



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY



Tesina de grado para la Licenciatura en Bioquímica

EVALUACIÓN DE VECTORES PARA LA PRODUCCIÓN DE PROTEÍNAS RECOMBINANTES EN DIFERENTES SISTEMAS DE EXPRESIÓN

Estudiante: Gabriela Paradedá

Orientador: Ph.D. Agustín Correa

Co-orientadora: Ph.D. Claudia Ortega

Unidad de Proteínas Recombinantes

Institut Pasteur de Montevideo

Montevideo, abril 2023

Orientador	Ph.D. Agustín Correa
Co-orientadora	Ph.D. Claudia Ortega
Autora	Bach. Gabriela Paradedá
Tribunal	
Fecha	
Calificación	

Agradecimientos:

A mis tutores Ph.D. Claudia Ortega y Ph.D. Agustín Correa por darme la oportunidad de trabajar con ellos, por la confianza, disposición y compromiso, me llevo mucho de esta experiencia, gracias.

A mis amigas y amigos, en especial a Paula, por creer en mí siempre y por su apoyo incondicional.

A mi hermano, que con su gran curiosidad e inventiva inspiró en mí un gran amor por la Ciencia y con quien compartimos la vida aprendiendo el uno del otro.

A mi papá, que hace un año ya no está, pero que puedo verlo con una sonrisa amplia, cálida y llena de orgullo.

A mi mamá, por su paciencia, ternura y amor inagotable, un pilar en mi vida que me apoya en cada paso.

Gracias.

Tabla de Contenidos

1	Resumen.....	7
2	Palabras clave.....	8
3	Abreviaturas.....	9
4	Introducción.....	11
4.1	Sistemas de Expresión Procariotas–Sistema <i>Escherichia coli</i>	15
4.1.1	Diseño del Vector de Expresión en <i>Escherichia coli</i>	16
4.1.1.1	Promotor.....	16
4.1.1.2	Vector pET Basado en el Promotor T7 e inducción con IPTG.....	16
4.1.1.3	Secuencias de Regulación de la Transcripción y Traducción	18
4.1.1.4	Marcador de Selección	18
4.1.1.5	Número de Copias del Plásmido	19
4.1.1.6	Clonación del Gen Blanco	19
4.1.2.	Cepas Hospedadoras de <i>E.coli</i>	20
4.1.3	Otras Consideraciones a Tener en Cuenta en la Búsqueda de Estrategias para Optimizar la Expresión de PRs en <i>E.coli</i>	22
4.1.3.1	Expresión Génica de Codones Sesgados.....	22
4.1.3.2	Localización Proteica en el Microambiente Bacteriano	23
4.1.3.3.	Evaluación de las Condiciones de Cultivo	25
4.1.3.4	Disminución o Tolerancia de la Toxicidad en el Hospedero por Proteínas Tóxicas.....	25
4.1.3.5	Reducción de la Degradación Proteolítica	26
4.1.3.6	Co-expresión de Chaperonas Moleculares	26
4.1.3.7	Consideración de la Formación de Cuerpos de Inclusión	26
4.1.3.8	Introducción de Proteínas o Tags de Fusión.....	27
4.2.	Sistemas de Expresión Eucariotas - Sistema de Expresión EXPI 293F.....	29
4.2.1	Modificaciones Post-traduccionales.....	29
4.2.1.1	Glicosilación y Perfiles de Glicosilación	29
4.2.2.	Diseño de Vectores para Sistemas de Expresión Basados en Células de Mamífero	31
4.2.2.1	Promotores.....	32
4.2.2.2	Intrones	33
4.2.2.3	Señales de Poliadenilación.....	33
4.2.2.4	Terminadores de la Transcripción	34
4.2.2.5	Elementos de Control Traduccional.....	34

4.2.2.6 Codones de Terminación.....	34
4.2.2.7 Marcadores de Selección	35
4.2.2.8 Péptidos Señal y Proteínas de Fusión	35
4.2.3 Transfección Transitoria vs. Transfección Estable.....	37
4.2.4 Metodologías de Transfección.....	38
4.2.5 Líneas Celulares de Mamífero.....	39
4.2.6 Medios de Cultivo Celular para la Producción Proteica.....	44
4.2.7 Optimización de la Expresión Proteica en Células de Mamífero	45
4.2.8 Purificación de Proteínas.....	46
4.2.9 Proteínas Recombinantes Producidas en la Presente Investigación.....	47
5. Objetivos.....	52
Objetivos generales.....	52
Objetivos específicos	52
6. Materiales y Métodos	53
6.1 Diseño de los Vectores de Expresión para el Sistema de Procariota- <i>E.coli</i>	53
6.2 Diseño de los Vectores de Expresión para el Sistema de Expresión de Mamífero EXPI 293F	54
6.3 Protocolos	55
6.3.1 Equipamiento utilizado en la presente investigación.....	55
6.3.2 Generación de Stock Celular de Plásmidos (Gliceroles).....	55
6.3.3. Expresión de PRs en el Sistema Procariota	56
6.3.3.1 Comparación de los Niveles de Expresión Citoplasmática de GFP en <i>E.coli</i>	56
6.3.3.2 Clonado y Expresión del mNb6 en el Sistema Procariota	57
6.3.3.3 Expresión del Vector pT7v2-mNb6 Clonado en Células Shuffle Electrocompetentes ...	59
6.3.4 Expresión de PRs en el Sistema Eucariota.....	60
6.3.4.1 Comparación de los Niveles de Expresión de GFP Utilizando Vectores con Distintos Péptidos Señal para la Secreción de GFP al Medio en Células Expi293F.....	60
6.3.4.2 Comparación de dos Métodos de Transfección Distintos en Células Expi293F	61
6.3.4.3 Comparación de los Niveles de Expresión de mNb6 Utilizando Distintas Concentraciones de ADN Plasmídico en Células Expi 293F.....	63
6.3.5 Ensayo ELISA para Evaluar la Unión mNb6-RBD	64
7. Resultados.....	66
7.1 Expresión de PRs en el Sistema Procariota	66
7.1.1. Comparación de los Niveles de Expresión Citoplasmática de GFP en <i>E.coli</i> al Utilizar los Vectores pT7-GFP y pT7V2-GFP	66

7.1.2 Clonado y Expresión de mNb6 a partir del Vector pT7v2-mNb6 Transformado en <i>E.coli</i>	68
7.1.3 Expresión del Vector pT7v2-mNb6 Clonado en Células Shuffle Electrocompetentes	69
7.2 Expresión de PRs en el Sistema Eucariota	72
7.2.1 Comparación de los Niveles de Expresión de GFP Utilizando Vectores con Distintos Péptidos Señal para la Secreción de GFP al Medio en Células de Mamífero Expi293	72
7.2.2 Comparación de dos Métodos de Transfección Distintos en el Sistema Eucariota (pCMV-IL2-GFP-ext)	74
7.2.3 Comparación de los Niveles de Expresión de mNb6 Transfectando Células EXPI 293F a Distintas Concentraciones de ADN Plasmídico	75
7.3 Ensayo ELISA para Evaluar la Unión mNb6-RBD en ambos Sistemas de Expresión.....	78
8. Conclusiones.....	79
9. Referencias Bibliográficas	83
10. Anexos	87
Anexo 1 Comparación de las soluciones utilizadas	87
Anexo 2 Preparación de Geles de Poliacrilamida para SDS-PAGE.....	88
Anexo 3 Preparación de placas LB-agar-Ampicilina	89

Resumen

La expresión de Proteínas Recombinantes se ha vuelto una herramienta invaluable en áreas académicas, industriales y biomédicas. Su producción ofrece una vía simple, rápida, económicamente rentable y eficiente para obtener proteínas puras, solubles y biológicamente activas. Entender la estructura molecular de las proteínas individuales permite dirigir preguntas importantes sobre su función fisiológica, así como también las rutas bioquímicas y de regulación en las que están implicadas. El desarrollo de la ingeniería genética ha favorecido el avance de la industria biotecnológica aplicada a la salud humana y animal, por lo que cada vez se utilizan más las Proteínas Recombinantes como agentes terapéuticos en el tratamiento de enfermedades, en el desarrollo de vacunas y como herramientas en ensayos de diagnóstico.

A pesar del crecimiento de diversos sistemas de expresión y de los considerables avances de las tecnologías de expresión y purificación, en muchas ocasiones resulta imprescindible la aplicación de estrategias de optimización. Estas tienen como finalidad lograr un rendimiento apropiado, obtener una proteína pura, soluble, estable y biológicamente activa, teniendo en cuenta, a su vez, la rentabilidad en tiempos y costos de producción. Si bien no existe un protocolo universal que asegure esto, existen reglas básicas a considerar. Estas reglas no necesariamente garantizarán la obtención de las características antes mencionadas, pero sí permitirán aproximarse a estos resultados. Por esto, es importante, a la hora de planificar la expresión de una proteína blanco, que se evalúen sus distintas particularidades, como su origen, propiedades biológicas, futura aplicación, el empleo de un hospedador adecuado para su producción, el compartimiento celular de expresión, su toxicidad en el hospedador, la presencia de codones poco frecuentes, de puentes disulfuro y/o modificaciones post-traduccionales complejas. Asimismo, se deberían evaluar distintos parámetros para su correcta expresión, como el diseño óptimo de los vectores de expresión, las condiciones de crecimiento, la inclusión de proteínas de fusión, así como explorar diferentes metodologías de clonado y posibles estrategias de purificación.

Con el fin de mejorar el rendimiento proteico final, en el presente trabajo se comparó y evaluó la producción de Proteínas Recombinantes mediante la utilización de distintos vectores plasmídicos, tanto en un sistema de expresión procariota como en uno eucariota. A su vez, en el sistema eucariota, se evaluaron distintas concentraciones de ADN plasmídico y diferentes metodologías de transfección para lograr una optimización en los niveles de expresión.

Palabras Clave

Clonado RF
Hospedador celular
Modificaciones post-traduccionales
Péptido señal
Proteínas de fusión
Proteínas Recombinantes
Plásmido
Purificación de proteínas
Regiones promotoras
Región de inicio de la traducción
Sistemas de expresión
Transformación
Transfección transitoria
Transfección estable
Vector de clonación
Vector de expresión

3. Abreviaturas

ARNm: ARN mensajero

ARNr: ARN ribosomal

ARNt: ARN de transferencia

CDR: Región determinante de la complementariedad

Células BHK: células de riñón de hámster bebé

Células COS: células derivadas de riñón de mono africano verde

Células CHO: células de ovario de hámster chino

Células HEK 293: células de riñón de embrión humano

Células SP2/0-Ag14: células de mieloma de ratón

Clonado RF: clonado Restriction free

Dsb: Enzimas formadoras de enlaces disulfuro, del inglés *Disulfide bridge-forming enzymes*

DsbC: disulfuro isomerasa C

E.coli: *Escherichia Coli*

GFP: Proteína verde fluorescente, del inglés *green fluorescent protein*

Gor: Glutación oxidoreductasa

GST: glutación S- transferasa

HRP: Peroxidasa de rábano

IgG: Inmunoglobulina G

IMAC: Cromatografía de afinidad con iones inmovilizados, del inglés *immobilized metal affinity chromatography*

IPTG: isopropil- β -D-tiogalactopiranosido

IRES: Sitio interno de entrada al ribosoma

MBP: proteína de unión a maltosa

NusA: sustancia A de N-utilización

ORI: Origen de replicación

PBS: Buffer fosfato salino

PCR: Reacción en cadena de la polimerasa

PDI: Proteínas disulfuro isomerasas

PEI: polímero catiónico polietilenimina

PR: Proteína Recombinante

PRs: Proteínas Recombinantes

RBD: dominio de unión al receptor

RBS: Sitio de unión al ribosoma

RE: Retículo endoplasmático

Región Fc: fragmento cristalizabile

SDS: Dodecil sulfato sódico

Secuencia SD: Secuencia Shine-Delgarno

SMC: Sitio de multiclonado o polilinker

SUMO: proteína pequeña modificadora de tipo ubiquitina

SV40: Virus de simio 40

TBS: Tris buffer salino

TIR: Región de inicio de la traducción

TrxA: tiorredoxina A

TrxB: Tiorredoxina B

UTR: regiones no traducidas de los genes

VPA: Ácido Valproico

4. Introducción

Las Proteínas Recombinantes (PRs) son producidas en el laboratorio mediante ingeniería genética, cuando se expresa un gen clonado que codifica para la proteína de interés, en una especie o una línea celular distinta a la original. Esta expresión heteróloga involucra la identificación del gen correspondiente de la proteína a expresar y su inserción en un vector que será luego introducido en un hospedador, el cual aportará la maquinaria celular para codificar y expresar dicha proteína.

La elección de las condiciones de expresión apropiadas, del sistema de expresión y del vector, es crucial en el proceso de expresión de PRs y depende, en gran medida, de la finalidad del estudio, ya sea para obtener grandes cantidades de proteína o para investigar sus propiedades funcionales. Por eso, el desarrollo de cada experimento requiere de un análisis exhaustivo de las características propias de la proteína y del sistema de expresión (Correa y Oppezzo, 2015).

En primera instancia, se debe evaluar el origen y las propiedades biológicas de la proteína, sus propiedades físico-químicas, sus niveles de expresión, su localización celular, la existencia de modificaciones post-traduccionales y cómo impactan en su actividad biológica, así como también la aplicación prevista. A su vez, debe elegirse el sistema de expresión, es decir, el organismo hospedero junto con el vector de expresión que portará el gen que codifica para la PR, y evaluar si la proteína puede ser tóxica o contiene codones poco frecuentes para el huésped seleccionado. En base a estos parámetros, es importante considerar también las condiciones de crecimiento —como temperatura, pH, aireación y el medio de cultivo— que influyen no solo en los niveles de biomasa sino también en los rendimientos de PR (Correa y Oppezzo, 2015; Jia y Jeon, 2016; Terpe, 2006).

Un vector es una molécula de ADN de doble cadena, con la capacidad de albergar un fragmento de ADN exógeno. De acuerdo con su uso, puede clasificarse como vector de clonación o de expresión. Los vectores de clonación transfieren y replican fragmentos de ADN exógeno que llevan insertado; estos pueden ser plásmidos, fagos, fagémidos, cósmidos o cromosomas artificiales bacterianos o de levadura. Para actuar como tales, deben contener todos los elementos necesarios para replicarse junto con el fragmento de ADN que transportan y mantenerse estables dentro de un huésped celular; estos elementos básicos son: un origen de replicación, que es la secuencia que provee un único sitio de reconocimiento a las proteínas que identifican el sitio de inicio de la replicación (ORI) y que permitirá al plásmido replicarse de manera autónoma e independiente al genoma del hospedero, un sitio múltiple de clonado (SMC), que es una secuencia específica, reconocida por diversas enzimas de restricción que permitirá

insertar el ADN del gen de interés, un marcador de selección, que permitirá seleccionar las células hospedadoras que contengan el vector recombinante. Por otro lado, los vectores de expresión pueden ser plásmidos o fagos y son aquellos que se utilizan para producir un transcrito o la PR producto de ese transcrito. Contienen, además de todos los elementos esenciales para mantenerse estables dentro de la célula hospedadora antes descritos, todos los elementos necesarios para la regulación de la transcripción y traducción del gen blanco, el promotor que promoverá la transcripción del gen de interés, la secuencia génica de la proteína que se desea expresar, el sitio de unión al ribosoma (RBS) y la secuencia de terminación de la traducción. Puede incluir también secuencias de tags o proteínas de fusión, sitios de corte proteolítico que permitirán remover el tag de la PR, y secuencias de péptidos señal que permitirán dirigir a la PR a determinado compartimiento del huésped celular (Figura 1) (Sandoval Rodríguez *et al.*, 2013).



Figura 1. Representación esquemática de los componentes básicos en un vector de expresión. (Imagen adaptada de Mesa-Pereira *et al.*, 2018)

Generalmente, para la expresión de PRs se utilizan plásmidos, que son moléculas de ADN circular, doble hebra, auto-replicantes, cuyo tamaño varía de entre menos de 1 kb a más de 500 kb y que son mantenidos en bacterias como entidades extracromosómicas. (Glick *et al.*, 2010). Existen varias estrategias para optimizar los niveles de producción proteica a nivel del plásmido como, por ejemplo, la adición de un péptido señal en la secuencia del vector, que permite la translocación de la proteína a un ambiente más favorable para su expresión o también el uso de proteínas o tags de fusión, que consisten en la unión de una corta secuencia de aminoácidos, polipéptido o proteína a la PR. En general, la fusión del tag se construye a nivel del ADN y la quimera se transcribe y expresa como una sola entidad dentro de la célula. Esta estrategia se utiliza con frecuencia para facilitar el proceso de purificación, favorecer el plegamiento, incrementar la solubilidad y prevenir la formación de cuerpos de inclusión, disminuir la degradación proteolítica y dirigir la translocación al periplasma o al medio extracelular (Correa y Opezzo, 2015, 2015; Sørensen y Mortensen, 2005a).

La obtención de cantidades suficientes de una proteína blanco en un estado soluble, puro y homogéneo desde su hospedador natural es muy poco frecuente. En este sentido, la ingeniería genética permite producir y expresar proteínas en forma recombinante, utilizando diferentes hospederos de expresión, como bacterias, levaduras, hongos filamentosos, algas unicelulares, plantas, células de insectos, células de mamífero, e incluso sistemas libres de células (Tabla 1).

Los sistemas procarióticos, en general, son los de elección para la mayoría de los propósitos debido a su fácil manipulación y bajo costo, aunque en ocasiones presentan grandes limitaciones en la producción de proteínas de origen eucariota. Muchas proteínas eucariotas sufren una gran variedad de modificaciones post-traduccionales, necesarias para su correcto plegamiento, estabilidad y actividad biológica que, generalmente, los hospederos procariotas no realizan. Aunque muchas proteínas eucariotas biológicamente activas han sido producidas con éxito en sistemas de expresión procarióticos, ocasionalmente es necesario utilizar sistemas de expresión eucarióticos cuando se persigue expresar proteínas eucariotas con actividad biológica (Jia y Jeon, 2016; Correa y Opezzo, 2011; Zemella *et al.*, 2015; Demain y Vaishnav, 2009).

Tabla 1. Aplicaciones, ventajas y desventajas de los principales sistemas de expresión utilizados en la producción de Proteínas Recombinantes (Información adaptada de *Overview of Protein Expression, Protein Biology Resource Library de ThermoFisher Scientific*).

Sistema de expresión	Aplicaciones más comunes	Ventajas	Desventajas
Bacterias	Análisis funcionales. Generación de fragmentos de anticuerpos. Ensayos funcionales. Interacciones de proteínas.	Escalable. Bajo costo. Condiciones de cultivo simples.	Puede requerir optimización proteica específica. Puede ser difícil expresar proteínas de algunos mamíferos.
Levaduras	Análisis estructurales. Generación de anticuerpos. Análisis funcionales. Interacciones proteicas.	Procesamiento de proteínas eucariotas. Escalable. Requerimientos de medio simples.	Se requiere fermentación para muy altos rendimientos. Las condiciones de crecimiento pueden requerir optimización.
Células de insecto	Ensayos funcionales. Análisis estructurales. Expresión de proteínas intracelulares. Expresión de complejos proteicos. Producción de virus.	Procesamiento proteico similar a mamíferos. Puede usarse en cultivos estáticos o en suspensión.	Condiciones de cultivo más demandantes que los sistemas procariotas. La producción de vectores de baculovirus recombinantes demanda tiempo.
Células de mamífero	Ensayos funcionales. Análisis estructurales. Producción de anticuerpos. Expresión de proteínas complejas. Interacciones proteicas. Producción de virus.	Alto nivel de procesamiento proteico. Las proteínas pueden expresarse tanto de manera transitoria como estable. Robustos sistemas de expresión transitoria optimizados para la producción proteica rápida y de muy alto rendimiento.	Rendimientos en gramos por litro solo posibles en cultivos en suspensión. Condiciones de cultivo más demandantes.
Libres de células	Proteínas tóxicas. Introducciones de tags o aminoácidos no naturales. Ensayos funcionales. Interacciones proteicas. Detección de inhibidores de la traducción.	Sistema abierto, disponible para agregar componentes no naturales. Expresión rápida. Formato simple.	La producción a gran escala no es sostenible.

4.1. Sistemas de Expresión Procariotas-Sistema de *Escherichia coli*

La enterobacteria *Escherichia coli* (*E.coli*) es el sistema de expresión más utilizado y fue el primer hospedador empleado en la producción de PRs hace más de 50 años. Las principales razones del extensivo uso de esta bacteria son: su fácil manipulación, accesible infraestructura de implementación, bajo costo, rápido crecimiento, alto rendimiento de producción de la proteína blanco y un conjunto creciente de herramientas e información genética. Sin embargo, cuando se quiere producir una proteína de origen eucariota en *E. coli*, se ha estimado que aproximadamente solo un 30 % de los genes clonados pueden ser expresados, el resto es degradado, forma cuerpos de inclusión o se expresa en forma soluble, pero agregada. Todo esto ocurre especialmente en el caso de las proteínas de membrana o las que requieren de modificaciones post-traduccionales para su plegamiento y/o función. En general, muchas de las modificaciones post-traduccionales encontradas en eucariotas como la N- o la O-glicosilación, amidación, hidroxilación, miristoilación, palmitación o sulfonación, no pueden ser realizadas por *E.coli*, lo cual limita la aplicación de estos sistemas de expresión heterólogos.

Otros factores que pueden impactar en la expresión de proteínas blanco solubles pueden ser: la correcta formación de puentes disulfuro, la presencia en la secuencia blanco de codones poco frecuentes para *E.coli* o la secuencia de la región de inicio de la traducción (TIR). Mediante herramientas bioinformáticas se puede predecir la secuencia base de estas modificaciones y estimar con buena precisión la presencia o ausencia de modificaciones post-traduccionales, sitios de fosforilación, localización proteica, entre otros (Correa y Oppezzo., 2011; Correa y Oppezzo, 2015).

La producción de PRs en *E.coli*, ya sea para análisis bioquímicos, terapéuticos o estudios estructurales, requiere que se cumplan dos pasos cruciales, por un lado, la expresión soluble de la proteína blanco y, por el otro, la purificación y estabilidad de la molécula funcional. Se han realizado considerables esfuerzos para mejorar la producción de PRs solubles y funcionales. Estos avances incluyen el desarrollo de diferentes cepas de expresión, una amplia variedad de plásmidos bajo el control de diferentes promotores con distintas fuerzas, fusión del gen de interés con tags moleculares que pueden ayudar en la purificación y/o producción de la proteína y, co-expresión de la misma junto con chaperonas moleculares o moduladores del plegamiento que pueden mejorar el plegado y estabilidad de la PR, así como la expresión proteica en una localización celular distinta, o también mutaciones en el gen blanco (Correa y Oppezzo, 2011).

4.1.1. Diseño del Vector de Expresión en *E.coli*

Los elementos que forman parte del vector de expresión incidirán directamente en la expresión del gen clonado. Aunque existe una gran variedad de vectores de expresión comerciales, en muchas ocasiones se debe diseñar un nuevo vector o modificar un vector comercial con el fin de alcanzar determinado objetivo. En ese caso, un óptimo diseño del vector resultará de vital importancia.

4.1.1.1. Promotor

El promotor está compuesto por una secuencia específica, localizada donde se encuentra el punto de inicio de la transcripción y contiene la información necesaria para activar o desactivar el gen que regula. A la hora de expresar PRs en un hospedador determinado, la elección del promotor tiene un gran impacto en el rendimiento final, y su secuencia es un elemento central que afecta la fuerza y duración de la transcripción y, a la vez, el rendimiento proteico. La expresión recombinante necesita, para ser más fuerte, presentar un nivel de expresión basal muy bajo, de manera de evitar un efecto detrimental en el metabolismo de la célula y su inducción debe ser simple y rentable.

Con el fin de garantizar un mayor control de la expresión recombinante, numerosos vectores de expresión incluyen promotores regulables que dirigen la expresión del gen de interés, y contienen secuencias operadoras reconocidas por reguladores transcripcionales, muchas veces codificados en el mismo vector de expresión. De esta forma, los reguladores transcripcionales, mayormente represores, bloquean la transcripción del gen clonado en ausencia de una señal inductora. La señal de inducción puede ser un cambio de temperatura, pH o fuerza iónica en el medio de cultivo, aunque en la mayoría de los casos el inductor es un metabolito que se agrega al medio cuando se quiere inducir la expresión. Los promotores comerciales más utilizados en la expresión proteica son el promotor T7, promotor T5, promotores híbridos como *trc* y *tac*, el promotor AraBAD, todos ellos son controlados mediante inductores químicos. También existen promotores comerciales como *CspA* y *pL/pR* que son controlados por temperatura (Correa y Oppezzo, 2011).

4.1.1.2. Vector pET Basado en el Promotor T7 e Inducción con IPTG

El vector comercial pET basado en T7 es el más utilizado en la expresión recombinante en *E.coli*. Este se basa en la alta selectividad de la ARN polimerasa T7 del fago T7 para conducir la producción de PRs. La ARN polimerasa T7 solo puede transcribir genes bajo el control del promotor T7 y se ha visto que puede transcribir 8 veces más rápido que otras polimerasas de *E.coli*, produciendo altos rendimientos de proteína. El promotor T7 es considerado un promotor

fuerte y la expresión de PRs puede alcanzar más de un 50 % del total de las proteínas celulares. Como *E.coli* carece de esta polimerasa, se han desarrollado cepas como la BL21(DE3), las cuales contienen una copia cromosómica del gen de la polimerasa T7, bajo control del promotor lac, derivado del promotor lacUV5.

El promotor LacUV5 contiene mutaciones puntuales que aumentan la fuerza del promotor y lo hacen menos sensible a la represión catabólica. En esta vía, el promotor solo es controlado por el represor lac, LacI, lo que permite la inducción con un análogo de la lactosa no hidrolizable, el isopropil- β -D-tiogalactopiranosido (IPTG) (Figura 2), incluso en presencia de glucosa. La adición de IPTG libera la represión causada por la unión de LacI al operador lac, resultando en la expresión de la polimerasa T7 que, a su vez, transcribe el gen blanco con la producción concomitante de PR (Correa y Oppezzo, 2011).

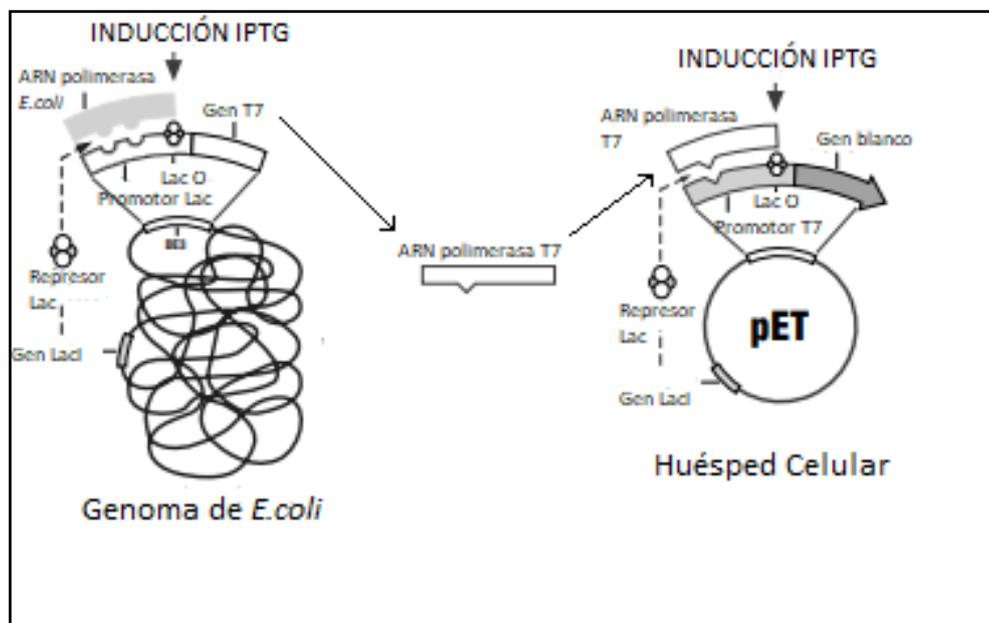


Figura 2. Esquema del control del promotor lac, mediante la inducción con IPTG Imagen adaptada de Novagen. pET System Manual (2006). 11th Ed.

(<https://kirschner.med.harvard.edu/files/protocols/Novagenpetsystem.pdf>)

La concentración del inductor, así como el tiempo de inducción, influirán sensiblemente en los niveles de expresión. Los promotores fuertes son los más utilizados para alcanzar altos niveles de expresión, ya que tienen una alta afinidad y estabilidad de unión a la ARN polimerasa, y su adición supone un incremento sustancial del gasto metabólico de las células hacia la producción de la proteína heteróloga, lo cual ralentizará el metabolismo del hospedero. Por este motivo, la adición del inductor se realiza preferiblemente en el momento en que las células hayan alcanzado la fase exponencial media-tardía del crecimiento, momento en el cual los cultivos, siendo aún relativamente frescos, presentan una alta densidad celular. La adición del inductor

en un estadio muy temprano o demasiado tardío de la fase de crecimiento del cultivo repercutirá negativamente en los rendimientos de la PR (Galloway *et al.*, 2003).

4.1.1.3. Secuencias de Regulación de la Transcripción y Traducción

Otros factores fundamentales a tener en cuenta a la hora de escoger el diseño del vector de expresión son las secuencias terminadoras de la transcripción y las secuencias reguladoras de la transcripción y traducción.

Las secuencias terminadoras de la transcripción son muy importantes ya que la transcripción sobre la secuencia de un promotor puede inhibir su función, fenómeno conocido como oclusión del promotor. Esta interferencia puede ser prevenida con la apropiada ubicación del terminador de la transcripción al finalizar la secuencia codificadora. De forma similar, la ubicación de un terminador de la transcripción en posición río arriba (*upstream*) al promotor será útil para minimizar transcripciones no deseadas. Las secuencias terminadoras de la transcripción formarán, además, una estructura secundaria por complementariedad de bases en el ARNm que protegerá de la acción de exonucleasas, extendiendo su vida media e incrementando sustancialmente los niveles de proteína producida (Makrides, 1996).

Por otro lado, dentro de las secuencias reguladoras de la traducción, en la región del sitio de unión al ribosoma (RBS), posicionado río abajo (*downstream*) del promotor, se encuentra una secuencia altamente conservada, llamada Shine-Delgarno (SD), que precede al codón de inicio de la traducción (AUG) en los ARNm procariotas. Esta secuencia es complementaria con el extremo 3' del ARNr 16S y su hibridación permite el ensamblaje de la maquinaria de inicio de la traducción. La distancia entre la secuencia SD y el codón de inicio de la traducción y la estructura secundaria del RBS influyen directamente en la eficiencia del inicio de la traducción, pudiendo dificultar el acceso del ARNm a la subunidad ribosomal 30S e inhibir la traducción. (Makrides, 1996).

4.1.1.4. Marcador de Selección

Otro elemento indispensable en un vector de expresión es el marcador de selección. Este es un gen que generalmente confiere resistencia a un antibiótico o genera un fenotipo particular con el cual puede seleccionarse la célula que incorporó el vector. Entre los genes de selección más comunes se encuentran los de resistencia a los antibióticos, como ampicilina y kanamicina. Algunos vectores plasmídicos poseen un segundo marcador de selección, como, por ejemplo, el gen *lacZ*, que se encuentra en el operón Lac de *E. coli*, que permite identificar a las células que introdujeron el vector de acuerdo a su color. También existen marcadores, como el gen reportero de la proteína verde fluorescente (*green fluorescent protein*, GFP) que permite que la célula

transformada emita fluorescencia cuando se expone a luz ultravioleta (Glick *et al.*, 2010, Sandoval Rodríguez *et al.*, 2013).

4.1.1.5. Número de Copias del Plásmido

Otra consideración importante a la hora de planificar la expresión recombinante de una proteína es la expresión basal del plásmido seleccionado, lo que se conoce como número de copias del plásmido y está determinado por el origen de replicación, el cual establece las dosis de genes accesibles para la expresión proteica. El número de copias para plásmidos de expresión en *E.coli* comúnmente se encuentra entre bajo (2 a 20) y alto (20 a 40). En general, se desea un alto número de copias para la máxima expresión de genes, pero, a veces, esto puede conducir en una recarga metabólica que afecta negativamente la biomasa y el rendimiento final. Debe considerarse que la presencia de promotores fuertes en plásmidos con un alto número de copias originará interferencias en la viabilidad celular (Correa y Oppezzo, 2011).

4.1.1.6. Clonación del Gen Blanco

Por lo expuesto, la obtención de una PR soluble, depende de una matriz compleja de muchas variables, entre ellas, la adecuada elección de una estrategia de clonado apropiada es fundamental. Esta selección dependerá de varios factores, como el número de genes blanco a ser expresados, las variantes del gen de interés a ser evaluadas, el uso de múltiples vectores conteniendo diferentes blancos o promotores, el uso de diferentes hospedadores de expresión, así como el equipamiento de laboratorio y presupuesto disponible. (Ortega *et.al.*, 2019)

En la presente investigación se seleccionó el método de clonado Restriction Free (RF) que permite la introducción precisa de fragmentos de ADN en cualquier posición deseada en un plásmido circular, independientemente de la secuencia. Varias manipulaciones moleculares pueden ser realizadas mediante clonado RF, además de la inserción del gen de interés en un vector deseado. Esto incluye el clonado simultáneo de varios fragmentos de ADN en diferentes posiciones de un vector destino, el clonado paralelo de algunos productos de PCR en diferentes vectores, como también la introducción simultánea de deleciones, inserciones o mutaciones en varias posiciones (Ortega *et.al.* 2019).

En el clonado RF (Figura 3), se diseñan 2 primers quiméricos (ChiFor y ChiRev) que llevan en sus extremos 5` secuencias complementarias al sitio de inserción en el vector destino (usualmente 30 nucleótidos), seguido de secuencias blanco específicas (usualmente 20-25 nucleótidos) en su extremo 3`. Las secuencias superpuestas son diseñadas para tener similares temperaturas de melting para mejorar la exactitud de la integración. Estos primers se usan en una primera PCR para la amplificación exponencial del gen blanco, generando un megaprimer,

que corresponde al gen de interés flanqueado en ambos extremos por secuencias superpuestas al sitio de integración en el vector destino. El megaprimero es purificado y usado en una segunda PCR (reacción RF) en combinación con el vector destino circular para la amplificación lineal de todo el plásmido y la integración del gen de interés en el sitio deseado. Solo las polimerasas PCR de alta fidelidad con actividad de no-desplazamiento de la cadena se pueden usar para el clonado RF a efectos de evitar la introducción de errores y la formación de largos concatámeros. El plásmido destino parental es removido por el tratamiento con una enzima que corta solo secuencias metiladas/hemimetiladas como la DpnI y el nuevo plásmido doble hebra sintetizado, que contiene 2 mellas (nicks), es utilizado para la transformación en *E.coli* donde el ADN mellado puede ser sellado por la actividad enzimática endógena.

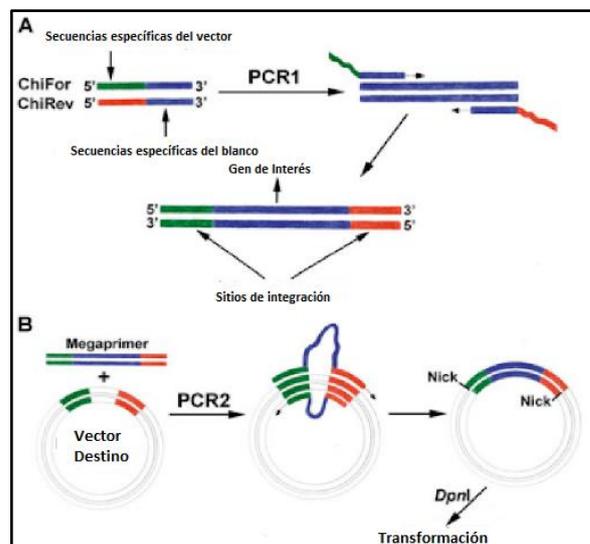


Figura 3. Esquema de la metodología de clonado RF. Imagen adaptada de Ortega *et.al.*, 2019.

4.1.2. Cepas Hospedadoras de *E.coli*

La selección de una cepa de *E.coli* puede tener un profundo impacto en la producción de PRs, ya que provee el trasfondo genético en donde ocurre la síntesis proteica. Las cepas de expresión generalmente carecen de las principales proteasas naturales, mantienen la estabilidad del vector de expresión y suministran los elementos génicos necesarios para que la expresión sea eficiente y regulable. La selección de la cepa de expresión se basa en las características de la proteína blanco, como, por ejemplo, si esta contiene puentes disulfuro, si es tóxica para el huésped, o si tiene codones poco frecuentes causados por el origen heterólogo taxonómico de la proteína blanco. Algunas cepas de *E.coli*, empleadas como hospederos para la expresión de PRs, tienen mutaciones en los genes de determinadas enzimas con actividad disulfuro reductasa para favorecer la formación de puentes disulfuro, necesarios para la conformación nativa de la proteína de interés. Otras cepas pueden co-expresar chaperonas para favorecer la expresión de PRs solubles y disminuir la formación de cuerpos de inclusión. En la actualidad se dispone de

una gran variedad de cepas de *E. coli* comerciales que han sido manipuladas genéticamente para diferentes aplicaciones. En la Tabla 2 se caracterizan brevemente algunas cepas comerciales de uso común para la expresión recombinante. (Jia y Jeon, 2016; Terpe, 2006; Correa y Oppezzo, 2015; Correa y Oppezzo, 2011).

Tabla 2. Cepas de *E. coli* empleadas comúnmente para la expresión de PRs (adaptado de Jia y Jeon, 2016; Terpe, 2006; Correa y Oppezzo, 2015; Correa y Oppezzo, 2011).

Cepa	Características	Uso recomendado
BL21	Deficiente en proteasas lon y OmpT. Carece de T7 ARN polimerasa.	Expresión recombinante de genes bajo control de promotores reconocidos por la ARN polimerasa de <i>E. coli</i> : lac, tac, trc y T5.
BL21(DE3)	Derivada de BL21. Contiene el profago λ DE3 que porta el gen de la T7 ARN polimerasa bajo control del promotor lacUV5, inducible por IPTG.	Expresión recombinante de genes bajo control de promotores T7, T7-lac y promotores reconocidos por la ARN polimerasa de <i>E. coli</i> .
BL21(DE3)pLysS	Derivada de BL21. Contiene el profago λ DE3 que porta el gen de la T7 ARN polimerasa bajo control del promotor lacUV5. Contiene el plásmido pLysS que porta el gen de la T7 lisozima (LysS), un inhibidor de la T7 ARN polimerasa que previene la expresión basal en cultivos no inducidos con IPTG.	Expresión recombinante con vectores basados en promotores T7, donde se requiera un mayor control de la expresión basal.
BL21(DE3)pLysE	Derivada de BL21. Contiene el profago λ DE3 que porta el gen de la T7 ARN polimerasa bajo control del promotor lacUV5. Contiene el plásmido pLysE, que expresa mayores niveles de T7 lisozima que pLysS.	Expresión recombinante con vectores basados en promotores T7, donde se requiera un elevado control de la expresión basal. Especialmente recomendada para la expresión de proteínas tóxicas para <i>E. coli</i> .
Rossetta(DE3) y Rossetta 2 (DE3)	Derivada de BL21. Contiene el profago λ DE3 que porta el gen de la T7 ARN polimerasa bajo control del promotor lacUV5. Contienen el plásmido pRARE o pRARE2, que portan genes para los ARNt de los codones AGG, AGA, AUA, CUA, CCC, CGA (pRARE) y AGA, AGG, AUA, CUA, GGA, CCC, CGG (pRARE2).	Expresión recombinante de genes eucariotas que contienen codones raros para <i>E. coli</i> . Se emplean además para la expresión recombinante de genes de especies procariontas alejadas filogenéticamente de <i>E. coli</i> , que también presenten codones raros.
BL21 CodonPlus y BL21 CodonPlus (DE3)	Derivadas de BL21-Gold. Contienen los genes para los ARNt de los codones AGG, AGA, AUA, CUA y CCC.	Expresión recombinante de genes que contienen codones raros para <i>E. coli</i> , tanto procariontas como eucariotas
Origami(DE3)	Derivada de K12. Contiene el profago λ DE3 que porta el gen de la T7 ARN polimerasa bajo control del promotor lacUV5. Contiene mutaciones en los genes trxB (tioredoxina reductasa) y gor (glutación reductasa), incrementando la formación de puentes disulfuro en el citoplasma.	Expresión recombinante de proteínas que requieran la formación de puentes disulfuro para alcanzar su conformación nativa.
Shuffle	Derivadas de K12 o B. Carece de las reductasas trxB y gor y posee una mutación supresora adicional (ahpC), lo que permite la formación de puentes disulfuro en el citoplasma. Además, expresan isomerasa DsbC dentro del citoplasma lo cual mejora la fidelidad de los puentes disulfuro en el citoplasma. Estricto control de la expresión por parte de lacI que permite la expresión de genes potencialmente tóxicos.	Expresión recombinante de proteínas que requieran la formación de puentes disulfuro para alcanzar su conformación nativa y expresión de genes potencialmente tóxicos.

4.1.3. Otras consideraciones a tener en cuenta en la búsqueda de estrategias para optimizar la expresión de PRs en *E.coli*

4.1.3.1. Expresión Génica de Codones Sesgados

Aunque el código genético es universal, tanto hospederos procariontes como eucariotes harán un uso preferencial de determinados codones sinónimos, lo cual está determinado por la abundancia relativa de los ARN de transferencia (ARNt) complementarios a cada codón. Por ello, debido a la naturaleza heteróloga de la proteína blanco, el gen blanco puede tener codones que están en baja abundancia en el hospedero. Para *E.coli* pueden ser codones raros algunos genes provenientes de eucariotes, arqueas e incluso eubacterias alejadas filogenéticamente, donde la presencia de estos codones puede tener una fuerte influencia en la expresión génica (Sørensen y Mortensen, 2005a).

La expresión de genes con codones raros puede conllevar a errores traduccionales por sustituciones aminoacídicas, disminución en la velocidad de traducción, cambio del marco de lectura, detención del crecimiento, terminación prematura de la traducción y bajo rendimiento de la producción proteica, entre otros. Hay evidencias que señalan que la estructura secundaria de la proteína podría estar relacionada con la frecuencia de uso de los ARNt. En *E.coli* así como en eucariotes, la proteína nascente se pliega co-traduccionalmente. En el túnel ribosomal un pequeño cambio en la tasa de traducción del ARNm podría impactar fuertemente en el desarrollo de la estructura secundaria de la proteína nascente. Estos cambios en la tasa de traducción dependen de los niveles relativos de ARNt (Agnov, *et al.*, 2008).

Existen dos estrategias alternativas para solucionar la presencia de codones raros en los genes de interés. La primera consiste en la síntesis del gen entero con codones optimizados para el sistema de expresión a utilizar. Esta estrategia comúnmente es exitosa, pero, a la vez, es costosa. La segunda estrategia consiste en la co-transformación del hospedero con un plásmido que porte los genes para los ARNt deficitarios. Existen disponibles varias cepas comerciales para co-expresar ARNt para codones raros, como Rossetta 2 y Rosetta2(DE3), BL21CodonPlus(DE3), que llevan incorporados plásmidos con un origen de replicación distinto a los usados en la mayoría de los vectores de expresión, lo cual garantiza la compatibilidad y estabilidad de ambos vectores. Estas estrategias pueden aumentar la velocidad de traducción, pero se ha visto también que en algunos casos conducen a la agregación y mal plegado de la proteína, lo cual sugiere que las pausas en la traducción son necesarias para el correcto plegado proteico (Correa y Oppezzo, 2011).

4.1.3.2. Localización Proteica en el Microambiente Bacteriano

Las PRs en principio pueden ser dirigidas a tres localizaciones celulares distintas: citoplasma, periplasma y medio extracelular (medio de cultivo). Cada una de estas localizaciones tendrá ventajas y desventajas.

Expresión Recombinante Dirigida al Citosol

Normalmente, la expresión y mantenimiento en el citosol es preferible debido a los mayores rendimientos de PR, la mayor simplicidad de los vectores de expresión o en algunos casos, la ventaja que ofrecen los cuerpos de inclusión formados. Sin embargo, en ocasiones, resulta una mejor opción, la exportación de la proteína hacia el espacio periplasmático o el medio extracelular. Algunas proteínas no logran recuperar su plegamiento nativo y actividad biológica tras un procedimiento de purificación desnaturalizante a partir de cuerpos de inclusión. En otros casos el ambiente fuertemente reductor del citoplasma bacteriano impide el correcto plegamiento de proteínas heterólogas cuya estructura terciaria depende de la formación de puentes disulfuro (Sørensen y Mortensen, 2005a; Makrides, 1996).

Expresión Recombinante Dirigida al Periplasma y Formación de Puentes Disulfuro

El espacio periplasmático ofrece un ambiente oxidante favorable para la formación de puentes disulfuro, disminuye la posibilidad de degradación proteolítica dado el menor número de proteasas celulares presentes y simplifica la purificación, al concentrar el producto deseado en un compartimiento celular mucho menos rico en proteínas celulares que el citosol (Sørensen y Mortensen, 2005a; Makrides, 1996).

Los puentes disulfuro son esenciales para el correcto plegamiento, estabilidad y/o función de las proteínas, por lo tanto, es importante tenerlos en cuenta cuando se expresan genes blanco. La presencia de ellos puede ser predicha por servidores web para estimar si la proteína blanco puede requerir de modificaciones post-traduccionales. Estos puentes son promovidos por la acción de las proteínas disulfuro isomerasas (PDI) en eucariotas y sus análogas Dsb en procariontes; se forman en ambientes oxidantes como el retículo endoplasmático (RE) de eucariotas o el periplasma de procariontes, donde el estado de oxido-reducción de estos ambientes es regulado, a su vez, por oxidoreductasas como tiorredoxina reductasa (trxB) y glutatión oxidoreductasa (gor). Debido a que el periplasma es un ambiente más oxidante que el citoplasma de *E.coli*, dirigir la proteína hacia el periplasma es una buena estrategia para expresarla con la correcta formación de sus puentes disulfuro. Dirigir la expresión proteica hacia este compartimento puede lograrse añadiendo un péptido señal en el extremo N-terminal de la secuencia blanco que, dependiendo de su naturaleza, puede utilizar distintas vías de traslocación que transportan proteínas a través de la membrana plasmática como precursores no plegados.

Una gran variedad de péptidos señal procariontes se han utilizado con éxito, sin embargo, el proceso de translocación es complejo y la presencia del péptido señal no siempre asegura un transporte eficiente al periplasma, lo cual puede repercutir en los rendimientos obtenidos. Estas vías de traslocación son: la vía Sec (traslocación post-traducciona l relativamente lenta) o SRP (traslocación co-traducciona l rápida). También existe otra vía llamada Twin-Arginina (vía Tat) que cataliza la traslocación de proteínas en su estado plegado. De todas formas, un inconveniente común de la expresión en el periplasma es que, si el gen blanco es expresado en tasas elevadas, la maquinaria de traslocación de la bacteria puede saturarse, generando efectos tóxicos en las células y reduciendo el rendimiento final de la proteína blanco. Una forma de resolver este problema es utilizando cepas de *E.coli* comerciales como Lemo21(DE3) (New England Biolabs), que permite ajustar finamente las tasas de transcripción del promotor T7 para evitar la saturación de la maquinaria de translocación (Correa *et al.*,2015; Makrides, 1996; Correa *et al.*, 2014).

Otra forma de conseguir la correcta formación de puentes disulfuro nativos es cambiar el estado redox del citoplasma de *E.coli* a un ambiente más oxidante. El ambiente reductor encontrado en el citoplasma de *E.coli* es mantenido por la acción de vías que involucran glutatión oxido-reductasa (*gor*) y tiorredoxina reductasa (*trxB*), por lo que se han desarrollado cepas de *E.coli* mutantes simples (*trxB*-) y dobles (*trxB*-/*gor*), como AD484 y Origami, respectivamente. Estas cepas potencian la producción de proteínas que contienen puentes disulfuro en el citoplasma de *E.coli*. Otra cepa que mejora considerablemente la formación de puentes disulfuro en el citoplasma de *E.coli* es la cepa comercial Shuffle, esta expresa la isomerasa DsbC en el citoplasma y contiene mutaciones *trxB*-/*gor*. Una forma alternativa de expresar la isomerasa DsbC en el sistema de expresión es fusionando su secuencia con la secuencia blanco de la proteína de interés. Por otro lado, se ha utilizado la co-expresión de la proteína Erv1p de levaduras, que es un catalizador de la formación de puentes disulfuro, sin la necesidad de interrumpir las vías reductoras del hospedador (Correa y Oppezzo, 2011; Correa y Oppezzo, 2015).

Expresión Recombinante Dirigida al Medio Extracelular

La secreción de la PR al medio extracelular comparte muchas ventajas logradas con la translocación al periplasma y favorece sustancialmente el proceso de purificación para algunas PRs. Por lo general, la secreción de la PR se logra empleando los sistemas de secreción extracelular existentes en el hospedero para otras proteínas celulares. *E.coli* normalmente secreta muy pocas proteínas y sus sistemas de secreción pueden ser ineficientes, por lo que se suelen obtener bajos rendimientos de PR. Otros sistemas de expresión bacterianos, como *B.subtilis*, son capaces de secretar grandes cantidades de proteína al medio de cultivo favoreciendo su purificación en estado soluble y nativo. En comparación con los hospederos

bacterianos, los sistemas de expresión eucariotas como levaduras o células de mamífero incrementan significativamente el rendimiento de PRs secretadas al medio (Terpe, 2006).

4.1.3.3. Evaluación de las Condiciones de Cultivo

Las condiciones de crecimiento (temperatura, pH, aireación), el medio de cultivo o el sistema de fermentación, influyen no solo en los niveles de biomasa, sino también en los rendimientos de PR. Entre los medios de cultivo para la expresión proteica los más utilizados son: medio Luria (LB), medio Terrific Broth (TB), medio de extracto de levadura y triptona (2YT), super caldo (SB), medio de autoinducción, entre otros.

La temperatura de cultivo contribuirá en la actividad de toda la maquinaria metabólica de la célula, incluyendo los mecanismos de transcripción, traducción, proteólisis, expresión, actividad de chaperonas, etc. En muchas ocasiones será recomendable cultivar hasta la fase logarítmica a 37°C y luego realizar la inducción de la expresión a 15°C-25°C, con lo cual se disminuirá la actividad de proteasas y la velocidad de traducción, y se aumentará la expresión de chaperonas. Todo ello favorecerá el correcto plegamiento y estabilidad de la PR, evitando, en muchos casos, la formación de cuerpos de inclusión y, consecuentemente, hará posible incrementar el rendimiento de la proteína soluble. Por lo tanto, es altamente recomendado evaluar diferentes temperaturas de inducción cuando se buscan condiciones óptimas para la expresión soluble de proteínas. Pero es necesario tener en cuenta que menores temperaturas también pueden disminuir la biomasa, lo cual podría obstaculizar el rendimiento final (Correa y Oppezzo, 2011; Sørensen y Mortensen, 2005b).

4.1.3.4. Disminución o Tolerancia de la Toxicidad en el Hospedero por Proteínas Tóxicas

La expresión recombinante de algunos genes heterólogos puede resultar extremadamente tóxica para el hospedero, interfiriendo con su fisiología y afectando el crecimiento e incluso la muerte del cultivo celular. La toxicidad de estos genes puede variar desde moderada hasta muy elevada, según el nivel de expresión y, por consiguiente, la cantidad de PR que es capaz de interferir con la fisiología del hospedero y afectar su crecimiento o su muerte. Varias estrategias son empleadas para lograr expresar y producir estas proteínas tóxicas, entre ellas, el empleo de promotores altamente regulados, la selección de cepas hospederas más resistentes al producto potencialmente tóxico, el empleo de promotores antisentido, la generación de proteínas de fusión que neutralicen la toxicidad, la utilización de vectores de bajo número de copias, empleo de sistemas de expresión libres de células, entre otros (Terpe, 2006).

4.1.3.5. Reducción de la Degradación Proteolítica

La proteólisis es un proceso selectivo y altamente regulado que juega un rol vital en la fisiología celular. *E.coli* contiene un gran número de proteasas localizadas en el citoplasma, periplasma y membranas interna y externa, que participan en actividades metabólicas y en la eliminación selectiva de proteínas anormales. La maquinaria proteolítica actúa como consecuencia de síntesis incompletas, mutaciones por sustitución de aminoácidos, síntesis excesiva de subunidades de complejos multiméricos, daño post-traducciona l por oxidación o ataque de radicales libres, entre otros. En principio, una proteína heteróloga podría fácilmente ser reconocida como anormal y ser degradada por la célula; de hecho, en muchos casos, la degradación proteolítica disminuye apreciablemente los rendimientos de PR. Varias estrategias han sido aplicadas a fin de minimizar la degradación proteolítica durante la expresión de PRs, como el empleo de cepas deficientes en proteasas (como la BL21 de *E.coli*), construcción de proteínas de fusión, translocación fuera del citosol, crecimiento o inducción a bajas temperaturas (la mayoría de las proteasas tienen un actividad óptima a 37°C), co-expresión con chaperonas, sustitución de aminoácidos específicos en el sitio de clivaje de la proteasa, optimización de las condiciones de fermentación, entre otras (Makrides, 1996).

4.1.3.6. Co-Expresión de Chaperonas Moleculares

La co-expresión de chaperonas moleculares que colaboran con el plegamiento puede incrementar la estabilidad conformacional y solubilidad de la proteína de interés o prevenir su degradación proteolítica. La co-expresión puede lograrse utilizando uno o varios vectores de expresión compatibles y, en general, se utiliza para mejorar la solubilidad de la proteína blanco o, en algunos casos, para colaborar en la formación y actividad óptima de complejos proteicos. Se ha sugerido que la producción de algunas PRs, podría abrumar a las chaperonas del hospedador, conduciendo a la acumulación de proteína blanco, como cuerpos de inclusión. Para minimizar la agregación y mejorar la solubilidad de muchas PRs puede utilizarse el conjunto de chaperonas DnaK/DnaJ/GrpE (KJE) o GroEL/GroEs (ELS), ClpB, IbpA e IbpB y el factor desencadenante asociado al ribosoma (Correa y Oppezzo, 2011).

4.1.3.7. Consideración de la Formación de Cuerpos de Inclusión

Usualmente uno de los principales objetivos de la expresión de PRs es la obtención de grandes cantidades de proteína de interés nativa y biológicamente activa. Pero en ocasiones, la maquinaria metabólica de la célula hospedera puede generar una respuesta de estrés, caracterizada por la formación de agregados insolubles en el citosol o periplasma de *E.coli*, conocidos como cuerpos de inclusión.

En las últimas décadas se ha visto, en varios casos, que los cuerpos de inclusión pueden estar hechos de proteínas nativas y activas. Esto abre la posibilidad del uso de aplicaciones posteriores donde la agregación proteica no es un impedimento, sino que facilita la producción y purificación. Por estos motivos, en ocasiones la formación de cuerpos de inclusión es explotada, por ejemplo, cuando la PR activa en el citosol es tóxica para el hospedador, por lo que una posible estrategia es introducir un tag que reduzca la solubilidad de la PR y permita dirigirla hacia el interior de los cuerpos de inclusión (Jia y Jeon, 2016; Sørensen y Mortensen, 2005b).

Sin embargo, en la mayoría de los casos, maximizar la obtención de la PR en forma soluble resulta más recomendable. Diversas estrategias han sido desarrolladas con este fin y en la mayor parte de los casos buscan evitar, minimizar o resolver la situación de estrés metabólico que ocurre en la célula. El cultivo o inducción a bajas temperaturas, el empleo de cepas de expresión comerciales, como ArticExpress (DE3), la modificación de los medios y/o condiciones de cultivo, la co-expresión de chaperonas, el uso de proteínas de fusión o la disminución del tamaño e hidrofobicidad del fragmento a expresar, son algunas de las estrategias más utilizadas para incrementar la solubilidad en la expresión de PRs (Jia y Jon, 2016; Sørensen y Mortensen, 2005b; Makrides, 1996).

4.1.3.8. Introducción de Proteínas o Tags de Fusión

Mediante ingeniería genética el gen blanco puede ser fácilmente clonado con diferentes tags de fusión que pueden aumentar la solubilidad proteica, el rendimiento final y facilitar la purificación proteica o procesamiento posterior. Estos pueden clasificarse en tags de afinidad o tags que mejoran la solubilidad.

Los tags de afinidad son tags cortos que pueden introducirse en N-ter o C-ter del blanco y son reconocidos por matrices o moléculas que sirven para la purificación por afinidad de la proteína fusionada. Entre los tags de afinidad, el HisTag es uno de los más utilizados para purificar PRs de *E.coli*. Este consiste en un pequeño tag de 0.84 Kda de 6-10 histidinas en tándem que puede interactuar reversiblemente con iones metálicos, comúnmente, Níquel o Cobalto, inmovilizados en una matriz metálica quelante. Esto permite condiciones de elución suave para el uso de un competidor como el imidazol. La cromatografía de afinidad de metales inmovilizados, (IMAC, del inglés: *immobilized metal affinity chromatography*), es la más utilizada como método de purificación de PRs. El HisTag tiene varias ventajas, como su pequeño tamaño y su naturaleza inerte, que lo hacen compatible con la mayoría de las aplicaciones posteriores. Debido a la estructura terciaria del HisTag no es necesaria la coordinación con el metal, por lo que es posible purificar la proteína mediante IMAC en condiciones desnaturizantes y subsecuentemente la proteína blanco es replegada (Correa y Oppezzo, 2015).

Un tag de afinidad alternativo es el Strep-tag-II; este también es un tag pequeño que consiste en 8 residuos aminoacídicos; al igual que el His-tag, es biológicamente inerte, proteolíticamente estable y no interfiere con el plegamiento proteico. Este sistema es altamente específico y es reconocido por streptactina contenida en el medio. La alta especificidad de esta unión hace a este tag una buena opción cuando se trabaja con proteínas que son expresadas en bajas cantidades (Correa y Oppezzo, 2015).

El Twin-Strep-Tag es una variante del Strep-tag-II que consiste en dos secuencias de Strep-tag-II conectadas por un enlace corto. Demostró ser el más apropiado para la purificación de muestras diluidas (Correa y Oppezzo, 2015).

Los tags que mejoran la solubilidad son proteínas extremadamente solubles y estables con actividades chaperonas o características termoestables que pueden ser fusionadas con el blanco para transferir algunas de estas características, mejorando el plegado y/o rendimiento final de la proteína blanco. Usualmente estos tags son expresados como fusiones N-terminales, y, luego de la escisión y remoción de la proteína de fusión, la proteína blanco se mantiene soluble. Entre las proteínas que mejoran la solubilidad podemos encontrar la proteína de unión a maltosa (MBP), la glutatión S- transferasa (GST), la tiorredoxina A (TrxA), la disulfuro isomerasa C (DsbC), la proteína pequeña modificadora de tipo ubiquitina (SUMO) y la sustancia A de N-utilización (NusA). Una característica atractiva de MBP y GST es que pueden usarse tanto como tags de afinidad o como mejoradores de solubilidad. Como Trx, SUMO, DsbC y NusA no son fáciles de purificar; estas suelen conjugarse con tags de afinidad como HisTag o Strep-tag-II para mejorar la purificación proteica (Correa y Oppezzo, 2015).

Como los tags de fusión pueden interferir con estudios estructurales y funcionales de la proteína expresada, usualmente es necesario remover el tag luego de la purificación de la proteína blanco. Esto es posible gracias a la incorporación de una secuencia aminoacídica corta, flexible e hidrofóbica, entre el tag de fusión y la proteína de interés, que pueda ser reconocida por una proteasa específica, como enteroquinasa, factor Xa, trombina, proteasa PreScission y la proteasa del virus grabado del tabaco (TEV) (Correa y Oppezzo, 2011).

4.2. Sistemas de Expresión Eucariotas - Sistema de Expresión EXPI 293F

Como se mencionó anteriormente, si bien las ventajas de expresar PRs en *E.coli* son numerosas, en ocasiones, cuando se busca expresar proteínas desafiantes, es necesario utilizar sistemas de expresión eucariotas, como levaduras, células de insectos, células de mamífero o incluso sistemas libres de células, ya descritos en la Tabla 1. Además, dependiendo de las cantidades y calidad requeridas de las proteínas blanco, una sola línea celular del hospedador puede ser el sistema de expresión óptimo, cuando otros fallan en la expresión del mismo blanco (Ortega *et al.*, 2018).

4.2.1. Modificaciones Post-Traduccionales

La obtención de un producto soluble y homogéneo puede no ser el único requerimiento en la producción de PRs, como es el caso de muchas proteínas terapéuticas, las cuales requieren algún tipo de procesamiento post-traduccionales para su actividad y estabilidad. Estas modificaciones aminoacídicas incluyen glicosilación, fosforilación, acetilación, sulfonación, acilación, gama-carboxilación y la adición de ácidos grasos en C14 y C16 por miristoilación y palmitación, respectivamente (Glick *et al.*, 2010). Entre estas, la glicosilación es la más común y compleja modificación de proteínas eucariotas de superficie y secretadas (Khan *et al.*, 2016).

4.2.1.1. Glicosilación y Perfiles de Glicosilación

La glicosilación es la adición enzimática de oligosacáridos a cadenas polipeptídicas nacientes en el RE. A su vez, el oligosacárido adherido se modifica aún más por la acción de glucosidasas y glicosiltransferasas dentro del RE y el aparato de Golgi. Las reacciones de modificación que ocurren en el RE están altamente conservadas entre eucariotas inferiores y superiores, mientras que las reacciones que tienen lugar dentro del aparato de Golgi varían entre especies y tipos de células (Khan *et al.*, 2016).

Las glicosilaciones más comunes implican la unión de azúcares específicos al grupo hidroxilo de una serina o una treonina (O-glicosilación) o a un grupo amino de una asparagina (N-glicosilación). La N-glicosilación suele ser esencial para la expresión y plegamiento de la glicoproteína; ocurre co-traduccionalmente en el RE, en la secuencia consenso Asn-X-Ser/Thr/Cys-X (donde X puede ser cualquier aminoácido excepto prolina) y luego la proteína con el núcleo de glicano se dirige hacia el aparato de Golgi para ser más procesada. La elaboración del núcleo de glicano conduce a un gran número de estructuras posibles, las que pueden clasificarse como: ricas en manosa, complejas o híbridas. Por otro lado, la O-glicosilación, ocurre en el aparato de Golgi, generalmente en regiones del polipéptido que contienen un gran número de residuos de serina, treonina y prolina, conocidos como dominios de mucina. Como la O-

glicosilación ocurre en el Golgi, esta modificación tiene poca relación con las primeras etapas del plegamiento proteico, por lo que los sitios que se encuentran O-glicosilados pueden ser eliminados de la proteína mediante mutagénesis dirigida al sitio de residuos aceptores Ser/Thr (Nettleship, 2012).

Aproximadamente, el 50% de las proteínas nativas humanas son glicoproteínas y el perfil de glicosilación de estas células es extremadamente diferente al de otras, en especial aquellas muy distantes en términos evolutivos, como bacterias, levaduras, plantas o insectos (Figura 4). Por esto es que se han propuesto numerosas estrategias para adaptar las vías de glicosilación y humanizar el perfil de glicosilación de distintos huéspedes de expresión (Khan *et al.*, 2016).

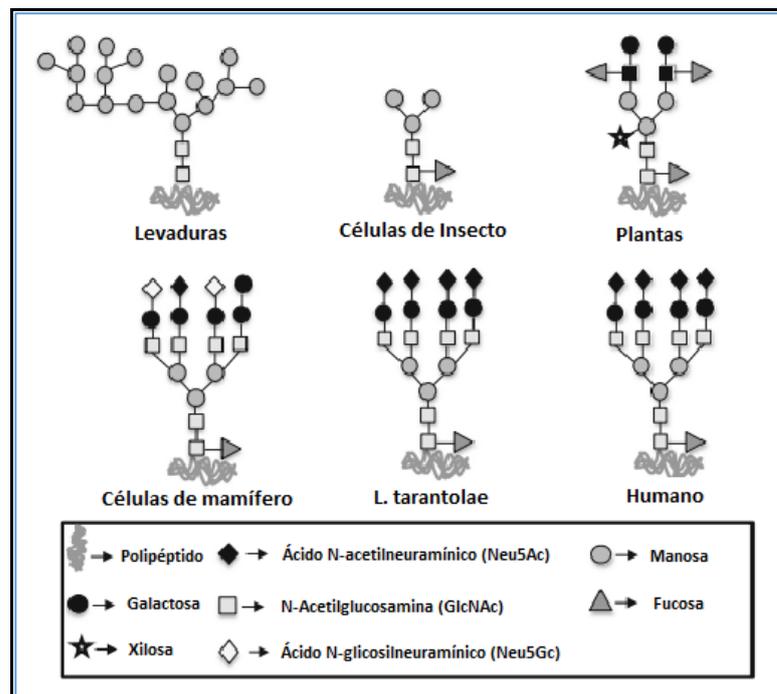


Figura 4. Comparación del perfil de glicosilación de glicoproteínas recombinantes derivadas de levaduras, células de insectos, plantas, células de mamífero, *L. tarantolae* y células humanas, adaptada de Khan *et al.*, 2017.

Los residuos de glicano influyen en el plegamiento, especificidad de sitio, homeostasis celular y regulación inmunológica de las glicoproteínas (Khan *et al.*, 2016). El correcto plegamiento de proteínas en células eucariotas, influenciado en gran parte por la glicosilación, requiere la asistencia de las chaperonas BiP y calnexina en el RE, que unen polipéptidos nacientes, y de las disulfuro isomerasas que catalizan la formación de enlaces disulfuro entre residuos de cisteína adyacentes. Esto no solo es importante para que la proteína adquiera una configuración para su actividad óptima, sino también para proteger regiones de ella que puedan ser reconocidas por proteasas que la degraden. Desafortunadamente, no existe un hospedador

celular eucariota universal que realice las modificaciones correctas en cada proteína. En algunos casos, un hospedador celular puede generar una proteína con un patrón de glicosilación que la vuelva extremadamente antigénica o que carezca de su correcta función. De todas formas, aunque una PR no cumpla con las propiedades estrictas que son necesarias para ser un agente terapéutico, puede ser útil para su uso en la investigación o en la industria.

Por lo expuesto, el uso de sistemas de células de mamífero ha aumentado mucho en los últimos tiempos, ya que producen de manera confiable proteínas eucariotas con un complemento completo de modificaciones post-traduccionales y, actualmente, alrededor de la mitad de las proteínas terapéuticas comerciales disponibles son producidas en células de mamífero. Sin embargo, existen varios desafíos importantes para la producción de altos niveles de proteínas heterólogas (Glick *et al.*, 2010).

4.2.2. Diseño de Vectores para Sistemas de Expresión Basados en Células de Mamífero

Los requerimientos básicos para la expresión de PRs en hospedadores eucariotas son similares a aquellos requeridos en procariontes y, al igual que en estos, es crucial la optimización de los vectores de expresión. La mayoría de los vectores construidos para la expresión de genes heterólogos en células de mamífero están basados en genomas de virus que infectan células de mamífero. El primer vector para este fin se basó en el virus de simio 40 (SV40), que puede replicarse en varias especies de células de mamífero. Su genoma es una molécula de ADN doble hebra de 5.2 kb, la cual carga genes de expresión temprana en el ciclo de infección, que funcionan en la replicación del ADN viral (genes tempranos), y genes que se expresan luego en el ciclo de infección, que funcionan en la producción de proteínas de la cápside viral (genes tardíos). Para su uso como vectores de clonado, algunos de los genes tempranos y tardíos son removidos y reemplazados con el gen blanco bajo control de señales de expresión de mamíferos apropiadas. Aunque muchos vectores de clonado están basados en ADN SV40, su uso está restringido a pequeños insertos porque solo una cantidad limitada de ADN puede ser empaquetada en la cápside viral. Otros vectores que pueden acomodar grandes cantidades de ADN clonado son derivados de adenovirus o papiloma-virus bovino (Glick *et al.*, 2010).

La regulación de la transcripción en los genomas eucariotas implica la interacción coordinada de múltiples elementos genéticos. El inicio de la transcripción está mediado por las interacciones de los factores de transcripción con sus elementos promotores y potenciadores asociados. Estos últimos son secuencias que pueden ubicarse miles de bases *downstream* o *upstream* del promotor que mejoran la actividad transcripcional cuando se unen a factores de transcripción. (Makrides, 2003).

4.2.2.1. Promotores

Algunos promotores utilizados en la expresión de PRs en mamíferos son transcripcionalmente activos en una amplia gama de tipos de células y tejidos. La mayoría, sin embargo, exhibe selectividad tisular, una propiedad que debe considerarse cuidadosamente antes de la construcción de vectores de expresión para la producción proteica (Glick *et al.*, 2010).

Los promotores utilizados para la producción de PRs en células de mamífero deben ser fuertes y, preferiblemente, activos en una amplia gama de tipos celulares, para permitir la evaluación cualitativa y cuantitativa de la PR. Suelen dividirse en dos clases, los que funcionan de manera constitutiva y los que están regulados por inducción o desrepresión (Tabla 3). Los promotores inducibles deben exhibir un mínimo nivel de actividad transcripcional basal y ser capaces de una inducción sustancial con un inductor no tóxico de una manera simple y rentable. Estos son preferibles para la producción de proteínas que pueden ser tóxicas para la célula huésped, para el estudio de la regulación génica durante el desarrollo en animales transgénicos y para aplicaciones experimentales y terapéuticas de transferencia génica (Makrides, 2003).

Tabla 3. Algunos promotores constitutivos e inducibles para la expresión recombinante en células eucariotas. (Makrides, 2003)

Promotor	Fuente	Propiedades
SV40	virus simio 40	Expresión constitutiva; en algunas líneas celulares inducible con éster de forbol. Amplio rango de hospedadores y tipos de células. En las líneas celulares COS que expresan el antígeno T, se logra un alto número de copias del vector.
hCMV-IE mCMV-IE	Genes promotores tempranos inmediatos de citomegalovirus humano y de ratón	Expresión constitutiva de alto nivel. Amplio rango de hospedadores y tipos de células.
VRS-LTR	Repetición terminal larga del virus del sarcoma de Rous	Expresión constitutiva de alto nivel en líneas celulares murinas y aviares.
MMTV-LTR	Virus del tumor mamario de ratón	Inducible con glucocorticoides. Nivel moderado de inducción transcripcional.
MoMLV-LTR	Virus de la leucemia murina de Moloney	Inducción transcripcional de moderada a fuerte.
Ad2MLP-TPL	Adenovirus principal promotor tardío y líder tripartito	Expresión constitutiva de alto nivel. Amplia gama de huéspedes.
centro	Gen de la ubiquitina C humana	Expresión constitutiva de alto nivel en una amplia gama de tejidos y tipos de células.

Promotor	Fuente	Propiedades
hEF-1 α	Gen de la subunidad 1 α del factor de elongación de la traducción humana	Expresión constitutiva de alto nivel. Amplio rango de hospedadores y tipos de células.
mPGK	Gen de la fosfoglicerato quinasa de ratón	Expresión constitutiva de alto nivel. Amplio rango de hospedadores y tipos de células.
mMT-I	Gen de la metalotioneína I de ratón	Inducible con Cd ⁺⁺ , Zn ⁺⁺ , ésteres de forbol. Promotor "con fugas"
hMT-II	Gen de la metalotioneína II humana	Inducible con Cd ⁺⁺ , Zn ⁺⁺ , ésteres de forbol. Promotor "con fugas"
hMT-IIA (mutante)	Gen de la metalotioneína II humana	Inducible con Cd ⁺⁺ , Zn ⁺⁺ , ésteres de forbol. Alta inducibilidad, baja actividad basal.
hIFN- α	Gen de interferón- α humano	Inducible con virus.
β -actina	Gen de la β -actina de pollo	Expresión constitutiva de alto nivel en una amplia gama de tejidos y tipos de células.
β -globina	gen de la globina β	Tejido específico para células eritroides β adultas. Posibles aplicaciones de la terapia génica.

4.2.2.2. Intrones

La mayoría de los genes de los eucariotas superiores contienen intrones que se eliminan durante el procesamiento del ARN. Las construcciones genómicas con intrones se expresan de manera más eficiente que las que carecen de ellos, un efecto que se cree que es debido a una mayor tasa de poliadenilación del ARN y al transporte nuclear acoplado al empalme de ARN. Aunque muchas construcciones de ADNc que carecen de intrones se pueden expresar de manera eficiente en células de mamífero, la inclusión de intrones puede aumentar la expresión de 10 a 20 veces. La ubicación de los intrones en el extremo 3' de la unidad de transcripción puede conducir a un empalme aberrante, por lo tanto, es preferible colocar los intrones en el extremo 5' del marco de lectura abierto. Sin embargo, el uso de intrones exige un diseño de vector cuidadoso, ya que es posible que las señales de corte y empalme crípticas puedan causar un procesamiento aberrante de la transcripción de ARNm, lo que resulta en niveles de expresión reducidos y productos proteicos defectuosos (Makrides, 2003).

4.2.2.3. Señales de Poliadenilación

La mayoría de los ARNm nacientes de eucariotas poseen una cola poli(A) en sus extremos 3', un elemento importante para la estabilidad y traducción del ARNm, que se agrega durante la escisión del transcrito primario y una reacción de poliadenilación acoplada. Las señales para la poliadenilación de los ARNm de mamíferos consisten en una secuencia AAUAAA altamente

conservada, ubicada unos 20 a 30 nucleótidos *upstream* del extremo 3' del ARNm, y una secuencia rica en GU no conservada inmediatamente *downstream* del sitio de poliadenilación (Makrides, 2003).

4.2.2.4. Terminadores de la Transcripción

La transcripción continua de un promotor a través de una segunda unidad de transcripción reduce la expresión del segundo gen, un fenómeno conocido como interferencia transcripcional. Por ello, para minimizar esta interferencia se pueden insertar señales de terminación de la transcripción entre dos unidades de transcripción. (Makrides, 2003).

4.2.2.5. Elementos de Control Traduccional

Para la expresión óptima de ADNc eucariota es necesario considerar cuidadosamente varias características estructurales, como el contexto de nucleótidos alrededor del codón de inicio de la traducción y las regiones no traducidas (UTR) 5' y 3', que están involucradas en muchos procesos post-transcripcionales, que controlan la localización del ARNm, estabilidad y eficiencia de traducción. El inicio de la traducción en organismos eucariotas superiores depende de la secuencia Kozak, una secuencia específica de nucleótidos alrededor del codón de inicio (AUG) en el ARNm. Esta secuencia se coloca en el extremo 5' del gen de interés y, además, se agrega un codón de stop para asegurar la terminación de la traducción en el lugar correcto. Por otro lado, contiene una secuencia en 5' y en 3' de regiones no traducidas (UTRs), importantes para una eficiente traducción y para la estabilidad del ARNm. Sin embargo, se ha documentado en procariontes y eucariotes, que las regiones ricas en GC pueden formar estructuras de horquilla estables capaces de inhibir el inicio de la traducción. Una solución para este potencial problema es la eliminación de la región 5'-UTR antes de la inserción de ADNc en los vectores de expresión, pero con la desventaja de que la región 5'-UTR podría contener elementos potenciadores de la traducción. Por otro lado, la regulación de la traducción de ciertos ARNm está mediada por elementos ricos en AU que se unen a proteínas ubicadas en la región 3'-UTR, y además, secuencias específicas presentes en esta región son capaces de influenciar la desestabilización del ARNm. (Makrides, 2003).

4.2.2.6. Codones de Terminación

La terminación de la traducción en genes de mamífero puede modularse mediante nucleótidos adicionales a los de los codones de terminación de los trinucleótidos. Tanto *in vitro* como *in vivo*, la base que sigue al codón de terminación influye en la eficiencia de la terminación

de la traducción, y los tetranucleótidos con una purina en la cuarta posición son más efectivos como señales de terminación que aquellos con una pirimidina (Makrides, 2003).

4.2.2.7. Marcadores de Selección

Además de la presencia de un marcador de selección para la propagación de vectores y/o las etapas de clonado en bacterias, los vectores de expresión de mamíferos contienen marcadores para la selección de células transfectadas para la preparación de líneas celulares estables y la amplificación de genes. En su mayoría, los sistemas que son usados para seleccionar células de mamífero transfectadas son los mismos que los usados en otros hospederos celulares eucarióticos. De hecho, varios marcadores de selección se han adaptado para células eucariotas. Los esquemas de selección son diseñados no solo para identificar células transfectadas, sino también para incrementar la producción de proteínas heterólogas por amplificación del número de copias del vector de expresión (Makrides, 2003) (Tabla 4).

Tabla 4. Marcadores de selección más utilizados en vectores de expresión de mamífero. Tabla extraída y modificada de Li *et al.* (2010).

Marcador de selección	Reactivo de selección
Marcadores de selección metabólicos	
Dihidrofolato reductasa (DHFR)	Metionina sulfoximina (MSX)
Glutamina sintetasa (GS)	Metotrexato (MTX)
Marcadores de selección antibióticos	
Puromicina acetiltransferasa	Puromicina
Blasticidina deaminasa	Blasticidina
Histidinol Deshidrogenasa	Histidinol
Higromicina fosfotransferasa	Higromicina
Gen de resistencia a Zeocina	Zeocina
Gen de resistencia a Bleomicina	Bleomicina
Aminoglicósido fosfotransferasa	Neomicina (G418)

4.2.2.8. Péptidos Señal y Proteínas de Fusión

Entre las principales ventajas de las células de mamífero se encuentra el hecho de que poseen una avanzada maquinaria de secreción. Esto posibilita la producción de PRs en forma

soluble y permite dirigirlas al medio extracelular, lo que beneficia los pasos subsiguientes de purificación.

Generalmente, en los eucariotas superiores, la vía clásica es la responsable de la secreción de la mayoría de las proteínas; en esta vía, la proteína que porta un péptido señal en su extremo N-terminal es sintetizada en los ribosomas. Conforme crece la cadena polipeptídica, la región del péptido señal conduce la translocación de la misma hacia la membrana del RE (Makrides, 2003; Owji *et al.*, 2018).

Las proteínas recién sintetizadas ingresan a la vía secretora a través del RE, donde se modifican covalentemente y adquieren sus estructuras terciarias y, en la mayoría de los casos, cuaternarias adecuadas. El correcto plegamiento de las proteínas es catalizado en la luz del RE por las peptidil prolil cis/trans isomerasas, donde también hay chaperonas, PDI y otras enzimas que llevan a cabo modificaciones postraduccionales (Chevet *et al.*, 2001). Una vez sintetizada la pre-proteína, una endopeptidasa separa el péptido señal, produciendo una proteína madura que queda dentro de una vesícula secretora que luego es dirigida hacia el Golgi, donde se producen más modificaciones post-traduccionales. Finalmente, la vesícula se fusiona con la membrana celular para secretar la proteína al medio extracelular (Makrides, 2003; Owji *et al.*, 2018).

Que una proteína sea dirigida o no hacia la ruta de secreción depende de la presencia del péptido señal. Este es una secuencia corta, de entre 16 y 30 residuos, presente en el extremo N-terminal de la pre-proteína. Típicamente consta de tres regiones: una región amino-terminal cargada positivamente (región N), una región hidrofóbica central (región H) y una región polar carboxi-terminal (región C), seguida por el sitio de escisión de la peptidasa señal. La región de aminoácidos hidrofóbicos (entre 5 y 16 residuos) provoca una tendencia hacia la formación de una alfa-hélice que promoverá la translocación de la proteína. La correcta elección del péptido señal tendrá un gran impacto en la productividad en una determinada línea celular, estos péptidos suelen ser intercambiables y se han utilizado ampliamente para posibilitar la secreción de proteínas en células de mamífero (Makrides, 2003, Román *et al.*, 2016a; Owji *et al.*, 2018, Zhang *et al.*, 2005).

Por otro lado, se deben considerar cuidadosamente varios factores en el diseño de la secuencia de las proteínas de fusión. A menudo se fusiona un tag de afinidad en el extremo N-terminal de una proteína para facilitar la purificación, sin embargo, un tag en C-terminal tiene la ventaja de que solo se pueden purificar proteínas completamente traducidas, estrategia que requiere que el extremo C-terminal sea estructuralmente accesible. Por último, la separación de la proteína de fusión de la proteína diana es facilitada mediante un sitio de escisión de proteasa diseñado entre las secuencias de la proteína blanco y la proteína de fusión. (Makrides, 2003).

4.2.3 Transfección Transitoria vs. Transfección Estable

La transfección se refiere a la introducción de ADN o ARN doble hebra en células eucariotas y, generalmente, el término alude a células de mamífero específicamente. La introducción de ADN en una célula permite la expresión o la producción de proteínas utilizando la maquinaria propia de las células, mientras que la incorporación de ARN en una célula se utiliza para regular la producción de una proteína específica. El sitio de acción del ARN transfectado es el citoplasma, mientras que el ADN debe ser transportado al núcleo para que la transfección sea eficaz. Allí, el ADN plasmídico no se incorpora al genoma y puede expresarse transitoriamente por un período corto de tiempo (transfección transitoria) o puede ser incorporado en el ADN genómico (transfección estable), donde el cambio se pasa de célula a célula durante la división celular (Figura 6).

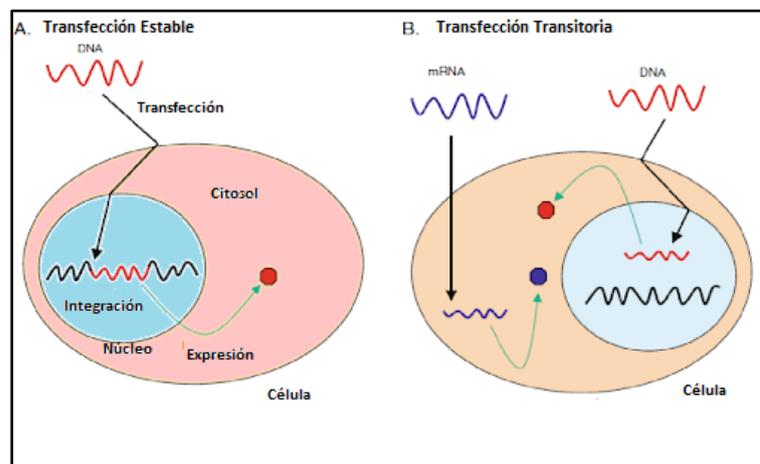


Figura 6. Esquema de los dos tipos de transfección. (a) Transfección estable: El ADN foráneo (rojo) es enviado al núcleo por el pasaje a través de la célula y la membrana nuclear y es integrado en el genoma del hospedero (negro), y expresado de manera sostenida. (b) Transfección transitoria: el ADN foráneo es enviado al núcleo, pero no se integra al genoma del hospedador. El ARNm (azul) es enviado al citosol donde es traducido. Los hexágonos corresponden a las proteínas expresadas. Imagen modificada de Artola *et al.* (2015).

La elección del tipo de transfección, estable o transitoria, así como los métodos que permiten la entrada del ADN foráneo a la célula, depende del tipo de célula y del propósito. El método ideal debe contemplar una alta eficiencia de transfección, baja toxicidad celular, mínimos efectos en la fisiología celular y, además, deben ser sencillos y reproducibles.

Para grandes cantidades de producción proteica son preferidas las células estables ya que proveen altos rendimientos con calidad consistente. De todas formas, en el estado temprano de un proyecto, donde se requieren miligramos de proteína, la transfección transitoria es preferida,

debido a su velocidad y menor costo. Además, provee una oportunidad de enfocarse en experimentos de ingeniería de proteínas, diseño de constructos o mejoras en los parámetros de crecimiento. Los sistemas transitorios ofrecen un producto con calidad adecuada para evaluación preclínica, acelerando así el estado de *Prueba de Concepto* cuando se proyectan múltiples candidatos. Esto puede lograrse antes de utilizar cantidades significativas de tiempo y recursos en la generación de líneas celulares estables (Tae *et al.*, 2010).

4.2.4 Metodologías de Transfección

Existen varias metodologías de transfección, tanto biológicas como -la transfección mediada por virus-, químicas, o físicas. Dentro de las metodologías químicas se encuentra la transfección a través de polímeros catiónicos, fosfato cálcico, lípidos catiónicos o aminoácidos catiónicos; donde los químicos cargados positivamente se unen a los ácidos nucleicos con carga negativa. De esta manera, se forman complejos con carga positiva que son atraídos hacia la membrana plasmática de la célula hospedadora, con carga negativa, e ingresan a ella por endocitosis; luego el ADN foráneo es translocado al núcleo. También se pueden introducir ácidos nucleicos a células de mamífero por medio de métodos físicos, como, por ejemplo, la micro-inyección directa, por biolística, por irradiación por láser, sonoporación, nanopartículas magnéticas, así como la electroporación descrita anteriormente para los sistemas procariotas (Kim y Eberwine, 2010).



Figura 7. Dispositivo Expert Stx utilizado para la transfección por el sistema MaxCyte.

La tecnología MaxCyte (Figura 7), aplicada en la presente investigación, es un ejemplo de método físico de transfección y se basa en la electroporación de flujo, la cual aprovecha la permeabilidad reversible de las membranas celulares en presencia de carga eléctrica, para crear un método eficaz como vía de administración para pequeñas moléculas, como ácidos nucleicos y proteínas.

Otra de las metodologías utilizadas en este trabajo, ha sido la transfección mediante el polímero catiónico polietilenimina (PEI) que, debido a su alta eficiencia de transfección, bajo costo y habilidad de escape de los endosomas, es una de las más utilizadas para la producción de PRs en sistemas eucariotas. El PEI funciona como vector para introducir el plásmido de expresión dentro de la célula, ya que los ácidos nucleicos libres presentan baja estabilidad química por su degradación enzimática y su eliminación por el sistema fagocítico. Entonces el

PEI es capaz de condensar estos ácidos nucleicos a través de interacciones electrostáticas, formando partículas que protegen el material génico y permiten la internalización en la célula por la vía endocítica. Al entrar a las células se neutralizan las cargas del complejo ADN/PEI, debido a una reorganización de los fosfolípidos, causando la liberación del ADN en el citoplasma (Figura 8).

Estas sustancias constan de una cabeza cargada positivamente, un segmento hidrófobo y un grupo enlazante; la cabeza positiva puede ser monovalente o multivalente, según la cantidad de grupos amino que contenga. Se han estudiado varias modificaciones, tanto de la naturaleza y densidad de los grupos catiónicos como del segmento hidrófobo, para mejorar su desempeño en la entrega de ácidos nucleicos. Si bien es considerado un estándar entre los métodos de transfección, en ocasiones, su citotoxicidad y baja biodegradabilidad limitan su aplicación, sobre todo cuando se quieren expresar proteínas para uso terapéutico (Díaz-Ariza *et al.*, 2018; Portolano *et al.*, 2014). En este estudio se ha utilizado PEI Max, una versión de PEI altamente eficiente, de rendimiento superior, baja citotoxicidad y compatible con una amplia gama de líneas celulares.

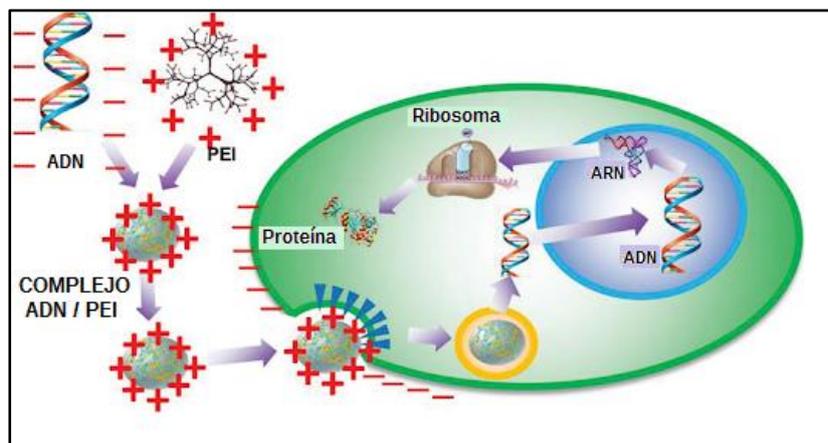


Figura 8. Formación del complejo ADN/PEI e internalización a la célula por vía endocítica. Imagen extraída de Yang *et.al.* (2017).

4.2.5 Líneas Celulares de Mamíferos

Existe una gran diversidad de líneas celulares de mamífero que se han desarrollado como hospederos para la producción heteróloga de PRs. Las células derivadas de riñón de mono Africano verde (COS), células de riñón de hámster bebé (BHK) y las células de riñón de embrión humano (HEK 293) son usadas para la expresión transitoria de genes y permiten la producción rápida de pequeñas cantidades de proteínas heterólogas, con el fin de evaluar su potencial como candidatos a drogas o probar la integridad de los constructos durante varias etapas del desarrollo del vector. Existen importantes características de estas células, como su receptividad a la transfección, su habilidad para crecer en cultivos en suspensión, sin la adición de suero que

contenga proteínas de origen animal y su adecuación a la producción de alta densidad y gran escala. Su habilidad de crecer en medio libre de suero no solo reduce costos, sino que también facilita la purificación de la proteína blanco y reduce los riesgos de contaminación con material de origen animal. Las células de ovario de hámster Chino (CHO) y las células de mieloma de ratón (SP2/0-Ag14) son las más comúnmente utilizadas para la expresión estable de genes, cuando se requieren altos rendimientos de proteína heteróloga (Glick, 2010).

Una gran cantidad de PRs están actualmente aprobadas para uso terapéutico, la mayoría producidas en células CHO (Tabla 5), que han sido adaptadas para crecer en cultivos en suspensión de alta densidad, (Glick, 2010). A su vez, en las últimas décadas ha habido un gran aumento del uso de la línea celular HEK 293 para la producción de proteínas bioterapéuticas (Tabla 6). Estas células se cultivan fácilmente en cultivos en suspensión, sin suero, se reproducen rápidamente, son aptas para varios métodos de transfección y son muy eficientes en la producción de PRs (Dumont *et al.* 2016).

Tabla 5. Ejemplos de biofármacos producidos en Líneas Celulares de mamífero CHO, aprobados en Estados Unidos y Europa. (Tabla extraída y modificada de Dumont *et al.*, 2016)

Línea celular	Tipo de biofármaco producido	Ejemplos
CHO	Anticuerpos monoclonales	Adalimumab, Bevacizumab, Brentuximab vedotin, Denosumab, Golimumab, Ibritumomab tiuxetan, Ipilimumab, Obinutuzumab, Omalizumab, Panitumumab, Pertuzumab, Rituximab, Siltuximab, Tocilizumab, Trastuzumab, Vedolizumab, Ado-trastuzumabemtansine, Ustekinumab
	Citoquinas	Darbepoetin, alfa, Interferón beta-1a, Epoetin alfa
	Enzimas	Agalsidasa beta, Alglucosidasa alfa, Alteplasa, ADNasa Humana, Tenecteplasa
	Proteínas de fusión – FC	Abatacept, Aflibercept, Alefacept, Belatacept, Rilonacept, Ziv-aflibercept
	Hormonas	Follitropina alfa, Follitropina beta, Hormona leutinizante, Proteína 1 Osteogénica, Tirotropina alfa.

Tabla 6. Ejemplos de biofármacos producidos en Líneas Celulares Humanas HEK-293, aprobados en Estados Unidos y Europa. (Tabla extraída y modificada de Dumont *et al.*, 2016)

Línea celular	Biofármaco producido	Indicación terapéutica
HEK-293	Drotrecogin alfa	Septicemia severa/Shock séptico
	rFVIII Fc	Hemofilia A
	Dulaglutide	Diabetes tipo 2
	Human-cl rhFVII	Hemofilia A

Varias mejoras se han realizado en las líneas celulares de mamífero para potenciar su productividad, incrementando el crecimiento celular, la estabilidad del vector, la expresión génica y la secreción proteica. Las condiciones en biorreactores a gran escala pueden ser estresantes para células de mamífero. El agotamiento de nutrientes y acumulación de desechos celulares tóxicos pueden limitar la viabilidad y densidad celular, y generar que las células respondan al estrés induciendo su apoptosis. A menudo, inhibidores químicos de las vías de muerte celular son utilizados, pero se han creado líneas celulares en las cuales las vías apoptóticas son inhibidas. Cuando las células perciben estrés, una respuesta inicial es la activación del supresor tumoral proteína p53, un factor de transcripción que induce la expresión de genes que codifican proteínas en la vía apoptótica. Un método para mejorar el crecimiento celular y viabilidad, bajo condiciones de cultivo en biorreactores, es prevenir que p53 active la vía responsable de la muerte celular. La proteína de ratón doble mutante 2 (MDM2) se une a la proteína p53 y previene que actúe como factor de transcripción, y también marca p53 para la degradación. Las células HEK 293 y CHO, cuando son transfectadas con plásmidos que contienen un gen MDM2 regulable y son cultivadas bajo condiciones que imitan las etapas tardías del cultivo celular y un medio limitado en nutrientes, permiten que los cultivos que expresan MDM2 tengan altas densidades celulares y un crecimiento celular más lento. (Glick, 2010).

Muchos cultivos de células de mamífero no son capaces de permitir altas densidades celulares, porque se acumulan productos metabólicos tóxicos en el medio e inhiben el crecimiento celular. A pesar de los esfuerzos para realizar la optimización de las condiciones de cultivo, inevitablemente nutrientes esenciales para el óptimo crecimiento celular, incluyendo oxígeno, son reducidos. Bajo condiciones de bajo oxígeno, muchas células, incluyendo las células CHO, secretan productos de desecho de ácido láctico, mientras luchan para obtener

energía de la glucosa. Bajo estas condiciones, el piruvato es convertido en lactato, por la lactato deshidrogenasa, en lugar de entrar al ciclo del ácido tricarboxílico, donde luego es oxidado por la actividad de la piruvato carboxilasa, enzima que tiene un bajo nivel de actividad en ausencia de oxígeno. Para contrarrestar la acidificación del medio por la secreción de lactato, típicamente se agregan componentes alcalinos; pero estos a su vez contribuyen en la reducción del crecimiento celular debido al incremento de la osmolaridad del medio. Un enfoque más efectivo podría ser lograr al disminuir la expresión de lactato deshidrogenasa o al incrementar la expresión y/o actividad de piruvato carboxilasa en el hospedador celular (Glick, 2010).

Las líneas celulares derivadas de riñón de embrión humano y ovario de hámster chino son las preferidas para la producción de proteínas en investigación. Varias líneas celulares diferentes han mostrado ser capaces de ser transfectadas transitoriamente para la expresión de PRs, incluyendo células HEK-293, CHO, COS y BHK. La productividad de la transfección transitoria depende de múltiples factores, incluyendo la proporción de la población celular transfectada, la capacidad de las células de crecer en el tiempo en un medio dado, la eficiencia de la transcripción y traducción en el hospedero celular y el procesamiento de PRs a través del aparato post-traducciona. La expresión génica transitoria en células HEK tiene un atractivo considerable por su habilidad de producir grandes cantidades de proteína en un período razonable de tiempo, muy alta eficiencia de transfección y traducción y procesamiento confiable, en comparación con otras células de mamífero. Muchas variables, incluyendo la salud y densidad celular en la transfección, condiciones de crecimiento, recipiente de cultivo, constructo de expresión, duración de expresión y método de transfección, deben ser optimizados para permitir altos niveles de expresión transitoria, a medida que la producción proteica disminuye con el tiempo debido a la dilución del vector de expresión por la división celular. (Glick, 2010).

La línea celular HEK-293 fue creada cuando células de embrión humano fueron transformadas con fragmentos de ADN de adenovirus tipo 5, resultando en una línea celular que replica rápidamente y tolera concentraciones reducidas de suero en el medio. Las líneas celulares HEK-293T y HEK-293EBNA están integradas por un gen que codifica el virus simio 40 (SV40) y el antígeno-1 nuclear de Epstein-Barr (EBNA), respectivamente. Las mismas producen estas proteínas para permitir la replicación génica episomal y segregación cuando los elementos oriSV40 y oriP están presentes en el plásmido de expresión. Esto previene la dilución del plásmido cuando la célula se duplica y resulta en la mejora del rendimiento de producción, lo cual hace posible lograr buenos rendimientos proteicos en menos de una semana, ofreciendo una alternativa interesante al tiempo consumido en la generación de líneas celulares estables (Glick, 2010).

Las células HEK-293 y CHO son mantenidas tanto en suspensión como en cultivos adherentes, pero el cultivo en suspensión es particularmente favorable por la fácil manipulación

y escalado. Ambas líneas celulares pueden producir proteínas con modificaciones post-traduccionales eucariotas en formas funcionales a altos niveles. Típicamente, en suspensión tienen un tiempo de duplicación menor a 24 horas y pueden crecer en densidades mayores a 5 millones de células por mililitro, según las condiciones de cultivo. Se han reportado, en ambos tipos de células, producciones de gramos de proteína por litro de cultivo, generados por transfección transitoria (Hunter *et al.*, 2019; Bollin *et al.*, 2011).

El tipo de producto requerido usualmente influye en la selección de la línea celular (Figura 9); por ejemplo, se prefieren líneas CHO para anticuerpos monoclonales, mientras que las HEK293 se escogen normalmente para la producción de virus. La selección de la línea celular también depende de su disponibilidad, la existencia de un banco GMP y acceso a licencias. La aplicación y cantidad de producto requerida determinará los sistemas de cultivo (desde placas multi-well hasta biorreactores a gran escala). A su vez, la metodología de transfección depende altamente de la línea celular, la escala de trabajo y la aplicación final del producto (Gutiérrez-Granados *et al.*, 2018).

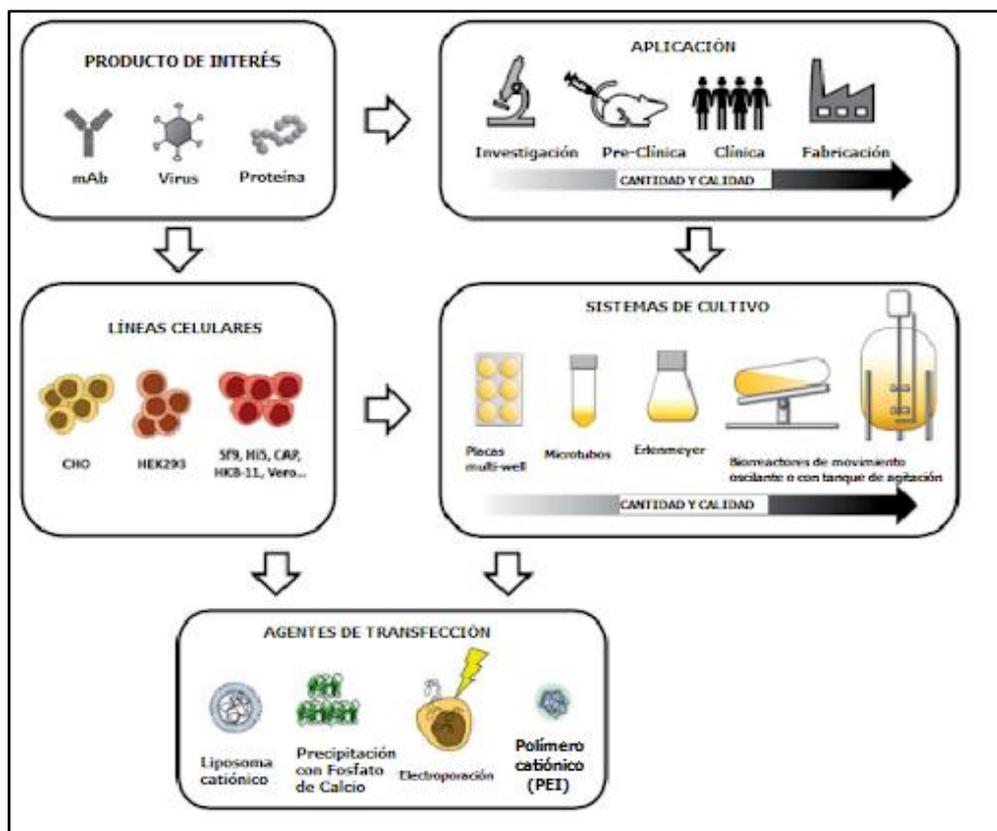


Figura 9. Elementos de transfección transitoria. Imagen extraída y modificada de Gutiérrez-Granados *et al.*, 2018.

4.2.6 Medio de Cultivo Celular para la Producción Proteica

Los medios de cultivo actuales para la producción de células de mamífero tienen ahora formulaciones que han progresado y son libres de suero o son definidos químicamente; en general, contienen entre 50 a 100 componentes agrupados como sustratos energéticos, quelantes, aminoácidos, surfactantes, buffers pH, vitaminas, derivados de ácidos nucleicos, elementos traza, sales, ácidos grasos y lípidos. El beneficio extra de los medios de cultivo libres de suero es que simplifican los procesos de purificación posteriores. Los medios libres de suero consisten en medio basal nutricionalmente completo, suplementado con una mezcla de hormonas determinada empíricamente, factores de crecimiento, factores adjuntos y proteínas de unión. Estas mezclas están disponibles comercialmente y se refiere a ellas como reemplazantes de suero. Proteínas hidrolasas derivadas de vegetales han mostrado ser capaces de reemplazar componentes del suero y derivados de animales. De todas formas, las hidrolasas por su naturaleza son complejas y no completamente definidas. (Hunter *et al.*, 2019).

El último avance en la tecnología de medios ha sido desarrollar medios de cultivo que no solo tengan componentes libres de derivados animales, sino que también estén químicamente definidos, lo cual implica desafíos significativos. Los componentes del medio de cultivo frecuentemente interactúan debido a la complejidad de las rutas metabólicas celulares. Además, los requerimientos nutricionales son definidos por el sistema en que las células son utilizadas. No se puede esperar que un medio diseñado para la transfección transitoria en células CHO sea el mejor para la selección clonal de una línea CHO estable. El medio para la expresión transitoria debe contener los nutrientes apropiados para soportar el crecimiento de cultivos de alta densidad expresando PRs, mientras que aquel que se desarrolla para las líneas celulares clonales debe proveer el crecimiento de células a extremadamente bajas densidades (Hunter *et al.*, 2019).

Uno de los parámetros más críticos a considerar cuando se desarrolla un medio de cultivo celular es el método de transfección a ser usado. Existen varias formulaciones comerciales disponibles que pueden ser utilizadas tanto en células HEK-293 como en CHO, que permiten niveles de expresión de PRs en el rango de muchos cientos de miligramos por litro. De todas formas, pueden tener varios componentes que podrían inhibir la transfección. El citrato de hierro III y el sulfato-dextrano inhiben el mediador PEI, pero son capaces de mejorar la producción proteica cuando se agregan posteriormente. Otros componentes pueden incrementar los niveles de expresión proteica ayudando en la transfección transitoria eficiente, como lo son N, N dimetilacetamida (DMA), acetato de litio (LiAc) y dimetil sulfóxido (DMSO). Una mejora adicional para la expresión proteica, tanto para líneas productivas estables como transitorias, se logra por suplementación de la fase de expresión mediante un proceso de agregado de pequeñas moléculas. En particular, el uso de inhibidores de la histona desacetilasa, como butirato de sodio

y ácido valproico, han mostrado incrementar los títulos. La mayoría de estos inhibidores modifican patrones de acetilación o metilación de histonas del huésped, incrementando así los niveles de transcripción celular y potenciando la producción proteica muchas veces. La concentración y tiempos de adición deben ser determinados empíricamente puesto que pueden resultar tóxicos para la célula y detener el crecimiento celular. Se ha visto, además, que otras moléculas pequeñas como LiAc y cafeína, tienen impactos positivos en la producción proteica. (Hunter *et.al.*, 2019).

Finalmente, la composición del medio puede también afectar las modificaciones post-traduccionales de las proteínas. De estas modificaciones, la glicosilación ha sido la más ampliamente estudiada debido a su capacidad de influenciar en las tasas de eliminación de fármacos, actividad específica, solubilidad, inmunogenicidad, y otros procesamientos de las proteínas. Varios nutrientes, precursores de azúcares, aminoácidos, metales traza, y hormonas, han mostrado influir en la distribución de glicanos. La inanición de glucosa ha mostrado que reduce la incidencia de glicosilación en residuos de asparagina, mientras que la suplementación con el azúcar galactosa en presencia de cloruro de manganeso ($MnCl_2$) puede alterar los tipos y abundancia de glicosilación. El impacto de la composición del medio cambia las modificaciones post-traduccionales, por lo que debe tenerse en cuenta durante el desarrollo, ya que puede tener un gran impacto en la actividad biológica de la proteína blanco (Hunter *et al.*, 2019).

4.2.7. Optimización de la Expresión Proteica en Células de Mamífero

Gran parte de la tecnología actual en sistemas de mamíferos se ha enfocado en procesos industriales, especialmente en la producción de biofármacos en escala de muchos kilogramos. A pesar de estos avances, seleccionar el proceso de expresión de PRs no es simple. Variables como la línea celular, el medio de cultivo y el método de cultivo impactan en el rendimiento, las características físicas y la actividad biológica de las proteínas expresadas, como se describe en la (Figura 10). La optimización de estos factores no puede realizarse de forma aislada; optimización génica, selección del vector, tipo de presión de selección, tipo de molécula a ser expresada, expresión estable o transitoria, tipo celular, selección del medio, producción a escala y modificaciones post-traduccionales, todas ellas afectan al rendimiento proteico final y a la actividad biológica (Hunter, *et al.*, 2019).

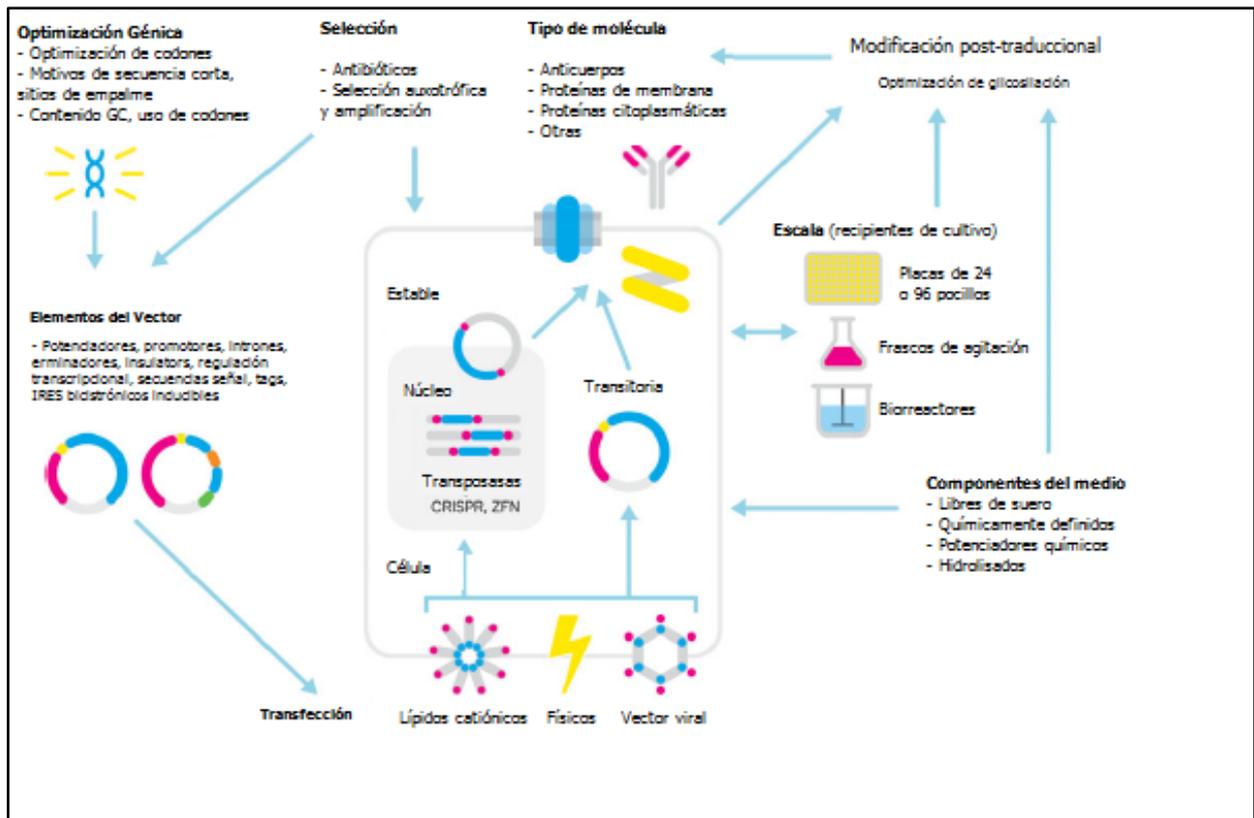


Figura 10. Resumen de los elementos que necesitan considerarse para el desarrollo de un sistema de expresión mamífero rápido y económico. Imagen modificada de Hunter *et al.*, 2019.

4.2.8. Purificación de Proteínas

En base a todas las consideraciones previas, luego de producidas las PRs deben ser purificadas. En la presente tesina, la mayor parte de las purificaciones de los extractos fueron realizadas mediante la técnica cromatográfica IMAC. Esta técnica separa las proteínas en base a una interacción fuerte y reversible entre algunos aminoácidos de la superficie de la proteína de interés, especialmente repetidos cortos de histidinas en serie que coordina metales divalentes como Ni^{2+} , Cu^{2+} , Zn^{2+} o Co^{2+} , inmovilizados en la matriz cromatográfica. También se realizaron purificaciones proteicas mediante cromatografía por afinidad con proteína A; esta proteína se une específicamente a la cadena pesada de las IgG en la región Fc, por lo que las columnas cromatográficas de proteína A unen IgG permitiendo purificar aquellas proteínas que contienen el fragmento Fc de IgG.

Otro método de purificación ampliamente utilizado y que se implementó en este trabajo, es la cromatografía gel filtración o cromatografía de exclusión molecular, una técnica de purificación muy extendida debido a su sencillez y eficiencia de resolución en mezclas complejas de macromoléculas. En este tipo de cromatografía se separan moléculas en virtud de sus diferencias de tamaño. La capacidad separadora reside en el gel, cuya matriz consta de un gran número de

esferas porosas microscópicas. Estas esferas están constituidas por largas cadenas de polímeros unidas entre sí por enlaces químicos para formar una red tridimensional.

Cuanto mayor es el tamaño de la molécula, menor es su capacidad para acceder al interior de las esferas del gel, por lo que va a ser excluida, corriendo por fuera de la matriz y, por lo tanto, se requiere menor volumen de líquido para extraerla de la columna. Cuanto más pequeñas son las moléculas, más se retrasarán eluyendo después. Teniendo en cuenta que para las proteínas globulares, en general, el tamaño está relacionado con el peso molecular, este tipo de cromatografía separa las sustancias en función de dichos pesos moleculares, obteniéndose primero las de mayor tamaño.

A diferencia de otras cromatografías líquidas, en la gel filtración las moléculas no interaccionan con la matriz de la columna y pueden utilizarse diferentes tipos de buffers sin que se altere la separación de las moléculas

4.2.9. Proteínas Recombinantes Producidas en la Presente Investigación

Durante este estudio se evaluó y comparó la expresión de dos proteínas: la proteína verde fluorescente GFP (por sus siglas en inglés, *Green Fluorescent Protein*) y el nanobody mNb6, tanto en un sistema procarionta con células de *E.coli* como en un sistema eucariota con células Expi293F.

La GFP es una proteína producida por la medusa *Aequorea victoria*, que emite bioluminiscencia en la zona verde del espectro visible (Figura 11). El gen que codifica esta proteína se utiliza habitualmente en biología molecular como marcador. Es una proteína de 27 kDa monomérica, de unos 238 aminoácidos, que forma una estructura terciaria de barril beta formado por 11 hebras beta unido a una hélice alfa central que atraviesa el barril en toda su longitud. En esta hélice hay tres aminoácidos consecutivos que forman un cromóforo natural, de forma que cuando la GFP es iluminada con luz ultravioleta, produce una brillante fluorescencia verde. Absorbe la luz azul con un máximo de longitud de onda de 395 nm, emitiendo bioluminiscencia en la zona verde del espectro con un máximo de 508 nm, si es excitada a 395 nm y, por el contrario, si es excitada a 475 nm la emisión ocurrirá a 503 nm, sin la adición de ningún cofactor o sustrato (Sacchetti *et al.*, 2000).

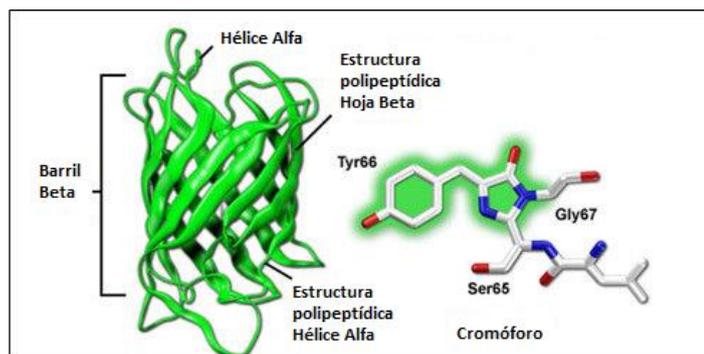


Figura 11. Estructura de la proteína verde fluorescente. Imagen adaptada de Abdul-Jalil *et al.* A. (2018).

La GFP se ha empleado ampliamente como gen reportero en diversos campos como la bioquímica, microbiología, ingeniería genética y fisiología. Desde el punto de vista biotecnológico, su utilización reside en que esta proteína se comporta como una señal luminosa capaz de expresarse en células mediante las técnicas rutinarias de transgénesis. Al llevar la fluorescencia incorporada en su estructura, la fluorescencia de la GFP puede producirse y mantenerse espontáneamente en aquellas células vivas que incluyan el gen que la codifica, sin necesidad de añadir otros agentes o cromóforos. Su secuencia puede fusionarse a otras proteínas, proporcionando a estas un dominio fluorescente extra, a modo de marca o etiqueta luminosa, que sirve para poder seguir su actividad *in vivo*, seleccionar y aislar aquellas células que producen la proteína fusionada a GFP o cuantificar la cantidad de dicha proteína producida en un momento dado (Sacchetti *et al.*, 2000).

Por otro lado, en el contexto de la pandemia COVID-19, los sostenidos esfuerzos de la comunidad científica para desarrollar vacunas, agentes terapéuticos y profilácticos, han producido innumerables estudios. Entre ellos, el uso de nanobodies ha tenido un gran impulso como agentes inmunoterapéuticos.

Los nanobodies son glicoproteínas homodiméricas derivadas de camélidos y de algunos peces cartilaginosos que presentan una porción Fc igual a la de los anticuerpos convencionales de tipo IgG, pero que carecen de cadena liviana y del dominio CH1 y que, además, su sitio de unión al antígeno está conformado únicamente por el dominio variable de la cadena pesada. La región variable de los nanobodies se denomina VHH y presenta únicamente tres regiones CDRs, a diferencia de los anticuerpos convencionales en los que el parátrope está conformado por 6. Pese a esto, se ha observado que la afinidad que exhiben los nanobodies es similar a la que presentan los anticuerpos convencionales (González-Sapienza *et al.*, 2017) (Figura 12).

Estas proteínas presentan una mayor solubilidad, como consecuencia de cambios en aminoácidos conservados involucrados en la interacción hidrofóbica VH-VL por residuos

hidrofílicos. Además, contienen residuos de cisteínas no canónicas en los CDRs lo que aumenta su estabilidad y la frecuencia de *hot spots* de hipermutabilidad. Sumado a esto, el hecho de que su CDR3 es más largo, amplía su capacidad de reconocimiento hacia cavidades profundas. Su pequeño tamaño (de aproximadamente 15 Kda), junto con su parátrope reducido con el CDR3 extendido, son características que hacen a los nanobodies capaces de acceder a epítopes que no son usualmente accesibles para los anticuerpos convencionales. El reconocimiento de epítopes cóncavos u ocultos (como son los sitios activos de las enzimas y epítopes crípticos virales), así como su alta capacidad de distinguir cambios conformacionales de una molécula, conllevan una importante gama de aplicaciones. Mantienen su funcionalidad intacta al ser sometidos a condiciones fisicoquímicas exigentes, como ser bajas temperaturas durante períodos prolongados, incubaciones a moderadas y altas temperaturas, y a la presencia de agentes químicos desnaturizantes en altas concentraciones. Esta resistencia a la desnaturalización y el plegamiento eficiente, luego de ser sometidos a las diferentes condiciones mencionadas, son producto del carácter hidrofílico aumentado que presentan y de su identidad de monodominio. Además, su estabilidad puede mejorarse mediante la introducción de residuos de cisteína en determinadas porciones que repercuten en la formación de enlaces disulfuro adicionales, presentándose, en consecuencia, resistentes a la degradación por pepsina o quimiotripsina (Gonzalez-Sapienza, 2017).

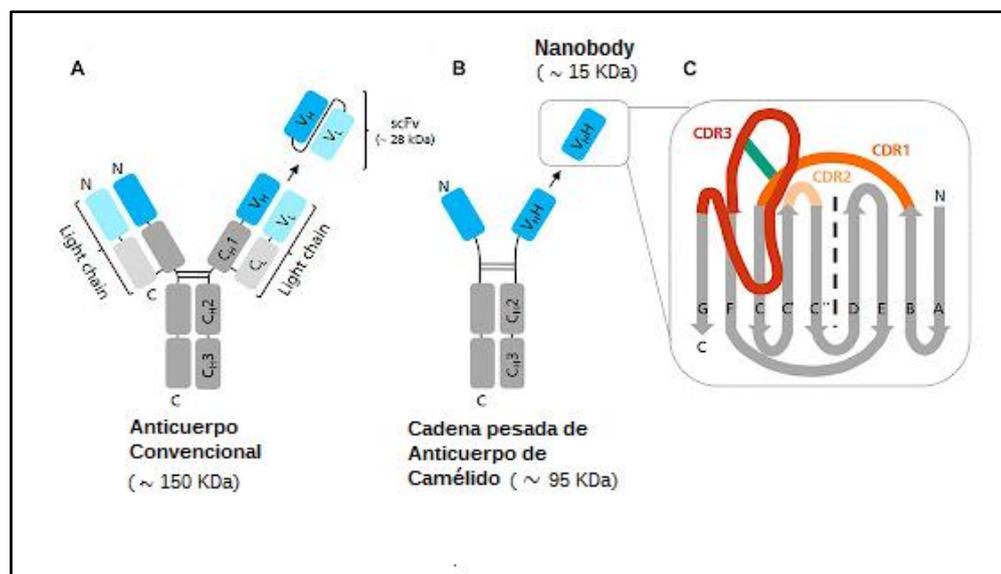


Figura 12. Representación gráfica de diferentes estructuras de anticuerpos y formatos de expresión recombinante. (a) Anticuerpo convencional IgG. (b) Anticuerpo de cadena pesada de camélido (HCAb). (c) Esquema del dominio V_HH. Imagen modificada de Haueis, *et al.*(2022).

Los nanobodies tienen un importante potencial en el área terapéutica dada su baja inmunogenicidad, la capacidad de acceder a tejidos más fácilmente debida a su pequeño tamaño, su comportamiento estable, la carencia de la porción Fc, que implica la no asociación de funciones efectoras, y una secuencia que comparte alto grado de identidad con la familia VH3 en humanos. Es posible expresar los nanobodies fácilmente y con un alto rendimiento en microorganismos, líneas celulares animales y plantas. Su producción en procariontes implica bajos costos y fácil manipulación, lo que, junto con el resto de las características antes descritas, posiciona a los nanobodies como una excelente opción para innumerables aplicaciones (Sun *et al.*, 2021, Muyldermans, 2021).

En un estudio realizado en el año 2020, (Schoof *et al.*, 2020), los investigadores sintetizaron nanobodies neutralizantes que bloquean la interacción entre Spike y ACE2. Los viriones de coronavirus están unidos por una membrana que contiene la glicoproteína homodimérica Spike, que es responsable de la entrada del virus al huésped celular. La porción de Spike expuesta a la superficie está compuesta de dos dominios, S1 y S2. El dominio S1 se une a la enzima convertidora de angiotensina 2 del receptor del huésped celular (ACE2) y S2 cataliza la fusión de las membranas viral y celular. Dentro de S1 se encuentra el dominio de unión al receptor (RBD), que se une directamente a ACE2. El RBD se une a Spike por una región flexible, que presenta un estado inaccesible hacia abajo o un estado accesible hacia arriba. La unión a ACE2 requiere que el RBD ocupe un estado activo y permita la escisión por las proteasas del huésped, lo que desencadena un cambio conformacional en S2 que permite la entrada del virus a la célula. En los viriones del SARS-CoV-2, Spike intercambia entre una conformación abierta activa con, al menos, un RBD en el estado hacia arriba y una conformación cerrada inactiva con todos los RBD en el estado hacia abajo (Schoof *et al.*, 2020).

En particular, los autores mencionados, encontraron que el nanobody (Nb6) se une a la proteína Spike en una conformación completamente inactiva, con sus dominios de unión al receptor bloqueados en un estado inaccesible hacia abajo, incapaz de unirse a ACE2. A su vez, consiguieron aumentar aún más la afinidad de ese nanobody mediante maduración de la afinidad (mNb6) y la incorporación de un diseño trivalente, generando el mNb6-tri, que consiguió una afinidad femtomolar por Spike y neutralización picomolar de la infección por SARS-CoV-2. Además, mNb6-tri conserva la función luego de la aerosolización, la liofilización y el tratamiento térmico (Schoof *et al.*, 2020).

La combinación de estabilidad, potencia y epítipo diverso de estos nanobodies anti-Spike proporcionan una estrategia profiláctica y terapéutica con gran potencial y, aunque los anticuerpos monoclonales de alta afinidad están liderando el camino como agentes terapéuticos potenciales, son costosos de producir mediante la expresión en células de mamífero y deben ser administrados por vía intravenosa por profesionales de la salud. Además, se necesitan grandes

dosis para su uso profiláctico, ya que solo una pequeña fracción de anticuerpos sistémicos atraviesa las capas de células epiteliales que recubren las vías respiratorias. Por el contrario, los nanobodies se pueden producir de forma económica en bacterias y levaduras. Su estabilidad inherente permite la administración en aerosol directamente en los epitelios nasales y pulmonares (Schoof *et al.*, 2020).

5. Objetivos

Objetivo General

El objetivo general del presente trabajo fue evaluar el efecto de determinados parámetros en los niveles de expresión de Proteínas Recombinantes en diferentes hospederos, uno procarionta y otro eucariota, con el fin de optimizar condiciones de expresión y, consecuentemente, incrementar rendimientos y disminuir costos de producción.

Objetivos Específicos

1. Comparar y evaluar los rendimientos de la proteína GFP utilizando 2 vectores con diferencias en sus regiones promotoras y regiones de inicio de la traducción, en el sistema de expresión de *E.coli*.
2. Evaluar la producción del nanobody mNb6 en células de *E.coli* y estudiar su funcionalidad.
3. Comparar y evaluar el efecto de 3 péptidos señal en la expresión de la proteína GFP en el sistema Expi293F.
4. Comparar y evaluar dos métodos de transfección diferentes en la expresión de la proteína GFP en células Expi293F.
5. Comparar y evaluar la expresión del nanobody mNb6 en células Expi293F, al utilizar 3 concentraciones distintas del vector de expresión y estudiar la funcionalidad del nanobody producido en este sistema.

6. Materiales y Métodos

6.1. Diseño de los Vectores de Expresión para el Sistema de Procarionota – *E.coli*

Para la expresión de PRs en el citoplasma de *E.coli*, se utilizaron 2 vectores: el pT7-GFP y el pT7v2-GFP. El primero, derivado del plásmido pET-32a, se utilizó como control y pertenece al conjunto de vectores construido previamente por la Unidad de Proteínas Recombinantes (UPR). El segundo, pT7v2-GFP, posee el mismo diseño que el control, pero con las secuencias del promotor y de TIR modificadas, ya que en un artículo recientemente publicado (Shilling *et.al.*, 2020) los autores encontraron fallas en el diseño de los vectores de expresión de la serie pET. Aunque esta serie funciona, esas fallas se producen en los módulos genéticos que controlan la transcripción y el inicio de la traducción, por lo que limitan los rendimientos en la expresión proteica. Por un lado, la secuencia consenso del promotor T7 convencional está truncada en el promotor T7lac y, por otro, la región de inicio de la traducción (TIR) se cree que consiguió por la adición de una fusión genética especial, y trabajos recientes indican que los TIRs nativos han co-evolucionado con los ribosomas de *E.coli* (Mortimer, 2014 y Scharff, 2011 como se citó en Shilling *et.al.*, 2020). Por esto, los investigadores, restauraron el promotor T7 conservado y los TIRs fueron sometidos a evolución sintética, lo cual permitió incrementar los rendimientos de producción proteica en la mayoría de los plásmidos de expresión pET.

Para la obtención del vector pT7v2-GFP, se le sintetizó en Genscript (www.genscript.com) utilizando el vector pT7-GFP como punto de partida. Ambos vectores poseen un gen de resistencia a la ampicilina, el operador Lac y el promotor Lacl, para permitir la inducción del gen de interés con IPTG y un sitio de escisión proteica muy específico: el TEV (sitio de reconocimiento de la proteasa del virus grabado del tabaco). A su vez, ambos contienen el gen de GFP en el sitio de inserción y las secuencias nucleotídicas que flanquean ese sitio se mantienen en ambos vectores, por lo que puede utilizarse el mismo megaprimer con el ADN blanco mediante el clonado RF.

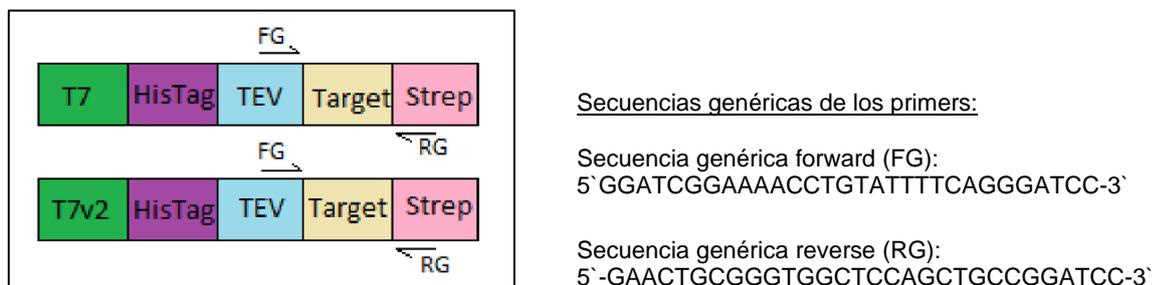


Figura 13. Representación esquemática de los vectores pT7-GFP y pT7v2-GFP para expresión de PRs en *E.coli*: T7: promotor T7; T7v2: promotor T7 modificado; TEV: Sitio de reconocimiento de la proteasa del virus grabado del tabaco; Strep: Strep-tag-II.

6.2. Diseño de los Vectores de Expresión para el Sistema de Expresión de Mamífero-EXPI293 F

En el sistema de expresión basado en células de mamífero (EXPI293F), se comparó la producción de PRs utilizando 3 péptidos señal: la cadena kappa de Inmunoglobulina murina, la Interleuquina 2 humana (IL-2) y el Interferón α -2 (IFN α 2), los cuales han demostrado ser altamente eficientes a la hora de permitir la translocación proteica y, en consecuencia, permitir la secreción al medio extracelular, incrementando significativamente la expresión de PR al medio de cultivo (Román *et al.*, 2016b).

El vector control, pCMV-kappa-GFP-Ex-Fc, forma parte del conjunto de vectores previamente construido por UPR. Está bajo control del promotor humano CMV y permite la secreción de la PR al medio de cultivo, ya que contiene un péptido señal derivado de la cadena liviana kappa de Inmunoglobulina murina. También posee una fusión C-terminal con el fragmento Fc de IgG1 humana, para permitir la purificación de proteínas secretadas mediante resinas de afinidad de proteína A. La porción Fc puede eliminarse de la proteína blanco mediante la escisión con la proteasa 3C.

Los otros dos vectores utilizados en el presente trabajo fueron pCMV-INF-GFP-Ex-Fc y pCMV-IL2-GFP-Ex-Fc, ambos iguales al vector control, pero con la secuencia del péptido señal INF- α e IL-2, en lugar de la secuencia kappa, donde la inserción de las diferentes secuencias líder se realizó en Genscript. Se utilizó un cuarto vector, idéntico al pCMV-GFP-Ex-Fc, pero con la secuencia del nanobody mNb6 en el sitio de inserción, en lugar de la de GFP, el vector pCMV-kappa-mNb6-Ex-Fc, también sintetizado por Genscript. Todos los vectores se transfectaron en la línea celular Expi293F.



Secuencias genéricas de los primers:

Secuencia genérica forward (FG):
5' GGATCGGAAAACCTGTATTTTCAGGGATCC-3'

Secuencia genérica reverse (RG):
5'-GAACTGCGGGTGGCTCCAGCTGCCGGATCC-3'

Figura 14. Representación esquemática de los vectores pCMV-GFP-Ex-Fc (pCMV-mNb6-Ex-Fc), pCMV-INF-GFP-Ex-Fc y pCMV-IL2-GFP-Ex-Fc. CMV: promotor/potenciador inmediato/temprano del citomegalovirus humano; Kappa: péptido señal derivado de la secuencia de la cadena liviana kappa de Ig murina; INF- α : péptido señal derivado de la secuencia de Interferón alfa; IL-2: péptido señal derivado de la secuencia de Interleuquina 2; TEV: Sitio de reconocimiento de la proteasa del virus grabado del tabaco; Strep: Strep-tag-II, 3C: sitio de reconocimiento de la proteasa 3C del rinovirus humano; Fc: fragmento Fc de IgG1 humana.

6.3. Protocolos

6.3.1. Equipamiento Utilizado en la Presente Investigación

- Para todas las transformaciones por electroporación se utilizó el electroporador Electro Square Porator BTX mod. ECM 830 y en las transfecciones, mediante MaxCyte, se utilizó el dispositivo Expert Stx.

- Para todas las cromatografías se utilizó el equipo AKTApure (GE *Healthcare Life Sciences*), que utiliza el software de control UNICORN.

- Las medidas de Absorbancia y densidad óptica fueron realizadas con el espectrofotómetro/fluorómetro DeNovix DS-11 FX, exceptuando las medidas de absorbancia de los ensayos ELISA donde se utilizó el lector de microplacas Thermofisher multiskan FC.

- El sonicado de todas las muestras se realizó mediante el sonicador digital Sonifier S-450.

- Todas las reacciones de PCR se realizaron utilizando el termociclador 2720 thermal cycler de Applied Biosystems.

- La transferencia en el análisis Western Blot se realizó utilizando una celda Mini Trans-Blot de Biorad

- La detección de proteínas del análisis Western Blot se realizó utilizando el equipo Amersham ImageQuant 800 UV de Cytiva.

6.3.2. Generación del Stock Celular de Plásmidos (Gliceroles)

Se transformaron los vectores plasmídicos pT7-GFP, pT7v2-GFP, *pCMV-GFP-Ex-Fc*, *pCMV-INF-GFP-Ex-Fc* y *pCMV-IL2-GFP-Ex-FC* en células DH5- α , una cepa muy versátil que permite una gran eficiencia de incorporación de plásmidos, genera insertos estables y que se utiliza para amplificar diferentes construcciones.

La transformación de las células DH5- α electrocompetentes se realizó mediante electroporación, Para esto se incorporó en hielo 1 μ L de cada vector a 20 ng/ μ L a una alícuota de 50 μ L de células DH5- α . Luego se electroporó, configurando el electroporador con los siguientes parámetros: modo LV, Voltaje 500V, longitud del pulso 8 ms y 1 pulso. Rápidamente, se agregó a cada transformación un 1 mL de medio de cultivo LB y se incubó en termomixer a 37°C, durante 1 hora. Transcurrido ese tiempo, se sembraron en placas de agar con ampicilina (cuya preparación luce en anexo 3) y se incubaron durante toda la noche, a 37°C.

Para almacenar los plásmidos transformados en las células DH5- α (*pCMV-GFP/DH5- α* , *pCMV-INF-GFP/DH5- α* , *pCMV-IL2-GFP/DH5- α* , *pT7-GFP/DH5- α* y *pT7V2-GFP/DH5- α*),

transcurrida la incubación, se tomó una colonia aislada de cada una de las 5 placas y se cultivaron en 4 mL de medio LB con 1 $\mu\text{L}/\text{mL}$ de ampicilina y luego se incubaron con agitación a 37°C, durante toda la noche. Al día siguiente, se verificó el crecimiento de los 5 cultivos (por el estado de turbidez), se tomaron 500 μL de cada uno y se les agregó 300 μL de glicerol 50 %. Por último, se congelaron y almacenaron los 5 gliceroles en freezer a -80°C .

6.3.3. Expresión de PRs en el Sistema Procariota

6.3.3.1. Comparación de los Niveles de Expresión Citoplasmática de GFP en *E.coli*

Se transformaron los vectores pT7–GFP y pT7V2–GFP en la cepa de expresión BL21(DE3) y se indujo la expresión de GFP. La transformación se realizó mediante electroporación de manera similar que en la sección anterior y se largó un precultivo de 5 mL para los cultivos de expresión. Al día siguiente se tomaron 1,5 mL de cada pre-cultivo junto con 150 mL de medio 2YT y 1 $\mu\text{L}/\text{mL}$ de ampicilina y se incubaron a 37°C, con agitación. A partir de las 2 horas de incubación, se comenzó a medir la Densidad Óptica, cuando alcanzó 0,6–0,7 aproximadamente a las 3 horas— se bajó la temperatura a 28°C durante 10 minutos y luego se realizó la inducción con IPTG 0,5 mM. Se incubó toda la noche, a 28°C, con agitación. Transcurridas 12 horas, se centrifugaron los cultivos durante 10 minutos a 4°C y 4000xg. Se descartó el sobrenadante y se resuspendió el pellet en 30 mL de agua destilada. Se centrifugó nuevamente por 10 minutos, a 4°C y 4000xg, y se pesó el pellet. Se resuspendieron ambos pellet en 20 mL de buffer A de IMAC cuya formulación se encuentra en el anexo 1 y se agregó a cada solución 0,1 mg/mL de lisozima (stock 10 mg/mL) para romper las paredes celulares y PSMF 0,1mM (stock 100 mM) —un inhibidor de proteasas, para luego almacenar en freezer a -80°C .

Sonicado, Centrifugación y Filtración de las Muestras

Las muestras congeladas a -80°C se descongelaron y fueron sometidas a sonicación para romper sus paredes y membranas celulares mediante ultrasonido, el equipo fue configurado en ciclos de 3 pulsos de 20 segundos con 1 minuto de descanso entre pulsos para que enfríe la muestra y 2 pulsos de 20 segundos con 1 minuto de descanso, y una amplitud de 50 %. Luego, las muestras sonicadas se centrifugaron a 18.000xg durante 60 minutos y el sobrenadante de cada una se filtró, con un filtro de 45 nm, para evitar el tapado de la columna con agregados proteicos o impurezas.

Purificación de GFP mediante IMAC

Para purificar la GFP mediante IMAC, en primer lugar, se equilibraron las columnas HiTrap HP (que utilizan Ni^{2+} como ligando unido a una matriz de agarosa) de 1 mL (GE *Healthcare Life Sciences*) con 10 volúmenes de buffer A, en hielo, utilizando una bomba peristáltica (la

composición de los buffers A y B utilizados se encuentra en el Anexo 1). Luego, se tomaron 20 μL de cada una de las muestras para colectar la fracción antes de pasar por la columna (AC), se cargaron las columnas con las muestras y se colectaron 20 μL de la fracción no unida (NU) a la columna, en los dos casos. Se conectó la columna al sistema AKTApure y se realizó un lavado con buffer A, hasta que la DO, a 280 nm, fue menor a 10 mU. Posteriormente, se realizó un segundo lavado (W) de la columna aumentando la concentración de imidazol a 40 mM, para eluir contaminantes que interaccionen más débilmente con la matriz. Se eluyó la proteína realizando un escalón de 100 % buffer B, del fraccionador se colectaron las fracciones del lavado (W) y el pool de la elución (E). Se midió la absorbancia a 280 nm, para determinar luego el rendimiento de GFP para las dos muestras.

Electroforesis SDS-PAGE de las Proteínas Purificadas

Previo a la corrida electroforética, para romper los puentes disulfuro de la proteína GFP y dividirla en las cadenas polipeptídicas de las que está compuesta, a 20 μL de cada una de las muestras, se les agregó 7 μL de buffer de carga de azul de bromofenol beta-mercaptoetanol (5X) y se incubaron a 95°C durante 5 minutos.

Luego, las muestras preparadas fueron sembradas junto con el marcador de peso molecular (Pierce PageRuler™ Plus Prestained Protein Ladder) en los pocillos. Este mismo marcador se utilizó en todas las corridas electroforéticas posteriores. Se conectaron los electrodos a la cuba de electroforesis y a la fuente de alimentación y se aplicó la corriente eléctrica (30 mA y 160 V, cuando la muestra atravesó el gel concentrador se subió el voltaje a 200V).

Una vez terminada la electroforesis se tiñó el gel para observar los resultados mediante tinción con Coomassie Brilliant blue R 250, un colorante aniónico que se une de manera inespecífica a proteínas.

Todas las muestras de las diferentes purificaciones proteicas, que lucen en los próximos apartados, se visualizaron de manera similar por electroforesis SDS-PAGE.

6.3.3.2. Clonado y Expresión de mNb6 en el Sistema Procariota

El clonado del nanobody mNb6 se realizó mediante la técnica de clonado RF. Para esto, se realizó una primer PCR para generar el megaprimer de mNb6, mezclando, en hielo, 1 μL ADN mNb6, 2 μL For genérico, 2 μL Rev genérico, 0,5 μL dNTPs, 10 μL buffer HF, 1 μL de ADN polimerasa Phusion (2 U/ μL) junto con 33,5 μL Agua. La reacción de PCR se desarrolló configurando en el termociclador el siguiente programa: 35 segundos a 98°C, 30 ciclos de 15 segundos a 98°C, 30 segundos a 62°C y 30 segundos a 72°C, y 4 minutos a 72°C.

Finalizada la PCR, se mezclaron los 50 μL del producto con 10 μL buffer TriTrack 6x y se sembró en un gel de agarosa 1 % (compuesto por 30 mL TAE 1x, 0,3 g agarosa, 3 μL de BrEt); sembrando en el gel de la siguiente forma: en el carril 12, 5 μL del marcador de peso molecular (Thermo Scientific™ GeneRuler 100 bp Plus DNA Ladder) y en los carriles 2, 3 y 4, 20 μL del producto de PCR a cada uno. Se aplicó la corriente, con un voltaje de 130V, y se esperó a que migren las bandas.

Una vez corrido el gel se visualizaron las bandas con luz UV y se cortaron las correspondientes al megaprimer de mNb6 y se purificó el megaprimer con QIA Quick gel extraction (QIAGEN), según las indicaciones del fabricante, con la excepción que la elución del megaprimer purificado se realizó con 30 μL de agua ultrapura y no con el buffer de elución. Luego se midió absorbancia a 260 nm y se almacenó a -20°C .

Al día siguiente, se realizó una segunda PCR para conseguir la amplificación lineal de todo el plásmido (tanto para pT7 como para pT7v2) e integración del gen de mNb6. Para esto, en hielo, y por duplicado (ya que son dos vectores), se mezclaron 110 ng de megaprimer mNb6, 25 ng de vector, 0,5 μL dNTPs, 10 μL buffer HF, 1 μL de ADN polimerasa Phusion (2 U/ μL) y agua hasta un volumen final de 50 μL .

La reacción de PCR se desarrolló configurando en el termociclador el siguiente programa: 3 minutos a 95°C , 32 ciclos de 30 segundos a 95°C , 1 minuto a 68°C y 4 minutos a 72°C , y 7 minutos a 72°C .

Finalizada la PCR se digirió el producto con la enzima de restricción DpnI para remover el plásmido parental. Para esto se le agregó 1 μL de DpnI a cada uno de los productos de la PCR y se incubó por 1 hora a 37°C . Luego se incubó durante 10 minutos a 80°C para desactivar la DpnI y se almacenó a -20°C .

Posteriormente, se realizó la transformación de pT7–mNb6 y pT7v2–mNb6 en células electrocompetentes DH5-alfa, de manera similar a la descrita anteriormente, pero, en este caso, se incorporaron 7 μL de cada vector a una alícuota 70 μL de células electrocompetentes DH5- α . Luego de la incubación se centrifugó la transformación a 8000xg durante 40 segundos, se descartaron 300 μL del sobrenadante y el resto se resuspendió y se sembró con perlas en placas de agar con ampicilina para luego incubar durante toda la noche a 37°C .

Al día siguiente, se realizó una Colony PCR, para esto, con el tip de una pipeta se tomaron 5 colonias aisladas por placa (pT7–mNb6 y pT7v2–mNb6). Se esparció cada colonia en tubos de PCR y el resto que quedó en el tip se incubó en falcons de 15 mL con 4 mL de medio LB y 1 $\mu\text{L}/\text{mL}$ de ampicilina. Estos 10 cultivos se incubaron toda la noche a 37°C , con agitación.

A los 10 tubos de PCR con colonias aisladas se les agregó la mezcla para PCR, conformada por 0,5 μL de dNTPs, 0,8 μL Rev, 0,8 μL For, 4 μL buffer HF, 0,5 μL de ADN polimerasa Phusion (2 U/ μL) y 13,4 μL agua. Además, se agregó una mezcla adicional a uno de los tubos de PCR para realizar un control positivo con el vector pT7–GFP. Posteriormente, la reacción de PCR se desarrolló configurando en el termociclador el siguiente programa: 4 minutos a 95°C, 30 ciclos de 30 segundos a 95°C, 30 segundos a 65°C y 30 segundos a 72°C, y 5 minutos a 72°C.

Finalizada la PCR se mezclaron los 20 μL del producto de cada PCR con 4 μL buffer TriTrack 6x y se sembraron en un gel de agarosa 1 % (compuesto por 45 mL TAE 1x, 0,45 g agarosa, 4,5 μL de BrEt), sembrando 12 μL de PCR pT7–mNb6 en los carriles 1 a 5, 2,5 μL de marcador de PM en el carril 6, 12 μL de PCR pT7v2–mNb6 en los carriles 7 a 10 y 12 μL de control positivo con miniprep pT7–GFP, en el carril 11.

Luego se corrieron los geles de agarosa y se obtuvo un clon positivo para T7v2–mNb6 del cual se procedió a extraer el ADN plasmídico con el kit de Thermo Scientific Gene Jet Plasmid Miniprep (#k0502), de acuerdo a las instrucciones del fabricante, con la excepción de que se eluyó el ADN con 50 μL de agua ultrapura (en lugar de buffer de elución).

6.3.3.3 Expresión del Vector pT7v2-mNb6 Clonado en Células Shuffle Electrocompetentes

Se transformó el vector pT7v2–mNb6 en la cepa de expresión Shuffle (electrocompetente), de manera similar a las secciones anteriores. Luego de la incubación, se tomaron 250 μL de la transformación junto con 20 mL de medio 2YT, 1 $\mu\text{L}/\text{mL}$ de ampicilina y 1 mL de glucosa 20% y se incubó a 37° C con agitación, durante toda la noche (pre-cultivos). Al día siguiente, se tomaron 8 mL del pre-cultivo junto con 800 mL de medio 2YT y 1 $\mu\text{L}/\text{mL}$ de ampicilina y se incubaron a 37°C, con agitación. La inducción se realizó de manera similar a la descrita anteriormente y el pellet obtenido se resuspendió en 30 mL de buffer A de IMAC y se le agregaron 0.1 mg/mL de lisozima (stock 10 mg/mL) y PMSF 0.1mM (stock 100 mM), para luego almacenar a –80°C. Luego las muestras se descongelaron, sonicaron, centrifugaron y filtraron como se describió anteriormente.

Purificación de mNb6 mediante IMAC

La purificación del nanobody se realizó mediante IMAC, con columnas HisTrap HP (GE Healthcare Life Sciences) de 1 mL, de manera similar a la descrita en la purificación de GFP.

Purificación de mNb6 mediante Cromatografía Gel Filtración

Para lograr un mayor grado de purificación del nanobody se realizó una segunda cromatografía, pero esta vez mediante la técnica de gel filtración, utilizando una columna Superdex 75 10/300 GL de HealthCare, la cual permite separar proteínas de entre 10 y 100 kDa.

Para esto, se lavó y equilibró la columna en el AKTA con agua destilada y buffer PBS (la composición del buffer PBS se encuentra en el anexo 1). Luego, se inyectaron 250 μL del pool de elución de IMAC de mNb6 diluidos en 250 μL de PBS. Posteriormente, se eluyó la proteína realizando un escalón de 100 % buffer PBS. Del fraccionador se colectó el pool de elución y el pico máximo del cromatograma.

6.3.4. Expresión de PRs en el Sistema Eucariota

6.3.4.1. Comparación de los Niveles de Expresión de GFP Utilizando Vectores con Distintos Péptidos Señal para la Secreción de GFP al Medio en Células Expi293F

Para obtener el ADN, a efectos de realizar la transfección de los distintos vectores en las células EXPI293F, a los plásmidos transformados en las células DH5- α (pCMV-GFP/DH5- α , pCMV-INF-GFP/DH5- α , pCMV-IL2-GFP/DH5- α) se les extrajo el ADN plasmídico mediante el kit para MAXIPREPs de Invitrogen (ADNp Kit pure link hipure plasmid filter DNA purification), con el cual se obtienen grandes cantidades de ADN de alta pureza.

Para esto, 2 mL de los cultivos descritos anteriormente para la obtención de gliceroles (pCMV-GFP/DH5- α , pCMV-INF-GFP/DH5- α , pCMV-IL2-GFP/DH5- α), se cultivaron junto con 200 mL de medio de cultivo 2YT y 1 $\mu\text{L}/\text{mL}$ de ampicilina. Se incubaron con agitación, a 37°C, durante toda la noche. Al día siguiente, se centrifugaron los cultivos durante 10 minutos, a 4°C y 4000xg.

Se descartó el sobrenadante y el pellet celular se resuspendió en 10 mL de buffer de resuspensión con ARNasa, luego se realizaron las Maxipreps siguiendo las indicaciones del fabricante, con la excepción de que se realizaron más pasos de lavado con Etanol 75 % para perder menos cantidad de ADN y el pellet se resuspendió en 220 μL de agua ultrapura. Luego se midió la absorbancia a 260 nm de cada una de las soluciones, en el NanoDrop, y se determinó la concentración de cada uno de los vectores purificados, los que fueron almacenados a –

En paralelo a la extracción del ADN plasmídico de estos 3 vectores, se comenzó con el cultivo de las células Expi293F, que fueron descongeladas, establecidas y transfectadas, siguiendo las instrucciones de la guía de usuario para sistemas de expresión Expi293 proporcionada por el fabricante (*ThermoFisher Scientific*).

Para la transfección, las condiciones de expresión establecidas fueron 100 mL de cultivo (en matraces de 500 mL, 5 % CO₂, 37°C, con agitación, a una concentración de 2x10⁶ células totales en medio de expresión Expi293 y aproximadamente 98 % de viabilidad el día de la transfección) para cada uno de los vectores. Siempre manipulando en cuarto de cultivo y flujo laminar.

Antes de realizar la transfección las células se centrifugaron a 200xg durante 8 minutos para eliminar posibles detritos celulares y se resuspendió el pellet en 5 mL de medio, luego se realizó

la transfección en matraces de 125 mL, introduciendo en cada uno 3 $\mu\text{g}/\text{mL}$ de ADN plasmídico junto con 6 $\mu\text{g}/\text{mL}$ de PEI Max (1 mg/mL), para cada una de las condiciones. Se incubaron los cultivos durante 3 horas y luego se transfirieron a matraces de 500 mL y se les agregó 95 mL de medio a cada uno. Transcurridas 12 horas, se agregó ácido valproico (VPA) 3.75 mM a cada matraz (un inhibidor de la actividad de la histona desacetilasa que mejora el rendimiento de PRs). A los 4 días post-transfección se contaron las células y se controló la viabilidad de cada transfección. Luego se centrifugaron los cultivos a 200xg durante 8 minutos, se almacenó el sobrenadante a -20°C , el pellet se resuspendió en 100 mL de medio y se le agregó 3 g/L de glucosa, y el cultivo se incubó por 3 días más. Transcurrido este tiempo, las células se centrifugaron durante 8 minutos a 1200xg, los 100 mL de cada sobrenadante se agregaron al almacenado a -20°C . Posteriormente, los 200 mL resultantes de cada uno de los cultivos se filtraron con un filtro de 45 nm.

Purificación de GFP Mediante Proteína A

Se utilizaron columnas HiTrap Protein A de 1 mL (GE *Healthcare Life Sciences*), las cuales unen IgG y permiten la purificación de las proteínas que contienen el fragmento Fc de IgG.

Estas columnas se equilibraron con 10 volúmenes de buffer fosfato (A), en hielo, utilizando una bomba peristáltica (la composición de los buffers A y B se encuentra en el anexo 1). Luego se cargaron las columnas con las muestras. Se conectó la columna al sistema AKTApure y se realizó un lavado con buffer A. Para colectar la muestra eluida se agregó fosfato 1 M en el fraccionador para que la misma no se encuentre en un medio demasiado ácido. Luego se eluyó la proteína con 100 % buffer citrato (buffer B). Del fraccionador se colectaron 40 μL de la fracción del pico máximo del cromatograma para su posterior análisis por SDS-PAGE y todo el pool de elución. Se midió la absorbancia del pool a 280 nm, para determinar luego el rendimiento de GFP para cada uno.

6.3.4.2. Comparación de Dos Métodos de Transfección Distintos en Células Expi293F

Se evaluó y comparó la expresión de GFP realizando en paralelo dos métodos de transfección transitoria distintos, utilizando el vector pCMV-IL2-GFP-Ex-FC en células EXPI293F. Por un lado, se transfectaron las células utilizando el método de transfección con el polímero catiónico PEI y, por el otro, con el sistema MaxCyte.

Para ambas transfecciones las condiciones de expresión fueron las mismas que las descritas anteriormente, con la excepción de que el volumen de cada uno de los cultivos fue de 70 mL. Previo a las transfecciones, los cultivos se centrifugaron a 200xg durante 8 minutos.

Para la transfección con PEI, el pellet se resuspendió en 5 mL de medio, luego se realizó la transfección en matraz de 125 mL, introduciendo 210 μg de ADN plasmídico junto con 420 μL de

PEI Max. Para la transfección por MaxCyte, como la cubeta de electroporación está diseñada para un volumen total de 400 μL , una tercera parte del pellet se resuspendió en buffer EP (buffer de electroporación, que protege a las células durante el proceso de transfección), hasta un volumen final de 400 μL (se transfectaron 3 veces más células con PEI Max que con MaxCyte debido al volumen de la cubeta y luego se realizaron las correcciones pertinentes). Posteriormente, se realizó la electroporación en el dispositivo Expert Stx, utilizando el programa para células HEK.

Ambos cultivos se incubaron durante 3 horas (con agitación para el cultivo transfectado con PEI Max y de manera estática para el cultivo transfectado con MaxCyte). Transcurrido este lapso de tiempo, el cultivo transfectado mediante PEI Max se transfirió a un matraz de 500 mL y se le agregó 65 mL de medio fresco precalentado y, al cultivo transfectado mediante MaxCyte, se le agregó 17,5 mL de medio de expresión Expi293 precalentado y suplementado con glutamina 200 mM.

Luego de pasadas 12 horas, se agregó 525 μL de ácido valproico al matraz de la transfección por PEI Max y 150,40 μL al matraz de la transfección por MaxCyte.

A los 4 días post-transfección, se contaron las células y se controló la viabilidad de cada transfección; se centrifugaron los cultivos a 200xg durante 8 minutos.

Luego, se congeló a -20°C y almacenó el sobrenadante; el pellet se resuspendió en 70 mL de medio —para la transfección con PEI— y en 20 mL —para la transfección por MaxCyte— y se le agregó a cada matraz 3 g/L de glucosa. Transcurridos 3 días, las células se centrifugaron durante 8 minutos a 1200xg y el sobrenadante se agregó al sobrenadante almacenado anteriormente; los 140 mL resultantes del cultivo transfectado con PEI Max y los 40 mL del cultivo transfectado por MaxCyte se filtraron con un filtro de 45 nm.

Purificación de GFP por IMAC y Electroforesis SDS-PAGE

Las proteínas se purificaron mediante IMAC, de manera similar a la descrita anteriormente en el ensayo de evaluación de vectores en *E.coli* para la expresión de GFP, pero utilizando, en este caso, columnas HiTrap Excel (que permiten purificar proteínas marcadas con poli-histidinas) de 1 mL (GE *Healthcare Life Sciences*). Luego, se prepararon las muestras con 40 μL de cada uno de los picos agregando 10 μL de azul de bromofenol beta-mercaptoetanol a cada una y calentando a 95°C durante 5 minutos, para luego realizar la corrida electroforética y visualización del gel de la misma manera que se describió en los apartados anteriores.

6.3.4.3. Comparación de los Niveles de Expresión de mNb6 Utilizando Distintas Concentraciones de ADN Plasmídico en Células Expi293F

Para comparar la expresión del nanobody mNb6 a 3 concentraciones distintas del vector, las condiciones de expresión establecidas fueron las mismas que las descritas anteriormente en el sistema Expi293F, y el volumen de cultivo fue de 70 mL por condición.

Para la transfección se utilizó el vector pCMV–Ex-mNb6–Fc (el mismo vector pCMV–Ex-GFP–Fc, que se utilizó como control anteriormente, pero con el gen de mNb6 insertado en lugar de GFP) y se transfectaron 210, 140 y 70 μg de ADN en cada cultivo. El procedimiento para la transfección y expresión del nanobody fue el mismo que el realizado para la expresión de GFP en la línea celular Expi293, solo se cambiaron las concentraciones de ADN y el volumen de cultivo. Por lo tanto, se disminuyó el volumen de PEI Max añadido a 420, 280 y 140 μL para las transfecciones de 210, 140 y 70 μg de ADN, respectivamente y el volumen de ácido valproico disminuyó a 525 μL por transfección.

Purificación de mNb6 Mediante IMAC y Electroforesis SDS-PAGE

Los 140 mL de sobrenadante obtenidos de cada cultivo se filtraron con un filtro de 45 nm y se procedió a purificar mNb6 mediante IMAC, de manera similar a la descrita anteriormente en el ensayo de evaluación de vectores en *E.coli* para la expresión de GFP, pero utilizando, en este caso, columnas HiTrap Excel de 1 mL (GE *Healthcare Life Sciences*). Posteriormente, se midió la absorbancia a 280 nm del pool de elución, para determinar el rendimiento de mNb6 para cada condición y se prepararon y visualizaron las muestras mediante electroforesis SDS-PAGE de manera similar a la descrita en los apartados anteriores.

Cromatografía Gel Filtración de mNb6

Al igual que para el sistema procarionta, para lograr un mayor grado de purificación del nanobody se realizó una cromatografía gel filtración, la cual se realizó de manera similar a la descrita anteriormente.

Análisis Western Blot

Para confirmar la identidad de las Proteínas Recombinantes en las últimas dos experiencias se realizó un análisis Western Blot. Para esto, luego de realizar la corrida electroforética SDS-PAGE, se transfirió el gel a un recipiente con buffer de transferencia (buffer de corrida + 20 % etanol), se armó el cassette y se corrió la transferencia en celda Mini Trans-Blot de Biorad a 400 mA en cámara fría, a 4°C, durante 1 hora.

Una vez realizada la transferencia, se detectaron las proteínas totales inmovilizadas y expuestas en la superficie de la membrana, mediante tinción con rojo de Ponceau, para luego

bloquear la membrana, incubándola durante toda la noche a 4°C con la solución de bloqueo, cuya composición luce en el Anexo 1. Al día siguiente se hicieron 5 lavados en intervalos de 5 minutos, con agitación suave con TBS y 0,05 % Tween y, posteriormente, se incubó la membrana con los anticuerpos (dilución 1:2000), para la proteína GFP se utilizaron 2,5 μL de Anti-His y para la mNb6 se utilizaron 2,5 μL de anti-IgG, ambos anticuerpos acoplados a peroxidasa de rábano (HRP), durante 2 horas a 4°C. Se lavó nuevamente la membrana con TBS 1x - 0,05 % Tween, mediante 5 lavados en intervalos de 5 minutos.

Se detectaron las proteínas a través de una reacción de oxidación de luminol que emite luz y que luego es capturada en una lámina de rayos X en el equipo Amersham ImageQuant 800 UV. La oxidación del luminol se produce por la reacción de la mezcla del kit Pierce ECL Western Blotting Substrate, comercializado por Thermo Scientific, con la HRP conjugada al anticuerpo. Para esto se llevó a cabo el siguiente protocolo: se realizó una solución de 500 μL de solución 1 y 500 μL de solución 2 y se distribuyó de manera uniforme en toda la membrana; luego la membrana se colocó en la bandeja del equipo Amersham ImageQuant 800 UV, de Cytiva, para obtener una imagen digitalizada de los resultados

6.3.5. Ensayo ELISA para Evaluar la Unión mNb6-RBD

Para determinar la funcionalidad del nanobody producido en ambos sistemas de expresión se estudió la unión del mNb6 al dominio de unión al receptor RBD, mediante el ensayo ELISA (del inglés, *enzyme linked immunosorbent assay*). La estrategia fue utilizar un RBD etiquetado con un Strep-tag-II, que es reconocido por la streptactina conjugada a HRP, por lo que al suministrarle un sustrato colorimétrico es posible detectar la señal que produce RBD cuando se une a mNb6.

La inmovilización del nanobody se realizó depositando, por duplicado, 100 μL de mNb6 0.5 μM (del pico de la cromatografía gel filtración diluido en PBS) en la microplaca para ELISA. A su vez, se realizaron dos controles. El primero, inmovilizando 100 μL de mNb6 0.5 μM , el segundo, depositando 100 μL de PBS en la microplaca; ambos por duplicado, para determinar tanto si la proteína sola o el PBS solo reaccionan con la streptactina-HRP.

Al día siguiente se realizaron 3 lavados, de 200 μL cada uno, con PBS. Luego se bloqueó la microplaca depositando 100 μL de PBS-BSA (0.5 % BSA) en cada pocillo y se incubó durante toda la noche a 4°C. Al día siguiente, se realizó un lavado con 200 μL de PBS a cada uno de los pocillos. Luego, se incubaron 100 μL de RBD-Strep (1 μM) en PBS-Tween 0.1 % y se incubó a temperatura ambiente durante 30 minutos, para todos los pocillos excepto para el control positivo. Transcurrido ese tiempo, se realizaron 3 lavados con 200 μL de PBS-Tween 0.1 %, cada uno. Para realizar la detección se agregaron 100 μL de streptactina conjugada a HRP, diluída 2000

veces en PBS-Tween 0,1 %), a cada pocillo y se incubó durante 20 minutos. Para el revelado, transcurrido el tiempo de incubación, se agregaron 100 μL del sustrato TMB (3,3', 5,5'-tetrametilbencidina) a cada pocillo y luego se quencheó la reacción con 50 μL de ácido sulfúrico (para inhibir la actividad enzimática y estabilizar el producto final de reacción que tiene color, con un viraje de amarillo a azul). Por último, se midió la absorbancia a 280 nm de cada muestra.

7. Resultados

7.1 Expresión de PRs en el Sistema Procarionta

En el sistema procarionta se compararon los niveles de expresión citoplasmática de GFP, utilizando los vectores pT7-GFP y el pT7v2-GFP, para comparar el rendimiento final de GFP conseguido tras transformar cada uno de estos vectores en células de *E.coli* de la cepa BL21(DE3). Además, en este sistema se incluyó el clonado, la expresión y purificación del nanobody mNb6, a partir de la transformación del constructo pT7v2-mNb6, en la cepa Shuffle de *E.coli*, para posteriormente evaluar su producción y funcionalidad a efectos de validar el nuevo vector pT7v2.

7.1.1. Comparación de los Niveles de Expresión Citoplasmática de GFP en *E.coli* al Utilizar los Vectores pT7-GFP y pT7V2-GFP

En primera instancia nos dispusimos a comparar el efecto que producen modificaciones en la región promotora y en la región de inicio de la traducción, de los vectores pT7-GFP y pT7v2-GFP, en los niveles de expresión citoplasmática de GFP en *E.coli*. Para ello, células electrocompetentes de *E.coli* de la cepa BL21(DE3) fueron transformadas con los vectores mencionados en cultivos de 150 mL y luego GFP fue purificada mediante IMAC.

Para visualizar e identificar la proteína expresada, se realizó una electroforesis SDS-PAGE de las purificaciones de GFP. Obteniéndose los siguientes resultados:

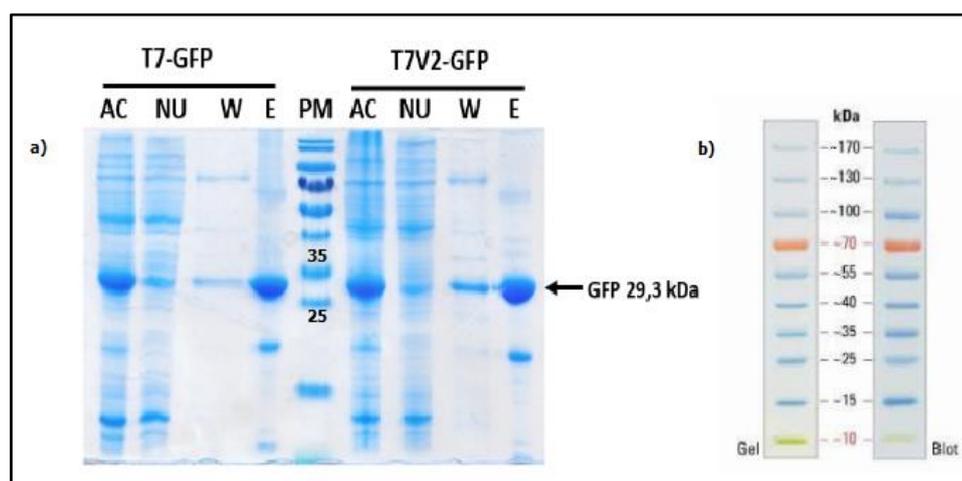


Figura 15. (a) AC: muestra antes de pasar a través de la columna; N.U: fracción de la muestra que no se unió a la columna; W: fracción de la muestra que se unió débilmente a la columna, luego del lavado; E: fracción de la muestra eluida; PM: marcador de peso molecular. (b) Referencia del marcador de peso molecular utilizado en todas las electroforesis realizadas en la presente investigación. (n=1).

Para ambas construcciones (pT7-GFP y pT7v2-GFP), observamos que:

En la fracción de la muestra, antes de pasar a través de la columna Hi-Trap (AC) que corresponde a la mezcla proteica en el citosol de la bacteria, se visualiza una banda más gruesa, alrededor de los 30 kDa, que correspondería a GFP (29.3 kDa).

En la fracción (N.U), que corresponde a la fracción de muestra que no se unió a la columna luego que la misma fue lavada con buffer A, se observa que la banda gruesa observada en el carril anterior se hace considerablemente más delgada y permanecen las otras bandas, lo que indicaría que la mayor parte de la GFP de la muestra se unió a la columna a través de la cola de histidinas.

La fracción (W), que corresponde a un segundo lavado con imidazol a 40 mM, se observa una débil banda a la altura del peso molecular de la GFP que indicaría, al igual que en el paso anterior, una leve pérdida de GFP unida a la columna.

Por último, en la fracción eluida (E), se puede observar una banda robusta de peso molecular de aproximadamente 30 kDa, que corresponde a la GFP purificada (29,3 kDa). A su vez, también se observa en este mismo carril una banda de aproximadamente 20 kDa, que podría corresponder a un contaminante de *E.coli* o a algún producto de degradación de GFP.

Cabe destacar que en ambos casos fue visible un fuerte color verde derivado de la GFP, lo cual indica que se obtuvo proteína soluble y funcional.

Por lo expuesto, de este análisis podemos concluir que:

⇒ **La proteína GFP fue purificada con éxito, resultando soluble y funcional**

Posteriormente, se midió la absorbancia a 280 nm de las fracciones con GFP purificada y se obtuvieron los siguientes resultados:

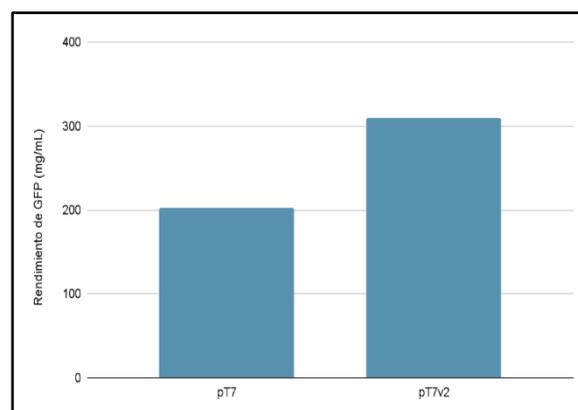


Figura 16. Comparación de los niveles de expresión citoplasmática de GFP, al transformar células *E.coli* de la cepa BL21(DE3) con los vectores pT7–GFP y pT7v2–GFP. (n=1).

⇒ La muestra **pT7** corresponde a la expresión de GFP a partir del vector control (pT7-GFP), y el rendimiento obtenido fue de **202 mg/L**.

⇒ La muestra **pT7v2** corresponde a la expresión de GFP a partir del vector con el promotor T7 y el TIR optimizados (pT7v2-GFP); el rendimiento obtenido fue de **309 mg/L**.

De estos resultados podemos concluir que:

⇒ **Se obtuvo casi un 53 % más de rendimiento de la proteína GFP cuando fue expresada en *E.coli* utilizando el vector optimizado (pT7v2)**

7.1.2. Clonado y Expresión de mNb6 a partir del Vector pT7v2-mNb6 Transformado en *E.coli*

Para validar los resultados obtenidos para GFP, generados en el ensayo anterior, nos dispusimos a clonar otra Proteína Recombinante de estructura diferente, como es el caso de un nanobody, con el fin de evaluar su expresión y funcionalidad dada por la unión a un blanco específico. Para esto, seleccionamos el nanobody mNb6, el cual se une al RBD; unión que nos indicará posteriormente la correcta funcionalidad de la proteína expresada.

Para clonar el gen de mNb6 se utilizó el método de clonado RF, con el fin de generar los constructos pT7-mNb6 y pT7v2-mNb6. Posteriormente, para determinar el correcto clonado de mNb6 e integración al plásmido correspondiente, se transformaron células DH5-alfa con cada constructo, para luego determinar qué células integraron correctamente el plásmido que porte el gen de mNb6, mediante Colony PCR.

Si bien se realizó este experimento en reiteradas ocasiones y se obtuvieron colonias para ambas transformaciones, en la colony PCR no se obtuvieron clones positivos para la transformación con pT7-mNb6, por lo no fue posible afirmar el correcto clonado de mNb6 utilizando el vector pT7. Por este motivo, se mostrarán únicamente los resultados obtenidos a partir del constructo pT7v2-mNb6, donde sí obtuvimos clones positivos.

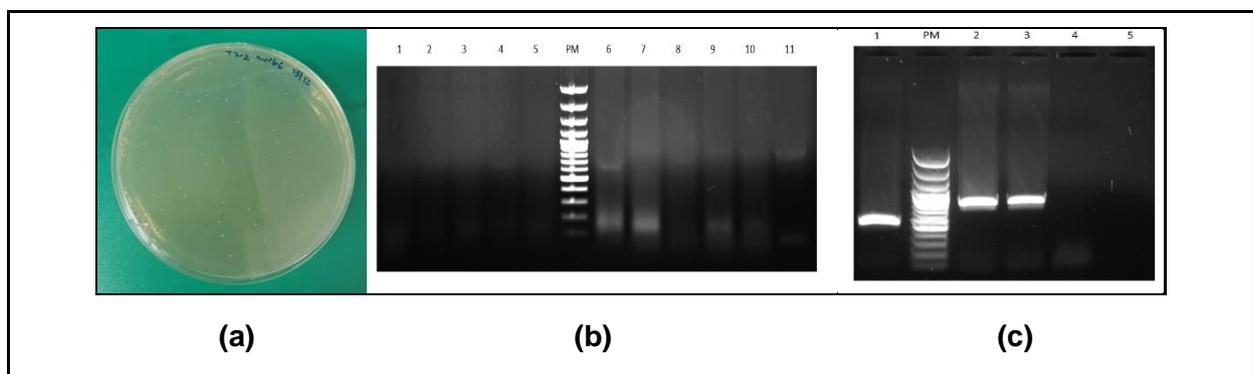


Figura 17. (a) Placa con colonias de células DH5-alfa transformadas con pT7v2-mNb6 (b) Colony PCR pT7-mNb6 y pT7v2-mNb6 en DH5-alfa. Carriles 1-5: PCR de colonias transformadas con pT7-mNb6. PM: marcador de peso molecular. Carriles 6-10: PCR de colonias transformadas con pT7v2-mNb6. Carril 11: control positivo con miniprep de pT7v2-GFP. (c) Carril 1. PCR de miniprep del clon positivo (carril 6 de la

colony PCR, 500pb). PM: marcador de peso molecular. Carril 2: control positivo con miniprep de pT7–GFP (700pb). Carril 3: control positivo con miniprep de pT7v2–GFP (700pb). Carril 4: control negativo de la PCR con la mezcla de reacción sin ADN. (n=1).

En el carril 6 de la figura 3.b), se observa un clon positivo para la transformación de pT7v2– mNb6 en células DH5–alfa, concordante con el peso molecular esperado para mNb6 de, aproximadamente, 500 pares de bases (pb). Se observa, a su vez, en los carriles 6,7, 9 y 10, bandas de unos pocos pares de bases que corresponden a primers que han hibridado formando dímeros (primer dimer). En el control positivo se observa una banda de aproximadamente 700 pb que corresponde a la secuencia de la GFP. En la figura 3.c), luego de extraer el ADN del clon positivo, se observa una banda de 500 pB, correspondiente a la secuencia de mNb6, que confirma su correcto clonado. A su vez, en los controles positivos, se observa correctamente en cada uno la banda, en aproximadamente 700 pb, que corresponden a la secuencia de GFP, y en el carril 4 no se observa banda lo cual corrobora que la mezcla de reacción no fue contaminada con ADN.

Los resultados expuestos muestran que:

⇒ La secuencia mNb6 fue correctamente expresada utilizando el vector pT7v2.

7.1.3. Expresión del Vector pT7v2–mNb6 Clonado en Células Shuffle Electrocompetentes

Para evaluar el rendimiento de mNb6, el constructo clonado fue expresado en células Shuffle, en un cultivo de 800 mL, y mNb6 se purificó mediante IMAC, y luego, para conseguir un mayor grado de pureza de la proteína y evaluar su estado oligomérico, se realizó una cromatografía gel filtración y se obtuvo el siguiente resultado:

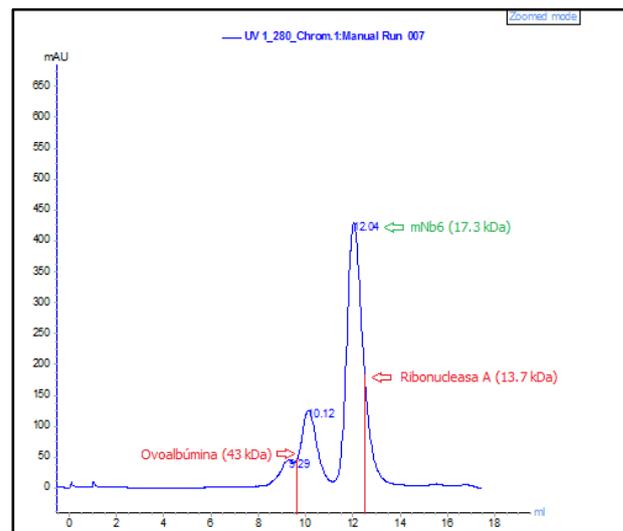


Figura 18. Cromatograma de gel filtración T7v2–mNb6. El pico máximo del cromatograma (con un volumen de elución de 12.04 mL) corresponde a la fracción eluida de mNb6. (n=1).

En tanto que las proteínas estándar Ovoalbúmina (43 kDa) y Ribonucleasa A (13.7 kDa) eluyen en dicha columna a 9.5 y 12.5 mL, respectivamente (marcado con línea roja), podemos inferir que la proteína se comporta mayoritariamente como monómero, siendo soluble y homogénea.

Posteriormente, las purificaciones por IMAC y gel filtración fueron analizadas mediante electroforesis SDS-PAGE, obteniéndose los siguientes resultados:

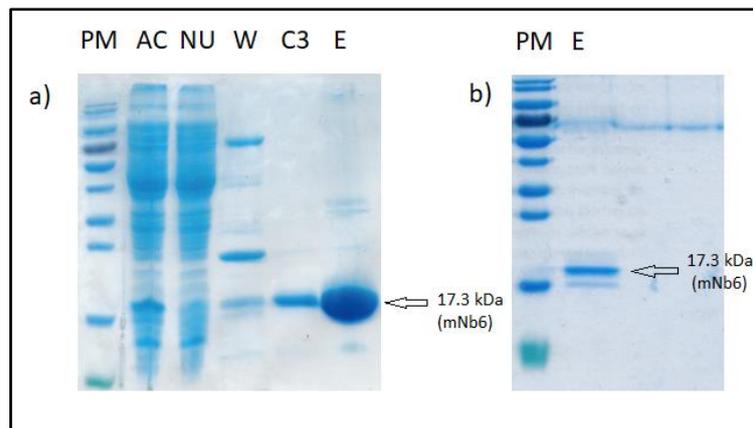


Figura 19. (a) SDS-PAGE de IMAC pT7v2-mNb6. PM: marcador de peso molecular; AC: fracción de la muestra antes de pasar por la columna Hi-Trap; N.U: fracción de la muestra que no se unió a la columna; W: fracción de la muestra que se unió débilmente a la columna, luego del lavado; C3: fracción que presentó un pequeño pico en el cromatograma luego de comenzar un gradiente con buffer B; E: fracción de la muestra eluida. (b) SDS-PAGE de gel filtración pT7v2-mNb6. PM: marcador de peso molecular; E: fracción de la muestra eluida. (n=1).

Para la SDS-PAGE de la IMAC la fracción de la muestra antes de pasar a través de la columna Hi-Trap, AC se trata de la mezcla proteica en el citosol de la bacteria y se observa una banda un poco más robusta alrededor de los 17 kDa, que correspondería a mNb6 (17.3kDa), en la fracción N.U. es la fracción de muestra que no se unió a la columna Hi-Trap, luego de que la misma fue lavada con buffer A, se observa que la banda robusta se hace considerablemente más delgada y permanecen las otras bandas, lo que indicaría que la mayor parte del mNb6 de la muestra, se unió a la columna a través de la cola de histidinas. La fracción W corresponde a un segundo lavado; al aumentar levemente la concentración de imidazol a 50 mM se observa una débil banda a la altura del peso molecular de la mNb6 que indicaría, al igual que en el paso anterior, una leve pérdida de mNb6 unido a la columna. La fracción C3 se colectó del fraccionador, ya que al aumentar el gradiente de B durante la IMAC (para eluir la proteína mediante un gradiente) se comenzó a observar un rápido aumento de la absorbancia en el cromatograma y, de acuerdo con los resultados de la electroforesis, se observa una leve pérdida de mNb6 en esta fracción. Por último, en la fracción eluida (E), se puede observar una banda robusta de peso molecular de aproximadamente 17 kDa, que corresponde al mNb6 purificado (17,3 kDa).

Por otro lado, en la SDS-PAGE, correspondiente al pico de elución del cromatograma de la cromatografía gel filtración, se observa una banda que corresponde al mNb6 purificado (17.3 kDa). En la parte superior del gel se observa una banda de 70 kDa, que atraviesa todos los carriles y no está presente en el gel A, por lo que posiblemente se trate de contaminación en el buffer de corrida electroforética.

De estos resultados, podemos concluir que:

⇒ Se purificó con éxito el nanobody mNb6, obteniéndose una proteína soluble y funcional que se comporta mayoritariamente como monómero.

Finalmente, se tomaron las medidas de absorbancia a 280 nm del nanobody purificado mediante IMAC, obteniéndose el siguiente resultado:

⇒ El rendimiento del nanobody mNb6, producido a partir de la transformación del vector pT7v2-mNb6 en la cepa Shuffle, fue de 32,5 mg/mL.

7.2. Expresión de PRs en el Sistema Eucariota

En el sistema eucariota se evaluaron y compararon 3 vectores de expresión diferentes, cada uno conteniendo un péptido señal distinto para la secreción de la proteína GFP al medio: un péptido señal derivado de la cadena liviana kappa de inmunoglobulina murina (pCMV-kappa-GFP-Fc, el cual se utilizó como control ya que es un vector utilizado rutinariamente en la Unidad de Proteínas Recombinantes-UPR); el mismo vector pCMV-kappa-GFP-Fc donde se sustituyó el péptido señal por uno derivado del Interferón alfa 2, y otro de Interleuquina 2 (pCMV-INF α -GFP-Fc y pCMV-IL2-GFP-Fc, respectivamente), ambos de origen humano. El fin ha sido comparar la expresión de GFP y determinar cuál de los vectores produce un mayor rendimiento proteico cuando la transfección se realiza en células Expi293F. En este sistema también se compararon dos métodos de transfección diferentes, utilizando el vector pCMV-IL2-GFP-Fc. Uno de esos métodos, basado en la introducción del vector mediante un polímero catiónico y el otro mediante electroporación de flujo con tecnología MaxCyte, con la finalidad de determinar qué método es más eficiente para la transfección en esta línea celular. Además, se evaluó la expresión del nanobody mNb6 en células Expi293F, utilizando el vector de expresión pCMV-kappa-mNb6-Fc a tres concentraciones distintas del ADN, con el fin de determinar la relación más efectiva en cuanto a la cantidad de ADN a utilizar y el rendimiento proteico obtenido. Finalmente, para una de estas tres condiciones, se estudió la funcionalidad del nanobody mediante el ensayo de ELISA al evaluar su unión al RBD.

7.2.1 Comparación de los Niveles de Expresión de GFP Utilizando Vectores con Distintos Péptidos Señal para la Secreción de GFP al Medio en Células de Mamífero Expi293F

Con el objetivo de comparar y evaluar el efecto de distintos péptidos señal en la expresión de la proteína GFP en el sistema Expi293F y de incrementar los niveles de expresión de esta proteína, se utilizaron tres vectores, cada uno con un péptido señal diferente para transfectar las células EXPI293F en suspensión. Dichas células se transfectaron a una concentración de 2×10^6 cel/mL y se utilizaron 150 mL de cultivo para cada vector.

Para la expresión de GFP secretada al medio, se utilizaron los vectores pCMV-kappa-Fc-GFP —que es el vector control y porta el péptido señal de la cadena liviana kappa de inmunoglobulina murina—, y pCMV-INF α -Fc-GFP —que porta el péptido señal INF α —, y pCMV-IL2-Fc-GFP —que porta el péptido señal IL2—, en cada transfección y para confirmar la identidad de la proteína obtenida, luego de la purificación por IMAC de cada muestra, se analizaron los picos obtenidos de estas cromatografías mediante electroforesis SDS-PAGE y análisis Western Blot (Figura 8), obteniéndose los siguientes resultados:

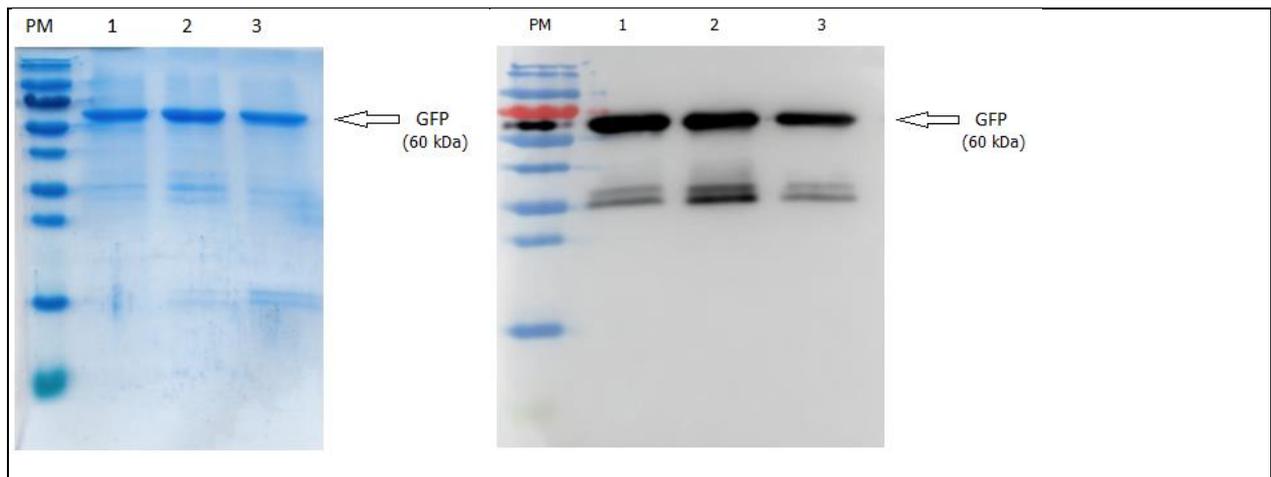


Figura 20. (a) Electroforesis SDS-PAGE para la identificación de GFP: PM: marcador de peso molecular, (1) GFP a partir del vector pCMV-kappa-GFP, (2) GFP a partir del vector pCMV-IL2-GFP, (3) GFP a partir del vector pCMV-INF α -GFP. (b) Análisis Western Blot con anticuerpo anti HisTag: PM: marcador de peso molecular, (1) GFP a partir del vector pCMV-kappa-GFP, (2) GFP a partir del vector pCMV-IL2-GFP, (3) GFP a partir del vector pCMV-INF α -GFP. (n=1).

Para todas las muestras observamos que en la electroforesis SDS-PAGE, los picos máximos del cromatograma, correspondientes a la fracción eluida de la IMAC, se comportan de la misma forma. Para todas las muestras se observa una banda de aproximadamente 60 kDa correspondiente a GFP fusionada al Fc de la IgG1, presente en el vector (peso teórico, 56.8 kDa).

Por otro lado, mediante la utilización del anticuerpo anti-His Tag en el análisis por Western Blot, se confirmó la presencia de GFP, al visualizar las mismas bandas en el gel que para la SDS-PAGE, para cada una de las muestras.

Se observa que, en ambos análisis, existen bandas de peso molecular de aproximadamente 35 kDa que pueden corresponder a productos de degradación de la proteína.

De los resultados de estos dos últimos análisis concluimos que:

⇒ La proteína GFP expresada en los tres vectores fue purificada con éxito.

Por último, se midió la absorbancia a 280 nm de cada una de las purificaciones, obteniéndose los siguientes resultados:

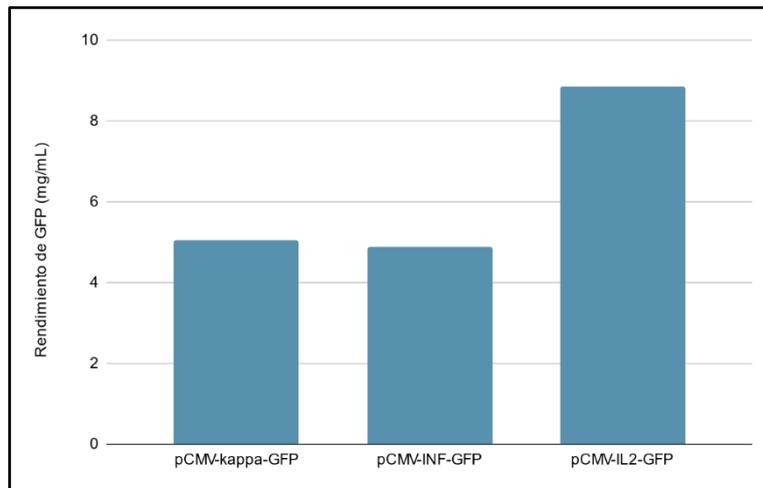


Figura 21. Comparación de los niveles de expresión de GFP secretada al medio (n=1):

- 1) Para la expresión de GFP a partir del vector control (pCMV-kappa-Fc-GFP), el rendimiento fue de **5,04 mg/L**.
- 2) Para la expresión de GFP a partir del vector pCMV-INF α -Fc-GFP, el rendimiento fue de **4,89 mg/L**.
- 3) Para la expresión de GFP a partir del vector pCMV-IL2-Fc-GFP, el rendimiento fue de **8,84 mg/L**.

Al comparar la expresión de la proteína GFP secretada al medio, tras la transfección de células EXP1293F, en tres condiciones distintas, la primera utilizando el vector control (pCMV-kappa-FC-GFP), la segunda con el vector que porta el péptido señal INF α (pCMV-INF α -FC-GFP), y la tercera con el vector que porta el péptido señal IL2 (pCMV-IL2-FC-GFP), encontramos que:

⇒ El vector pCMV-IL2-FC-GFP, que porta el péptido señal IL2, condujo a casi el doble de expresión de GFP, comparado con la expresión de esta proteína utilizando los vectores que portan los péptidos señal kappa e INF- α (pCMV-kappa-FC-GFP y pCMV-INF α -FC-GFP, respectivamente).

7.2.2. Comparación de Dos Métodos de Transfección Distintos en el Sistema Eucariota (pCMV-IL2-GFP-ext)

Visto y considerando que en el experimento anterior se obtuvo una mayor cantidad de proteína GFP utilizando el vector pCM-IL2-GFP-ext, nos dispusimos a evaluar el rendimiento de esta proteína, usando este vector, mediante dos métodos de transfección distintos: uno basado en el polímero catiónico PEI Max y el otro en el sistema de electroporación MaxCyte. Posteriormente, las purificaciones obtenidas de la IMAC fueron cuantificadas mediante la medida

de absorbancia a 280 nm y analizadas mediante electroforesis SDS-PAGE. Para un cultivo de 70 mL de cada transfección se obtuvieron los siguientes resultados:

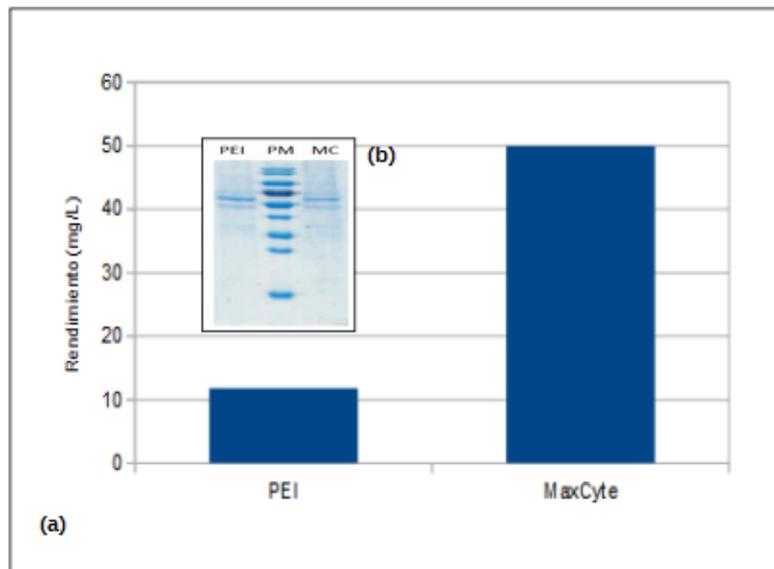


Figura 22. (a) Comparación de los niveles de expresión de GFP: Utilizando el método de transfección basado en PEI, se obtuvo un rendimiento de GFP de **11,79 mg/L** y utilizando el método de transfección MaxCyte, se obtuvo un rendimiento de GFP de **49,93 mg/L**. (b) PEI: GFP obtenida de la transfección de pCMV-IL2-GFP en células EXPI293 mediante PEI; PM: marcador de peso molecular; MC: GFP obtenida de la transfección de pCMV-IL2-GFP en células EXPI293 mediante sistema MaxCyte. (n=1).

⇒ Se obtuvo más de un 76 % de expresión de la proteína GFP utilizando el método de transfección MaxCyte.

Además, para todas las muestras observamos que en la electroforesis SDS-PAGE, las purificaciones por IMAC se comportan de la misma forma, observándose una banda de aproximadamente 60 kDa correspondiente a GFP fusionada al Fc de la IgG1, presente en el vector (Peso Teórico, 56,8 kDa).

7.2.3. Comparación de los Niveles de Expresión de mNb6 Transfectando Células EXPI293F a Distintas Concentraciones de ADN Plasmídico

Con el fin de determinar si es posible utilizar menor concentración de ADN en la transfección y consecuentemente disminuir costos, se compararon los niveles de expresión de mNb6 secretado al medio, transfectando en células EXPI293F 3 concentraciones de ADN plasmídico (210, 140 y 70 μg de pCMV-kappa-mNb6), utilizando 420, 280 y 140 μL de PEI Max (1 mg/mL), respectivamente. Para un cultivo de 70 mL de cada transfección, luego de la purificación mediante IMAC, se obtuvieron los siguientes resultados:

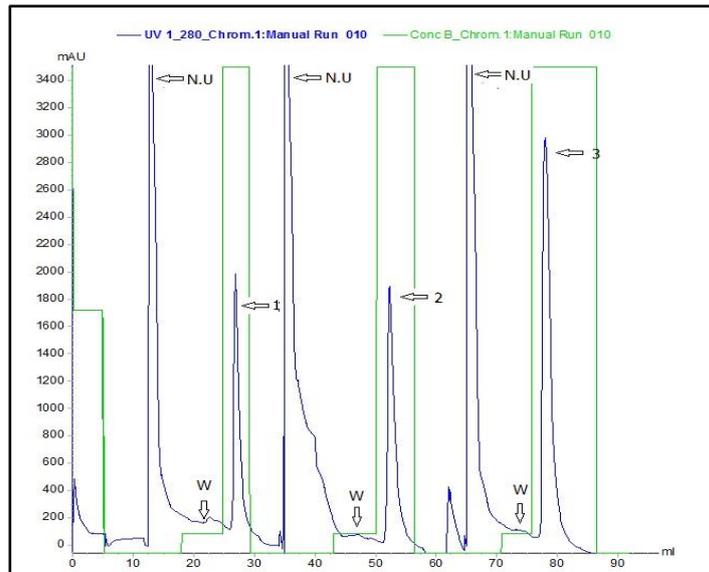


Figura 23. Cromatograma IMAC de mNb6 expresado en células de mamífero, comparando distintas concentraciones del vector pCMV-kappa-mNb6 transfectado en células EXPI293F. (n=1).

N.U: fracción no unida a la columna, W: lavado con Imidazol 40mM: **1** - elución de mNb6 (de transfección con 70 µg de ADN de mNb6 con Imidazol 500 mM). **2** - elución de mNb6 (de transfección con 140 µg de ADN de mNb6) con Imidazol 500 mM; **3** - elución de mNb6 (de transfección con 210 µg de ADN) con Imidazol 500 mM.

Para conseguir un mayor grado de pureza de la proteína y evaluar su estado oligomérico, se realizó una cromatografía gel filtración y se obtuvo el siguiente resultado:

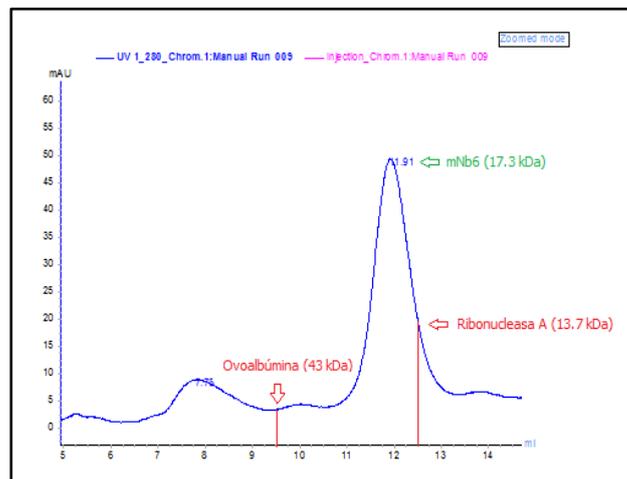


Figura 24. Cromatograma de gel filtración mNb6 (210 µg ADN) (Expi). El pico máximo del cromatograma (con un volumen de elución de 11.91 mL) corresponde a la fracción eluida de mNb6. (n=1).

En tanto que las proteínas estándar Ovalbúmina (43 kDa) y Ribonucleasa A (13.7 kDa) eluyen en dicha columna a 9.5 y 12.5 mL, respectivamente (marcado con línea roja), podemos inferir que la proteína se comporta mayoritariamente como monómero, resultando soluble y homogénea.

Las purificaciones obtenidas de la IMAC fueron analizadas mediante electroforesis SDS-PAGE y análisis Western Blot, obteniéndose los siguientes resultados:

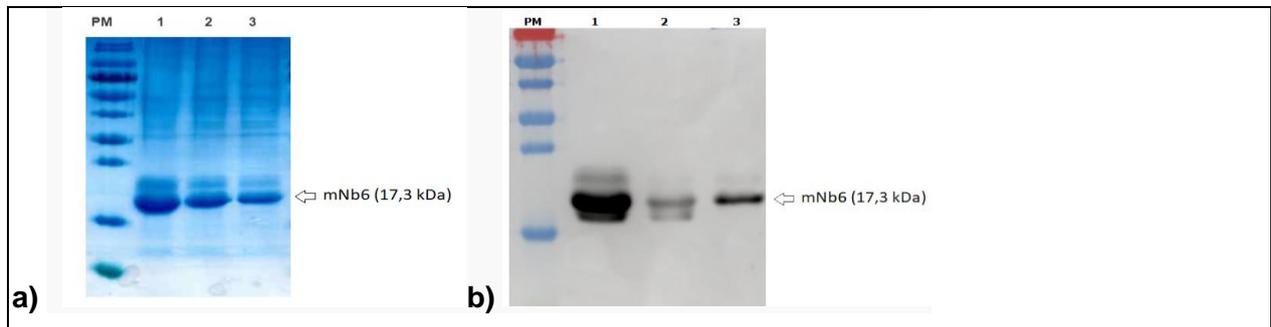


Figura 25. a) SDS-PAGE: PM marcador de peso molecular, carril 1: mNb6 de transfección con 210 µg de ADN, carril 2: mNb6 de transfección con 140 µg de ADN, carril 3: mNb6 de transfección con 70 µg de ADN. b) Western Blot con anticuerpo Anti-IgG: PM marcador de peso molecular, carril 1: mNb6 de transfección con 210 µg de ADN, carril 2: mNb6 de transfección con 140 µg de ADN, carril 3: mNb6 de transfección con 70 µg de ADN. (n=1).

De los ensayos anteriores, podemos concluir que:

⇒ Se purificó con éxito el nanobody mNb6, obteniéndose una proteína soluble y homogénea que se comporta mayoritariamente como monómero.

Luego, se midió la absorbancia a 280 nm de cada una de las muestras purificadas por IMAC, obteniéndose los siguientes resultados:

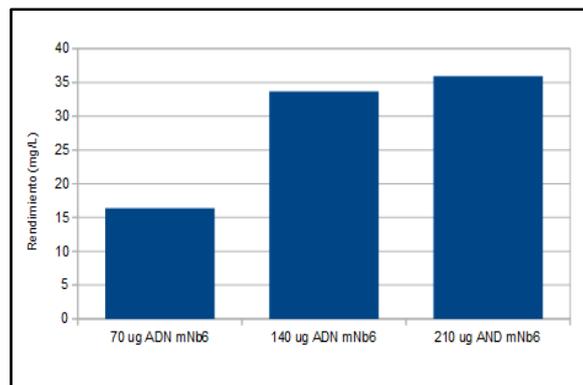


Figura 26. Comparación de los niveles de expresión de mNb6 (n=1):

- 1) Para la expresión de mNb6 a partir de 70 µg de pCMV-k-mNb6, el rendimiento fue de **16,37 mg/L**.
- 2) Para la expresión de mNb6 a partir de 140 µg de pCMV-k-mNb6, el rendimiento fue de **33,66 mg/L**.
- 3) Para la expresión de mNb6 a partir de 210 µg de pCMV-k-mNb6, el rendimiento fue de **35,92 mg/L**.

⇒ Utilizando una tercera parte menos de ADN que la que se utiliza regularmente en los protocolos de transfección con PEI Max, se observa una disminución de tan solo un 6 % en el rendimiento final de mNb6. Mientras que, utilizando dos terceras partes menos de ADN, se obtuvo casi un 46 % menos rendimiento de mNb6.

7.3 Ensayo ELISA para Evaluar la Unión mNb6-RBD en Ambos Sistemas de Expresión

Para analizar la unión del mNb6 al dominio de unión al receptor RBD, se realizó un ensayo de ELISA. En este ensayo se detectó la señal que produjo la unión de mNb6 al RBD etiquetado con un Strep-tag-II que se une a streptactina (conjugada a HRP), obteniéndose los siguientes resultados:

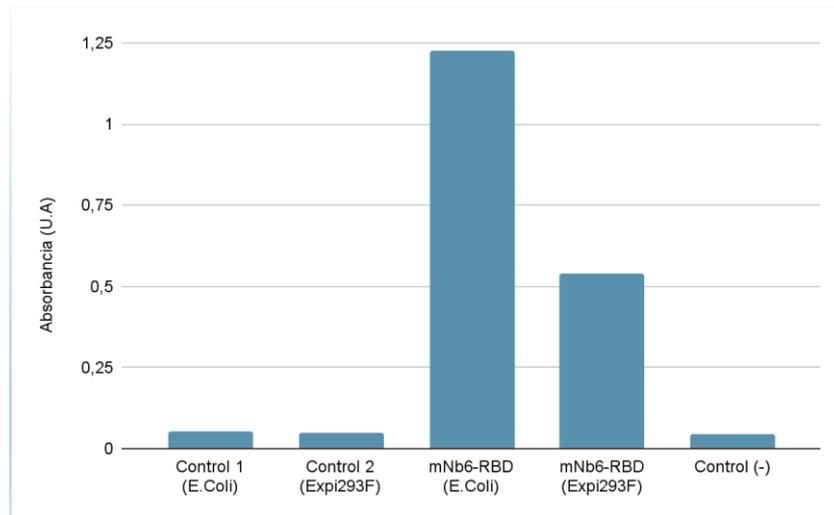


Figura 27. Resultados del ensayo Elisa mNb6-RBD. Control (1) *E.coli*: mNb6 expresado en sistema procariota y PBS. Control (2) Expi: mNb6 expresado en sistema eucariota y PBS. mNb6-RBD-*E.coli*: mNb6 expresado en sistema procariota unido a RBD-Strep. mNb6-RBD-Expi: mNb6 expresado en sistema eucariota unido a RBD-Strep. Control (-): PBS y RBD-Strep. (n=1).

En la figura 13, se observa como la streptactina-HRP no reacciona con ninguna de las proteínas expresadas (controles 1 y 2) ni con PBS (control negativo). Por otro lado, cuando RBD-Strep está presente, se observa una fuerte señal, evidenciado la interacción entre mNb6 y RBD, siendo este último el que es detectado por streptactina-HRP ya que contiene un Strep-Tag-II.

Los resultados muestran más del doble de absorbancia para el sistema de *E.coli*; esto puede obedecer a que se subestimó la concentración de T7v2-mNb6 o se sobreestimó la concentración de pCMV-k.mNb6, lo que condujo a que la concentración de pT7v2-mNb6 utilizada fue mayor que la de pCMV-k-mNb6.

Estos resultados nos permiten concluir lo siguiente:

⇒ Se observa para ambas muestras la correcta funcionalidad del nanobody, ya que hay un aumento de la absorbancia producido por la unión de mNb6 a RBD, aumento que se debe al cambio de color producido por la unión del Strep-tag-II del RBD a la streptactina conjugada a HRP.

8. Conclusiones

Obtener una proteína soluble y funcional mediante técnicas de ingeniería genética es un gran desafío dada la cantidad de parámetros que deben considerarse a la hora de planificar su expresión, en búsqueda de optimizar los resultados obtenidos. La elección de las condiciones de expresión adecuadas depende, en primer lugar, de la aplicación prevista de la proteína blanco y la optimización de estos parámetros busca no solo aumentar la cantidad o la calidad de la proteína en cuanto a su solubilidad y funcionalidad, sino también requiere evaluar la rentabilidad de los tiempos y costos de producción. Además, se requiere una constante actualización de estos conocimientos dado el creciente aumento de información y tecnologías disponibles.

En esta investigación se exploraron diferentes diseños de vectores en dos sistemas de expresión diferentes: uno procariota basado en *E.coli*, el otro eucariota basado en células de mamífero Expi293F en suspensión, derivadas de la línea celular HEK293, para evaluar la expresión recombinante de las proteínas GFP y mNb6.

Uno de los parámetros clave en la búsqueda de optimizar las condiciones de expresión de una Proteína Recombinante es el diseño del vector de expresión. En este diseño se deben tener en cuenta no solo los factores que maximicen la expresión proteica, sino también las metodologías de clonado adecuadas, así como las posibles estrategias de purificación posteriores.

En la presente investigación se seleccionó el método de clonado Restriction Free (RF) para clonar el nanobody mNb6 en el sistema de expresión procariota. La gran ventaja que ofrece este método, además de ser sencillo, rápido y económico, es que puede utilizarse en cualquier vector y que la inserción del ADN blanco puede realizarse en cualquier posición en el vector destino, independientemente de la secuencia. De esta forma, se evita que secuencias extra puedan permanecer en el gen de interés y, en el caso que varios vectores contengan la misma secuencia de inserción, el mismo megaprimer puede ser utilizado en todos ellos.

La mayoría de las purificaciones en el presente trabajo se realizaron mediante cromatografía de afinidad IMAC, dado que todos los vectores de expresión utilizados contenían un His-tag de 6 histidinas en tándem en la porción N-terminal. A su vez, el diseño de todos los vectores de expresión utilizados aquí permite evaluar dos métodos de purificación, ya que la proteína blanco contendrá un His-tag N-terminal y si no se incluye el codón de stop al final del gen blanco, un Strep-tag-II en C-terminal, por lo que pueden ser purificados mediante IMAC y/o resinas de afinidad streptactina. Este tag de purificación adicional en la proteína blanco puede ser útil en el caso de que aparezcan intermediarios de degradación, por lo que acoplado la purificación IMAC con la streptactina se obtendrá un producto purificado con los extremos N y C terminales intactos.

Además, cuando buscamos producir proteínas que se expresan en baja abundancia, la purificación por IMAC presenta inconvenientes en cuanto a que se obtienen muchos contaminantes del hospedador. Por ello, incluir el Strep-tag II en la secuencia del vector de expresión es una buena estrategia en caso de que queramos producir una proteína de baja expresión. A su vez, los vectores de expresión del sistema eucariota contenían en el C-terminal una tercera fusión, el fragmento Fc de IgG1 humana, permitiendo en uno de los ensayos la purificación de la proteína secretada mediante resinas de afinidad de proteína A.

Como último método de purificación, en dos de los ensayos del presente trabajo, realizamos una segunda cromatografía mediante la técnica de gel filtración, técnica que separa las moléculas de acuerdo a diferencias en su tamaño. Esto nos permitió lograr un mayor grado de purificación y determinar el estado oligomérico de las proteínas obtenidas.

En el desarrollo de esta tesis, para el sistema de expresión procariota se utilizaron las cepas BL21(DE3) y Shuffle, cepas ampliamente utilizadas en la expresión recombinante en *E.coli*, para comparar los niveles de expresión citoplasmática de la proteína GFP, utilizando por un lado un vector control, el pT7-GFP, de la serie pET, el cual se utiliza amplia y tradicionalmente en la expresión de PRs en *E.coli*, y un segundo vector, al que llamamos pT7v2-GFP, cuyo diseño es igual al anterior pero contiene las secuencias del promotor y la región de inicio de la traducción optimizadas. Los resultados mostraron que ambos constructos, tras la transformación en la cepa BL21(DE3), condujeron a la expresión de la proteína GFP pura, en su forma monomérica, soluble y funcional. Además, al utilizar el vector optimizado, pT7v2, se logró obtener un aumento mayor al 50 % en la producción de GFP. Posteriormente, este vector fue validado con éxito al expresar una proteína de distinta estructura, el nanobody mNb6, obteniéndose de la transformación del vector pT7v2-mNb6 en la cepa Shuffle una proteína soluble, homogénea y funcionalmente activa. Por lo expuesto, podemos decir que utilizar este vector optimizado en *E.coli* puede aumentar ampliamente el rendimiento proteico.

Para la expresión de PRs en el sistema eucariota, se utilizó la línea celular EXPI293F, derivada de las células HEK adaptadas en suspensión que son altamente transfectables. El sistema de transfección Expi293F es un sistema de expresión transitorio de alto rendimiento y es una herramienta importante para generar PRs en un corto período de tiempo, ya que las células crecerán con un tiempo de duplicación de 24-25 horas durante el crecimiento en fase logarítmica a alta densidad en medios libres de componentes de origen animal. Además, esta línea celular es susceptible a varios métodos de transfección y produce PRs en cantidades satisfactorias para análisis funcionales y estructurales, y proporciona niveles de expresión superiores en comparación con las líneas celulares estándar 293, como HEK293, en la expresión transitoria de proteínas

Un factor fundamental, que influye significativamente en el rendimiento proteico obtenido, es la adición de un péptido señal en el diseño del vector de expresión, ya que este péptido permitirá la translocación de la proteína a un ambiente más favorable para su expresión. Al comparar la expresión de la proteína GFP secretada al medio, tras la transfección de células EXPI293F utilizando vectores que portaban la secuencia de distintos péptidos señal (secuencia señal kappa, IL2 e $\text{INF}\alpha$) —de las cuales se sabe por estudios previos que son eficaces en cuanto a la translocación proteica al medio extracelular— encontramos que dentro de estos tres, el vector que porta la secuencia del péptido señal IL2 es más eficaz en la expresión de la proteína GFP, ya que condujo a la expresión de casi el doble de proteína. Por lo tanto, diseñar un vector que porte esta secuencia mejorará significativamente los rendimientos proteicos obtenidos en la expresión recombinante de una proteína de interés.

Otro punto importante a tener en cuenta a la hora de diseñar estrategias que optimicen la expresión recombinante es la metodología de transfección. En este trabajo, dado que en el sistema de expresión EXPI293F, obtuvimos una mayor cantidad de proteína GFP utilizando el vector que portaba el péptido señal IL2, (pCM-IL2-GFP-ext). Es por eso que decidimos evaluar y comparar el rendimiento de esta proteína, en este sistema, mediante dos métodos de transfección distintos. Uno de ellos es utilizado rutinariamente en estos sistemas de expresión y se basa en la introducción del ADN a las células, a través de la formación de complejos con el polímero catiónico PEI Max. El otro se basa en la introducción de ADN mediante una técnica más actual, la electroporación de flujo denominada MaxCyte. En esta experiencia se encontró que el método de transfección basado en MaxCyte fue más eficaz en la expresión de GFP en este sistema, al producir un rendimiento 4.2 veces superior que para el método de transfección tradicional mediante el polímero catiónico PEI.

Por último, un factor que jugará un rol fundamental en el diseño del esquema de expresión es el presupuesto del laboratorio, por lo que, en ocasiones, es necesario realizar un balance que permita optimizar la relación costo-efectividad. Uno de los reactivos necesarios, el ADN plasmídico, puede obtenerse mediante el uso de kits comerciales como ser Maxi-prep. Según el volumen a transfectar, los requerimientos de ADN pueden ser muy importantes, en tanto pueden aumentar costos y tiempo de trabajo.

En este sentido, se evaluó y comparó la producción del nanobody mNb6 expresado en el sistema EXPI293F, utilizando tres concentraciones distintas de ADN plasmídico (pCMV-k-mNb6-ext) en la transfección, mediante el polímero catiónico PEI, para determinar si es posible hacer uso de una menor cantidad de ADN en la transfección y, en consecuencia, disminuir los costos y tiempos de producción. En esta experiencia se obtuvo en los tres casos una proteína en su forma monomérica, soluble y funcional, y se determinó que, con una tercera parte menos de ADN plasmídico de la que se utiliza regularmente en los protocolos de transfección con PEI, se pueden

obtener rendimientos muy similares. Esto resulta útil para cuando la cantidad de ADN es un factor limitante, dado que se pueden lograr resultados muy parecidos invirtiendo menor cantidad de ADN y, con lo cual se optimiza la relación costo-beneficio.

En esta investigación se evaluaron algunos parámetros que pueden influir en la expresión recombinante y se reflejan la compleja matriz de factores que deben evaluarse a la hora de planificar un esquema de producción de Proteínas Recombinantes efectivo. Adicionalmente, los resultados obtenidos colaboran en esta búsqueda para la optimización al explorar la efectividad de distintas metodologías, la importancia del diseño de vectores de expresión en los rendimientos finales y la rentabilidad en costos de producción.

9. Referencias Bibliográficas

- Abdul-Jalil, A. A., Khalaf, S. N., and Dawod, S. S. (2018). Modification of the green fluorescent protein properties using the site direct mutagenesis effects by replacing Y66W. *Biochemical and Cellular Archives*. 18. 1671-1678.
https://www.connectjournals.com/file_html_pdf/2881602H_1671A.pdf
- Agnov, E., Hillier, C. J., Kincaid, R. L., Lyon, J. A. (2008). Heterologous protein expression in enhanced by harmonizing the codon usage frequencies of the target gene with those of the expression host. *PLoS One*. 14; 3(5): 2189. <https://pubmed.ncbi.nlm.nih.gov/18478103/>
- Artola, J., Villamonte, G., Gonzales, H. (2015). DNA transfection of animal cells as a tool used in animal biotechnology. *The Biologist*. 12. 126.
https://www.researchgate.net/publication/312297399_DNA_TRANSFECTION_OF_
- Bollin, F., Dechavanne, V, Chevalet, L. (2011). Design of Experiment in CHO and HEK transient transfection condition optimization. *Protein Expr Purif*. Jul;78(1):61-8.
<https://pubmed.ncbi.nlm.nih.gov/21354312/>
- Chevet, E., Cameron, P. H., Pelletier, M., Thomas, D., Bergeron, J. (2001). The endoplasmic reticulum: integration of protein folding, quality control, signaling and degradation, *Current Opinion in Structural Biology*. Volume 11, Issue 1, , Pages 120-124, ISSN 0959-440X,
[https://doi.org/10.1016/S0959-440X\(00\)00168-8](https://doi.org/10.1016/S0959-440X(00)00168-8).
- Correa, A., Oppezzo, P. (2011). Tuning different expression parameters to achieve soluble recombinant proteins in E. coli: advantages of high-throughput screening. *Biotechnol J*. Jun;6(6):715-30. doi: 10.1002/biot.201100025. <https://pubmed.ncbi.nlm.nih.gov/21567962/>
- Correa, A., Ortega, C., Obal, G., Alzari, P., Vincentelli, R., Oppezzo, P. (2014). Generation of a vector suite for protein solubility screening. *Front Microbiol*. Feb 25;5:67. doi: 10.3389/fmicb.2014.00067. <https://pubmed.ncbi.nlm.nih.gov/24616717/>
- Correa, A., Oppezzo, P. (2015). Overcoming the solubility problem in E. coli: available approaches for recombinant protein production. *Methods Mol Biol*. 1258:27-44.
<https://pubmed.ncbi.nlm.nih.gov/25447857/>
- Demain, A. L., Vaishnav, P. (2009). Production of recombinant proteins by microbes and higher organisms. *Biotechnol Adv*. May-Jun;27(3):297-306. doi: 10.1016/j.biotechadv.2009.01.008. Epub 2009 Jan 31. PMID: 19500547.
- Díaz-Ariza, I., Sierra, C. A., Pérez-Pérez, L. D. (2018). Desarrollo de vectores génicos basados en polímeros sintéticos: pei y pdmaema. *Rev. Colomb. Cienc. Quím. Farm.*, Vol. 47(3), 350-374.
- Dumont, J., Ewart, D., Mei, B., Estes, S., Kshirsagar, R. (2016). Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives. *Critical Reviews in*

- Biotechnology*, 36:6, 1110-1122, DOI: 10.3109/07388551.2015.1084266.
<https://pubmed.ncbi.nlm.nih.gov/26383226/>
- Galloway, C. A., Sowden, M. P., Smith, H. C. (2003). Increasing the yield of soluble recombinant protein expressed in *E. coli* by induction during late log phase. *Biotechniques*. Mar;34(3):524-6, 528, 530. doi: 10.2144/03343st04. <https://pubmed.ncbi.nlm.nih.gov/12661158/>
- Glick, B. R., Pasternak, J., Patten, C. L. *Molecular Biotechnology, Principles and Applications of Recombination DNA*, 4th Ed. <https://www.wiley.com/en-gb/Molecular+Biotechnology:+Principles+and+Applications+of+Recombinant+DNA,+4th+Edition-p-9781555814984>
- González-Sapienza, G., Rossotti, M. A., Tabares-da Rosa, S. (2017). Single-Domain Antibodies As Versatile Affinity Reagents for Analytical and Diagnostic Applications. *Front Immunol*. Aug 21;8:977. doi: 10.3389/fimmu.2017.00977. PMID: 28871254. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5566570/>
- Gutiérrez-Granados, S., Cervera, L., Kamen, A., Gòdia, F. (2018). Advancements in mammalian cell transient gene expression (TGE) technology for accelerated production of biologics. *Critical Reviews in Biotechnology*, 38:6, 918-940, DOI: 10.1080/07388551.2017.1419459
- Hauéis, L., Stech, M., & Kubick, S. (2022). A Cell-free Expression Pipeline for the Generation and Functional Characterization of Nanobodies. *Frontiers in bioengineering and biotechnology*. (10). <https://doi.org/10.3389/fbioe.2022.896763>
- Hunter, M., Yuan, P., Vavilala, D., Fox, M. (2019). Optimization of Protein Expression in Mammalian Cells. *Curr Protoc Protein Sci*. Feb;95(1):e77. doi: 10.1002/cpps.77. Epub 2018 Sep 28. PMID: 30265450.
- Jia, B., Jeon, C. O. (2016). High-throughput recombinant protein expression in *Escherichia coli*: current status and future perspectives. *Open Biol*. Aug;6(8):160196. doi: 10.1098/rsob.160196. PMID: 27581654; PMCID: PMC5008019.
- Khan, A. H., Bayat, H., Rajabibazl, M., Sabri, S., Rahimpour, A. (2017). Humanizing glycosylation pathways in eukaryotic expression systems. *World J Microbiol Biotechnol*. Jan;33(1):4. doi: 10.1007/s11274-016-2172-7. Epub 2016 Nov 11. PMID: 27837408.
- Kim, T, K., Eberwine, J. H. (2010). Mammalian cell transfection: the present and the future. *Anal Bioanal Chem*. Aug;397(8):3173-8. doi: 10.1007/s00216-010-3821-6. Epub 2010 Jun 13. PMID: 20549496; PMCID: PMC2911531.
- Li, F., Vijayasankaran, N., Shen, A. Y., Kiss, R., Amanullah, A. (2010). Cell culture processes for monoclonal antibody production. *MAbs*. Sep-Oct;2(5):466-79. doi: 10.4161/mabs.2.5.12720. Epub 2010 Sep 1. PMID: 20622510; PMCID: PMC2958569.
- Makrides, S. C. (1996). Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol Rev*. Sep;60(3):512-38. doi: 10.1128/mr.60.3.512-538.1996. PMID: 8840785; PMCID: PMC239455.

- Makrides, S. C. (2003). Vectors for gene expression in mammalian cells. 2003. *New Comprehensive Biochemistry*. 38, 9–26. [https://doi.org/10.1016/S0167-7306\(03\)38002-0](https://doi.org/10.1016/S0167-7306(03)38002-0)
- Mesa-Pereira, B., Rea, M. C., Cotter, P. D., Hill, C., Ross R. P. (2018). Heterologous Expression of Biopreservative Bacteriocins With a View to Low Cost Production; *Frontiers in Microbiology*; (9). <https://www.frontiersin.org/article/10.3389/fmicb.2018.01654>; DOI=10.3389/fmicb.2018.01654; ISSN=1664-302X
- Muyldermans, S. (2021). Applications of Nanobodies. *Annual review of animal biosciences*, 9, 401–421. <https://doi.org/10.1146/annurev-animal-021419-083831>
- Nettleship, J. E. (2012). Structural Biology of Glycoproteins. In (Ed.), *Glycosylation*. IntechOpen. <https://doi.org/10.5772/48154>
- Ortega, C., Prieto, D., Abreu, C., Oppezzo, P., Correa, A. (2018). Multi-Compartment and Multi-Host Vector Suite for Recombinant Protein Expression and Purification. *Front Microbiol*. Jun 27;(9):1384. doi: 10.3389/fmicb.2018.01384. PMID: 29997597; PMCID: PMC6030378.
- Ortega, C., Abreu, C., Oppezzo, P., Correa, A. (2019). Overview of High-Throughput Cloning Methods for the Post-genomic Era. *Methods Mol Biol*. 2019;2025:3-32. doi: 10.1007/978-1-4939-9624-7_1. PMID: 31267446.
- Owji, H., Nezafat, N., Negahdaripour, M., Hajiebrahimi, A., Ghasemi, Y. (2018). A comprehensive review of signal peptides: Structure, roles, and applications. *Eur J Cell Biol*. Aug;97(6):422-441. doi: 10.1016/j.ejcb.2018.06.003. Epub 2018 Jun 22. PMID: 29958716.
- Portolano, N., Watson, P. J., Fairall, L., Millard, C. J., Milano, C. P., Song, Y., Cowley, S. M. Schwabe, J. W. (2014). Recombinant protein expression for structural biology in HEK 293F suspension cells: a novel and accessible approach. *J Vis Exp*. Oct 16;(92):e51897. doi: 10.3791/51897. PMID: 25349981; PMCID: PMC4420617.
- Román, R., Miret, J., Martínez-Monge, I., Roura, A., Casablanca, A., Lecina, M., Cairó, J. (2016a). Study on vector elements for efficient selection, expression and secretion in HEK293 cells. *New Biotechnology*. 33. S189. 10.1016/j.nbt.2016.06.1374.
- Román, R., Miret, J., Scalia, F., Casablanca, A., Lecina, M., Cairó, J. J. (2016b). Enhancing heterologous protein expression and secretion in HEK293 cells by means of combination of CMV promoter and IFN α 2 signal peptide. *Journal of Biotechnology*, (239), 57-60, ISSN 0168-1656. <https://doi.org/10.1016/j.jbiotec.2016.10.005>
- Sacchetti, A., Ciccocioppo, R., & Alberti, S. (2000). The molecular determinants of the efficiency of green fluorescent protein mutants. *Histology and histopathology*. 15(1), 101–107. <https://doi.org/10.14670/HH-15.101>
- Sandoval Rodríguez, A., Mena Enríquez, M., Márquez Aguirre, A. (2013). Vectores de clonación y expresión. En: Salazar Montes A, Sandoval Rodríguez A, Armendáriz Borunda J. (Eds.). *Biología molecular. Fundamentos y aplicaciones en las ciencias de la salud*. Mc Graw Hill.

- Shilling, P. J., Mirzadeh, K., Cumming, A. J., Widesheim, M., Köck, Z., Daley, D. O. (2020). Improved designs for pET expression plasmids increase protein production yield in *Escherichia coli*. *Commun Biol.* May 7;3(1), 127-138: 214. doi: 10.1038/s42003-020-0939-8. PMID: 32382055; PMCID: PMC7205610.
- Schoof, M., Faust, B., Saunders, R.A., Sangwan. S., Rezelj, V., Hoppe, N., Boone, M., Billesbølle, C.B., Puchades, C., Azumaya, C.M., Kratochvil, H.T., Zimanyi, M., Deshpande, I., Liang, J., Dickinson, S., Nguyen, H.C., Chio, C.M., Merz, G.E., Thompson, ... Manglik, A. (2020). An ultrapotent synthetic nanobody neutralizes SARS-CoV-2 by stabilizing inactive Spike. *Journal List.* Dec 18;370(6523):1473-1479. doi: 10.1126/science.abe3255. Epub 2020 Nov 5. PMID: 32817938; PMCID: PMC7430568.
- Sørensen, H. P., Mortensen, K. K. (2005a). Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *J Biotechnol.* Jan 26;115(2):113-28. doi: 10.1016/j.jbiotec.2004.08.004. PMID: 15607230.
- Sørensen, H. P., Mortensen, K. K. (2005b). Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microb Cell Fact.* Jan 4;4(1):1. doi: 10.1186/1475-2859-4-1. PMID: 15629064; PMCID: PMC544838.
- Sun, S., Ding, Z., Yang, X., Zhao, X., Zhao, M., Gao, L., Chen, Q., Xie, S., Liu, A., Yin, S., Xu, Z., Lu, X. (2021). Nanobody: A Small Antibody with Big Implications for Tumor Therapeutic Strategy. *International journal of nanomedicine.* 16, 2337–2356. <https://doi.org/10.2147/IJN.S297631>
- Terpe, K. (2006). Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol.* Sep;72(2):211-22. doi: 10.1007/s00253-006-0465-8. Epub 2006 Jun 22. PMID: 16791589.
- Yang, S., Zhou, X., Li, R., Fu, X., & Sun, P. (2017). Optimized PEI-based Transfection Method for Transient Transfection and Lentiviral Production. *Current protocols in chemical biology.* 9(3), 147–157. <https://doi.org/10.1002/cpch.25>
- Zemella, A., Thoring, L., Hoffmeister, C., Kubick, S. (2015). Cell-Free Protein Synthesis: Pros and Cons of Prokaryotic and Eukaryotic Systems. *ChemBioChem.* 16 , 2420.
- Zhang, L., Leng, Q., Mixson, A. J. (2005). Alteration in the IL-2 signal peptide affects secretion of proteins in vitro and in vivo. *J Gene Med.* Mar;7(3):354-65. doi: 10.1002/jgm.677. PMID: 15619290.

10. Anexos

Anexo 1. Composición de las Soluciones Utilizadas

Buffer A (buffer de lavado IMAC con columnas HiTrap HP y Excel):

- Tris 50 mM pH 8
- NaCl 300 mM
- Imidazol 20 mM

Buffer B (buffer de elución IMAC con columnas HiTrap HP y Excel):

- Tris 50 mM pH 8
- NaCl 300 mM
- Imidazol 0,5 M

Buffer Fosfato pH 7, 50 mM (buffer A de lavado con columna HiTrap Protein A)

- Na_2HPO_4
- NaH_2PO_4

Buffer Citrato pH 3, 0,1 M (buffer B de elución con columna HiTrap Protein A)

- $\text{C}_6\text{H}_8\text{O}_7 \cdot \text{H}_2\text{O}$ (ácido cítrico monohidrato)
- $\text{C}_6\text{H}_5\text{O}_7\text{Na}_3 \cdot 2\text{H}_2\text{O}$ (citrato trisódico dihidrato)

Buffer fosfato salino (PBS)

- pH 7,4
- NaCl
- KCl
- Na_2HPO_4
- KH_2PO_4
- Buffer salino Tris (TBS)
- $(\text{HOCH}_2)_3\text{CNH}_2$ (Tris)
- NaCl

Solución de bloqueo (Western Blot):

- TBS 1x
- 0,05 % Tween

- 5 % de leche en polvo

Solución de lavado (Western Blot):

- TBS 1x
- 0,05 % Tween

Buffer de corrida SDS-PAGE

- Tris 25 mM
- Glicina 192 mM
- SDS 0,1 %
- pH 8,3-8,6

Anexo 2. Preparación de Geles de Poliacrilamida para SDS-PAGE

En este trabajo utilizamos geles de acrilamida de 15 %, realizados según el siguiente protocolo (para conseguir 2 geles):

Para obtener 5 mL de mezcla en el gel separador se agregaron en orden los siguientes componentes:

- 0,4 mL de agua destilada
- 2,5 mL de acrilamida 30 %
- 2 mL de buffer Tris pH 8.8
- 50 μ L de SDS 10 %
- 50 μ L de persulfato de amonio (PSA) 10 %
- 2 μ L de TEMED

Una vez armado el soporte para los geles, se cargó cada uno de ellos con la mezcla anterior y se dejó polimerizar por 15 minutos, aproximadamente.

Luego que el gel separador polimerizó, se prepararon 3 mL del gel concentrador agregando en orden los siguientes componentes:

- 2.1 mL de agua destilada
- 0,5 mL de acrilamida 30 %
- 0,38 mL de buffer Tris pH 6,8
- 30 μ L de SDS 10 %
- 30 μ L de PSA 10 %
- 3 μ L de TEMED

Luego se cargó cada gel con la solución anterior e inmediatamente se agregaron los peines para formar los pocillos del gel y se dejó polimerizar por 15 minutos.

Anexo 3. Preparación de Placas LB-Agar-Ampicilina

Para realizar las placas de LB-Agar y ampicilina se realizó el siguiente procedimiento:

Se fundieron 300 mL de medio LB-Agar en microondas y se dejó enfriar. Transcurridos unos 20 minutos, se le agregaron 1 $\mu\text{L}/\text{mL}$ de ampicilina. Luego, se repartió la mezcla en placas de Petri plásticas, generando capas de LB-agar-ampicilina de aproximadamente 0,7 cms. Por último, se permitió que solidifiquen a temperatura ambiente, evitando que se genere condensación y se almacenaron en cámara de frío, a 4°C.