UNIVERSIDAD MAYOR
para espíritus emprendedores

Facultad de Ciencias
ESCUELA DE BIOTECNOLOGÍA

Diseño racional de péptidos antigénicos de *P. salmonis* y expresión en un sistema heterólogo

Héctor Eduardo Alfredo Vásquez Aguirre

Proyecto de Tesis para optar al título de Biotecnólogo

Santiago - Chile

Año 2023

UNIVERSIDAD MAYOR
para espíritus emprendedores

Facultad de Ciencias
ESCUELA DE BIOTECNOLOGÍA

Diseño racional de péptidos antigénicos de *P. salmonis* y expresión en un sistema heterólogo

Héctor Eduardo Alfredo Vásquez Aguirre

Proyecto de Tesis para optar al título de Biotecnólogo

Tutor: Dr. Sebastián Reyes Cerpa; Co-tutor: Dr. Jaime Andrés Rivas Pardo

Laboratorio de Genómica Microbiana

Centro de Genómica y Bioinformática, Universidad Mayor

Santiago - Chile
Año 2023

INFORMACIÓN DEL LABORATORIO Y/O CENTRO DE INVETIGACIÓN

Nombre del laboratorio y/o Centro de Investigación: Laboratorio de Genómica Microbiana del Centro de Genómica y Bioinformática de la Universidad Mayor.

Fuente de Financiamiento: Proyecto FONDECYT REGULAR 1221064.

ABREVIATURAS

ADN	Ácido desoxirribonucleico
ADNp	Ácido desoxirribonucleico plasmídico
ADNr	Ácido desoxirribonucleico ribosómico
ARNr	Ácido ribonucleico ribosómico
ARNt	Ácido ribonucleico de transferencia
BLAST	Herramienta de Búsqueda de Alineamiento Local Básico
Blastp	Herramienta de Búsqueda de Alineamiento Local Básico para secuencias proteicas
COG	Grupos Ortólogos de Proteínas
E/PE	Fase exponencial/Fase post-exponencial
EF-Ts	Factor de elongación termoestable
E-up	Proteínas más abundantes en fase exponencial
FAO	Organización de las Naciones Unidas para la Alimentación y la Agricultura
FPE	Fracción proteica de la expresión
GST	Proteína de transferencia de glutatión
HAP	Neumonía adquirida en el hospital

HF	Alta fidelidad
Hsp	Proteína de choque térmico
hVDAC1	Canal aniónico dependiente de voltaje de mitocondrias humanas
IPTG	Isopropil-β-d-tiogalactopiranósido
L1	Loop largo no estructurado de superficie número 1 de la proteína barril-beta de membrana externa de *P. salmonis* (WP_016209892.1)
LB	Medio Luria-Bertani
MBP	Proteína de unión a maltosa
NCBI	Nacional Center for Biotechnology Information
NEB	New England BioLabs
NMR	Resonancia magnética nuclear
nr	No-redundante
OD_{600}	Densidad óptica a 600 nm
OMAP	Proteínas asociadas a la membrana externa
OmpA	Proteína A de membrana externa
ONG	Organización no gubernamental

PBS	Solución salina tamponada con fosfato
PCR	Reacción en cadena de la polimerasa
PCR-RFLP	Reacción en cadena de la polimerasa con el polimorfismo de longitud del fragmento de restricción
qPCR	Reacción en cadena de la polimerasa cuantitativo
SCFV	Proteína de tamaño pequeño con circulación corta
SDS-PAGE	Electroforesis en gel de poliacrilamida con dodecilsulfato de sodio
Sernapesca	Servicio Nacional de Pesca y Acuicultura del Gobierno de Chile
SRS	Síndrome rickettsial salmonídeo
SUMO	Proteína similar a ubiquitina
T3SS	Sistema de secreción tipo III
T4BSS	Sistema de secreción tipo IVB
TAE	Disolución tampón formada por tris, acetato y ácido etilendiaminotetraacético
UV	Ultravioleta

ÍNDICE DE CONTENIDOS

INFORMACIÓN DEL LABORATORIO Y/O CENTRO DE INVETIGACIÓN..... i

ABREVIATURAS.....ii

ÍNDICE DE CONTENIDOS..... i

ÍNDICE DE TABLAS iv

ÍNDICE DE FIGURAS v

RESUMEN vii

SUMMARY viii

1. INTRODUCCIÓN. 1

1.1. Salmonicultura chilena 1

1.2. Piscirickettsia salmonis..... 2

1.3. Generación de anticuerpos y antígenos..... 4

2. HIPÓTESIS Y OBJETIVOS..... 8

2.1. Hipótesis científica 8

2.2. Objetivo general..... 8

2.3. Objetivos específicos..... 8

3. MATERIALES Y MÉTODOS..... 9

3.1. Materiales. 9

3.1.1. Recursos bioinformáticos utilizados..... 9

3.1.2. Plásmidos y cepas bacterianas utilizados..... 9

3.2. Métodos. 10

3.2.1. Selección péptido antigénico para construcción in silico de vector de clonamiento. ... 10

3.2.1.1. Búsqueda bibliográfica de proteínas de superficie abundantes en contexto de infección de organismos filogenéticamente cercanos a *P. salmonis.* 10

3.2.1.2. Identificación de las proteínas homólogas en *P. salmonis* y filtro por conservación entre especies. 10

3.2.1.3. Modelamiento de las proteínas seleccionados post-filtro por conservación e identificación de regiones expuestas al medio extracelular. 11

3.2.1.4. Selección péptido antigénico de *P. salmonis*, construcción y síntesis de pUC57 conteniendo la secuencia nucleotídica del péptido seleccionado. 12

3.2.2. Subclonamiento del péptido antigénico L1 desde vector de clonamiento pUC57 a vector de expresión pFN18A. 12

3.2.2.1. Ensayo de diluciones de ADNp para estandarización en visualización por electroforesis en gel de agarosa TAE. 12

3.2.2.2. Preparación de células *E. coli* XL10, BLR (DE3) pLysS y C41 (DE3) quimiocompetentes. 13

3.2.2.3. Propagación de XL10 pUC57-L1 y XL10 pFN18A-10, seguido de purificación de pUC57-L1B y pFN18A-L1. 13

3.2.2.4. Amplificación, digestión y purificación de L1 desde pUC57-L1B..... 15

3.2.2.5. Ensayos de digestión de pFN18A $(L)_8$ y purificación de pFN18A sin inserto. 16

3.2.2.6. Ligación enzimática de L1 con pFN18A para propagación y purificación. 17

3.2.2.7. Secuenciación de resultados de purificados de propagación de pFN18A-L1..... 18

3.2.3. Ensayos de expresión del péptido antigénico L1 y selección de cepa bacteriana de expresión. 18

3.2.3.1. Transformación de cepas bacterianas de expresión con plásmido pFN18A-L1..... 18

3.2.3.2. Ensayos de expresión de L1 en BLR (DE3) pLysS pFN18A-L1 y C41 (DE3) pFN18A-L1. 19

3.2.3.3. Visualización de los ensayos de expresión proteica por electroforesis SDS-PAGE. 20

4. PRESENTACIÓN Y ANÁLISIS DE RESULTADOS 22

4.1. Identificación de proteínas de superficie abundantes en contexto de infección de *P. salmonis* por homología y selección de sus péptidos antigénicos 22

4.1.1. Selección de proteínas de superficie abundantes en contexto de infección de organismos filogenéticamente cercanos a *P. salmonis* mediante búsqueda bibliográfica. 22

4.1.1.1. *Piscirickettsia salmonis* 22

4.1.1.2. *Coxiella burnetii* 22

4.1.1.3. *Legionella pneumophila* 23

4.1.1.4. *Pseudomonas aeruginosa* 24

4.1.2. Selección de antígenos de superficie de *P. salmonis* por homología de proteínas abundantes en contexto de infección de organismos filogenéticamente cercanos 26

4.1.2.1. Búsqueda de proteínas abundantes en contexto de infección por homología en *P. salmonis* mediante alineamiento local de secuencias proteicas 26

4.1.2.2. Filtro de conservación aminoacídica de proteínas de *P. salmonis* para selección en base a especificidad de posterior péptido a expresar. 29

4.1.2.3. Modelamiento de proteínas candidatas 33

4.1.3. Selección de péptido antigénico de superficie de *P. salmonis* y, diseño y síntesis de vector de clonamiento contiendo la secuencia nucleotídica del péptido..... 38

4.2. Expresión de péptido antigénico de superficie L1 de *P. salmonis* mediante el subclonamiento de su secuencia nucleotídica en vector de expresión pFN18A 40

4.2.1. Subclonamiento del péptido antigénico L1 en vector de expresión pFN18A 40

4.2.2. Ensayos de expresión de Halo-péptido y selección cepa de expresión para péptido antigénico L1 48

5. DISCUSIÓN DE RESULTADOS 55

6. CONCLUSIONES 64

7. IMPLICANCIAS, RECOMENDACIONES Y/O PROYECCIONES FUTURAS..... 65

8. REFERENCIAS 68

ÍNDICE DE TABLAS

Tabla 1. Listado de bases de datos utilizadas..... 9

Tabla 2. Listado de programas y recursos bioinformáticos utilizados. 9

Tabla 3. Listado de plásmidos y cepas bacterianas a comparar utilizados. 9

Tabla 4. Listado de proteínas abundantes en contexto de infección bacteriano de organismos filogenéticamente cercanos a *P. salmonis* comparando la ubicación subcelular de cada proteína..... 25

Tabla 5. Resultados alineamiento local de secuencias proteicas por Blastp contra *P. salmonis* de las proteínas de superficie abundantes en contexto de infección de *C. burnetii*, *L. pneumophila* y *P. aeruginosa*..... 28

Tabla 6. Comparación estadística de los resultados de alineamientos locales mediante Blastp de las proteínas de superficie de *P. salmonis* propuestas como abundantes en contexto de infección por homología. 33

ÍNDICE DE FIGURAS

Figura 1. Diagrama de flujo para el objetivo específico 1. 12
Figura 2. Diagrama de flujo para el objetivo específico 2. 20
Figura 3. Modelo proteico predicho por SWISS-MODEL de la proteína barril-beta de membrana externa de *P. salmonis (*WP_016209892.1). 34
Figura 4. Modelo proteico predicho por SWISS-MODEL de la proteína de la familia OmpA de *P. salmonis* (WP_017376173.1). 35
Figura 5. Modelo proteico predicho por SWISS-MODEL de la proteína de membrana externa de la familia TolC de *P. salmonis* (WP_032126547.1), en oligo-estado de monómero..... 36
Figura 6. Modelo proteico predicho por SWISS-MODEL de la proteína de membrana externa de la familia TolC de *P. salmonis* (WP_032126547.1), en oligo-estado de homotrímero. 37
Figura 7. Segunda mejor coincidencia de modelo proteico predicho por SWISS-MODEL de la proteína barril-beta de membrana externa de *P. salmonis (*WP_016209892.1), visualizado mediante Mol* Viewer. 38
Figura 8. Mapa genético del vector de clonamiento pUC57 sintetizado conteniendo la secuencia nucleotídica del péptido L1. 39
Figura 9. Resultados electroforesis en gel de agarosa TAE de los productos de PCR de colonia de XL10 pUC57-L1 y de los productos de PCR de pUC57-L1B. 41
Figura 10. Mapa genético del vector de expresión pFN18A, conteniendo la secuencia codificante para HaloTag..... 42
Figura 11. Resultados electroforesis en gel de agarosa TAE 1,0% de los productos de las digestiones enzimáticas de pFN18A $(L)_8$ para obtención del vector de expresión sin inserto..... 44
Figura 12. Mapa genético del vector de expresión pFN18A-L1, vector de expresión del péptido antigénico L1. 45
Figura 13. Resultados electroforesis en gel de agarosa TAE de los productos de PCR de colonia de XL10 pFN18A-L1 y de los productos de PCR de pFN18A-L1..... 47
Figura 14. Resultados electroforesis en gel de agarosa TAE de la corroboración de integridad de pFN18A-L1 y pUC57-L1B. 48
Figura 15. Resultados SDS-PAGE del ensayo de expresión de BLR (DE3) pLysS pFN18A-L1 a 20°C y 37°C. 50
Figura 16. Resultados SDS-PAGE del ensayo de expresión de BLR (DE3) pLysS pFN18A-L1 a 20°C..... 51

Figura 17. Resultados SDS-PAGE del ensayo de expresión de BLR (DE3) pLysS pFN18A-Halo y de BLR (DE3) pLysS pFN18A-L1 a 37°C y 20°C. 52
Figura 18. Resultados SDS-PAGE del ensayo de expresión de C41 pFN18A-L1 a 20°C. 54

RESUMEN

El presente trabajo de tesis se enfocó en el desarrollo de un sistema innovador para la generación de péptidos antigénicos específicos de *Piscirickettsia salmonis*, agente etiológico del Síndrome Rickettsial Salmonídeo (SRS), una enfermedad que causa importantes pérdidas económicas en la industria salmonera chilena. Este patógeno afecta principalmente a los salmones del Atlántico (*Salmo salar*), siendo un factor crítico en la producción acuícola en Chile. El objetivo principal fue identificar y seleccionar un péptido antigénico específico de *P. salmonis* mediante la revisión bibliográfica y el diseño *in silico*, para la posterior utilización de este péptido en la generación de anticuerpos específicos para futuras aplicaciones biotecnológicas, como el diagnóstico, prevención y tratamiento de SRS.

Se aplicó una estrategia basada en la comparación de bases de datos y el uso de herramientas de biología computacional para encontrar el péptido más adecuado para el desarrollo de antígenos y su posterior uso en diagnóstico de SRS. Mediante el modelamiento proteico y alineamiento múltiple de secuencias de las proteínas de superficie abundantes en contexto de infección de *P. salmonis*, con énfasis en la búsqueda de las porciones expuestas al medio extracelular, fue posible encontrar un péptido que podría servir como marcador específico de *P. salmonis*, denominado como péptido L1 (Loop largo no estructurado número 1). La secuencia nucleotídica del péptido fue predicha y visualizada *in silico* integrada en un vector de clonamiento pUC57 para su posterior síntesis y subclonamiento en un vector de expresión pFN18A, resultado corroborado mediante la secuenciación de los productos de purificación de los ADNp obtenidos. Además, el péptido fue expresado utilizando sistemas heterólogos de expresión en *Escherichia coli* con cepas BLR (DE3) pLysS y C41 (DE3), aunque se identificaron desafíos relacionados con la concentración y la degradación proteica. A pesar de las dificultades, la versatilidad del sistema HaloTag, plataforma integrada en el vector de expresión, puede permitir futuras mejoras en la purificación del péptido. Este trabajo representa un paso importante hacia la generación de anticuerpos específicos contra *P. salmonis*, además, abre oportunidades para investigaciones y aplicaciones biotecnológicas en la salud animal, incluyendo el desarrollo de fármacos para tratar SRS, y para lograr diagnósticos más rápidos y accesibles.

SUMMARY

The present thesis work focused on the development of an innovative system for generating specific antigenic peptides of *Piscirickettsia salmonis*, the etiological agent of Salmonid Rickettsial Syndrome (SRS), a disease that causes significant economic losses in the Chilean salmon industry. This pathogen primarily affects Atlantic salmon (*Salmo salar*), playing a critical role in aquaculture production in Chile. The main objective was to identify and select a specific antigenic peptide through literature review and *in silico* design for the subsequent use of this peptide in the generation of specific antibodies for future biotechnological applications, such as the diagnosis, prevention, and treatment of SRS.

A strategy based on database comparison and the use of computational biology tools was applied to find the most suitable peptide for the development of antigens to address SRS. Through protein modeling and multiple sequence alignment of surface proteins in the context of *P. salmonis* infection, with an emphasis on the search for portions exposed to the extracellular environment, it was possible to find a peptide that could serve as a specific marker for *P. salmonis*, referred to as peptide L1 (Long Unstructured Loop number 1). The nucleotide sequence of the peptide was predicted and visualized *in silico*, integrated into a pUC57 cloning vector for subsequent synthesis and subcloning into a pFN18A expression vector, with the success of the process confirmed through sequencing of the purified DNA products. Additionally, the peptide was evaluated in a heterologous system in *Escherichia coli* with expression strains BLR (DE3) pLysS and C41 (DE3), although challenges related to protein concentration and degradation were identified. Despite the difficulties, the versatility of the HaloTag system, integrated into the expression vector, may allow for future improvements in peptide purification. This achievement marks a significant step towards generating specific antibodies against *P. salmonis*. Furthermore, it opens opportunities for research and biotechnological applications in animal health, including drug development for treating SRS, and for achieving faster and more accessible diagnostics.

1. INTRODUCCIÓN.

1.1. Salmonicultura chilena

La seguridad alimentaria es esencial para mejorar el estatus socioeconómico y combatir la malnutrición en los países subdesarrollados (1). La malnutrición en humanos resalta la importancia de las proteínas animales en la dieta regular. La pesca, en particular la acuicultura, desempeña un papel crucial al proporcionar proteínas altamente nutritivas y asequibles a nivel mundial. Estos recursos acuáticos son una valiosa fuente de vitaminas, proteínas, micronutrientes y minerales, especialmente para las personas con bajos recursos. La Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO) reconoce el potencial continuo de la acuicultura para promover el desarrollo económico y mejorar el bienestar humano (2). Durante las últimas décadas, en Chile, la industria acuícola ha experimentado un notable crecimiento, consolidándose como uno de los ejes centrales para el desarrollo económico del país (3).

Dentro de la acuicultura chilena, el cultivo y exportación de salmónidos es un pilar fundamental, en donde se han logrado exportaciones anuales de salmones valoradas en más de MM US$5.000 en los últimos años (años 2018, 2019 y 2021), y solo el 2020 las cifras de exportaciones anuales de salmones fueron valoradas en MM US$4.303, representando un 14,3% menos que el año anterior debido principalmente a los efectos de la pandemia por COVID-19 (4). Además, en términos de cosecha de salmones, el año 2020 se cosecharon sobre 1 millón de toneladas, siendo el valor más alto de cosecha desde el año 2007. La producción de salmones en Chile y los valores proporcionados están constituidos principalmente por las especies *Salmo salar* (salmón del Atlántico), *Oncorhynchus kisutch* (salmón coho) y *Oncorhynchus mykiss* (trucha arcoíris), siendo estas las principales especies cultivadas en la salmonicultura chilena (5). Dentro de las 1.043.144 toneladas de salmones cosechadas para el año 2020, el salmón del Atlántico constituye un 74,82% del total (780.475 toneladas), seguido por un 16,95% del salmón coho (176.776 toneladas) y un 8,23% de la trucha arcoíris (85.893 toneladas) (5). Mientras que, para el siguiente año, los centros con las especies salmón Atlántico aumentaron en un 7%, salmón coho aumentó en un 14% y la trucha arcoíris aumentó en un 13% (6).

El sector salmonicultor chileno está en constante trabajo para mejorar los indicadores sanitarios dentro de sus centros de cultivo, estos intentos de mejoras se realizan mediante la innovación y adopción de tecnologías que permitan incrementar el desempeño de los centros de cultivo. En 2020 la industria salmonera chilena pagó un total de MM CL$11.230 y realizó una inversión en Investigación y Desarrollo

(I+D) de MM CL$1.417, logrando una disminución de un 14,22% en la tasa de utilización de antibióticos en el total de la industria salmonera chilena para el año 2020 en comparación con el año anterior (298 g/ton de antibiótico en 2020 y 343 g/ton de antibiótico en 2019). Esto para lograr cumplir el acuerdo realizado por la ONG Monterey Bay Aquarium en 2019 de reducir en un 25% el uso de antibióticos en la acuicultura para el año 2050. Sin embargo, el desarrollo y éxito de la industria salmonera está directamente relacionado con el control de la mortalidad asociada a enfermedades derivadas del cultivo intensivo de peces (7). La aparición de enfermedades infecciosas es actualmente el problema más serio que debe enfrentar la acuicultura a nivel global (8). Durante el ciclo productivo, los peces están expuestos a diversos microorganismos, tales como virus, bacterias y parásitos, los que ocasionan brotes infecciosos y mortalidades (9,10). La susceptibilidad para ser infectados y enfermar aumenta bajo condiciones de cultivo intensivo, dada la presencia de estímulos que provocan alteraciones fisiológicas que gatillan estrés en los peces y una menor capacidad en la respuesta inmune protectora (11). En Chile, el principal problema que enfrenta la industria acuícola es el Síndrome Rickettsial Salmonídeo (SRS), causado por la bacteria Gram negativa *Piscirickettsia salmonis*. En la última década, *P. salmonis* es responsable de más del 80% de la mortalidad debida a causas infecciosas en salmón del Atlántico y trucha arcoíris (12–14).

1.2. Piscirickettsia salmonis

P. salmonis es una bacteria Gram negativa, predominantemente cocoide, aeróbica y no encapsulada, con un diámetro de aproximadamente 0,5 a 1,5 µm. Es el agente causante del Síndrome Ricketsial Salmonídeo (SRS) o Piscirickettsiosis, una enfermedad que afecta principalmente a peces, especialmente salmónidos (15). La infección por *P. salmonis* causa una septicemia crónica y sistémica en los peces afectados, lo que se manifiesta en una variedad de síntomas internos, incluyendo descoloración renal, esplenomegalia e hígado pálido con la formación de numerosos nódulos de aproximadamente 5 a 6 mm de diámetro (16,17).

Una característica fundamental de este patógeno es su capacidad para infectar, sobrevivir y replicarse principalmente dentro de vacuolas citoplasmáticas de macrófagos y leucocitos polimorfonucleares, sin inducir un efecto citopático característico. Al ser un patógeno intracelular, *P. salmonis* emplea diversos mecanismos de subversión para evadir las respuestas inmunitarias del hospedero. Estos mecanismos incluyen la evasión del reconocimiento inmunitario del hospedero y la modulación y supresión de sus respuestas inmunitarias propias. Su principal mecanismo de patogenicidad radica en la evasión de la

degradación fagolisosomal, lo que le permite mantenerse y multiplicarse dentro de las células del sistema inmunitario del hospedero (18). Estas características hacen de *P. salmonis* un patógeno altamente adaptado para interactuar con su hospedero y evadir las defensas inmunológicas, lo que representa un desafío significativo para el control y tratamiento de la piscirickettsiosis en la acuicultura. El entendimiento de los mecanismos de patogenicidad de esta bacteria es crucial para el desarrollo de estrategias efectivas para su control y para proteger la salud de las poblaciones de peces en ambientes acuáticos.

El control y profilaxis frente a *P. salmonis* se ha basado en el uso de vacunas y antibióticos, sin embargo, ambas estrategias no han resultado efectivas (19). Las vacunas solo inducen una protección a corto plazo y mayormente mediada por anticuerpos (19–21), mientras que los antibióticos, principalmente Florfenicol y Oxitetraciclina, han sido utilizados como una herramienta para mantener altas tasas de producción, lo que se traduce en un uso excesivo bajo condiciones de alta incertidumbre y baja efectividad (22). Los antibióticos suelen ser ineficientes frente a infecciones bacterianas intracelulares, dada la escasa biodisponibilidad intracelular del fármaco, generalmente subterapéutica, que favorece la sobrevida de los patógenos y promueve el desarrollo de resistencias (23–26), generando un impacto sobre los ecosistemas marinos y representando una seria amenaza para la salud pública a nivel global (22).

Para diagnosticar SRS se han implementado diferentes métodos de diagnóstico, entre ellos se cuentan cultivos bacterianos (27–29), tinciones (30), PCR de punto final (31), qPCR (32), PCR-RFLP (33) y técnicas moleculares basadas en anticuerpos (34). Sin embargo, todas estas estrategias requieren equipamiento complejo y personal calificado no pudiendo ser realizadas en los mismos centros de cultivo. En ocasiones, entre la toma de muestra y la confirmación de diagnóstico pueden pasar más de 48 horas, dada la ubicación remota y aislada de muchos centros de cultivo. Por otro lado, los protocolos de toma de muestra en animales acuáticos establecidos por el Servicio Nacional de Pesca y Acuicultura del Gobierno de Chile (Sernapesca) señala que se debe privilegiar la extracción de peces orillados, moribundos, desorientados o mortalidad fresca del día (35). Lo anterior implica un muestreo cuando ya el brote se encuentra en curso y los resultados son obtenidos cuando ya no existe la posibilidad de una toma de decisiones y una respuesta apropiada, que permitan contener tempranamente los brotes infecciosos, llevando a un uso inadecuado y excesivo de antibióticos, como única alternativa para enfrentar los brotes de SRS. Por lo tanto, es clara la necesidad de implementar una estrategia de diagnóstico que pueda ser realizada rutinariamente, de bajo costo, de fácil manipulación y que permita

una toma de decisiones y respuesta temprana, cuando aún es posible contener los brotes, evitando las pérdidas económicas asociadas.

1.3. Generación de anticuerpos y antígenos

Los anticuerpos desempeñan un papel fundamental como herramienta molecular en el reconocimiento de proteínas y su uso está ampliamente extendido en la comunidad biomédica para diversos fines, como la investigación básica, diagnóstico clínico e, inclusive, terapias. Una de las características clave de los anticuerpos es su capacidad para desarrollar una alta especificidad hacia su antígeno proteico objetivo (36). Son altamente diversos en secuencia y tienen la capacidad de reconocer una amplia variedad de antígenos extraños, incluyendo proteínas, péptidos y combinaciones de estos (37).

En la actualidad, todos los anticuerpos y macromoléculas derivadas de anticuerpos que se están desarrollando para una amplia gama de aplicaciones terapéuticas requieren del diseño de antígenos y de la ingeniería de proteínas. Los enfoques de ingeniería se basan en nuestro conocimiento de la estructura de las proteínas y, especialmente, en cómo estas estructuras están relacionadas con su función. Las diferencias en la estructura de las proteínas pueden tener un impacto significativo en su capacidad de interactuar con las células y los receptores clave que desencadenan una respuesta inmune (38), como también en el desarrollo de anticuerpos específicos contra estos antígenos.

Es importante destacar que la actividad biológica de un péptido está estrechamente ligada a su conformación, es decir, la orientación espacial de sus grupos funcionales. Los péptidos pueden adoptar diferentes estructuras secundarias que se estabilizan mediante enlaces de hidrógeno, interacciones electrostáticas e hidrofóbicas, enlaces disulfuro y/o ciclización (39). Por lo tanto, el diseño de antígenos se basa en los motivos estructurales de la proteína nativa con el objetivo de generar anticuerpos específicos para el patógeno (40).

Una de las principales aplicaciones del diseño de antígenos es su utilidad en el desarrollo de estrategias de diagnóstico, por ejemplo, frente a infecciones, donde es esencial detectar la presencia de anticuerpos específicos del patógeno en los fluidos de los pacientes infectados, como el suero sanguíneo. Por lo tanto, existe un gran interés en desarrollar métodos confiables para identificar rápidamente posibles antígenos específicos (41). Actualmente, la identificación y selección de péptidos antigénicos se basan en la construcción de bibliotecas de péptidos del patógeno involucrados en la infectividad y transmisión. Esto se logra mediante la búsqueda de secuencias de ADN que codifican

para estos péptidos en bases de datos establecidas, como la del *National Center for Biotechnology Information* (NCBI) (42,43).

En el caso de las bacterias, sus proteínas de superficie desempeñan un papel crucial en diversas interacciones con el entorno, como la comunicación entre células, el transporte de nutrientes y la resistencia a los antibióticos. Además, estas proteínas son fundamentales durante el ciclo infectivo, ya que están involucradas en la defensa contra las respuestas del hospedero y en la inducción de toxicidad. Por lo tanto, algunas de estas proteínas de superficie se consideran candidatas prometedoras para el desarrollo de tratamientos antimicrobianos y técnicas de diagnóstico contra organismos patógenos (44).

Un ejemplo de tratamientos contra organismos patógenos es la generación de vacunas, en donde existen dos enfoques principales para identificar proteínas de superficie bacteriana para su desarrollo. Uno de ellos es la vacunología inversa, que implica el clonamiento y expresión de todas las proteínas codificadas en la secuencia del genoma de un organismo que se predice *in silico* que están expuestas en la superficie o que son secretadas. Luego, cada proteína se somete a un cribado mediante inmunización de alto rendimiento para evaluar su capacidad de generar anticuerpos en una especie animal definida que puedan neutralizar al organismo objetivo (44,45). Sin embargo, este método tarda mucho tiempo en su realización y requiere análisis genómicos complejos. Como alternativa, se han identificado proteínas de superficie bacteriana utilizando técnicas de proteómica. Sin embargo, los pasos requeridos para la separación de proteínas de superficie y el tratamiento previo de péptidos pueden ser laboriosos en términos de tiempo y recursos, además, la presencia de proteínas o péptidos citoplasmáticos puede afectar la reproducibilidad (44).

Por otro lado, la técnica del despliegue de fagos es una poderosa herramienta molecular para identificar rápidamente mimotopos de patógenos a partir de anticuerpos antibacterianos. Los antígenos correspondientes pueden ser identificados mediante análisis de homología de los mimotopos que abarcan todo el genoma bacteriano. Sin embargo, en la práctica es difícil identificar antígenos basándose únicamente en mimotopos debido a la gran cantidad de proteínas homólogas que pueden ser obtenidas mediante herramientas de análisis de secuencias, como el análisis de homología de secuencias de entrada cortas utilizando la Herramienta de Búsqueda de Alineamiento Local Básico (BLAST)(46). Cabe destacar que las proteínas de superficie bacteriana representan solo una fracción de todas las proteínas presentes en las bacterias y pueden ser identificadas mediante la búsqueda en bases

de datos que indican su localización subcelular (44), como, por ejemplo, la base de datos de UniProtKB (47).

Por otro lado, dada la cantidad de proteínas identificadas de *P. salmonis* que no han sido caracterizadas y/o tienen funciones desconocidas en las bases de datos disponibles (48), es un gran desafío establecer proteínas candidatas que participen en roles de patogenicidad de este organismo. Es por esto, que la búsqueda de proteínas homólogas por similitud de secuencias en organismos filogenéticamente cercanos a *P. salmonis*, puede ser una estrategia para encontrar posibles candidatos de proteínas no descritas que pueden ser antígenos nativos, sin la necesidad del uso de procedimientos y equipamiento costoso, cómo el caso de la técnica de despliegue de fagos. La homología es un concepto esencial en el análisis computacional de secuencias de proteínas y ADN, que se refiere a la ascendencia evolutiva común. Se infiere homología cuando dos secuencias o estructuras comparten más similitudes de las esperadas por azar, lo que sugiere que surgieron de un ancestro común en lugar de forma independiente. Esta relación de similitud excesiva se explica mediante la ascendencia común, lo que implica que las secuencias son homólogas (49). Por lo tanto, al encontrar, mediante una revisión exhaustiva de la bibliografía vigente, una proteína abundante en contexto de infección de un organismo filogenéticamente cercano a *P. salmonis* que sea candidata como potencial antígeno de superficie se puede buscar por alineamiento de secuencias las posibles proteínas homólogas dentro de *P. salmonis* para extraer un péptido antigénico adecuado para la generación de futuras estrategias para el diagnóstico o tratamiento de SRS.

En resumen, la infección por *P. salmonis*, también conocida como SRS, representa un desafío importante para la acuicultura de salmónidos debido a las pérdidas económicas y el impacto en el bienestar animal que conlleva. En este contexto, surge la necesidad de desarrollar nuevas estrategias de diagnóstico y tratamientos eficaces para combatir esta enfermedad.

La presente investigación se enfoca en el desarrollo de un diseño y expresión de péptidos antigénicos abundantes en la superficie de *P. salmonis* durante la infección. Para lograr este objetivo, se empleará una combinación de revisión bibliográfica y diseño *in silico* para identificar y seleccionar los péptidos más adecuados para su posterior expresión y evaluación. La pregunta de investigación que guía este estudio es: ¿Es posible identificar, seleccionar y expresar péptidos antigénicos de *P. salmonis* que sean abundantes en su superficie durante la infección, utilizando una estrategia basada en la revisión bibliográfica y el diseño *in silico*?

Además, el diseño racional de fármacos representa una estrategia innovadora en el desarrollo de compuestos terapéuticos, aprovechando el conocimiento profundo del sistema en estudio para diseñar específicamente moléculas con propiedades farmacológicas deseadas. Sus ventajas incluyen el uso de avances tecnológicos, como la informática y técnicas experimentales avanzadas, lo que permite un enfoque más preciso y eficiente en la búsqueda de nuevos fármacos (50). En este trabajo de tesis, el diseño racional será fundamental debido a su capacidad para optimizar la generación de péptidos antigénicos específicos de *P. salmonis*. El entendimiento detallado de las interacciones moleculares y el uso de recursos informáticos avanzadas permite la selección más precisa de péptidos con potencial para generar anticuerpos específicos, agilizando así el desarrollo nuevas herramientas para el diagnóstico y la prevención de SRS en la salmonicultura. Esto contribuiría a mejorar la salud de las poblaciones de peces y a fortalecer la sostenibilidad de la industria acuícola.

2. HIPÓTESIS Y OBJETIVOS

2.1. Hipótesis científica

Péptidos inmunogénicos de superficie en contexto de infección de *P. salmonis* pueden ser identificados mediante revisión bibliográfica, diseñados *in silico* y expresadas en *E. coli*, como estrategia para el futuro desarrollo de anticuerpos específicos para aplicaciones biotecnológicas.

2.2. Objetivo general

Generar un péptido inmunogénico de *P. salmonis* mediante un diseño racional que le permita ser un candidato para el desarrollo de anticuerpos específicos.

2.3. Objetivos específicos

2.3.1. Identificar y seleccionar un péptido antigénico específico de *P. salmonis* mediante el análisis comparativo de bases de datos y herramientas de biología computacional.

2.3.2. Generar el péptido inmunogénico de superficie de *P. salmonis* mediante un sistema de expresión heterólogo en *E. coli* para su preparación para ser utilizado para el desarrollo de anticuerpos específicos.

3. MATERIALES Y MÉTODOS

3.1. Materiales.

3.1.1. Recursos bioinformáticos utilizados.

Tabla 1. Listado de bases de datos utilizadas.

Nombre	Identificador	Referencia
NCBI	https://www.ncbi.nlm.nih.gov/	(43,51–66)
UniProtKB	https://www.uniprot.org/	(47)
RefSeq de NCBI	https://www.ncbi.nlm.nih.gov/refseq/	(67)

Tabla 2. Listado de programas y recursos bioinformáticos utilizados.

Nombre	Identificador	Referencia
BLAST	https://blast.ncbi.nlm.nih.gov/Blast.cgi	(46)
Clustal Omega	https://www.ebi.ac.uk/Tools/msa/clustalo/	(68)
Swiss-Model	https://swissmodel.expasy.org/	(69–73)
Mol* Viewer	https://molstar.org/viewer/	(74)
SnapGene	https://www.snapgene.com/	SnapGene® software (de Dotmatics; disponible en snapgene.com)

3.1.2. Plásmidos y cepas bacterianas utilizados.

Tabla 3. Listado de plásmidos y cepas bacterianas a comparar utilizados.

Plásmidos		
Nombre	**Descripción**	**Identificador**
pUC57	Vector de clonamiento	https://www.ncbi.nlm.nih.gov/nuccore/Y14837.1
pFN18A	Vector de expresión	https://www.ncbi.nlm.nih.gov/nuccore/186470146
Cepas de *E. coli*		
Nombre	**Descripción**	**Referencia**
XL10	Cepa de clonamiento	(75)
BLR (DE3) pLysS	Cepa de expresión	(76)
C41 (DE3)	Cepa de expresión	(77)

3.2. Métodos.

3.2.1. Selección péptido antigénico para construcción in silico de vector de clonamiento.

3.2.1.1. Búsqueda bibliográfica de proteínas de superficie abundantes en contexto de infección de organismos filogenéticamente cercanos a *P. salmonis.*

Con el objetivo de buscar y seleccionar proteínas de superficie abundantes en contexto de infección de organismos filogenéticamente cercanos a *P. salmonis*, se realizó una revisión bibliográfica en la base de datos de NCBI (78). En esta búsqueda fueron consideradas especies bacterianas cuya similitud de secuencia del ADNr 16S con *P. salmonis* les permite ser consideradas filogenéticamente cercanas, en base al estudio de Mauel M. J. *et al*. (1999), entre ellas encontramos a *Coxiella burnetii*, *Legionella pneumophila*, *Pseudomonas aeruginosa*, *Wolbachia persica* y *Francisella tularensis* (53). También se utilizó la secuencia del genoma de la cepa de referencia LF-89 de *P. salmonis* (ATCC VR-1361) para obtener información sobre sus proteínas (51). Además, se revisó la literatura sobre *C. burnetii*, *L. pneumophila* y *P. aeruginosa*, identificándose proteínas seroactivas, es decir las que están presentes en muestras de sueros sanguíneos de animales infectados, y otras proteínas que pueden estar implicadas en la patogenia de estas bacterias. Se prestaron especial atención a las proteínas relacionadas con procesos de invasión y supervivencia intracelular, la motilidad celular y el tráfico intracelular. Para la elección de las proteínas candidatas a comparar en el posterior análisis se utilizaron las referencias mencionadas en la Tabla 1 de la base de datos de NCBI, las que fueron principalmente seleccionadas en base a ser abundantes en contexto de infección del organismo, según lo indicado en los estudios revisados (Figura 1).

3.2.1.2. Identificación de las proteínas homólogas en *P. salmonis* y filtro por conservación entre especies.

A cada una de las proteínas seleccionadas mediante la revisión bibliográfica, se le hizo una identificación de sus secuencias aminoacídicas y de su ubicación subcelular en el organismo correspondiente, a través de la búsqueda en la base de datos UniProtKB (47) y de RefSeq de NCBI (67). Luego, se realizó un filtro por su ubicación (Figura 1), seleccionando las que estuvieran localizadas en la membrana celular externa de cada organismo. Estas candidatas fueron sometidas a alineamiento local de secuencias mediante la herramienta bioinformática de BLAST (46) contra las proteínas ya descritas de *Piscirickettsia salmonis* ATCC VR-1361 (taxid:1227812) con la base de datos de secuencias

de proteínas no-redundantes (nr), con el objetivo de encontrar e identificar las posibles homólogas en *P. salmonis* en base a la similitud de secuencias y cobertura de los alineamientos obtenidos (Figura 1), seleccionando la mejor coincidencia para cada una de las sometidas al análisis de acuerdo a los criterios mencionados.

Al tener identificadas las posibles proteínas homólogas de *P. salmonis* de aquellas de superficie candidatas abundantes en contexto de infección de organismos filogenéticamente cercanos, se realizó un segundo filtro en base a conservación (Figura 1) entre otros organismos mediante alineamientos locales de secuencia de cada una por BLAST (46) con la base de datos de secuencias de proteínas nr, contra los organismos filogenéticamente cercanos a *P. salmonis*, por un lado y, por otro lado, contra toda la base de datos ofrecida por BLAST excluyendo al filo *Proteobacteria* (taxid:1224), considerándose como organismos de cercanía genética y que habita en entornos similares a *P. salmonis*, por lo tanto, utilizándose con el objetivo de buscar proteínas de organismos que no fuesen cercanos al organismo.

El filtro por conservación se realizó mediante un análisis estadístico de los resultados obtenidos en Blastp a través del recurso informático Microsoft Excel, obteniendo los valores de media, mediana, valor máximo y desviación estándar de a lo más 100 coincidencias para cada proteína analizada. Los criterios para definir la conservación idónea de las proteínas utilizados fueron tener una media de identidad de secuencia menor a 35% con un valor máximo menor a 45% de identidad para lo obtenido en ambos Blastp realizados a cada una de ellas.

3.2.1.3. Modelamiento de las proteínas seleccionados post-filtro por conservación e identificación de regiones expuestas al medio extracelular.

A las proteínas candidatas seleccionadas posterior al filtro por conservación, se les realizaron alineamientos múltiples de secuencias mediante la herramienta Clustal Omega (68), contra proteínas con similitudes altas encontradas en los análisis anteriores. Además, se modeló cada una de las candidatas mediante la herramienta bioinformática SWISS-MODEL (69–73) (Figura 1), para el análisis de los dominios de cada proteína y su comparación con las regiones conservadas en las secuencias obtenidas por el alineamiento, para así seleccionar la candidata más idónea para el posterior análisis, en base a sus dominios conservados y expuestos al medio extracelular por el organismo.

3.2.1.4. Selección péptido antigénico de *P. salmonis*, construcción y síntesis de pUC57 conteniendo la secuencia nucleotídica del péptido seleccionado.

La proteína seleccionada se visualizó mediante el programa Mol* Viewer (74), para la identificación de los péptidos antigénicos expuestos al medio (Figura 1) y la extracción especifica de las secuencias proteicas de estos para su incorporación *in silico* en un vector de clonamiento pUC57, utilizando el software SnapGene Viewer de Dotmatics (Figura 2). Finalmente, se sintetizó el plásmido resultante (pUC57-L1) mediante la externalización del servicio al Centro de Síntesis Genético de GenScript Biotech Corporation (Figura 2).

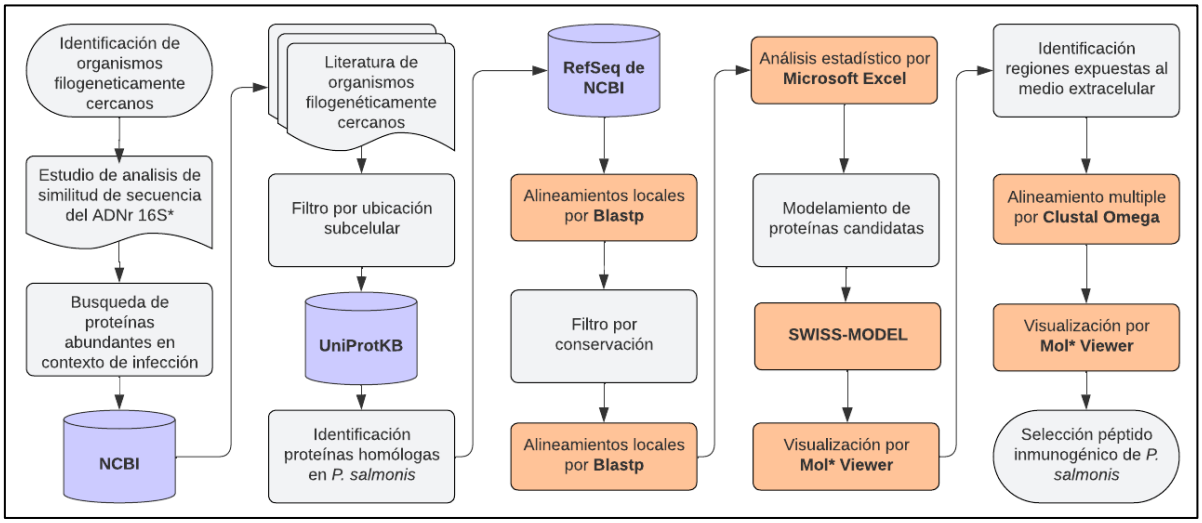

Figura 1. Diagrama de flujo para el objetivo específico 1.
Pasos metodológicos utilizados para el desarrollo del objetivo número 1. Procesos de color azul representan bases de datos utilizadas y en color naranjo los recursos bioinformáticos utilizados.
*****: Para identificar los organismos filogenéticamente cercanos a *P. salmonis*, se utilizó el estudio de Mauel M. J. et al. (1999) (53).

3.2.2. Subclonamiento del péptido antigénico L1 desde vector de clonamiento pUC57 a vector de expresión pFN18A.

3.2.2.1. Ensayo de diluciones de ADNp para estandarización en visualización por electroforesis en gel de agarosa TAE.

Se realizaron ensayos de concentración idónea de ADNp para su visualización por electroforesis en gel de agarosa TAE utilizando como ADN el plásmido pFN18A $(L)_8$. Se hicieron diluciones seriadas de

pFN18A $(L)_8$ con 0 ng, 157 ng, 314 ng, 628 ng, 942 ng, 1.256 ng, 1.570 ng y 2.198 ng de plásmido, en un primer ensayo. Luego, en un segundo ensayo se utilizaron 0 ng, 79 ng, 157 ng, 236 ng, 314 ng, 471 ng, 628 ng, 785 ng y 942 ng de ANDp. Las diluciones fueron mezcladas con solución colorante de carga de gel de ADN para su posterior visualización con luz UV mediante un transiluminador UV con sistema de documentación de geles, luego de una electroforesis en gel de agarosa TAE 1,0%, utilizando como marcador de peso molecular el Ladder AccuRuler 1Kb DNA RTU de Maestrogen, con un programa de 80 minutos a 80 V.

3.2.2.2. Preparación de células *E. coli* XL10, BLR (DE3) pLysS y C41 (DE3) quimiocompetentes.

Se crecieron pre-inóculos de *E. coli* XL10, BLR (DE3) pLysS y C41 (DE3) en 5 mL de medio Luria-Bertani (LB) más antibiótico de selección para cada cepa (100 µg/mL de ampicilina, 30 µg/mL de tetraciclina y 100 µg/mL de ampicilina, respectivamente) a 37°C en tubos cónicos de centrifuga de 15 mL durante 18 horas. Luego, se diluyeron los cultivo en razón 1:100 en 20 mL de caldo LB más antibiótico de selección para cada cepa a 37°C en tubos cónicos de centrifuga de 50 mL hasta llegar a una densidad óptica a 600 nm (OD_{600}) de 0,7 determinada por espectrofotometría en un lector de microplacas Tecan Nanoquant Infinite M200 Pro. Después, se enfriaron los cultivos en hielo durante 10 min y se centrifugaron las células a 10.015 x g por 10 min a 4°C, para descartar el sobrenadante y resuspender el pellet en 12,5 mL de $CaCl_2$ 0,1 M frio y estéril. Finalizando con una nueva centrifugación a 10.015 x g por 10 min a 4°C, para descartar nuevamente el sobrenadante y resuspender el pellet en 5 mL de $CaCl_2$ 0,1 M y glicerol al 10%. Las células quimiocompetentes se almacenaron a -80°C para el uso a posterior.

3.2.2.3. Propagación de XL10 pUC57-L1 y XL10 pFN18A-10, seguido de purificación de pUC57-L1B y pFN18A-L1.

Para la transformación de las células bacterianas con el vector de clonamiento, se agregaron 2 µL del vector pUC57-L1 sintetizado por Genscript a una concentración de 20 ng/µL a 100 µL de células *E. coli* XL10 quimiocompetentes mezclando con golpes suaves al tubo de microcentrífuga. Se incubaron las células con el ADNp en hielo por 30 minutos y se les dio un shock térmico por 45 segundos a 42°C en un sistema de calentamiento en seco para bloques metálicos (Thermo-block), inmediatamente después, se transfirieron los tubos a hielo para dejar recuperando a las células del shock durante 15 minutos. Se agregaron 250 µL de medio LB a temperatura ambiente y se incubó durante 1 hora a 37°C. Luego de la incubación, se realizó un cultivo de las células transformadas en placas de agar-LB con 100 µg/mL de ampicilina, utilizando 50 µL del caldo de células concentradas por centrifugación a 726 x g

por 8 minutos y a temperatura ambiente, las cuales fueron sembradas en césped. Las placas con células transformadas se incubaron por 18 horas a 37°C.

Para corroborar la presencia del vector pUC57-L1 en las células transformadas se seleccionaron 5 colonias de cada placa y se realizó una reacción en cadena de la polimerasa (PCR) de punto final de colonias utilizando 10 µM de los partidores pUC/M13 fwd (5'-GGTTTTCCCAGTCACGAC-3') y pUC/M13 rev (5'-CAGGAAACAGCTATGAC-3') flanqueantes al inserto de L1, utilizando la GoTaq G2 Green Master Mix de Promega, en volúmenes de reacción de 25 µL y con los siguientes parámetros de ciclo: 13 minutos a 95°C; seguido por 35 ciclos de 30 segundos a 95°C, 6 segundos a 49°C y 30 segundos a 72°C; finalizando con 5 minutos a 72°C y tiempo ilimitado a 4°C previa utilización de los productos de PCR. Luego, estos productos se sometieron a electroforesis en gel de agarosa TAE 2,0%, con un programa de 90 V por 50 minutos y utilizando de marcador molecular el PCRBIO Ladder III de PCR Biosystems, siendo visualizados con luz UV mediante un transiluminador UV con sistema de documentación de geles.

Cada colonia aislada seleccionada de cada placa se respaldó en placas agar-LB con ampicilina 100 µg/mL y se transfirieron, al mismo tiempo, a tubos cónicos de centrifuga de 50 mL con 15 mL de caldo LB y ampicilina 100 µg/mL, los cuales fueron cultivados por 16 horas a 37°C. Se centrifugaron los cultivos a 4°C por 10 minutos a 10.015 x g, rescatando el pellet y desechando el sobrenadante. El ADNp se purificó utilizando el kit Wizard Plus SV Minipreps DNA purification system. Brevemente, el pellet de células *E. coli* XL10 pUC57-L1 se resuspendió en 250 µL de solución de resuspensión. Luego, se agregaron 250 µL de solución de lisis, homogenizando 4 veces por inversión, seguido de la adición de 10 µL de proteasa alcalina, mezclado por inversión 4 veces e incubando la reacción de la proteasa a temperatura ambiente durante 5 minutos. Se agregaron 350 µL de solución de neutralización y se homogenizó por inversión 4 veces más. Se centrifugaron las soluciones a temperatura ambiente durante 10 minutos a 15.805 x g. Se traspasó el sobrenadante resultante de la centrifugación a columnas de recolección y se centrifugó nuevamente a temperatura ambiente durante 1 minuto a 15.805 x g. Finalmente, se le realizó un lavado a la columna de recolección con 250 µL de solución de lavado y centrifugó a temperatura ambiente durante 2 minutos a 15.805 x g. Se transfirió la columna de recolección conteniendo el ADNp a un tubo de microcentrífuga estéril de 1,5 mL y se agregaron 50 µL de agua libre de nucleasas precalentada a 65°C en un Thermo-block, centrifugando los tubos a temperatura ambiente durante 1 min a 15.805 x g y cuantificando el resultado de la purificación a 260/280 nm mediante un lector de microplacas Tecan

Nanoquant Infinite M200 Pro, utilizando como blanco de medición la misma agua libre de nucleasas precalentada utilizada en la purificación. Los ADNp purificados se almacenaron a -20°C.

Para el caso de la propagación de XL10 pFN18A-L1 y purificación de pFN18A-L1, se realizó el mismo proceso detallado anteriormente pero la transformación se hizo con 5 µL del producto de la ligación enzimática de pFN18A $(L)_8$ con el inserto contenedor de la secuencia del péptido L1 (véase sección 3.2.2.6) y, para la revisión de la integridad de la secuencia de L1 mediante PCR de punto final de colonias y PCR de punto final (posterior a la purificación del ADNp), se utilizaron 10 µM de los partidores T7 prom (5'-TAATACGACTCACTATAGGG-3') y T7 ter (5'-GCTAGTTATTGCTCAGCGG-3'), utilizando la GoTaq G2 Green Master Mix de Promega, en volúmenes de reacción de 25 µL y con los siguientes parámetros de ciclo: 13 minutos a 95°C; seguido por 35 ciclos de 30 segundos a 95°C, 6 segundos a 52°C y 30 segundos a 72°C; finalizando con 5 minutos a 72°C y tiempo ilimitado a 4°C previo utilización de los productos de PCR. Además, los productos del PCR de punto final de colonia se sometieron a electroforesis en gel de agarosa TAE 2,0%, con un programa de 90 V por 60 minutos y utilizando de marcador molecular el PCRBIO Ladder III de PCR Biosystems. Mientras que, los productos de PCR de punto final post-purificación de ADNp se sometieron a electroforesis en gel de agarosa TAE 1,0%, con un programa de 80 V por 80 minutos y utilizando de marcador molecular el Ladder AccuRuler 1Kb DNA RTU de Maestrogen. Siendo ambos geles de electroforesis visualizados con luz UV mediante un transiluminador UV con sistema de documentación de geles.

3.2.2.4. Amplificación, digestión y purificación de L1 desde pUC57-L1B.

Se realizó por duplicado un PCR de punto final a pUC57-L1B con la utilización de 10 µM de los partidores pUC/M13 fwd (5'-GGTTTTCCCAGTCACGAC-3') y pUC/M13 rev (5'-CAGGAAACAGCTATGAC-3'), utilizando la GoTaq G2 Green Master Mix de Promega, en volúmenes de reacción de 25 µL y con los siguientes parámetros de ciclo: 13 minutos a 95°C; seguido por 35 ciclos de 30 segundos a 95°C, 6 segundos a 49°C y 30 segundos a 72°C; finalizando con 5 minutos a 72°C y tiempo ilimitado a 4°C previo utilización de los productos de PCR. Visualizando los resultados de la amplificación por electroforesis en gel de agarosa TAE 2,0% con el PCRBIO Ladder III como marcador de peso molecular, en un transiluminador UV con sistema de documentación de geles con luz azul para no dañar el ADN. Se cortaron las bandas del gel correspondientes al producto de PCR con un bisturí y se disolvió en solución de unión de membrana en una razón de 1 mg de gel por 1 µL de solución a 65°C en un Thermo-block dentro de un tubo de microcentrífuga de 1,5 mL, luego se siguió el siguiente protocolo de purificación

de ADN mediante el kit Wizard SV gel and PCR Clean-Up System de Promega: se transfirió la mezcla de gel disuelto con el producto de PCR a una minicolumna SV ensamblada en una columna de recolección y se incubó durante 1 min a temperatura ambiente, luego se centrifugó a 15.805 x g durante 1 min, descartando el fluido acumulado en la columna de recolección; posteriormente, se realizó un lavado de ADN unido a la minicolumna mediante la adición de 700 µl de solución de lavado de membrana y la centrifugación a 15.805 x g durante 1 min, desechando el fluido recolectado, seguido de un segundo lavado pero con 500 µL de solución de lavado de membrana y centrifugación a 15.805 x g durante 5 min, para una última centrifugación sin fluido a 15.805 x g durante 1 min; finalmente, se realizó una elución de ADN transfiriendo la minicolumna con el ADN unido a un tubo de microcentrífuga de 1,5 mL, en donde se incubo por 1 min a temperatura ambiente con 50 µL de H_2O libre de nucleasas calentada previamente a 65°C, para una última centrifugación a 15.805 x g durante 1 min, obteniendo el ADN purificado.

El ADN purificado mediante el protocolo mencionado anteriormente se cuantificó mediante un lector de microplacas Tecan Nanoquant Infinite M200 Pro y se sometió a una primera digestión enzimática simple durante 18 horas a 37°C de ~500 ng del producto de PCR purificado utilizando 10 U de *Bgl*II de NEB en buffer de digestión NEBuffer™ 3.1 y un volumen de reacción de 30 µL. Se purificaron los productos de la digestión mediante el protocolo Wizard SV gel and PCR Clean-Up System, enriqueciendo la columna de recolección con ambos productos de digestión. Al producto de la purificación se le realizó una segunda digestión enzimática simple durante 18 horas a 37°C de 25 µL del volumen obtenido en la purificación, utilizando 10 U de *Kpn*I de NEB en buffer de digestión NEBuffer™ r2.1 y un volumen de reacción de 30 µL. Finalmente, se purificaron los productos de la digestión mediante el protocolo Wizard SV gel and PCR Clean-Up System enriqueciendo la columna de recolección con ambos productos de digestión, se cuantificó el ADN obtenido mediante un lector de microplacas Tecan Nanoquant Infinite M200 Pro y se almacenó a -20°C para la posterior ligación enzimática con pFN18A lineal sin inserto.

3.2.2.5. Ensayos de digestión de pFN18A $(L)_8$ y purificación de pFN18A sin inserto.

Se realizaron 3 ensayos diferentes de digestión enzimática para pFN18A $(L)_8$, plásmido suministrado por el Laboratorio de Genómica Microbiana del Centro de Genómica y Bioinformática de la Universidad Mayor. El primero fue una digestión doble durante 18 horas a 37°C de 1.000 ng de pFN18A $(L)_8$ con 10 U de *BamH*I y 10 U de *Kpn*I con el NEBuffer™ r2.1, utilizando un volumen final de reacción de 20 µL

enraizado con H_2O libre de nucleasas y comparando la digestión 1.000 ng de pFN18A $(L)_8$ sin digerir, como control negativo. El segundo ensayo se realizó con los mismos parámetros del primero, pero por cuadriplicado y digerido por 40 min a 37°C. Finalmente, el tercer ensayo se realizó utilizando 2 controles negativos de 1.000 ng y 2.000 ng de pFN18A $(L)_8$ sin digerir y se realizaron las siguientes digestiones: 1.000 ng pFN18A $(L)_8$ con 10 U *BamH*I y 10 U *Kpn*I, en NEBuffer™ r2.1 para un volumen final de reacción de 10 µL; 1.000 ng pFN18A $(L)_8$ con 10 U *BamH*I-HF y 10 U *Kpn*I-HF, en rCutSmart™ Buffer para un volumen final de reacción de 10 µL; 2.000 ng pFN18A $(L)_8$ con 10 U *BamH*I y 10 U *Kpn*I, en NEBuffer™ r2.1 para un volumen final de reacción de 20 µL; 2.000 ng pFN18A $(L)_8$ con 20 U *BamH*I y 20 U *Kpn*I, en NEBuffer™ r2.1 para un volumen final de reacción de 20 µL; y 2.000 ng pFN18A $(L)_8$ con 20 U *BamH*I-HF y 20 U *Kpn*I-HF, en rCutSmart™ Buffer para un volumen final de reacción de 20 µL. Todas las enzimas y buffer de digestión utilizados provinieron de Bio England BioLabs (NEB).

Los 3 ensayos de digestiones enzimáticas se sometieron a electroforesis en gel de agarosa TAE 1,0% independientes, programadas por 110 minutos a 80 V y utilizando el marcador de peso molecular Ladder AccuRuler 1Kb DNA RTU de Maestrogen. Los primeros dos ensayos se visualizaron y fotografiaron en un transiluminador UV con sistema de documentación de geles de una electroforesis en gel con luz UV, mientras que el último ensayo fue con luz azul, para no dañar la integridad del ADN. Las bandas de gel del último ensayo de digestión correspondientes a pFN18A $(L)_8$ sin inserto se cortaron con un bisturí y se disolvieron en solución de unión de membrana en una razón de 1 mg de gel por 1 µL de solución a 65°C en un Thermo-block dentro de un tubo de microcentrífuga de 1,5 mL, para luego purificar el ADN contenido en la mezcla mediante el protocolo de purificación Wizard SV gel and PCR Clean-Up System de Promega (véase sección 3.2.2.4). Se cuantificó la concentración del producto de la purificación a 260/280 nm mediante un lector de microplacas Tecan Nanoquant Infinite M200 Pro y se almacenó ADN a -20°C para su posterior uso en la ligación enzimática con inserto que codifica para la secuencia del péptido L1.

3.2.2.6. Ligación enzimática de L1 con pFN18A para propagación y purificación.

Se realizó una ligación enzimática durante 18 horas a temperatura ambiente del producto de la purificación del vector pFN18A $(L)_8$ sin inserto y del producto de la purificación del inserto que codifica para la secuencia del péptido L1, utilizando 50 ng del vector y 16 ng del inserto, en un buffer de ligación de ADN T4 de NEB con 1 µL de ADN ligasa T4 de NEB y un volumen de reacción final de 20 µL. La reacción de ligación se inactivó a 65°C por 10 min. Finalmente, se realizó la transformación de células

quimiocompetentes *E. coli* XL10 con 5 µL del producto de la ligación enzimática (véase sección 3.2.2.3), para la posterior propagación de colonias y purificación del constructo.

Después de la propagación de *E. coli* XL10 pFN18A-L1 y purificación de pFN18A-L1 (véase sección 3.2.2.3), se realizó una electroforesis en gel de agarosa TAE 1,0% con los resultados de la purificación de ADNp de pFN18A-L1 y de pUC57-L1, con el objetivo de comprobar la integridad de los plásmidos. La electroforesis se programó a 80 V durante 80 min y se utilizó como marcador de peso molecular el Ladder AccuRuler 1Kb DNA RTU de Maestrogen, siendo el resultado de la electroforesis visualizado y fotografiado con luz UV mediante un transiluminador UV con sistema de documentación de geles.

3.2.2.7. Secuenciación de resultados de purificados de propagación de pFN18A-L1.

Se realizó la secuenciación de los productos de la purificación de pFN18A-L1 utilizando el servicio externo del Centro de Secuenciación Automática de ADN de la Pontificia Universidad Católica de Chile (Figura 2), donde se utilizó la tecnología ABI PRISM 3100 para secuenciar 5 plásmidos pFN18A-L1 utilizando los partidores Halo fwd (5'-GGTCTGAATCTGCTGCAAGAAG-3') y T7 ter (5'-GCTAGTTATTGCTCAGCGG-3'). Se analizaron y compararon las secuencias nucleotídicas obtenidas de la secuenciación mediante la herramienta bioinformática SnapGene, realizando alineamiento múltiples de las secuencias con Clustal Omega proporcionado por SnapGene.

3.2.3. Ensayos de expresión del péptido antigénico L1 y selección de cepa bacteriana de expresión.

3.2.3.1. Transformación de cepas bacterianas de expresión con plásmido pFN18A-L1.

Para la transformación de las cepas de expresión con los vectores de expresión (Figura 2), se agregaron 2 µL del plásmido pFN18A-L1 o con 2 µL de plásmido pFN18A-Halo, en ambos casos se utilizó una concentración de 20 ng/µL agregados a 100 µL de cada célula (*E. coli* BLR (DE3) pLysS y C41 (DE3)) quimiocompetente mezclando con golpes suaves al tubo de microcentrífuga. Se incubaron las células con el ADNp en hielo por 30 minutos y se les dio un shock térmico por 45 segundos a 42°C, inmediatamente después, se transfirieron los tubos a hielo para dejar recuperando a las células del shock durante 15 minutos. Se agregaron 250 µL de medio LB a temperatura ambiente y se incubó durante 1 hora a 37°C. Luego de la incubación, se realizó un cultivo de las células transformadas en placas de agar-LB con 100 µg/mL de ampicilina, utilizando 50 µL del caldo de células concentradas por centrifugación a 736 x g por 8 minutos y a temperatura ambiente. Las placas con células transformadas

se incubaron por 18 horas a 37°C, obteniéndose 3 placas transformadas de: BLR (DE3) pLysS con pFN18A-L1, BLR (DE3) pLysS con pFN18A-Halo y C41 (DE3) con pFN18A-L1.

El vector de expresión pFN18A-Halo fue suministrado por el Laboratorio de Genómica Microbiana del Centro de Genómica y Bioinformática de la Universidad Mayor.

3.2.3.2. Ensayos de expresión de L1 en BLR (DE3) pLysS pFN18A-L1 y C41 (DE3) pFN18A-L1.

Para el ensayo de expresión de L1 en BLR (DE3) pLysS pFN18A-L1, se cultivaron 2 colonias de la placa de transformación de BLR (DE3) pLysS con pFN18A-L1 en 2 tubos cónicos de centrifuga de 15 mL diferentes para cada colonia con 5 mL de caldo LB y 30 µg/mL de tetraciclina durante 18 horas a 37°, para ser utilizados como pre-inóculos para los cultivos de expresión. De cada cultivo de pre-inoculo se extrajeron 450 µL para inocular dos tubos cónicos de centrifuga de 50 mL diferentes con 15 mL de caldo LB y 30 µg/mL de tetraciclina, para alcanzar un 3,0% de volumen de células del cultivo final. Cada cultivo se incubó a 37°C hasta llegar a una OD_{600} de 0,7 determinada por espectrofotometría en un lector de microplacas Tecan Nanoquant Infinite M200 Pro. A los 2 cultivos realizados se les extrajo un volumen de 1 mL al inicio de la incubación, el cual fue denominado testigo tiempo 0 (T0) y a uno de los dos cultivos por colonia se le indujo la expresión de proteínas con IPTG 1 mM, obteniéndose dos cultivos inducidos y dos sin inducir (SI), de los cuales uno de cada uno se dividió en dos tubos distintos para ser cultivados en condiciones diferentes, uno a 37°C y el otro a 20°C. Luego, con el cultivo bacteriano de cada tubo, se extrajeron de cada cultivo volúmenes de testigo de 1 mL a la hora de incubación (T1), a 3 horas de incubación(T2) y a 18 horas (T3) de cultivo.

A cada testigo extraído de los cultivos se les hizo un tratamiento de centrifugación a 16.642 x g por 5 min a 4°C, para descartar el sobrenadante y resuspender el pellet en 500 µL de PBS, repitiendo el proceso de lavado una vez más de la misma forma. Finalmente, a cada testigo se resuspendió en 500 µL de PBS y se les realizó una sonicación con 6 ciclos de 1 seg de ultrasonido y 3 seg de reposo con un sonicador ultrasónico a baja frecuencia (~40 kHz), se les realizó una centrifugación a 16.642 x g por 10 min a 4°C y se recuperó el sobrenadante para ser almacenados a 4°C previo visualización de los resultados mediante electroforesis SDS-PAGE. Los sobrenadantes recuperados finales son denominados fracción proteica de la expresión (FPE).

En el caso del ensayo de expresión de L1 en C41 (DE3) pFN18A-L1, se realizó el mismo proceso detallado anteriormente pero solo una colonia fue evaluada la expresión en crecimiento solo a 20°C. Mientras

que, para los controles de expresión de Halo en BLR (DE3) pLysS pFN18A-Halo, se realizó el mismo proceso que se hizo con BLR (DE3) pLysS pFN18A-L1, pero solo se tomaron volúmenes de testigo de 1 mL al tiempo 0 y a las 18 horas de cultivo (T0 y T3).

3.2.3.3. Visualización de los ensayos de expresión proteica por electroforesis SDS-PAGE.

A cada FPE para cada condición y según los requerimientos de cada gel de poliacrilamida que se realizó, se le extrajo 9 µL y se agregó en un tubo de microcentrífuga de 0,6 mL con 1 µL del ligando HaloTag Alexa Fluor 488 de Promega, siendo incubada la reacción de ligación de Halo Tag con Alexa Fluor durante 15 min a temperatura ambiente y en total oscuridad. Luego, se mezclaron las incubaciones con buffer de carga SB para SDS-PAGE y nuevamente se incubaron por 10 min a 95°C. Finalmente, las FPE tratadas se sometieron a electroforesis según las especificaciones de condiciones de cada gel de poliacrilamida preparado, por 150 minutos a 120 V, utilizando el marcador de peso molecular PageRuler™ Prestained Protein Ladder Plus de Thermo Scientific. Cada gel se visualizó y fotografiado con luz azul mediante un transiluminador UV con sistema de documentación de geles, para luego, ser teñidos con tratamiento de solución de azul de Coomassie durante 18 horas y posterior retiro del exceso de tinción con solución de destinción de 40% metanol y 10% ácido acético, ambos tratamientos con agitación en un agitador de bandeja, para así ser visualizados y fotografiados con luz blanca mediante un transiluminador UV con sistema de documentación de geles (Figura 2).

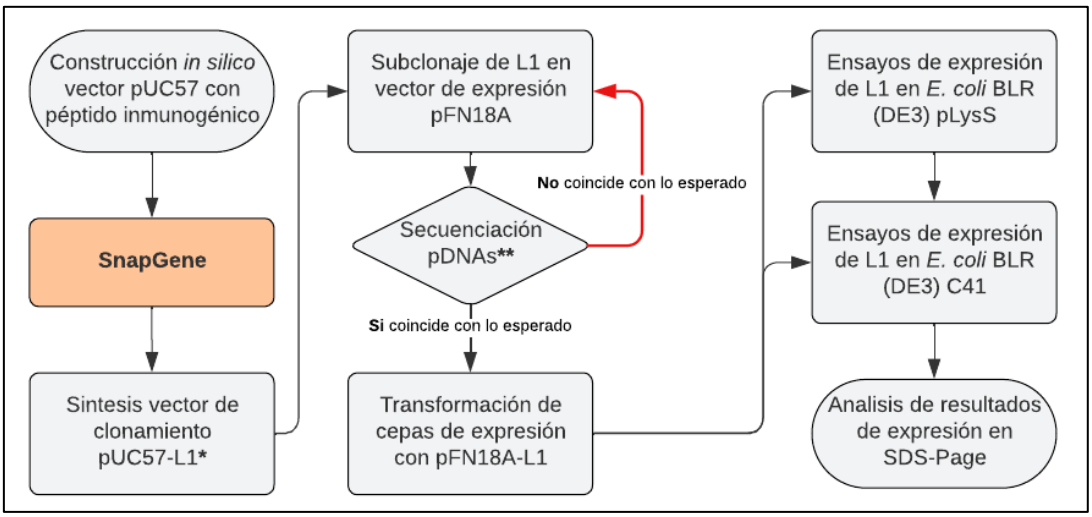

Figura 2. Diagrama de flujo para el objetivo específico 2.
Pasos metodológicos utilizados para el desarrollo del objetivo número 2. Procesos de color naranjo representan los recursos bioinformáticos utilizados.
*****: La síntesis del vector de clonamiento pUC57-L1 fue realizado mediante externalización del servicio por Genscript.

**: La secuenciación de los plásmidos obtenidos fue realizada mediante externalización del servicio por el Centro de Secuenciación automática de ADN de la Pontificia Universidad Católica de Chile.

4. PRESENTACIÓN Y ANÁLISIS DE RESULTADOS

4.1. Identificación de proteínas de superficie abundantes en contexto de infección de *P. salmonis* por homología y selección de sus péptidos antigénicos

4.1.1. Selección de proteínas de superficie abundantes en contexto de infección de organismos filogenéticamente cercanos a *P. salmonis* mediante búsqueda bibliográfica.

4.1.1.1. *Piscirickettsia salmonis*

A través de la revisión en la literatura existente sobre *P. salmonis*, se logró encontrar que mediante la secuenciación del genoma de la cepa de referencia LF-89 de *P. salmonis* (ATCC VR-1361) se obtuvo un genoma contenido en un cromosoma circular de 3.184.851 pb y en tres plásmidos, pPSLF89-1 (180.124 pb), pPSLF89-2 (33.516 pb) y pPSLF89-3 (51.573 pb). Con un total de 2.850 genes que codifican proteínas, 56 ARNt y seis copias del ARNr 5S-16S-23S (51). Según el análisis de cercanía filogenéticas tomando en cuenta la similitud de la secuencia del ADNr 16S de *P. salmonis* realizado por Mauel M. J. *et al*. (1999), se obtiene que los organismos más cercanos son *Coxiella burnetii*, *Legionella pneumophila*, *Pseudomonas aeruginosa*, *Wolbachia persica* y *Francisella tularensis* con similitudes de 89,5%, 89,2%, 86,8%, 86,5% y 86,3% respectivamente (53).

4.1.1.2. *Coxiella burnetii*

C. burnetii es una Gram negativa con un ciclo de vida bifásico que consta de una variante de células pequeñas ambientalmente estables y una variante de células grandes replicativas de aproximadamente 0,3 y 2,0 µm de tamaño, respectivamente. Esta bacteria se replica dentro de una vacuola parasitófora en monocitos/macrófagos eucariontes *in vivo* y produce un sistema de secreción tipo IVB (T4BSS) que secreta proteínas efectoras en la célula hospedera para controlar numerosos eventos de infección. *C. burnetii* es un patógeno intracelular obligado que causa fiebre Q aguda y crónica (54). Entre los trabajos revisados para la búsqueda de proteínas de superficie abundantes de microorganismos relacionados filogenéticamente a *P. salmonis*, destaca el estudio de Deringer, J. R. *et al*. (2011) donde se identificaron 9 proteínas seroactivas de *C. burnetii* Nine Mile, para las cuales siete de ellas mostraron buena reactividad con sueros inmunes, CBU0932 (glicerol quinasa), CBU1241 (malato deshidrogenasa), CBU1396 (succinato-CoA ligasa), CBU1385 (EF-Ts), CBU0299 (Ribonucleasa PH), CBU0750 (arabinosa-5-fosfato isomerasa) y CBU0103 (peptidasa, familia M20A), mientras que por otro lado se detectaron dos con reactividades bajas, CBU0140 (proteína de división celular FtsA) y CBU0858 (NAD^+ sintasa

dependiente de glutamina). CBU1290 (DnaK) pertenece a la familia de proteínas de choque térmico 70 (Hsp70) y está altamente conservada dentro de los organismos procariontes y eucariontes. Hsp70 funciona como chaperona molecular y en el transporte de proteínas a través de membranas. DnaK puede estar asociado a la membrana en bacterias patógenas y participa en la unión a la superficie de la célula hospedera, facilitando la invasión celular (56). CBU1260 (OmpA) fue identificada como la primera invasina de *Coxiella*, el dominio transmembrana previsto de CBU1260 es típico de la familia de la proteína A de membrana externa (OmpA), la cual se encuentra en varias bacterias y media la adhesión y/o internalización dentro de las células hospederas (57). En el estudio de Zhang, G. *et al*. (2004) se clonaron e identificaron los antígenos inmunodominantes de *C. burnetii*, identificándose 2 proteínas clonadas previamente (Com1 y Hsp60) y 18 proteínas inmunorreactivas nuevas en su tiempo. La mayoría de los genes clonados codificaban proteínas potencialmente expuestas a la superficie y se identificó un único gen del sistema de secreción tipo IV, IcmG (ORF 02827) (58), que mediante la búsqueda en la base de datos de UniProt se pudo identificar como CBU1626.

Mediante la búsqueda de la ubicación subcelular de cada proteína mencionada anteriormente mediante la base de datos de UniProt, se obtiene que CBU1290 (DnaK) y CBU1260 (OmpA) se pueden encontrar en la membrana externa de *C. burnetii*, mientras que CBU1626 (IcmG) es una proteína transmembrana (Tabla 4).

4.1.1.3. *Legionella pneumophila*

L. pneumophila es el patógeno causante de neumonía aguda (legionelosis), enfermedad extrapulmonar (pericarditis o endocarditis) y la fiebre de Pontiac. Las especies de Legionella son Gram negativas con forma de bacilo y residen en la superficie o en el agua potable, transmitiendo la enfermedad a humanos mediante aerosol, generalmente. Estas bacterias se multiplican intracelularmente en macrófagos alveolares. Los neutrófilos y monocitos reclutados, como también las enzimas bacterianas, producen una inflamación alveolar destructiva (59).

Las proteínas asociadas a la membrana externa (del inglés, *Outer Membrane Associated Proteins*, OMAP) son actores importantes en la adhesión de patógenos, la comunicación, la detección ambiental, la transducción de señales y la virulencia. En el estudio de Aurass, P., *et al*. (2016) se identificaron las diez OMAP más abundantes, de las cuales aproximadamente un 70% están presentes en en *L. pneumophila*. Cinco de ellas, LPG0791 (Mip), LPG1841 (Com1), LPG0844 (Ttg2D), LPG0712 (YjeA) y LPG1571 (TolB), se identificaron en cantidad comparable en E/PE (fase exponencial/fase post-

exponencial) y las otras cinco fueron E-up (proteínas más abundantes en la fase exponencial), como LPG0570 (OmpH), LPG0123 (DsbA), LPG0298 (SurA), LPG1331 (HtrA) y LPG1765 (LolA). Este grupo incluye un serie de proteínas de membrana externa ya conocidas y/o predichas (por medio de LocateP) y todas contienen un péptido señal de la señal predicha de la peptidasa I (SPI). La OMAP más abundante detectada fue la chaperona OmpH de tipo prefoldina, que también es conocido como Skp (LPG0507) (61). LPG1974 es la proteína principal de membrana externa de *L. pneumophila* que ha demostrado notables homologías en la longitud de la secuencia primaria y en la identidad de muchos aminoácidos con el canal aniónico dependiente de voltaje de mitocondrias humanas (VDAC1 humano, porina 31HL). Los genes que codifican proteína similares a eucariotas permiten ayudar a *L. pneumophila* en su adaptación al hospedador eucariota. LPG1974 forma poros dependientes de voltajes y selectivos de aniones similares a los poros formados por hVDAC1, sugiriendo también, que LPG1974 es un homólogo estructural y funcional de hVDAC1 (62).

A través de la búsqueda de la ubicación subcelular por la base de datos UniProtKB de cada una de las proteínas mencionadas, se proponen como candidatas a LPG0791 (Mip), LPG1841 (Com1), LPG0507 (Skp) y LPG1974, debido a que están clasificadas como proteínas de membrana externa de *L. pneumophila* (Tabla 4).

4.1.1.4. *Pseudomonas aeruginosa*

P. aeruginosa es una de las principales causas de las infecciones nosocomiales, incluida la neumonía adquirida en el hospital (HAP), infecciones del torrente sanguíneo, del tracto urinario, heridas y quemaduras (63). Este organismo es una gamma-proteobacteria Gram negativa ubicua metabólicamente versátil que prospera en el suelo y los hábitats acuáticos. De los muchos factores de virulencia empleados por *P. aeruginosa* para causar enfermedades, el que atrae más la atención es el sistema de secreción tipo III (T3SS) debido a su ubicuidad en los aislados clínicos y su vínculo epidemiológico de causar peores resultados clínicos y mayor tasa de mortalidad en humanos (63,64). En el estudio de Crouzet, M., *et al*. (2017) se cuantificaron 12 OMAP, las cuales la mayoría se acumuló en exceso en células unidas. Entre ellas, PA1178 (OprH) aumenta la estabilidad de las membranas externas de *P. aeruginosa* al interactuar directamente con el lipopolisacárido y PA2853 (Oprl) se adhiere a las células epiteliales de la tráquea y el intestino delgado de pollos. El resto de las proteínas cuantificadas fueron PA0291 (OprE), PA0833, PA1041, PA1777 (OprF), PA2760, PA3647, PA3790 (OprC), PA3819 y PA4974 (OmpH) (66).

Mediante la búsqueda de cada una de las proteínas en UniProt, se obtuvo que las candidatas para análisis son PA0291, PA1178, PA1777, PA2853, PA3819, PA3790 y PA4974, de acuerdo con la información que está disponible sobre la ubicación subcelular de cada una de ellas (Tabla 4).

Tabla 4. Listado de proteínas abundantes en contexto de infección bacteriano de organismos filogenéticamente cercanos a *P. salmonis* comparando la ubicación subcelular de cada proteína.

Organismo	Nombre de locus ordenado	Acceso UniProt	Nombre recomendado (Anotación UniRule)	Ubicación subcelular (UniProt)	Componente celular (Anotación GO)
Coxiella burnetii	CBU0932	Q83D14	Glicerol quinasa	No disponible	Citosol (GO:0005829)
Coxiella burnetii	CBU1241	Q83C87	Malato deshidrogenasa	No disponible	No disponible
Coxiella burnetii	CBU1396	P53591	Succinato-CoA ligasa (formador de ADP), subunidad alfa	No disponible	Complejo succinato-CoA ligasa, formador de ADP (GO:0009361)
Coxiella burnetii	CBU1385	Q9X5U9	Factor de elongación Ts (EF-Ts)	Citoplasma	Citoplasma (GO:0005737)
Coxiella burnetii	CBU0299	Q83EL9	Ribonucleasa PH	No disponible	No disponible
Coxiella burnetii	CBU0750	Q83DI1	Arabinosa-5-fosfato isomerasa	No disponible	No disponible
Coxiella burnetii	CBU0103	Q83F46	Peptidasa, familia M20A	No disponible	No disponible
Coxiella burnetii	CBU0140	Q83F13	Proteína de división celular FtsA	Membrana celular interna	Sitio de división celular (GO:0032153)
					Lado citoplasmático de la membrana plasmática (GO:0009898)
Coxiella burnetii	CBU0858	Q83D82	NAD(+) sintasa dependiente de glutamina	No disponible	Citoplasma (GO:0005737)
Coxiella burnetii	CBU1290	O87712	Proteína chaperona DnaK	Citoplasma	Citosol (GO:0005829)
				Membrana celular	Membrana plasmática (GO:0005886)
Coxiella burnetii	CBU1260	Q83C69	Proteína no caracterizada CBU_1260 (OmpA)	Membrana celular externa	Membrana celular externa (GO:0009279)
Coxiella burnetii	CBU1626	Q83B87	IcmG (Sistema de secreción tipo IV)	Transmembrana	Membrana (GO:0016020)
Legionella pneumophila	LPG0791	Q5ZXE0	Proteína de membrana externa MIP	Membrana celular externa	Membrana celular externa (GO:0009279)
Legionella pneumophila	LPG1841	Q5ZUF7	Proteína de membrana externa de 27 kDa	No disponible	No disponible
Legionella pneumophila	LPG0844	Q5ZX88	Proteína de tolerancia al tolueno Ttg2D	No disponible	No disponible
Legionella pneumophila	LPG0712	Q5ZXL9	Similar a endo-1,4-beta-xilanasa	No disponible	No disponible
Legionella pneumophila	LPG1571	Q5ZV69	Proteína del sistema Tol-Pal TolB	Periplasma	Espacio periplásmico (GO:0042597)
Legionella pneumophila	LPG0123	Q5ZZ88	Proteína de intercambio tiol:disulfuro	Periplasma	Espacio periplásmico (GO:0042597)
Legionella pneumophila	LPG0298	Q5ZYR3	Chaperona SurA	Periplasma	Espacio periplásmico delimitado por membrana externa (GO:0030288)
Legionella pneumophila	LPG1331	Q5ZVV9	Serina endoproteasa periplasmática similiar a DegP	Periplasma	Espacio periplásmico (GO:0042597)
Legionella pneumophila	LPG1765	Q5ZUM8	Proteína transportadora de lipoproteínas de membrana externa	Periplasma	Espacio periplásmico (GO:0042597)
Legionella pneumophila	LPG0507	Q5ZY65	Proteína OmpH de membrana externa (Skp)	No disponible	No disponible
Legionella pneumophila	LPG1974	Q5ZU34	Proteína principal de membrana externa	No disponible	No disponible
Pseudomonas aeruginosa	PA1178	G3XD11	PhoP/Q y proteína de membrana externa H1 inducible en bajo Mg^{2+}	No disponible	Membrana celular externa (GO:0009279)
Pseudomonas aeruginosa	PA2853	P11221	Lipoproteína principal de membrana externa	Membrana celular externa	Membrana celular externa (GO:0009279)
				Ancla lipídica	Membrana externa (GO:0019867)
Pseudomonas aeruginosa	PA0291	G3XDA5	Porina de membrana externa inducida anaeróbicamente OprE	No disponible	Membrana externa (GO:0019867)
Pseudomonas aeruginosa	PA0833	Q9I5A7	Proteína que contiene un dominio similar a OmpA	No disponible	Membrana (GO:0016020)
Pseudomonas aeruginosa	PA1041	Q9I4T3	Proteína de membrana externa probable	No disponible	Membrana (GO:0016020)
Pseudomonas aeruginosa	PA1777	P13794	Proteina de membrana externa F	Membrana celular externa	Membrana celular externa (GO:0009279)
				Proteína de membrana multipaso	Membrana externa (GO:0019867)
					Complejo de poros (GO:0046930)
Pseudomonas aeruginosa	PA2760	Q9I083	Proteína de membrana externa probable	No disponible	Membrana (GO:0016020)
Pseudomonas aeruginosa	PA3647	Q9HXY5	Proteína similar a Skp	No disponible	Espacio periplásmico delimitado por membrana externa (GO:0030288)
Pseudomonas aeruginosa	PA3790	G3XD89	Porina de membrana externa de transporte de cobre OprC	Membrana celular externa	Membrana celular externa (GO:0009279)
Pseudomonas aeruginosa	PA3819	Q9HXI3	Proteína que contiene el dominio de cremallera de glicina 2TM	No disponible	Membrana externa (GO:0019867)
Pseudomonas aeruginosa	PA4974	G3XCK3	Proteína canal TolC	No disponible	Membrana externa (GO:0019867)

Se utilizó la base de datos de UniProtKB para la búsqueda de la información obtenida para cada proteína. Resaltado en naranjo, las coincidencias idóneas para cada proteína analizada.

4.1.2. Selección de antígenos de superficie de *P. salmonis* por homología de proteínas abundantes en contexto de infección de organismos filogenéticamente cercanos

4.1.2.1. Búsqueda de proteínas abundantes en contexto de infección por homología en *P. salmonis* mediante alineamiento local de secuencias proteicas

Con el objetivo de buscar el grado de similitud de cada proteína candidata se realizaron alineamientos locales de sus secuencias aminoacídicas contra las secuencias de proteínas ya descritas de *Piscirickettsia salmonis* ATCC VR-1361 (taxid:1227812) mediante la herramienta Blastp utilizando la base de datos de secuencias de proteínas no-redundantes (nr) y todos los demás valores por defecto, reflejándose los resultados en la Tabla 5.

En el caso de las proteínas de *C. burnetii*, se obtuvieron dos emparejamientos o coincidencias en *P. salmonis* de 76,77% y 76,31% de identidad con cobertura de 99% para CBU1290 o la chaperona molecular DnaK; cuatro coincidencias de 25,10%, 23,53%, 25,30% y 24,90% de identidad con 100% de cobertura, una coincidencia de 26,59% de identidad con 69% de cobertura y dos coincidencias de 31,94% y 36,11% de identidad con un 24% de cobertura para CBU1269; cinco coincidencias de 26,32%, 41,18%, 41,51%, 41,51% y 22,30% de identidad con coberturas de 44%, 20%, 21%, 21% y 55%, respectivamente, para CBU1626 o IcmG del sistema de secreción tipo IV de *C. burnetii*.

Por otro lado, para las proteínas de *L. pneumophila*, se obtuvieron dos coincidencias para LPG0791 o la proteína de membrana externa MIP, correspondientes a proteínas con 38,40% de identidad y 99% de cobertura y, para el caso de la proteína LPG1841, se obtuvieron tres coincidencias de 35%, 96%, 35%, 53% y 27,53% de identidad con coberturas de 81%, 81% y 92%, respectivamente.

Finalmente, en los alineamientos locales de las proteínas de *P. aeruginosa* contra *P. salmonis*, se obtuvo para la proteína PA1777 u OprF, dos coincidencias de 28,95% de identidad y 97% de cobertura, una coincidencia de 30,09% de identidad y 32% de cobertura, y otra coincidencia de 29,70% de identidad con un 28% de cobertura; para PA3819 o proteína que contiene el dominio de cremallera de glicina 2TM, se obtuvieron dos coincidencias de 34,41% y 33,33% de identidad con 51% de cobertura; para PA4974 o TolC, se obtuvieron dos coincidencias con cobertura de 89% e identidad de 32,26% y 31,34%, y dos coincidencias con cobertura de 93% e identidad de 28,16% y 27,72%.

En los casos de las proteínas LPG0507 y LPG1974 de *L. pneumophila*, como también para las proteínas PA1178, PA2853, PA0291 y PA3790 de *P. aeruginosa*, no se encontraron similitudes significativas

mediante sus alineamientos locales mediante Blastp contra *P. salmonis* y con los parámetros establecidos, por lo tanto, se descartaron de la selección.

Tomando en consideración todas las coincidencias obtenidas y mencionados anteriormente, se seleccionó la proteína de *P. salmonis* más idónea como potencial homóloga para cada proteína comparada, tomando como referencia los valores de porcentaje de identidad, cobertura, puntaje máximo y E-value, como calificativos de idoneidad. Las proteínas seleccionadas de *P. salmonis* para un posterior análisis (Tabla 5, filas resaltadas en color naranjo) son la chaperona molecular DnaK- (WP_016210369.1), la proteína barril-beta de membrana externa (WP_016209892.1), la proteína hipotética (WP_016209681.1), la isomerasa cis-trans de peptidil-propil tipo FKBP (WP_017376766.1), la proteína de la familia DsbA (WP_016210800.1), la proteína de la familia OmpA (WP_017376173.1), la proteína de superficie de la familia RT0821/Lpp0805 (WP_027242773.1) y la proteína de membrana externa de la familia TolC (WP_032126547.1).

Tabla 5. Resultados alineamiento local de secuencias proteicas por Blastp contra *P. salmonis* de las proteínas de superficie abundantes en contexto de infección de *C. burnetii*, *L. pneumophila* y *P. aeruginosa*.

Proteína de referencia de otros organismos			Proteínas de *Piscirickettsia salmonis* obtenidas del alineamiento de secuencias						
Nombre de locus ordenado	NCBI Reference Sequence (RefSeq)	Nombre recomendado (Anotación UniRule)	NCBI Reference Sequence (RefSeq)	Nombre descriptivo (NCBI)	Puntaje máximo	Cobertura	E-value	Porcentaje identidad	Longitud de accesión
CBU1290	WP_005770882.1	Proteína chaperona DnaK	WP_016210369.1*	Chaperona molecular Dnak	984,0	99%	0,0	76,77%	644
			WP_017378060.1	Chaperona molecular Dnak	979,0	99%	0,0	76,31%	644
CBU1260	WP_010958108.1	Proteína no caracterizada CBU_1260 (OmpA)	WP_016209892.1*	Proteína barril-beta de membrana externa	60,5	100%	8,0E-12	25,10%	208
			WP_026063589.1	Proteína barril-beta de membrana externa	56,2	100%	2,0E-10	23,53%	208
			WP_017377000.1	Proteína barril-beta de membrana externa	45,4	100%	1,0E-06	25,30%	189
			WP_016209896.1	Proteína barril-beta de membrana externa	42,4	100%	1,0E-05	24,90%	189
			WP_017377818.1	Proteína barril-beta de membrana externa	39,3	69%	1,0E-04	26,59%	214
			WP_016210293.1*	Proteína de transporte de proteínas de membrana	38,5	24%	4,0E-04	31,94%	426
			WP_017377245.1	Proteína de transporte de proteínas de membrana	38,5	24%	4,0E-04	36,11%	426
CBU1626	WP_010958333.1	IcmG (Sistema de secreción tipo IV)	WP_016209681.1	Proteína hipotética	55,1	44%	1,0E-09	26,32%	326
			WP_027242630.1	Proteína hipotética	52,0	20%	1,0E-08	41,18%	326
			WP_017378269.1	Proteína hipotética	51,6	21%	1,0E-08	41,51%	301
			WP_016209971.1*	Proteína hipotética	51,6	21%	2,0E-08	41,51%	301
			WP_027242653.1	Proteína hipotética	50,4	55%	4,0E-08	22,30%	308
LPG0791	WP_011213317.1	Proteína de membrana externa MIP	WP_017376766.1	Isomerasa cis-trans de peptidil-prolil tipo FKBP	172,0	99%	5,0E-54	38,40%	259
			WP_016210710.1	Isomerasa cis-trans de peptidil-prolil tipo FKBP	170,0	99%	2,0E-53	38,40%	259
LPG1841	WP_010947567.1	Proteína de membrana externa de 27 kDa	WP_016210800.1*	Proteína de la familia DsbA	110,0	81%	2,0E-29	35,96%	275
			WP_017376827.1	Proteína de la familia DsbA	109,0	81%	3,0E-29	35,53%	275
			WP_027242548.1	Proteína de la familia DsbA	102,0	92%	5,0E-27	27,53%	250
LPG0507	WP_010946255.1	Proteína OmpH de membrana externa (Skp)	Sin similitud significativa	N/A	N/A	N/A	N/A	N/A	N/A
LPG1974	WP_010947690.1	Proteína principal de membrana externa	Sin similitud significativa	N/A	N/A	N/A	N/A	N/A	N/A
PA1178	WP_003082431.1	PhoP/Q y proteína de membrana externa H1 inducible en bajo Mg^{2+}	Sin similitud significativa	N/A	N/A	N/A	N/A	N/A	N/A
PA2853	WP_003090944.1	Lipoproteína principal de membrana externa	Sin similitud significativa	N/A	N/A	N/A	N/A	N/A	N/A
PA0291	WP_003084265.1	Porina de membrana externa inducida anaeróbicamente OprE	Sin similitud significativa	N/A	N/A	N/A	N/A	N/A	N/A
PA1777	WP_003087843.1	Proteina de membrana externa F	WP_017376173.1*	Proteína de la familia OmpA	106,0	97%	5,0E-27	28,95%	326
			WP_122940232.1	Proteína de la familia OmpA	106,0	97%	6,0E-27	28,95%	326
			WP_016209425.1	Proteína de la familia OmpA	52,0	32%	1,0E-08	30,09%	213
			WP_017377882.1	Proteína de la familia OmpA	51,2	28%	3,0E-08	29,70%	213
PA3790	WP_003119294.1	Porina de membrana externa de transporte de cobre OprC	Sin similitud significativa	N/A	N/A	N/A	N/A	N/A	N/A
PA3819	WP_003092848.1	Proteína que contiene el dominio de cremallera de glicina 2TM	WP_027242773.1*	Proteína de superficie de la familia de RT0821/Lpp0805	42,7	51%	3,0E-06	34,41%	162
			WP_032126730.1	Proteína de superficie de la familia de RT0821/Lpp0805	40,8	51%	1,0E-05	33,33%	162
PA4974	WP_003095699.1	Proteína canal TolC	WP_032126547.1	Proteína de membrana externa de la familia TolC	186	89%	3,0E-54	32,26%	480
			WP_027243158.1	Proteína de membrana externa de la familia TolC	184	89%	3,0E-53	31,34%	486
			WP_230387406.1*	Proteína de membrana externa de la familia TolC	178	93%	2,0E-51	28,16%	442
			WP_230383380.1	Proteína de membrana externa de la familia TolC	175	93%	2,0E-50	27,72%	441

Se utilizó un BLAST proteína-proteína (Blastp) con la base de datos estándar de secuencias de proteínas no redundantes (nr), usando como organismo objetivo *Piscirickettsia salmonis* ATCC VR-1361 (taxid:1227812) y los parámetros de algoritmo por defecto. Resaltado en naranjo, las coincidencias idóneas para cada proteína analizada.

*, resultados coincidentes con lo obtenido por un Blastp con base de datos experimental de secuencias de proteínas no redundantes agrupadas (clustered nr), usando como organismo objetivo *Piscirickettsia salmonis* ATCC VR-1361 (taxid:1227812) y los parámetros de algoritmo por defecto, obteniéndose como antepasado de clúster a g-proteobacteria para todos los casos.

4.1.2.2. Filtro de conservación aminoacídica de proteínas de *P. salmonis* para selección en base a especificidad de posterior péptido a expresar.

Para la selección de las mejores proteínas candidatas para la expresión recombinante por *Escherichia coli*, se realizaron alineamientos utilizando la herramienta Blastp para cada proteína contra *C. burnetii*, *L. pneumophila*, *P. aeruginosa* y *P. salmonis*, como también, un Blastp de cada proteína sin especificaciones para analizar su conservación en otros organismos que no sean cercanos filogenéticamente a *P. salmonis*. En el caso de las proteínas con mayor identidad obtenidas en los Blastp con los tres organismos filogenéticamente más cercanos a *P. salmonis* según el estudio de Mauel M. J. *et al.* (1999), se seleccionaron las mejores coincidencias para cada proteína y se realizó un alineamiento múltiple de sus secuencias mediante Clustal Omega, buscando las regiones más conservadas en las respectivas secuencias proteicas.

En el caso de la chaperona molecular DnaK (WP_016210369.1) se obtuvo, mediante el Blastp excluyendo al filo *Proteobacteria* (taxid:1224) (Tabla S1), una gran variedad de proteínas de distintos organismos con altos porcentajes de identidad de secuencias, desde 68% hasta un 77%, con coberturas desde 93% a 100%. Incluyendo proteínas de organismos con género *Streptococcus*, *Euryarchaeota*, *Halobacteria*, *Stretosporangium*, *Pseudomonas*, *Priestia*, *Citrobacter*, *Drosophila*, entre otros. Para corroborar su conservación en organismos filogenéticamente lejanos a *P. salmonis* se realizó un Blastp excluyendo al dominio *Bacteria* (taxid:2) para el cual se obtiene, nuevamente, una gran cantidad de proteínas de distintos organismos con una alta identidad de secuencia (+60%), por lo tanto, la secuencia proteica de la chaperona molecular DnaK de *P. salmonis* es altamente conservada en muchos organismos sin incluir a microorganismos bacterianos. En el Blastp contra los tres organismos filogenéticamente cercanos a *P. salmonis* (Tabla S2), se obtuvieron proteínas con identidades de entre 74,6% a 76,6%, con coberturas desde un 98% a un 100%.

Con relación a la proteína barril-beta de membrana externa (WP_016209892.1), mediante el Blastp excluyendo al filo *Proteobacteria* (taxid:1224) (Tabla S3), se obtuvieron únicamente 19 proteínas con baja identidad de secuencias que no superan el 33%, de organismos pertenecientes a los géneros *Abscondita*, *Tanacetum*, *Geothrix*, *Sulfuvarum*, *Elizabethkingia*, *Flavobacterium*, *Mycobacterium* y *Finegoldia*. Para su Blastp con los tres organismos filogenéticamente cercanos como objetivo (Tabla S3)

se obtuvieron 36 proteínas con identidades de secuencias menores al 30,21%, de las cuales 11 de las proteínas son pertenecientes a *C. burnetii* y 25 a *L. pneumophila*. Se seleccionaron las mejores coincidencias para cada organismo, las proteínas beta-barril de membrana externa de *C. burnetii* (WP_040947705.1) y de *L. pneumophila* (WP_061513559.1) con identidades de 27,63% y 26,53%, respectivamente, y se les realizó un alineamiento múltiple de sus secuencias proteicas por Clustal Omega incluyéndose su homólogo hipotético en *P. salmonis* (WP_016209892.1) y la segunda proteína idónea en el análisis de homología desde las proteínas abundantes en contexto de infección (Tabla 5), correspondiente a la proteína beta-barril de membrana externa (WP_026063589.1) de *P. salmonis*. obteniéndose tres regiones más conservadas dentro del alineamiento tomando como referencia la secuencia proteica de WP_016209892.1, la primera, desde el residuo 2 hasta el número 46; la segunda, desde el residuo 79 hasta el 99; y la tercera, desde el residuo 154 hasta el 187.

Mediante el Blastp excluyendo al filo *Proteobacteria* (taxid:1224) de la proteína hipotética WP_016209681.1 de *P. salmonis* (Tabla S4) se obtuvo solamente una proteína, correspondiente a la proteína hipotética KFE69_05580 de la especie *bacterium* SCSIO 12844 con un 34,85% de identidad y una cobertura de 39%. Para su Blastp con los tres organismos filogenéticamente cercanos como objetivo (Tabla S4) se obtuvieron 86 proteínas con identidad de secuencia entre 26,3% y 47,2%, y coberturas menores a 48%. Por lo tanto, se realizó alineamiento múltiple mediante Clustal Omega donde se seleccionaron las proteínas IcmG/DotF del sistema de secreción tipo IVB de *L. pneumophila* (WP_027228891.1), de *P. aeruginosa* (WP_070610851.1) y de *C. burnetii* (WP_005772726.1), con identidades de 31,79%, 47,17% y 24,09%, respectivamente, y se incluyó la proteína hipotética de *P. salmonis* analizada (WP_016209681.1) con los dos siguientes similitudes de la lista de proteínas idóneas en el análisis de homología desde las proteínas abundantes en contexto de infección (Tabla 5), es decir las proteínas WP_027242630.1 y WP_017378269.1. La región más conservada dentro de la secuencia proteica dentro del alineamiento múltiple se ubica desde el residuo 275 hasta el 324, tomando como referencia la secuencia de WP_027242630.1.

Con respecto a la isomerasa cis-trans de peptidil-prolil tipo FKBP (WP_017376766.1), mediante el Blastp excluyendo al filo *Proteobacteria* (taxid:1224) (Tabla S5), se obtuvieron una gran variedad de proteínas de distintos organismos con identidades desde 38,62% a 46,98%, destacándose proteínas pertenecientes a *Nitrospirota bacterium*, *Deltaproteobacteria bacterium*, *Acidobacteriia bacterium*, entre otros organismos. Además, se realizó un Blastp excluyendo al dominio Bacteria (taxid:2) obteniéndose un gran número de proteínas de diversos organismos con una similitud media (40% a

46%), por lo tanto, esta proteína es conservada en distintos organismos, incluyendo no bacterianos. En el Blastp contra los tres organismos filogenéticamente cercanos a *P. salmonis* (Tabla S6), se obtuvieron proteínas de identidades entre 33% a 41%, en donde destacan para cada organismo las proteína peptidil-prolil cis-trans-isomerasa de tipo FKBP de *C. burnetii* (WP_011996666.1) con un 40,56% de identidad, un potenciador de infectividad de macrófagos Mip de *L. pneumophila* (WP_059427894.1) con un 38,80% de identidad y un precursor Mio de peptidil-prolil cis-trans isomerasa de *P. aeruginosa* (CRR57457.1) con un 39,68% de identidad.

Mediante el Blastp excluyendo al filo *Proteobacteria* (taxid:1224) de la proteína de la familia (WP_016210800.1) de *P. salmonis* (Tabla S7), se obtuvieron proteínas de organismos pertenecientes, principalmente, a los generos *Bacillus*, *Campylobacter*, *Cichorium*, *Drosophila*, *Prestia*, *Photobacterium* y *Nitrospira*, con identidades rondando entre 28,25% e inclusive 100%, para el caso de la proteína de formación de enlaces disulfuro DsbA (PRP61050.1) de *Bacillus halotolerans*. Además, a la proteína analizada se le realizó un Blastp excluyendo al dominio *Bacteria* (taxid:2) obteniéndose un gran número de proteínas de diversos organismos con una similitud baja-media (28% a 35%), por lo tanto, esta proteína es conservada en distintos organismos, incluyendo no bacterianos. En el Blastp contra los tres organismos filogenéticamente cercanos a *P. salmonis* (Tabla S8), se obtuvieron diversas proteínas con identidades que llegan hasta 55% en algunos casos, aproximadamente, destacándose la proteína que contiene un dominio de tiorredoxina (MXH37503.1) de *P. Aeruginosa*, las proteínas de la familia DsbA de *C. burnetii* (WP_284185834.1) y de *L. pneumophila* (WP_044497326.1).

Para la proteína de la familia OmpA (WP_017376173.1) se realizó un Blastp excluyendo al filo *Proteobacteria* (taxid:1224) (Tabla S9), obteniéndose proteínas con valor máximo de identidad de 33,04%, destacándose organismos de los géneros *Desulfuromonas*, *Geobacter*, *Citrifermentans* y *Helicobacter*, entre otros. Por otro lado, al realizarle un Blastp contra los tres organismos filogenéticamente cercanos a *P. salmonis* (Tabla S10) se obtuvieron entre las primeras 100 mejores coincidencias solo proteínas de *P. aeruginosa*, destacándose como la proteína con mayor identidad la porina de membrana externa OprF (WP_070338528.1), con un 30,06% de identidad de secuencias y un 97% de cobertura.

En el caso de la proteína de superficie de la familia de RT0821/Lpp0805 (WP_027242773.1) de *P. salmonis*, mediante el Blastp excluyendo al filo *Proteobacteria* (taxid:1224) (Tabla S11), se obtuvieron una gran variedad de proteínas de distintos organismos con identidades desde 32% a 50%,

destacándose proteínas pertenecientes a *Dehalococcoidia bacterium*, *Pelobacter seleniigenes*, *Nitrospinota bacterium* y *Halobacteria archaeon*, entre otros organismos. Para el caso del Blastp contra los tres organismos filogenéticamente cercanos a *P. salmonis* (Tabla S12), se obtuvieron 19 proteínas de hasta un 49,41% de máximo de valor de identidad de secuencias.

Finalmente, para la proteína de membrana externa de familia TolC (WP_032126547.1), se le realizó un Blastp excluyendo al filo *Proteobacteria* (taxid:1224) (Tabla S13), obteniéndose proteínas pertenecientes a los organismos de géneros *Bacillus*, *Halobacterium*, *Aureobasidium*, *Capitella*, *Virgibacillus*, entre otros, con identidades de secuencias entre un 28,84% a un 41,12%. En el Blastp contra los tres organismos filogenéticamente cercanos a *P. salmonis* (Tabla S14), se obtuvieron una gran cantidad de proteínas con un valor de identidad máximo del 35,92%, destacándose las proteínas de membrana externa de familia TolC para cada organismo, WP_010957349.1 con un 35,16% de identidad para *C. burnetii*, WP_062727197.1 con un 35,92% de identidad para *L. pneumophila* y MBH4410191.1 con un 35,31% de identidad para *P. aeruginosa*. A las secuencias de las proteínas mencionadas anteriormente e incluyendo a la proteína de membrana externa de familia TolC de *P. salmonis* (WP_032126547.1) se les realizó un alineamiento múltiple de secuencias de proteínas mediante Clustal Omega, obteniéndose dos regiones conservadas evidentes. Tomando como referencia la secuencia de la proteína de *P. salmonis*, la primera región conservada está ubicada desde el residuo número 168 hasta el 252 y la segunda desde el residuo número 348 hasta el 475.

Tomando en consideración los resultados mencionados anteriormente y el análisis estadístico de cada uno de los alineamientos locales realizados, reflejado en la Tabla 6, se seleccionaron 3 proteínas candidatas consideradas como poco conservadas entre organismos filogenéticamente cercanos y lejanos a *P. salmonis*. Las proteínas seleccionadas (color naranjo en Tabla 6) fueron la proteína barril-beta de membrana externa (WP_016209892.1), la proteína de la familia OmpA (WP_017376173.1) y la proteína de membrana externa de la familia TolC (WP_032126547.1).

Tabla 6. Comparación estadística de los resultados de alineamientos locales mediante Blastp de las proteínas de superficie de *P. salmonis* propuestas como abundantes en contexto de infección por homología.

Proteínas de *Piscirickettsia salmonis* obtenidas del alineamiento de secuencias		Porcentaje de identidad de las proteínas obtenidas de Blastp excluyendo a *Proteobacteria*					Porcentaje de identidad de las proteínas obtenidas de Blastp contra organismos filogenéticamente cercanos				
NCBI Reference Sequence (RefSeq)	Nombre descriptivo (NCBI)	Media	Mediana	Valor máximo	Desviación estándar	N° de prot. comparadas	Media	Mediana	Valor máximo	Desviación estándar	N° de prot. comparadas
WP_016210369.1	Chaperona molecular Dnak	72,72%	72,95%	77,09%	0,01751	100	75,09%	74,73%	76,60%	0,00554	100
WP_016209892.1	Proteína barril-beta de membrana externa	26,69%	26,19%	32,98%	0,03169	19	27,07%	27,63%	30,21%	0,01386	36
WP_016209681.1	Proteína hipotética	34,85%	34,85%	34,85%	0,0	1	35,35%	30,59%	47,17%	0,07926	86
WP_017376766.1	Isomerasa cis-trans de peptidil-prolil tipo FKBP	42,47%	42,26%	46,98%	0,01577	100	37,49%	37,84%	40,65%	0,01794	99
WP_016210800.1	Proteína de la familia DsbA	35,53%	33,49%	100,00%	0,08151	100	37,24%	36,28%	55,88%	0,04340	100
WP_017376173.1	Proteína de la familia OmpA	29,20%	29,24%	33,04%	0,01485	100	29,27%	29,21%	30,06%	0,00266	100
WP_027242773.1	Proteína de superficie de la familia de RT0821/Lpp0805	41,50%	41,43%	50,00%	0,03612	100	40,84%	43,54%	49,41%	0,07751	19
WP_032126547.1	Proteína de membrana externa de la familia TolC	33,12%	32,93%	41,12%	0,02430	100	33,34%	32,17%	35,92%	0,01937	100

4.1.2.3. Modelamiento de proteínas candidatas

Las proteínas candidatas resultantes de los análisis por alineamiento local por Blastp y por alineamiento múltiple por Clustal Omega se sometieron a modelamiento proteico mediante el recurso bioinformático SWISS-MODEL del Swiss Institute of Bioinformatics, buscando analizar las estructuras tridimensionales formadas identificando las regiones de residuos aminoacídicos expuestas al medio extracelular. En un primer lugar, para la proteína barril-beta de membrana externa (WP_016209892.1) se obtuvo como mejor modelo proteico el reflejado en la Figura 3, el cual fue logrado a través del templado de la proteína de la familia de la meta-vía de degradación del fenol (A0A1L6TDB8.1.A), modelo conseguido por AlphaFold v2 (DB: A0A1L6TDB8_PISSA, gen: A0A1L6TDB8_PISSA, organismo: *P. salmonis*). El modelo proteico corresponde a una biounidad de oligo-estado de monómero con una identidad de secuencia de 96,63%, cobertura del 100% y un estimado de calidad del modelo global (GMQE) de 0,66.

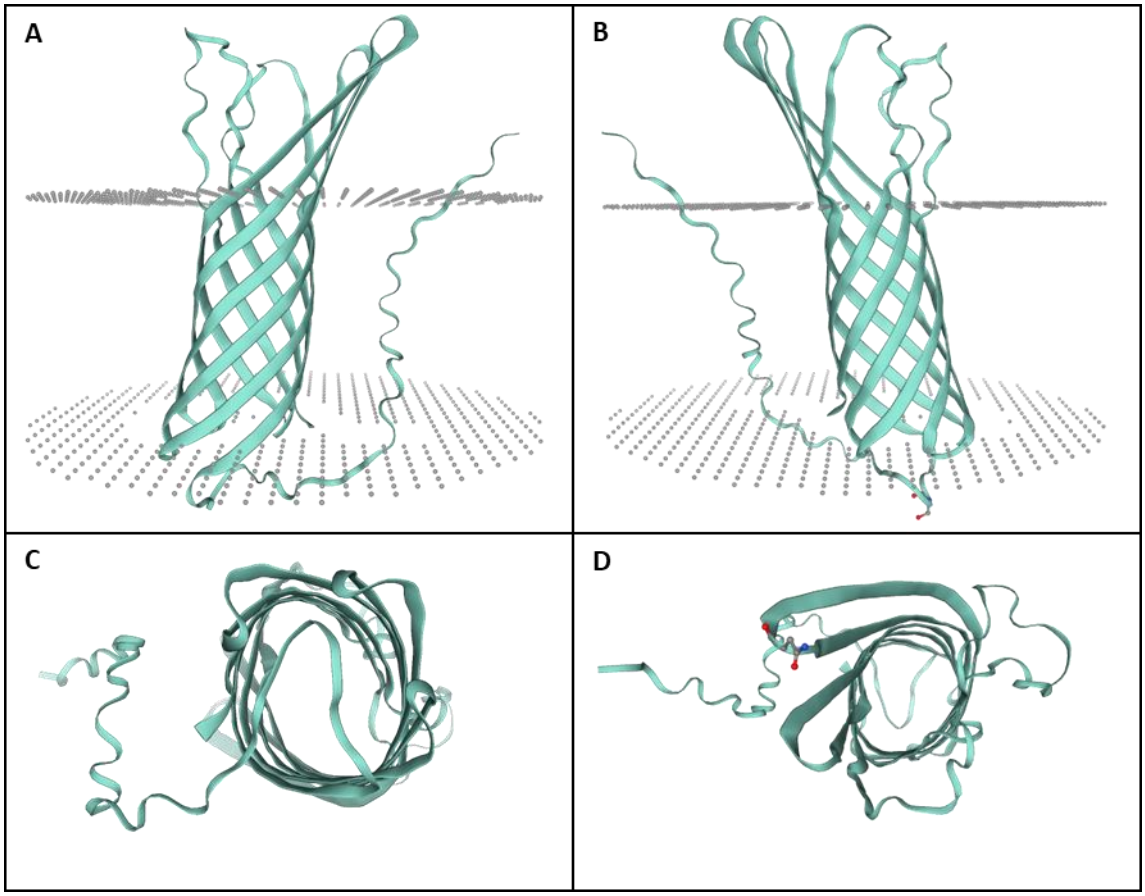

Figura 3. Modelo proteico predicho por SWISS-MODEL de la proteína barril-beta de membrana externa de *P. salmonis (*WP_016209892.1).
A, B, C y D, corresponden al mismo modelo proteico predicho visto desde diferentes ángulos. Puntos grises de A y B, corresponden a la membrana externa e interna, para la zona superior e inferior de cada imagen, respectivamente, predicho por SWISS-MODEL por ser un modelo proteico que tiene segmentos transmembrana en base al templado utilizado.

Por otro lado, para la proteína de la familia OmpA (WP_017376173.1) se obtuvo como mejor modelo proteico el reflejado en la Figura 4, el cual fue logrado a través del templado de una proteína probablemente obsoleta según SWISS-MODEL, con entrada de UniProtKB desconocida (A0A3M8FMR5.1.A), modelo conseguido por AlphaFold v2 (DB: A0A3M8FMR5, gen: desconocido, organismo: desconocido). El modelo proteico corresponde a una biounidad de oligo-estado de monómero con una identidad de secuencia de 98,77%, cobertura del 100% y un GMQE de 0,74.

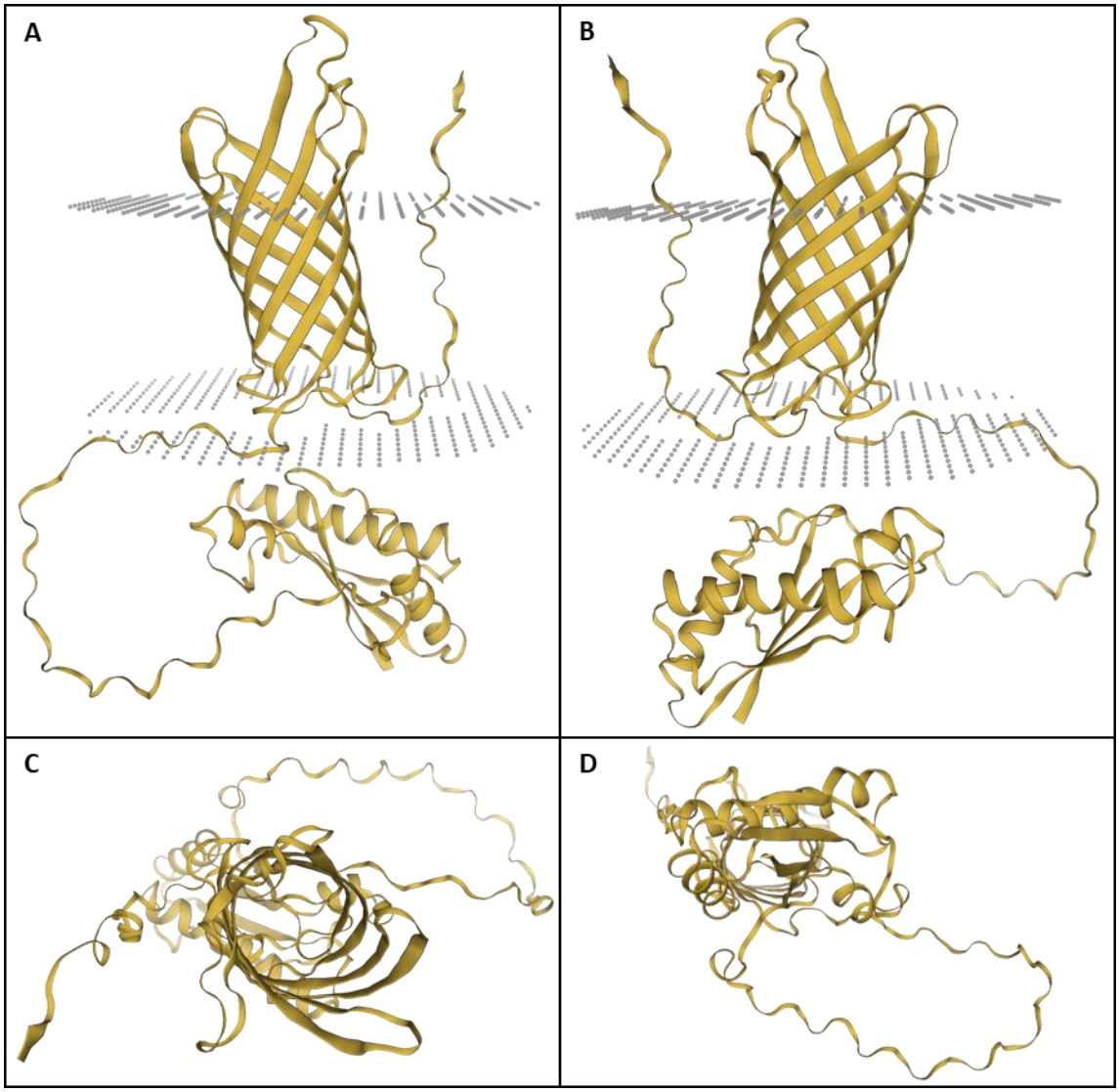

Figura 4. Modelo proteico predicho por SWISS-MODEL de la proteína de la familia OmpA de *P. salmonis* (WP_017376173.1).
A, B, C y D, corresponden al mismo modelo proteico predicho visto desde diferentes ángulos. Puntos grises de A y B, corresponden a la membrana externa e interna, para la zona superior e inferior de cada imagen, respectivamente, predicho por SWISS-MODEL por ser un modelo proteico que tiene segmentos transmembrana en base al templado utilizado.

Finalmente, para la proteína de membrana externa de la familia TolC (WP_032126547.1) se obtuvo como mejor modelo proteico el reflejado en la Figura 5, el cual fue logrado a través del templado de una TolC del sistema de secreción tipo I (A0A1L6TAB0.1.A), modelo conseguido por AlphaFold v2 (DB: A0A1L6TAB0_PISSA, gen: A0A1L6TAB0_PISSA, organismo: *P. salmonis*). El modelo proteico corresponde a una biounidad de oligo-estado de monómero con una identidad de secuencia de 98,13%, cobertura del 100% y un GMQE de 0,85. Además, se obtuvo una segunda mejor coincidencia que hace una predicción de que la proteína forma par de un oligo-estado de homotrímero visualizado en la Figura

6, en base al templado utilizado de la proteína TolC de membrana externa (2vde.1.A), obtenido a través cristalografía de rayos X del estado abierto del componente de membrana exterior TolC de la bomba de eflujo multi-droga. Modelo que logró una identidad de secuencia del 34,54%, un GMQE de 0,65 y un QMEANDisCo Global de 0,70 ± 0,05.

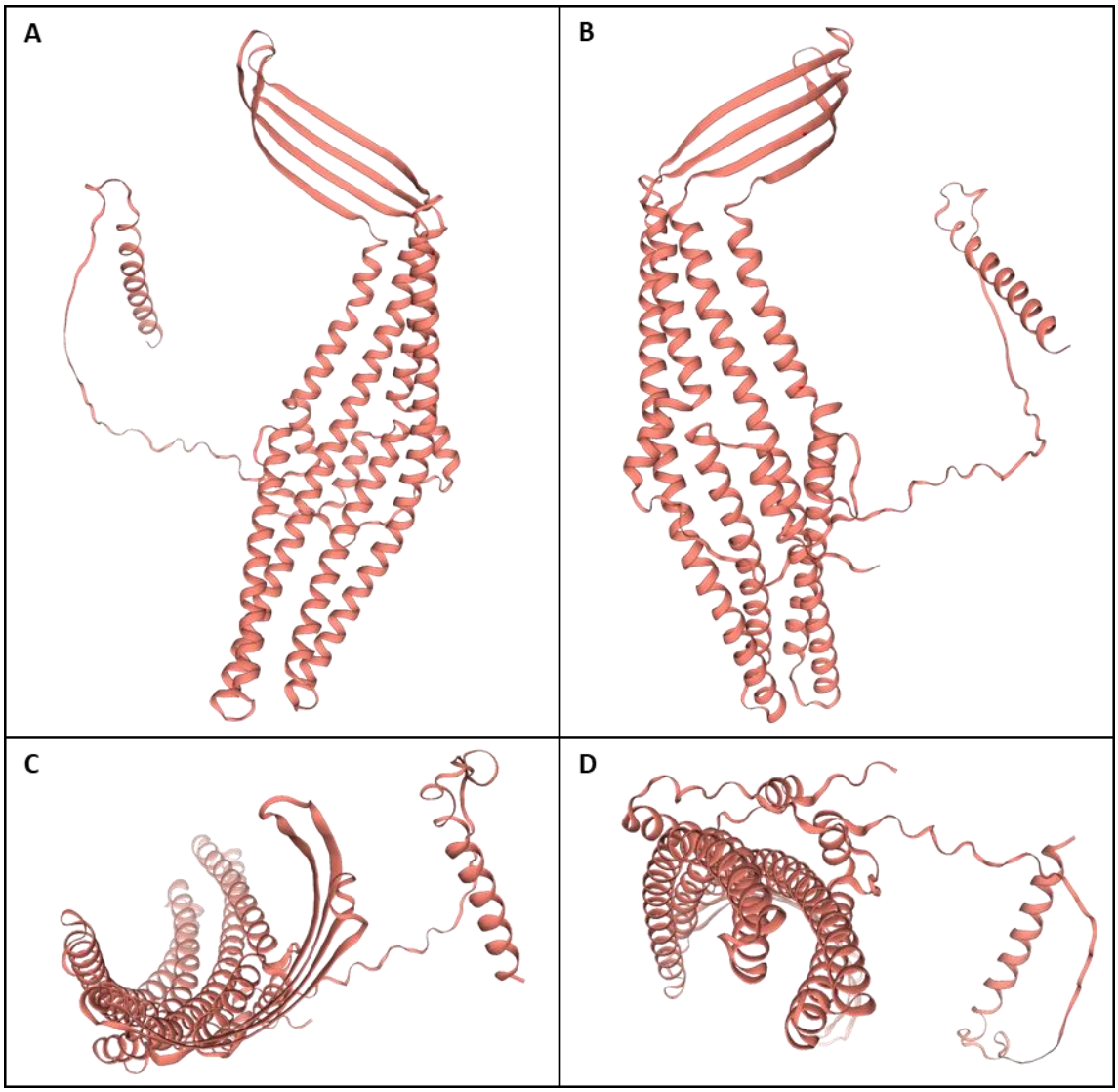

Figura 5. Modelo proteico predicho por SWISS-MODEL de la proteína de membrana externa de la familia TolC de *P. salmonis* (WP_032126547.1), en oligo-estado de monómero.
A, B, C y D, corresponden al mismo modelo proteico predicho visto desde diferentes ángulos.

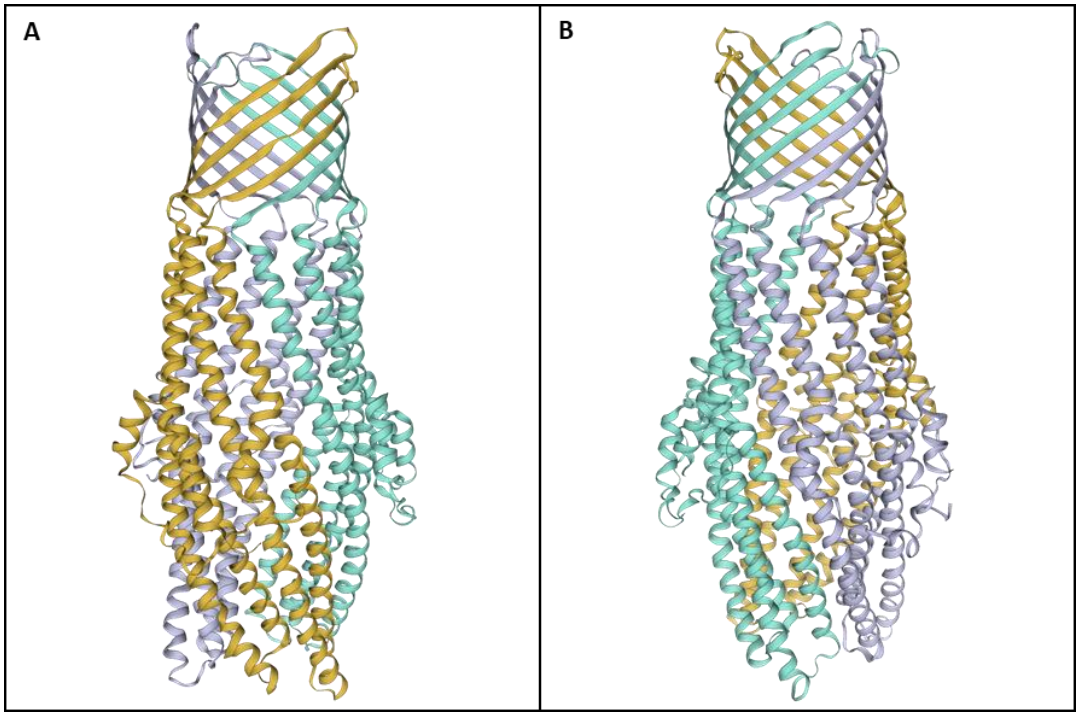

Figura 6. Modelo proteico predicho por SWISS-MODEL de la proteína de membrana externa de la familia TolC de *P. salmonis* (WP_032126547.1), en oligo-estado de homotrímero.
A y B, corresponden al mismo modelo proteico predicho visto desde diferentes ángulos.

Adicionalmente, para la proteína barril-beta de membrana externa de *P. salmonis* (WP_016209892.1) se obtuvieron 7 modelos distintos en base a templados no generados por la tecnología de AlphaFold, si no que, generados en base a espectroscopia resonancia magnética nuclear (NMR) y/o cristalografía de rayos X, de los cuales la mejor coincidencia en base a identidad de secuencias estaba descrita como proteína de membra externa A (OmpA) con el templado 1g90.1.A, obteniéndose un 18,67% de identidad de secuencia, un QMEAN de -6,23 y un GMQE de 0,40. El PDB del templado fue 1g90, correspondiente al dominio transmembrana de la proteína de membrana externa A obtenida desde *E. coli* por espectroscopia NMR. Está estructura fue publicada por Arora, A., *et al.* (2001) y se describen las zonas de residuos aminoacídicos azules y naranjas mostradas en la Figura 7 como el fin extracelular de la estructura, compuesta de 4 loops largos no-estructurados (79). A través de la comparación de las secuencias marcadas como loops (Figura 7) y de las regiones conservadas obtenidas por alineamiento múltiple, se obtiene que las regiones conservadas corresponden a la estructura barril-beta de la proteína (región coloreada en gris) y que la región extracelular es la menos conservada dentro de los organismos analizados, por lo tanto, se puede desprender que son loops variables entre especies permitiendo una mejor especificidad para ser candidatas como péptidos antigénicos.

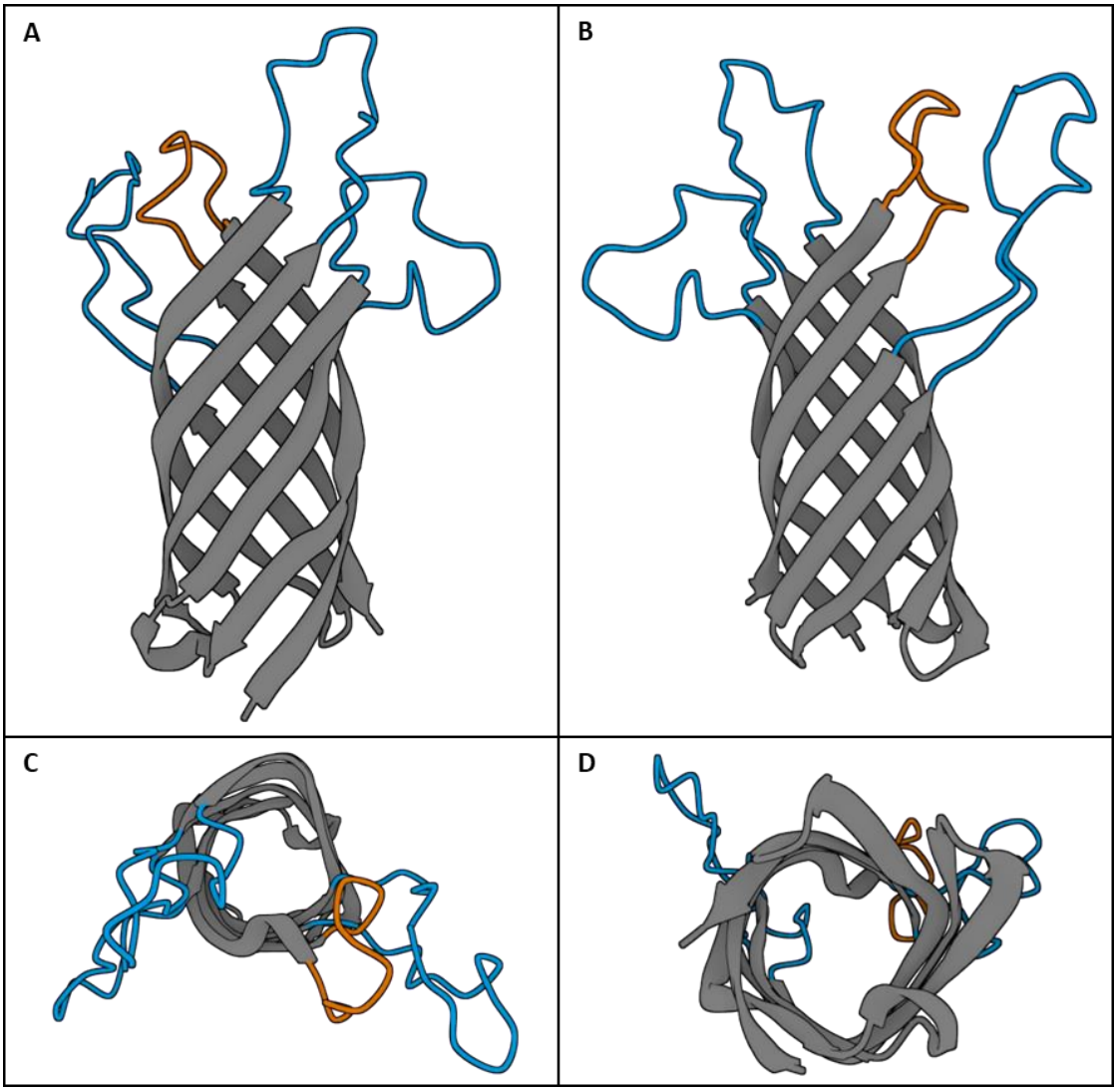

Figura 7. Segunda mejor coincidencia de modelo proteico predicho por SWISS-MODEL de la proteína barril-beta de membrana externa de *P. salmonis (*WP_016209892.1), visualizado mediante Mol* Viewer.
A, B, C y D, corresponden al mismo modelo proteico predicho visto desde diferentes ángulos.

4.1.3. Selección de péptido antigénico de superficie de *P. salmonis* y, diseño y síntesis de vector de clonamiento contiendo la secuencia nucleotídica del péptido

Se sintetizó un vector de clonamiento pUC57 conteniendo la secuencia nucleotídica de L1, correspondiente al péptido antigénico extraído de la región coloreada en naranjo de la Figura 7. La secuencia nucleotídica se obtuvo a través de la visualización del modelo proteico obtenido por SWISS-MODEL en el programa bioinformático Mol* Viewer. En la Figura 8 se observa el vector sintetizado, para el cual, utilizando los sitios de restricción para *BamH*I y *Kpn*I dentro de pUC57 se integró la

secuencia que contiene L1 con los mismos sitios de restricción en cada extremo (Figura 8-B) y se incluyó rio arriba de L1 un sitio de restricción para *Bgl*II con la secuencia codificante para un sitio TEV.

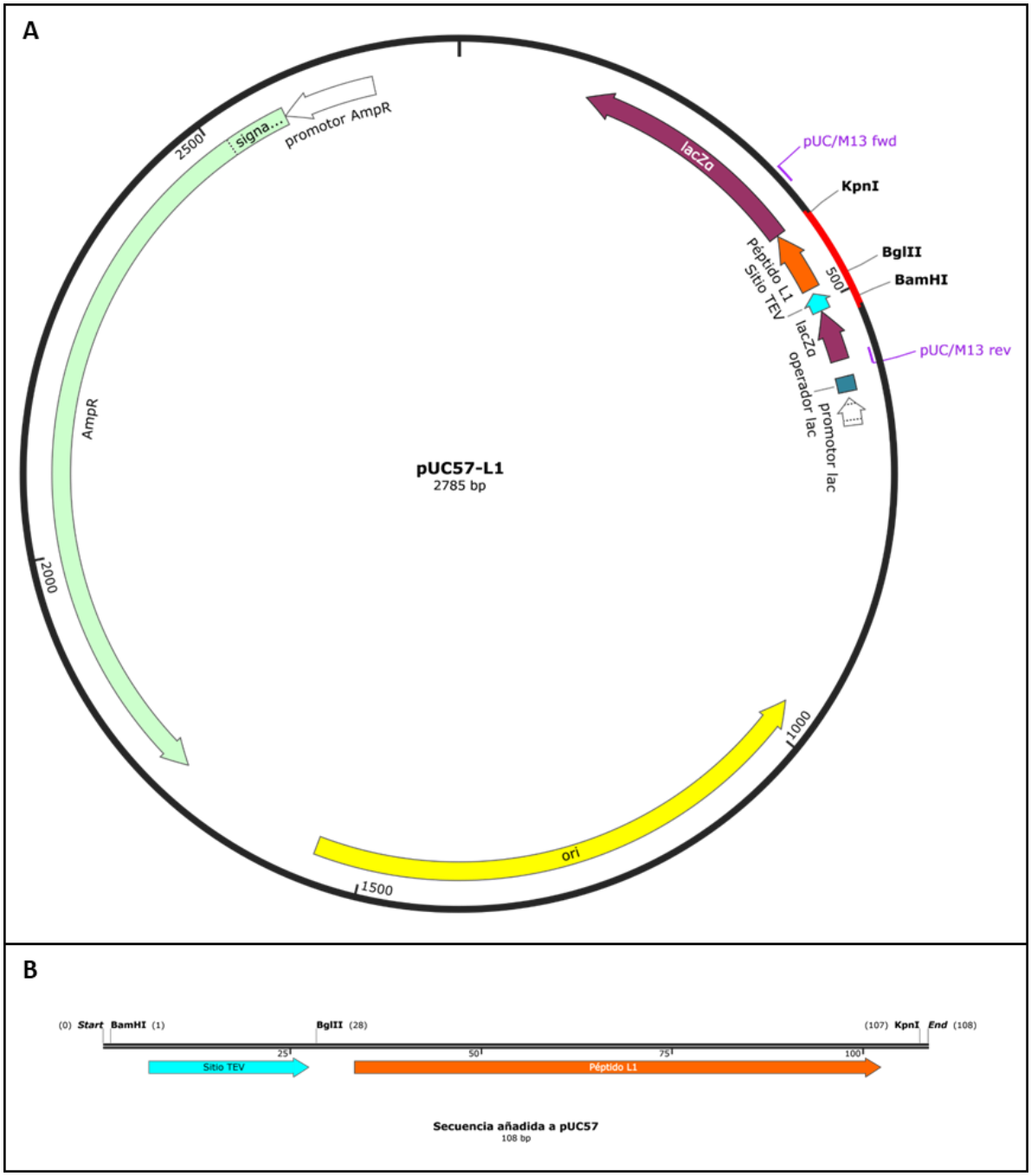

Figura 8. Mapa genético del vector de clonamiento pUC57 sintetizado conteniendo la secuencia nucleotídica del péptido L1.

A) Representación gráfica circular de pUC57-L1 mediante su visualización por la herramienta SnapGene de Dotmatics. En letras moradas se visualizan partidores específicos que hibridan en la secuencia nucleotídica del vector (pUC/M13 fwd y pUC/M13 rev), mientras que la región marcada en rojo del vector corresponde a la región diseñada y añadida al vector conteniendo la secuencia nucleotídica de L1.

B) Representación gráfica lineal de la región que contiene L1 y que fue diseñada y añadida a pUC57 para su síntesis mediante su visualización por la herramienta SnapGene de Dotmatics. Se resaltan los sitios de restricción *BamH*I, *Bgl*II y *Kpn*I en la zona superior de la representación que permiten su inserción en pUC57 y, un posterior manejo de las secuencias mediante herramientas de biología molecular.

4.2. Expresión de péptido antigénico de superficie L1 de *P. salmonis* mediante el subclonamiento de su secuencia nucleotídica en vector de expresión pFN18A

4.2.1. Subclonamiento del péptido antigénico L1 en vector de expresión pFN18A

Se realizó una transformación de *E. coli* XL10 con el plásmido previamente diseñado y sintetizado de pUC57-L1 mediante el protocolo de transformación de bacterias quimiocompetentes detallado en la metodología. Se obtuvieron varias colonias visiblemente separadas en las placas sembradas post transformación y cultivo bacteriano, de las cuales de una placa se seleccionaron 4 colonias distantes dentro de la placa para una corroboración de la presencia del plásmido dentro de las bacterias mediante PCR de colonias con los partidores pUC/M13 fwd y pUC/M13 rev. Se sometieron los resultados a electroforesis en gel de agarosa TAE (Figura 9), evidenciándose la presencia de bandas de peso molecular estimado entre 200 y 250 pares de bases según el marcador de peso molecular utilizado (Figura 9-A). Luego, se cultivaron en caldo las colonias seleccionadas para una posterior purificación del ADNp de interés, los cuales se identificaron mediante PCR, nuevamente con los partidores pUC/M13 fwd y pUC/M13 rev, para comprobar la presencia de pUC57-L1. Al visualizar mediante electroforesis en gel de agarosa TAE los resultados del PCR del ADNp purificado, se evidencian bandas correspondientes a un tamaño molecular estimado entre 200 y 250 pb (Figura 9-C), siendo de un tamaño similar al esperado de 223 pb del producto de PCR de pUC57-L1 con los partidores utilizados (Figura 9-B). Ambas bandas obtenidas en la electroforesis (Figura 9-C, carriles II y III) se purificaron y se eluyeron en una misma columna de purificación, buscando enriquecer la columna y obtener una mayor concentración de ADN del inserto L1. Los ADNp que se purificaron se denominaron como pUC57-L1B.

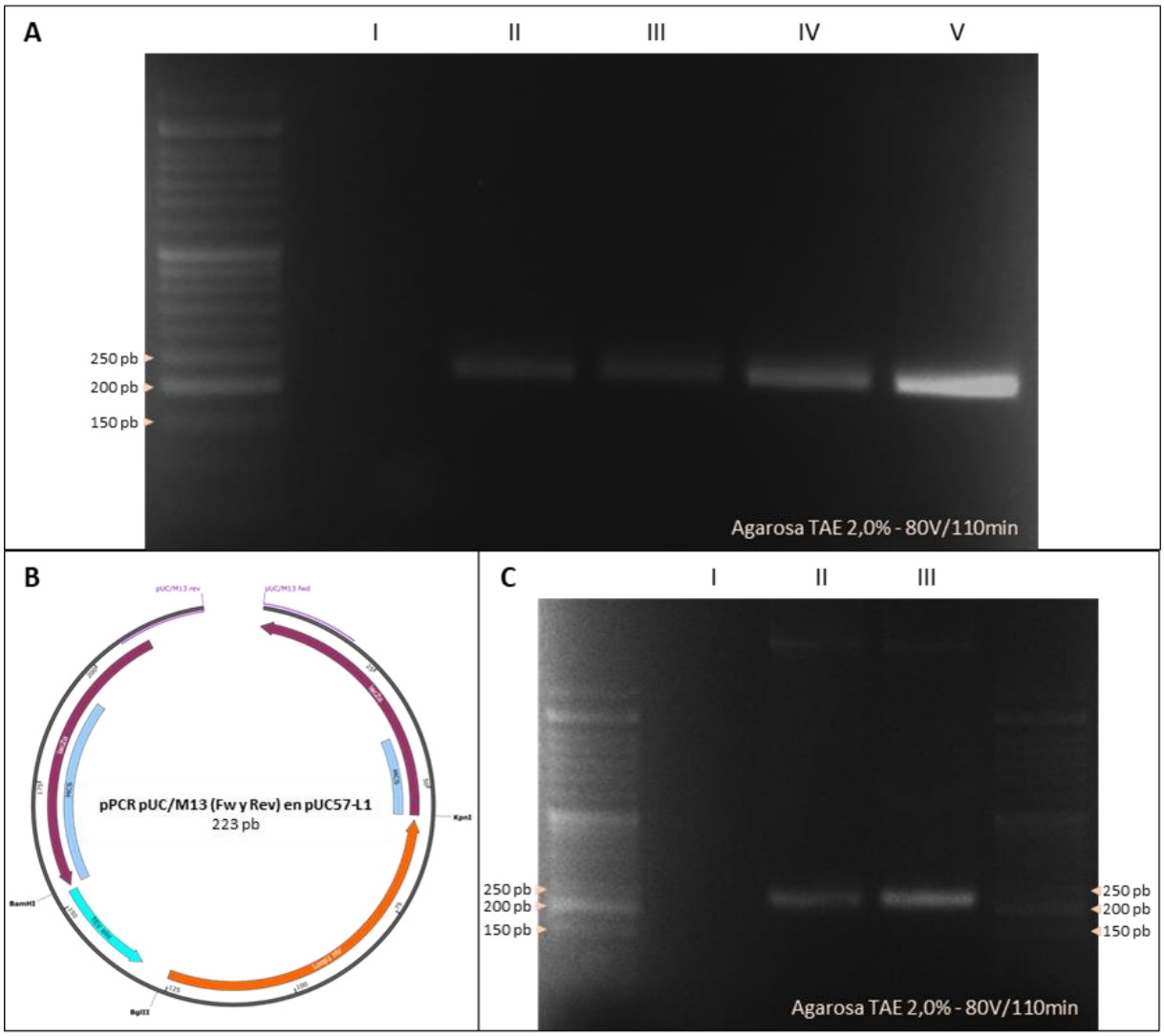

Figura 9. Resultados electroforesis en gel de agarosa TAE de los productos de PCR de colonia de XL10 pUC57-L1 y de los productos de PCR de pUC57-L1B.

A) Visualización con luz UV de electroforesis en gel de agarosa TAE 2,0%, utilizando el marcador de peso molecular en el primer pocillo del gel (de izquierda a derecha). A-I, corresponde al resultado de PCR del control sin templado (NTC); A-II, A-III, A-IV y A-V, corresponden a los resultados de PCR con templados de colonias de XL10 transformadas con pUC57-L1 utilizando los partidores pUC/M13 fwd y pUC/M13 rev.

B) Representación gráfica lineal del producto de PCR esperado de pUC57-L1 con los partidores pUC/M13 fwd y pUC/M13 rev, mediante su visualización por la herramienta SnapGene de Dotmatics. Se destaca el tamaño espero del producto de PCR de 223 pares de bases.

C) Visualización con luz azul de electroforesis en gel de agarosa TAE 2,0%, utilizando el marcador de peso en el primer pocillo del gel (de izquierda a derecha). A-I, corresponde al resultado de PCR del control sin templado (NTC); A-II y A-III, corresponden a los resultados de PCR con templados de pUC57-L1B utilizando los partidores pUC/M13 fwd y pUC/M13 rev.

Para la posterior expresión del péptido antigénico se seleccionó como vector de expresión a pFN18A, el cual, como principal característica, presenta la secuencia codificante de la proteína de fusión HaloTag (Figura 10), una herramienta importante para la purificación del péptido de interés. Se utilizó para la obtención del vector de expresión lineal sin inserto, el vector pFN18A $(L)_8$ realizando digestiones enzimáticas en los sitios de restricción para *BamH*I y *Kpn*I, manteniendo estos mismos sitios intactos.

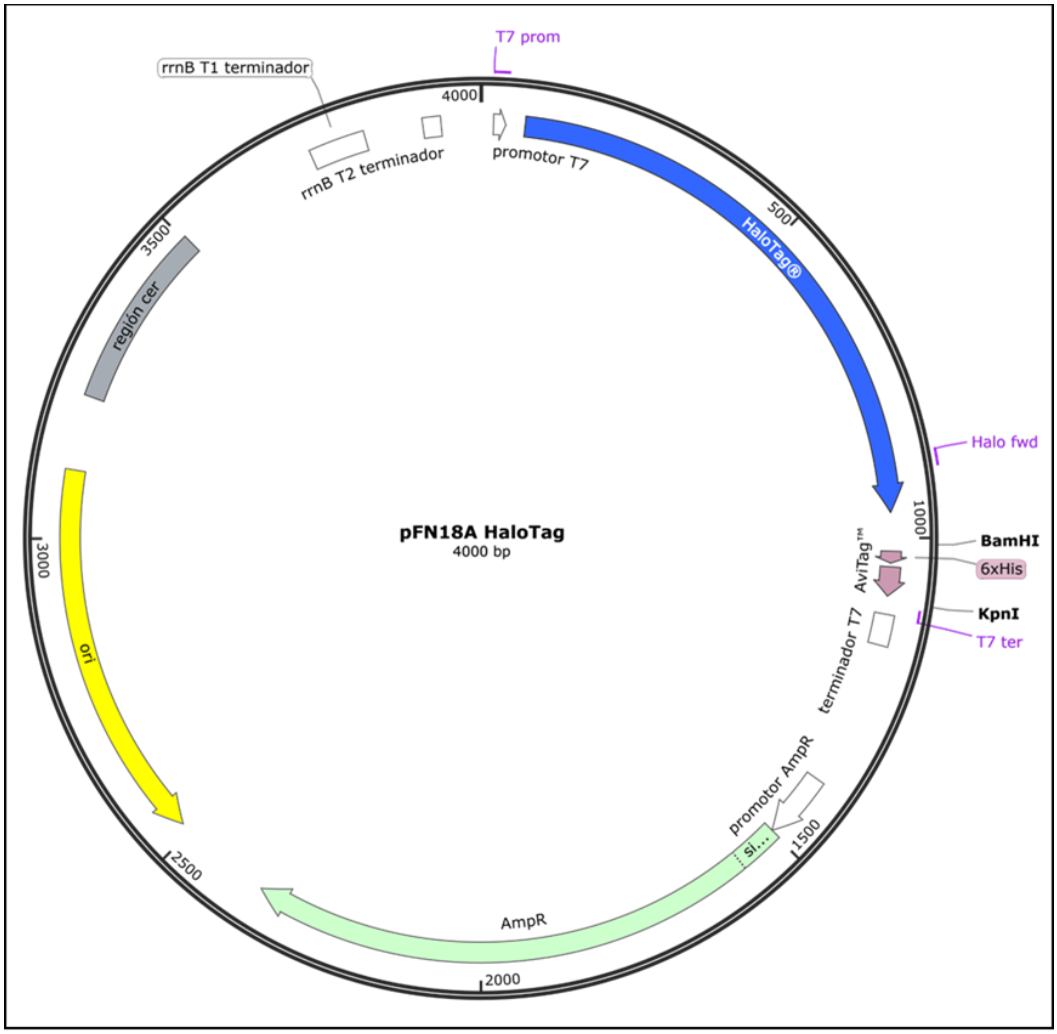

Figura 10. Mapa genético del vector de expresión pFN18A, conteniendo la secuencia codificante para HaloTag. Representación gráfica circular de pFN18A mediante su visualización por la herramienta SnapGene de Dotmatics. En letras moradas se visualizan partidores específicos que hibridan en la secuencia nucleotídica del vector (Halo fwd y T7 ter), región de secuencia nucleotídica donde se posicionaría la secuencia de L1 post-subclonamiento. Mientras que la flecha marcada en azul se observa la región codificante para la proteína de fusión HaloTag con el promotor T7 rio arriba de ella.

Para la obtención de pFN18A sin inserto, se realizaron 3 ensayos de digestiones enzimáticas reflejándose los resultados en la Figura 11. En primer lugar, se hicieron digestiones enzimáticas dobles

de pFN18A $(L)_8$ con las enzimas *BamH*I y *Kpn*I, durante 40 minutos (Figura 11-C) y 18 horas (Figura 11-B), en donde no se obtuvieron los resultados esperados de liberación de vector e inserto. Finalmente, se realizaron digestiones durante 18 horas, pero utilizando enzimas de digestión de alta fidelidad (Figura 11-A), en donde se obtuvo un inserto liberado de 1,5 Kb a 2,0 Kb y, además, el vector sin inserto de ~4,0 Kb (3.916 pb esperadas) reflejándose en la Figuras 11-A-IV y 11-A-VII. Posteriormente, se purificaron las bandas correspondientes al vector sin inserto y se eluyeron ambas bandas en una misma columna de purificación, esperando enriquecer la columna y obtener una mayor concentración del purificado de ADN.

Además, para la obtención del inserto con la secuencia nucleotídica de L1 para un posterior subclonamiento en el vector de expresión, se realizaron digestiones simples del producto obtenido de la purificación en banda de lo resultante en la Figura 11-C, primero con *Bgl*II y su buffer de digestión correspondiente, ADN digerido que fue purificado para luego realizar la digestión con la enzima *Kpn*I y su buffer correspondiente. El resultado del ADN digerido con ambas enzimas fue purificado y almacenado para la posterior ligación enzimática con el vector de expresión pFN18A sin inserto.

Los ADN resultantes de la purificación de los productos de digestión con las enzimas, por un lado, *Bgl*II y *Kpn*I del producto de PCR pUC57-L1B con los partidores pUC/M13 fwd y pUC/M13 rev y, por otro lado, *BamH*I y *Kpn*I del producto de la digestión enzimática de pFN18A $(L)_8$, se sometieron a ligación con la enzima T4 esperándose un vector de expresión pFN18A con un inserto conformado por la secuencia nucleotídica del péptido L1 (Figura 12), denominado pFN18A-L1.

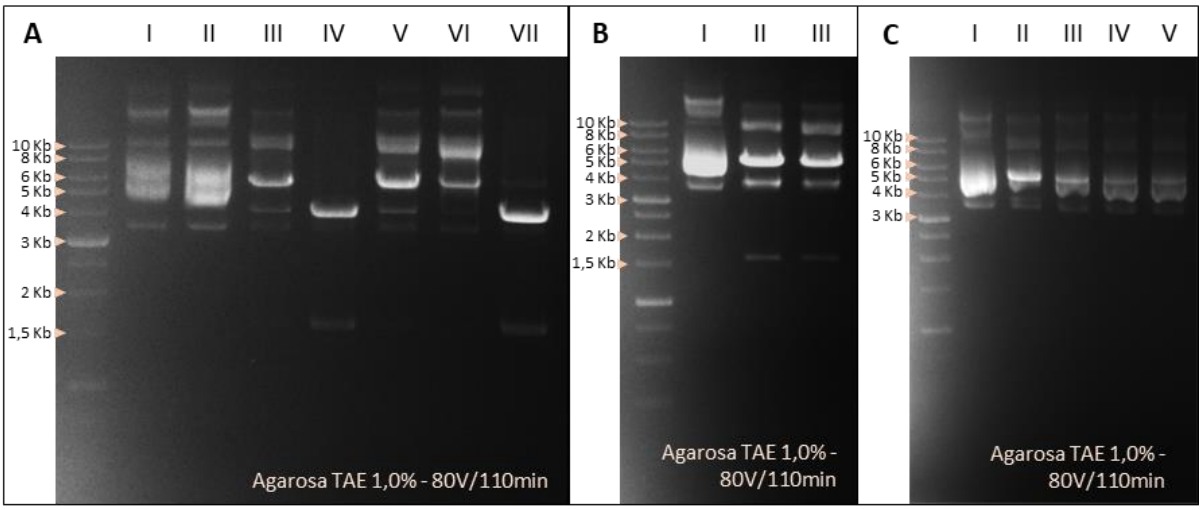

Figura 11. Resultados electroforesis en gel de agarosa TAE 1,0% de los productos de las digestiones enzimáticas de pFN18A $(L)_8$ para obtención del vector de expresión sin inserto.

A) Visualización con luz UV del ensayo de digestiones enzimáticas de pFN18A $(L)_8$ a 37°C por 18 horas comparando enzimas de alta fidelidad (HF) con enzimas sin alta fidelidad a diferentes volúmenes de reacción de digestión, utilizando el marcador de peso en el primer pocillo del gel (de izquierda a derecha). A-I y A-II, controles sin digerir de 1.000 ng y 2.000 ng de pFN18A $(L)_8$, respectivamente; A-III y A-VI, resultado digestión doble con 10 U de *BamH*I y 10 U de *Kpn*I en 1.000 ng y 2.000 ng de pFN18A $(L)_8$, respectivamente; A-IV, resultado digestión doble con 10 U de *BamH*I-HF y 10 U de *Kpn*I-HF en 1.000 ng de pFN18A $(L)_8$; A-V, resultado digestión doble con 20 U de *BamH*I y 20 U de *Kpn*I en 2.000 ng de pFN18A $(L)_8$; y A-VII, resultado digestión doble con 20 U de *BamH*I-HF y 20 U de *Kpn*I-HF en 2.000 ng pFN18A $(L)_8$.

B) Visualización con luz UV de la digestión enzimática doble de pFN18A $(L)_8$ a 37°C por 18 horas con enzimas sin alta fidelidad, utilizando el marcador de peso molecular en el primer pocillo del gel (de izquierda a derecha). B-I, control sin digerir de 1.000 ng de pFN18A $(L)_8$; B-II y B-III, resultados digestiones dobles con 10 U de *BamH*I y 10 U de *Kpn*I en 1.000 ng de pFN18A $(L)_8$.

C) Visualización con luz UV de la digestión enzimática doble de pFN18A $(L)_8$ a 37°C por 40 minutos con enzimas sin alta fidelidad, utilizando el marcador de peso molecular en el primer pocillo del gel (de izquierda a derecha). C-I, control sin digerir de 1.000 ng de pFN18A $(L)_8$; C-II, C-III, C-IV y C-V, resultados digestiones dobles con 10 U de *BamH*I y 10 U de *Kpn*I en 1.000 ng de pFN18A $(L)_8$.

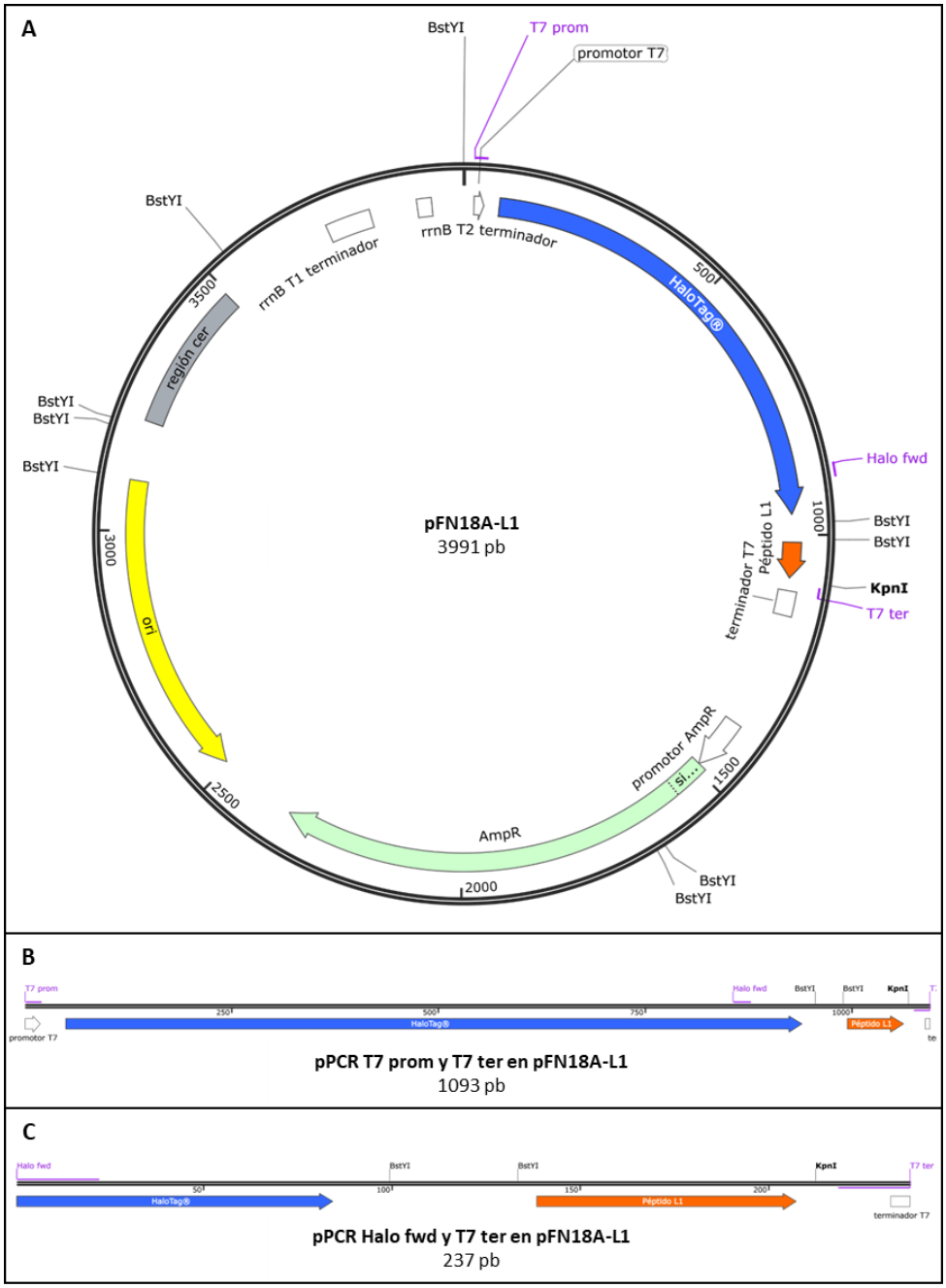

Figura 12. Mapa genético del vector de expresión pFN18A-L1, vector de expresión del péptido antigénico L1.
A) Representación gráfica circular de pFN18A-L1 mediante su visualización por la herramienta SnapGene de Dotmatics. En letras moradas se visualizan partidores específicos que hibridan en la secuencia nucleotídica del vector (Halo fwd y T7 ter), región de secuencia nucleotídica donde se posiciona la secuencia del péptido antigénico L1. Mientras que la flecha marcada en azul se observa la región codificante para la proteína de fusión HaloTag con el promotor T7 rio arriba de ella y rio abajo se encuentra la región codificante para L1, marcado como una flecha naranja.
B) Representación gráfica lineal del producto de PCR de pFN18A-L1 con los partidores T7 prom y T7 ter, mediante su visualización por la herramienta SnapGene de Dotmatics.
C) Representación gráfica lineal del producto de PCR de pFN18A-L1 con los partidores Halo fwd y T7 ter, mediante su visualización por la herramienta SnapGene de Dotmatics.

Posterior a la ligación enzimática y purificación del vector resultante, se transformaron bacterias quimiocompetentes *E. coli* XL10 con pFN18A-L1, en donde se obtuvieron varias colonias visiblemente separadas en las placas sembradas post transformación y cultivo bacteriano, de las cuales se seleccionaron 10 colonias desde dos placas para una corroboración de la presencia del plásmido dentro de las bacterias mediante PCR de colonias con los partidores T7 prom y T7 ter. Se visualizaron los resultados del PCR de colonias a través de electroforesis en gel de agarosa TAE presentados en la Figura 13 (solo se presentan 11 resultados de los 20 PCR de colonia realizados), evidenciándose la presencia de bandas de peso molecular estimado entre 1.000 y 1.200 pares de bases según el marcador de peso molecular utilizado (Figura 13-A), lo que corrobora el peso de la banda esperada del producto de PCR según la simulación *in silico* de este (Figura 12-B). Luego, se cultivaron en caldo las colonias seleccionadas para una posterior purificación del ADNp de interés, los cuales se sometieron a identificación mediante PCR, nuevamente con los partidores T7 prom y pUC/M13 ter, para comprobar la presencia de pFN18A-L1. Al visualizar mediante electroforesis en gel de agarosa TAE los resultados del PCR del ADNp purificado, se evidencian bandas correspondientes a un tamaño molecular estimado entre 1.000 y 1.500 pb (Figura 13-B), correspondiendo al tamaño estimado de 1.093 pb del producto de PCR de pFN18A-L1 con los partidores utilizados (Figura 12-B).

Finalmente, se visualizaron mediante electroforesis en gel de agarosa TAE los ADNp obtenidos en los pasos del subclonamiento para comprobar la integridad de los ADN, tomando en consideración las concentraciones idóneas de carga en gel obtenidas mediante el ensayo de diluciones de ADNp (Figura S1). Por lo tanto, en la Figura 14 se puede apreciar la integridad de pFN18A-L1 y de pUC57-L1B. Además, como comprobante del éxito del subclonamiento se secuenciaron los ADNp purificados de 5 colonias de *E. coli* XL10 pFN18A-L1, mediante un servicio de secuenciación externo, obteniéndose una secuencia consenso que tiene el 100% de coincidencia con la secuencia del péptido antigénico L1 a través de sus alineamientos múltiples por Clustal Omega de los resultados de la secuenciación con los partidores Halo fwd y T7 ter, cada uno independientemente, incluyendo los sitios de restricción esperados intactos en la secuencia post-subclonamiento dentro de pFN18A-L1.

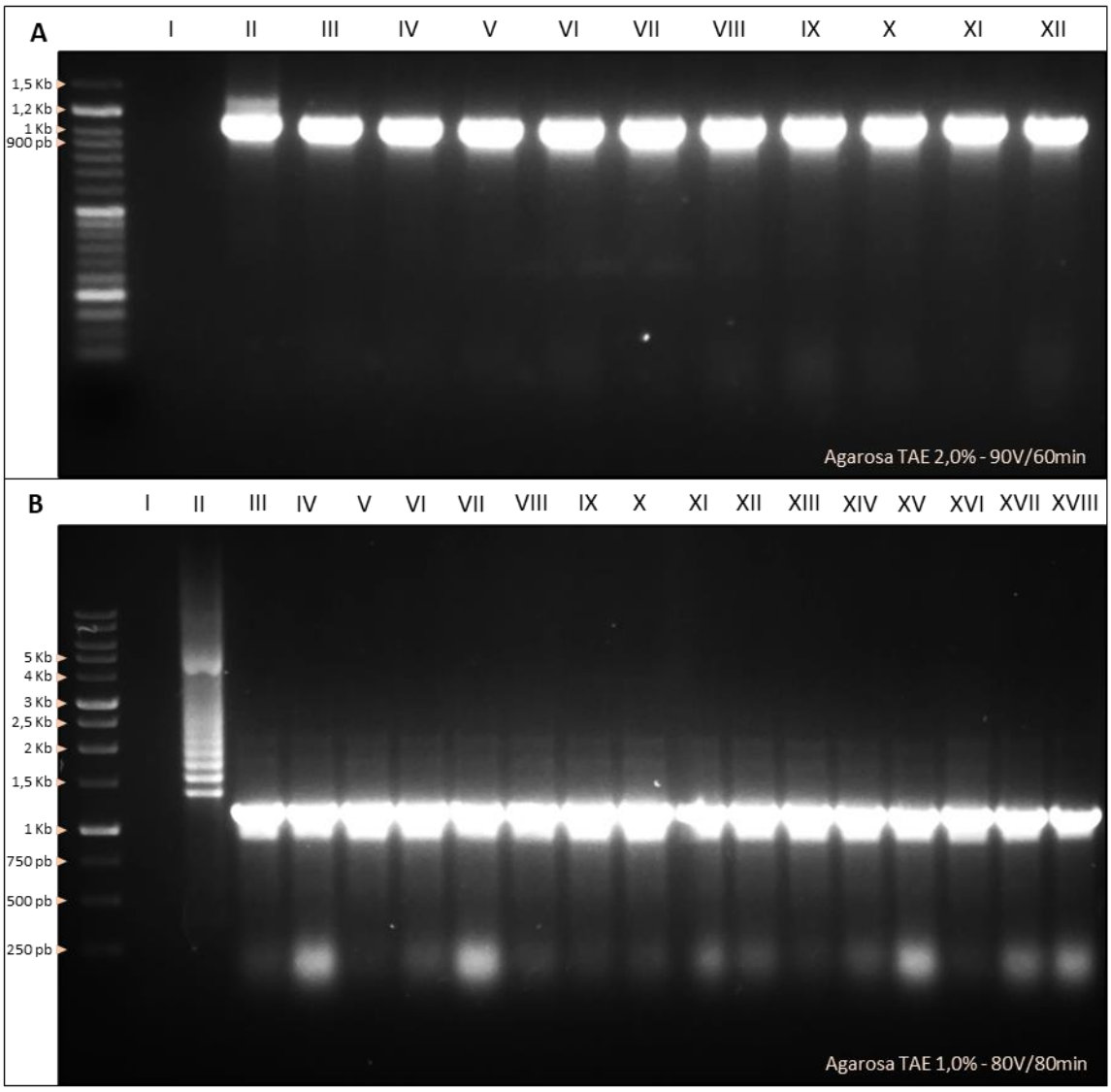

Figura 13. Resultados electroforesis en gel de agarosa TAE de los productos de PCR de colonia de XL10 pFN18A-L1 y de los productos de PCR de pFN18A-L1.

A) Visualización con luz UV de electroforesis en gel de agarosa TAE 2,0%, utilizando el marcador de peso molecular en el primer pocillo del gel (de izquierda a derecha). A-I, corresponde al resultado de PCR del control sin templado (NTC); A-II a A-XII, corresponden a los resultados de PCR con templados de colonias de *XL10* transformadas con pFN18A-L1 utilizando los partidores T7 prom y T7 ter.

B) Visualización con luz UV de electroforesis en gel de agarosa TAE 1,0%, utilizando el marcador de peso molecular en el primer pocillo del gel (de izquierda a derecha). A-I, corresponde al resultado de PCR del control sin templado (NTC); A-II, corresponde al resultado de PCR con los partidores T7 prom y T7 ter utilizando como templado a pFN18A $(L)_8$ como control negativo; A-III a A-XVIII, corresponden a los resultados de PCR con templados de pFN18A-L1 utilizando los partidores T7 prom y T7 ter.

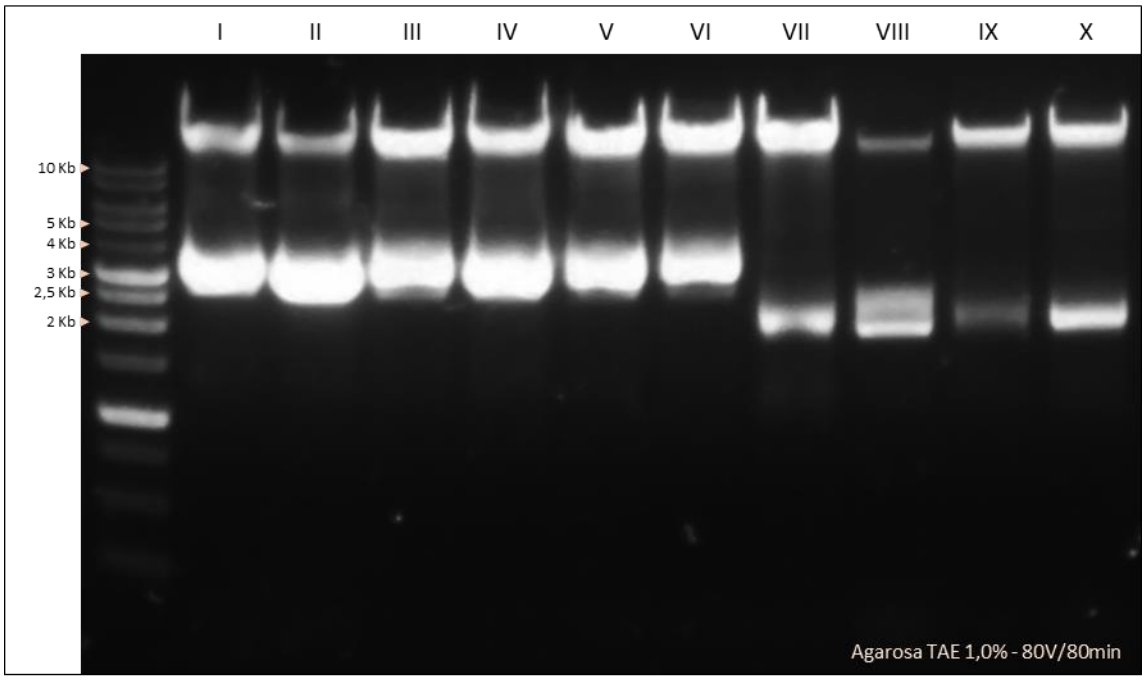

Figura 14. Resultados electroforesis en gel de agarosa TAE de la corroboración de integridad de pFN18A-L1 y pUC57-L1B.
Visualización con luz UV de electroforesis en gel de agarosa TAE 1,0%, utilizando el marcador de peso molecular en el primer pocillo del gel (de izquierda a derecha). I a VI, pFNA18A-L1 purificados posterior subclonamiento; VII a X, pUC57-L1 purificados posterior transformación de *E. coli* XL10 con pUC57-L1.

4.2.2. Ensayos de expresión de Halo-péptido y selección cepa de expresión para péptido antigénico L1

Teniendo comprobada la integridad de la secuencia nucleotídica del péptido antigénico L1 dentro del vector de expresión pFN18A, se procedió a transformar con pFN18A-L1 las cepas de expresión BLR (DE3) pLysS y C41 (DE3), con el fin de evaluar la expresión del péptido en conjunto con la proteína de fusión HaloTag. En primer lugar, se realizó un ensayo de expresión BLR (DE3) pLysS pFN18A-L1 comparando las temperaturas de cultivo de 20°C y 37°C, a diferentes tiempos de cultivo post-inducción con IPTG 1 mM (Figura 15). En la mayoría de los casos, se observaron dos bandas fluorescentes claras en la fotografía del gel expuesto a luz azul. Una de estas bandas era más intensa y tenía un peso molecular estimado de 40 kDa, mientras que la otra era más tenue y tenía un peso molecular mayor, pero menor a 55 kDa.

En general, se encontró una leve expresión basal en el cultivo sin inducción tanto en el inicio (T0) como a las 18 horas de cultivo. Sin embargo, se observó una mayor expresión del péptido a 20°C, alcanzando un máximo a las 18 horas de cultivo (Figura 15, T3-20 imagen inferior).

Se llevó a cabo un segundo ensayo de expresión de BLR (DE3) pLysS pFN18A-L1, esta vez comparando controles sin inducción en todos los tiempos evaluados a 20°C (Figura 16). Los resultados fueron similares al primer ensayo, mostrando una expresión basal evidente en todos los tiempos sin inducción, excepto a las 18 horas (T3) con inducción mediante IPTG, donde se observó una banda fluorescente fuerte de aproximadamente 40 kDa de peso molecular.

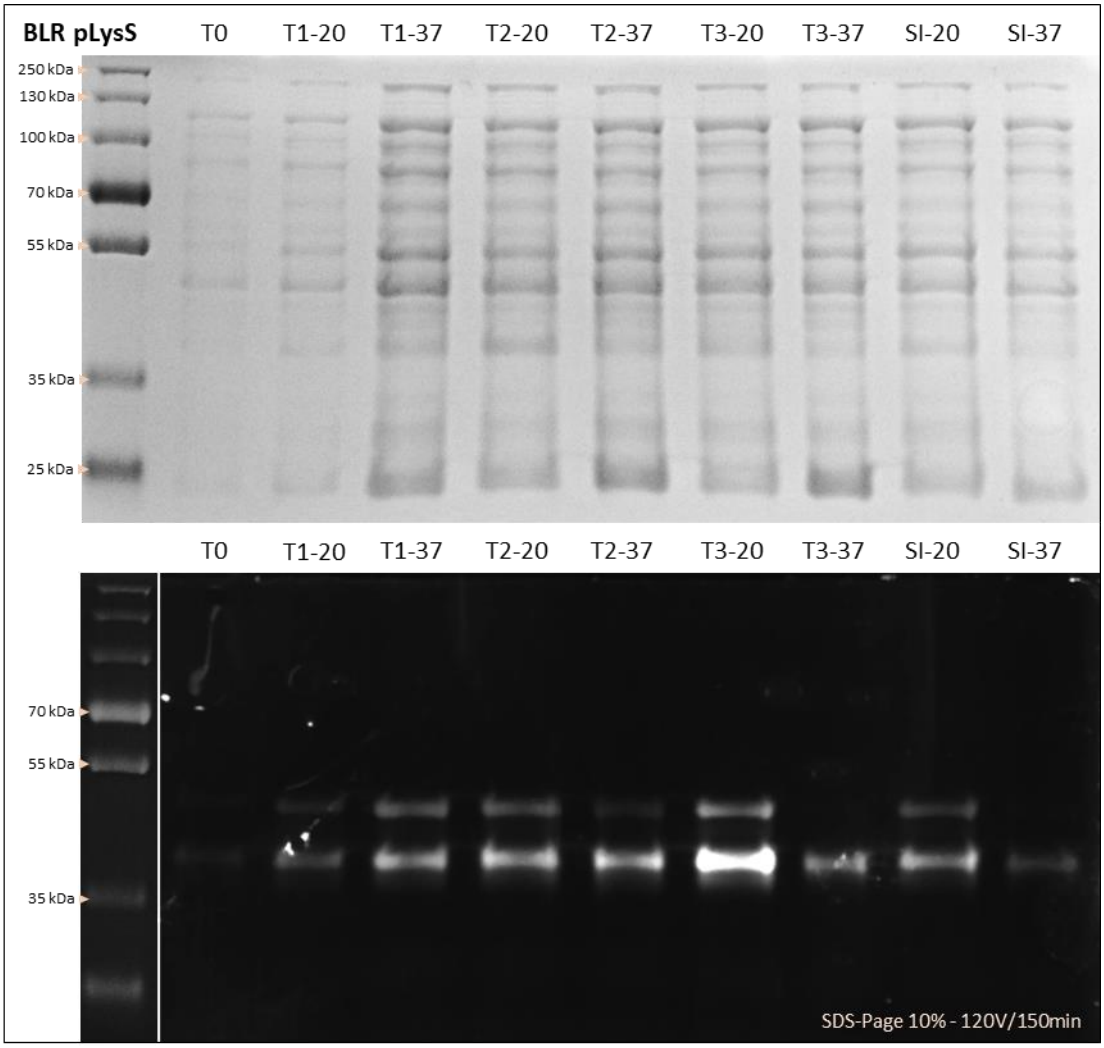

Figura 15. Resultados SDS-PAGE del ensayo de expresión de BLR (DE3) pLysS pFN18A-L1 a 20°C y 37°C. Visualización con luz blanca posterior a tinción con azul de Coomassie (fotografía superior) y luz azul previo-tinción (fotografía inferior) de SDS-PAGE, utilizando el marcador de peso molecular en el primer pocillo del gel (de izquierda a derecha). T0, testigo del cultivo a tiempo 0 posterior al inoculo inicial; T1, testigo del cultivo a las 1 hora post-inducción con IPTG 1 mM; T2, testigo del cultivo a las 3 horas post-inducción con IPTG 1 mM; T3, testigo del cultivo a las 18 horas post-inducción con IPTG 1 mM; SI, testigo del cultivo sin inducción a las 18 horas posterior al inoculo inicial; las temperaturas de cultivo son representadas en la figura con 20 y 37 posterior al tiempo de cultivo para 20°C y 37°C, respectivamente.

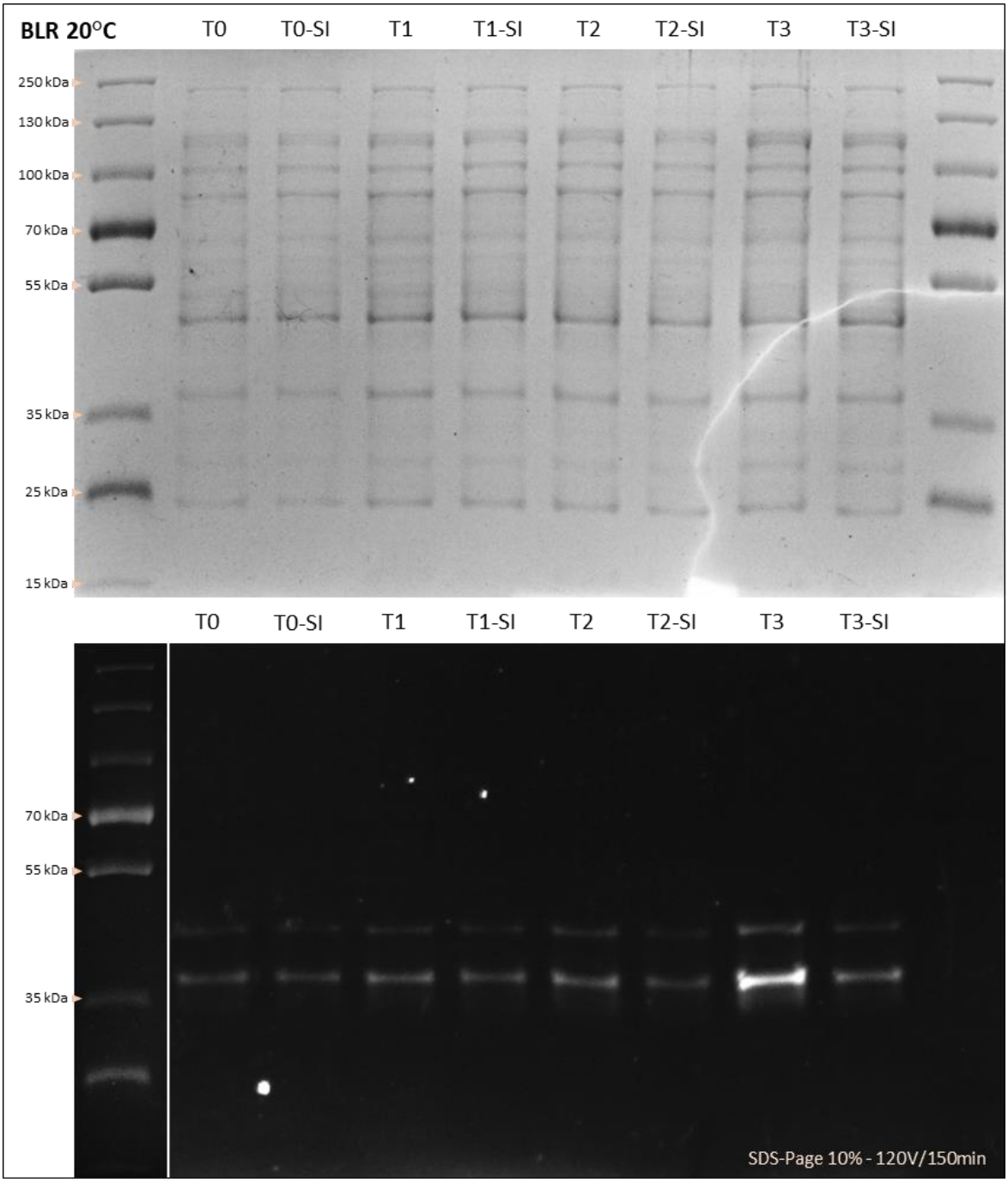

Figura 16. Resultados SDS-PAGE del ensayo de expresión de BLR (DE3) pLysS pFN18A-L1 a 20°C.
Visualización con luz blanca posterior a tinción con azul de Coomassie (fotografía superior) y luz azul previo-tinción (fotografía inferior) de SDS-PAGE, utilizando el marcador de peso molecular en el primer pocillo del gel (de izquierda a derecha). T0, testigo del cultivo a tiempo 0 posterior al inoculo inicial; T1, testigo del cultivo a las 1 hora post-inducción con IPTG 1 mM; T2, testigo del cultivo a las 3 horas post-inducción con IPTG 1 mM; T3, testigo del cultivo a las 18 horas post-inducción con IPTG 1 mM; SI, testigo del cultivo sin inducción al tiempo correspondiente.

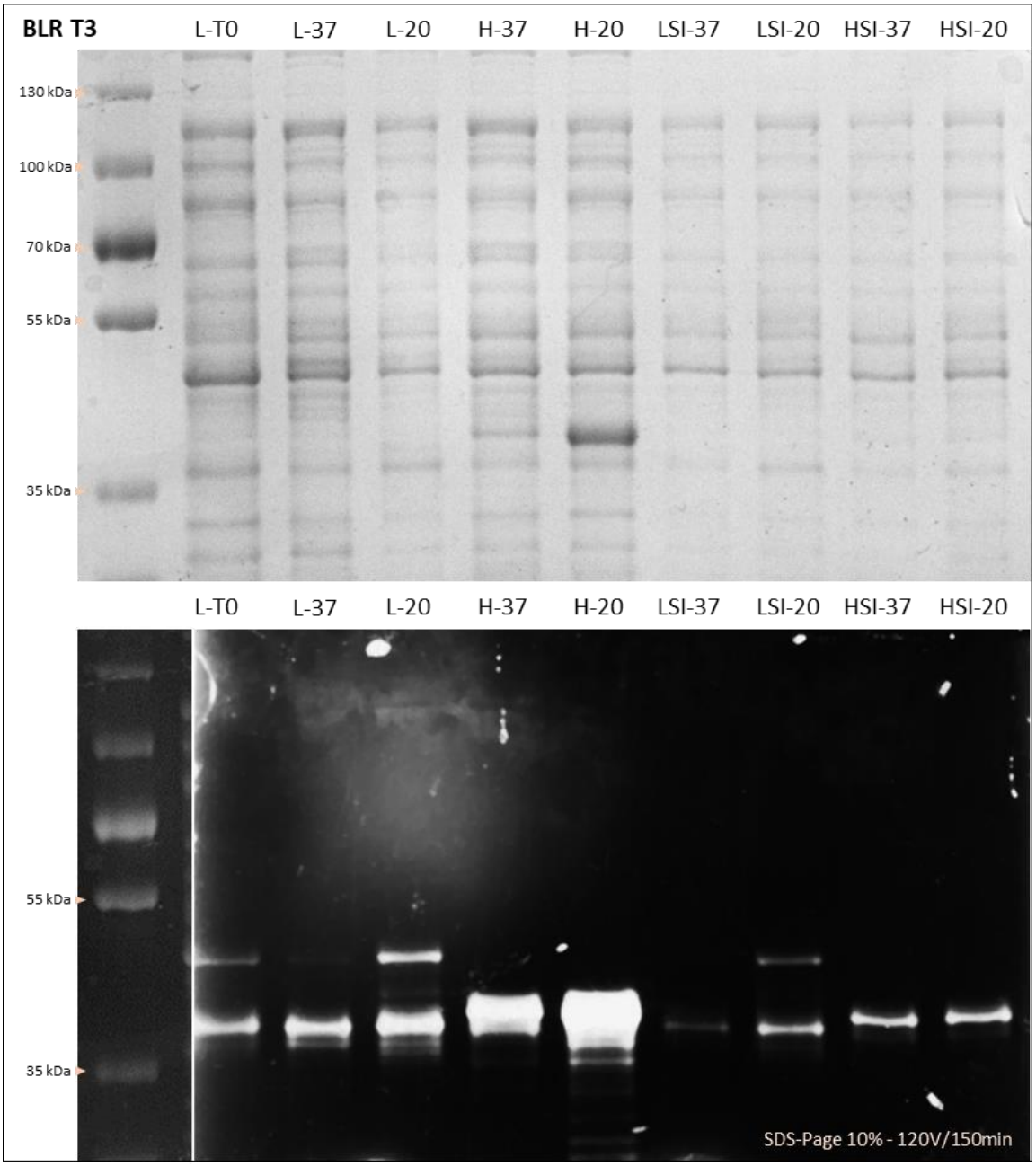

Figura 17. Resultados SDS-PAGE del ensayo de expresión de BLR (DE3) pLysS pFN18A-Halo y de BLR (DE3) pLysS pFN18A-L1 a 37°C y 20°C.
Visualización con luz blanca posterior a tinción con azul de Coomassie (fotografía superior) y luz azul previo-tinción (fotografía inferior) de SDS-PAGE, utilizando el marcador de peso molecular en el primer pocillo del gel (de izquierda a derecha). Las cepas bacterianas con cada plásmido son representadas en la figura con L y H, para BLR(DE3) pLysS pFN18A-L1 y de BLR(DE3) pLysS pFN18A-Halo, respectivamente. T0, testigo del cultivo a tiempo 0 posterior al inoculo inicial (el resto de los testigos se tomaron después de las 18 horas de cultivo posterior al inoculo inicial); SI, testigo del cultivo sin inducción (el resto de los testigos se indujeron con IPTG 1mM); las temperaturas de cultivo son representadas en la figura con 20 y 37 posterior al tiempo de cultivo para 20°C y 37°C, respectivamente.

Finalmente, se llevó a cabo una evaluación comparativa de las FPE más importantes de los primeros dos ensayos, tomando como control la FPE de BLR (DE3) pLysS Halo, que corresponde a la expresión del HaloTag por sí solo sin el péptido antigénico (Figura 17). Se obtuvo una mayor expresión de HaloTag por sí solo a 20°C y se pudo inferir que la banda superior observada en la visualización con luz azul correspondía a la expresión del HaloTag junto con el péptido antigénico, mientras que la banda de aproximadamente 40 kDa en los primeros ensayos pertenecía al HaloTag por sí solo.

Por otro lado, para comparar la expresión con otra cepa *de E. coli*, se realizó un nuevo ensayo de expresión, utilizando la cepa de expresión C41 (DE3) en lugar de BLR (DE3) (Figura 18). Se obtuvieron valores similares a los obtenidos con BLR (DE3), pero con una mayor expresión tanto de HaloTag por sí solo como de HaloTag junto con el péptido antigénico, al comparar con el control de expresión de BLR (DE3) pLysS pFN18A-L1 (Figura 18, T3-H). Además, se observó un mayor rendimiento de expresión del péptido antigénico a las 3 horas de expresión a 20°C (Figura 18, T2). Estos resultados también se apreciaron después de la tinción con azul de Coomassie del gel y la visualización con luz blanca, donde se observaron bandas evidentes en las FPE correspondientes a los pesos moleculares de las mismas bandas visualizadas por fluorescencia.

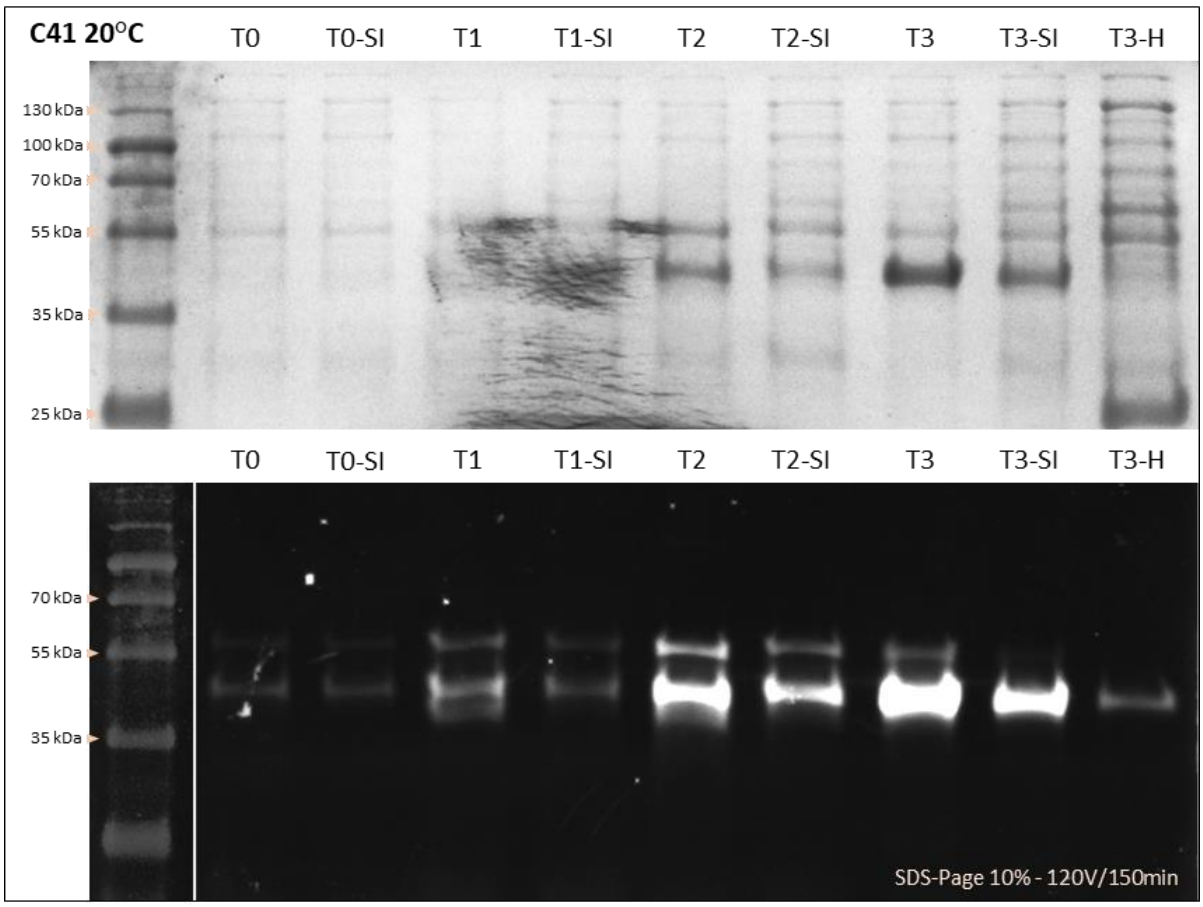

Figura 18. Resultados SDS-PAGE del ensayo de expresión de C41 pFN18A-L1 a 20°C.
Visualización con luz blanca posterior a tinción con azul de Coomassie (fotografía superior) y luz azul previo-tinción (fotografía inferior) de SDS-PAGE, utilizando el marcador de peso molecular en el primer pocillo del gel (de izquierda a derecha). T0, testigo del cultivo a tiempo 0 posterior al inoculo inicial; T1, testigo del cultivo a las 1 hora post-inducción con IPTG 1 mM; T2, testigo del cultivo a las 3 horas post-inducción con IPTG 1 mM; T3, testigo del cultivo a las 18 horas post-inducción con IPTG 1 mM; SI, testigo del cultivo sin inducción a las 18 horas posterior al inoculo inicial; H, control de expresión de Halo sin el péptido L1 mediante el cultivo de BLR(DE3) pLysS pFN18A-Halo a 20°C y el tiempo correspondiente.

5. DISCUSIÓN DE RESULTADOS

Uno de los principales desafíos para lograr una inmunización eficiente contra bacterias es su alta diversidad genética. Esta diversidad se atribuye a tasas elevadas de mutación y eventos de recombinación, junto con un mecanismo de reparación de ADN poco eficiente y una alta capacidad para la transferencia genética. Como consecuencia, las poblaciones bacterianas pueden evolucionar y adaptarse específicamente a su entorno en cada hospedero, incluso en respuesta a tratamientos con antibióticos (80). Por lo tanto, se requiere el desarrollo de vacunas que abarquen esta amplia y dinámica diversidad bacteriana. Los antígenos seleccionados para el desarrollo de vacunas contra patógenos bacterianos, usualmente son proteínas que corresponden a factores de virulencia con potencial inmunogénico. Una alta inmunogenicidad resultará en una fuerte reacción inmune, donde la neutralización de los factores de virulencia dificultará el proceso de infección bacteriana (81). Lo anterior es una de las razones que ha orientado la revisión bibliográfica realizada como primera parte de esta tesis, hacia la búsqueda de proteínas que fueran abundantes durante el ciclo infectivo desarrollado por los organismos filogenéticamente cercanos a *P. salmonis* y, además, relevantes en procesos de infección y patogenicidad de estos organismos.

La ubicación subcelular de una proteína está estrechamente correlacionada con su función biológica (82). Ciertas proteínas de membrana desempeñan roles vitales en muchos procesos celulares, por ejemplo, las proteínas receptoras de membrana transmiten señales entre el entorno interno y externo, las proteínas de transporte mueven moléculas e iones a través de la membrana, y las moléculas de adhesión celular, como las proteínas involucradas en la respuesta inmunitaria, permiten a las células identificarse e interactuar entre sí (83). Asimismo, estas proteínas desempeñan un papel fundamental en la infección de las células del hospedero, ya que están implicadas en la evasión de las respuestas del sistema inmunitario y en la inducción de toxicidad. Por consiguiente, ciertas proteínas de superficie son consideradas candidatas prometedoras para el desarrollo de vacunas (44). En nuestro trabajo, al momento de realizar la identificación y selección por revisión bibliográfica de proteínas abundantes en contexto de infección de organismos filogenéticamente cercanos a *P. salmonis*, las proteínas candidatas se filtraron de acuerdo a su ubicación subcelular, logrando obtener solo proteínas localizadas en la membrana celular externa de la bacteria y posiblemente expuestas al medio extracelular y, así, poder trabajar con proteínas involucradas en procesos patogénicos y que fuesen de fácil acceso para su detección en futuros métodos de diagnóstico.

Una de las principales limitaciones para nuestro diseño racional de péptido inmunogénicos de *P. salmonis* ha sido la ausencia de una anotación de secuencias de proteínas de esta bacteria en bases de datos públicas, lo que nos ha llevado a que nuestra estrategia para la identificación de proteínas de membrana, expuestas y abundantes en contexto de infección en *P. salmonis*, necesite la identificación de proteínas homólogas en bacterias filogenéticamente relacionadas y cuyas secuencias puedan ser identificadas en las anotaciones genómicas de esta bacteria.

Para lograr lo anterior, realizamos alineamientos locales de las secuencias aminoacídicas de las proteínas candidatas en las bacterias filogenéticamente relacionadas con *P. salmonis*, previamente identificadas en la búsqueda bibliográfica, las cuales se analizaron mediante la herramienta BLAST para predecir posibles proteínas homólogas en *P. salmonis* y, así, predecir proteínas involucradas en sus procesos patogénicos, aunque no estén descritas del todo, y correlacionar que son proteínas abundantes en contexto de infección de esta bacteria. Aun así, aunque se busquen proteínas homólogas, para generar antígenos que puedan originar especificidad en un posterior proceso de generación de anticuerpos, es importante identificar cierta conservación de la proteína en *P. salmonis* con relación al resto de los organismos, sobre todo del entorno en donde se desempeña esta bacteria.

En general, las evidencias de estudios en ratones indican que los anticuerpos contra los antígenos conservados son altamente efectivos contra múltiples cepas de un patógeno, en comparación a los anticuerpos contra los antígenos externos en constante cambio (84). Por lo tanto, buscar proteínas o péptidos más conservados entre especies podría ser una ventaja para la posterior generación de anticuerpos de uso diagnostico o terapéutico contra SRS. Y es por esto, que se realizó un filtro por conservación de las proteínas candidatas a través de alineamientos locales de secuencias por BLAST de ellas contra los organismos filogenéticamente cercanos y, en paralelo, contra toda la base de datos excluyendo los organismos filogenéticamente cercanos a *P. salmonis*. Obteniéndose una cantidad más reducida de proteínas candidatas a evaluar en un posterior análisis en base a modelamiento proteico y búsqueda de regiones expuestas al medio extracelular.

Al momento de evaluar la calidad de un modelamiento de proteína como los realizados mediante SWISS-MODEL, se puede predecir a través de la similitud de secuencias entre el objetivo y el templado. Una identidad de secuencia entre ambos por encima del 30% es un indicador bueno de la precisión esperada en un modelo (85). Para el caso de las proteínas modeladas mediante SWISS-MODEL utilizando como templado proteínas obtenidas mediante la tecnología de AlphaFold, se lograron

identidades de secuencias mayores al 96% y coberturas de 100% para las tres proteínas sometidas a este modelamiento (Figuras 3, 4 y 5), por lo tanto, se puede inferir con estos resultados que los modelamientos realizados son confiables en relación a la identidad de secuencias, tomando en consideración que la tecnología de predicción de proteínas de AlphaFold no está completamente resuelta en términos de predecir la posición relativa de los dominios dentro de una cadena, como los dominios pueden cambiar su conformación relativa en respuesta a estímulos o como estos dominios hacen cambios en la transición de un estado desordenado a uno ordenado (86). Aun así, para motivos de este estudio, solo se necesita predecir el modelo proteico a modo general, para diferenciar los dominios o secciones dentro de la proteína que estuviesen expuestos al medio extracelular para la posterior extracción de las secuencias nucleotídicas de estas zonas para expresarlas y utilizarlas como péptidos antigénicos por si solos y no como la proteína nativa en su totalidad.

La proteína seleccionada por el análisis bibliográfico y por homología de proteínas de superficie abundantes en contexto de infección de organismos filogenéticamente cercanos a *P. salmonis*, fue la proteína barril-beta de membrana externa de secuencia de referencia de NCBI WP_016209892.1. Esta proteína fue encontrada en base a homología por proteína no caracterizada CBU_1260 de *C. burnetii* (55), la cual es descrita como la primera invasina identificada en *Coxiella*. El gen de CBU_1260 contiene un dominio transmembrana típico de la familia de proteínas de la membrana externa (OmpA), que se encuentra en varias bacterias y facilita la adhesión y/o internalización en las células hospederas (57).

Mediante el modelamiento proteico de la proteína barril-beta de membrana externa de *P. salmonis* por SWISS-MODEL se obtuvo, en un primer lugar a través de la herramienta de AlphaFold (Figura 3), una proteína compuesta por laminas betas en la región de membrana exponiendo 4 estructuras peptídicas largas al medio extracelular. Además, por la primera coincidencia de modelamiento sin incluir la herramienta AlphaFold (Figura 7), se logró identificar estas estructuras como loops largos no-estructurados definidas como fin extracelular de la proteína (79), tomando en consideración el alineamiento múltiple de secuencias de esta proteína con las proteínas con identidades más cercanas encontradas mediante el análisis de conservación por alineamiento local (Blastp), se identifican estos loops largos no-estructurados como la zonas de más alta variabilidad genética entre las proteínas analizadas, confiriéndole características de especificidad del organismos en términos inmunogénicos. Es importante mencionar que las proteínas obtenidas en el modelamiento por AlphaFold en SWISS-MODEL se encontraron en La Base de Datos de Estructuras de Proteínas AlphaFold (AlphaFold DB), la cual es una amplia base de datos pública de modelos de estructuras de proteínas altamente precisos.

Estos modelos son el resultado de AlphaFold2, un algoritmo de inteligencia artificial desarrollado por DeepMind (87).

Por otro lado, respecto a las otras proteínas candidatas para el estudio correspondientes a la proteína de membrana externa de la familia TolC (WP_032126547.1) y a la proteína de la familia OmpA (WP_017376173.1),se decidió no incluirlas en el subclonamiento y expresión peptídica, en una primera instancia, debido a poseer estructuras pequeñas expuestas al medio extracelular (Figuras 4 y 5), pudiendo provocar complicaciones al momento de expresar las proteínas y de utilizarlas para estudios posteriores basados en generación de anticuerpos para estos péptidos.

Las secuencias de los 4 loops largos no estructurados son grandes candidatos, tanto en conjunto como por si solos, como posibles péptidos antigénicos para ser utilizados en este estudio. Sin embargo, en base a los recursos disponibles se seleccionó solo uno, correspondiente a la zona pintada en naranjo de la Figura 7. Siendo el primer loop en orden de la secuencia suministrada por la base de datos RefSeq de NCBI. En términos de una futura utilización del péptido para aplicaciones biotecnológicas, por ejemplo, la generación de anticuerpos, se espera que las proteínas nativas presenten una mayor inmunogenicidad que los péptidos cortos. Sin embargo, existen ejemplos de péptidos que pueden estimular la producción de anticuerpos tanto contra el péptido como contra regiones de la proteína nativa. Esto se debe a que los péptidos pueden adoptar estructuras estables que imitan las de la proteína completa, lo que explica esta aparente contradicción (88). Por lo tanto, se seleccionó un único péptido dentro de la proteína para su expresión heteróloga buscando, también, esclarecer en futuras investigaciones la enigmática de esta contradicción, evaluando la inmunogenicidad de este péptido en comparación a, por ejemplo, la proteína nativa.

La secuencia nucleotídica del loop largo no-estructurado seleccionado para su expresión, fue incluido en el diseño *in silico* y síntesis externalizada dentro del vector de clonamiento pUC57, el cual es una plataforma ampliamente utilizada en la actualidad en el ámbito de clonamiento y propagación en biología molecular e ingeniería genética. pUC57 contiene una casete LacZ, que pierde su funcionalidad cuando se clona una secuencia de ADN en su sitio de clonamiento múltiple, permitiendo así la selección por el método de Azul/Blanco después de la transformación en *E. coli* (89). Por lo tanto, la secuencia del péptido fue integrada en el sitio de clonamiento múltiple, insertándose de forma *in silico* mediante los sitios de restricción de las enzimas *BamH*I y *Kpn*I.

Es importante tener en consideración al momento del diseño experimental para la producción de proteínas o péptidos recombinantes, los distintos factores que pueden afectar en la solubilidad, plegamiento, formación de cuerpos de inclusión, características estructurales y de inmunogenicidad. Un gran porcentaje de las proteínas que se fabrican utilizando técnicas recombinantes de sistemas de expresión se pueden diseñar añadiendo residuos adicionales a la proteína diana en los extremos N- o C- terminal de esta misma, durante el paso de clonamiento de la secuencia de la proteína. Estos residuos adicionales denominados etiquetas, se diferencian de una proteína de fusión en que, generalmente, este último se refiere a la unión de extremo a extremo de dos proteínas, mientras que las etiquetas suelen ser una unión de un péptido más corto a una proteína e incluye regiones conectoras (92). La adición de etiquetas de afinidad o de proteínas de fusión en el diseño de la expresión de la proteína heteróloga permite tener variados efectos positivos en términos de las propiedades bioquímicos de la proteína diana, como mejorar el rendimiento de la proteína, prevenir la proteólisis, facilitar el plegamiento y aumentar la solubilidad (93).

Algunas tecnologías de etiquetas de afinidad más nuevas no se han evaluado rigorosamente, por lo tanto, es importante realizar las experimentaciones correspondientes para su validación y, además, estas nuevas tecnologías contienen propiedades únicas que resultan útiles en términos de purificaciones problemáticas (94). Un ejemplo de estas nuevas tecnologías es HaloTag, la cual es una proteína de fusión de 33 kDa basada en una deshalogenasa de haluro-alcanos, enzima presente en algunas bacterias, pero que es modificada mediante ingeniería genética de modo de generar intermediarios covalentes entre el sitio activo y el alcano substrato. Esto le permite una serie de posibilidades en términos de purificación y localización de proteínas recombinantes, tanto en sistemas *in vitro* como *in vivo* (95). Esta plataforma ofrece varias ventajas para investigaciones que requieren alta precisión y especificidad, teniendo una gran variedad de ligandos a HaloTag que permiten el aislamiento, imagenología molecular, inmovilización y purificación rápida de las proteínas de interés (96). Sin embargo, aunque el tamaño relativamente grande de HaloTag puede ser desventajoso en comparación con otras etiquetas de afinidad más pequeñas, puede convertirse en una ventaja cuando se aplica integrándose con proteínas de tamaño pequeño con circulación corta (SCFV), mejorando la circulación de ellas (95–97). Es por esto, que en este estudio se decidió por utilizar la tecnología de HaloTag como proteína de fusión que acompaña al péptido antigénico producido, incluyendo al vector de expresión pFN18A como receptor final del subclonamiento de la secuencia de loop de interés.

Mediante la secuenciación del vector de expresión resultante del subclonamiento de pFN18A y el péptido antigénico, se logró encontrar una región pequeña de nucleótidos (aproximadamente de 20 pb) entre la secuencia codificante de Halo y la de L1 no esperado en el diseño *in silico* del vector. Está región pertenecía al plásmido utilizado para el subclonamiento, pFN18A $(L)_8$, y podría generar complicaciones en estudios posteriores relacionados con la evaluación de la inmunogenicidad y especificidad de posibles anticuerpos generados. Aun así, mediante la utilización de técnicas de PCR con el diseño de partidores específicos para sectores del vector, es posible remover la región deseada sin generar alteraciones en las zonas circundantes o en las mismas secuencias de interés (98), como también, pudiéndose agregar más regiones antigénicas de la proteína de interés o añadir componentes estructurales para generar una mejor respuesta inmunogénica del péptido producido.

Sin embargo, a pesar de la presencia de esta región inesperada, la realización del subclonamiento fue exitosa basado en los resultados obtenidos en la secuenciación, lográndose comprobar la integridad de la secuencia nucleotídica del péptido antigénico insertado en el correspondiente vector de expresión y manteniendo los sitios de restricción rio arriba y abajo de L1 esperados en el diseño y ensayo *in silico* del subclonamiento de pFN18A con L1 desde el vector de clonamiento pUC57-L1. Además, al no tener un experimento de comprobación de resultados en la digestión enzimática de lo obtenido post-amplificación por PCR del inserto de pUC57-L1B, debido a los pequeños tamaños en peso molecular de las bandas esperadas, la realización de la secuenciación se hace clave para poder comprobar el éxito de este paso experimental pudiéndose visualizar lo mencionado anteriormente respecto a la integridad de los sitios de restricción esperados.

Teniendo el vector de expresión pFN18A-L1 purificado y entrando en contexto de la expresión de proteínas heterólogas, una plataforma ampliamente utilizada y estudiada son cepas de *E. coli* que contengan sistemas en cascada inducible por isopropil-β-d-tiogalactopiranósido (IPTG) y de producción de la polimerasa de ARN T7 (76). Las cepas DE3 llevan el gen de la polimerasa de ARN T7 (RNAP) bajo el control del promotor lacUV5, una versión más fuerte del promotor nativo lacZ inducible con IPTG. Esto permite la expresión de cualquier gen de interés bajo el control de un promotor T7 (99). Y es por esto por lo que las cepas BLR (DE3) y C41 (DE3) utilizadas pueden presentar ventajas de base al momento de expresar proteínas heterólogas con la ayuda de la inducción por IPTG, como se realizó en el estudio.

En el caso de la cepa C41 (DE3), presenta ventajas al reducir la toxicidad de la producción de proteínas de membrana, generando una mayor producción de estas proteínas (99), lo que sugiere ser una interesante plataforma para la expresión de proteínas de membrana en este estudio, pero no corrobora una mejora en la producción de péptidos aislados de estas proteínas. Sin embargo, esta cepa de *E. coli* porta mutaciones que los hacen más tolerantes a las proteínas tóxicas en general (77,100), siendo capaz de producir proteínas de membrana que estén involucradas en labores de patogenicidad del organismo como, por ejemplo, las proteínas con las que se trabajó en este estudio.

Por otro lado, BLR (DE3) pLysS contiene el plásmido pLysS que confiere una reducción de la expresión basal mediante el gen para la lisozima T7, un inhibidor natural de la ARN T7. Esto proporciona un mecanismo eficiente para inhibir la pequeña cantidad de ARN T7 sintetizada en ausencia de inductor, debido a su transcripción estocástica desde el promotor lacUV5. Sin embargo, incluso con este control más estricto, se sabe que la expresión no regulada todavía ocurre (101,102), corroborándose en este estudio (Figuras 15, 16 y 17). Y, además, pLysS confiere resistencia al cloranfenicol (101), por lo tanto, permite añadir una prueba por selección adicional al momento de cultivar esta cepa.

En los ensayos de expresión de Halo-L1, se observó una expresión basal de Halo y/o de Halo-péptido en condiciones sin inducción por IPTG, inclusive con la utilización de la cepa de expresión BLR (DE3) pLysS, corroborando la aseveración de que la expresión no regulada todavía ocurre con presencia de la lisozima T7 (101). Además, se logró inferir en los ensayos de expresión de BLR (DE3) pLysS pFN18A-L1 que existe una mayor expresión de Halo y/o Halo-péptido a 20°C que, a una temperatura mayor, en base a los resultados de SDS-PAGE. Esto se puede deber a que la expresión de un gen bajo el promotor T7 puede ser contraproducente a temperaturas más elevadas, debido a que se ha registrado, en el estudio de Namdev P. *et al*. (2019), bajas en la fluorescencia de expresión a 37°C de hasta un ~60% en comparación con temperaturas más bajas de cultivo con inducción por IPTG (15°C) (103). Pudiendo el aumento de temperatura afectar tanto en la expresión por si de las proteínas bajo el promotor T7, o afectar negativamente en la fluorescencia emitida por la proteínas expresadas.

Además, en los resultados del análisis de la expresión de L1 en las distintas cepas utilizadas, al observar con luz azul se obtuvieron 2 bandas fluorescentes de distintos pesos moleculares para cada FPE, estas bandas corresponden a la proteína de fusión unida al ligando Alexa Fluor 488, el cual es un fluoróforo con fotoactivación reversible, por lo tanto, puede someterse a múltiples ciclos de activación y desactivación antes de ser degradado por exposición prolongada a la luz (104). Como se mencionó

anteriormente, la banda fluorescente de menor peso molecular concuerda con el peso esperado de HaloTag solitariamente y la otra banda de mayor peso a HaloTag conteniendo el péptido L1, aun así, esta última presentaba una menor intensidad en la fluorescencia, como también, una menor intensidad en su visualización con tinción de Coomassie. Esto pudo deberse a que algunos péptidos son objetivos de degradación cuando se expresan heterólogamente como proteínas de fusión en *E. coli*, lo que puede limitar los rendimientos después del aislamiento y la purificación (102,103). La degradación de proteínas representa un desafío significativo en la producción de proteínas recombinantes en *E. coli*, afectando la cantidad y calidad de los productos obtenidos. Para abordar este problema, es posible optimizar las condiciones de cultivo y manipular el sistema de expresión del hospedero mediante enfoques de ingeniería de bioprocesos (105), mejorando la tasa de producción versus proteólisis con un análisis más profundo del medio de cultivo y las condiciones del crecimiento bacteriano para la expresión heteróloga. De todos modos, no se puede inferir con certeza lo que pudo provocar el fenómeno de doble banda debido a que no existen antecedentes respecto a esto.

Una solución viable para la problemática de la degradación del péptido es la utilización de un sistema de etiquetamiento proteico tipo "sándwich", en donde se expresa el péptido de interés con una proteína de fusión rio arriba y otra rio abajo, envolviéndolo y evitando su degradación. Sin embargo, una de las principales desventajas de usar este sistema SPI es el aumento significativo en el tiempo necesario para los procedimientos de clonamiento y aislamiento, en comparación con el clonamiento y aislamiento de una fusión de proteínas tradicional (106). Esta estrategia podría desarrollarse incluyendo a HaloTag y adicionando una proteína de fusión diferente rio arriba del péptido, por ejemplo, la proteína similar a la ubiquitina (SUMO), la proteína de unión a maltosa (MBP), la proteína de transferencia de glutatión (GST) o proteínas inteínas (107).

Por otro lado, el protocolo de purificación de HaloTag sugerido por el fabricante consta de la unión de la proteína marcada con HaloTag a una resina HaloLink, la elución de la proteína objetivo mediante el tratamiento con la proteasa TEV, y la posterior eliminación de la proteasa TEV marcada con His-tag mediante la resina HisLink (105,106). Este proceso permite obtener en la purificación únicamente al péptido de interés, liberándose de la proteína de fusión HaloTag, permitiendo así, la obtención del péptido aún con los problemas de expresión que se originaron respecto a las dos bandas fluorescentes visualizadas. No obstante, esta última actividad de purificación afecta el rendimiento de la preparación, necesitándose así, mayores volúmenes de cultivo para alcanzar concentraciones del péptido óptimas para su utilización en estudios posteriores. Aun así, la plataforma HaloTag permite la captura y

purificación eficientes de la proteína objetivo, incluso a bajos niveles de expresión, como en el estudio de Ohana R. *et al*. (2011), en donde se logró la purificación eficiente de quinasas humanas funcionales a partir de células HEK293T (108). Como también, HaloTag ofrece diferentes herramientas para ser utilizadas en la purificación eficiente del péptido como, por ejemplo, las microperlas de polimetilmetacrilato magnético que al ser utilizadas en un proceso de purificación experimentan un acoplamiento covalente que conduce a una inmovilización estable y espacialmente definida con el ligando de HaloTag, logrando inmovilizar aproximadamente 13 µg de proteína por cada 1 mg de microperlas en minutos según el estudio de Motejadded H. *et al*. (2010) (109).

Con la llegada de la secuenciación completa de los genomas de diferentes especies, la selección de antígenos basada en el genoma ha desempeñado un papel importante en el descubrimiento de antígenos y en el diseño de vacunas. Un enfoque que se ha utilizado para explorar los genomas bacterianos patógenos se denomina vacunología inversa y permite investigar el repertorio antigénico potencial completo de un organismo a partir de su secuencia genómica (45). La vacunología inversa implica el clonamiento y expresión de todas las proteínas en la secuencia del genoma de un organismo que se predice *in silico* que están expuestas en la superficie o son secretadas. Luego, cada proteína se examina mediante inmunización de alto rendimiento en ratones para determinar su capacidad de inducir la producción de anticuerpos capaces de neutralizar al organismo objetivo (45). En base a esto y buscando una alternativa más económica y en un menor tiempo de obtención, es que se realizó una selección de algunas proteínas candidatas a generar una mayor inmunogenicidad en base a su implicancia en la infección y ubicación celular. Además, una limitación en el desarrollo de vacunas basadas en múltiples proteínas es la complejidad del proceso de producción, el tiempo y los costos involucrados. Una alternativa más rentable a la mezcla de proteínas es una vacuna basada en epítopos, que incluye epítopos de diversas proteínas. Estas vacunas basadas en epítopos pueden presentarse como un conjunto de péptidos o, como alternativa, los epítopos pueden incluirse en un polipéptido. Estos epítopos pueden abarcar un mayor número de blancos proteicos y, en consecuencia, la diversidad inmunogenética de las bacterias (81). Por lo que, podría ser un desafío interesante para futuras investigaciones, la integración de las secuencias de los 4 loops o regiones no estructuradas que presenta la proteína seleccionada en un mismo inserto para el vector de expresión, logrando así un polipéptido abarcando toda la zona expuesta al medio extracelular de la proteína. O, es más, integrar más péptidos antigénicos de otras potenciales proteínas inmunogénicas en una misma secuencia polipeptídica a expresar, para así probar y evaluar su inmunogenicidad en conjunto para un organismo.

6. CONCLUSIONES

1) En el presente trabajo de tesis, se realizó un diseño racional de péptidos inmunogénicos con el objetivo de identificar y expresar un péptido con potencial de ser usado en diagnóstico o en el desarrollo de vacunas dirigidas a *P. salmonis*. Como resultado se seleccionó un péptido que corresponde a un loop largo no estructurado de superficie de la proteína barril-beta de membrana externa de *P. salmonis* (WP_016209892.1) (denominado como L1). La estrategia experimental incluyó la revisión bibliográfica de organismos filogenéticamente cercanos, que reportaran proteínas de membrana externa abundantes, para luego identificar regiones conservadas de ellas para el posterior modelamiento de la estructura de las secuencias proteicas de *P. salmonis* y que correspondieran a dominios expuestos al medio extracelular del organismo.

2) Un avance significativo de este trabajo de tesis es la expresión del péptido seleccionado en un sistema de expresión heterólogo en *E. coli*. Este logro es crucial para la generación de anticuerpos específicos dirigidos contra esta bacteria posibles de utilizar en el desarrollo de nuevas técnicas de diagnóstico específicas. La capacidad de producir este péptido de en un sistema de expresión heterólogo abre la puerta a futuras investigaciones y aplicaciones biotecnológicas, incluyendo el desarrollo de kits de diagnóstico más precisos y la potencial creación de terapias dirigidas contra *P. salmonis*. Estos resultados destacan la importancia de la ingeniería genética y la biotecnología en la generación de herramientas clave para abordar problemas de salud en la acuicultura y la microbiología.

7. IMPLICANCIAS, RECOMENDACIONES Y/O PROYECCIONES FUTURAS

El trabajo presentado en esta tesis científica tiene el potencial de generar repercusiones en diferentes ámbitos. Desde una perspectiva científica y de investigación, los resultados obtenidos podrían contribuir al entendimiento de los mecanismos de expresión de proteínas heterólogas y su relación con la degradación de las proteínas producidas en sistemas de expresión en bacterias *E. coli*. Además, la posterior generación de anticuerpos específicos contra los péptidos producidos puede ser valiosa en la investigación básica para comprender mejor la función de proteínas específicas y los procesos biológicos subyacentes. Esto podría desentrañar nuevos mecanismos celulares y contribuir al avance del conocimiento científico, sobre todo en un organismo con limitadas investigaciones relacionadas al respecto como *P. salmonis*. En adición, la metodología empleada en este trabajo para el diseño y generación de péptidos antigénicos abre puertas hacia la generación de vacunas más eficientes y específicas, en el escenario que actualmente las vacunas comercialmente disponibles no logran entregar la protección necesaria, lo cual podría tener un impacto notable en el campo de la inmunización y la salud pública. Desde una perspectiva económica, la optimización de la producción de proteínas recombinantes en sistemas de expresión y la selección de péptidos inmunogénicos de manera *in silico*, podría reducir costos asociados a la investigación y desarrollo de nuevos productos biofarmacéuticos.

Por otro lado, la generación de anticuerpos contra los péptidos inmunogénicos generados podría ser valiosa para la identificación y validación de biomarcadores. Estos biomarcadores podrían utilizarse en el diagnóstico temprano y seguimiento de enfermedades, mejorando la precisión y eficacia de los tratamientos contra *P. salmonis* y, de la misma manera, podría permitir la investigación detallada de mecanismos inmunológicos y respuestas específicas del sistema inmunológico de peces infectados por este patógeno. Pudiendo contribuir a una mejor comprensión de cómo se desarrollan y regulan las respuestas inmunológicas del pez.

Los péptidos con demostrada inmunogenicidad y capacidad para inducir anticuerpos específicos podrían ser empleados como objetivos en la concepción y formulación de fármacos. Estos péptidos podrían ser sometidos a modificaciones para optimizar su afinidad de unión a las metas terapéuticas, integrándose en tratamientos altamente dirigidos y de gran selectividad. Asimismo, los péptidos específicos generados podrían ser considerados como candidatos para la elaboración de vacunas. Los anticuerpos resultantes podrían brindar inmunidad frente a enfermedades infecciosas, en este caso

SRS. Además, en el ámbito de la salud animal, la generación de anticuerpos específicos podría tener aplicaciones en la investigación y diagnóstico de SRS en poblaciones de peces infectados. Esta perspectiva cobra relevancia en el ámbito veterinario de los cultivos de salmones, por ejemplo, contribuyendo a mejorar la salud, bienestar y calidad en el producto final de los animales de la industria salmonera.

En resumen, la generación de péptidos para la posterior producción de anticuerpos tiene un amplio rango de implicancias en diversos campos, desde la investigación biomédica hasta la terapéutica y la salud animal. Su aplicabilidad en el desarrollo de diagnósticos precisos, terapias específicas y fármacos innovadores hace de esta área una vía prometedora para avanzar en la comprensión y abordaje de la infección por *P. salmonis*, como también, para ampliar el conocimiento sobre este patógeno.

A modo de protección intelectual, se identifican potenciales productos patentables derivados de esta investigación, especialmente relacionados con los péptidos generados y las secuencias nucleotídicas utilizadas para la generación de estos. Estos avances podrían tener aplicaciones en la generación de anticuerpos específicos, el diseño de vacunas, el diagnóstico y la generación de terapias específicas contra SRS. Además, los resultados obtenidos en este estudio tienen el potencial de ser publicados en revistas científicas especializadas, contribuyendo al avance del conocimiento en las áreas de expresión de proteínas recombinantes, la generación de péptidos inmunogénicos, la búsqueda de proteínas de superficie de *P. salmonis* incluyendo sus modelamientos computacionales y las aplicaciones en biotecnología que conllevan. Además, profundizando el estudio y siguiendo con la investigaciones planteadas para un futuro, se pueden publicar artículos científicos que describan el proceso de generación de los péptidos, sus propiedades inmunogénicas y su capacidad para inducir la producción de anticuerpos específicos. Estos artículos podrían ser publicados en revistas científicas especializadas en microbiología, biotecnología o inmunología.

En base a los resultados presentados en este trabajo de tesis, se sugiere considerar las siguientes recomendaciones para futuras investigaciones. Sería beneficioso explorar en mayor profundidad la optimización de condiciones de cultivo para reducir la degradación de proteínas en sistemas de expresión heteróloga, lo que podría resultar en una mayor producción de proteínas recombinantes de alta calidad para una posterior purificación, utilizando las herramientas de la tecnología HaloTag, por ejemplo. Asimismo, investigar la aplicabilidad de las técnicas desarrolladas en otros sistemas de expresión y para diferentes proteínas heterólogas podría ampliar aún más el impacto de esta

investigación, utilizando la metodología de búsqueda y selección de los péptidos inmunogénicos utilizada.

Finalmente, el presente trabajo de tesis es un primer paso para la creación de un sistema colorimétrico avanzado para detectar de manera específica la presencia de *P. salmonis*. Este método se basará en la interacción de anticuerpos policlonales monoespecíficos con las proteínas superficiales de la bacteria, basándose en la expresión y purificación de los péptidos inmunogénicos generados en este trabajo y la posterior generación de los anticuerpos específicos a partir de estos péptidos. Esto establece un punto de partida esencial para el diseño de un kit de diagnóstico que permitirá la detección inmediata y en el lugar de la infección por *P. salmonis*, a través de una reacción de cambio de color. En la actualidad, los métodos de diagnóstico disponibles no son idóneos para responder rápidamente a brotes infecciosos, lo que enfatiza la necesidad de estrategias más accesibles y sencillas que permitan a la industria monitorizar el estado de salud de sus cultivos de manera constante.

La detección temprana de cultivos infectados mediante esta técnica podría optimizar la aplicación de antibióticos, utilizándolos de manera eficiente y en el momento preciso para lograr una respuesta efectiva. Esto contribuiría a reducir el impacto medioambiental y económico del patógeno en la industria salmonera. Además, esta aproximación fomenta la prevención de infecciones en otros cultivos de salmón y facilita el desarrollo de un kit de uso sencillo y sin complicaciones para los profesionales de la acuicultura. De esta manera, se elimina la necesidad de personal altamente capacitado y equipos sofisticados, permitiendo un monitoreo ágil y efectivo de la salud de los cultivos, en línea con las demandas de la industria.

8. REFERENCIAS

1. Bozsik N, Cubillos JPT, Stalbek B, Vasa L, Magda R. Food security management in developing countries: Influence of economic factors on their food availability and access. PLoS One. 2022 Jul 1;17(7 July).

2. Pradeepkiran JA. Aquaculture role in global food security with nutritional value: A review. Vol. 3, Translational Animal Science. Oxford University Press; 2019. p. 903–10.

3. Flores-Kossack C, Montero R, Köllner B, Maisey K. Chilean aquaculture and the new challenges: Pathogens, immune response, vaccination and fish diversification. Fish Shellfish Immunol. 2020 Mar 1;98:52–67.

4. Intesal. VI Informe de Sustentabilidad SalmonChile. 2020.

5. Servicio nacional de pesca y acuicultura de Chile (Sernapesca). Informe Sanitario De Salmonicultura En Centros Marinos Año 2020. 2020.

6. Servicio nacional de pesca y acuicultura de Chile (Sernapesca). Informe Sanitario Con Información Sanitaria De Agua Dulce Y Mar Año 2021. Vol. 2021. 2021.

7. Adams A, Thompson KD. Biotechnology offers revolution to fish health management. Trends Biotechnol. 2006;24(5):201–5.

8. Food and Agriculture Organization of the United Nations (FAO). Report of the FAO/MSU/WB First Multi-Stakeholder Consultation on a Progressive Management Pathway to Improve Aquaculture Biosecurity (PMP/AB), Washington, D.C., United States of America, 10–12 April 2018. Vol. 1254, FAO Fisheries and Aquaculture Report. 2019.

9. Figueroa C, Veloso P, Espin L, Dixon B, Torrealba D, Elalfy IS, et al. Host genetic variation explains reduced protection of commercial vaccines against Piscirickettsia salmonis in Atlantic salmon. Vol. 10, Scientific Reports. Nature Publishing Group UK; 2020.

10. Lafferty KD, Harvell CD, Conrad JM, Friedman CS, Kent ML, Kuris AM, et al. Infectious diseases affect marine fisheries and aquaculture economics. Ann Rev Mar Sci. 2015;7(September 2014):471–96.

11. Tort L. Stress and immune modulation in fish. Dev Comp Immunol. 2011;35(12):1366–75.

12. Servicio nacional de pesca y acuicultura de Chile (Sernapesca). Informe Sanitario De Salmonicultura En Centros Marinos Año 2017. 2018.

13. Servicio nacional de pesca y acuicultura de Chile (Sernapesca). Informe Sanitario De Salmonicultura En Centros Marinos Año 2018. 2019.

14. Marshall SH, Conejeros P, Zahr M, Olivares J, Gómez F, Cataldo P, et al. Immunological characterization of a bacterial protein isolated from salmonid fish naturally infected with Piscirickettsia salmonis. Vaccine. 2007;25(11):2095–102.

15. Fryer JL, Hedrick RP. Piscirickettsia salmonis: A Gram-negative intracellular bacterial pathogen of fish. J Fish Dis. 2003;26:251–62.

16. Branson EJ, Nieto Diaz-Muñoz D. Description of a new disease condition occurring in farmed coho salmon , Oncorhynchus kisuteh (Walbaum), in South America. J Fish Dis. 1991;14(2):147–57.

17. Nourdin-Galindo G, Sánchez P, Molina CF. Comparative Pan-Genome Analysis of Piscirickettsia salmonis Reveals Genomic Divergences within Genogroups. Front Cell Infect Microbiol. 2017;7(459):1–16.

18. Rozas-Serri M. Why Does Piscirickettsia salmonis Break the Immunological Paradigm in Farmed Salmon ? Biological Context to Understand the Relative Control of Piscirickettsiosis. Front Immunol. 2022;13:1–23.

19. Maisey K, Montero R, Christodoulides M. Vaccines for piscirickettsiosis (salmonid rickettsial septicaemia, SRS): the Chile perspective. Expert Rev Vaccines. 2017;16(3):215–28.

20. Rozas M, Enríquez R. Piscirickettsiosis and Piscirickettsia salmonis in fish: A review. J Fish Dis. 2014;37(3):163–88.

21. Evensen Ø. Immunization strategies against Piscirickettsia salmonis infections: Review of vaccination approaches and modalities and their associated immune response profiles. Front Immunol. 2016;7(NOV):1–9.

22. Fundación Oceana. Uso de antibióticos en la salmonicultura chilena: causas, efectos y riesgos asociados. 2018.

23. Abed N, Saïd-Hassane F, Zouhiri F, Mougin J, Nicolas V, Desmaële D, et al. An efficient system for intracellular delivery of beta-lactam antibiotics to overcome bacterial resistance. Vol. 5, Scientific Reports. Nature Publishing Group; 2015.

24. Maya S, Indulekha S, Sukhithasri V, Smitha KT, Nair S V., Jayakumar R, et al. Efficacy of tetracycline encapsulated O-carboxymethyl chitosan nanoparticles against intracellular infections of Staphylococcus aureus. Int J Biol Macromol. 2012;51(4):392–9.

25. Smitha KT, Nisha N, Maya S, Biswas R, Jayakumar R. Delivery of rifampicin-chitin nanoparticles into the intracellular compartment of polymorphonuclear leukocytes. Int J Biol Macromol. 2015;74:36–43.

26. Abed N, Couvreur P. Nanocarriers for antibiotics: A promising solution to treat intracellular bacterial infections. Int J Antimicrob Agents. 2014;43(6):485–96.

27. Yañez AJ, Silva H, Valenzuela K, Pontigo JP, Godoy M, Troncoso J, et al. Two novel blood-free solid media for the culture of the salmonid pathogen Piscirickettsia salmonis. J Fish Dis. 2013;36(6):587–91.

28. Mauel MJ, Ware C, Smith PA. Culture of Piscirickettsia salmonis on enriched blood agar. Journal of Veterinary Diagnostic Investigation. 2008;20(2):213–4.

29. Yañez AJ, Valenzuela K, Silva H, Retamales J, Romero A, Enriquez R, et al. Broth medium for the successful culture of the fish pathogen Piscirickettsia salmonis. Dis Aquat Organ. 2012;97(3):197–205.

30. Fryer JL, Lannan CN. Rickettsial infections of fish. Annu Rev Fish Dis. 1996;6(9):3–13.

31. Marshall S, Heath S, Henríquez V, Orrego C. Minimally Invasive Detection of Piscirickettsia salmonis in Cultivated Salmonids via the PCR. Applied and Enviromental Microbiology. 1998;64(8):3066–9.

32. Karatas S, Mikalsen J, Steinum TM, Taksdal T, Bordevik M, Colquhoun DJ. Real time PCR detection of Piscirickettsia salmonis from formalin-fixed paraffin-embedded tissues. J Fish Dis. 2008;31(10):747–53.

33. Mandakovic D, Glasner B, Maldonado J, Aravena P, González M, Cambiazo V, et al. Genomic-based restriction enzyme selection for specific detection of Piscirickettsia salmonis by 16S rDNA PCR-RFLP. Front Microbiol. 2016;7(643):1–12.

34. Lannan CN, Erwing SA, Fryer JL. A Fluorescent Antibody Test for Detection of the Rickettsia Causing Disease in Chilean Salmonids. J Aquat Anim Health. 1991;3(4):229–34.

35. Servicio nacional de pesca y acuicultura de Chile (Sernapesca). Norma Técnica N°1: Procedimientos para el muestreo de animales acuáticos. Programa Sanitario General Laboratorios de Diagnóstico de Enfermedades de Animles Acuáticos Chile; 2019 p. 1–24.

36. Millward SW, Agnew HD, Lai B, Lee SS, Lim J, Nag A, et al. In situ click chemistry: From small molecule discovery to synthetic antibodies. Integrative Biology (United Kingdom). 2013;5(1):87–95.

37. Lee JH, Yin R, Ofek G, Pierce BG. Structural Features of Antibody-Peptide Recognition. Front Immunol. 2022 Jul 7;13.

38. Tontini M, Romano MR, Proietti D, Balducci E, Micoli F, Balocchi C, et al. Preclinical studies on new proteins as carrier for glycoconjugate vaccines. Vaccine. 2016 Jul 29;34(35):4235–42.

39. Trier N, Hansen P, Houen G. Peptides, antibodies, peptide antibodies and more. Vol. 20, International Journal of Molecular Sciences. MDPI AG; 2019.

40. Saylor K, Gillam F, Lohneis T, Zhang C. Designs of Antigen Structure and Composition for Improved Protein-Based Vaccine Efficacy. Vol. 11, Frontiers in Immunology. Frontiers Media S.A.; 2020.

41. Ricci AD, Brunner M, Ramoa D, Carmona SJ, Nielsen M, Agüero F. APRANK: Computational Prioritization of Antigenic Proteins and Peptides From Complete Pathogen Proteomes. Front Immunol. 2021 Jul 15;12.

42. Li L, Gao M, Li J, Zu S, Wang Y, Chen C, et al. Methods to Identify Immunogenic Peptides in SARS-CoV-2 Spike and Protective Monoclonal Antibodies in COVID-19 Patients. Small Methods. 2021 Jul 1;5(7).

43. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, et al. Database resources of the national center for biotechnology information. Nucleic Acids Res. 2022 Jan 7;50(D1):D20–6.

44. Hu YF, Zhao D, Yu XL, Hu YL, Li RC, Ge M, et al. Identification of bacterial surface antigens by screening peptide phage libraries using whole bacteria cell-purified antisera. Front Microbiol. 2017 Jan 26;8(JAN).

45. Delany I, Rappuoli R, Seib KL. Vaccines, reverse vaccinology, and bacterial pathogenesis. Cold Spring Harb Perspect Med. 2013;3(5).

46. Altschup SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic Local Alignment Search Tool. Vol. 215, J. Mol. Biol. 1990.

47. Bateman A, Martin MJ, Orchard S, Magrane M, Ahmad S, Alpi E, et al. UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Res. 2023 Jan 6;51(D1):D523–31.

48. Ortiz-Severín J, Travisany D, Maass A, Cambiazo V, Chávez FP. Global proteomic profiling of piscirickettsia salmonis and salmon macrophage-like cells during intracellular infection. Microorganisms. 2020 Dec 1;8(12):1–26.

49. Pearson WR. An introduction to sequence similarity (“homology”) searching. Curr Protoc Bioinformatics. 2013;(SUPPL.42).

50. Mavromoustakos T, Durdagi S, Koukoulitsa C, Simcic M, Papadopoulos MG, Hodoscek M, et al. Strategies in the Rational Drug Design. Vol. 18, Current Medicinal Chemistry. 2011.

51. Pulgar R, Travisany D, Zuñiga A, Maass A, Cambiazo V. Complete genome sequence of Piscirickettsia salmonis LF-89 (ATCC VR-1361) a major pathogen of farmed salmonid fish. J Biotechnol. 2015;212:30–1.

52. Fryer JL, Hedrick RP. Piscirickettsia salmonis: A Gram-negative intracellular bacterial pathogen of fish. J Fish Dis. 2003;26:251–62.

53. Mauel MJ, Giovannoni SJ, Fryer JL. Phylogenetic analysis of Piscirickettsia salmonis by 16S, internal transcribed spacer (ITS) and 23S ribosomal DNA sequencing. Dis Aquat Organ. 1999;35(2):115–23.

54. Shaw EI, Voth DE. Coxiella burnetii: A pathogenic intracellular acidophile. Microbiology (United Kingdom). 2019 Jan 1;165(1):1–3.

55. Seshadri R, Paulsen IT, Eisen JA, Read TD, Nelson KE, Nelson WC, et al. Complete genome sequence of the Q-fever pathogen Coxiella burnetii. Proc Natl Acad Sci U S A. 2003;100(9):5455–60.

56. Deringer JR, Chen C, Samuel JE, Brown WC. Immunoreactive Coxiella burnetii Nine Mile proteins separated by 2D electrophoresis and identified by tandem mass spectrometry. Microbiology (N Y). 2011;157(2):526–42.

57. Martinez E, Cantet F, Fava L, Norville I, Bonazzi M. Identification of OmpA, a Coxiella burnetii Protein Involved in Host Cell Invasion, by Multi-Phenotypic High-Content Screening. PLoS Pathog. 2014;10(3):1–22.

58. Zhang G, Kiss K, Seshadri R, Hendrix LR, Samuel JE. Identification and Cloning of Immunodominant Antigens of Coxiella burnetii. Infect Immun. 2004;72(2):844–52.

59. Winn WCJr. Legionella. In: Medical Microbiology 4th edition [Internet]. Baron S. 1996. Available from: https://www.ncbi.nlm.nih.gov/books/NBK7619/

60. Dybwad M, Aarskaug T, Fykse EM, Madslien EH, Blatny JM. Complete genome sequences of six Legionella pneumophila isolates from two collocated outbreaks of legionnaires' disease in 2005 and 2008 in Sarpsborg/Fredrikstad, Norway. Genome Announc. 2016;4(6):15–6.

61. Aurass P, Gerlach T, Becher D, Voigt B, Karste S, Bernhardt J, et al. Life stage-specific proteomes of Legionella pneumophila reveal a highly differential abundance of virulence-associated Dot/Icm effectors. Molecular and Cellular Proteomics. 2016;15(1):177–200.

62. Younas F, Soltanmohammadi N, Knapp O, Benz R. The major outer membrane protein of Legionella pneumophila Lpg1974 shows pore-forming characteristics similar to the human

mitochondrial outer membrane pore, hVDAC1. Biochim Biophys Acta Biomembr. 2018;1860(8):1544–53.

63. Berube BJ, Rangel SM, Hauser AR. Pseudomonas aeruginosa: breaking down barriers. Curr Genet. 2016;62(1):109–13.

64. Tümmler B, Klockgether J. Recent advances in understanding Pseudomonas aeruginosa as a pathogen. F1000Res. 2017;6(0):1–10.

65. Stover CK, Pham XQ, Erwin AL, Mizoguchi SD, Warrener P, Hickey MJ, et al. Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen. Nature. 2000;406(6799):959–64.

66. Crouzet M, Claverol S, Lomenech AM, Le Sénéchal C, Costaglioli P, Barthe C, et al. Pseudomonas aeruginosa cells attached to a surface display a typical proteome early as 20 minutes of incubation. PLoS One. 2017;12(7):1–24.

67. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, et al. Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 2016;44(D1):D733–45.

68. Madeira F, Pearce M, Tivey ARN, Basutkar P, Lee J, Edbali O, et al. Search and sequence analysis tools services from EMBL-EBI in 2022. Nucleic Acids Res. 2022 Jul 5;50(W1):W276–9.

69. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: Homology modelling of protein structures and complexes. Nucleic Acids Res. 2018 Jul 2;46(W1):W296–303.

70. Bienert S, Waterhouse A, De Beer TAP, Tauriello G, Studer G, Bordoli L, et al. The SWISS-MODEL Repository-new features and functionality. Nucleic Acids Res. 2017 Jan 1;45(D1):D313–9.

71. Guex N, Peitsch MC, Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. Electrophoresis. 2009;30(SUPPL. 1).

72. Studer G, Rempfer C, Waterhouse AM, Gumienny R, Haas J, Schwede T. QMEANDisCo—distance constraints applied on model quality estimation. Bioinformatics. 2020;36(6):1765–71.

73. Bertoni M, Kiefer F, Biasini M, Bordoli L, Schwede T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. Sci Rep. 2017 Dec 1;7(1).

74. Sehnal D, Bittrich S, Deshpande M, Svobodová R, Berka K, Bazgier V, et al. Mol∗Viewer: Modern web app for 3D visualization and analysis of large biomolecular structures. Nucleic Acids Res. 2021 Jul 2;49(W1):W431–7.

75. Greener AL, Jerpseth BD. Highly transformable bacterial cells and methods for producing the same. United States; US patent 6.706,525 B1, 2006.

76. Goffin P, Dehottay P. Complete genome sequence of Escherichia coli BLR(DE3), a recA-deficient derivative of E. coli BL21(DE3). Genome Announc. 2017;5(22).

77. Schlegel S, Genevaux P, de Gier JW. De-convoluting the Genetic Adaptations of E.coli C41(DE3) in Real Time Reveals How Alleviating Protein Production Stress Improves Yields. Cell Rep. 2015 Mar 17;10(10):1758–66.

78. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, et al. Database resources of the national center for biotechnology information. Nucleic Acids Res. 2022 Jan 7;50(D1):D20–6.

79. Arora A, Abildgaard F, Bushweller JH, Tamm LK. Structure of outer membrane protein A transmembrane domain by NMR spectroscopy. Nat Struct Biol. 2001;8(4):334–8.

80. Ailloud F, Didelot X, Woltemate S, Pfaffinger G, Overmann J, Bader RC, et al. Within-host evolution of Helicobacter pylori shaped by niche-specific adaptation, intragastric migrations and selective sweeps. Nat Commun. 2019 Dec 1;10(1).

81. Calado CRC. Antigenic and conserved peptides from diverse Helicobacter pylori antigens. Biotechnol Lett. 2022 Mar 1;44(3):535–45.

82. Yu CS, Lin CJ, Hwang JK. Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n -peptide compositions . Protein Science. 2004 May;13(5):1402–6.

83. Li H, Liu L, Zhang W jia, Zhang X, Zheng J, Li L, et al. Analysis of the Antigenic Properties of Membrane Proteins of Mycobacterium tuberculosis. Sci Rep. 2019 Dec 1;9(1).

84. Kaminski DA, Lee FEH. Antibodies against conserved antigens provide opportunities for reform in influenza vaccine design. Vol. 2, Frontiers in Immunology. 2011.

85. Fiser A. Template-based protein structure modeling. Vol. 673, Methods in molecular biology (Clifton, N.J.). 2010. p. 73–94.

86. David A, Islam S, Tankhilevich E, Sternberg MJE. The AlphaFold Database of Protein Structures: A Biologist's Guide. Vol. 434, Journal of Molecular Biology. Academic Press; 2022.

87. Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, et al. AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. 2022 Jan 7;50(D1):D439–44.

88. Camacho CJ, Katsumata Y, Ascherman DP. Structural and thermodynamic approach to peptide immunogenicity. PLoS Comput Biol. 2008;4(11).

89. Goulev Y, Matifas A, Heyer V, Reina-San-Martin B, Charvin G. Cosplay: An expandable toolbox for combinatorial and swift generation of expression plasmids in yeast. PLoS One. 2019 Aug 1;14(8).

90. Struhl K. Subcloning of DNA Fragments. In: Current Protocols in Molecular Biology. 1991. p. 1–11.

91. Shuldiner AR, Tanner K. Rapid (Ligase-Free) Subcloning of Polymerase Chain Reaction Products. In: PCR Protocols. 1993. p. 229–40.

92. Malhotra A. Chapter 16 Tagging for Protein Expression. 1st ed. Vol. 463, Methods in Enzymology. Elsevier Inc.; 2009. 239–258 p.

93. Arnau J, Lauritzen C, Petersen GE, Pedersen J. Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. Protein Expr Purif. 2006;48(1):1–13.

94. Kimple ME, Brill AL, Pasker RL. Overview of affinity tags for protein purification. Curr Protoc Protein Sci. 2013;(SUPPL.73):608–16.

95. England CG, Luo H, Cai W. HaloTag Technology: A Versatile Platform for Biomedical Applications. Bioconjug Chem. 2015 Jun 17;26(6):975–86.

96. Los G V., Encell LP, McDougall MG, Hartzell DD, Karassina N, Zimprich C, et al. HaloTag: A novel protein labeling technology for cell imaging and protein analysis. ACS Chem Biol. 2008 Jun 20;3(6):373–82.

97. Chen W, Younis MH, Zhao Z, Cai W. Recent biomedical advances enabled by HaloTag technology HHS Public Access. Vol. 46, Biocell. 2022.

98. Rashtchian A. Novel methods for cloning and engineering genes using the polymerase chain reaction. Vol. 6, Current Opinion in Biotechnology. 1995.

99. Heyde SAH, Nørholm MHH. Tailoring the evolution of BL21(DE3) uncovers a key role for RNA stability in gene expression toxicity. Commun Biol. 2021 Dec 1;4(1).

100. Schreiber C, Müller H, Birrenbach O, Klein M, Heerd D, Weidner T, et al. A high-throughput expression screening platform to optimize the production of antimicrobial peptides. Microb Cell Fact. 2017 Feb 13;16(1).

101. Rosano GL, Morales ES, Ceccarelli EA. New tools for recombinant protein production in Escherichia coli: A 5-year update. Vol. 28, Protein Science. Blackwell Publishing Ltd; 2019. p. 1412–22.

102. Zhang X, Studier FW. Mechanism of Inhibition of Bacteriophage T7 RNA Polymerase by T7 Lysozyme. J Mol Biol. 1997;269:10–27.

103. Namdev P, Dar HY, Srivastava RK, Mondal R, Anupam R. Induction of T7 Promoter at Higher Temperatures May Be Counterproductive. Indian Journal of Clinical Biochemistry. 2019 Jul 1;34(3):357–60.

104. Cook A, Walterspiel F, Deo C. HaloTag-Based Reporters for Fluorescence Imaging and Biosensing. Vol. 24, ChemBioChem. John Wiley and Sons Inc; 2023.

105. Singha TK, Gulati P, Mohanty A, Khasa YP, Kapoor RK, Kumar S. Efficient genetic approaches for improvement of plasmid based expression of recombinant protein in Escherichia coli: A review. Vol. 55, Process Biochemistry. Elsevier Ltd; 2017. p. 17–31.

106. Lamer T, Vederas JC. Simplified cloning and isolation of peptides from “sandwiched” SUMO-peptide-intein fusion proteins. BMC Biotechnol. 2023 Dec 1;23(1).

107. Lamer T, van Belkum MJ, Wijewardane A, Chiorean S, Martin-Visscher LA, Vederas JC. SPI "sandwich": Combined SUMO-Peptide-Intein expression system and isolation procedure for improved stability and yield of peptides. Protein Science. 2022 May 1;31(5).

108. Ohana RF, Hurst R, Vidugiriene J, Slater MR, Wood K V., Urh M. HaloTag-based purification of functional human kinases from mammalian cells. Protein Expr Purif. 2011 Apr;76(2):154–64.

109. Motejadded H, Kranz B, Berensmeier S, Franzreb M, Altenbuchner J. Expression, one-step purification, and immobilization of HaloTag TM fusion proteins on chloroalkane-functionalized magnetic beads. Appl Biochem Biotechnol. 2010;162(7):2098–110.