

UCH-FC

DOC - Microbiología

VFA6

C.1

**CARACTERIZACIÓN ESTRUCTURAL Y  
EVOLUTIVA DEL COMPLEJO  
FTSW/FTSI DE *ESCHERICHIA COLI***

Tesis

Entregada a la

Universidad de Chile

En Cumplimiento Parcial de los Requisitos

Para Optar al Grado de

Doctor en Ciencias

Mención en Microbiología

Facultad de Ciencias

por

**Felipe Villanelo Lizana**

Enero, 2018

Director de Tesis

Dr. Octavio Monasterio Opazo

FACULTAD DE CIENCIAS  
UNIVERSIDAD DE CHILE  
INFORME DE APROBACIÓN  
TESIS DE DOCTORADO

Se informa a la Escuela de Postgrado de la Facultad de Ciencias que la  
Tesis de Doctorado presentada por el candidato

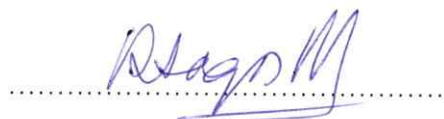
**Felipe Villanelo Lizana**

Ha sido aprobada por la comisión de Evaluación de la tesis como requisito  
para optar al grado de Doctor en Ciencias mención Microbiología, en el  
examen de Defensa Privada de Tesis rendido el día Martes 24 de Octubre  
de 2017 y aprobado el 3 de Enero de 2018.

Director de Tesis:  
Octavio Monasterio Opazo



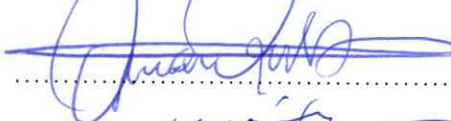
Co Director de Tesis:  
Dra. Rosalba Lagos Mónaco



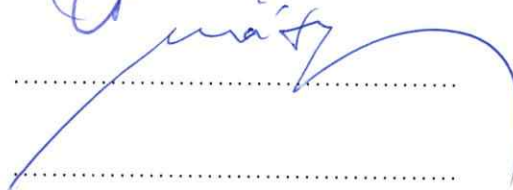
Comisión de Evaluación de Tesis:  
Dr. Oscar León Decap



Dr. Omar Orellana Orellana



Dr. Claudio Vásquez Guzmán



Dr. Danilo González Nilo





Un niño tímido e inquieto que siempre buscó bichitos en el pasto y desarmó sus juguetes (y otras cosas) debajo de la cama. Leí todos los libros de mi casa (que eran pocos) y jugué con cada mascota gatuna que tuvimos (que fueron muchas). Pasé por mil pasatiempos, por mil carreras futuras a estudiar hasta que me quedé en la Bioquímica.

La Bioquímica me llevó al puerto principal, donde encontré amistades, comidas, fiestas, estructuras de proteínas, el sonido del mar, y por cierto, el primogénito, venido el pobre sin que nadie le advirtiera a donde lo traían.

Después de vagar por el puerto, por casas añejas, por pegas precarias, por gente pintoresca, caí en la capital, en busca de un doctorado.

Allí me adentré en las proteínas de membrana, en las dinámicas, luché contra experimentos mañosos y con organizaciones políticas. Las últimas me ganaron la pelea si, porque aquí me tienen militando (depués de renegar de las militancias), entre simulaciones y muebles caseros.

Dicen que la capital me regaló un doctorado (aunque de regalado bien poco), pero yo creo que fue la yapa no más. El regalo fueron las 3 mujeres que hoy me acompañan, más el primogénito por cierto, que de nuevo sin que nadie le advirtiera tuvo que aprender a ser hermano mayor.

Alguien me dijo una vez que la felicidad solo sirve si puedes compartirla. Hoy, con mi hermosa familia no soy ni más ni menos que completamente feliz.

No podría dedicar esto a Andrea, Amanda, Aurora ni a César  
porque tendrían que ser mil hojas  
con más dibujos que figuras  
con más risas que tablas  
con más poemas que experimentos  
con más soles que números  
y aunque mil hojas fueran  
llenas de sudor y trabajo  
para ellas necesitaría mil estrellas  
llenas de amor y gratitud que les adornaran para siempre  
sus almas llenas de luz

## Agradecimientos

Quisiera partir esta sección declarando que es mucha la gente a la que tengo que agradecer, así que si alguien me falta o no me alcanza el espacio que no se enoje. Han sido muchos años en esto y ha tenido varias etapas y subetapas y pucha que he necesitado ayuda.

Primero a Andrea por llegar a mi vida así de pronto, junto con este proceso. Sino fuera por toda tu ayuda, apañe y paciencia, no estaría escribiendo nada de esto. Además me regalaste dos cabritas chicas lindas y locas, Amanda y Aurora, que gracias a sus risas pude no volverme loco mientras peleaba con experimentos imposibles y me obligaron a no pasar de largo leyendo papers o corriendo dinámicas. Gracias por todos los días de risas, llantos, enojos y abrazos, gracias por cada segundo que me regalan las tres.

César, el primogénito. A él, le debo un agradecimiento muy especial, porque probablemente tuvo que bancarse lo más difícil, que me fuera de su lado. Tantos años peleando para que vivieramos juntos y se me ocurre irme a hacer un doctorado a otra ciudad. Gracias por no perder nuestra relación, a pesar de tus pocos años cuando esto empezó, siempre dejaste que te fuera a ver y después viniste religiosamente a verme, aunque fueras chico, aunque fuera lejos, aunque probablemente no entendieras muy bien que estaba haciendo lejos. Gracias por aceptar a tus hermanas que llegaron de pronto en medio de este proceso. Gracias por ser tú, así como eres, por estar lejos en distancia, pero siempre cerca en el corazón.

Por cierto agradecer a mi suegra, mi suegro, mi super suegra y mis cuñados Miguel y Pedro, por todo el apañe en este tiempo. A mi mamá, papá y hermanos, David y Cristian, y mi tía Any, que quisas nos vimos menos, pero siempre estuvieron ahí para lo que necesité.

Agradecer a los compas de BEM, grandes amistades que no solo carretearon (y pucha que carreteamos) sino que ayudaron en experimentos, presentaciones y demases. Stolen, Chica, Pablo, Gisela, Pauli, Luis, Alexis, Esteban, Maca, Don Pedro. Gracias también al profesor Monasterio y la doctora Rosalba por aceptarme en el laboratorio.

Agradecer a la "militancia", pues, el crecimiento político acompañó mi proceso doctoral. A los compas de Conocimiento Colectivo, a los de Valparaíso, a los de SOL, a los del GAP, a los del Frente Amplio que ojalá siga creciendo. Al profe Babul que se da el trabajo de escribirme para saludar mis artículos, aunque no siempre estamos de acuerdo. A Gonzalo Gutiérrez y las conversas largas para arreglar la ciencia.

Por último, pero no menos importante, al Dlab. Partiendo por el jefe, que más que jefe es un amigo, que me ha esperado todo este tiempo y ha depositado en mí un montón de confianza. A los compas del Dlab, un lugar donde se juntan locos de todo tipo, a Claudia, Jorge y Joáco que me aguantan de jefe. A Guti, Jota, Toño, Alejo, Seba, Nacho, Paolo. Mención especial para Carlos, Alberto y Yerko, que siempre se preocuparon de cuanto me faltaba, de si necesitaba algo, de apañar para charlas, marchas, juegos o cuanta lesera se me ocurriera.

A las secres de postgrado, Claudia y Janet, que se merecen un altar. A todos y todas los que contribuyeron a este trabajo, que por fin ha llegado a su final feliz.

## Resumen

La división bacteriana es un proceso esencial que ocurre gracias a un conjunto de proteínas que forman el divisoma. Las proteínas de membrana FtsW y FtsI son parte de esta maquinaria de división, que actúan como un complejo en la fase tardía, luego de la segregación del cromosoma y de la incorporación del complejo FtsQ/FtsB/FtsL. Justamente este complejo es uno de los que menos información molecular se tiene y su función específica permanece bajo un halo de dudas.

FtsI es una proteína de un paso transmembrana, de la familia de las proteínas de unión a penicilina, con actividad enzimática transpeptidasa. En tanto FtsW, es una proteína de 10 pasos transmembrana, sin función demostrada y sin estructura conocida. Tampoco tiene -hasta ahora- homólogos que permitan suponer su plegamiento. Tampoco se conoce la forma en que estas proteínas interactúan, ni cómo su función específica contribuye al proceso de división.

En esta tesis se estudiaron en profundidad ambas proteínas, desde los puntos de vista estructural y evolutivo. A partir de estos análisis se sustenta la hipótesis de que FtsW es un miembro lejano de la familia de las transglicosilasas y su función más probable es extender las cadenas de peptidoglicano, a partir del precursor lipídico Lípido II. Luego FtsI es capaz de entrecruzar dichas cadenas. Con ambas actividades concentradas en un complejo, es posible extender la cadena de peptidoglicano en el sentido de la constricción, para crear los nuevos polos de las células hijas.

Durante esta tesis fue posible además obtener modelos estructurales para ambas proteínas, que en el caso de FtsW fue refrendado por algunos experimentos y para FtsI, incluyó el modelado de su región transmembrana. Además se sugieren regiones y residuos de probable contacto entre ambas proteínas.

Finalmente se analizó el devenir evolutivo de los genes que codifican para ambas proteínas, en el reino de las bacterias, concluyendo que el complejo FtsW/FtsI está muy conservado y podría ser un elemento que co-evolucionó junto a la pared de peptidoglicano.

## Abstract

The bacterial division is an essential process that is executed by a complex of proteins, called divisome. FtsW and FtsI are membrane proteins part of this machinery, that form a complex. This complex is formed in the late phase of division after the chromosome segregation and FtsQ/FtsB/FtsL complex binding. However, the FtsW/FtsI complex have a low number of molecular evidences about its function and structure.

FtsI has only one transmembrane span domain. It belongs to the penicillin binding protein family and has transpeptidase activity. FtsW is a ten-transmembrane span protein, without a demonstrated function or structure. It doesn't have any homolog that could suggest a specific folding. The way both proteins interact each other to form the complex is also a mystery.

In this Thesis both proteins was studied in deep, from structure to evolution. These analyses support the hypothesis of FtsW being a far-member of the transglycosilase family, so its most probable function is to extend the peptidoglycan chains from the precursor Lipid II. After that, FtsI can crosslink those chains. As the complex has both enzymatic activities, it is possible to extend the peptidoglycan chain in the direction of constriction, creating the new poles of the daughter cells.

During this Thesis it was possible to obtain structural models for both proteins. The FtsW model was supported by some experiments and the FtsI model include its transmembrane domain. Moreover, specific residues and regions of both proteins are proposed to form the contact interface in the complex.

Finally, the evolution of the corresponding genes was analyzed in all the bacteria kingdom. The FtsW/FtsI is very conserved and probably is an element that evolved together the peptidoglycan wall itself.

# Índice

<b>Resumen</b>	<b>v</b>
<b>Abstract</b>	<b>vi</b>
<b>1. Introducción</b>	<b>1</b>
1.1. La división bacteriana . . . . .	1
1.2. Etapa tardía o periplásmica . . . . .	4
1.3. Maquinaria de síntesis de peptidoglicano . . . . .	8
1.4. La controversia de la función de FtsW . . . . .	10
1.5. Presentación del problema . . . . .	14
<b>2. Objetivo General</b>	<b>16</b>
<b>3. Objetivos Específicos</b>	<b>16</b>
<b>4. Materiales y métodos</b>	<b>17</b>
4.1. Obtención de secuencias . . . . .	17
4.2. Alineamientos múltiples . . . . .	17
4.3. Análisis filogenético . . . . .	17
4.4. Identificación de dominios . . . . .	18



4.5. Análisis de propiedades de secuencias . . . . .	18
4.6. Creación de Logos . . . . .	18
4.7. Software interactivo de análisis de proteínas de membrana . . . . .	18
4.8. Mutaciones correlacionadas y predicción de contactos . . . . .	19
4.9. Proyección helicoidal . . . . .	19
4.10. Biología molecular . . . . .	19
4.11. SCAM . . . . .	20
<b>5. Resultados</b>	<b>22</b>
<b>6. Objetivo Específico 1</b>	<b>22</b>
6.1. Análisis de FtsW . . . . .	23
6.1.1. Conservación . . . . .	23
6.1.2. Empaquetamiento . . . . .	31
6.1.3. Mutaciones correlacionadas . . . . .	32
6.1.4. Predicción estructural . . . . .	35
6.1.5. Estabilidad estructural . . . . .	46
6.2. Ensayos de accesibilidad a cisteína . . . . .	48
6.2.1. Síntesis de mutantes . . . . .	49
6.2.2. Curvas de crecimiento . . . . .	49

6.2.3. Resultados SCAM . . . . .	50
6.3. Análisis de FtsI . . . . .	54
6.3.1. Conservación . . . . .	54
6.3.2. Modelamiento estructural . . . . .	54
<b>7. Objetivo Específico 2</b>	<b>58</b>
7.1. Mutaciones correlacionadas . . . . .	58
7.2. Modelamiento estructural . . . . .	60
7.3. Objetivo Específico 3 . . . . .	63
7.3.1. Presencia del complejo FtsW/FtsI . . . . .	63
7.3.2. Clasificación de genes <i>ftsW/rodA</i> . . . . .	64
7.3.3. Modelamiento evolutivo del complejo FtsW/FtsI . . . . .	67
<b>8. Discusión y Conclusiones</b>	<b>71</b>
<b>Bibliografía</b>	<b>79</b>

## Índice de figuras

1. Figura 1. División bacteriana . . . . .	2
2. Figura 2 Complejo FtsB/FtsL/FtsQ . . . . .	6

3.	Figura 3. Proteínas FtsW y FtsI . . . . .	9
4.	Figura 4. Molécula de Lípido II . . . . .	11
5.	Figura 5. Síntesis de Lípido II . . . . .	12
6.	Figura 6. Topología de FtsW . . . . .	23
7.	Figura 7. Logos de secuencia para zonas TM de FtsW . . . . .	26
8.	Figura 8. Proyección helicoidal para zonas TM de FtsW . . . . .	28
9.	Figura 9. Logos de secuencia para <i>loops</i> de FtsW . . . . .	30
10.	Figura 10. Logo de secuencia de <i>loop</i> 7-8 de FtsW . . . . .	31
11.	Figura 11. Empaquetamiento de FtsW . . . . .	32
12.	Figura 12. Predicción de contactos para FtsW . . . . .	34
13.	Figura 13. Modelo <i>ab initio</i> de FtsW . . . . .	36
14.	Figura 14. Estructuras cristalográficas de enzimas transglicosilasas . . . . .	38
15.	Figura 15. Distribución del largo de los <i>loops</i> de FtsW . . . . .	41
16.	Figura 16. Mapas de contactos para enzimas transglicosilasas . . . . .	42
17.	Figura 17. Alineamiento esquemático entre 3WAK, 5F15 y FtsW . . . . .	43
18.	Figura 18. Modelo comparativo de FtsW . . . . .	45
19.	Figura 19. Dinámica molecular para modelo comparativo de FtsW . . . . .	47
20.	Figura 20. Reactivos usados para SCAM . . . . .	48
21.	Figura 21. Curvas de crecimiento de mutantes FtsW cisteína-única . . . . .	51

22.	Figura 22. Western blots con análisis de SCAM . . . . .	52
23.	Figura 23. Caracterización de la región TM de FtsI . . . . .	55
24.	Figura 24. Estructura de FtsI . . . . .	57
25.	Figura 25. Modelo de interacción entre FtsW y FtsI . . . . .	62
26.	Figura 26. Distribución de la cantidad de <i>ftsW</i> y <i>ftsI</i> por genoma . . . . .	64
27.	Figura 27. Árbol filogenético de las secuencias proteicas de FtsW/RodA . . . . .	68
28.	Figura 28. Árbol filogenético de las secuencias proteicas de FtsI/MrdA . . . . .	69
29.	Figura 29. Comparación de dendrogramas de las proteínas FtsW y FtsI . . . . .	70
30.	Figura 30. Conservación de residuos en flipasa MurJ . . . . .	72
31.	Figura 31. Modelo de interacción alternativo entre FtsW y FtsI . . . . .	76

## Índice de tablas

1.	Tabla 1. Valores de polaridad y empaquetamiento para zona TM de FtsW . . . . .	27
2.	Tabla 2. Coeficiente de correlación de Matthews para zonas TM de FtsW . . . . .	29
3.	Tabla 3. Conteo de contactos predichos entre regiones de FtsW . . . . .	35
4.	Tabla 4. Mutaciones diseñadas para ensayos de SCAM . . . . .	50
5.	Tabla 5. Conteo de contactos físicos predichos entre regiones de FtsW y FtsI . . . . .	58
6.	Tabla 6. Contactos físicos predichos entre FtsW y FtsI . . . . .	59
7.	Tabla 7. Clasificación de genes <i>ftsW/rodA</i> . . . . .	66

8. Tabla 8. Clasificación de genes *ftsI/mrdA* . . . . . 67

# 1. Introducción

## 1.1. La división bacteriana

La división celular en bacterias es un proceso complejo, sincronizado y fuertemente regulado, que permite la formación de células hijas mediante la constricción de las membranas interna y externa, gracias a la formación de un anillo proteico en el lado citoplasmático y de un septo de peptidoglicano (PG) en el periplásmico. Al igual que otros procesos celulares, una serie de interacciones proteína-proteína específicas de un conjunto de complejos proteicos son los encargados de llevar a cabo la división celular [1, 2].

En la división celular de *Escherichia coli* participan al menos una docena de proteínas (FtsZ, FtsA, ZipA, FtsE, FtsX, FtsK, FtsQ, FtsB, FtsL, FtsW, FtsI y FtsN), muchas de las cuales son esenciales para el proceso de división y forman parte del llamado divisoma (Figura 1). Luego de la replicación del cromosoma bacteriano, el primer paso de la división celular es la formación de un anillo de polímeros de FtsZ (anillo Z) que se localiza en la cara interna de la membrana celular en la mitad longitudinal de la bacteria. FtsZ es el principal elemento del citoesqueleto bacteriano, es homólogo estructural de la tubulina eucarionte [3, 4], tiene actividad GTPásica y se cree que ejerce la fuerza motriz necesaria para constreñir la membrana plasmática durante la división [5, 6]. Además, su dominio C-terminal presenta zonas flexibles implicadas tanto en la interacción con proteínas del protoanillo (ZipA y FtsA) como en la actividad GTPásica [7, 8].

Al inicio de la división celular y la formación del divisoma, los polímeros de FtsZ interactúan directamente con las proteínas ZipA y FtsA para formar la primera estructura macromolecular: el protoanillo. ZipA es una proteína bitópica de membrana, mientras que FtsA presenta una hélice anfipática que le permite interactuar con la membrana de manera periférica. Al igual que FtsZ, ambas proteínas poseen homología con elementos del citoesqueleto eucarionte: ZipA es homólogo de MAP-Tau [9], mientras que FtsA es homólogo de actina [10]. La función de estas proteínas se centra principalmente en la organización y estabilización del anillo Z en la cara

interna de la membrana plasmática y a pesar que sus funciones son redundantes, esto asegura el anclaje del anillo Z a la misma y la formación del protoanillo, lo cual es la etapa esencial del proceso de división.

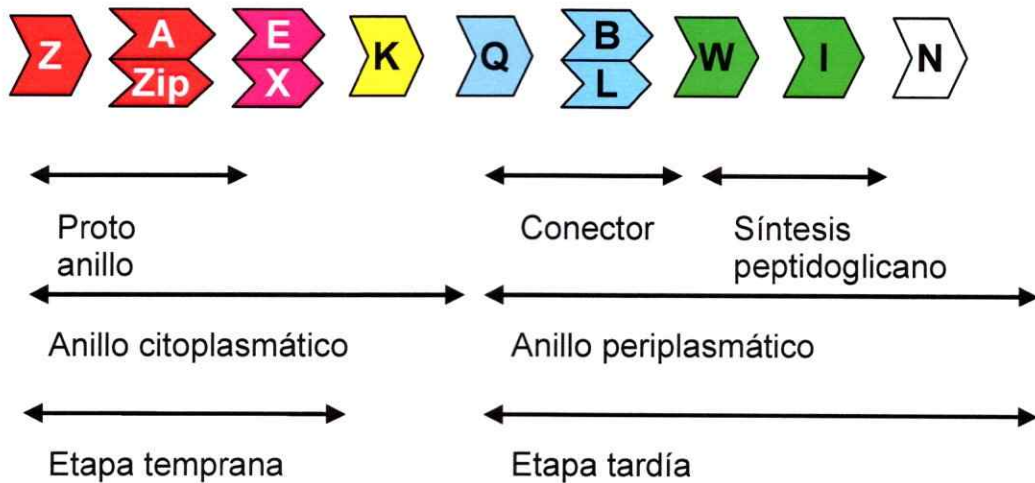


Figura 1: Esquema de las proteínas que participan en la división celular de *E. coli*, sus interacciones y su distribución durante el proceso. El protoanillo es formado por la interacción entre FtsZ, FtsA y ZipA, cuyo ensamblaje es un evento temprano de la división. La adición de FtsE/FtsX y FtsK forma el anillo citoplasmático. En una etapa tardía, se forma el conector periplasmático integrado por el complejo FtsQ/FtsB/FtsL y las proteínas involucradas en la síntesis de PG: FtsW y FtsI, seguido de FtsN. Los nombres son abreviados excluyendo el término Fts. Adaptado de [2].

Una vez formado el protoanillo, éste es estabilizado por el complejo FtsE/FtsX, localizado en la membrana interna donde actúa como un canal ABC [11]. En este momento se une FtsK al dominio 1C de FtsA [12, 13], donde acopla la segregación del cromosoma al proceso de división celular. Así el DNA replicado puede ser segregado a las dos células hijas por medio de FtsK. Esta última es una proteína de membrana de 1.329 aminoácidos compuesta de un dominio N-terminal formado por cuatro segmentos transmembrana (TM) que se localizan junto con FtsZ; un dominio central o *linker* estabiliza la interacción con proteínas de la división y; el dominio C-terminal citoplasmático que contiene un motivo de unión a ATP y es requerido para

la separación del cromosoma recién duplicado (segregación). *In vitro* el dominio C-terminal de FtsK forma complejos hexaméricos que jalan al DNA de manera ATP-dependiente y activan la recombinación de los sitios *dif*, lo cual gatilla la segregación del cromosoma bacteriano ([14, 15]). La localización final de FtsK en el protoanillo y la segregación de los cromosomas permiten comenzar con el reclutamiento de las proteínas periplasmáticas requeridas para continuar con la división bacteriana. Cabe mencionar que FtsK es parte del sistema de respuesta SOS de las bacterias [16].

Luego de un retardo temporal (que en *E. coli* pueden llegar a ser 20 min), comienza la llamada fase tardía, donde se reclutan tres proteínas bitópicas de membrana sin actividad enzimática conocida, pero esenciales para la división. Estas proteínas que localizan en el septo celular se denominan FtsQ, FtsB y FtsL en *E. coli* (sus homólogos en *Bacillus subtilis* son DivIB, DivIC y FtsL, respectivamente). Cada una contiene tres dominios: uno citoplasmático corto, una región TM y otro periplasmático. Se ha visto que las tres proteínas se ensamblan antes de su interacción con el anillo citoplasmático [17], pero FtsL y FtsB pueden hacerlo concertadamente de forma independiente a FtsQ, y sin la necesidad de otro componente previo a su reclutamiento en el sitio de división [18].

Una vez reclutado el complejo FtsQ/FtsB/FtsL, se localizan en el sitio de división dos proteínas involucradas en la síntesis de péptidoglicano: FtsW, un presunto translocador de precursores de péptidoglicano [19] requerido para la localización de FtsI [20], y FtsI, proteína de unión a penicilina (PBP) específica de la septación, llamada también PBP3 (Penicillium Binding Protein 3) [18].

Finalmente, la última proteína en reclutarse al sitio de división es FtsN, una proteína de ensamblaje tardío y de función desconocida que requiere de todos los componentes río arriba para ser reclutada al sitio de septación [18, 21]. Un modelo propuesto para la función de FtsN sugiere que ésta se extiende en el espacio periplasmático actuando como un puente entre la capa de PG y el anillo citoplasmático [22]. El reclutamiento de FtsN estabiliza al anillo [2] y a las otras proteínas del divisoma ya ensambladas. Una vez ensamblados todos los componentes



del divisoma, la célula se divide en dos células hijas.

## 1.2. Etapa tardía o periplásmica

Tradicionalmente se ha dividido todo el proceso de división en 2 etapas: la etapa temprana, que comprende desde la polimerización de FtsZ hasta el reclutamiento de FtsK y la consecuente segregación de los cromosomas. La etapa tardía, ocurre con un desfase temporal respecto a la primera (en este tiempo se debe reparar el DNA si está defectuoso) y comprende el reclutamiento del complejo FtsQ/FtsB/FtsL (Figura 2), que a su vez recluta a FtsW/FtsI, finalizando con FtsN y otras enzimas modificadoras del PG, pero que no son esenciales para la división como AmiC y PBP2B [23].

FtsQ es la primera de estas proteínas en unirse, interaccionando directamente con FtsK a través de su dominio POTRA [24]. Aunque en *E. coli*, esta proteína es esencial para la viabilidad celular, en *B. subtilis* es esencial sólo en condiciones de temperaturas extremas [25, 26]. En *B. subtilis* (donde se conoce como DivIB) sólo se requiere de su dominio periplasmático para su localización al sitio de división [27], mientras que en *E. coli* se requieren ambos dominios (citoplasmático y periplasmático)[28].

El dominio periplásmico de DivIB en *Bacillus stearothermophilus* tiene un plegamiento autónomo: dominio  $\alpha$ , dominio  $\beta$  y un probable dominio  $\gamma$  sin estructura. El dominio  $\alpha$  es el más cercano a la membrana citoplasmática y coincide con una predicción teórica de dominio POTRA (abreviación del inglés, *polypeptide-transport-associated*) asociado a funciones de chaperona en otras proteínas [24]. Además, mediante ensayos de doble híbrido e inmunoprecipitación se ha demostrado que FtsQ es capaz de interactuar con varias de las proteínas de la división, entre ellas FtsI, FtsL, FtsN, FtsB y FtsW [29]. Estos resultados sugieren que el rol de DivIB (FtsQ) podría ser de tipo chaperona al estabilizar a otras proteínas de la división y sus interacciones correspondiente [30, 31], aunque esto no se ha demostrado experimentalmente. A la fecha, solo se ha mostrado en *B. subtilis* que DivIB juega un rol crucial en la estabilidad de FtsL frente a

altas temperaturas [32].

FtsB y FtsL son dos pequeñas proteínas de 103 y 121 residuos respectivamente, que se unen a FtsQ y presentan una conformación principalmente helicoidal, desde su única región TM hasta cerca del extremo C-terminal (Figura 2). Pueden formar heterodímeros u oligómeros *in vivo* a través de sus motivos de cierre de leucinas en la región periplásmica y parte de sus dominios TMs [32, 33]. Debido a la inestabilidad intrínseca de FtsL en ausencia de FtsB en *B. subtilis* [34], ambas proteínas se asocian antes de la formación del complejo periplasmático [30], y lo mismo ha sido descrito en *E. coli* [17]. La estructura cristalográfica de FtsQ de *E. coli* [35] muestra que varias mutaciones en la región C-terminal de esta proteína desacoplan el reclutamiento del heterodímero FtsB/FtsL, lo que sugiere que el ensamblaje del multicomplejo FtsQ/FtsB/FtsL podría estar mediado por el dominio C-terminal de cada proteína, propuesta que ha sido confirmada usando mutantes truncas para FtsB y FtsL [36, 37].

A pesar que su función es aparentemente sólo estructural, la formación del complejo ternario FtsQ/FtsB/ FtsL parece ser un punto de control clave en la división celular. En *B. subtilis* se ha descrito que la expresión de FtsL está altamente regulada tanto a nivel transcripcional por la proteína de la iniciación de la replicación del DNA, DnaA [38], como a nivel post-traducciona mediante proteólisis a través de la acción de la metaloproteasa YluC [39]. Esta proteasa, homóloga de la proteasa RseP de *E. coli*, participa en mecanismos de proteólisis de intramembrana regulados (RIP) mediante el reconocimiento de estructuras anómalas en hélices TMs. Todo esto enmarcado dentro de un complejo tipo de regulación post-transcripcional [40]. La proteólisis de FtsL mediada por YluC altera los niveles celulares de FtsL, lo cual regula la formación del complejo FtsQ/FtsB/FtsL y a la vez la actividad del divisoma [39, 41].

El complejo FtsQ/FtsB/FtsL es importante como factor estructural clave para conectar los eventos citoplásmicos (fase temprana, anillo Z) con los eventos periplásmicos (fase tardía, síntesis de PG). Con la información disponible [35, 36, 37] se ha generado un modelo de la estructura de este complejo [42], postulando dos posibles configuraciones, un heterotrímero y un heterohexámero (2:2:2). De este trabajo destaca la cantidad de residuos hidrofóbicos que

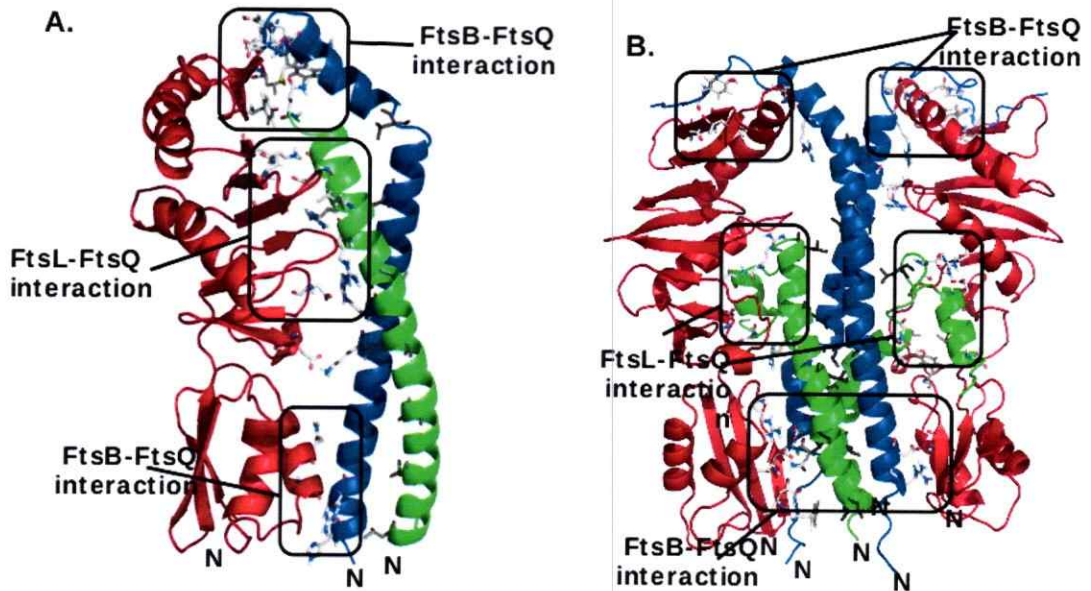


Figura 2: Modelos de la región periplásmica del complejo FtsB(azul)/FtsL(verde)/FtsQ(rojo) generados por modelamiento molecular. Se muestran en alambres las cadenas laterales de los residuos identificados en las interfases de interacción entre proteínas. La parte inferior del esquema corresponde a la región N-terminal de cada proteína que conecta con su correspondiente hélice TM (no resuelto por el modelo). **A.** Complejo con estequiometría 1:1:1. **B.** Complejo con estequiometría 2:2:2. Tomado de [42]

quedan disponibles para interactuar con otras proteínas como FtsW y FtsI. También quedan disponibles algunas regiones del dominio POTRA de FtsQ, así como de su dominio C-terminal, para poder interactuar con estas mismas proteínas [42].

Luego del complejo FtsQ/FtsB/FtsL, se unen al divisoma las proteínas FtsW y FtsI. FtsI es una PBP tipo B, con un dominio transpeptidasa (TP), que participa directamente en la formación de PG, entrecruzando cadenas de éste, siendo su actividad enzimática esencial para la división [43]. Una vez localizada FtsI, puede interactuar con otras PBPs como PBP1B [44, 23] que no son esenciales, pero que poseen además actividad transglicosilasa, ausente en FtsI. Estudios han confirmado que para que FtsI se localice en el septo de división son suficientes sus primeros 56 residuos N-terminales [45], aunque estudios posteriores muestran que incluso

bastaría con la región TM [46, 47].

Por su parte, FtsW es una proteína con 10 segmentos TM, que se asoció desde su descripción como compañera de FtsI, sin haber claridad sobre su función [48, 49]. Se ha reportado que tiene actividad "flipasa" sobre la molécula de Lípido II [19], precursor de PG y sustrato de FtsI, pero este resultado ha sido cuestionado [50]. Mediante mutaciones y proteínas truncas se han identificado algunas regiones como probables de interacción con FtsI, como el "loop" periplasmático entre las regiones TM9 y TM10 [51, 52].

Ambas proteínas necesitan que las otras proteínas de la división estén presentes [53, 20], y se necesitan de la una a la otra para localizar en el septo de división [18], aunque FtsW podría localizar sin necesidad de FtsI [20]. A su vez FtsW y FtsI son necesarias para la correcta localización de FtsN [16]. Se ha demostrado mediante inmunoprecipitación y ensayos de transferencia de energía resonante (FRET) *in vivo* que estas dos proteínas pueden formar un complejo entre ellas y luego localizar en el septo de división [54], tal como el complejo FtsQ/FtsB/FtsL [17].

En este complejo FtsW/FtsI, las evidencias indican que FtsW sería quien contiene el determinante para localizar en el divisoma, así como de la unión al complejo FtsQ/FtsB/FtsL, aunque aun no existen pruebas de esto último. De hecho, existen apenas dos evidencias respecto a la interacción entre ambos complejos en *E. coli*, contándose una muy indirecta, donde la sobreexpresión de FtsQ aumenta los efectos negativos (longitud de bacterias que no se dividen) de mutantes de FtsI que no localizan en el septo [55]. La otra es que ambos complejos, son capaces de reclutar al otro, cuando sólo uno de ellos puede localizar en el septo (por fusión con ZapA en cepas carentes de FtsA) [18].

Por el modelo de interacción y reclutamiento de las proteínas en el divisoma de *E. coli* se acepta que luego del complejo FtsQ/FtsB/FtsL debe reclutarse FtsW/FtsI, para finalizar con FtsN. Diversos estudios de interacción usando experimentos de doble híbrido muestran múltiples interacciones entre estas proteínas [29, 56]. Sin embargo, hay poca información directa de la interacción de estos dos subcomplejos, que a su vez se forman de manera independiente del

divisoma antes de unirse a éste.

### 1.3. Maquinaria de síntesis de peptidoglicano

En el final de la fase tardía de la división bacteriana se recluta lo que se conoce como la maquinaria de síntesis de PG, compuesta por las proteínas FtsW y FtsI. Este complejo tiene por función la síntesis de pared celular en el sitio de septación, de modo de que ésta vaya extendiéndose en el sentido de la constricción, y formar pared en los nuevos polos de las bacterias hijas. Sin embargo, este evento por sí solo no impulsa físicamente la separación. La polimerización/despolimerización de FtsZ en el citoplasma y la participación de todas las otras proteínas del divisoma contribuyen a este fenómeno y aun se debate cuál de ellos es el principal (revisado en [57]).

FtsW es una proteína de membrana multipaso de 414 residuos y 10 regiones TM (Figura 3). Su gen está codificado en la región *dcw* (division and cell wall) y existe en prácticamente todas las bacterias que poseen una pared de PG [58, 49]. Pertenece a la familia de proteínas SEDS (del inglés: *shape, elongation, division and sporulation*) que incluye a otras dos proteínas: RodA, que participa en la elongación de *E. coli* y otras bacterias cilíndricas [59] y SpoVE, que participa en el proceso de esporulación de *B. subtilis* y otras bacterias Gram-positivo [60], compartiendo un 30 % de identidad entre ellas [48]. Por largos años solo se especuló con la función de estas proteínas [61, 62] y no hubo ningún estudio destinado a esclarecer su función, hasta que se logró determinar que FtsW puede traslocar Lípido II [19, 63], precursor de PG. A pesar que este resultado ha sido cuestionado, cualquiera sea la función de FtsW, muy probablemente sus homólogos RodA y SpoVE comparten esa función.

FtsI, una PBP tipo B con función transpeptidasa tiene 588 residuos y es esencial para la división [43]. Comparte homología con todas las transpeptidasas bacterianas, aunque carece de actividad trasglicosilasa como otras enzimas de esta familia. Está compuesta por tres dominios (Figura 3): el dominio de unión a membrana que consiste en una hélice TM; el dominio no

catalítico o dominio de función desconocida y el dominio catalítico. Este último es sensible a ser inactivado por antibióticos beta-lactámicos.

Ambas proteínas, FtsW y FtsI, forman parte crucial del divisoma [16] y pueden formar un complejo entre ellas antes de unirse al resto de las proteínas de la división [54]. De las dos, FtsW parece ser quien contiene el determinante para unirse al resto del divisoma, pues en ausencia de FtsI puede localizar en el septo de división y no al revés [20]. Sin embargo, no se conoce ningún mecanismo o con cuál otra proteína interactúa directamente. De hecho, en un trabajo donde investigan el efecto *in vivo* de una gran cantidad de mutantes y proteínas FtsW truncas, ninguna de ellas falla en la localización en el septo [51].

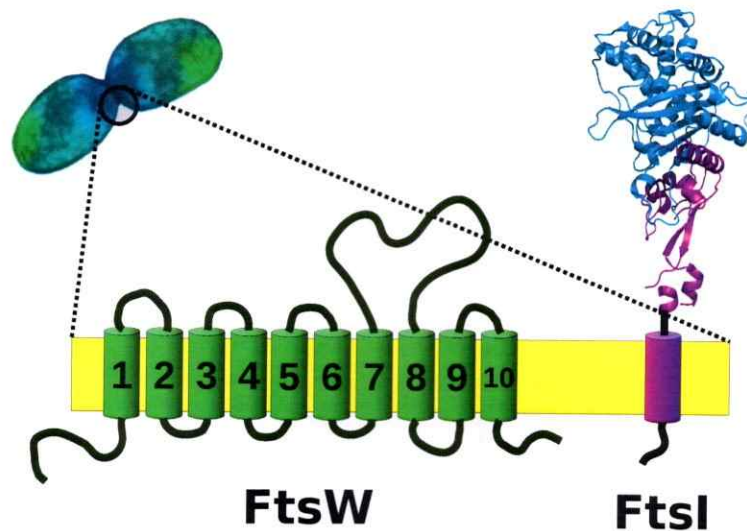


Figura 3: Representación esquemática de las proteínas FtsW y FtsI. El rectángulo amarillo representa la membrana interna, donde la parte superior a éste corresponde al periplasma y la parte inferior al citoplasma. Las regiones TM de FtsW aparecen numeradas. Se muestra la región periplásmica de FtsI modelada, utilizando como templado estructuras cristalográficas de otras PBP. En rosado se muestra la región TM de FtsI (residuo 19 al 39) y el dominio central (residuos 40 al 236); en celeste el dominio catalítico (residuo 237 al 577).

FtsI, por sí sola es incapaz de localizar en el septo y requiere de FtsW para hacerlo. Se ha encontrado que es la región TM de FtsI la que posee todo lo necesario para ser reconocida

por FtsW, aunque no se puede descartar la participación de regiones del dominio no catalítico, aunque en menor grado de importancia [45, 47]. Este dominio no catalítico sí es necesario para interactuar con FtsN, la próxima y última proteína esencial para la división bacteriana [46].

Respecto al mecanismo de interacción entre FtsW y FtsI, se sabe que la primera recluta a la segunda al sitio de división a través de la zona TM de FtsI, aunque no se ha logrado determinar cuál o cuáles de las 10 zonas TM de FtsW son necesarias para ello. La única evidencia relacionada es que el doble mutante P268A/P275A no puede reclutar a FtsI *in vivo*, pero sí es capaz de localizarse en el sitio de división [51], aunque lo último no pudo ser confirmado con certeza en un estudio posterior, donde fue analizado por ensayos de doble híbrido [54]. Estos residuos de prolina mutados se encuentran en una región bastante conservada entre las proteínas de la familia y se localizan en el "loop- periplasmático entre las regiones TM9 y TM10, hacia el extremo C-terminal de la proteína.

#### 1.4. La controversia de la función de FtsW

Desde su descripción se ignoró por completo la función molecular de FtsW, RodA o cualquiera de sus homólogos. Todas pertenecen a una misma familia de proteínas, que no tiene homología con ninguna otra. La falta de trabajos bioquímicos en este tipo de proteínas parecía indicar dificultades metodológicas. Solo se especuló con alguna función relacionada con Lípido II (Figura 4), el sustrato de las PBP, que acompañan a FtsW/RodA.

Sin embargo, la falta de una función pareció no importar a quienes trabajan en el tema, pues hay un trabajo [51] donde se encontró un mutante de FtsW, R243Q (sobre el "loop" periplasmático que conecta los TM 7 y 8), que es letal e impide la división y fue relacionado con pérdida de función. Función, curiosamente, desconocida en ese momento.

Recién en 2011, se publicó un trabajo donde se determinaba *in vitro* que la función de FtsW sería traslocar, de manera específica, un precursor de PG de un lado de la bicapa a otro. O sea FtsW, y por lo tanto sus homólogas, serían flipasas de Lípido II (Figura 5). En un trabajo

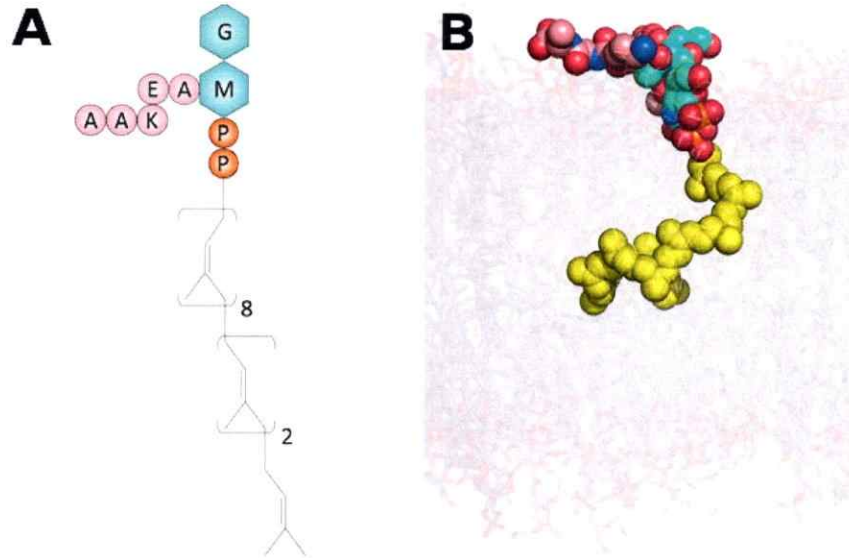


Figura 4: Molécula de Lípido II. **A.** Representación esquemática. **B.** Modelo molecular, embebido en una membrana. La cola lipídica aparece en amarillo; grupos fosfato en naranja; azúcares en celeste; y el pentapéptido en rosado. Adaptado de [64].

posterior de este mismo laboratorio, se profundiza en los residuos que serían clave para esta función, identificando la zona TM4 como clave para la especificidad, mencionando los residuos R145 y K153 [63]. En este trabajo también se determina que FtsW puede traslocar fosfolípidos, sin especificidad, pero por un mecanismo distinto, que no necesita la TM4.

El precursor Lípido II, es sintetizado en la cara interna de la membrana citoplasmática de la bacteria (Figura 5) gracias a un conjunto de enzimas, que lo sintetizan modularmente. Primero una molécula de glicopentapéptido se une a un *carrier* lipídico derivado del isopreno, el undecaprenol. A esta molécula (Lípido I) la enzima MurG le agrega una unidad de GlcNAc, para formar Lípido II. Para que las proteínas que sintetizan PG - como las PBPs - usen el Lípido II, éste debe estar disponible para unirse a los respectivos sitios activos que se encuentran en el periplasma. Por ello es necesaria una proteína que lleve a cabo su traslocación, dado que la molécula de Lípido II es muy grande y polar como para hacerlo de manera espontánea [65]. Esto sería llevado a cabo por FtsW.



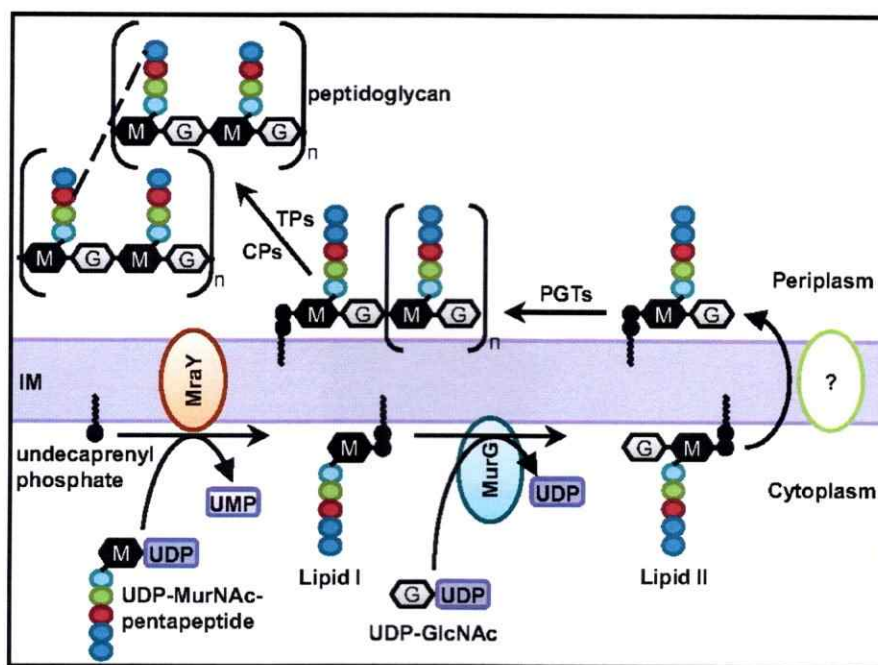


Figura 5: Síntesis de Lípido II en *E. coli*. Tomado de [66]

Esta función de flipasa tiene cierto sentido, pensando en la formación del complejo FtsW/FtsI, y sus homólogos RodA/PBP2 y SpoVE/SpoVD, ya que indicaría la formación de un complejo, no sólo estructural, sino que también funcional. A partir de esto se puede generalizar que el complejo proteína SEDS/PBP tiene la función de extender la pared bacteriana a partir del precursor sintetizado citoplasmáticamente [67]. Cada complejo sería específico para cada condición fisiológica (división, elongación o esporulación).

Sin embargo, investigaciones del grupo liderado por Natividad Ruiz en Ohio State University, llevadas prácticamente en paralelo con el grupo holandés responsable de los trabajos anteriores, indican que sería otra la proteína encargada de esta función [66]. MurJ es un transportador ABC de 14 pasos TM, perteneciente a la superfamilia de exportadores MOP (*multidrug/oligosaccharidyl-lipid/polysaccharide*). A esta familia pertenecen otras flipasas de oligosacáridos unidos a poliprenoles, como Wzx del sistema de biosíntesis de antígeno O. De hecho son homólogas y se construyó un modelo estructural a partir de la estructura de Wzx, modelo refrendado por experimentos de accesibilidad de cisteínas [68].

Con ensayos *in vivo* usando DAP (aminoácido únicamente usado para la síntesis de peptidoglicano) marcado con tritio y ColM (colicina que corta la sección soluble de Lípido II, solo si está en la cara externa) demostraron que al inhibir la función de MurJ, el Lípido II se vuelve resistente al tratamiento con ColM, lo que indica que éste se queda en la cara interna de la membrana [50]. De hecho, estos mismos experimentos en cepas *knock out* para FtsW indican que aun hay traslocación de Lípido II, descartando que FtsW sea la traslocadora de este lípido y de paso contradiciendo las evidencias anteriores [19, 63].

Con esta controversia presente, durante septiembre de 2015 aparece un completo trabajo [69] de modelamiento estructural *ab initio* de familias de proteínas sin estructura conocida ni homólogos con estructura conocida. En este trabajo se extiende la funcionalidad del programa de modelamiento Rosetta [70], incluyendo análisis de perfiles de secuencia y mutaciones correlacionadas, logrando modelar 58 estructuras de familias de proteínas procarióticas. Una de estas familias de proteínas es SEDS, y los autores modelan la proteína FtsW basándose remotamente en la estructura de una glicosiltransferasa eucarionte. Esta proteína usada como molde [71], cataliza la transferencia de oligosacáridos desde un carrier lipídico (dolicol) hasta residuos de asparagina en proteínas con una secuencia consenso en particular.

A pesar que la similitud de secuencia es muy baja (35%), según los autores hay una coincidencia entre el mapa de mutaciones correlacionadas en FtsW y el mapa de contactos en el cristal de la transglicosilasa para 7 de sus 10 zonas TM [69]. Esto, complementado con el algoritmo *ab initio* de Rosetta [70] para modelar fragmentos pequeños, les permite modelar FtsW en toda su extensión. Dado el foco del trabajo donde describen otros tantos modelamientos, los autores no profundizan en detalles ni en refinamientos estructurales necesarios.

Un año después de este trabajo, aparecen dos artículos que detectan experimentalmente una función transglicosilasa en RodA, el cercano homólogo de FtsW [72, 73]. En Meeske *et al.* 2016, incluso se indentifican dos residuos fundamentales para la función, Trp105 y Asp280, conservados en FtsW. En ambos trabajos los autores discuten que probablemente la función de toda proteína de la familia SEDS es la actividad transglicosilasa, que en conjunto con la

actividad transpeptidasa de sus compañeras PBP tipo B (como FtsI), conforman la maquinaria elemental para la síntesis de PG.

### 1.5. Presentación del problema

El complejo FtsW/FtsI es parte esencial de la maquinaria de división celular. Ambas proteínas forman un complejo *in vivo* independiente del resto del divisoma, que por lo menos en cuanto a la función de FtsI, se sabe que entrecruza cadenas de PG en los nuevos polos de las bacterias hijas en formación. Sin embargo, la estructura de este complejo, las regiones que intervienen en su interacción y lo que es muy relevante, la real función de FtsW, son temas aun no resueltos y que podrían dar información clave en el entendimiento del proceso de división bacteriana.

En cuanto a la interacción de ambas proteínas, no se sabe qué región de FtsW recluta a FtsI, y cual sería el rol del *loop* entre las regiones TM9 y TM10 que fue sugerido como importante para esta interacción. Para abarcar estas interrogantes desde un punto de vista molecular, inevitablemente surge la necesidad de considerar temas más básicos, concernientes a la estructura de estas proteínas y la relación con su función.

Producto de la controversia respecto a la función específica de FtsW, el tener información estructural respecto al complejo completo, o a las regiones peptídicas involucradas en la interacción, podría ser relevante. Si efectivamente en FtsW hay un poro que cruce de lado a lado de la membrana, o una región abierta hacia la parte lipídica de la bicapa (como en SecY) [74], podría indicar la posibilidad de que trasloque alguna molécula. Por el contrario, la ausencia de este poro y la identificación de posibles residuos catalíticos, podrían apoyar la hipótesis de que FtsW es en realidad una transglicosilasa, tal como se demostró para RodA [72].

Así también el tema podría ser abordado desde un punto de vista evolutivo, al tener en cuenta que -en teoría- cada proteína de la familia SEDS actúa en compañía de una PBP tipo B. Funciones complementarias requieren necesariamente residuos y regiones que hayan co-evolucionado. Estos análisis estructural y evolutivo, podrían también indicar si es que FtsW tiene las caracte-

rísticas necesarias para ser una flipasa, o si es que siempre está acompañado de una PBP.

## **2. Objetivo General**

Obtener un modelo tridimensional para el complejo FtsW/FtsI y analizar la relación de éste con su función

## **3. Objetivos Específicos**

1. Analizar estructuralmente las proteínas FtsW y FtsI, proponiendo un modelo estructural para cada una
2. Analizar la formación del complejo FtsW/FtsI y sus implicancias funcionales
3. Estudiar las relaciones evolutivas del complejo FtsW/FtsI en el reino de las bacterias

## 4. Materiales y métodos

### 4.1. Obtención de secuencias

Las secuencias de FtsW, FtsI, RodA y MrdA, de *E. coli* fueron obtenidas de la base de datos KnowledgeBase Uniprot. Las secuencias de los genomas analizados y de las secuencias proteicas correspondientes se obtuvieron de la base de datos Microbial Genome Database for Comparative Analysis (MBGD) [75]. Las comparaciones entre secuencias se hicieron con el programa seqret de la *suite* EMBOSS usando valores por defecto para puntajes, y la matriz de sustitución BLOSUM62.

### 4.2. Alineamientos múltiples

Los alineamientos múltiples de secuencia se realizaron con el programa Muscle [76], de manera iterativa hasta alcanzar un alineamiento estable, usando la matriz de sustitución BLOSUM62.

### 4.3. Análisis filogenético

Para todos los dendrogramas mostrados se hicieron muestreos de 1000 subgrupos de secuencias, a partir de alineamientos múltiples. Con cada uno de estos subgrupos se construyó un dendrograma usando el programa Fasttree [77]. Con los 1000 dendrogramas se construyó un árbol consenso usando el programa consense de la *suite* para análisis evolutivos Phylip [78]. La evaluación del modelo de ajuste evolutivo se hizo con programa ProtTest [79].

#### 4.4. Identificación de dominios

Para la identificación de dominios presentes en cada secuencia, se utilizaron las descripciones provistas por Pfam. Estas definiciones, escritas como modelos ocultos de Markov (HMM), se alinearon con cada secuencia de interés utilizando la suite de programas HMMER 3.0 [80].

#### 4.5. Análisis de propiedades de secuencias

La predicción de zonas TM se hizo con el programa Phobius [81]. La predicción de las zonas TM que podrían conformar un poro se hizo con el programa Memsat-SVM [82], disponible a través del servidor web del University College London.

#### 4.6. Creación de Logos

Para la creación de logos de secuencia a partir de alineamientos múltiples, se utilizó el programa Skylign [83], que utiliza información de modelos ocultos de Markov para extraer información de la entropía en cada columna.

#### 4.7. Software interactivo de análisis de proteínas de membrana

Durante el desarrollo de esta Tesis, se elaboró un software *ad hoc* de análisis de proteínas de membrana interactivo. Este programa sirvió para centralizar un conjunto de herramientas de análisis existentes, así como otras que se desarrollaron específicamente. El programa denominado *memprot* se escribió en el lenguaje *python* con el intérprete versión 2.7.9. Este programa usa las herramientas ya descritas, así como otras de licencia GPL entre las que podemos mencionar a *gnuplot* para algunos gráficos, la suite de análisis MDAnalysis [84] para los mapas de contactos y *numpy* para cálculos matemáticos avanzados.

#### 4.8. Mutaciones correlacionadas y predicción de contactos

Se utilizaron 3 programas distintos para calcular mutaciones correlacionadas o predecir contactos físicos a partir de estos. Gremlin [85], PSICOV [86] y Cmat [87]. Los resultados se normalizan para que vayan de cero (mínimo) a uno (máximo). Se promedian por cada par de residuos para obtener un mapa consenso.

#### 4.9. Proyección helicoidal

Los análisis de proyección helicoidal de cada TM, se realizan asumiendo una hélice  $\alpha$  de geometría teórica ideal, con  $100^\circ$  de rotación entre cada residuo consecutivo, vista desde su extremo C-terminal. Los gráficos que se muestran al interior de cada hélice, corresponden al histograma polar de cualquier propiedad numérica de los residuos. Se calcula representando como coordenada polar el valor de dicha propiedad para cada residuo en función del ángulo correspondiente en la hélice. Los valores para cada residuo de una hélice completa se promedian polarmente, para obtener la dirección del vector de dicha propiedad. Esto para cada secuencia en el alineamiento. Luego se hace un histograma de todas las direcciones de los vectores de todas las secuencias, y este histograma se grafica polarmente al interior de cada hélice.

#### 4.10. Biología molecular

*Construcción de mutantes.* Se construyó un derivado del gen *ftsW* con un tag FLAG [88] adecuado para su visualización por WB y un promotor que permita su transcripción. Este gen es un mutante doble C174S/C361S, que no posee residuos de cisteínas. Este gen se denomina *ftsW* $\Delta$ Cys y sobre este se construyeron los mutantes descritos en la Tabla 1 (Genescript, USA), todos sobre un plasmidio pUC57 con un promotor *lac* para su adecuada expresión usando el inductor IPTG.



Estos plasmidios se usaron para transformar cepas *E. coli* EC912 [20]. Esta cepa tiene el gen cromosomal de FtsW interrumpido por un gen de resistencia a kanamicina y un gen extra de FtsW en otra ubicación del cromosoma, controlado por un promotor inducible por arabinosa y reprimible por glucosa.

#### 4.11. SCAM

*Cultivos bacterianos y transformaciones.* Las bacterias se mantuvieron a  $-80\text{ }^{\circ}\text{C}$  en solución de glicerol 50% v/v. Se utilizó medio Luria-Bertani (LB), con o sin agar, suplementado con ampicilina (50  $\mu\text{g/ml}$ ), cloranfenicol (50  $\mu\text{g/ml}$ ), kanamicina (50  $\mu\text{g/ml}$ ), arabinosa (0,2% v/v), glucosa (0,2%  $\mu\text{g/ml}$ ) o IPTG (0,1 mM), según fuera necesario. Para la transformación de las bacterias con los respectivos plasmidios se usó quimiotransformación. Medio LB suplementado según fuera necesario fue inoculado e incubado *overnight*. A partir de ese pre-inóculo se inocularon 100 ml de cultivo LB, también suplementados, e incubados a  $37\text{ }^{\circ}\text{C}$  hasta alcanzar  $\text{DO}_{600}$  entre 0,4-0,5. El volumen total de cultivo se incubó en hielo durante 30 min y posteriormente se centrifugó durante 10 min, a  $4\text{ }^{\circ}\text{C}$  y 7.000 rpm.

El sobrenadante se eliminó y las células se suspendieron en 50 mL de solución de  $\text{CaCl}_2$  0,1 M. La suspensión bacteriana nuevamente se centrifugó y las células suspendidas en 500  $\mu\text{l}$  de solución de  $\text{CaCl}_2$  0,1 M y glicerol 30% v/v. Este volumen se fraccionó en 50  $\mu\text{l}$  por reacción y se mantuvieron a  $-80\text{ }^{\circ}\text{C}$  para su uso posterior. Para la transformación, una alícuota de células quimiocompetentes se descongeló en baño agua-hielo, luego se añadió plasmidio (entre 100 y 200  $\mu\text{g}$  de plasmidio total). La suspensión de bacterias se agitó suavemente e incubó en hielo durante 30 min. Luego las células se incubaron a  $42\text{ }^{\circ}\text{C}$  por 60 seg e inmediatamente se pusieron en hielo por 5 min. Luego se agregó 900  $\mu\text{l}$  de medio LB sin antibióticos y se incubaron durante 60 min para la recuperación de las células y expresión de la resistencia. Las curvas de crecimiento se hicieron mediante mediciones continuas y agitación a  $37\text{ }^{\circ}\text{C}$ .

*SCAM.* Se creció cada cepa en las condiciones descritas hasta que llegaron a una  $\text{DO}_{600}$  de

1.0. Se colectaron, lavaron, y se suspendieron en PBS. Se dividieron en 3 tubos iguales por cepa, donde uno se usa para tratamiento (T) y los otros para control negativo (C-) y positivo (C+). A la muestra T se agregó 5 mM de *N-etil-maleimida* (NEM; Sigma, USA) y se dejó con agitación suave a temperatura ambiente por 5 min. Pasado el tiempo se agregó 500 uL de PBS para detener la reacción. Se centrifugaron todas las células y se suspendieron en buffer de lisis (urea 6M, SDS 1 %, Tris-HCl 15 mM pH 7,4). Las células se resuspendieron por pipeteo y se agregó 20 mM de *metil-polietilenglicol-maleimida* (MalPEG; Sigma, USA ) en *dimetil formamida* (DMF) a T y C+. Se agitaron suavemente por 5 min a temperatura ambiente, protegidos de la luz. Se agregó buffer de carga y se incubaron por 30 min a 42 °C. Con estas muestras se corrió un *western blot* (WB), usando anticuerpo monoclonal anti tag FLAG (ThermoFisher, USA) para revelar seguido de anticuerpo secundario acoplado a peroxidasa de rábano (HRP).

## 5. Resultados

### 6. Objetivo Específico 1

Para analizar estructuralmente proteínas de membrana existe un conjunto de herramientas descritas en la bibliografía que incluyen programas de predicción, inclusión de datos evolutivos o de evidencias experimentales. Sin embargo, no existe un programa capaz de integrar esos resultados en un análisis integrado. Por ello, durante el desarrollo de esta Tesis se elaboró un programa de análisis de proteínas de membrana, que integra diversos *softwares* existentes y otros métodos desarrollados *ad hoc* en una plataforma gráfica, amigable e interactiva, que permite además leer los formatos más usados en bioinformática (Uniprot, FASTA, CLUSTAL, PIR, etc.) y es altamente portable. Este programa denominado temporalmente *memprot*, permite leer secuencias con o sin información de la topología de membrana o de estructura secundaria, que se carga de manera gráfica en la pantalla (Figura 6). El esquema de la topología es además interactivo, pues se puede colorear por distintas propiedades (hidrofobicidad, packing, conservación) y se pueden seleccionar residuos o regiones para otros análisis específicos. El programa puede cargar alineamientos múltiples y proyectar la conservación sobre la secuencia. También dispone de algunos visualizadores específicos, como uno del alineamiento múltiple (donde se pueden identificar las regiones TM) y otro de la proyección en el eje Z (o proyección tipo *wheel*) de las zonas TM. El programa permite también cargar información de mutaciones correlacionadas generadas externamente y predecir las zonas TM en caso de no disponer de una definición primaria.

Utilizando este programa se realizaron la mayoría de los análisis presentados en esta Tesis, así como las figuras de representación. Los algoritmos específicos diseñados se explican junto a los respectivos resultados.

## 6.1. Análisis de FtsW

### 6.1.1. Conservación

FtsW es una proteína multipaso, con 10 zonas transmembrana que pertenece a la familia de proteínas SEDS. En esta familia hay sólo proteínas procarióticas, bastante similares entre ellas, con 10 zonas TM cada una y función molecular desconocida a la fecha. Se presume que todas estas proteínas podrían tener la misma función en contextos fisiológicos diferentes, como la división en FtsW, la elongación en RodA o la esporulación en SpoVD.

Para comenzar el análisis de conservación se colectaron 1.507 genes de *ftsW/rodA* pertenecientes a la familia SEDS, extrayendo las correspondientes secuencias proteicas y alineándolas. De este alineamiento múltiple se calculó el nivel de conservación entre ellas, el cual se proyectó sobre una topología 2D de la proteína (Figura 6)

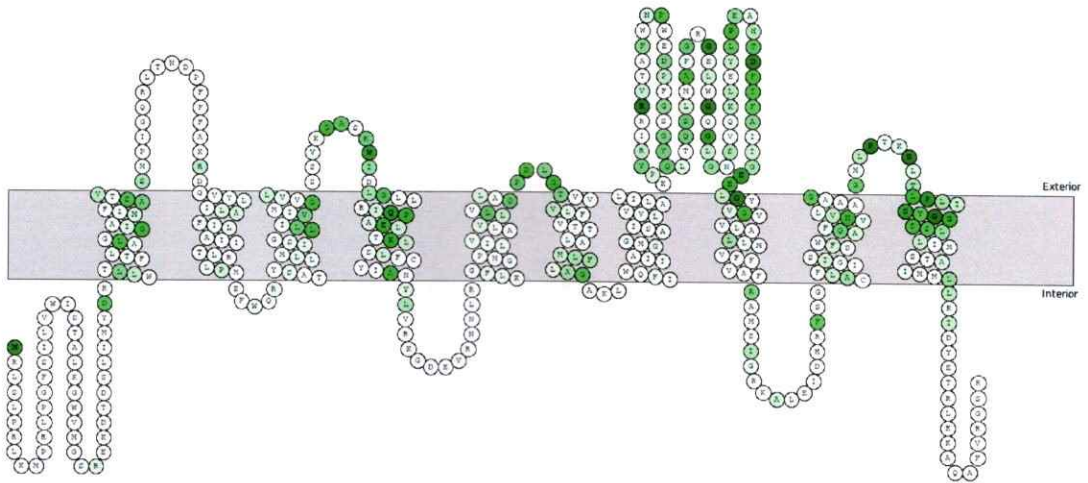


Figura 6: Topología 2D de FtsW (usando la definición de zonas TM de Lara *et al.* [49]), mostrando en escala de verde los valores de conservación para cada residuo, usando un alineamiento múltiple de 1.507 secuencias de proteínas de la familia SEDS. Mientras más oscuro el verde, mayor conservación. Figura se generó usando el software *memprot* desarrollado durante esta Tesis.

De un análisis visual se identificaron 5 regiones TMs más conservadas que las otras: TM4, TM8 y TM10 con algunos aminoácidos altamente conservados (verde oscuro); y TM1 y TM9 con varios residuos conservados. A su vez, los loops periplasmáticos, con excepción del loop 1-2, están bastante conservados.

La mayoría de los residuos conservados en las zonas TMs son glicinas, lo que tiene sentido considerando que este residuo es importante en la interacción hélice-hélice dentro de la membrana [89]. Además se puede observar algunos residuos polares como Gln147, Glu150 y Lys153 en TM4, con los dos últimos que podrían estar formando un par iónico; N359 en TM9 y la región 378-383 en TM10.

Por su parte, en los loop periplasmáticos se observa Arg137-Trp138 en el loop 3-4 y Arg246, Asp297, además de varios residuos de glicina en el loop 7-8. Cabe destacar que los residuos Trp138 y Asp297 altamente conservados coinciden con los residuos Trp105 y Asp280 identificados como esenciales para la función en RodA [72].

La mayoría de los residuos con alto nivel de conservación están orientados hacia el periplasma, ya sea en loops o en zonas TM. Esta observación es interesante considerando que una de las posibles funciones de FtsW es la de flipasa específica de Lípido II. Si este fuera el caso, se hubiese esperado algún residuo o loop conservado en la cara citoplasmática de la proteína, que sirviera para el reconocimiento específico de este lípido, cosa que no se observa.

### *Zonas transmembrana*

Para analizar la conservación o características evolutivas específicas cada zona TM, hace falta tener un alineamiento múltiple óptimo de estas regiones. Los programas de alineamiento múltiple existentes se basan en alinear globalmente la totalidad de las secuencias dadas, sin procurar entre regiones particulares al interior de cada secuencia. Esto puede significar una desventaja en proteínas de membrana, considerando que las zonas TMs tienen características distintivas, tanto de polaridad [90] como de conservación [91, 92]. De hecho se han construido matrices de sustitución específicas para estas regiones [91, 92].

A partir de estas características diferenciales y de la disponibilidad de matrices de sustitución específicas, se han diseñado algunos programas específicos para el alineamiento múltiple de proteínas de membrana. Sin embargo, estos programas tienen características no aptas para el trabajo de esta Tesis, como: buscar homólogos de manera automática en PSI-Coffee [93]; disponer solo de una versión *web* con una cantidad limitada de secuencias en PRALINE<sup>TM</sup> [94], basarse en la disponibilidad de estructuras tridimensionales en 3D-Coffee [95], o no tener el programa disponible en AlignME [96].

Debido a la falta de un programa adecuado para este trabajo, en el curso de esta Tesis se diseñó un algoritmo nuevo para optimizar alineamientos múltiples de proteínas de membrana helicoidales, utilizando matrices de sustitución específicas, modelos ocultos de Markov [97] y predicción de zonas TM usando el programa Phobius [81], que hasta ahora es el que ha generado mejores resultados. Una innovación del programa es que si bien utiliza predicciones para las secuencias sin información topológica, complementa estas predicciones con la definición de regiones TM establecida para un miembro de referencia, que en este caso es FtsW de *E. coli*. La definición de la topología TM de FtsW fue obtenida experimentalmente [49], por lo tanto es más confiable que las predicciones descritas y tiene prioridad en el algoritmo desarrollado, a la hora de buscar un inicio y fin consenso de cada TM.

Utilizando este programa fue posible obtener sub-alineamientos específicos para cada región TM y analizar diferencialmente su conservación. Una forma visual de hacer esto es generar *logos* de secuencias, que codifican en el tamaño de cada símbolo la conservación de cada residuo en cada posición o columna. En la Figura 7 se muestran los logos para cada una de las 10 regiones TM de FtsW. Se puede observar un bajo nivel de conservación para TMs especialmente en TM1, TM2, TM5 y TM6. Además se ven algunos residuos particulares, como los ya mencionados Gln147, Glu150 y Lys153 en TM4, N359 en TM9 y la región 378-383 en TM10. Esto corrobora el análisis visual mostrado en la Figura 6.

Las zonas TM, a pesar que pueden tener residuos puntuales altamente conservados como los recién descritos, son en general poco conservadas. Sin embargo, ciertas características específicas

podrían conservarse, independiente de la identidad del residuo, como la hidrofobicidad. Para explorar la conservación de la hidrofobicidad de cada TM se utilizó la escala de hidrofobicidad de Engelman, Steitz y Goldman [98] promediada en los residuos de cada TM. En la Tabla 1 se

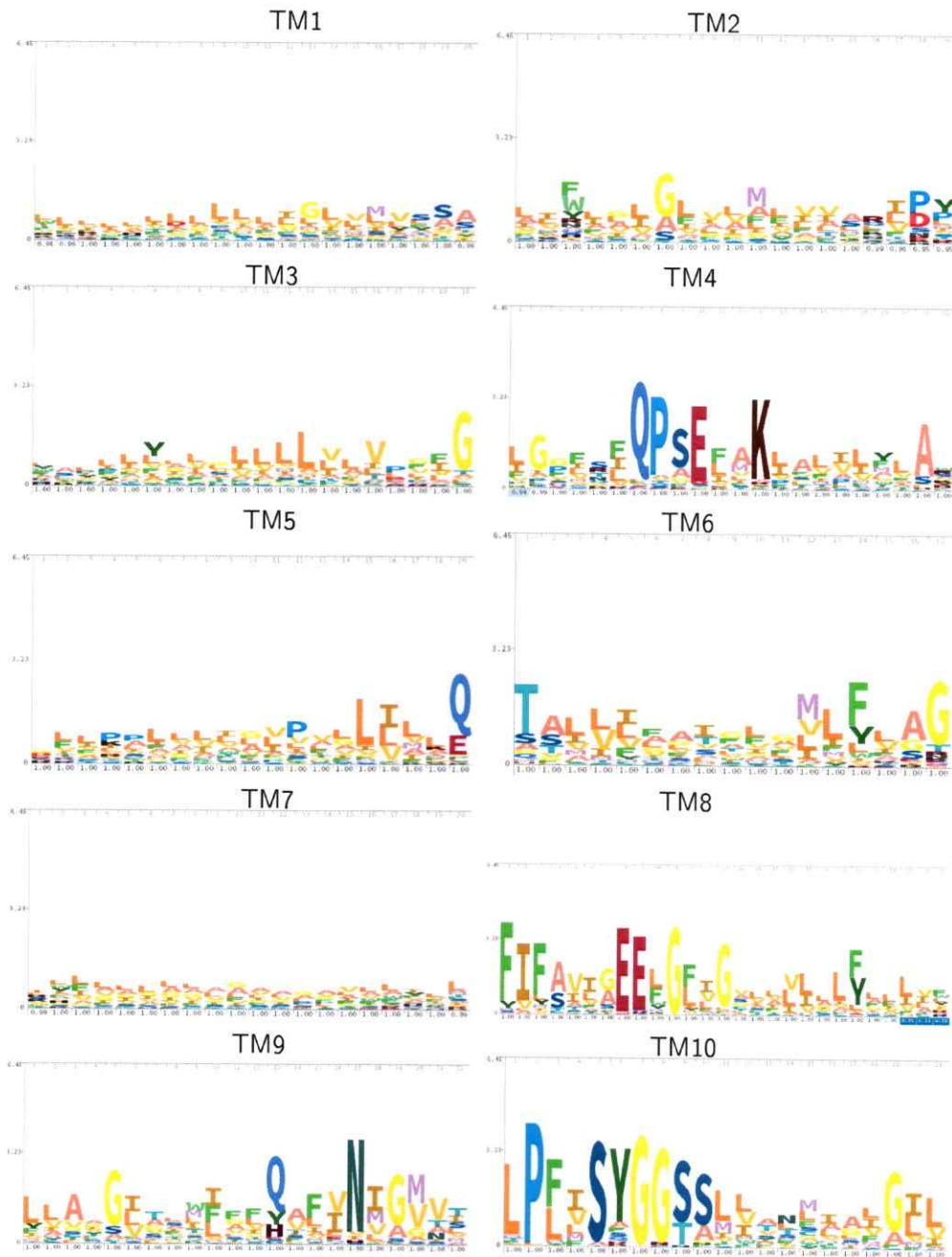


Figura 7: Logos de secuencia para las 10 zonas TM de FtsW

puede ver un análisis general de cada TM donde se promedia la polaridad de todos los residuos de ésta, junto con su desviación estándar.

TM	Polaridad	Empaquetamiento	Ocurrencia
TM1	0.127 ± 0.044	0.458 ± 0.051	80.4 %
TM2	0.140 ± 0.034	0.465 ± 0.053	94.5 %
TM3	0.119 ± 0.027	0.466 ± 0.048	93.2 %
TM4	0.208 ± 0.028	0.471 ± 0.036	96.1 %
TM5	0.147 ± 0.030	0.457 ± 0.049	91.8 %
TM6	0.107 ± 0.027	0.500 ± 0.049	85.6 %
TM7	0.135 ± 0.040	0.490 ± 0.059	86.2 %
TM8	0.091 ± 0.017	0.455 ± 0.037	99.7 %
TM9	0.130 ± 0.018	0.502 ± 0.034	93.0 %
TM10	0.124 ± 0.015	0.536 ± 0.030	71.9 %

Tabla 1: Valores promedio de índice de polaridad e índice de empaquetamiento para cada TM a partir de alineamiento múltiple.

Dado que al promediar la polaridad de cada residuo se podría estar perdiendo información estructural, se proyectó este índice sobre un mapa helicoidal que considera las dimensiones teóricas de una hélice  $\alpha$ . En esta proyección es posible tener una idea de si existen superficies con cierta polaridad conservada, en lugar de solo el promedio, independiente de la naturaleza de cada aminoácido. Este análisis se realizó con el programa *memprot* y se presenta en la Figura 8 (barras amarillas al centro) y vemos que algunas TM como 1, 3, 4 y 9 muestran claras tendencias de orientación en cuanto a su índice de polaridad. Esto podría indicar que esa superficie está orientada hacia los lípidos de la membrana.

Es interesante destacar que el programa desarrollado en esta Tesis permite hacer un análisis en función de la presencia o ausencia de cada zona TM (Tabla 2). Un análisis de este tipo permite por un lado saber cuáles regiones TM son eventualmente prescindibles para la función, y por otro, hacer correlaciones entre ellas en cuanto a su presencia, independiente de su secuencia peptídica. De la ocurrencia de cada TM se observa que en general todas están muy conservadas, siendo las menos las TM extremas. Esto tiene sentido si pensamos en procesos de corte y empalme de genes, que podrían generar proteínas truncas, principalmente en los extremos.



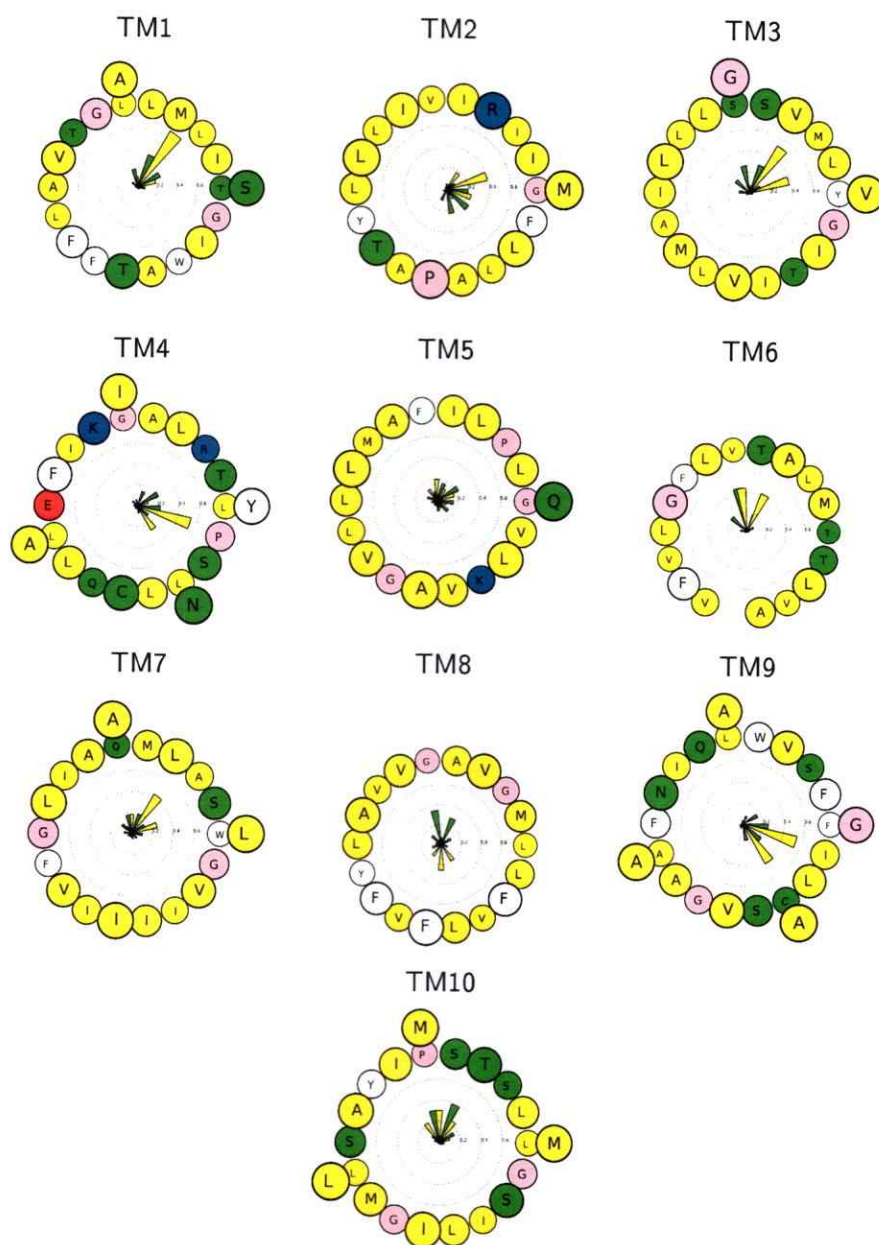


Figura 8: Proyección helicoidal de cada zona TM de FtsW indicando el índice de polaridad (verde) y el índice de empaquetamiento (amarillo). El ángulo cero corresponde al primer residuo de cada TM, donde los residuos subsiguientes se ubican a  $a+100$  grados del anterior, en función de las dimensiones de una hélice  $\alpha$  perfecta.

Por su parte, la correlación entre la ocurrencia de distintas TMs podría mostrar una posible interacción debido a su dependencia mutua. En la Tabla 2 vemos que TM3, TM5, TM9 y TM10 están correlacionadas, de modo que si una de estas falta, otra TM específica la reemplazará. Esto podría indicar una interacción directa entre estas TMs.

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9
TM2	0,10								
TM3	-0,09	0,17							
TM4	0,03	0,01	-0,01						
TM5	-0,09	0,13	0,84	-0,02					
TM6	-0,02	-0,01	0,48	-0,05	0,42				
TM7	-0,07	0,04	0,21	0,00	0,19	0,07			
TM8	0,01	-0,01	-0,01	0,06	-0,02	0,05	0,02		
TM9	-0,12	0,14	0,93	-0,04	0,81	0,48	0,19	-0,01	
TM10	0,06	0,08	0,36	0,01	0,36	0,21	0,13	0,00	0,39

Tabla 2: Coeficiente de correlación de Matthews para cada par de zonas TM. En blanco correlaciones no significativas (menores de 0.20), en amarillo levemente correlacionadas (0.20-0.29), en escala de rojos (de más claro a más intenso) correlación media (0.30-0.79), en rojo intenso correlación alta (0.79-1.00)

### *Análisis de loops*

A partir de la Figura 6 se identifica que los loops más conservados están en la región periplásmica, a excepción del loop 1-2, que está pobremente conservado. Esto es interesante dado que en general los loops cortos de proteínas de membrana están pobremente conservados [99]; sin embargo loop 3-4, loop 5-6 y loop 9-10 que tienen solo 11, 4, 9 residuos, respectivamente, están muy conservados, no solo en cuanto a identidad de residuos, sino que también en cuanto a su longitud (Figura 9 y Figura 15).

El loop 3-4 contiene un residuo de triptófano (Trp138) conservado en el 100 % de las secuencias que corresponde con el residuo Trp102 de RodA, que fue identificado como esencial para la función transglicosilasa [72]. Aparecen también uno o dos residuos de arginina conservados, en las posiciones inmediatamente anteriores a Trp138. Estos residuos podrían eventualmente tener que ver con la unión de los grupos fosfato del Lípido II.

El loop 5-6 tiene una conservación casi exacta de sus únicos 4 residuos. Esto podría indicar la necesidad de un posicionamiento preciso entre TM5 y TM6, así como un rol funcional para

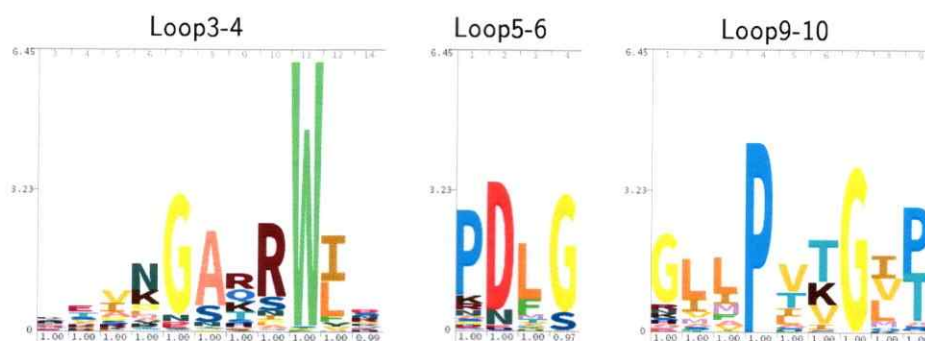


Figura 9: Logos de secuencia para los loops cortos más conservados en FtsW

alguno de estos residuos. Pro196 y Gly199 podrían ser importantes para la conformación del loop, y Asp197 podría tener algún rol funcional, como coordinar iones magnesio, que probablemente estén unidos a los grupos fosfato.

El loop 9-10 tiene una longitud también conservada y en cuanto a residuos, se pueden encontrar glicinas y prolinas altamente conservadas. Se presume que este loop es importante para posicionar TM9 y TM10, ya que no tiene residuos con grupos funcionales que pudieran ser importantes para la función o interacción específica con el sustrato.

Finalmente el loop 7-8, posee 58 residuos y también es muy conservado (Figura 10). Dada la longitud es probable que este loop tenga algún tipo de plegamiento, y por lo tanto se realizó una predicción de estructura secundaria. Esta muestra que hay tres probables hélices  $\alpha$ , una al principio, que eventualmente podría ser una continuación de TM7, una al centro y una al final, que podría ser una hélice continua con TM8. También se observa una posible estructura de hebra  $\beta$  en la región central. Cabe destacar la numerosa presencia de residuos de glicina conservados. Búsquedas por homología de esta región o de patrones de plegamiento (*threading*) no dieron resultados, por lo que se infiere un plegamiento único, que probablemente posee un importante grado de flexibilidad.

Justo antes de comenzar la tercera hélice predicha de este loop, se encuentra el residuo Asp297, que corresponde con el residuo Asp262 identificado en RodA como vital para la función



Figura 10: Logo de secuencia del loop 7-8. En la parte inferior se muestra la predicción de estructura secundaria, cilindros amarillos para hélices  $\alpha$  y flechas verdes para hebras  $\beta$ .

[72], nuevamente un residuo cargado que podría tener que ver con la interacción con el sustrato Lípido II. Así también existen otros residuos cargados altamente conservados.

### 6.1.2. Empaquetamiento

Existe una estrategia interesante que permite analizar proteínas de membrana a pesar de las dificultades que presenta su modelamiento completo. La identificación de aminoácidos relevantes en el empaquetamiento de hélices TMs [100] que permitan predecir el arreglo u ordenamiento de estas regiones en la proteína en su estado nativo. De esta manera, teniendo un mapa de la topología 2D de la proteína (Figura 11) es posible predecir cuáles residuos podrían estar dirigidos hacia las cadenas lipídicas que componen la membrana y cuáles interaccionan con otras hélices de la misma proteína, permitiendo y facilitando la predicción de la estructura terciaria de la proteína estudiada. Si dicho análisis es acompañado con estudios de conservación de aminoácidos se podrían encontrar posiciones relevantes en el plegamiento de FtsW.

Para realizar el estudio se deben otorgar valores de empaquetamiento a cada aminoácido. Estos se obtienen desde el análisis de estructuras cristalográficas de proteínas de membrana que se encuentran en la bases de datos PDB [100]. Este valor indica de manera indirecta el promedio de las distancias del aminoácido específico con el resto de la proteína, es decir, mientras menos distancia existe entre un determinado aminoácido y el resto de la proteína, mayor será su valor de empaquetamiento.

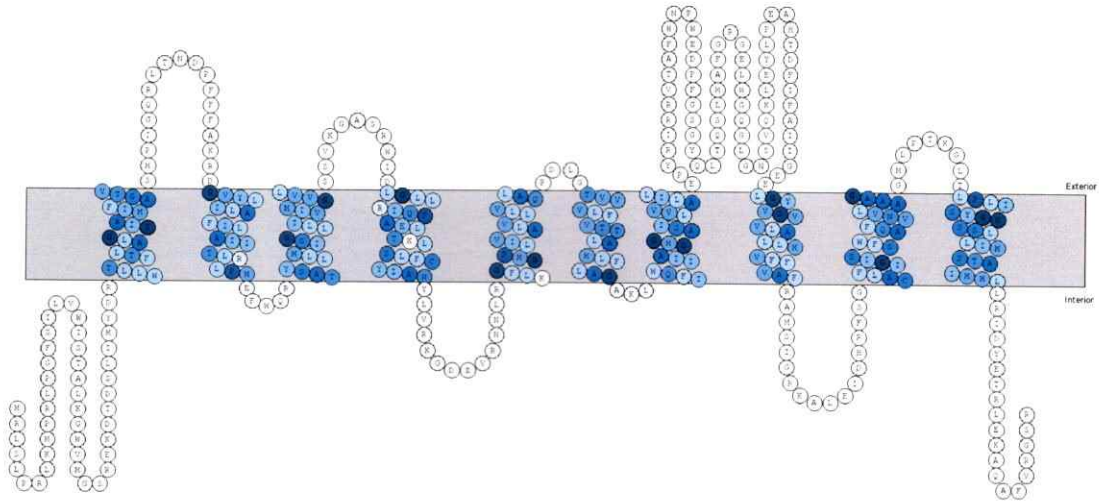


Figura 11: Topología 2D de FtsW (usando la definición de zonas TM de [49]), mostrando en escala de azul los valores de empaquetamiento para cada residuo de las zonas TM.

Al calcular las superficies de cada hélice TM modelada como una hélice perfecta, se pueden calcular valores de empaquetamiento promediadas o suavizadas para cada residuo. Esta información se puede complementar además con datos de conservación. Se sabe que los residuos en zonas TM que interactúan con otras zonas TM, son mucho más conservados que los residuos que enfrentan a los lípidos, y por ello complementar este análisis con datos de conservación robustece las observaciones.

Al igual que con el índice de polaridad, es posible proyectar el valor de empaquetamiento en una visión helicoidal de cada TM, asumiendo una geometría perfecta. De este modo es posible identificar caras de TMs preferidas a nivel de conservación. La Figura 8 presenta este análisis, con las barras verdes al interior de cada círculo.

### 6.1.3. Mutaciones correlacionadas

El estudio de mutaciones correlacionadas corresponde al análisis estadístico combinado entre cada par posible de posiciones dentro de un alineamiento múltiple de secuencias [101]. Este

tipo de análisis permite identificar posibles interacciones entre residuos, tanto intra-proteína, como inter-proteínas, pues el encontrar una mutación correlacionada, implica que al cambiar un aminoácido del par, el otro también cambia para mantener la interacción. Estas posibles interacciones pueden ser contactos directos entre esos residuos [102], importantes para la estructura o para la función, o ser residuos que se encuentren lejos físicamente, pero que por las características de la proteína mantienen una dependencia mutua. Estos análisis es posible hacerlos sólo con alineamientos múltiples con un alto número de secuencias, usualmente superior a 100.

Existen varias técnicas estadísticas para eliminar el ruido que generan pares de residuos que no están correlacionados. Además es posible incluir más información adicional de la secuencia en base a modelos estadísticos entrenados con datos empíricos. Estos métodos pretenden, no sólo encontrar pares de residuos correlacionados, sino que predecir específicamente contactos físicos, usando datos experimentales de estructuras proteicas. En este trabajo se usaron dos de los mejores métodos descritos para predecir contactos físicos, Gremlin [85] y PSICOV [86], así como Cmat [87], un método que se basa exclusivamente en las secuencias, usando un novedoso método de eliminación del ruido de fondo.

Utilizando el alineamiento múltiple descrito anteriormente para FtsW, se calcularon las mutaciones correlacionadas usando los 3 métodos descritos. Interesantemente la mayoría de los pares con alto puntaje coinciden para los 3 métodos, particularmente para Gremlin y PSICOV, lo que se refleja en el gráfico consenso mostrado en la Figura 12.

A partir de estos gráficos es posible identificar la región entre el loop 7-8 y la TM10 con gran cantidad de contactos con otras regiones y entre ellas. Las zonas TMs que muestran contactos en diagonal podrían corresponder a TMs que interactúan. Así es posible identificar un posible cluster de hélices TMs entre TM3, TM4, TM5 y TM6. A este clúster se podría sumar, aunque más débilmente TM7. Por otro lado TM8, TM9 y TM10 podrían formar otro clúster. TM1 y TM2 parecen interactuar entre ellas y con TM8-TM10. En cuanto a los loops, con excepción de loop 1-2, todos los loops peripásmicos están correlacionados entre ellos. Por el lado citoplasmático, el loop 2-3 y el loop 8-9 muestran correlación. Por su parte el gran loop 7-8

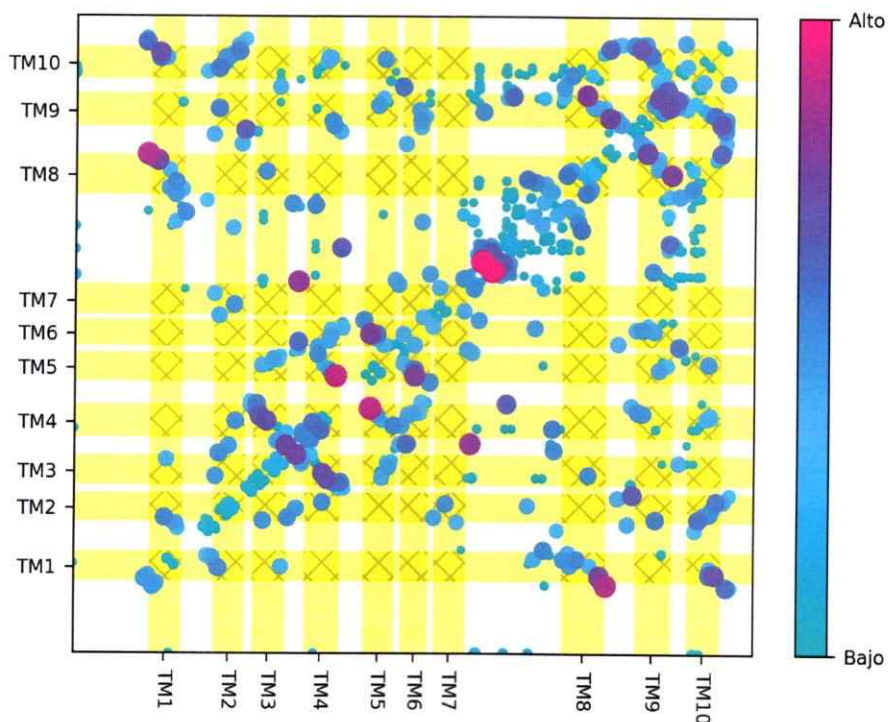


Figura 12: Mapa consenso de contactos predichos para FtsW. Las barras amarillas que cruzan el gráfico son las regiones TMs. El área de cada punto y el color son proporcionales al puntaje consenso dado para cada predicción o contacto. Se incluyen los resultados de los programas Gremlin [85], PSICOV [86], Cmat [87]. Los mapas tienen simetría diagonal.

muestra una mayor complejidad de interacciones producto de su probable plegamiento. TM9 y TM10 son las regiones TMs con mayor cantidad de posibles contactos con el resto de la proteína, por lo que es probable que tengan un rol central en el arreglo de hélices en la membrana.

Para cuantificar cuales son las regiones con mayor probabilidad de estar en contacto, se contabilizan el total de contactos entre dos regiones puntuales, considerando sólo 207 pares de contactos (la mitad del largo de la proteína). En la Tabla 3 se observan las regiones más interconectadas, lo que refuerza el análisis anterior.

Región	Región	Gremlin	PSICOV	Cmat
Loop7-8	Loop7-8	27	22	40
TM2	TM10	11	9	2
TM1	TM8	10	9	1
TM6	TM9	8	9	3
Loop7-8	TM8	7	5	7
TM9	TM10	6	9	5
TM4	TM6	6	6	1
TM1	Loop7-8	6	8	8
TM3	TM4	5	7	1
Loop8-9	TM9	5	5	4
TM9	Loop9-10	4	5	5
TM1	TM10	4	4	2

Tabla 3: Conteo de mutaciones correlacionadas y posibles contactos en función de pares de regiones en FtsW.

#### 6.1.4. Predicción estructural

##### *Ab initio*

A pesar de la complejidad y dificultad que implica modelar una proteína de membrana multi-paso, que no tiene ninguna homóloga con estructura conocida, como es el caso de FtsW/RodA hay métodos que permiten tener una aproximación. El modelamiento *ab initio* de proteínas de membrana ha concitado creciente interés, dada la dificultad de obtener cristales de proteínas de membrana en estado nativo y el consecuente escaso número de este tipo de estructuras disponibles en la base de datos PDB (menos del 2 % del total) [103]. La inclusión de parámetros evolutivos y descriptores de las hélices TM (como el comentado índice de empaquetamiento), permiten alimentar métodos de red neural que permiten predecir estructuras con confiabilidades aceptables, aunque siempre por debajo de sus símiles para proteínas solubles.

FILM3 [104] es un programa que utilizando la secuencia, datos evolutivos (como mutaciones correlacionadas) y descriptores de las hélices TM, ha conseguido mejorar la confiabilidad respecto a los otros programas existentes hasta el momento. Por lo tanto, se utilizó este programa y las mutaciones correlacionadas calculadas (Figura 12), para obtener un modelo aproximado de la



estructura de FtsW.

Se obtuvieron 400 modelos a partir del mismo set de mutaciones correlacionadas. Estos modelos se evaluaron con la herramienta proporcionada por FILM3, así como por el software maiden [105], programa que analiza modelos de proteínas de membrana usando potenciales estadísticos especialmente diseñados. Con esta clasificación, se escogió el mejor modelo y con el programa modeller se re-modelaron los loops, excepto el loop 7-8 por su longitud y aparente flexibilidad.

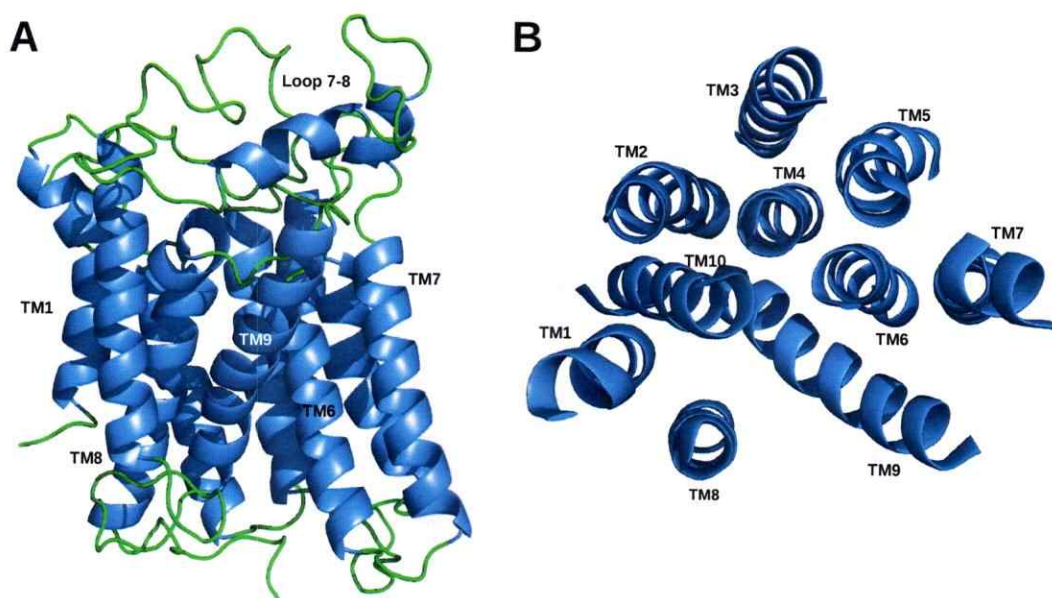


Figura 13: Modelo *ab initio* generado por el programa FILM3. **A.** Vista desde el plano de la membrana. **B** Vista desde el periplasma.

En la Figura 13 se muestra el mejor modelo obtenido para FtsW, usando las definiciones de zonas TM reportadas en [49]. Las TM3, TM4, TM5, TM6 y TM7 son modeladas como un manojo de hélices relativamente compacto. Las hélices TM1 y TM8 son modeladas juntas; TM9 y TM10 son modeladas como hélices inclinadas que pasan por dentro de la estructura, y conectan con TM2 ligeramente alejada de TM3. El loop 7-8 se modela desordenado por sobre la hélice TM9. Los loop 3-4, loop 5-6 y loop 9-10 aparecen relativamente cercanos, en la región

central de la estructura, confirmando la idea de que podrían participar de la función.

### *Modelamiento comparativo*

Un excelente método para detectar homologías remotas son los modelos ocultos de Markov (HMM). Estos son formas de representación estadística de los alineamientos múltiples que permiten detectar relaciones evolutivas, que escapan del alcance de métodos solo basados en secuencias. Usando este tipo de modelo se ha desarrollado un algoritmo de búsqueda de homología HMM-HMM llamado *HHpred* [106]. Buscando a partir de la secuencia de FtsW de *E. coli* se encuentra, aparte de otras proteínas de la familia SEDS, similitud con una ligasa de antígeno O. Este enzima, es una transglicosilasa que transfiere polímeros de antígeno O a un glicolípido acceptor [107]. El sustrato está, al igual que en el caso de FtsW, unido a un carrier de undecaprenil-fosfato, por lo que tiene sentido pensar que esta relación evolutiva es real.

A partir de este misma búsqueda, surgió el estudio que muestra que RodA de *E. coli* es una transglicosilasa [72], de una familia completamente distinta a las PBPs tipo A, que tienen un dominio transpeptidasa y uno transglicosilasa y, son inhibibles por monoemicina. Finalmente este estudio se complementa con uno de modelamiento tridimensional, que usó la estructura de una transglicosilasa de arquea como molde, basándose en la correspondencia de contactos experimentales de dicha proteína con los contactos predichos para FtsW [85].

Sin embargo, en el estudio donde se presenta el modelamiento tridimensional de FtsW también se modelan otras 10 proteínas, a través de un método semi-automático [85]. Por lo tanto, el modelamiento no contiene un análisis específico ni antes ni después del modelamiento. Por ello en esta Tesis se planteó como objetivo modelar FtsW utilizando como molde transglicosilasas de membrana, o sea, asumiendo dicha función para FtsW.

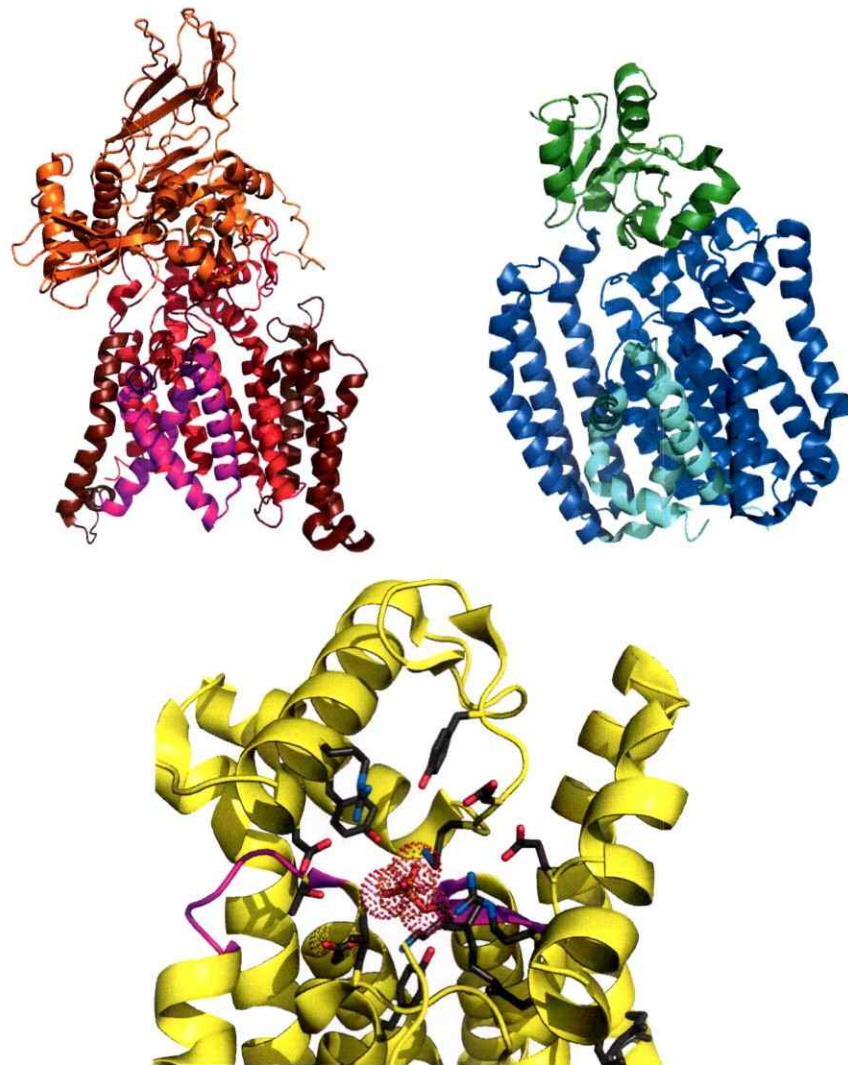


Figura 14: Visualización de las estructuras cristalográficas usadas como moldes para generar un modelo comparativo de FtsW. **A.** 3WAK. En rojo, clúster de TM1 a TM6; rojo oscuro, hélices TM entre los clústeres; violeta, clúster de TM10-TM12; naranja, dominio C-terminal **B.** 5F15. En azul claro, clúster de TM1 a TM6; azul opaco, hélices TM entre los clústeres; celeste, clúster de TM18-TM10; verde, dominio C-terminal. A y B están alineadas estructuralmente. **C.** Zoom del posible sitio activo en 5F15. Se destaca el grupo fosfato de una molécula de undecaprenil fosfato, y las cadenas laterales de los residuos que podrían participar en la unión. Los loops, loop3-4 y loop5-6 se destacan en violeta.

El trabajo de Kamisetty *et al.* [85] usó como molde la estructura de una oligosacaryltransferasa de arquea que transfiere oligosacáridos a un residuo de asparagina proteico, almacenada en la base de datos PDB (código 3WAK) [71]. Las oligosiltransferasas que tienen proteínas como aceptores del oligosacárido, son abundantes en eucariontes y arqueas y han sido clasificadas en la familia STT3 [108, 109]. En bacterias solo han sido identificadas en *Campylobacter jejuni* [110]. No obstante, en bacterias existen oligosiltransferasas que tienen aceptores distintos de proteínas, como lípidos. Este es el caso de la proteína ArnT, que transfiere una molécula de 4-amino-4-deoxi-L-arabinosa a Lípido A, para modificar la capa de lipopolisacárido [111]. Dado que la proteína ArnT de *Cupriavidus metallidurans* posee estructura tridimensional resuelta, se incluyó como posible molde (depositada en la base de datos PDB con el código 5F15).

Ambas proteínas tienen un plegamiento similar, con dos arreglos de hélices que interactúan entre ellos (Figura 14). Un primer clúster está conformado por 6 hélices TM ordenadas antiparalelamente, con un gran loop entre el TM1 y TM2. A su vez TM3 se caracteriza por ser una hélice desordenada, con un quiebre en la región central. Este clúster, en particular el loop 1-2 forma interacciones con un dominio C-terminal alfa-beta que ambas proteínas poseen, de unos 500 residuos. En este mismo clúster TM5 y TM6 forman un par antiparalelo cercano, con un loop muy corto que las conecta. El otro clúster de TMs es de 3 hélices, con una hélice quebrada, otra que cruza por el interior de la proteína y una última hélice larga e inclinada. Las últimas dos hélices interactúan con el otro clúster. Ambos clústeres están conectados por hélices TM de cantidad variable. Estas hélices variables son, una en 5F15 (TM7) y 3 en 3WAK (TM7, TM8 y TM9). La última de estas hélices desemboca en el periplasma, con una hélice paralela a la membrana que forma un loop desestructurado que conecta con el segundo clúster. Hacia el C-terminal, después del segundo cluster hay un número variable de TMs que no interactúan fuertemente con el resto de la proteína, 1 en 3WAK y 3 en 5F15.

Un primer paso del proceso de modelamiento comparativo es tener un alineamiento secuencia-estructura de la proteína a modelar, en este caso FtsW y los moldes. Sin embargo, las proteínas

seleccionadas 3WAK y 5F15, tienen 13 zonas TM cada una, dificultando la asignación de las regiones similares a FtsW. Además del gran dominio C-terminal que FtsW no posee, que sirve para reconocer el sustrato proteico en 3WAK o Lípido A en 5F15. Otra diferencia importante es la presencia en estas proteínas de un gran loop entre las TM1 y TM2, que contiene varios residuos, que entre otros, forman lo que se presume es el sitio activo (Figura 14C) por la unión de un átomo de manganeso, que podría estar en reemplazo de un átomo de magnesio, en 3WAK [71] y por la posición de una molécula de undecaprenil fosfato, co-cristalizada en 5F15 [111].

En FtsW y las otras proteínas de la familia SEDS, el loop 1-2 es mucho más corto y está pobremente conservado. Esto tiene sentido, considerando que este loop en las proteínas molde interactúa con el dominio C-terminal, ausente en FtsW. Sin embargo, deja abierto el problema del sitio activo.

Otros residuos que forman el sitio activo en las proteínas molde yacen sobre los loops periplasmáticos loop 3-4 y loop 5-6. Estos loops, junto con residuos en el mencionado loop 1-2 y el loop 9-10, forman un sitio a la altura de las cabezas polares de los fosfolípidos, con abundantes residuos cargados y aromáticos donde podría alojarse el grupo pirofosfato de Lípido II, tal como lo muestra la posición del undecaprenil fosfato en 5F15 [111].

Los loops 3-4, loop 5-6 y loop 9-10 están altamente conservados en FtsW y otras proteínas de la familia SEDS. La diferencia crítica está en el loop 7-8, que en las proteínas molde (loop 9-10 en 3WAK) es corto, no participa directamente en la unión del sustrato y es altamente flexible, dada su falta de estructura e incluso, no es posible resolver la posición de varios de sus residuos.

Como se analizó en la sección anterior, el loop 7-8 está altamente conservado en las proteínas SEDS y contiene algunos residuos candidatos a participar de la unión de sustrato. Entonces, aparecen dos posibles hipótesis: la primera es que el loop 7-8 de FtsW se corresponda con el loop 1-2 de las otras proteínas, dada la presencia de residuos importantes para la función, el largo y la capacidad de formar una estructura relativamente compleja. La otra hipótesis es que dada la pérdida del dominio C-terminal en FtsW, también se haya perdido el loop 1-2, y su rol

de unión del sustrato haya sido reemplazado por el loop 7-8, desestructurado en 3WAK (donde corresponde al loop 9-10), pero que en 5F15 también contiene algunos residuos que podrían estar implicados en la unión de los grupos fosfatos (Figura 14C). Esta segunda hipótesis implicaría que el loop 1-2 de FtsW viene de un evento evolutivo distinto y por lo tanto no participaría de la función.

Para resolver esta disyuntiva, se requiere analizar los otros loops que intervienen en la unión de sustrato y en la formación del posible sitio activo. Específicamente los loop 3-4 y loop 5-6 que están altamente conservados tanto en los moldes como en FtsW. Si se analiza la secuencia, no hay tanta similitud, sin embargo se conserva la longitud de esos loops (Figura 15), algo que podría estar relacionado con su función [112]. Además, en FtsW existen numerosos pares de residuos correlacionados entre los loop 3,4, loop 5-6 y loop 9-10, que podría indicar una homología entre estas regiones. Adicionalmente la TM3, tanto en 3WAK como en 5F15, muestra un quiebre que es clave para posicionar el loop 3-4, generado por un residuo de prolina altamente conservado, pero que en ocasiones puede ser reemplazado por glicina. En el TM3 de FtsW se

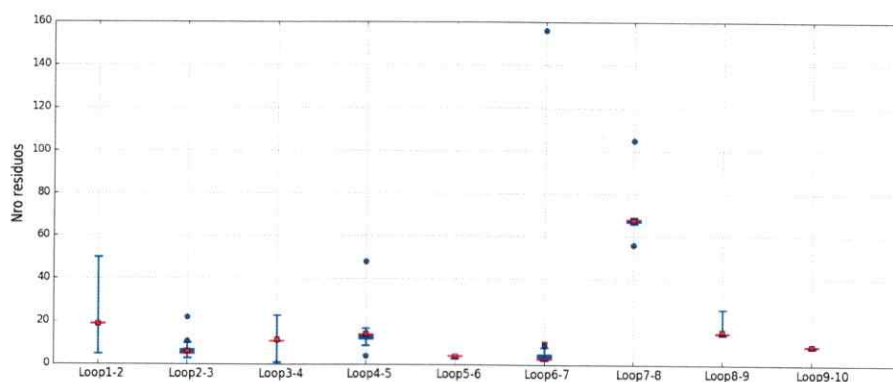


Figura 15: Distribución del largo de los loops de FtsW calculados a partir del alineamiento múltiple de trabajo. Los cuadrados rojos corresponden al promedio, las líneas rojas a la mediana. La caja azul corresponde a los datos entre el 25 y 75 % de la distribución. Las líneas azules son los valores que están entre el 5 y el 95 % de la distribución. Los puntos azules los datos fuera de este último rango.

encuentra un residuo de glicina también altamente conservado, además de un residuo de prolina dos posiciones antes, pero menos conservado. Por lo tanto tiene sentido pensar que los TMs, desde TM3 hasta TM6 son homólogos a los correspondientes TMs en 3WAK y 5F15, y forman parte del sitio de unión al sustrato, donde la gran cantidad de residuos cargados conservados podría ser clave en la interacción con el grupo pirofosfato de Lípido II, sea directamente o través de los iones magnesio que lo acompañan.

En tanto, el segundo clúster de TMs observado en las estructuras molde, también podría tener un correlato en FtsW. El loop 9-10 corresponde en longitud al loop de los moldes (Figura 15), pudiendo completar parte importante de la estructura. Sin embargo, un dato relevante lo aportan los mapas de contactos inter-residuos de los moldes (Figura 16), en comparación con los datos de predicción de contactos a partir de mutaciones correlacionadas en FtsW (Figura 12).

Al comparar los mapas de contacto con los mapas de predicciones de contactos de FtsW, vemos que en las regiones del TM3 al TM6 hay una similitud importante, especialmente en las interacciones entre hélices antiparalelas (puntos que forman una línea de pendiente negativa,

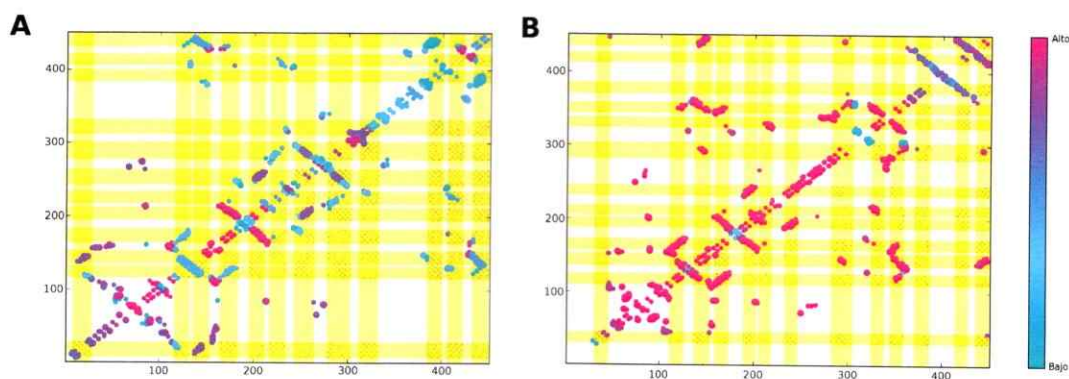


Figura 16: Mapas de contactos para 3WAK **A** y 5F15 **B**, generados al calcular pares de residuos cuyos carbonos  $\beta$  están a menos de  $6\text{\AA}$  en la estructura cristalográfica. El tamaño de los puntos es proporcional a la distancia entre los átomos, y el color corresponde al nivel de conservación. Cuadros amarillos son las zonas TMs.

como TM3 con TM4 o TM4 con TM5), o interacciones de hélices paralelas, como TM4 con TM6. En tanto en las regiones TM8-TM10, también se observan similitudes a niveles de contacto. No obstante, un punto de discordancia entre estos mapas es evidente en las TM1 a TM2, que en los moldes aparece estableciendo contactos con el primer clúster de hélices, pero en FtsW aparece estableciendo contacto con las TM8 y TM10, pertenecientes al segundo clúster. De hecho a nivel de contactos, TM1 de FtsW parece tener mayor similitud con TMs posteriores al segundo clúster, como TM12 o TM3, ausentes en FtsW.

Dados estos análisis, se plantea que en FtsW las TM1 y TM2 y el loop que las conecta, no corresponde con las TM1-TM2 de los moldes, se encuentra estructuralmente ligado al segundo clúster y no al primero. Esto se podría explicar porque al perderse el dominio alfa-beta C-terminal en las proteínas tipo SEDS, se perdió la presión evolutiva para mantener el gran loop 1-2 y por lo tanto, la función de las TM1 y TM2 que lo formaban. En cambio las actuales TM1 y TM2 de FtsW podrían venir de un proceso evolutivo independiente y ello explicaría su perfil de contactos diferente y su baja conservación, tanto en términos de secuencia (Figura 7) como en términos de su presencia (Tabla 1).

Finalmente se elaboró un modelo comparativo, alineando las zonas TMs como se muestra en la Figura 17, específicamente desde TM3 hasta TM10, tomando como patrón a las estructuras depositadas como 3WAK y 5F15, dejando fuera las hélices intermedias TM7 y TM8 de 3WAK. El loop 7-8 de FtsW se modeló basándose en las regiones iniciales del loop correspondiente en los moldes, que corresponden con hélices quebradas, que continúan luego de las respectivas

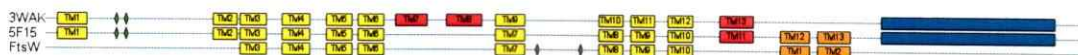


Figura 17: Alineamiento esquemático entre 3WAK, 5F15 y FtsW. En amarillo, zonas TM similares entre moldes y FtsW; en naranja, zonas TM de FtsW alineadas de manera discontinuada con TMs de 5F15; en rojo, zonas TM de moldes sin regiones TM similares en FtsW. En azul dominio de unión a sustrato, ausente en FtsW.



que según la predicción de estructura secundaria corresponden (Figura 10). El resto del loop 7-8 se modeló usando un servidor de *threading*. Por su parte los TM1 y TM2 se modelaron basadas en los TM12 y TM13 de 5F15 que si bien no comparten una homología detectable, corresponden en cierta medida en sus contactos con el resto de la estructura. El modelo final se seleccionó en base a potenciales estadísticos específicamente diseñados para proteínas de membrana [105] y se muestra en la Figura 18.

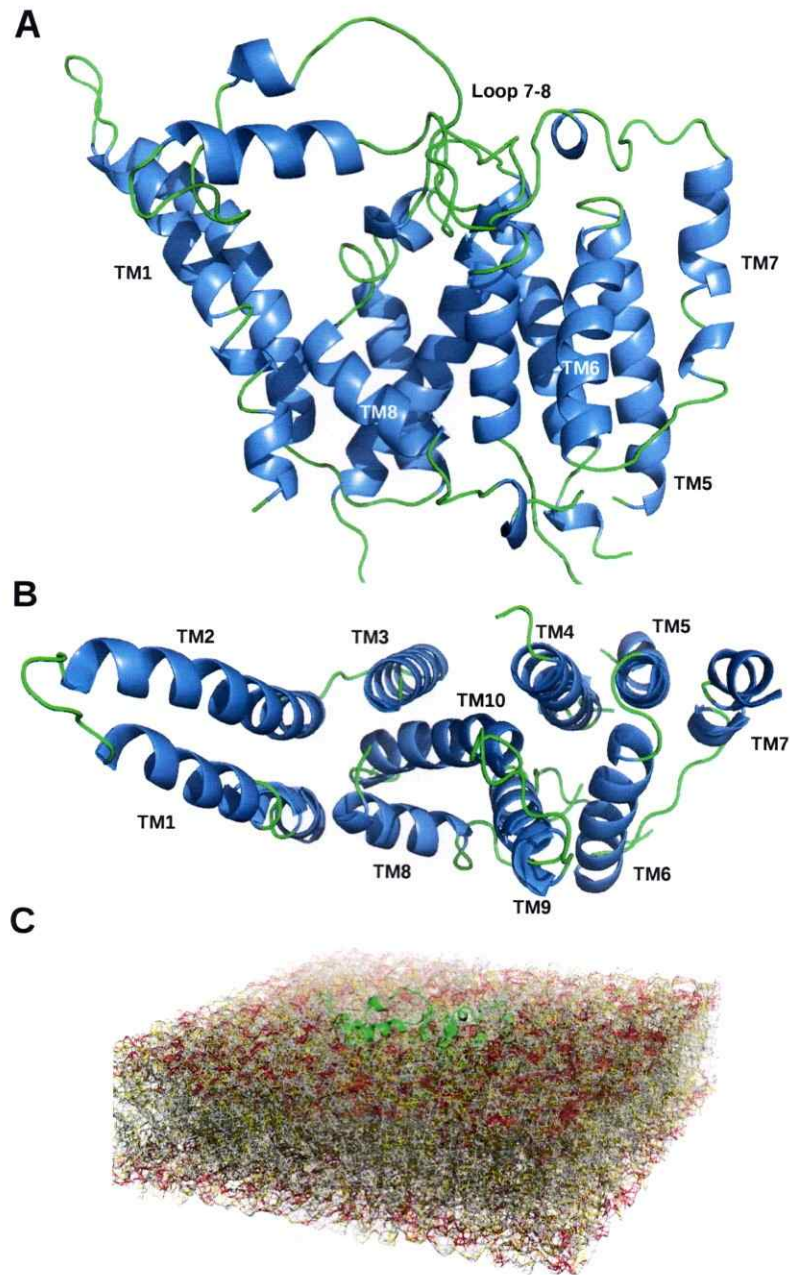


Figura 18: Modelo final obtenido por comparación con los moldes cristalográficos 3WAK y 5F15. En A se observa la estructura sola, en representación de *cartoons* mirada desde el plano de la membrana. Se indican algunas regiones. En B se observa la misma estructura desde el lado externo de la membrana. En C se observa la estructura (en verde) embebida en una bicapa de POPC.

### 6.1.5. Estabilidad estructural

Para estimar la estabilidad del modelo generado, este se embebió en un bicapa lipídica (Figura 18B) y se sometió a dinámica molecular. Luego de una etapa de minimización energética y otra de equilibrado (para ajustar presión y volumen) se calculó el movimiento promedio de la estructura.

En la Figura 19A se observa tanto el radio giro como la desviación cuadrática media de cada paso de la dinámica en función del estado inicial. Estos análisis permiten tener una idea general del movimiento o cambio de la estructura a medida que avanza la simulación. Al observar en RMSD se observa que en la primera mitad de la simulación, la proteína aun está ajustando su posición, pero en la segunda mitad ya parece mantenerse en torno a un valor estable de 4 Å. En tanto el radio de giro se mantiene estable en torno a los 2,4 nm, lo que indica que la estructura general de la proteína es estable y no tiende a desplegarse. Esto último no es trivial, considerando la muy baja similitud con los moldes empleados, y la inexactitud del posicionamiento de las interfases entre hélices TM. Ciertamente, mayores tiempos de simulación son ideales para este tipo de estructuras embebidas en una membrana lipídica, sin embargo con estos tiempos aun es posible tener una idea general de la estabilidad.

A partir de la segunda mitad de la dinámica, cuando se alcanza un meta-equilibrio, se puede realizar el mismo análisis de desviación cuadrática media, pero no promediando el movimiento de todos los átomos, sino promediando en función de cada residuo a lo largo del tiempo de simulación. Este análisis o fluctuación de la raíz cuadrática media (RMSF) (Figura 19B) permite analizar la movilidad en función de las regiones. Es posible observar una movilidad diferenciada entre las regiones TMs y los loops. Como era esperado, los loops más largos, tienen una mayor movilidad, como loop 1-2 y loop 7-8. A pesar de la movilidad general menor de las zonas TMs (Figura 19A, línea verde), hay algunas de estas hélices que presentan mayor movilidad como TM6, TM7 y TM10. Esto podría deberse a la mala orientación de algunos residuos respecto a la membrana o a la ambigua asignación de los límites de cada zona TM. Estos límites fueron

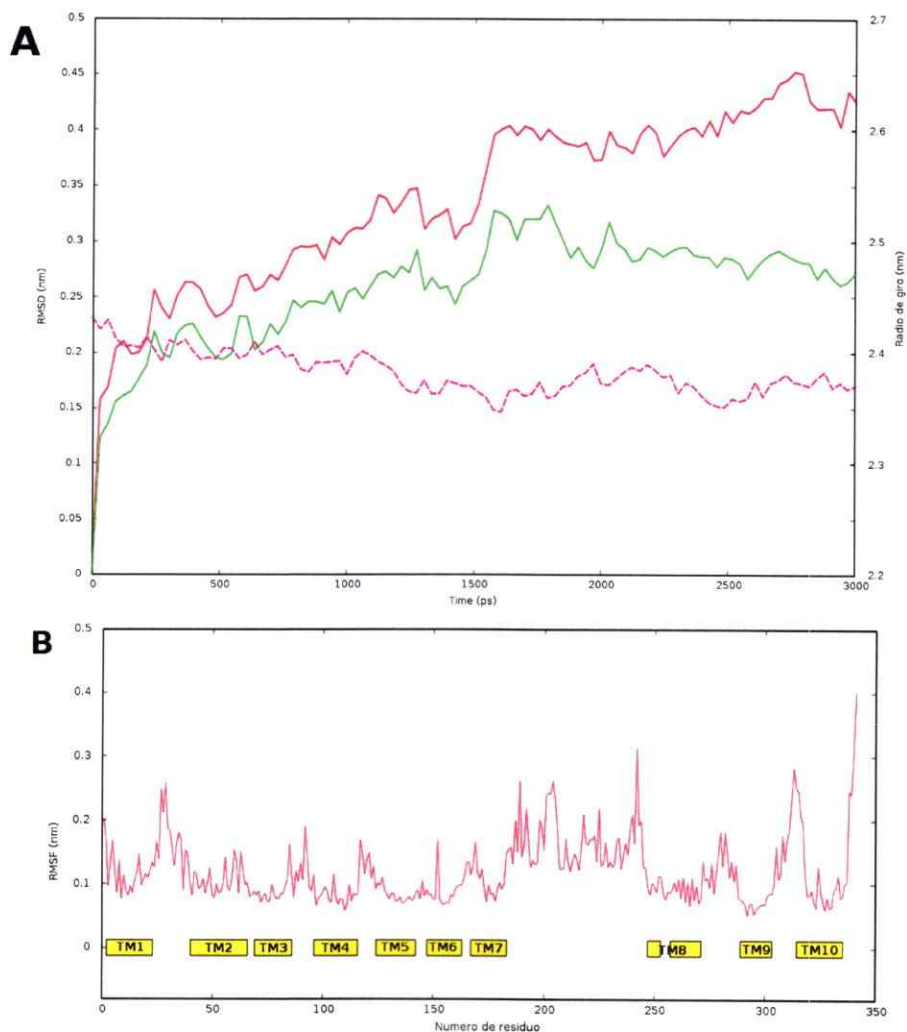


Figura 19: Estabilidad del modelo de FtsW embebido en una bicapa de POPC. **A.** En el eje Y de la izquierda se grafica el RMSD para FtsW completa (línea roja) y solo las zonas TM de FtsW (línea verde). En el eje Y de la izquierda se muestra el radio de giro de FtsW completa (línea morada entrecortada). **B.** RMSF de FtsW. Abajo del gráfico se indican las regiones TMs.

considerados como los estimados en Lara *et al.* [49], los que se obtuvieron usando trunco fusionados a enzimas reporteras (HRP). Sin embargo, este análisis podría ser un tanto ambiguo en la definición exacta de los límites de cada región, particularmente en hélices TM de estructura no-canónicas como lo muestran algunas hélices en los moldes, con quiebres internos o hélices paralelas a la membrana.

## 6.2. Ensayos de accesibilidad a cisteína

Con el objetivo de dar algún tipo de soporte experimental al modelo de FtsW presentado, se recurrió a la técnica llamada ensayo de accesibilidad a cisteína [113]. Este ensayo consiste en expresar mutantes de la proteína a analizar, con un solo residuo de cisteína y tratar las bacterias vivas que expresen dichos mutantes en baja concentración, con distintos reactivos capaces de reaccionar con el grupo tiol de la cisteína en función de su accesibilidad al solvente. De este modo, este ensayo se ha utilizado para analizar distintas proteínas de membrana [68, 114, 115]. Uno de los reactivos útiles para este ensayo es *N-etil-maleimida* (NEM) (Figura 20), que se une covalentemente a cisteínas que estén en sitios expuestos (independiente de su orientación respecto a la membrana). Una vez tratadas las bacterias con este reactivo (o respectivos controles) las bacterias se lisan y se desnaturan con urea 8 M para exponer todas las cisteínas independiente de su posición en la proteína, o sea independiente de si han reaccionado con NEM o no. Se trata la muestra con *metil-polietilenglicol-maleimida* (Mal-PEG) (Figura 20), que se une a todas las cisteínas que no sido bloqueadas por NEM (o sea por su posición en la estructura). La masa molar del último reactivo es 2 kDa, por lo que al unirse a la proteína aumenta su movilidad electroforética y es posible ver la diferencia en un Western blot (WB) al comparar con un control no tratado. Así, usando un conjunto de mutantes cisteína única y los resultados de accesibilidad obtenidos mediante este ensayo, es posible tener una idea de la topología 3D de la proteína FtsW y verificar si es compatible con el modelo tridimensional planteado en las secciones anteriores.

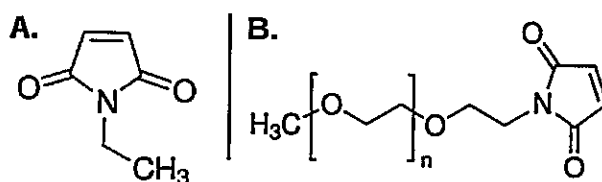


Figura 20: Reactivos a usar para el ensayo de accesibilidad a cisteínas. A. NEM. B. Mal-PEG.

Para el caso de FtsW, donde ya se conoce la topología de la proteína [49], este ensayo puede servir para identificar regiones intra-membrana accesibles a solvente. Esto podría significar la

presencia de un poro o canal a través de la membrana, lo que sería coherente con el rol de lipasa de Lípido II asignado a esta proteína [19]. Esta forma de indentificar un poro o canal, ya ha sido utilizada para la proteína MurJ, que también es indicada como traslocadora de Lípido II [68].

### 6.2.1. Síntesis de mutantes

Para el análisis del método SCAM es necesario generar mutantes de FtsW cisteína-única en distintas posiciones de su estructura. Los resultados SCAM pueden ser interpretados en términos de la accesibilidad de la región donde está la cisteína y por lo tanto la elección de la ubicación de ésta(s) es clave. En particular en el caso de las zonas TM, que en teoría deberían ser completamente inaccesibles, es posible interpretar un aumento de la accesibilidad como la posible presencia de un poro.

A partir de un análisis de conservación e hidrofobicidad se propone el diseño y construcción de mutantes cisteína-única para FtsW en las regiones transmembrana TM1, TM4, TM8, TM9 y TM10. Estas regiones son las más conservadas 7 y muestran indicios de tener algún rol funcional.

La posición específica para incorporar estas cisteínas, se propone considerando la mayor centralidad posible respecto a la posición en la bicapa, pero en un residuo que esté pobremente conservado, para minimizar la posibilidad de que dicho mutante interfiera con la función o genere una proteína inviable. Las mutaciones propuestas se muestran en la Tabla 4. Cabe destacar que estas mutaciones se deben hacer sobre un gen de *ftsW* libre de cisteínas. En FtsW existen dos residuos de cisteína que están pobremente conservados, por lo que se mutaron a serinas 6, para tener el *background* cisteína-libre y sobre éste sintetizar las mutantes descritas a continuación.

### 6.2.2. Curvas de crecimiento

Al analizar las curvas de crecimiento para todas las cepas mutantes, incluidos los respectivos controles, se ve que la expresión de las mutantes y supresión de la expresión del gen WT

TM	Mutación
TM1	A56C
TM4	T152C
TM8	A315C
TM9	S353C
TM10	S382C

Tabla 4: Mutaciones diseñadas para ensayos de SCAM

cromosomal (creciendo en glucosa/IPTG) no afecta significativamente el crecimiento de las cepas (Figura 21). Estos resultados indican que ninguna de las mutantes es tóxica para las bacterias y complementan bien a la proteína WT.

### 6.2.3. Resultados SCAM

Los experimentos de SCAM se analizan en función de la exposición al solvente de la cisteína analizada. Si una cisteína está expuesta al solvente podrá ser bloqueada por NEM, y esto a su vez impedirá que el reactivo MalPEG reaccione y aumente la masa aparente en el WB. Por lo tanto, una banda que se asemeje al control negativo (donde no se agrega ningún reactivo), corresponderá a una cisteína que está expuesta al solvente. Al contrario si una banda se asemeja al control positivo (MalPEG agregado a la proteína desnaturada), indica que ese residuo no está expuesto al solvente y por lo tanto no pudo ser bloqueado por NEM. A continuación los resultados obtenidos (Figura 22), así como el análisis por separado de cada mutante.

**$\Delta$ Cys.** Este doble mutante C174S/C361S no posee ningún residuo de cisteína, por lo tanto es insensible a la aplicación de los reactivos usados en SCAM. Se usa sólo como control, para identificar la marca correspondiente a la proteína FtsW, sin modificar (flecha negra)

**A56C.** Este mutante está ubicado en la TM1. Al tratar con NEM se observa el mismo patrón que con el control positivo, concluyéndose que este residuo no está expuesto al solvente. En el control negativo se observa una doble banda.

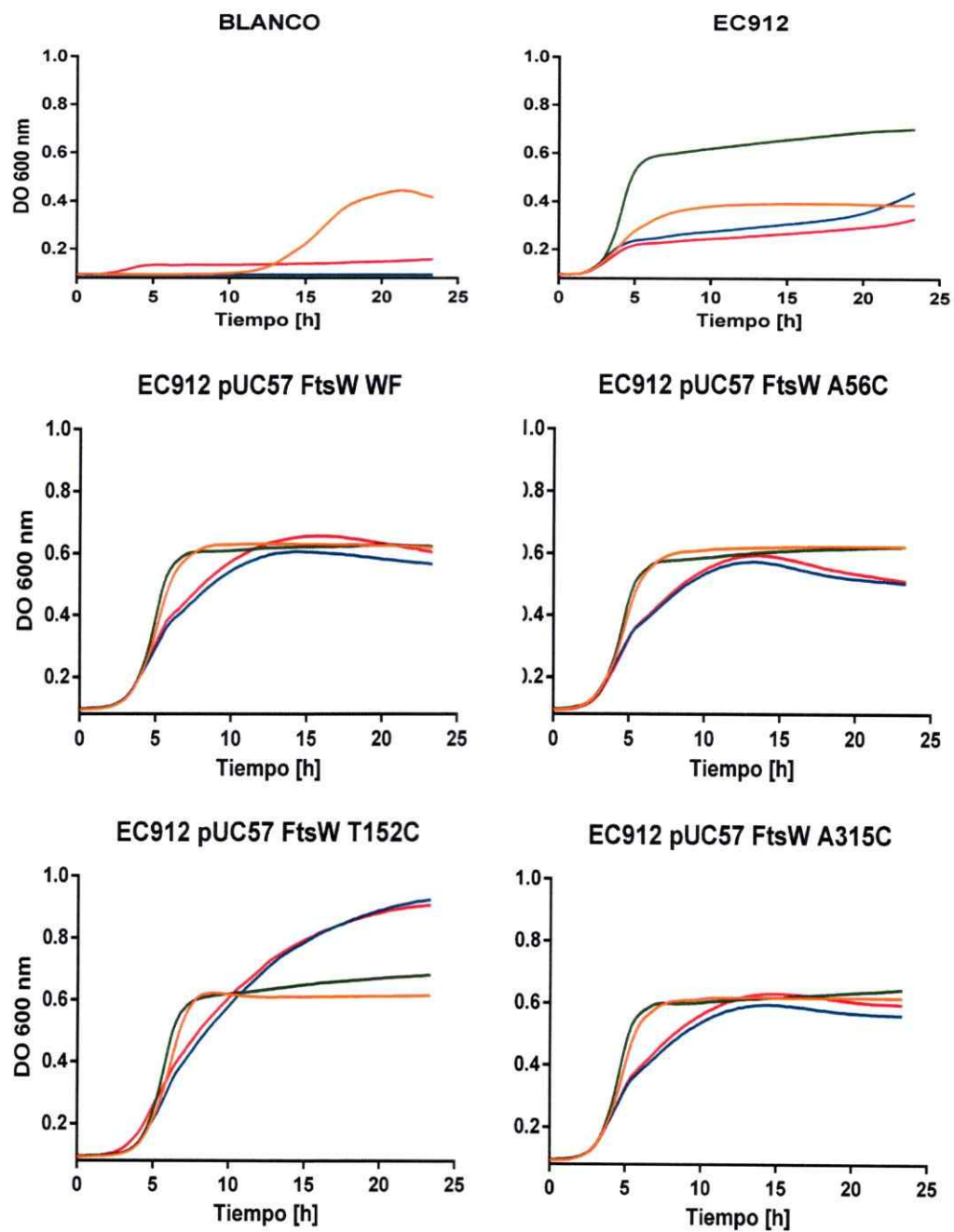


Figura 21: Curvas de crecimiento de cepas mutantes cisteína-única de FtsW. Leyenda en próxima página.



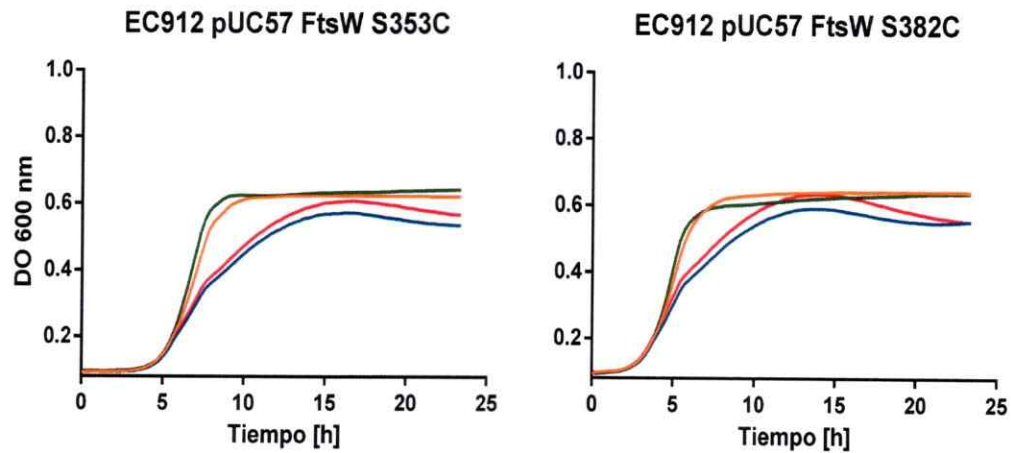


Figura 21: Para todos los gráficos, los colores corresponden a los medios de crecimiento: rojo, solo LB; verde, LB con arabinosa 0,02 %; azul, LB con 0,1 mM IPTG; naranja, LB con 0,1 mM IPTG y 0,02 % de glucosa. La cepa EC912 es el gen WT de *ftsW* en un plasmidio en *trans*. La cepa FtsW WF, es el gen *ftsW* libre de cisteínas (C158S/C346S). Todas las otras mutantes se hacen sobre este último background.

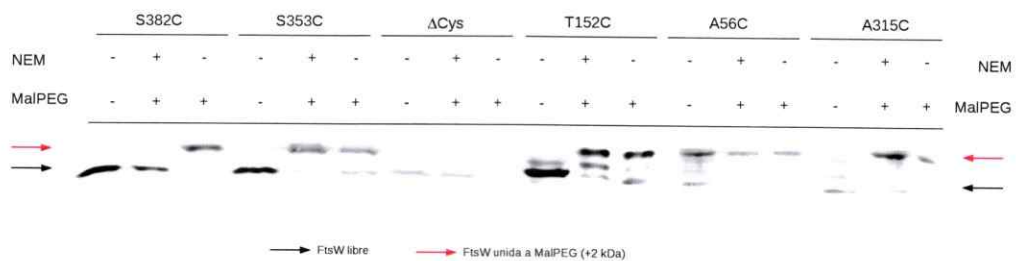


Figura 22: .

Resultados de WB con análisis de SCAM. Flecha negra indica marca de la proteína sin marcar con MalPEG. Flecha roja indica movilidad electroforética alterada por la unión de MalPEG.

**T152C.** Esta mutación está ubicada en el TM4. Al tratar con NEM se observa un patrón muy similar al control positivo, concluyéndose que este residuo no está expuesto al solvente. En todos los carriles se observa una banda de peso aparente intermedio, que no debería influir en los resultados, dado que se ve en todas las muestras. Aunque esto también podría indicar una exposición intermedia.

**A315C.** Esta mutación se encuentra en la TM8. Al tratar con NEM se observa un patrón muy similar al control positivo, concluyéndose que este residuo no está expuesto al solvente.

**A353C.** Esta mutación se encuentra en la TM9, una de las más conservadas y más polares. Al tratar con NEM se observa un patrón similar al control positivo, concluyéndose que este residuo no está expuesto al solvente.

**A382C.** Esta mutación se encuentra en la TM10, junto con TM9 una de las más conservadas y más polares. Al tratar con NEM se observa un patrón similar al control negativo, concluyéndose que este residuo sí está expuesto al solvente. Es la única mutación que revela esta característica.

Interesantemente, sólo uno de los mutantes probados aparece como expuesto al solvente. Para la función propuesta para FtsW como flipasa, estos resultados no son los esperados ya que la habilidad de traslocar un glicolípido desde una cara de la membrana a la otra, supone la presencia de un poco o región polar al interior de la estructura, que permitiera el paso de la región polar de Lípido II a través de la bicapa lipídica. La presencia de residuos cargados y conservados en TM4 (Figura 7) apoyaba esta hipótesis, así como la predicción de un posible poro en la TM9. Sin embargo, ninguna de estas hélices parece expuesta al solvente, solo TM10, que también está conservada.

En cambio la hipótesis de que FtsW sería una transglicosilasas, con homología remota con transglicosilasas eucariontes y de arqueas, parece hacer más sentido con estos resultados. De hecho, el modelo estructural presentado, a partir de esta homología remota (Figura 18) es bastante consistente con estos resultados, con la única excepción de la exposición del TM10. Esta última hélice, está en el modelo desarrollado en esta Tesis, oculto entre otras TMs, por lo que no concuerda un nivel alto de exposición.

### 6.3. Análisis de FtsI

#### 6.3.1. Conservación

A partir de los alineamientos múltiples obtenidos para FtsI, se estima el nivel de conservación de cada residuo. Este nivel se puede graficar en forma de color sobre la estructura tridimensional completa (que se describe en la próxima sección). En la Figura 24B es posible apreciar este coloreado. En general podemos ver un mayor nivel de conservación en el dominio PBP particularmente en el sitio activo y el bolsillo donde se aloja el sustrato. En cambio en la zona TM y en dominio no-catalítico (NC) hay solo algunos residuos puntuales conservados. En la zona TM destacan las argininas al principio y final (Arg23 y Arg41). En el dominio NC hay residuos conservados en algunas hélices, unas en la interfase con el dominio PBP y otras más alejadas.

La zona TM de FtsI (Figura 23) que no aparece en ninguna estructura cristalográfica, en general está poco conservada, salvo por un par de residuos de arginina que flanquean esta región. En un estudio con fusiones a GFP encuentran varios residuos en el TM, cuyas mutaciones impiden la correcta localización septal de FtsI, entre los que se encuentran Arg23, Phe24, Arg41 y Trp44, entre otros [53]. Sin embargo, ellos discuten sobre la poca conservación de esta región, y sugieren que la participación del dominio no catalítico podría ser mayor en otras especies de bacterias, o que habrían mutaciones correlacionadas con FtsW que enmascaran la conservación.

#### 6.3.2. Modelamiento estructural

FtsI posee solo una hélice TM y el resto de sus residuos se encuentran en el periplasma. A pesar de que no hay estructuras cristalográficas de esta proteína completa en *E. coli*, sí hay para la región soluble [116] y para varias homólogas como FtsI de *Pseudomonas aeruginosa* [117], de *Staphylococcus aureus* [118] o PBP2 de *Neisseria gonorrhoeae* [119]. En estos análisis de cristalografía se observan los dos dominios de las PBPs tipo B, uno catalítico o PBP, donde se une al PG [116] y el dominio N-terminal NC, de aun dudosa función. Se piensa que podría tener

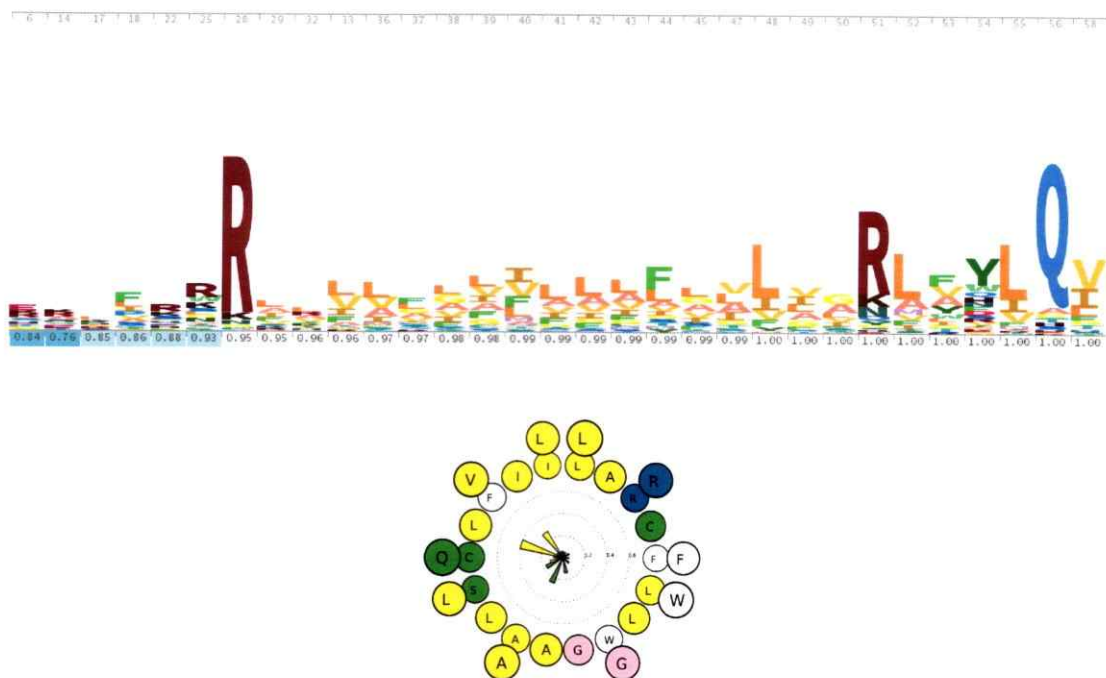


Figura 23: Caracterización de la región TM de FtsI. **A.** Logo de secuencia que muestra el nivel de conservación. **B.** Proyección helicoidal, que incluye en su interior histograma del vector de empaquetamiento (verde) y vector de hidrofobicidad (amarillo).

que ver con la interacción con otras proteínas del divisoma como FtsW o FtsN, ya que posee numerosos *loops* no definidos (Figura 24) que eventualmente se estabilizan al estar en contacto con otras proteínas [116]. Este dominio no se une al saco de PG y quedan dudas de si es un dominio de dimerización o no, pues los experimentos no son concluyentes [116]. Esto a pesar de que este dominio también se conoce como dominio de dimerización de PBPs.

En tanto la zona TM no es posible modelarla por comparación (pues no está incluida en ninguna de las estructuras cristalográficas citadas), pero no es un problema dada su simplicidad y los pocos residuos que quedan hacia el lado citoplasmático. Gracias a esto, en este trabajo se modeló la zona TM de FtsI, como una hélice recta, desde el residuo F18 hasta el residuo R60 en la interfase con el dominio NC. Esta región que denominamos proximal, entre la membrana

y el dominio catalítico, también fue modelada como una hélice  $\alpha$  continua, debido a que la predicción de estructura secundaria así lo indica.

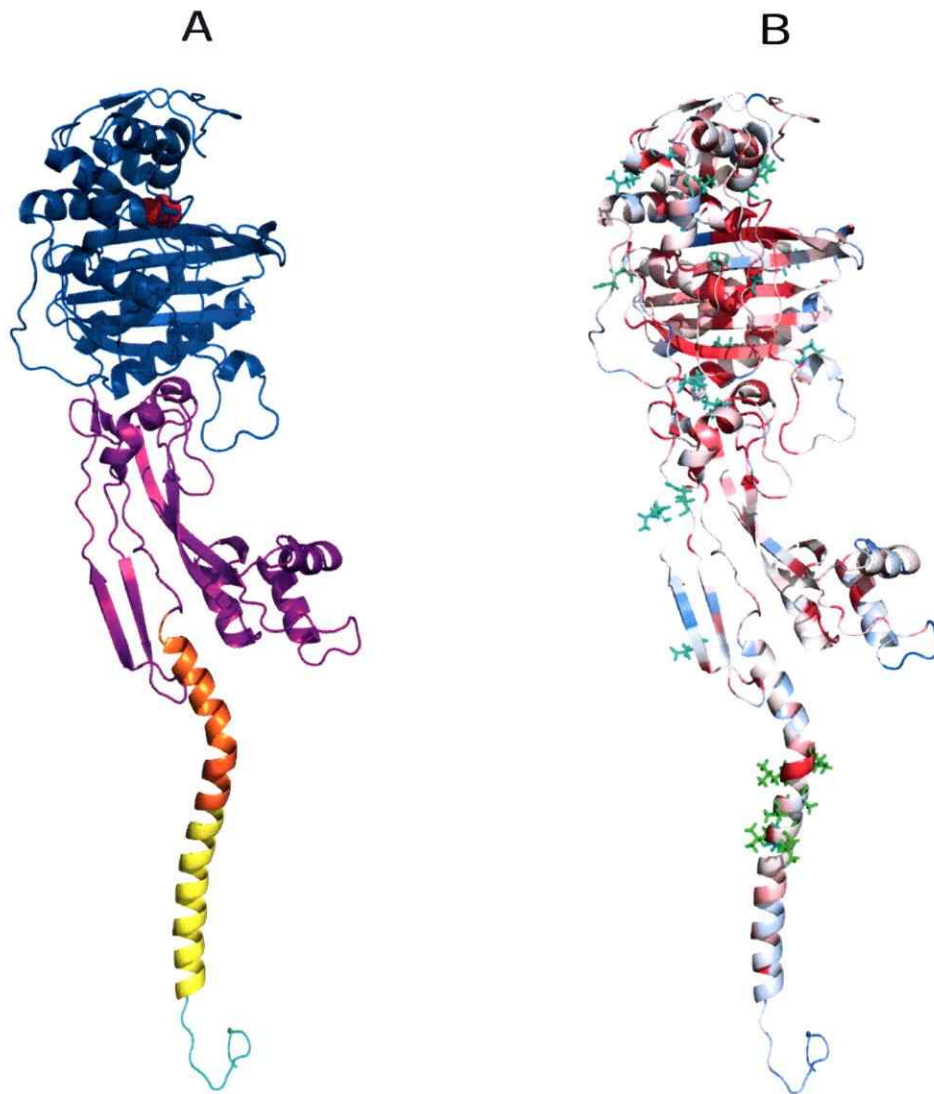


Figura 24: Estructura de FtsI de *E. coli*, modelada a partir de los moldes descritos en el texto. La región citoplasmática y TM fueron modeladas forzando una hélice  $\alpha$  continua desde el residuo 18. **A.** Estructura coloreada según sus dominios, en celeste región citoplasmática, en amarillo zona TM, en naranja zona proximal, en púrpura dominio no-catalítico, y en azul dominio PBP. En rojo se indentifica el residuo catalítico Ser307. **B.** La misma estructura anterior coloreada según su nivel de conservación. Mientras más roja la cinta mayor conservación, mientras más azul menos conservada. Las cadenas laterales representadas explícitamente corresponden a residuos identificados como correlacionados con residuos en FtsW. Mientras más verdes estos residuos más alta la probabilidad de que esté en contacto con FtsW.

## 7. Objetivo Específico 2

### 7.1. Mutaciones correlacionadas

Tal como se hace para una sola proteína, también es posible obtener mutaciones correlacionadas entre dos proteínas diferentes, a partir de un alineamiento múltiple de ambas secuencias concatenadas, que correspondan en cuanto al par que forman (que sean de la misma especie). Usando el programa Gremlin, que posee una variante específica para pares de proteínas, que se basa en mutaciones correlacionadas, pero además incluye información empírica de complejos proteicos. En la Tabla 5 se muestran los resultados para estas predicciones.

Región en FtsW	Región en FtsI	Número de contactos
TM4	PBP	14
TM9	PBP	14
Loop7-8	PBP	10
TM8	TM	6
TM9	TM	6
TM10	NC	6
TM8	Prox	5
TM4	NC	5
TM7	NC	5
TM9	Prox	5
Loop9-10	TM	3
Loop9-10	PBP	3

Tabla 5: Número de contactos físicos por regiones, predichos para FtsW-FtsI. Para FtsI la simbología de las regiones es: TM: Región TM; Prox: región periplásmica entre la zona TM y el dominio no-funcional; NC: dominio no-funcional; PBP: dominio transpeptidasa.

De este análisis (Tabla 5) se pueden identificar algunos puntos de probable interacción entre ambas proteínas. Llama la atención que la región de FtsI con mayor número de posibles contactos con FtsW sea su dominio PBP, que está demasiado lejos como para formar una interacción física con FtsW, particularmente con sus zonas TM. Una hipótesis para explicar esto, es que cambios en el tipo de sustrato que distintas bacterias usan, podría generar cambios en los sitios de unión en FtsW y a su vez en el sitio de unión de sustrato en FtsI, que está cerca del sitio activo, en

FtsW		FtsI		Probabilidad de contacto
Residuo	Región	Residuo	Región	
Val313	TM8	Leu39	TM	0.982
Phe271	Loop7-8 (H2)	Leu45	Prox	0.842
Tyr309	TM8	Val47	Prox	0.842
Val360	TM9	Leu38	TM	0.802
Val312	TM8	Val42	Prox	0.729
Tyr309	TM9	Ala43	Prox	0.717
Leu276	Loop7-8 (L2)	Val47	Prox	0.470
Met366	Loop9-10	Leu38	TM	0.450
Leu276	Loop7-8 (L2)	Val244	NC	0.346
Leu91	TM2	Gln203	NC	0.321
Leu316	TM8	Leu39	TM	0.318
Thr369	Loop9-10	Glu219	NC	0.317
Phe300	Loop7-8 (H3)	Val42	Prox	0.304

Tabla 6: Contactos físicos predichos para el par FtsW-FtsI, sin considerar el dominio PBP de FtsI. Se indica probabilidad de acierto entregada por el mismo programa. La simbología de regiones es la misma de la Tabla 5. Para loop 7-8 de FtsW, se agregó el elemento de estructura secundaria donde se encuentra el residuo.

el dominio PBP. Esta hipótesis es plausible dada la gran cantidad de bacterias consideradas en los alineamientos múltiples y el hecho que las zonas de FtsW predichas como relacionadas son bastante conservadas y conectan con loops identificados en la unión del sustrato.

Aparte del dominio PBP de FtsI, otras regiones establecen posibles contactos con FtsW, aunque en menor cantidad. Una de interés es la única zona TM de FtsI, que se espera que sea vital en la interacción con FtsW. Sin embargo, establece pocos contactos, con TM9 y el loop9-10, por lo menos a nivel de predicción. Este resultado es contradictorio, porque por un lado se corresponde con el hecho de esta zona en FtsI no está demasiado conservada (Figura 23A), pero por otro lado existe evidencia experimental que muestra, que al menos en *E. coli*, la zona TM sí contiene residuos importantes para la unión [53].

Dados los resultados de mutaciones correlacionadas, una de las hipótesis presentadas por Weiss *et al.* [53] para explicar la poca conservación queda sin base, que era la presencia de correlación con residuos en FtsW [53]. Además, los dos residuos de arginina identificados como



importantes en FtsI, y que sí están conservados (Figura 23A), deberían estar interactuando con residuos cargados negativamente en FtsW. Esos posibles puntos de contacto en FtsW, además deberían estar igualmente conservados y expuestos a la región externa de la proteína. Sin embargo, hay solo un residuo que cumple con estos requisitos, que es Glu305, al principio del TM8. Este residuo podría estar interactuando con Arg41 en el lado periplásmico de FtsI, aunque es discutible dado que estarían a la altura de los grupos fosfato de los fosfolípidos que eventualmente apantallarían una interacción de cargas como esa.

Del análisis individual de los posibles contactos (Tabla 6) se identifican otras regiones como TM8, TM9, loop 7-8 y loop 9-10. La mayoría de los residuos identificados se encuentran en la parte periplasmática, específicamente en FtsI entre su TM y el dominio NC, en la denominada región proximal, que se modeló como una hélice  $\alpha$  continua. Por lo tanto esta región parece clave para interactuar con FtsW, ya sea por sus TM8 y TM9, que están cercanas, y en menor medida el loop 7-8. Particularmente las regiones del loop 7-8 que estarían haciendo contacto con FtsI, es su hélice central (H2), y la hélice final (H3) que conecta con el TM8.

## 7.2. Modelamiento estructural

Dados los antecedentes recabados aparecen dos regiones en cada proteína que podrían estar involucradas en la formación del complejo FtsW/FtsI. En ambos casos las interacciones estarían en el lado periplásmico de la membrana. En FtsW, tanto TM8 como TM9, y en menor medida el loop 9-10 que conecta TM9 con TM10, muestran evidencias de que establecen contactos con FtsI. TM8 es una hélice particular, que aparece quebrada al centro, muy inclinada y con una hélice anterior, casi paralela a la bipaca. Esta hélice anterior denominada H3, dentro del loop 7-8, forma un ángulo de  $30^\circ$  con TM8 y posee gran cantidad de residuos conservados. Luego de TM8 viene un gran loop citoplasmático, pobremente conservado, que da origen a TM9, una hélice recta que cruza por el centro del manojo de hélices, pero deja una cara expuesta a la membrana. Esta región en FtsW estaría en contacto con la región periplásmica de FtsI que está entre la mitad de su única zona TM y el dominio NC. Toda esta parte de FtsI probablemente adopta

una configuración helicoidal y contiene residuos conservados y correlacionados con residuos en FtsW, particularmente en el rango 38-47.

Por otro lado, existe otro par de regiones en ambas proteínas que podrían estar involucradas en su interacción, pero con evidencia un poco más débil. El gran loop 7-8 de FtsW, altamente conservado y de estructura probablemente flexible, contiene residuos que eventualmente contactan a FtsI en su dominio NC. El dominio NC de las PBP tipo B, contiene algunos residuos conservados, pero no hay consenso sobre su función. Una de las funciones propuestas es la formación de interacciones con otras proteínas, lo que soporta la idea de que estaría contactando al loop 7-8 de FtsW. El loop 1-2 de FtsW, a pesar de estar menos conservado que los otros loops periplásmicos, también podría establecer contactos con este dominio de FtsI.

Al insertar en la membrana ambas proteínas juntas, hacen sentido las interacciones propuestas, ya que están a una distancia de interacción posible, particularmente la región proximal de FtsI y los TM8 y TM9 de FtsW. Por su parte, el dominio NC de FtsI parece muy lejos de la membrana como para interactuar con FtsW, sin embargo la flexibilidad del loop 7-8 hace pensar de que esta interacción sí podría generarse. De hecho, la proteína FtsI está anclada en un solo punto a la membrana, como un tallo, que eventualmente podría doblarse y acercarse al dominio NC a los loops periplásmicos de FtsW (Figura 25).

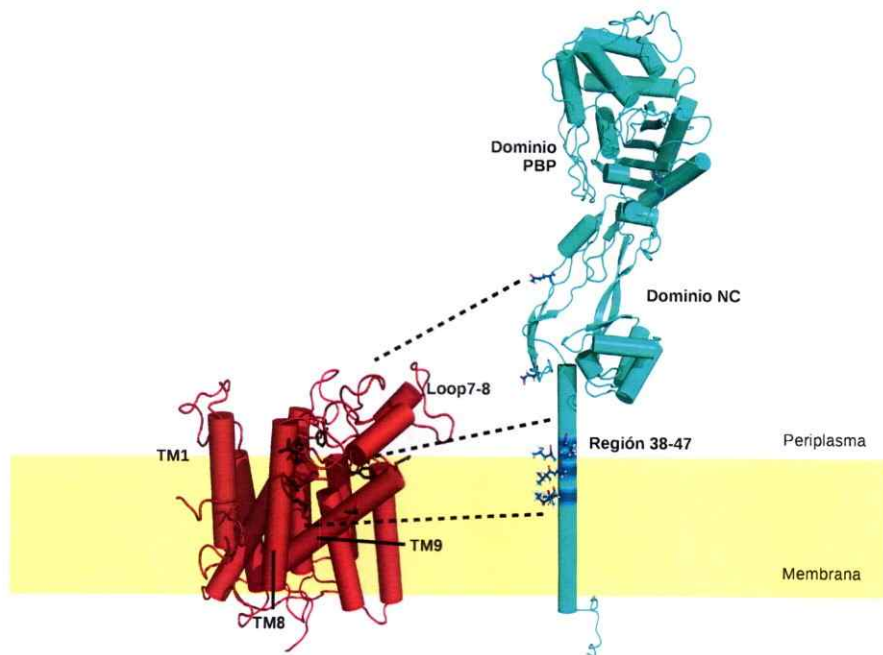


Figura 25: Representación esquemática de la propuesta de interacción entre FtsW (rojo y verde) y FtsI (celeste). En FtsI se omite el dominio PBP. La membrana lipídica se representa como un rectángulo amarillo, donde la parte inferior es el citoplasma y la superior el periplasma. Las líneas punteadas indican las posibles regiones de interacción. La interacción entre la región 38-47 de FtsI con los TM8 y TM9 de FtsW, sería la principal región de interacción. La interacción entre el loop1-2 y loop7-8 de FtsW con el dominio NC de FtsI también es una posible región de interacción, pero con menor cantidad de evidencias. Se indican algunos residuos importantes.

### 7.3. Objetivo Específico 3

#### 7.3.1. Presencia del complejo FtsW/FtsI

La interacción de entre FtsW y FtsI es indispensable para la división bacteriana en *E. coli*. Si bien este proceso es relativamente bien conservado, en la diversidad bacteriana podrían haber variantes a nivel de presencia o ausencia de genes y proteínas específicas. Para estudiar la conservación del complejo FtsW/FtsI se analizaron clústeres de genes ortólogos (COG) en la base de datos *Microbial Genome Database for Comparative Analysis* (MBGD). Esta base de datos identifica ortólogos mediante similitud de secuencias nucleotídicas en genomas bacterianos completamente secuenciados. Al analizar 868 genomas (que abarcan prácticamente todos los filos conocidos de bacterias) se identificaron 3 COGs correspondientes a FtsW y 3 para FtsI. Dada la similitud a nivel de secuencias y el hecho que también forma un complejo binario con una PBP tipo B, se analizaron también los COGs correspondientes a las proteínas RodA (1 clúster) y MrdA (2 clústeres) [48], homólogos de FtsW y FtsI, respectivamente. Sólo un bajo número de genomas no contiene ningún miembro de la familia SEDS (2,8%).

En este análisis se encontraron 1.507 genes de *ftsW/rodA* y 3.153 genes de *ftsI/mrdA*. Estos genes tienen una distribución bastante diversa, aunque la mayor proporción corresponde a pares *ftsW/ftsI* dentro de la región *dcw*, identificada en *E. coli* como un gran operón que contiene a la mayoría de los genes de la división celular y síntesis de PG [120]. Así también hay genes *ftsW/rodA* con más de un gen *ftsI/mrdA* en su cercanía. Para el caso de *rodA* y *mrdA*, ambos genes están contiguos en *E. coli* en una unidad transcripcional única, fuera del cluster *dcw* [121]. En la Figura 26 se muestra la distribución de *ftsW/rodA* en función del número total de estos genes en cada especie, así como de la cantidad de *ftsI/mrdA* en su vecindad.

Sin embargo existen muchos genes *ftsW/rodA* que no tienen ningún gen en su vecindad, lo que incluye hasta 10.000 bp río arriba y río abajo de cada gen. Esto hace difícil clasificarlos como genes *ftsW* o *rodA* dada su alta similitud de secuencia. Además, el método para asignar

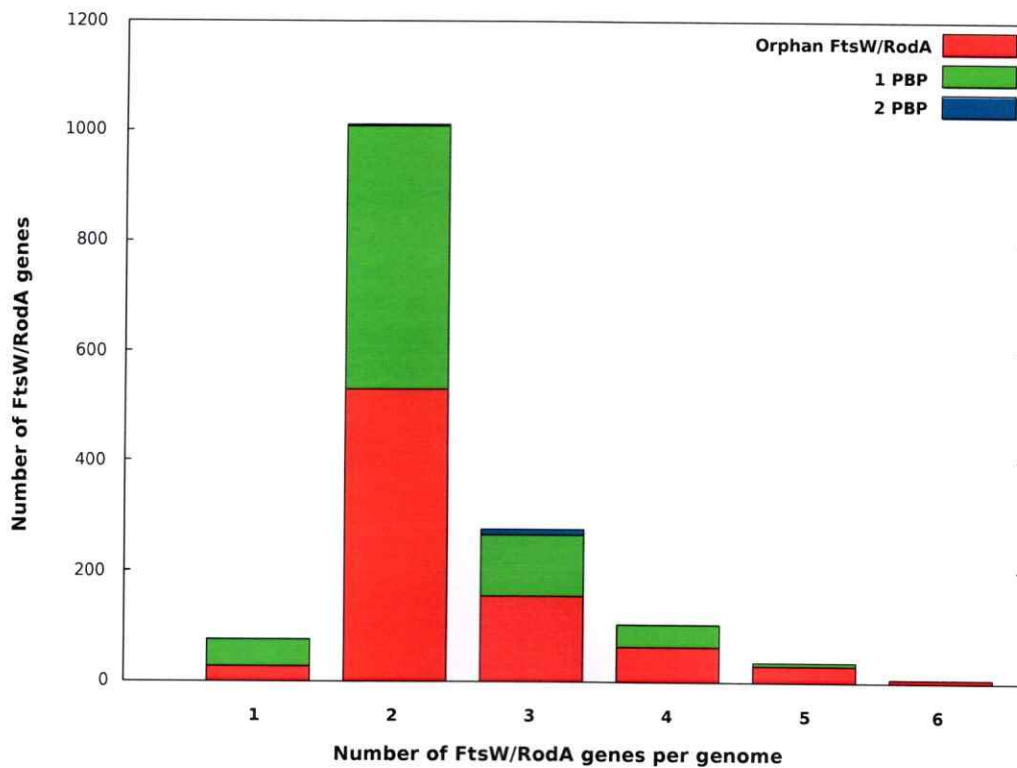


Figura 26: Distribución de la cantidad de *ftsW* por genoma, y de la cantidad de *ftsI* acompañantes.

los nombres de los genes en las bases de datos es con métodos automáticos que simplemente copian los nombres en función de la homología con otros genes. Así, este método puede llevar a ambigüedades, y eso explica que muchos de estos genes supuestamente *ftsW* o *rodA* en los COGs podrían no corresponder, dado el análisis hecho en esta Tesis.

### 7.3.2. Clasificación de genes *ftsW/rodA*

Dada la dificultad de clasificar *a priori* cada gen como *ftsW* o *rodA*, la ambigüedad de los nombres asignados automáticamente y su diversa distribución en los genomas estudiados, se realizó una novedosa clasificación de los genes *ftsW/rodA* en cada especie. Esta clasificación es

compleja, pero necesaria dado que los genes *ftsW/rodA* pueden estar ambos en cada especie, en más de una copia. Además en el caso de Gram positivos, como *B. subtilis*, se ha identificado el complejo SpoVD/SpoVE, homólogo a FtsW/FtsI (y por lo tanto a RodA/MrdA). Este complejo participaría en el proceso de esporulación de estas bacterias. Por lo tanto en este caso, las bacterias podrían tener genes que codificarían para 3 proteínas de la familia SEDS. Lo interesante de *B. subtilis* y otras bacterias Gram positivo, es que los genes *spoVE* y *spoVD* están ubicados en el clúster *dcw*, exactamente en la posición de *ftsW* y *ftsI* de otras bacterias, dejando dudas sobre si es la misma u otra proteína. El denominado *ftsW* de *B. subtilis* en cambio, es un gen que no se encuentra en una unidad transcripcional con otro gen, está al parecer aislado de otros genes relacionados con división, esporulación o síntesis de PG. Es lo que en este trabajo hemos denominado FtsW/RodA huérfana.

Para hacer un análisis evolutivo de la proteína FtsW es vital poder clasificar cada gen encontrado como *ftsW*, *rodA* o *spoVE*. Como se mencionó anteriormente, los tres genes son lo suficientemente similares como para que los programas automáticos los confundan. De hecho, todos poseen el dominio FtsW/RodA/SpoVE, encontrado en servidores de dominios como InterPro o Pfam, de donde también se suele tomar la información para asignar nombres a genes desconocidos.

A partir de los 1.507 genes *ftsW/rodA*, se identificaron sus contextos genómicos para clasificarlos de mejor manera. El análisis de los contextos genómicos permitió identificar 7 tipos de genes distintos definidos según la Tabla 7.

Sólo en el caso de FtsW I, FtsW II y RodA es posible encontrar con seguridad el par de genes necesario para formar el complejo FtsW/FtsI (o su homólogo correspondiente). Sin embargo, no es posible descartar que una proteína FtsW forme un complejo con una PBP de una región distinta, como es el ya mencionado caso de FtsW de *B. subtilis*. De hecho, un cuarto de todos los genes *ftsW/rodA* son huérfanos, lo que no significa que no formen un complejo con una PBP, sino que no es posible saber con certeza si lo hace o no, o con cual gen *ftsI/mrdA* se asocia.

Tipo	Contexto genómico	Observación	Proporción
<i>ftsW</i> I	En clúster <i>dcw</i> , incluyendo por lo menos un gen <i>ftsI</i> y al menos 3 genes del proceso de síntesis de PG	FtsW clásica	20,5 %
<i>ftsW</i> II	En clúster <i>dcw</i> no clásico, incluyendo por lo menos un gen <i>ftsI</i> y entre 1 y 3 genes del proceso de síntesis de PG		0,9 %
<i>ftsW</i> III	En clúster <i>dcw</i> clásico, pero sin gen <i>ftsI</i>		17,0 %
<i>ftsW</i> IV	En clúster <i>dcw</i> no clásico, sin gen <i>ftsI</i>		4,6 %
<i>ftsW</i> O	<i>ftsW/rodA</i> huérfana		26,7 %
<i>rodA</i> I	Contigua a gen de PBP tipo B ( <i>ftsI</i> o <i>mrda</i> ) y fuera del clúster <i>dcw</i>	RodA clásica	22,6 %
<i>rodA</i> O	Contigua a gen de PBP tipo B ( <i>ftsI</i> o <i>mrda</i> ) que ha perdido su zona TM		7,6 %

Tabla 7: Clasificación de genes *ftsW/rodA* usada para identificar los genes en los genomas estudiados.

Por su parte, los genes *pbp* tipo B son aproximadamente los mismos que *ftsW/rodA* (1.503), pero tienen una distribución distinta (Tabla 8). Los genes insertos en el clúster *dcw*, o sea clasificados inequívocamente como *ftsI* son el 34,4 % del total y los que están fuera de este clúster, pero contiguos a un gen *rodA* son clasificados como genes *mrda*, y corresponden al 29,7 %. Esta disparidad entre *ftsW* y *ftsI* se explica porque puede haber más de un gen *ftsI* en el clúster *dcw*. Por su parte el par *rodA/mrda* parece mucho mejor conservado. Cabe destacar que existe un porcentaje de genes *pbp* tipo B que han perdido su hélice TM, quedando solubles, pero representan apenas un 4,1 % del total, siendo un porcentaje muy menor incluso en PBPs huérfanas.

A partir de esta clasificación es posible inferir que el complejo FtsW/FtsI es un evento conservado, al menos en cuanto a la presencia de los genes correspondientes. Perfectamente el total de los genes clasificados como *ftsW* podría tener una contraparte *pbp* tipo B, si se considera que los genes aislados o huérfanos, igualmente podrían formar el complejo.

Tipo	Contexto genómico	Observación	Proporción
<i>ftsI</i>	En clúster <i>dcw</i> , incluyendo por lo menos un gen <i>ftsW</i> y al menos un gen del proceso de síntesis de PG	FtsI clásica	34,4 %
<i>ftsIx</i>	Como <i>ftsI</i> pero sin su zona TM	FtsI desanclada	1,3 %
<i>mrda</i>	Contigua a gen de <i>rodA</i>	MrdA clásica	29,7 %
<i>mrdaX</i>	Como <i>mrda</i> pero sin su zona TM	MrdA desanclada	0,8 %
<i>pbp O</i>	PBP tipo B huérfana		31,7 %
<i>pbp Ox</i>	PBP tipo B huérfana desanclada		2,0 %

Tabla 8: Clasificación de genes *ftsI/mrdA* usada para identificar los genes en los genomas estudiados.

### 7.3.3. Modelamiento evolutivo del complejo FtsW/FtsI

Con el objetivo de analizar la evolución de FtsW, FtsI y el complejo que forman, se hicieron alineamientos múltiples y a partir de éstos, dendogramas con todas las secuencias traducidas de los genes identificados la sección anterior. Estos dendogramas se convirtieron en árboles filogenéticos usando como raíz una secuencia de proteína de membrana no emparentada directamente (SecY para FtsW/RodA y betalactamasa para FtsI/MrdA).

En la Figura 27 se ve el árbol filogenético para todas las proteínas FtsW/RodA identificadas, indicando en los anillos externos su clasificación (según la Tabla 7) y los filios taxonómicos a los que pertenecen.

Del análisis del árbol filogenético de FtsW/RodA se pueden diferenciar zonas exclusivas con proteínas FtsW o RodA, en sus distintas variantes, que no se mezclan entre ellas. Esto es interesante, pues en la construcción del árbol se consideraron exclusivamente las secuencias aminoacídicas y no la clasificación realizada mediante el contexto genómico. Esto podría indicar que los genes de *ftsW* y *rodA* se diferenciaron temprano en la evolución, en varios episodios discretos y luego se transmitieron verticalmente en las distintas especies que fueron apareciendo. De hecho la distribución de filios taxonómicos (que se muestra en el anillo exterior) parece apoyar esta idea.



Los genes huérfanos (amarillos en el anillo interno en la Figura 27) aparecen intercalados en todo el árbol, tanto en regiones de *ftsW* y *rodA*. Esto no invalida la hipótesis anterior, pues estos genes huérfanos podrían corresponder a duplicados de *ftsW* y *rodA* o a transposiciones virales que se intercalan en regiones distintas del genoma y que podrían mantener su función, asociándose a las mismas u otras PBP o permanecer como huellas vestigiales del fenómeno de duplicación.

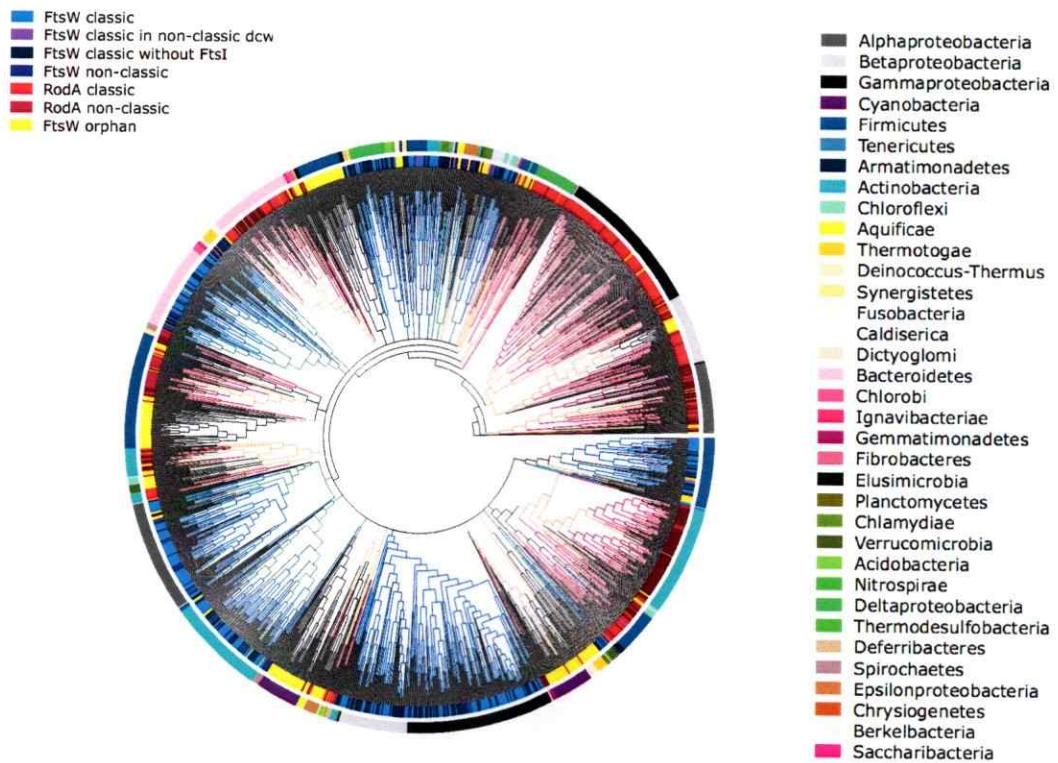


Figura 27: Árbol filogenético de las secuencias proteicas de FtsW/RodA. De los círculos que rodean el árbol, el interior indica el tipo de proteína, según la clasificación mostrada en la Tabla 8, el círculo exterior muestra los filos taxonómicos a los que pertenecen. Los colores de los clados al interior, corresponden a la distribución de tipo de proteína de los nodos hijos. Azul si todos los nodos son FtsW, rojos si son RodA, naranja si son RodA mezcladas con FtsW huérfanas, verde si son FtsW mezcladas con FtsW huérfanas, negros si es una mezcla más compleja.

El mismo tipo de análisis se puede realizar para los genes *ftsI/mrdA*. Estos genes codifican

para proteínas PBP tipo B, y en general poseen una zona TM, un dominio central y un dominio TP. Dado que el dominio TP es sensible a compuestos betalactámicos, está sometido a una presión evolutiva diferente del resto de la proteína. Por lo tanto, para poder realizar un análisis evolutivo, se quitó este dominio de la construcción de los árboles filogenéticos.

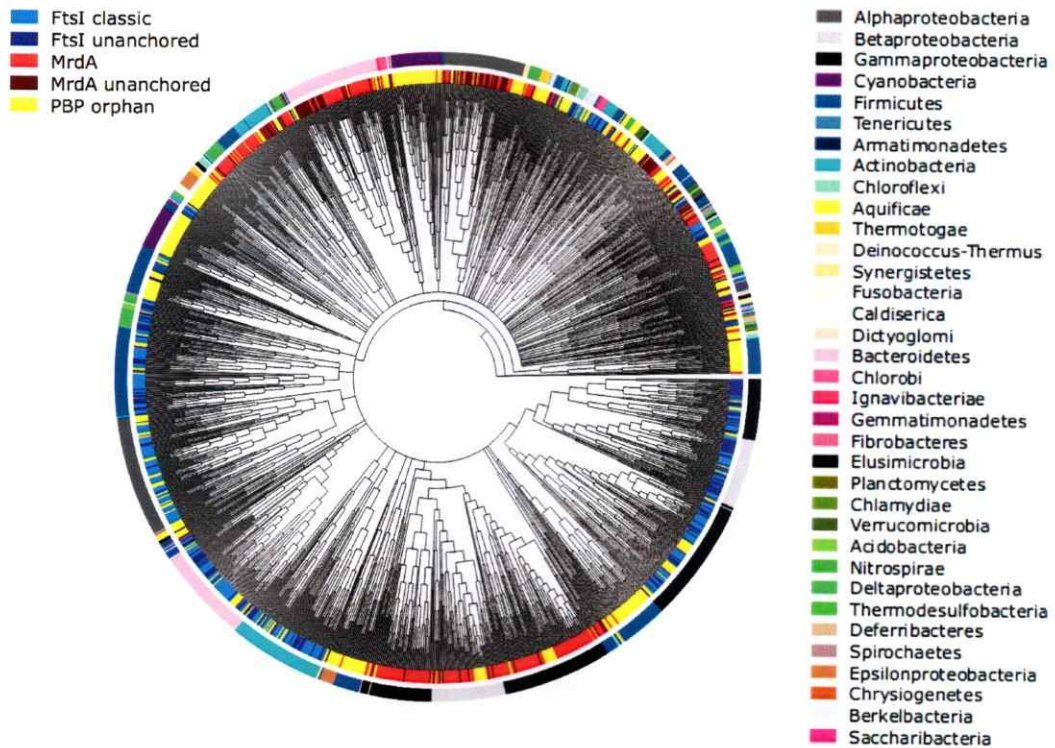


Figura 28: Árbol filogenético de las secuencias proteicas de FtsI/MrdA. De los círculos que rodean el árbol, el interior indica el tipo de proteína, según la clasificación mostrada en la Tabla 8, el círculo exterior muestra los filos taxonómicos a los que pertenecen.

En la Figura 28 se observa que, al igual que en el análisis anterior, las proteínas PBP tipo B parecen distribuidas en regiones exclusivas tanto para FtsI como para MrdA, a pesar de su similitud. Tal como en FtsW/RodA, en general solo las PBP huérfanas (en amarillo) aparecen entremezcladas con las otras, pudiendo deberse al mismo fenómeno discutido para FtsW/RodA. Sin embargo, en una región del gráfico (zona superior izquierda) parece haber una mezcla de proteínas FtsI con MrdA, con presencia de proteínas de los taxones Actinobacteria y Firmicutes,

entre otros. Dentro de estos taxones podemos encontrar bacterias que esporulan, y ello podría explicar la mezcla de proteínas, dada su distribución distintiva de genes.

Para profundizar en la evolución del complejo FtsW/FtsI, sin la interferencia de los genes *ftsW/rodA* que hayan perdido a su compañero PBP, se seleccionaron las secuencias de las proteínas tipo FtsW I, FtsW II y RodA, que poseen solo una PBP cercana. También se seleccionaron las proteínas PBP correspondientes. Con ambos grupos de proteínas se realizaron alineamientos múltiples y dendrogramas, que se muestran en la Figura 29.

En esta figura es posible ver, que la distribución de taxones es bastante similar entre FtsW y FtsI, confirmando la idea de que este complejo pudo haber evolucionado de manera conjunta.

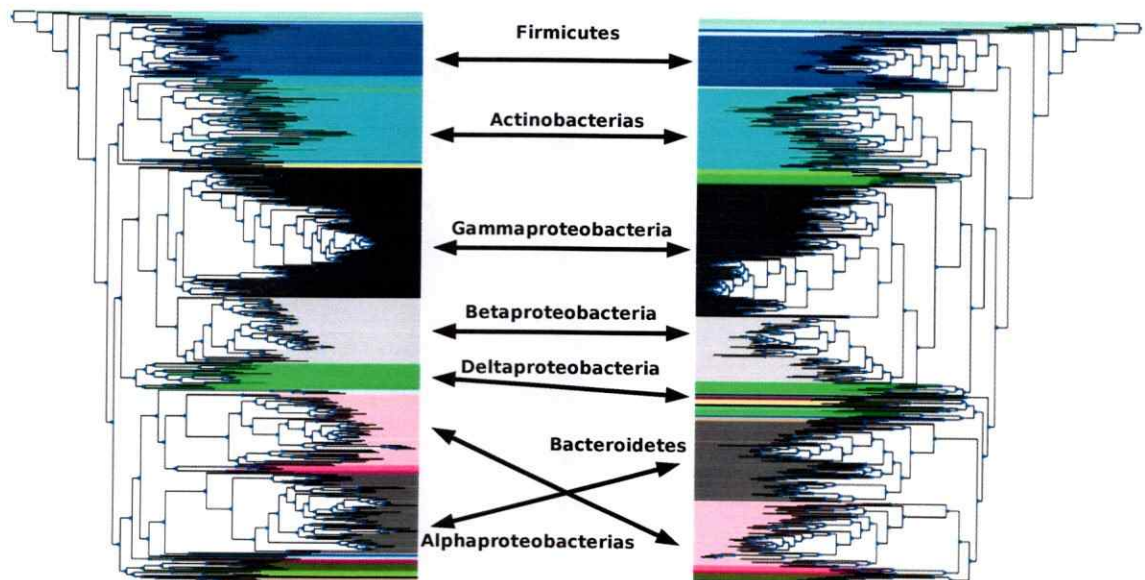


Figura 29: Comparación de dendrogramas de las proteínas FtsW y FtsI que podrían formar un único complejo por especie. Los colores indican filo taxonómico tal como en las figuras anteriores y se indican los más numerosos.

## 8. Discusión y Conclusiones

En esta Tesis se estudió el complejo que forman las proteínas FtsW y FtsI en *E. coli*, así como las proteínas por separado. Este complejo es esencial para la división bacteriana e interactúa con otras proteínas del divisoma.

### *FtsW y FtsI*

En un primer objetivo se estudió en profundidad cada proteína tanto a nivel de conservación de su secuencia como de su estructura, con el objetivo de obtener un modelo del complejo FtsW/FtsI, o al menos proponer regiones de interacción entre ambas proteínas. Con este mismo objetivo, en ambas se estudiaron las propiedades de sus zonas TM (10 hélices para FtsW; 1 hélice para FtsI), de modo de conocer su posible orientación en la membrana. Para FtsW se identificaron zonas TM claves en la posible interacción con varias otras hélices, haciendo las veces de nucleadoras de la estructura terciaria. Así también se identificaron loops periplásmicos, que parecen tener una función clave en FtsW, tanto por la conservación de sus residuos como de su longitud. El loop 7-8 por cierto, uno de los más conservados parece clave en la función.

Todos estos *loops* se encuentran orientados hacia el periplasma y son ricos en residuos cargados altamente conservados. Esto tiene sentido pensando en el sustrato de FtsW, la molécula de Lípido II. Esta molécula posee una cola lipídica (undecaprenol) unida a pirofosfato, que a su vez está unido a su parte glicopeptídica. Estos grupos fosfatos quedan alineados con las cabezas polares de los fosfolípidos, y en general están coordinados con iones magnesio que neutralizan la carga negativa. En otras proteínas que unen undecaprenil-fosfato, como en UppP de *E. coli*, los residuos cargados en esta región de interfase membrana-periplasma son claves en la unión de esta molécula [122, 123].

La controversial función de FtsW, fue considerada en estos análisis para explorar la compatibilidad de los resultados con una de las dos hipótesis planteadas para explicar su función. La primera hipótesis sobre la función de FtsW propone que es una flípasa de Lípido II que lo trasloca

desde la cara interna a la cara externa de la membrana bacteriana interna. Esta hipótesis fue considerada por buena parte de la comunidad experta en el área, incluso antes de la aparición de evidencia experimental [19, 63]. Sin embargo, estas evidencias no parecieron convencer a todos, dada la propuesta de una flipasa diferente, la proteína de membrana MurJ, que también contaba con evidencia experimental [66, 68, 50] y que comenzó a recibir la aceptación de la comunidad. En lo que concierne a esta Tesis, la posible función flipasa parece difícil de sostener por dos razones de peso: por un lado, las proteínas con función flipasas conocidas, tanto en *E. coli* como en otras especies, no tienen ningún tipo de similitud con FtsW, de hecho la tienen con la otra flipasa propuesta, MurJ. Por otro lado, FtsW y las proteínas de la familia SEDS, no tienen residuos conservados en la cara citoplasmática, sino más bien en la cara periplásmica, haciendo difícil concebir un mecanismo de reconocimiento para este lípido en caso de ser una flipasa. De hecho en MurJ, la distribución de residuos conservados parece estar más homogéneamente distribuída entre regiones citoplásmicas y periplásmicas (Figura 30).

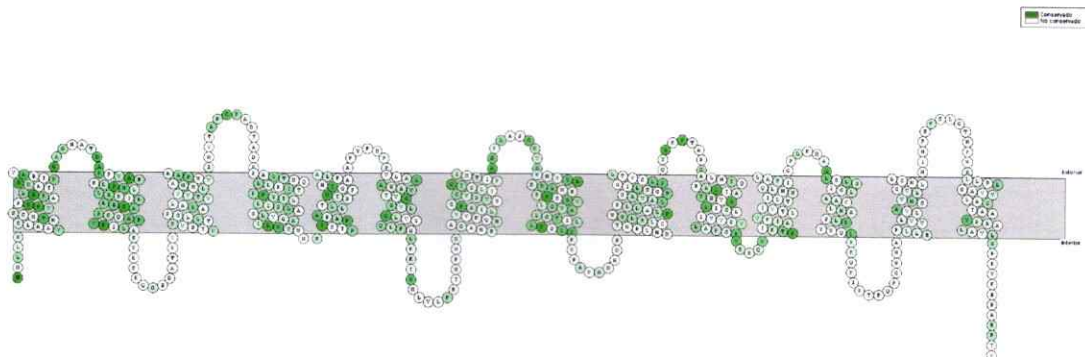


Figura 30: Conservación de residuos en flipasa MurJ. Los tonos de verde más oscuros indican mayor conservación. Imagen generada con el programa memprot, desarrollado en esta Tesis.

La segunda hipótesis considerada acerca de la función de FtsW, indica que podría ser una transglicosilasa, que usa Lípido II como dador del grupo glicopeptídico. Esto a partir de homologías remotas encontradas, usando modelos estadísticos de representación de secuencias [69]. Esta función tendría mucho sentido ya que de este modo, el complejo FtsW/FtsI sería suficiente para extender y entrecruzar las cadenas de peptidoglicano que crecen hacia el interior de la célula durante la división. Esta hipótesis cuenta con evidencia experimental a su favor, pero usando el

homólogo cercano de FtsW, RodA, la que probablemente cumple una función molecular idéntica a FtsW, pero en el contexto de la elongación, formando un complejo con MrdA, homóloga de FtsI [72].

La disponibilidad de estructuras tridimensionales de transglicosilasas de membrana, permitieron hacer comparaciones pertinentes con FtsW. De este modo, los análisis tanto estructurales como de secuencia hechos en esta Tesis apuntan a que hay varias similitudes con estas transglicosilasas. La estructura general parece coincidir con estas enzimas, tanto en modelos comparativos (Figura 18) como en modelos *ab initio* que no tienen en cuenta esta hipótesis (Figura 13). Interesantemente, estos últimos modelos muestran una coincidencia importante con los modelos comparativos, en cuanto al ordenamiento de las zonas TM dentro de la membrana, sus interacciones y la ubicación de los loops. Además, se encuentran niveles de similitud de secuencia en varias regiones de la proteína, como zonas TM, en general poco conservadas, y en loops periplásmicos, lo que le da sustento a nuestra hipótesis. Solo la región desde el N-terminal hasta el TM2 parece no coincidir con las enzimas transglicosilasas, lo que se condice con la poca conservación de esta región en las proteínas SEDS. Una observación interesante es que las regiones TM1 y TM2 parecen mejor modeladas en el modelo *ab initio* que en el comparativo, pues satisfacen mejor las predicciones de contactos entre residuos, ya que en el modelo comparativo estas regiones no tienen un patrón para comparar y solo se usaron un par de TMs extra observadas en 5F15, pero que claramente no corresponden evolutivamente con TM1 y TM2 de FtsW (Figura 17).

Para tener evidencia experimental que apoye o desmienta a los modelos propuestos se realizó el ensayo de accesibilidad a cisteínas (SCAM) en FtsW. De las zonas TM probadas para su accesibilidad, solo la TM10 se muestra como accesible al solvente y TM4 eventualmente presenta una accesibilidad baja (Figura 22). En conjunto, esta evidencia es difícilmente compatible con la hipótesis que FtsW sea una flipasa, dada la falta de un poro o canal por donde pase la región polar de Lípido II, lo que sí se observa en MurJ, con esta misma técnica [68]. Esto se complementa con la falta de conservación en zonas citoplasmáticas de la proteína. Respecto del

modelo estructural planteado se observa congruencia con los resultados de accesibilidad, con la excepción de TM10, que no debería estar expuesta según el modelo. Sin embargo, esta hélice es larga e inclinada y está próxima a la TM4 que está quebrada. Esto podría indicar que es posible que entre TM10 y TM4 se presente un sitio relativamente accesible hacia al periplasma, si la proteína es lo suficientemente flexible o las condiciones experimentales usadas para hacer el ensayo favorecieran una conformación de este tipo. De hecho, las proteínas usadas para este ensayo provienen de un gen en *trans* (de un plasmidio) que podría tener una expresión mayor a la proteína cromosomal. Este exceso de proteína podría no encontrar una cantidad equivalente de proteína FtsI para formar el complejo FtsW/FtsI u otras proteínas del divisoma, y esto a su vez podría generar una proteína con un plegamiento no óptimo, como ha sido observado para otras proteínas del divisoma [34]. Una proyección de esta Tesis es continuar los ensayos de SCAM con otras mutantes en otras regiones TM o en loops, para seguir validando la estructura tridimensional propuesta.

Para concluir esta parte de la Tesis se puede decir que el modelo estructural propuesto para FtsW, tiene consistencia estadística con las estructuras experimentales de proteínas de membrana disponibles y es estable en simulaciones moleculares una vez embebida en una bicapa lipídica. Aunque lo ideal es realizar estas simulaciones a tiempos mayores para verificar la estabilidad del modelo, trabajo que se plantea como una proyección a esta Tesis. Adicionalmente los resultados de SCAM, si bien no son concluyentes, no apoyan la hipótesis de que FtsW tenga un poro como en otras flipasas, y en buena medida son congruentes con el modelo estructural planteado. Por lo tanto, según los análisis *in silico* y experimentales hechos en esta Tesis, se concluye que la hipótesis de que FtsW es una transglicosilasa y no una flipasa, es la más plausible.

Esta conclusión acerca de la nueva función de FtsW, puede ser reforzada desde un punto de vista evolutivo. Las transglicosilasas que presentan esta homología remota con FtsW, miembros de la familia de proteínas STT3, contienen miembros en eucariontes y arqueas, pero no en bacterias. Estas proteínas presentan un núcleo de plegamiento común con los modelos propuestos para FtsW, pero con algunas características distintas como la presencia de un gran dominio

C-terminal de unión y reconocimiento al aceptor del grupo glicosídico, que en algunos casos es una proteína y en otros, son otros lípidos. Justamente el aceptor del grupo glicosídico en FtsW es diferente, es la capa de peptidoglicano, por lo que tiene sentido pensar en la pérdida de este dominio. Otra región, presente en las proteínas STT3 y ausente en FtsW, es el largo *loop* entre las zonas TM1 y TM2, que también interactúa con el dominio C-terminal, por lo que también tiene sentido su pérdida durante la evolución dado el cambio en el sustrato aceptor. Por lo tanto, es razonable pensar que ambas familias de proteínas, SEDS y STT3, tienen un ancestro común presente en proto-organismos previos a la separación entre bacterias, arqueas y eucariontes. Entonces, la diferenciación de estas proteínas pudo estar ligada a la diferenciación de las bacterias, que desarrollaron una pared de peptidoglicano, única para este tipo de células. Ciertamente, para explorar en detalle esta posibilidad hacen falta estudios filogenéticos y evolutivos más detallados.

Por su parte, FtsI es una proteína más estudiada que FtsW, tanto por su dominio de unión a penicilina, como por su homología con otras PBP tipo B, que incluso poseen estructura conocida. Por lo tanto, el foco de esta Tesis respecto a FtsI fue la conservación y estructura de su región más próxima a la membrana y por lo tanto a FtsW. Su única zona TM no es tan conservada, aunque en la región proximal, entre la zona TM y el dominio no-catalítico parecen haber residuos clave en la interacción con FtsW. Las mutaciones correlacionadas apuntan a esta región, así como algunas partes del dominio no-catalítico, que muestran la importancia de ellos para la interacción con FtsW. Llama la atención dos residuos de arginina altamente conservados en la zona TM de FtsI, ambos mirando la misma cara de la hélice, a la altura de las cabezas polares de los fosfolípidos. Por lo que se podría pensar que estos son importantes para la interacción con FtsW, pero no aparecen correlacionados con ningún residuo en sus respectivas compañeras SEDS. Además aparecen en la cara opuesta a otros residuos sí correlacionados con residuos en FtsW.



### Complejo FtsW/FtsI

A partir de los modelos construidos para ambas proteínas se proponen dos posibles regiones de interacción (Figura 25), que estarían dando origen al complejo FtsW/FtsI. Una es la región proximal de FtsI, con la zona TM8-TM9 de FtsW. Otras regiones serían el loop 7-8 y el loop 1-2 de FtsW con loops del dominio no-catalítico de FtsI. Esta última región ya había sido identificada como posible zona de interacción [53] entre ambas proteínas. Sin embargo, en los datos de mutaciones correlacionadas aparecen además varios residuos en el dominio no-catalítico y en el dominio PBP de FtsI haciendo contacto con FtsW, pero que están bastante lejanos espacialmente, según la configuración tipo tallo de FtsI. Esta estructura tipo tallo, es solo una suposición, pues nunca se ha cristalizado una PBP tipo B con su zona TM y por lo tanto no se podría descartar que su estructura en realidad este paralela a la membrana.

Esto implicaría que su zona proximal, que nunca se ha podido cristalizar, esté doblada en

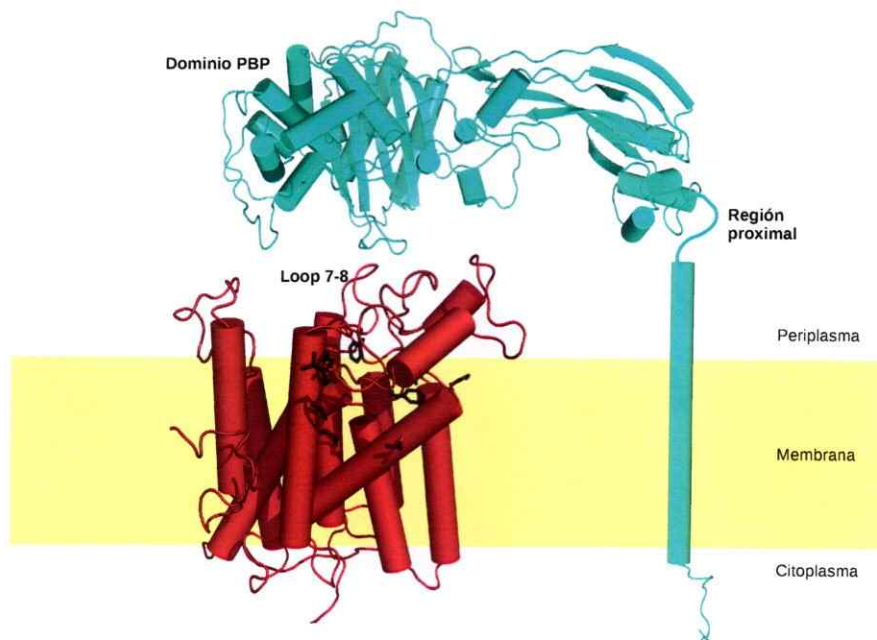


Figura 31: Posible posición de FtsI respecto a la membrana. En esta posición su dominio PBP quedaría a distancia de interacción con varios loops de FtsW, especialmente el loop 7-8.

casi 90° acercando los dominios que parecen más lejanos a la membrana a FtsW, permitiéndole establecer contactos con esta y acercar ambos sitios activos para trabajar coordinadamente (Figura 31). Como proyección para confirmar o refutar esta posibilidad se plantea la opción de hacer pruebas experimentales *in vivo*, como ensayos de entrecruzamiento entre mutantes FtsW y FtsI cisteína-únicas, usando entrecruzantes que reaccionen con grupos tioles en función de su distancia.

Se optó solo por proponer un modo de interacción y no un modelo tridimensional. Esto porque la propuesta de estructura para FtsW es lo suficientemente gruesa para que contenga errores en la posición de ciertos residuos particulares. Así también la estructura propuesta para el loop 7-8 está lejos de ser la óptima, dada su probable flexibilidad y su falta de patrones o moldes similares. Por lo tanto la construcción de un modelo tridimensional exacto entre ambas proteínas, solo obedecería a una sobre-interpretación de los datos, y a un nivel de detalle imposible de obtener con las técnicas empleadas en este trabajo. En un futuro, se podrían diseñar rutinas complejas de modelamiento-simulación de zonas TMs y loops para optimizar su estructura.

### ***Relaciones evolutivas de FtsW y FtsI***

Finalmente en el último objetivo se realizó un análisis evolutivo de ambas proteínas, tanto a nivel de secuencia génica como aminoacídica, para observar la relación entre ellas, el complejo que forman y las otras proteínas del divisoma. Ambas proteínas pertenecen a sus respectivas familias, SEDS para FtsW y PBP tipo B para FtsI. Otros miembros de estas familias son altamente conservados y también forman un complejo entre ellas.

El análisis genético muestra una complementación debido a la presencia de ambos genes en la mayoría de los genomas bacterianos estudiados. Como no siempre es posible identificar a nivel genético los participantes de la formación de cada complejo FtsW/FtsI, se propuso una nueva clasificación, tanto de los genes de *ftsW* (genes de la familia SEDS), como para *ftsI* (genes *pbp* tipo B) basada en el contexto genómico. Gracias a esta clasificación, se identificó que la mayoría de los genes *ftsW* están dentro del clúster *dcw* y es posible encontrar uno o más genes

*ftsI* en la misma región. Otro complejo que se incluyó en este análisis fue el de RodA/MrdA que se ha propuesto cumple la misma función molecular que FtsW/FtsI, pero en un contexto fisiológico distinto, como es la elongación. La clasificación propuesta permite diferenciar entre ambos tipos de complejo.

Como complemento, se analizó la relación evolutiva entre las secuencias proteicas, independiente del contexto genómico de sus respectivos genes, pudiendo construir árboles filogenéticos de ambas proteínas. Al cruzar estos resultados con la clasificación propuesta se logró ver que tanto FtsW como RodA son diferenciables en estos dendogramas (Figura 27), pudiendo haberse diferenciado en eventos evolutivos discretos. Lo mismo se observó para FtsI y MrdA, apoyando la idea que ambos complejos han evolucionado de manera separada, y en general como complejo. La presencia de genes huérfanos se observan intercalados en los dendogramas de ambos tipos de proteínas y podrían explicarse como eventos de duplicación/empalme en regiones diferentes de los genomas. Esto no refutaría la hipótesis de la coevolución de FtsW con FtsI.

Finalmente, a manera de conclusión, en esta Tesis se propone la conservación del complejo FtsW/FtsI debido a la presencia de sus respectivos genes en contextos genómicos específicos, y por las relaciones filogenéticas que se establecen a partir de sus secuencias proteicas. También se entrega evidencia que apoya la idea de que FtsW es una transglicosilasa de peptidoglicano y no una flipasa para Lípido II. Se propone una estructura para FtsW basada en este tipo de enzimas y los resultados experimentales apoyan este modelo. La estructura modelada de FtsW y de FtsI completa, permite proponer un modo de unión entre ambas, identificando regiones que consecuentemente serían importantes para todo el proceso de división bacteriana [42].

## Referencias

- [1] J. Errington, R. A. Daniel, and D. J. Scheffers, "Cytokinesis in Bacteria," *Microbiology and Molecular Biology Reviews*, vol. 67, no. 1, pp. 52-65, 2003.
- [2] M. Vicente and A. I. Rico, "Septum Enlightenment: Assembly of Bacterial Division Proteins," *Journal of Bacteriology*, vol. 188, no. 1, pp. 19-27, 2006.
- [3] H. P. Erickson, "FtsZ, a tubulin homologue in prokaryote cell division," *Trends in Cellular Biology*, vol. 7, no. 9, pp. 362-367, 1997.
- [4] J. Löwe and L. A. Amos, "Crystal structure of the bacterial cell-division protein FtsZ," *Nature*, vol. 391, no. 6663, pp. 203-206, 1998.
- [5] Z. Li, M. J. Trimble, Y. V. Brun, and G. J. Jensen, "The structure of FtsZ filaments *in vivo* suggests a force generating role in cell division," *The EMBO Journal*, vol. 26, no. 22, pp. 4694-4708, 2007.
- [6] M. Osawa, D. E. Anderson, and H. P. Erickson, "Reconstitution of contractile FtsZ rings in liposomes," *Science*, vol. 320, no. 5877, pp. 792-794, 2008.
- [7] R. Díaz-Espinoza, A. P. Garce, J. J. Arbildúa, F. Montecinos, J. Brunet, R. Lagos, and O. Monasterio, "Domain folding and flexibility of *Escherichia coli* FtsZ determined by tryptophan site-directed mutagenesis," *Protein Science*, vol. 16, pp. 1543-1556, 2007.
- [8] A. J. Martín-Galiano, R. M. Buey, M. Cabezas, and J. M. Andreu, "Mapping flexibility and the assembly switch of cell division protein FtsZ by computational and mutational approaches," *Journal of Biological Chemistry*, vol. 285, no. 29, pp. 22554-22565, 2010.
- [9] D. RayChaudhuri, "ZipA is a MAP-Tau homolog and is essential for structural integrity of the cytokinetic FtsZ ring during bacterial cell division," *EMBO Journal*, vol. 18, no. 9, pp. 2372-2383, 1999.
- [10] S. Pichoff and J. Lutkenhaus, "Unique and overlapping roles for ZipA and FtsA in septal ring assembly in *Escherichia coli*," *EMBO Journal*, vol. 21, no. 4, pp. 685-693, 2002.

- [11] K. L. Schmidt, N. D. Peterson, R. J. Kustus, M. C. Wissel, B. Graham, G. J. Phillips, and D. S. Weiss, "A predicted ABC transporter, FtsEX, is needed for cell division in *Escherichia coli*," *Journal of Bacteriology*, vol. 186, no. 3, pp. 785-793, 2004.
- [12] B. D. Corbin, B. Geissler, M. Sadasivam, and W. Margolin, "Z-ring-independent interaction between a subdomain of FtsA and late septation proteins as revealed by a polar recruitment assay," *Journal of Bacteriology*, vol. 186, no. 22, pp. 7736-7744, 2004.
- [13] A. I. Rico, M. García-Ovalle, J. Mingorance, and M. Vicente, "Role of two essential domains of *Escherichia coli* FtsA in localization and progression of the division ring," *Molecular Microbiology*, vol. 53, no. 5, pp. 1359-1371, 2004.
- [14] W. Steiner, G. Liu, W. D. Donachie, and P. Kuempel, "The cytoplasmic domain of FtsK protein is required for resolution of chromosome dimers," *Molecular Microbiology*, vol. 31, no. 2, pp. 579-583, 1999.
- [15] L. Aussel, F. X. Barre, M. Aroyo, A. Stasiak, A. Z. Stasiak, and D. Sherratt, "FtsK is a DNA motor protein that activates chromosome dimer resolution by switching the catalytic state of the XerC and XerD recombinases," *Cell*, vol. 108, no. 2, pp. 195-205, 2002.
- [16] L. Wang, M. K. Khattar, W. D. Donachie, and J. Lutkenhaus, "FtsI and FtsW are localized to the septum in *Escherichia coli*," *Journal of Bacteriology*, vol. 180, no. 11, pp. 2810-2816, 1998.
- [17] N. Buddelmeijer and J. Beckwith, "A complex of the *Escherichia coli* cell division proteins FtsL, FtsB and FtsQ forms independently of its localization to the septal region," *Molecular Microbiology*, vol. 52, pp. 1315-1327, Jun 2004.
- [18] N. W. Goehring, M. D. Gonzalez, and J. Beckwith, "Premature targeting of cell division proteins to midcell reveals hierarchies of protein interactions involved in divisome assembly," *Molecular Microbiology*, vol. 61, pp. 33-45, Jul 2006.
- [19] T. Mohammadi, V. van Dam, R. Sijbrandi, T. Vernet, A. Zapun, A. Bouhss, M. Diepeveen-de Bruin, M. Nguyen-Distèche, B. de Kruijff, and E. Breukink, "Identification of FtsW

- as a transporter of lipid-linked cell wall precursors across the membrane," *The EMBO Journal*, vol. 30, pp. 1425-32, apr 2011.
- [20] K. Mercer and D. S. Weiss, "The *Escherichia coli* cell division protein FtsW is required to recruit its cognate transpeptidase, FtsI (PBP3), to the division site," *Journal of Bacteriology*, vol. 184, no. 4, pp. 904-912, 2002.
- [21] A. I. Rico, M. García-Ovalle, P. Palacios, M. Casanova, and M. Vicente, "Role of *Escherichia coli* FtsN protein in the assembly and stability of the cell division ring," *Molecular Microbiology*, vol. 76, pp. 760-771, may 2010.
- [22] J. C. Yang, F. Van Den Ent, D. Neuhaus, J. Brevier, and J. Löwe, "Solution structure and domain architecture of the divisome protein FtsN," *Molecular Microbiology*, vol. 52, no. 3, pp. 651-660, 2004.
- [23] P. Müller, C. Ewers, U. Bertsche, M. Anstett, T. Kallis, E. Breukink, C. Fraipont, M. Terrak, M. Nguyen-Distèche, and W. Vollmer, "The essential cell division protein FtsN interacts with the murein (peptidoglycan) synthase PBP1B in *Escherichia coli*," *Journal of Biological Chemistry*, vol. 282, no. 50, pp. 36394-36402, 2007.
- [24] L. Sánchez-Pulido, D. Devos, S. Genevrois, M. Vicente, and A. Valencia, "POTRA: a conserved domain in the FtsQ family and a class of b-barrel outer membrane proteins," *Trends in Biochemical Sciences*, vol. 28, pp. 523-526, oct 2003.
- [25] B. Beall and J. Lutkenhaus, "Impaired cell division and sporulation of a *Bacillus subtilis* strain with the *ftsA* gene deleted," *Journal of Bacteriology*, vol. 174, no. 7, pp. 2398-2403, 1992.
- [26] V. L. Katis, R. G. Wake, and E. J. Harry, "Septal localization of the membrane-bound division proteins of *Bacillus subtilis* DivIB and DivIC is codependent only at high temperatures and requires FtsZ," *Journal of Bacteriology*, vol. 182, no. 12, pp. 3607-3611, 2000.

- [27] V. L. Katis and R. G. Wake, "Membrane-bound division proteins DivlB and DivlC of *Bacillus subtilis* function solely through their external domains in both vegetative and sporulation division," *Journal of Bacteriology*, vol. 181, no. 9, pp. 2710-2718, 1999.
- [28] L. M. Guzman, D. S. Weiss, and J. Beckwith, "Domain-swapping analysis of FtsI, FtsL, and FtsQ, bitopic membrane proteins essential for cell division in *Escherichia coli*," *Journal of Bacteriology*, vol. 179, no. 16, pp. 5094-5103, 1997.
- [29] V. D'Ulisse, M. Fagioli, P. Ghelardini, and L. Paolozzi, "Three functional subdomains of the *Escherichia coli* FtsQ protein are involved in its interaction with the other division proteins," *Microbiology*, vol. 153, pp. 124-138, Jan 2007.
- [30] E. J. Harry, B. J. Stewart, and R. G. Wake, "Characterization of mutations in *divlB* of *Bacillus subtilis* and cellular localization of the DivlB protein," *Molecular Microbiology*, vol. 7, no. 4, pp. 611-621, 1993.
- [31] S. L. Rowland, K. D. Wadsworth, S. A. Robson, C. Robichon, J. Beckwith, and G. F. King, "Evidence from artificial septal targeting and site-directed mutagenesis that residues in the extracytoplasmic  $\beta$  domain of DivlB mediate its interaction with the divisomal transpeptidase PBP 2B," *Journal of Bacteriology*, vol. 192, no. 23, pp. 6116-6125, 2010.
- [32] R. A. Daniel and J. Errington, "Intrinsic instability of the essential cell division protein FtsL of *Bacillus subtilis* and a role for DivlB protein in FtsL turnover," *Molecular Microbiology*, vol. 36, no. 2, pp. 278-289, 2000.
- [33] J. M. Ghigo and J. Beckwith, "Cell Division in *Escherichia coli*: Role of FtsL Domains in Septal Localization, Function, and Oligomerization," *Journal of Bacteriology*, vol. 182, pp. 116-129, Jan 2000.
- [34] S. A. Robson, K. A. Michie, J. P. Mackay, E. Harry, and G. F. King, "The *Bacillus subtilis* cell division proteins FtsL and DivlC are intrinsically unstable and do not interact with one another in the absence of other septosomal components," *Molecular Microbiology*, vol. 44, no. 3, pp. 663-674, 2002.

- [35] F. van den Ent, T. M. F. Vinkenvleugel, A. Ind, P. West, D. Veprintsev, N. Nanninga, T. den Blaauwen, and J. Löwe, "Structural and mutational analysis of the cell division protein FtsQ," *Molecular Microbiology*, vol. 68, pp. 110-123, apr 2008.
- [36] M. D. Gonzalez and J. Beckwith, "Divisome under construction: distinct domains of the small membrane protein FtsB are necessary for interaction with multiple cell division proteins," *Journal of Bacteriology*, vol. 191, pp. 2815-2825, apr 2009.
- [37] M. D. Gonzalez, E. a. Akbay, D. Boyd, and J. Beckwith, "Multiple interaction domains in FtsL, a protein component of the widely conserved bacterial FtsLBQ cell division complex," *Journal of Bacteriology*, vol. 192, pp. 2757-2768, jun 2010.
- [38] A. I. Goranov, L. Katz, A. M. Breier, C. B. Burge, and A. D. Grossman, "A transcriptional response to replication status mediated by the conserved bacterial replication protein DnaA," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 36, pp. 12932-12937, 2005.
- [39] M. Bramkamp, L. Weston, R. a. Daniel, and J. Errington, "Regulated intramembrane proteolysis of FtsL protein and the control of cell division in *Bacillus subtilis*," *Molecular Microbiology*, vol. 62, pp. 580-591, oct 2006.
- [40] Y. Akiyama, K. Kanehara, and K. Ito, "RseP (YaeL), an *Escherichia coli* RIP protease, cleaves transmembrane sequences," *The EMBO Journal*, vol. 23, pp. 4434-4442, nov 2004.
- [41] M. Bramkamp, "The putative *Bacillus subtilis* L,D-transpeptidase YciB is a lipoprotein that localizes to the cell poles in a divisome-dependent manner," *Archives of Microbiology*, vol. 192, pp. 57-68, jan 2010.
- [42] F. Villanelo, A. Ordenes, J. Brunet, R. Lagos, and O. Monasterio, "A model for the *Escherichia coli* FtsB/FtsL/FtsQ cell division complex," *BMC Structural Biology*, vol. 11, p. 28, jan 2011.
- [43] J. Pogliano, K. Pogliano, D. S. Weiss, R. Losick, and J. Beckwith, "Inactivation of FtsI inhibits constriction of the FtsZ cytokinetic ring and delays the assembly of FtsZ rings at



- potential division sites," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, pp. 559–64, jan 1997.
- [44] U. Bertsche, T. Kast, B. Wolf, C. Fraipont, M. E. G. Aarsman, K. Kannenberg, M. von Rechenberg, M. Nguyen-Distèche, T. den Blaauwen, J.-V. Höltje, and W. Vollmer, "Interaction between two murein (peptidoglycan) synthases, PBP3 and PBP1B, in *Escherichia coli*," *Molecular Microbiology*, vol. 61, pp. 675–90, aug 2006.
- [45] A. Piette, C. Fraipont, T. D. Blaauwen, M. E. G. Aarsman, S. Pastoret, and M. Nguyen-distè, "Structural determinants required to target penicillin-binding protein 3 to the septum of *Escherichia coli*," *Journal of Bacteriology*, vol. 186, no. 18, pp. 6110–6117, 2004.
- [46] M. C. Wissel and D. S. Weiss, "Genetic analysis of the cell division protein FtsI (PBP3): Amino acid substitutions that impair septal localization of FtsI and recruitment of FtsN," *Journal of Bacteriology*, vol. 186, no. 2, pp. 490–502, 2004.
- [47] M. C. Wissel, J. L. Wendt, C. J. Mitchell, and D. S. Weiss, "The transmembrane helix of the *Escherichia coli* division protein FtsI localizes to the septal ring," *Journal of Bacteriology*, vol. 187, no. 1, pp. 320–328, 2005.
- [48] M. Ikeda, T. Sato, M. Wachi, H. K. Jung, F. Ishino, Y. Kobayashi, and M. Matsubishi, "Structural similarity among *Escherichia coli* FtsW and RodA proteins and *Bacillus subtilis* SpoVE protein, which function in cell division, cell elongation, and spore formation, respectively," *Journal of Bacteriology*, vol. 171, pp. 6375–8, nov 1989.
- [49] B. Lara and J. A. Ayala, "Topological characterization of the essential *Escherichia coli* cell division protein FtsW," *FEMS Microbiology Letters*, vol. 216, pp. 23–32, oct 2002.
- [50] L.-T. Sham, E. K. Butler, M. D. Lebar, D. Kahne, T. G. Bernhardt, and N. Ruiz, "MurJ is the flippase of lipid-linked precursors for peptidoglycan biogenesis," *Science*, vol. 345, pp. 220–222, jul 2014.

- [51] S. Pastoret, C. Fraipont, T. D. Blaauwen, M. E. G. Aarsman, A. Thomas, R. Brasseur, and M. Nguyen-diste, "Functional analysis of the cell division protein FtsW of *Escherichia coli*," *Journal of Bacteriology*, vol. 186, no. 24, pp. 8370-8379, 2004.
- [52] P. Datta, A. Dasgupta, A. K. Singh, P. Mukherjee, M. Kundu, and J. Basu, "Interaction between FtsW and penicillin-binding protein 3 (PBP3) directs PBP3 to mid-cell, controls cell septation and mediates the formation of a trimeric complex involving FtsZ, FtsW and PBP3 in mycobacteria," *Molecular Microbiology*, vol. 62, pp. 1655-1673, dec 2006.
- [53] D. S. Weiss, J. C. Chen, J. M. Ghigo, D. Boyd, and J. Beckwith, "Localization of FtsI (PBP3) to the septal ring requires its membrane anchor, the Z ring, FtsA, FtsQ, and FtsL," *Journal of Bacteriology*, vol. 181, pp. 508-520, jan 1999.
- [54] C. Fraipont, S. Alexeeva, B. Wolf, R. Der Ploeg, M. Schloesser, T. Den Blaauwen, and M. Nguyen-Distèche, "The integral membrane FtsW protein and peptidoglycan synthase PBP3 form a subcomplex in *Escherichia coli*," *Microbiology*, vol. 157, no. 1, pp. 251-259, 2011.
- [55] M. Marrec-Fairley, A. Piette, X. Gallet, R. Brasseur, H. Hara, C. Fraipont, J. M. Ghuysen, and M. Nguyen-Distèche, "Differential functionalities of amphiphilic peptide segments of the cell-septation penicillin-binding protein 3 of *Escherichia coli*," *Molecular Microbiology*, vol. 37, pp. 1019-1031, sep 2000.
- [56] G. Karimova, N. Dautin, and D. Ladant, "Interaction network among *Escherichia coli* membrane proteins involved in cell division as revealed by bacterial two-hybrid analysis," *Journal of Bacteriology*, vol. 187, no. 7, pp. 2233-2243, 2005.
- [57] H. P. Erickson, D. E. Anderson, and M. Osawa, "FtsZ in bacterial cytokinesis: cytoskeleton and force generator all in one," *Microbiology and Molecular Biology Reviews*, vol. 74, pp. 504-528, dec 2010.

- [58] M. M. Khattar, K. J. Begg, and W. D. Donachie, "Identification of FtsW and characterization of a new *ftsW* division mutant of *Escherichia coli*," *Journal of Bacteriology*, vol. 176, no. 23, pp. 7140-7147, 1994.
- [59] A. O. Henriques, P. Glaser, P. J. Piggot, and C. P. Moran Jr., "Control of cell shape and elongation by the *rodA* gene in *Bacillus subtilis*," *Molecular Microbiology*, vol. 28, no. 2, pp. 235-247, 1998.
- [60] N. G. Stoker, J. K. Broome Smith, A. Edelman, and B. G. Spratt, "Organization and subcloning of the *dacA-rodA-pbpA* cluster of cell shape genes in *Escherichia coli*," *Journal of Bacteriology*, vol. 155, no. 2, pp. 847-853, 1983.
- [61] J. V. Holtje, "Growth of the stress-bearing and shape-maintaining murein sacculus of *Escherichia coli*," *Microbiology and Molecular Biology Reviews*, vol. 62, no. 1, pp. 181-203, 1998.
- [62] B. Lara, D. Mengin-Lecreulx, J. A. Ayala, and J. van Heijenoort, "Peptidoglycan precursor pools associated with *MraY* and *FtsW* deficiencies or antibiotic treatments," *FEMS Microbiology Letters*, vol. 250, pp. 195-200, sep 2005.
- [63] T. Mohammadi, R. Sijbrandi, M. Lutters, J. Verheul, N. I. Martin, T. Den Blaauwen, B. De Kruijff, and E. Breukink, "Specificity of the transport of Lipid II by *FtsW* in *Escherichia coli*," *The Journal of biological chemistry*, vol. 289, pp. 14707-14718, apr 2014.
- [64] A. Chugunov, D. Pyrkova, D. Nolde, A. Polyansky, V. Pentkovsky, and R. Efremov, "Lipid II forms potential "landing terrain" for lantibiotics in simulated bacterial membrane," *Scientific Reports*, vol. 3, pp. 1-11, 2013.
- [65] V. van Dam, R. Sijbrandi, M. Kol, E. Swiezewska, B. de Kruijff, and E. Breukink, "Transmembrane transport of peptidoglycan precursors across model and bacterial membranes," *Molecular Microbiology*, vol. 64, pp. 1105-1114, may 2007.

- [66] N. Ruiz, "Bioinformatics identification of MurJ (MviN) as the peptidoglycan lipid II flipase in *Escherichia coli*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 40, pp. 15553-15557, 2008.
- [67] W. Vollmer and U. Bertsche, "Murein (peptidoglycan) structure, architecture and biosynthesis in *Escherichia coli*," *Biochimica et Biophysica Acta - Biomembranes*, vol. 1778, no. 9, pp. 1714-1734, 2008.
- [68] E. K. Butler, R. M. Davis, V. Bari, P. A. Nicholson, and N. Ruiz, "Structure-function analysis of MurJ reveals a solvent-exposed cavity containing residues essential for peptidoglycan biogenesis in *Escherichia coli*," *Journal of Bacteriology*, vol. 195, pp. 4639-49, oct 2013.
- [69] S. Ovchinnikov, L. Kinch, H. Park, Y. Liao, J. Pei, D. E. Kim, H. Kamisetty, N. V. Grishin, and D. Baker, "Large-scale determination of previously unsolved protein structures using evolutionary information," *eLife*, vol. 4, no. September, pp. 1-25, 2015.
- [70] C. A. Rohl, C. E. Strauss, K. M. Misura, and D. Baker, "Protein structure prediction using Rosetta," *Methods Enzymol*, vol. 383, pp. 66-93, 2004.
- [71] S. Matsumoto, A. Shimada, J. Nyirenda, M. Igura, Y. Kawano, and D. Kohda, "Crystal structures of an archaeal oligosaccharyltransferase provide insights into the catalytic cycle of N-linked protein glycosylation," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, no. 44, pp. 17868-73, 2013.
- [72] A. J. Meeske, E. P. Riley, W. P. Robins, T. Uehara, J. J. Mekalanos, D. Kahne, S. Walker, A. C. Kruse, T. G. Bernhardt, and D. Z. Rudner, "SEDS proteins are a widespread family of bacterial cell wall polymerases," *Nature*, pp. 1-15, 2016.
- [73] H. Cho, C. N. Wivagg, M. Kapoor, Z. Barry, P. D. A. Rohs, H. Suh, J. A. Marto, E. C. Garner, and T. G. Bernhardt, "Bacterial cell wall biogenesis is mediated by SEDS and PBP polymerase families functioning semi-autonomously," *Nature Microbiology*, vol. 1, no. 10, p. 16172, 2016.

- [74] D. J. F. du Plessis, G. Berrelkamp, N. Nouwen, and A. J. M. Driessen, "The lateral gate of SecYEG opens during protein translocation," *The Journal of Biological Chemistry*, vol. 284, pp. 15805–14, jun 2009.
- [75] I. Uchiyama, M. Mihara, H. Nishide, and H. Chiba, "MBGD update 2015: Microbial genome database for flexible ortholog analysis utilizing a diverse set of genomic data," *Nucleic Acids Research*, vol. 43, no. D1, pp. D270–D276, 2015.
- [76] R. C. Edgar, "MUSCLE: Multiple sequence alignment with high accuracy and high throughput," *Nucleic Acids Research*, vol. 32, no. 5, pp. 1792–1797, 2004.
- [77] M. N. Price, P. S. Dehal, and A. P. Arkin, "FastTree 2 - Approximately maximum-likelihood trees for large alignments," *PLoS ONE*, vol. 5, no. 3, 2010.
- [78] J. Felsenstein, "PHYMLIP - Phylogeny inference package - v3.2," 1989.
- [79] F. Abascal, R. Zardoya, and D. Posada, "ProtTest: Selection of best-fit models of protein evolution," *Bioinformatics*, vol. 21, no. 9, pp. 2104–2105, 2005.
- [80] S. Eddy, "Profile hidden Markov models," *Bioinformatics*, vol. 14, no. 9, pp. 755–763, 1998.
- [81] L. Käll, A. Krogh, and E. L. L. Sonnhammer, "A combined transmembrane topology and signal peptide prediction method," *Journal of Molecular Biology*, vol. 338, no. 5, pp. 1027–1036, 2004.
- [82] T. Nugent and D. T. Jones, "Detecting pore-lining regions in transmembrane protein sequences," *BMC Bioinformatics*, vol. 13, no. 169, pp. 1471–2105, 2012.
- [83] T. J. Wheeler, J. Clements, and R. D. Finn, "Skylign: a tool for creating informative, interactive logos representing sequence alignments and profile hidden Markov models," *BMC Bioinformatics*, vol. 15, no. 1, p. 7, 2014.

- [84] N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, and O. Beckstein, "MDAnalysis: A toolkit for the analysis of molecular dynamics simulations," *Journal of Computational Chemistry*, vol. 32, no. 10, pp. 2319–2327, 2011.
- [85] H. Kamisetty, S. Ovchinnikov, and D. Baker, "Assessing the utility of coevolution-based residue-residue contact predictions in a sequence- and structure-rich era," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, no. 39, pp. 15674–9, 2013.
- [86] D. T. Jones, D. W. A. Buchan, D. Cozzetto, and M. Pontil, "PSICOV: Precise structural contact prediction using sparse inverse covariance estimation on large multiple sequence alignments," *Bioinformatics*, vol. 28, no. 2, pp. 184–190, 2012.
- [87] C.-S. Jeong and D. Kim, "Reliable and robust detection of coevolving protein residues," *Protein Engineering, Design & Selection*, vol. 25, no. 11, pp. 705–713, 2012.
- [88] A. Knappik and A. Pluckthun, "An improved affinity tag based on the FLAG peptide for the detection and purification of recombinant antibody fragments," *BioTechniques*, vol. 17, no. 4, pp. 754–761, 1994.
- [89] W. P. Russ and D. M. S., "The GxxxG motif: A framework for transmembrane helix-helix association," *Journal of Molecular Biology*, vol. 296, no. 3, pp. 911–919, 2000.
- [90] G. von Heijne, "Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule," *Journal of Molecular Biology*, vol. 225, no. 2, pp. 487–494, 1992.
- [91] D. T. Jones, W. R. Taylor, and J. M. Thornton, "A mutation data matrix for transmembrane proteins," *FEBS Letters*, vol. 339, no. 3, pp. 269–275, 1994.
- [92] P. C. Ng, J. G. Henikoff, and S. Henikoff, "PHAT: a transmembrane-specific substitution matrix. Predicted hydrophobic and transmembrane," *Bioinformatics (Oxford, England)*, vol. 16, no. 9, pp. 760–766, 2000.

- [93] J.-M. Chang, P. Di Tommaso, J.-F. Taly, and C. Notredame, "Accurate multiple sequence alignment of transmembrane proteins with PSI-Coffee," *BMC Bioinformatics*, vol. 13, no. Suppl 4, p. S1, 2012.
- [94] W. Pirovano, K. A. Feenstra, and J. Heringa, "PRALINETM: a strategy for improved multiple alignment of transmembrane proteins," *Bioinformatics (Oxford, England)*, vol. 24, pp. 492-7, feb 2008.
- [95] F. Armougom, S. Moretti, O. Poirot, S. Audic, P. Dumas, B. Schaeli, V. Keduas, and C. Notredame, "Espresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee," *Nucleic Acids Research*, vol. 34, pp. W604-8, jul 2006.
- [96] M. Stamm, R. Staritzbichler, K. Khafizov, and L. R. Forrest, "Alignment of helical membrane protein sequences using AlignMe," *PLoS ONE*, vol. 8, no. 3, 2013.
- [97] L. S. Johnson, S. R. Eddy, and E. Portugaly, "Hidden Markov model speed heuristic and iterative HMM search procedure," *BMC Bioinformatics*, vol. 11, no. 1, p. 431, 2010.
- [98] D. Engelman, T. Steitz, and A. Goldman, "Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins," *Annual Review of Biophysics and Biophysical Chemistry*, vol. 15, pp. 321-353, 1986.
- [99] E. Michalsky, A. Goede, and R. Preissner, "Loops In Proteins (LIP) a comprehensive loop database for homology modelling," *Protein Engineering Design and Selection*, vol. 16, no. 12, pp. 979-985, 2003.
- [100] W. Liu, M. Eilers, A. B. Patel, and S. O. Smith, "Helix packing moments reveal diversity and conservation in membrane protein structure," *Journal of Molecular Biology*, vol. 337, pp. 713-29, mar 2004.
- [101] U. Göbel, C. Sander, R. Schneider, and A. Valencia, "Correlated mutations and residue contacts in proteins," *Proteins: Structure, Function, and Bioinformatics*, vol. 18, no. 4, pp. 309-317, 1994.

- [102] W. R. Taylor, R. S. Hamilton, and M. I. Sadowski, "Prediction of contacts from correlated sequence substitutions," *Current Opinion in Structural Biology*, vol. 23, no. 3, pp. 473 479, 2013.
- [103] D. Frishman, *Structural Bioinformatics of Membrane Proteins*. Wien, Austria: Springer-Verlag/Wien, 1 ed., 2010.
- [104] T. Nugent and D. T. Jones, "Membrane protein orientation and refinement using a knowledge-based statistical potential," *BMC Bioinformatics*, vol. 14, no. 1, p. 276, 2013.
- [105] G. Postic, Y. Ghouzam, and J.-C. Gelly, "An empirical energy function for structural assessment of protein transmembrane domains," *Biochimie*, vol. 115, pp. 155 61, 2015.
- [106] J. Söding, "Protein homology detection by HMM-HMM comparison," *Bioinformatics*, vol. 21, no. 7, pp. 951 960, 2005.
- [107] G. Schmidt, D. Mannel, H. Mayer, H. Y. Whang, and E. Neter, "Role of a lipopolysaccharide gene for immunogenicity of the enterobacterial common antigen," *Journal of Bacteriology*, vol. 126, no. 2, pp. 579 586, 1976.
- [108] R. Zufferey, R. Knauer, P. Burda, I. Stagljar, S. te Heesen, L. Lehle, and M. Aebi, "STT3, a highly conserved protein required for yeast oligosaccharyl transferase activity *in vivo*," *The EMBO Journal*, vol. 14, no. 20, pp. 4949 60, 1995.
- [109] M. Igura, N. Maita, J. Kamishikiryo, M. Yamada, T. Obita, K. Maenaka, and D. Kohda, "Structure-guided identification of a new catalytic motif of oligosaccharyltransferase," *The EMBO Journal*, vol. 27, no. 1, pp. 234 243, 2008.
- [110] A. Dell, A. Galadari, F. Sastre, and P. Hitchen, "Similarities and differences in the glycosylation mechanisms in prokaryotes and eukaryotes," *International Journal of Microbiology*, vol. 2010, 2010.
- [111] V. I. Petrou, C. M. Herrera, K. M. Schultz, O. B. Clarke, J. Vendome, D. Tomasek, S. Banerjee, K. R. Rajashankar, M. Belcher Dufresne, B. Kloss, E. Kloppmann, B. Rost, C. S.



- Klug, M. S. Trent, L. Shapiro, and F. Mancina, "Structures of aminoarabinose transferase ArnT suggest a molecular basis for lipid A glycosylation," *Science*, vol. 351, no. 6273, pp. 608 612, 2016.
- [112] Y. Choi and C. M. Deane, "FREAD revisited: Accurate loop structure prediction using a database search algorithm," *Proteins: Structure, Function, and Bioinformatics*, vol. 78, no. 6, pp. NA NA, 2009.
- [113] A. Karlin and M. H. Akabas, "Substituted-Cysteine Accessibility Method," *Methods in Enzymology*, vol. 293, no. 8, pp. 123 145, 1998.
- [114] J. Kronengold, E. B. Trexler, F. F. Bukauskas, T. A. Bargiello, and V. K. Verselis, "Single-channel SCAM identifies pore-lining residues in the first extracellular loop and first transmembrane domains of Cx46 hemichannels," *The Journal of General Physiology*, vol. 122, no. 4, pp. 389 405, 2003.
- [115] M. Skerrett, E. Kasperek, F. L. Cao, J. H. Shin, J. Aronowitz, S. Ahmed, and B. J. Nicholson, "Application of SCAM (substituted cysteine accessibility method) to gap junction intercellular channels," *Cell Communication & Adhesion*, vol. 8, no. 4-6, pp. 179 185, 2001.
- [116] E. Sauvage, A. Derouaux, C. Fraipont, M. Joris, R. Herman, M. Rocaboy, M. Schloesser, J. Dumas, F. Kerff, M. Nguyen-Distehe, and P. Charlier, "Crystal structure of penicillin-binding protein 3 (PBP3) from *Escherichia coli*," *PLoS ONE*, vol. 9, no. 5, pp. 1 11, 2014.
- [117] S. Sainsbury, L. Bird, V. Rao, S. M. Shepherd, D. I. Stuart, W. N. Hunter, R. J. Owens, and J. Ren, "Crystal structures of penicillin-binding protein 3 from *Pseudomonas aeruginosa*: comparison of native and antibiotic-bound forms," *Journal of Molecular Biology*, vol. 405, pp. 173 84, jan 2011.
- [118] H. Yoshida, F. Kawai, E. Obayashi, S. Akashi, D. I. Roper, J. R. H. Tame, and S.-Y. Park, "Crystal structures of penicillin-binding protein 3 (PBP3) from methicillin-resistant

- Staphylococcus aureus* in the apo and cefotaxime-bound forms," *Journal of Molecular Biology*, vol. 423, pp. 351-64, oct 2012.
- [119] A. J. Powell, J. Tomberg, A. M. Deacon, R. A. Nicholas, and C. Davies, "Crystal structures of penicillin-binding protein 2 from penicillin-susceptible and -resistant strains of *Neisseria gonorrhoeae* reveal an unexpectedly subtle mechanism for antibiotic resistance," *Journal of Biological Chemistry*, vol. 284, no. 2, pp. 1202-1212, 2009.
- [120] M. Vicente, M. J. Gomez, and J. A. Ayala, "Regulation of transcription of cell division genes in the *Escherichia coli* *dcw* cluster," *Cellular and Molecular Life Sciences*, vol. 54, no. 4, pp. 317-324, 1998.
- [121] F. Ishino, W. Park, S. Tomioka, S. Tamaki, I. Takase, K. Kunugita, H. Matsuzawa, S. Asoh, T. Ohta, and B. G. Spratt, "Peptidoglycan synthetic activities in membranes of *Escherichia coli* caused by overproduction of penicillin-binding protein 2 and RodA protein," *Journal of Biological Chemistry*, vol. 261, no. 15, pp. 7024-7031, 1986.
- [122] R. T. Guo, T. P. Ko, A. P. C. Chen, C. J. Kuo, A. H. J. Wang, and P. H. Liang, "Crystal structures of undecaprenyl pyrophosphate synthase in complex with magnesium, isopentenyl pyrophosphate, and farnesyl thiopyrophosphate: Roles of the metal ion and conserved residues in catalysis," *Journal of Biological Chemistry*, vol. 280, no. 21, pp. 20762-20774, 2005.
- [123] H. Y. Chang, C. C. Chou, M. F. Hsu, and A. H. J. Wang, "Proposed carrier lipid-binding site of undecaprenyl pyrophosphate phosphatase from *Escherichia coli*," *Journal of Biological Chemistry*, vol. 289, no. 27, pp. 18719-18735, 2014.