNVHLLR	262	270	K263					K264	loop XL	3.8	SI	
1;2;3;4;5;10;11;9	FcpA	ATELKYGKDSPEYK	271	284	K275					K278	loop XL	9.1	SI	
1;2;10;11;9	FcpA	YGKDSPEYKHIVNVINRVISVAQ	276	300	K278					K284	loop XL	8.5	SI	
11;9	FcpA	KKEIMEQESLWK	63	74	K63					K64	loop XL	3.8	SI	
1;2	FcpA	KEIMEQESLWKNPDFK	64	79	K64					S71	loop XL	11.2	SI	
1;2;3;4;5;6;7;8;10;11;9	FcpA	NPDFKGYNKTFQELHQLSK	75	93	K79					K83	loop XL	8.7	SI	
4;5	FcpA	NPDFKGYNK	75	83	K79					Y81	loop XL	6.8	SI	
2	FcpA	KEEAEKK	126	132	K126					K131	loop XL	8.7	SI	
1;2;11;9	FcpA	QDALNYFSKAINHLIK	162	179	S169					K170	loop XL	3.8	SI	
1;2;9	FcpA	DSPEYKHIVNVINRVISVAQ	279	300	S280					K284	loop XL	6.3	SI	
1	FcpA	WYWQKVDRK	138	146	Y139					K142	loop XL	5.0	SI	
1;2	FcpA	EYPAHKYLASAYSFEENMIKK	229	249	K234					S238	loop XL	22.1	SI	
1;6;7;8	FcpA	YGKDSPEYKHIVNVINR	276	292	K278					K284	loop XL	6.7	SI	
1;2;3;5;6;7;8	FcpA	DSPEYKHIVNVINR	279	292	S280					K284	loop XL	9.7	SI	
1;2	FcpA	KTKGPDDMLFKYR	249	261	K249					K251	loop XL	19.7	SI	
3;4;5	FcpA	AKQDALNYFSK	160	170	K161	FcpA	EKMK			K157	loop XL	6.0	SI	
1	FcpA	RLKWYWQK	133	142	K137						dead end	n/c	SI	
3;11;9	FcpA	LKWYWQK	134	142	K137						dead end	n/c	SI	
11;9	FcpA	WYWQKVDR	138	145	K142						dead end	n/c	SI	
1;2	FcpA	EKMKAKQDALNYFSK	156	170	K157						dead end	n/c	SI	
1;2;3;4;5;6;7;8;10;11;9	FcpA	AINHLIKNPDLR	171	184	K179						dead end	n/c	SI	
1;2	FcpA	ERPEFKR	185	191	K190						dead end	n/c	SI	
1;2;11;9	FcpA	TKGPDDMLFK	250	259	K251						dead end	n/c	SI	
1;2;3;4;5;6;7;9	FcpA	GPDDMLFKYR	252	261	K259						dead end	n/c	SI	
1;2;5;11;9	FcpA	ATELKYGK	271	278	K275						dead end	n/c	SI	
1	FcpA	YGKDSPEYK	276	284	K278						dead end	n/c	SI	
3;5;6;7;8;10;11;9	FcpA	AKQDALNYFSKAINHLIKNPDLR	160	184	K170						dead end	n/c	SI	
3;7	FcpA	AKQDALNYFSK	160	170	S169						dead end	n/c	SI	
1;2;3;5;6;7;8;10;11;9	FcpA	QDALNYFSKAINHLIKNPDLR	162	184	K179						dead end	n/c	SI	
1;2	FcpA	EYPAHKYLASAYSFEENMIK	229	248	S238						dead end	n/c	SI	
1;2	FcpA	KTKGPDDMLFK	249	259	K251						dead end	n/c	SI	
1;2;3;4;5;9	FcpA	TKGPDDMLFKYR	250	261	K259						dead end	n/c	SI	
1	FcpA	GELVPENDDAELTEDQKKK	44	62	K60						dead end	n/c	SI	
1;2;3;4;5;10;11;9	FcpA	ATELKYGKDSPEYK	271	284	K278						dead end	n/c	SI	
1;2;10;11;9	FcpA	YGKDSPEYKHIVNVINRVISVAQ	276	300	K284						dead end	n/c	SI	
1;2;3;4;5;6;7;11;9	FcpA	EIMEQESLWKNPDFK	65	79	K74						dead end	n/c	SI	
1;2	FcpA	QDALNYFSK	162	170	S169						dead end	n/c	SI	
1;2	FcpA	YLASAYSFEENMIK	235	248	S238						dead end	n/c	SI	
1;2	FcpA	HIVNVINRVISVAQ	285	300	S297						dead end	n/c	SI	
1;2	FcpA	KEIMEQESLWK	64	74	S71						dead end	n/c	SI	
1;2;4;5;7	FcpA	TFQELHQLSK	84	93	S92						dead end	n/c	SI	
1;2	FcpA	LALSNYQSGVNTIMK	102	116	S109						dead end	n/c	SI	
1;2	FcpA	GELVPENDDAELTEDQK	44	60	T56						dead end	n/c	SI	
1;2	FcpA	TFANNQFR	94	101	T94						dead end	n/c	SI	
1;2	FcpA	KEIMEQESLWKNPDFK	64	79	S71						dead end	n/c	SI	
4;5	FcpA	NPDFKGYNK	75	83	Y81						dead end	n/c	SI	
1;2;3;4;5;6;7;9	FcpA	GYNKTFQELHQLSK	80	93	S92						dead end	n/c	SI	
1;2;11;9	FcpA	QDALNYFSKAINHLIK	162	179	K170						dead end	n/c	SI	
1	FcpA	YLASAYSFEENMIKK	235	249	K248						dead end	n/c	SI	
1;2;9	FcpA	DSPEYKHIVNVINRVISVAQ	279	300	K284						dead end	n/c	SI	
1;2;3;5;6;7;8;10;9	FcpA	TFQELHQLSKTFANNQFR	84	101	K93						dead end	n/c	SI	
3;5	FcpA	LALSNYQSGVNTIMKNR	102	118	K116						dead end	n/c	SI	
1	FcpA	WYWQKVDRK	138	146	K142						dead end	n/c	SI	
2	FcpA	NRDWVEQYR	117	125	Y124						dead end	n/c	SI	
1;2;3;4;5;6;7;10;11;9	FcpA	GELVPENDDAELTEDQKK	44	61	K60						dead end	n/c	SI	
1;2	FcpA	EIMEQESLWK	65	74	S71						dead end	n/c	SI	
2	FcpB	NVEALKGSLKY	267	277	K272	B0SJ12	HKYK	215	218		n/c	NO	n/c	
2;9	FcpB	NLGAKFK	123	129	K127	B0SKL2	GLSAYK	137	142		n/c	NO	n/c	
4;5;7;10	FcpB	IFDYNDRDSNKHYK	252	265	K262	B0SN26	KADNLLK	28	34		n/c	NO	n/c	
9	FcpB	GVGKYILSNVENTK	212	225	K215	B0SNS0	AKEKLR	161	166		n/c	NO	n/c	

1;2	FcpB	NLGAKFK	123	129	K127	B0SNX6	SFKK	225	228		n/c	NO	n/c	
1	FcpB	NLGAKFK	123	129	K127	B0STS3	KGDYAK	146	151		n/c	NO	n/c	
1	FcpB	IFDYNDRDSNKHYK	252	265	K262	B0STT3	KGTTAVPK	82	89		n/c	NO	n/c	
2	FcpB	YILSNVENTKSNPIR	216	230	K225	B0STT4	HYKK	402	405		n/c	NO	n/c	
1;11;9	FcpB	NVEALKGSLK	267	276	K272	B0STT4	HYKK	402	405		n/c	NO	n/c	
5	FcpB	KNVEALK	266	272	K266	B0SU03	VDDKEK	591	596		n/c	NO	n/c	
5	FcpB	SSKNLGAKFKKMELFFEGSNDPDPR	120	146	K122	B0STJ3	DPETGEIGKRETNTMPQWAGSCYYYLR	538	564		n/c	NO	n/c	
3;4;5;6;7;8;10;11;9	FcpB	KMELFFEGSNDPDPR	130	146	K130	B0SPP1	ADEKNESVSQSDATK	20	34	K24	inter XL	21.0	SI	
1;2;3;4;5;6;7;8;10;11;9	FcpB	NLGAKFK	123	129	K127	FcpB	GSLKY	273	277	K276	inter XL	10.8	SI	
1;2	FcpB	QFYFKNLDYFDK	236	247	K240	FcpB	NVEALKGSLK	267	276	K272	inter XL	9.5	SI	
1;2	FcpB	KNVEALK	266	272	K266	FcpB	NLGAKFK	123	129	K127	inter XL	17.1	SI	
2;3;4;5	FcpB	KNVEALK	266	272	K266	FcpB	GSLKY	273	277	K276	inter XL	16.6	SI	
3;10;11;9	FcpB	NVEALKGSLKY	267	277	K272	FcpB	NLGAKFK	123	129	K127	inter XL	13.3	SI	
1;2	FcpB	NVEALKGSLKY	267	277	K272	FcpB	KQFYFK	235	240	K235	inter XL	10.5	SI	
11;9	FcpB	GVGKYILSNVENTK	212	225	K215	FcpB	NLTKIR	152	157	K155	intra XL	16.3	SI	
1;2;3;5;11;9	FcpB	YILSNVENTKSNPIR	216	230	K225	FcpB	KQFYFK	235	240	K235	intra XL	14.9	SI	
10	FcpB	IFDYNDRDSNKHYKK	252	266	K262	FcpB	NVEALKGSLK	267	276	K272	intra XL	15.2	SI	
1;2;11;9	FcpB	NVEALKGSLK	267	276	K272	FcpB	KQFYFK	235	240	K235	intra XL	12.5	SI	
1;2	FcpB	IFDYNDR	252	258	Y255	FcpB	KNVEALK	266	272	K266	intra XL	16.8	SI	
1;2;3;4;5;6;7;8;10;11;9	FcpB	IFDYNDRDSNKHYK	252	265	K262	FcpB	KNVEALK	266	272	K266	loop XL	6.2	SI	
1;6;7;8	FcpB	DSNKHYK	259	265	K262	FcpB	KNVEALK	266	272	K266	loop XL	12.5	SI	
1;2	FcpB	IFDYNDRDSNK	252	262	S260	FcpB	KNVEALK	266	272	K266	loop XL	11.1	SI	
1;2	FcpB	IFDYNDR	252	258	Y255	FcpB	DSNKHYK	259	265	K262	loop XL	11.4	SI	
3;4;5	FcpB	SSKNLGAKFK	120	129	K122					K127	loop XL	14.4	SI	
1;2	FcpB	KMELFFEGSNDPDPRK	130	147	K130					S138	loop XL	22.8	SI	
1;2;3;4;5;7;10;11;9	FcpB	KMELFFEGSNDPDPR	130	146	K130					S138	loop XL	22.8	SI	
9	FcpB	GVGKYILSNVENTK	212	225	K215					S219	loop XL	12.4	SI	
1;2;3;5;9	FcpB	KQFYFKNLDYFDK	235	247	K235					K240	loop XL	7.7	SI	
1	FcpB	QFYFKNLDYFDKLFTK	236	251	K240					K247	loop XL	10.5	SI	
1;2;11	FcpB	LFTKIFDYNDRDSNK	248	262	K251					S260	loop XL	4.9	SI	
3;10;11	FcpB	IFDYNDRDSNKHYKKNVEALK	252	272	K262					K265	loop XL	14.4	SI	
1;2;3;4;5;6;7;8;10;11;9	FcpB	IFDYNDRDSNKHYKK	252	266	K262					K265	loop XL	14.4	SI	
3	FcpB	DSNKHYKKNVEALK	259	272	K262					K265	loop XL	14.4	SI	
4;5;6;7;10;11;9	FcpB	DSNKHYKK	259	266	K262					K265	loop XL	14.4	SI	
1;2	FcpB	HYKKNVEALK	263	272	K265					K266	loop XL	3.8	SI	
1;2	FcpB	KNVEALKGSLKY	266	277	K266					K272	loop XL	9.8	SI	
1	FcpB	KNVEALKGSLK	266	276	K266					K272	loop XL	9.8	SI	
7	FcpB	NVEALKGSLKY	267	277	K272					K276	loop XL	8.1	SI	
3	FcpB	SMADTEKELDDNISEVNKR	40	58	S40					K46	loop XL	9.8	SI	
1;2;3;4;5;6;7;10;11;9	FcpB	IFDYNDRDSNKHYK	252	265	S260					K262	loop XL	5.6	SI	
1;2	FcpB	IELFYQHDDYPVWGTPETPSEK	190	211	T204					S209	loop XL	13.4	SI	
1;2	FcpB	TLPHKTVLYK	72	81	T72					K76	loop XL	11.8	SI	
1;4;9	FcpB	SSKNLGAK	120	127	K122						dead end	n/c	SI	
1	FcpB	NLGAKFKKMELFFEGSNDPDPRK	123	147	K127						dead end	n/c	SI	
7	FcpB	FKKMELFFEGSNDPDPR	128	146	K129						dead end	n/c	SI	
1	FcpB	IELFYQHDDYPVWGTPETPSEKGVGK	190	215	K211						dead end	n/c	SI	
1;2;3;4;5;6;7;8;10;11;9	FcpB	YILSNVENTKSNPIR	216	230	K225						dead end	n/c	SI	
1;2	FcpB	KQFYFK	235	240	K235						dead end	n/c	SI	
1;2;9	FcpB	QFYFKNLDYFDK	236	247	K240						dead end	n/c	SI	
1;2;11;9	FcpB	NLDYFDKLFTK	241	251	K247						dead end	n/c	SI	
1	FcpB	LFTKIFDYNDR	248	258	K251						dead end	n/c	SI	
6;7;8	FcpB	DSNKHYK	259	265	K262						dead end	n/c	SI	
1;2	FcpB	KNVEALK	266	272	K266						dead end	n/c	SI	
1;2;11;9	FcpB	NVEALKGSLK	267	276	K272						dead end	n/c	SI	
3	FcpB	ELDDNISEVNKR	47	58	K57						dead end	n/c	SI	
1;2	FcpB	GKPSADGERCEAADK	82	96	K83						dead end	n/c	SI	
1;2	FcpB	MELFFEGSNDPDPRK	131	147	S138						dead end	n/c	SI	
1;2	FcpB	MELFFEGSNDPDPR	131	146	S138						dead end	n/c	SI	
1;2	FcpB	VISVIADVAPNGEPAHNDK	171	189	S173						dead end	n/c	SI	
1;2	FcpB	YILSNVENTK	216	225	S219						dead end	n/c	SI	
1;2	FcpB	SMADTEKELDDNISEVNK	40	57	S40						dead end	n/c	SI	
1;2	FcpB	IFDYNDRDSNK	252	262	S260						dead end	n/c	SI	
1	FcpB	ELDDNISEVNK	47	57	S53						dead end	n/c	SI	
1	FcpB	QEAQDNTCLHLEVFDFVGSEDGK	97	119	S115						dead end	n/c	SI	
1;2	FcpB	TYIYQNNFLLEDK	158	170	T158						dead end	n/c	SI	
1;2;3;4;5;7;10;11;9	FcpB	KMELFFEGSNDPDPR	130	146	S138						dead end	n/c	SI	
9	FcpB	GVGKYILSNVENTK	212	225	S219						dead end	n/c	SI	
1;2;3;5;9	FcpB	KQFYFKNLDYFDK	235	247	K240						dead end	n/c	SI	
1;2;11	FcpB	LFTKIFDYNDRDSNK	248	262	S260						dead end	n/c	SI	
1;2;3;4;5;6;7;8;10;11;9	FcpB	IFDYNDRDSNKHYKK	252	266	K265						dead end	n/c	SI	
1;2	FcpB	KNVEALKGSLKY	266	277	K272						dead end	n/c	SI	
7	FcpB	NVEALKGSLKY	267	277	K276						dead end	n/c	SI	
3	FcpB	SMADTEKELDDNISEVNKR	40	58	K46						dead end	n/c	SI	
1;2;3;4;5;6;7;10;11;9	FcpB	IFDYNDRDSNKHYK	252	265	K262						dead end	n/c	SI	
1;2	FcpB	NLDYFDK	241	247	Y244						dead end	n/c	SI	
3;5	FlaA1	SGAISTSAKVR	306	316	K314	FlaA1	VKLEQMLQK	297	305	K298	?	NO	n/c	
3;4;5;6;7;8	FlaA1	SGAISTSAKVR	306	316	K314	FlaA1	LVPGKPGDVK	109	118	K113	inter XL	25.2	SI	
3	FlaA1	FDENNKYSPVLNDPAKSPLVSR	84	105	K89	FlaA1	SYIDADNKK	159	167	K166	inter XL	23.6	SI	
3;4;5;6;7	FlaA1	YSPVLNDPAKSPLVSR	90	105	K99	FlaA1	SYIDADNKK	159	167	K166	inter XL	6.2	SI	
3;4;5;6;7;8	FlaA1	LVPGKPGDVKNVDAGK	109	126	K113	FlaA1	SGAISTSAKVR	306	316	K314	inter XL	25.2	SI	
3	FlaA1	VRDCGGSNGSSQNK	315	328	S321	FlaA1	NVDAGKVLAVK	119	131	K126	inter XL	NO	n/c	
3	FlaA1	VRDCGGSNGSSQNK	315	328	S321	FlaA1	LVPGKPGDVK	109	118	K113	inter XL	NO	n/c	
3;4;5	FlaA1	LVPGKPGDVKNVDAGK	109	126	K113	FlaA1	LVPGKPGDVKNVDAGK	109	126	K113	inter XL	20.2	SI	
3	FlaA1	LVPGKPGDVKNVDAGK	109	126	K113	FlaA1	VRDCGGSNGSSQNK	315	328	S321	inter XL	23.6	SI	
3	FlaA1	NVDAGKVLAVK	119	131	K126	FlaA1	NLIFKQFK	244	251	K248	intra XL	20.7	SI	
3	FlaA1	YSPVLNDPAKSPLVSR	90	105	K99	FlaA1	DGSGDMKFDENNK	77	89	K83	intra XL	22.3	SI	
3;5	FlaA1	EVKLVPGKPGDVK	106	118	K108	FlaA1	NVDAGKVLAVK	119	131	K126	intra XL	9.7	SI	
3;4;5;6;7	FlaA1	FDENNKYSPVLNDPAKSPLVSR	84	105	K89					K99	intra XL	22.9	SI	
3;4;5;6;7;8;10;11;9	FlaA1	KMEKVYGIEAPGNVK	168	182	K168					K171	loop XL	9.5	SI	
4;7	FlaA1	MEKVYGIEAPGNVK	169	182	K171					Y173	loop XL	6.8	SI	
3;4;5;10;11;9	FlaA1	NLIFKQFKIR	244	253	K248					K251	loop XL	9.7	SI	
3	FlaA1	VKLEQMLQKSGAISTSAKVR	297	316	K298					K305	loop XL	NO	n/c	
3;4;5;6;7;8;10;11;9	FlaA1	VKLEQMLQKSGAISTSAK	297	314	K298					S306	loop XL	NO	n/c	
3;5;7	FlaA1	LEQMLQKSGAISTSAKVR	299	316	K305					K314	loop XL	NO	n/c	
1;2;3;4;5;6	FlaA1	LEQMLQKSGAISTSAK	299	314	K305					S306	loop XL	NO	n/c	
3	FlaA1	DGSGDMKFNNKYSPVLNDPAK	77	99	K83					K89	loop XL	10.2	SI	
3;5	FlaA1	FNNKYSPVLNDPAK	84	99	K89					S91	loop XL	6.3	SI	
3;4;5	FlaA1	EVKLVPGKPGDVKNVDAGK	106	126	K108					K113	loop XL	14.2	SI	
3;4;5	FlaA1	EVKLVPGKPGDVK	106	118	K108					K113	loop XL	14.2	SI	
3	FlaA1	LVPGKPGDVKNVDAGKVLAVK	109	131	K113					K118	loop XL	9.1	SI	
3;4;5;6;7;8	FlaA1	QYKVSNGVATPVGNDIGALELK	26	47	Q25					K28	loop XL	NO	n/c	
3;4	FlaA1	YSPVLNDPAKSPLVSR	90	105	S91					K99	loop XL	NO	n/c	
3;5	FlaA1	NVDAGKVLAVK	119	131	K126	FlaA1	LVPGKPGDVK	109	118	K113	loop XL	6.3	SI	
3;5;7	FlaA1	MEKVYGIEAPGNVK	169	182	K171	FlaA1	SYIDADNKK	159	167	K166	loop XL	15.0	SI	
3	FlaA1	KMEKVYGIEAPGNVK	168	182	K171	FlaA1	SYIDADNKK	159	167	K166	loop XL	9.5	SI	

1;2;3;4;5;6;7;8;10;11;9	FlaA1	LVPGKPGDVKNVDAGK	109	126	K113					K118	loop XL	9.4	SI	
1;3;4;5;6;7;10;11;9	FlaA1	NVDAGKVLAVK	119	131	K126						dead end	n/c	SI	
3;4;5	FlaA1	SYIDADNKK	159	167	K166						dead end	n/c	SI	
2;3;5;10;9	FlaA1	VYGIEAPGNVKAISVWVLGR	172	191	K182						dead end	n/c	SI	
1;2;3;4;5;6;7;8;10;11;9	FlaA1	NLIFKQFK	244	251	K248						dead end	n/c	SI	
3;4;5	FlaA1	VKLEQMLQK	297	305	K298						dead end	n/c	SI	
2;3;4;5;6;7;8;10	FlaA1	SGAISTSAKVR	306	316	K314						dead end	n/c	SI	
6;7;8	FlaA1	LVPGKPGDVK	109	118	K113						dead end	n/c	SI	
3;4;5	FlaA1	VRDCGGSNGSSQNK	315	328	S321						dead end	n/c	SI	
3	FlaA1	VSNGVATPVGNDIGALELK	29	47	S30						dead end	n/c	SI	
3;4;5;6;7;8;10;11;9	FlaA1	KMEKVYGIEAPGNVK	168	182	K171						dead end	n/c	SI	
4;7	FlaA1	MEKVYGIEAPGNVK	169	182	Y173						dead end	n/c	SI	
3;4;5;6;7;8;10;11;9	FlaA1	VKLEQMLQKSGAISTSAK	297	314	S306						dead end	n/c	SI	
1;2;3;4;5;6	FlaA1	LEQMLQKSGAISTSAK	299	314	S306						dead end	n/c	SI	
3;4;5;6;7	FlaA1	FDENNKYSPVLNDPAKSPLVSR	84	105	K99						dead end	n/c	SI	
3;5	FlaA1	FNNKYSPVLNDPAK	84	99	S91						dead end	n/c	SI	
3;4;5	FlaA1	EVKLVPGKPGDVK	106	118	K113						dead end	n/c	SI	
1;2;3;4;5;6;7;8;10;11;9	FlaA1	LVPGKPGDVKNVDAGK	109	126	K118						dead end	n/c	SI	
3;4;5;6;7;8	FlaA1	QYKVSNGVATPVGNDIGALELK	26	47	K28						dead end	n/c	SI	
1;2	FlaA1	FQFTFPGNNVVTIRPPR	132	148	T135						dead end	n/c	SI	
6;7	FlaA2	SEAKYPGSEIK	228	238	K231	B0SMX0	LSSVMLEKEK	152	161		n/c	NO	n/c	
5	FlaA2	TVPEDGGDKIEK	89	100	K91	B0SRM7	ALGVSPSYLSK	22	32		n/c	NO	n/c	
3;5;6;7;8;9	FlaA2	VELYPPHEYIIKGK	118	131	K129	B0SJC6	NPYIKYDNR	298	306	K302	inter XL	20.0	SI	
3;4;5;6;7;8	FlaA2	VELYPPHEYIIKGK	118	131	K129	FlaA1	DGVNKDGSGDMK	72	83	K76	inter XL	27,6	SI	
10;9	FlaA2	SEAKYPGSEIK	228	238	K231	FlaA2	GLIKDVFK	79	88	K82	inter XL	11.0	SI	
6;7;9	FlaA2	VELYPPHEYIIKGK	118	131	K129	FlaA2	SEAKYPGSEIK	228	238	K231	intra XL	11.0	SI	
3;5;6;7;8;10;11;9	FlaA2	VELYPPHEYIIKGK	118	131	K129	FlaA2	YPGSEIKDNW	232	241	K238	intra XL	14.1	SI	
1;2;3;5;10;11;9	FlaA2	KLTATVPGFIPQSTR	174	188	K174	FlaA2	KIVLDDFEEAEDWR	47	60	K47	intra XL	14.7	SI	
9	FlaA2	GLIKDVFKTVPEDGGDKIEK	79	100	K82	FlaA2	KLTATVPGFIPQSTR	174	188	K174	intra XL	8.5	SI	
3;5;6;7;9	FlaA2	DVFKTVPEDGGDKIEK	83	100	K88	FlaA2	KLTATVPGFIPQSTR	174	188	K174	intra XL	11.5	SI	
1	FlaA2	TVPEDGGDKIEK	89	100	K91	FlaA2	GLIKDVFK	79	88	K82	intra XL	13.7	SI	
1	FlaA2	IEKNHILGVK	98	107	K100	FlaA2	TVPEDGGDK	89	97	T89	intra XL	18.7	SI	
1;2	FlaA2	ATTPLGETR	63	71	T64	FlaA2	TLKLVQR	72	78	K74	intra XL	21.8	SI	
4;5;10	FlaA2	DYKDVTHNIR	154	163	K156					T159	loop XL	8.8	SI	
1;2;4;5;10;11;9	FlaA2	KLTATVPGFIPQSTR	174	188	K174					T178	loop XL	13	SI	
1;2;3;5	FlaA2	TDRSEAKYPGSEIKDNW	225	241	K231					K238	loop XL	11.4	SI	
3;6;7;10;11;9	FlaA2	GLIKDVFKTVPEDGGDKIEK	79	100	K82					K88	loop XL	7.3	SI	
3	FlaA2	DVFKTVPEDGGDKIEK	83	100	K88					K91	loop XL	17.2	SI	
1;2;3;4;5;6;7;8;10;11;9	FlaA2	TVPEDGGDKIEKNHILGVK	89	107	K91					K100	loop XL	5.1	SI	
1;2;3;5;9	FlaA2	NHILGVKTHFAAK	101	113	K107					T108	loop XL	3.8	SI	
3;5;10;11;9	FlaA2	HTLFAKFR	146	153	T147					K151	loop XL	12.9	SI	
1;2;3;4;5;7;8;11;9	FlaA2	THFAAKGLDR	108	117	T108					K113	loop XL	13.5	SI	
10;9	FlaA2	GLDRVELYPPHEYIIKGK	114	131	Y121					K129	loop XL	19.5	SI	
1	FlaA2	GLDRVELYPPHEYIIK	114	129	Y121					Y126	loop XL	10.4	SI	
1	FlaA2	DTGGQNTANTANDSPLRK	28	47	T29						dead end	n/c	SI	
3;6;9	FlaA2	DTGGQNTANTANDSPLR	28	46	T29						dead end	n/c	SI	
3;4;5;6;7;8;10;11;9	FlaA2	VELYPPHEYIIKGK	118	131	K129						dead end	n/c	SI	
2;3;4;5;9	FlaA2	FRDYKDVTHNIR	152	163	K156						dead end	n/c	SI	
2	FlaA2	KIVLDDFEEAEDWRVK	47	62	K47						dead end	n/c	SI	
1;2;3;11;9	FlaA2	KIVLDDFEEAEDWR	47	60	K47						dead end	n/c	SI	
1;2;3;4;5;7;9;10;11	FlaA2	TDRSEAKYPGSEIK	225	238	K231						dead end	n/c	SI	
1;2;3;4;5;7;9;10;11	FlaA2	SEAKYPGSEIKDNW	228	241	K231						dead end	n/c	SI	
2;3;4;5	FlaA2	YPGSEIKDNW	232	241	K238						dead end	n/c	SI	
2;4;5	FlaA2	VKATTPLGETR	61	71	K62						dead end	n/c	SI	
3;5	FlaA2	TLKLVQR	72	78	K74						dead end	n/c	SI	
1;2;3;4;5;6;7;11;9	FlaA2	GLIKDVFK	79	88	K82						dead end	n/c	SI	
1;7	FlaA2	DVFKTVPEDGGDK	83	97	K88						dead end	n/c	SI	
1;2;3;4;5;10;9	FlaA2	TVPEDGGDKIEK	89	100	K91						dead end	n/c	SI	
1;2;5;7;11;9	FlaA2	IEKNHILGVK	98	107	K100						dead end	n/c	SI	
1;2	FlaA2	LTATVPGFIPQSTR	175	188	S186						dead end	n/c	SI	
3;4;5	FlaA2	DVTHNIR	157	163	T159						dead end	n/c	SI	
1;2	FlaA2	ATTPLGETR	63	71	T64						dead end	n/c	SI	
1	FlaA2	TVPEDGGDK	89	97	T89						dead end	n/c	SI	
1;2	FlaA2	THFAAK	108	113	T108						dead end	n/c	SI	
3;4;5	FlaA2	VELYPPHEYIIK	118	129	Y121						dead end	n/c	SI	
4;5;10	FlaA2	DYKDVTHNIR	154	163	T159						dead end	n/c	SI	
1;2;4;5;10;11;9	FlaA2	KLTATVPGFIPQSTR	174	188	S186						dead end	n/c	SI	
1;2;3;5	FlaA2	TDRSEAKYPGSEIKDNW	225	241	K238						dead end	n/c	SI	
3;6;7;10;11;9	FlaA2	GLIKDVFKTVPEDGGDKIEK	79	100	K88						dead end	n/c	SI	
3	FlaA2	DVFKTVPEDGGDKIEK	83	100	K91						dead end	n/c	SI	
1;2;3;4;5;6;7;8;10;11;9	FlaA2	TVPEDGGDKIEKNHILGVK	89	107	K100						dead end	n/c	SI	
1;2;3;5;9	FlaA2	NHILGVKTHFAAK	101	113	T108						dead end	n/c	SI	
3;5;10;11;9	FlaA2	HTLFAKFR	146	153	K151						dead end	n/c	SI	
1;2;3;4;5;7;8;11;9	FlaA2	THFAAKGLDR	108	117	K113						dead end	n/c	SI	
10;9	FlaA2	GLDRVELYPPHEYIIKGK	114	131	K129						dead end	n/c	SI	
1	FlaA2	GLDRVELYPPHEYIIK	114	129	Y126						dead end	n/c	SI	
1	FlaB1	MIINHNLAAINSHR	1	14	M1	B0SK36	KIQLFTNTK	651	661		n/c	NO	n/c	
1	FlaB1	MIINHNLAAINSHR	1	14	M1	B0SKW3	VLQDKKVSK	244	254		n/c	NO	n/c	
1	FlaB1	MIINHNLAAINSHR	1	14	M1	B0SMD7	SIGNYLFKGFR	25	35		n/c	NO	n/c	
1;2	FlaB1	MIINHNLAAINSHR	1	14	M1	B0SMR3	IANLEQFVSK	515	526		n/c	NO	n/c	
2	FlaB1	MIINHNLAAINSHR	1	14	M1	B0SR08	ATEIEALAVATAK	230	242		n/c	NO	n/c	
1;2	FlaB1	MIINHNLAAINSHR	1	14	M1	B0SR89	MGETLSVKIQSK	1	12		n/c	NO	n/c	
2	FlaB1	SNDAIGTLDAALTR	191	204	S191	B0SRI6	LSKQR	56	60		n/c	NO	n/c	
1	FlaB1	VLKFQNEEVSK	15	25	K17	B0SRL4	LKKGEMIHMMNIGR	221	234		n/c	NO	n/c	
3;4;5	FlaB1	VLKFQNEEVSK	15	25	K17	B0S0	MSMMIPPFLGADVMK	247	261		n/c	NO	n/c	
1;2;3;5;10;11;9	FlaB1	MIINHNLAAINSHR	1	14	M1	FlaB2	ALKFTQWDVDK	15	25	K17	inter XL	26,8	SI	
1;2;3;4;5;7;10;11;9	FlaB1	VLKFQNEEVSK	15	25	K17	FlaB2	MIINHNMSAIQSHR	1	14	M1	inter XL	26.8	SI	
9	FlaB1	AGDDASGLAVSEKMR	39	53	K51	FlaB4	MIINHNVSAIFAHR	1	14	M1	inter XL	25.7	SI	
1;2;3;10;11	FlaB1	MIINHNLAAINSHR	1	14	M1	FlaB4	TLKSNDANLSK	15	25	K17	inter XL	NO	n/c	
1;2;3;5;10;11;9	FlaB1	VLKFQNEEVSK	15	25	K17	FlaB4	MIINHNVSAIFAHR	1	14	M1	inter XL	26.8	SI	
1;2	FlaB1	NQILVQSGTAMLAQANVRPQGVLSLLR	255	281	S261	FlaB1	IRDADMAEETVAFTK	240	254	T249	inter XL	25.3	SI	
1;2;3;4;5;6;7;8;10;11;9	FlaB1	MIINHNLAAINSHR	1	14	M1	FlaB1	VLKFQNEEVSK	15	25	K17	inter XL	26,8	SI	
1	FlaB1	VLKFQNEEVSK	15	25	K17	FlaB1	VLKFQNEEVSK	15	25	K86	intra XL	27.3	SI	
2	FlaB1	SNDAIGTLDAALTR	191	204	S191	FlaB1	ISKQR	205	209	K207	intra XL	17.7	SI	
3;5;9	FlaB1	VLKFQNEEVSK	15	25	K17					S24	intra XL	12.5	SI	
1;2;3;5	FlaB1	LEHAAKGLMYENTQASESR	219	239	K224					S236	intra XL	20.7	SI	
1;9	FlaB1	MIINHNLAAINSHRVLKFQNEEVSK	1	25	M1					K17	intra XL	22.8	SI	
1	FlaB1	MIINHNLAAINSHR	1	14	M1					S12	intra XL	21.1	SI	
1;2	FlaB1	GLMYENTQASESR	225	239	S236					S238	intra XL	6.6	SI	
1	FlaB1	NQILVQSGTAMLAQANVRPQGVLSLLR	255	281	S261					S278	intra XL	22.5	SI	
1;2;3;4;5	FlaB1	FQNEEVSKNMEK	18	29	K25						dead end	n/c	SI	
1;2;3;4;10;9	FlaB1	AGDDASGLAVSEKMR	39	53	K51						dead end	n/c	SI	
1;2	FlaB1	SNDAIGTLDAALTR	191	204	S191						dead end	n/c	SI	
1;2	FlaB1	FQNEEVSK	18	25	S24						dead end	n/c	SI	
1;2	FlaB1	AGDDASGLAVSEK	39	51	S44						dead end	n/c	SI	
1;2	FlaB1	IRDADMAEETVAFTK	240	254	T249						dead end	n/c	SI	

1;2	FlaB1	ANLGAYFNR	210	218	Y215						dead end	n/c	SI	
3;5;9	FlaB1	VLKFQNEEVSK	15	25	S24						dead end	n/c	SI	
1;2;3;5	FlaB1	LEHAAKGLMYENTQASESR	219	239	S236						dead end	n/c	SI	
1;9	FlaB1	MIINHNLAAINSHRVLKFQNEEVSK	1	25	K17						dead end	n/c	SI	
1	FlaB1	MIINHNLAAINSHR	1	14	S12						dead end	n/c	SI	
1;2	FlaB1	GLMYENTQASESR	225	239	S238						dead end	n/c	SI	
1	FlaB1	NQILVQSGTAMLAQANVRPQGVLSLLR	255	281	S278						dead end	n/c	SI	
3	FlaB2	ALKMSEGAIK	172	181	K174	B0SJN5	KSTK	619	622		n/c	NO	n/c	
3	FlaB2	ADFALNKIMK	197	206	K203	B0SK30	KTSK	365	368		n/c	NO	n/c	
1	FlaB2	MIINHNMSAIQSHR	1	14	M1	B0SL66	TVMIAPMTTVMR	41	52		n/c	NO	n/c	
4	FlaB2	ADFALNKIMK	197	206	K203	B0SN61	ISADIFLER	223	231		n/c	NO	n/c	
3;5	FlaB2	MIINHNMSAIQSHR	1	14	M1	B0SQU6	SASLEIVKMMMK	84	95		n/c	NO	n/c	
1	FlaB2	MIINHNMSAIQSHR	1	14	M1	B0SSM7	TPFFIKDLPDR	497	507		n/c	NO	n/c	
1;2	FlaB2	ALKFTQWDVDK	15	25	K17	B0S0	MSMMIPPFLGADVMK	247	261		n/c	NO	n/c	
6	FlaB2	ALKMSEGAIK	172	181	K174	B0SKL4	ADLEASKSMVLQGK	254	267		n/c	NO	n/c	
3;4;5;6;7;8;10;11;9	FlaB2	AAIAKADFALNK	190	203	K196	FlaB2	ALKMSEGAIK	172	181	K174	intra XL	17.1	SI	
3;4;5;10;11;9	FlaB2	ADFALNKIMK	197	206	K203	FlaB2	ALKMSEGAIK	172	181	K174	intra XL	16.4	SI	
1;2;3;5;11;9	FlaB2	QILVQSGTAMLAQASLRPNSVLR	255	277	S260	FlaB2	IRDADMAEEMVALTTK	239	254	T252	intra XL	11.3	SI	
1;2;3;5;10;11;9	FlaB2	MIINHNMSAIQSHR	1	14	M1	FlaB2		15	25	K17	intra XL	26.8	SI	
1;2	FlaB2	MIINHNMSAIQSHRALKFTQWDVDK	1	25	M1					K17	intra XL	21.1	SI	
4	FlaB2	MIINHNMSAIQSHR	1	14	M1					S8	loop XL	17.2	SI	
3	FlaB2	ALKFTQWDVDK	15	25	K17						dead end	n/c	SI	
1;3;4;5;10;11;9	FlaB2	ALKMSEGAIK	172	181	K174						dead end	n/c	SI	
5	FlaB2	MSEGAIKIALSTPGKAAIAKADFALNK	175	203	K181						dead end	n/c	SI	
3;4;5;6;10;11;9	FlaB2	AAIAKADFALNK	190	203	K196						dead end	n/c	SI	
1;9	FlaB2	ADFALNKIMK	197	206	K203						dead end	n/c	SI	
3;5	FlaB2	INLAGDDASGLAVSEKLR	36	53	K51						dead end	n/c	SI	
1;2	FlaB2	GLMGAYENMQASESR	224	238	S235						dead end	n/c	SI	
1;2	FlaB2	INLAGDDASGLAVSEK	36	51	S44						dead end	n/c	SI	
1;2	FlaB2	FTQWDVDK	18	25	T19						dead end	n/c	SI	
3;4	FlaB3	IGTSAEFNKVKPLDGR	126	141	S129					K136	loop XL	18.0	SI	
3;5;9	FlaB3	ASDSLQVLDDAISKVNR	191	207	K204						dead end	n/c	SI	
3;4	FlaB3	NQILTKSGVAMLAQANLRPESVVK	255	278	K260						dead end	n/c	SI	
1;2	FlaB3	ASDSLQVLDDAISK	191	204	S192						dead end	n/c	SI	
1;2	FlaB4	MIINHNVSAIFAHR	1	14	M1	B0SLC6	TSIAIDTIINQK	176	187		n/c	NO	n/c	
1;2	FlaB4	ANSVIGLCDDALR	192	204	S194	B0SRP0	TIEDVK	87	92		n/c	NO	n/c	
1;2	FlaB4	MIINHNVSAIFAHR	1	14	M1	B0SJM1	LTELQKQLSK	57	66		n/c	NO	n/c	
2	FlaB4	MIINHNVSAIFAHR	1	14	M1	B0SQ28	SILESPTEIKS	151	161		n/c	NO	n/c	
1;2	FlaB4	MIINHNVSAIFAHR	1	14	M1	B0SRQ9	KEEEEVPLSK	342	351		n/c	NO	n/c	
2	FlaB4	MIINHNVSAIFAHR	1	14	M1	B0STH7	DQKLTEVLSK	592	601		n/c	NO	n/c	
1;2	FlaB4	ANSVIGLCDDALR	192	204	S194	B0STP4	VISK	266	269		n/c	NO	n/c	
3;5;10;9	FlaB4	INKAGDDASGLAVSEK	36	51	K38	FlaB4	MIINHNVSAIFAHR	1	14	M1	inter XL	16.9	SI	
1;9	FlaB4	MIINHNVSAIFAHRTLKSNDANLSK	1	25	M1	FlaB4	TLKSNDANLSK	15	25	K17	inter XL	17.2	SI	
1;2;3;4;5;10;11;9	FlaB4	MIINHNVSAIFAHR	1	14	M1	FlaB4	TLKSNDANLSK	15	25	K17	intra XL	22.8	SI	
11;9	FlaB4	MIINHNVSAIFAHR	1	14	M1	FlaB4	SNDANLSKDIEK	18	29	K25	intra XL	25.2	SI	
9	FlaB4	MIINHNVSAIFAHRTLK	1	17	M1	FlaB4	SNDANLSK	18	25	S18	intra XL	25.2	SI	
1;2;3	FlaB4	TLKSNDANLSKDIEK	15	29	K17					K25	loop XL	22.0	SI	
3;4;5;10;11;9	FlaB4	SNDANLSKDIEKLSSGMR	18	35	K25					K29	loop XL	7.4	SI	
5	FlaB4	DIEKLSSGMR	26	35	K29					S31	loop XL	8.4	SI	
3;4;5;10;11;9	FlaB4	INKAGDDASGLAVSEK	36	51	K38					S44	loop XL	10.2	SI	
1;2	FlaB4	MIINHNVSAIFAHR	1	14	M1					S8	loop XL	12.5	SI	
9	FlaB4	MIINHNVSAIFAHRTLKSNDANLSKDIEK	1	29	K17					K25	loop XL	9.1	SI	
1;2;3;4;5;9	FlaB4	MEHAAKGLMYENTQASESR	220	240	K225						dead end	n/c	SI	
2;3;10;9	FlaB4	TLKSNDANLSK	15	25	K17						dead end	n/c	SI	
1;2;3;4;5;10;9	FlaB4	SNDANLSKDIEK	18	29	K25						dead end	n/c	SI	
1	FlaB4	LSSGMRINKAGDDASGLAVSEK	30	51	K38						dead end	n/c	SI	
1;2	FlaB4	ANSVIGLCDDALR	192	204	S194						dead end	n/c	SI	
1;2;3;10;11;9	FlaB4	YQILTQAATSMLAQANMK	256	273	S265						dead end	n/c	SI	
1	FlaB4	RAEQNTEDGMSLIQTAEGYLQETHEIVQR	61	89	S71						dead end	n/c	SI	
9	FlaB4	MIINHNVSAIFAHRTLK	1	17	S18						dead end	n/c	SI	
1;2;3	FlaB4	TLKSNDANLSKDIEK	15	29	K25						dead end	n/c	SI	
5	FlaB4	DIEKLSSGMR	26	35	S31						dead end	n/c	SI	
3;4;5;10;11;9	FlaB4	INKAGDDASGLAVSEK	36	51	S44						dead end	n/c	SI	
1;2	FlaB4	MIINHNVSAIFAHR	1	14	S8						dead end	n/c	SI	

Tabla 13: Datos finales completos de XL-MS de filamentos purificados de *L. biflexa wt*, obtenidos al combinar los mejores 4 juegos de datos de los ensayos discutidos en la sección 3.3.3. Es a partir de estos datos que se construyen la tabla 2 y las figuras 28, 29 y 36 con la red de interacciones entre proteínas flagelares.

ANEXO 3 – Articulos científicos publicados durante esta tesis.

1- Crystallization of FcpA from *Leptospira*, a novel flagellar protein that is essential for pathogenesis.
San Martin F, Mechaly AE, Larrieux N, Wunder EA Jr, Ko AI, Picardeau M, Trajtenberg F, Buschiazzo A. Acta Crystallogr F Struct Biol Commun. 2017 Mar 1;73(Pt 3):123-129. doi: 10.1107/S2053230X17002096.

2- An asymmetric sheath controls flagellar supercoiling and motility in the *Leptospira* spirochete.
Gibson KH, Trajtenberg F, Wunder EA, Brady MR, San Martin F, Mechaly A, Shang Z, Liu J, Picardeau M, Ko A, Buschiazzo A, Sindelar CV. Elife. 2020 Mar 11;9:e53672. doi: 10.7554/eLife.53672.

3- Structural basis of flagellar filament asymmetry and supercoil templating by *Leptospira* spirochete sheath proteins
Megan R. Brady, Fabiana San Martin, Garrett E. Debs, Kimberley H. Gibson, Azalia Rodríguez, Rosario Durán, Elsio A. Wunder Jr, Albert I. Ko, Alejandro Buschiazzo, Charles V. Sindelar. bioRxiv preprint 2922 Mar03, doi: https://doi.org/10.1101/2022.03.03.482903

4- 3D cryo-EM imaging of bacterial flagella: Novel structural and mechanistic insights into cell motility.
Mondino S, San Martin F, Buschiazzo A. J Biol Chem. 2022 Jul;298(7):102105. doi: 10.1016/j.jbc.2022.102105.

5- Diving into the complexity of the spirochetal endoflagellum.
San Martin F, Fule L, Iraola G, Buschiazzo A, Picardeau M. Trends Microbiol. 2023 Mar;31(3):294-307. doi: 10.1016/j.tim.2022.09.010.