

**UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
FACULTAD DE CIENCIAS QUÍMICAS**



**EXPRESIÓN Y PURIFICACIÓN DE PROTEÍNAS RECOMBINANTES EN
CORYNEBACTERIUM GLUTAMICUM Y *ESCHERICHIA COLI*
UTILIZANDO LA PROTEÍNA SMBP Y UN MUTANTE DE CUSF
COMO PROTEÍNAS DE FUSIÓN**

POR

TERESA VARGAS CORTEZ

**COMO REQUISITO PARCIAL PARA OBTENER EL GRADO DE DOCTOR
EN CIENCIAS CON ORIENTACIÓN EN MICROBIOLOGÍA APLICADA**

DICIEMBRE, 2017

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
FACULTAD DE CIENCIAS QUÍMICAS



EXPRESIÓN Y PURIFICACIÓN DE PROTEÍNAS RECOMBINANTES EN
CORYNEBACTERIUM GLUTAMICUM Y *ESCHERICHIA COLI* UTILIZANDO LA
PROTEÍNA SMBP Y UN MUTANTE DE CUSF COMO PROTEÍNAS DE FUSIÓN

Por

TERESA VARGAS CORTEZ

Como requisito parcial para obtener el grado de **DOCTOR EN CIENCIAS** con
Orientación en Microbiología Aplicada

Diciembre, 2017

**EXPRESIÓN Y PURIFICACIÓN DE PROTEÍNAS RECOMBINANTES EN
CORYNEBACTERIUM GLUTAMICUM Y *ESCHERICHIA COLI* UTILIZANDO LA PROTEÍNA
SMBP Y UN MUTANTE DE CUSF COMO PROTEÍNAS DE FUSIÓN**

Aprobación de la tesis:

Dr. Xristo Zárate Kalfópulos

ASESOR

Dra. María Teresa Garza González

COMITÉ TUTORIAL

Dr. Jesús Alberto Gómez Treviño

COMITÉ TUTORIAL

Dr. José Rubén Morones Ramírez

COMITÉ TUTORIAL

Agradecimientos

Agradezco enormemente a todas las personas e instituciones que me han permitido el desarrollo de este importante proyecto.

A CONACyT por la beca de manutención nacional, beca mixta y apoyo otorgado. A la Escuela de Graduados en Ciencias de la Facultad de Ciencias Químicas de la UANL por el soporte brindado, el uso de instalaciones y beca de manutención brindada durante el primer año y medio de mis estudios de doctorado.

A mi asesor, Dr. Xristo Zárate Kalfópulos y mi comité tutorial, Dra. María Teresa Garza González, Dr. Jesús Alberto Gómez Treviño y Dr. José Rubén Morones Ramírez quienes contribuyeron enormemente en el mejoramiento y desarrollo del proyecto.

A la coordinadora del posgrado en Microbiología Aplicada, la Dra. María Elena Cantú Cárdenas por su enorme apoyo, sobre todo como mediadora con CONACyT.

Y un agradecimiento muy especial al Dr. Sheng Yang del Institute of Plant Physiology & Ecology de Shanghai Institutes for Biological Sciences, CAS por permitirme realizar mi estancia con su equipo de trabajo.

Nomenclatura

μL	Microlitro
μM	Micromolar
ADN	Ácido desoxirribonucleico
ARN	Ácido ribonucleico
ATCC	American Type Culture Collection
AXR3	Auxin response 3
BL21(DE3)	Cepa de E. coli para expresión de proteínas
bp	Pares de bases
Cas9	Endonucleasa específica
CgR	Péptido señal TAT de C. glutamicum
CRISPR	Clustered, Regularly, Interspaced, Short Palindromic Repeats
CusF	Proteína pequeña de E. coli parte del complejo CusCFBA
DH5α	Cepa de E. coli para producción y manutención de ADN
dNTP's	Desoxirribonucleótidos
DSB	Corte en doble hebra
EDTA	Ácido etilendiaminotetraacético
EK	Enteroquinasa
FPLC	Fast protein liquid chromatography
g	Gramos
G-CSF	Factor estimulante de crecimiento de granulocitos
GFP	Proteína verde fluorescente
GST	Glutation-S-Transferasa
h	Hora
IMAC	Cromatografía de afinidad a iones metálicos
IMB1	Imbibition-inducible 1
IPTG	Isopropil-B-D-tiogalactopiranosido
kbp	Kilopares de bases
kD	Kilodaltones
kV	Kilovoltios
L	Litro
LB	Medio Luria Bertani
M	Molar
MBP	Proteína de unión a maltosa
MCS	Sitio de clonación multiple
mg	Miligramo
MG1655	Cepa de E. coli, primera en ser totalmente secuenciada
mL	Mililitro
mm	Milímetros
mQ	MiliQ (grado de pureza)
nM	Nanomolar
nm	Nanometros

PAM	Protospacer adjacent motif
PCR	Reacción en cadena de la polimerasa
PiF3	Phytochrome-interacting factor 3
PixE	PatA subfamily protein
pmol	Picomoles
rpm	Revoluciones por minuto
SDS-PAGE	Electroforesis en gel de poliacrilamida
Sec	Vía se secreción general
SmbP	Proteína pequeña de unión a metales
Tat	Translocalización doble arginina (Twee Arginine Translocation)
TorA	Péptido señal de la proteína Trimetilamina-N-oxidoreductasa 1
U.V.	Luz ultravioleta
Ω	Ohms

TABLA DE CONTENIDO

CAPÍTULO 1	6
1. INTRODUCCIÓN	7
1.1 <i>Producción de proteínas recombinantes</i>	7
1.2 <i>Organismos hospederos de expresión</i>	8
1.3 <i>Proteínas de fusión</i>	10
1.4 <i>Uso de SmbP y CusF como proteínas de fusión</i>	13
1.5 <i>¿Cómo mejorar los sistemas de expresión y purificación existentes?</i>	15
1.5.1 Promotores	15
1.5.2 Péptidos señal	17
1.5.3 Inducción y condiciones de cultivo	20
1.5.4 Expresión libre de plásmidos	21
1.6 <i>Técnicas de edición</i>	21
1.6.1 CRISPR/Cas9	21
CAPÍTULO 2	23
2. ANTECEDENTES	23
CAPÍTULO 3	29
3. OBJETIVOS E HIPÓTESIS	29
3.1 <i>Objetivo general</i>	29
3.2 <i>Objetivos específicos</i>	29
3.3 <i>Hipótesis</i>	29
CAPÍTULO 4	30
4. METODOLOGÍA	30
4.1 <i>Material, equipo y ubicación del equipo</i>	30
4.2 <i>Estrategia general experimental</i>	30
4.3 <i>Mejoramiento del sistema CusF/E. coli para la expresión y purificación de proteínas recombinantes heterólogas</i>	31
4.3.1 Mutagénesis sitio dirigida del gen CusF	31
4.3.2 Construcción de vectores con CusF3H+	32
4.3.3 Expresión de proteínas de interés y GFP	34
4.3.4 Purificación de GFP marcada con las diferentes proteínas de fusión	36
4.3.5 Remoción de la proteína de fusión CusF3H+_GFP con enteroquinasa	38
4.4 <i>Desarrollo de los sistemas CusF3H+/C. glutamicum y SmbP/C. glutamicum para la expresión, secreción y purificación de proteínas recombinantes heterólogas</i>	39
4.4.1 Diseño de plásmidos para secreción de proteínas heterólogas en C. glutamicum	39
4.4.2 Construcción de plásmidos derivados de pZ8-ptac	42
4.4.3 Expresión y secreción de proteínas en C. glutamicum	44
4.5 <i>Diseño de sistema de expresión de proteínas recombinantes libre de plásmidos usando la técnica CRISPR/Cas9</i>	46
4.5.1 Diseño de pTCueo y pTYahl	46
4.5.2 Construcción de plásmidos pTCueo y pTYahl	49
4.5.3 Edición genómica de E. coli MG1655 y BL21(DE3)	55
CAPÍTULO 5	59
5. RESULTADOS	59
5.1 <i>Mutación y amplificación del gen CusF</i>	59

5.2	<i>Construcción de plásmidos empleando el vector pET30a(+) y el gen de CusF3H+</i>	60
	5.2.1 Construcción de pET30a-CusF3H+	60
	5.2.2 Construcción de pET30a-CusF3H+ con proteínas de interés	61
5.3	<i>Expresión piloto de proteínas marcadas con CusF3H+ y His-tag</i>	62
	5.3.1 Expresión piloto de CusF3H+-GFP	62
	5.3.2 Expresión piloto de CusF3H+_G-CSF	63
	5.3.3 Expresión piloto de CusF3H+_AXR3	64
	5.3.4 Expresión piloto de CuF3H+_IMB1 y His_IMB1	65
	5.3.5 Expresión piloto de CusF3H+_PIF3 y His_PIF3	66
	5.3.6 Expresión piloto de CusF3H+_PixE y His_PixE	67
5.4	<i>Purificación de proteínas</i>	68
	5.4.1 Purificación de CusFc-GFP	68
	5.4.2 Purificación de CusF3H+_GFP	70
	5.4.3 Purificación de His_GFP	71
	5.4.4 Remoción de CusF3H+ y obtención de GFP pura	72
5.5	<i>Construcción de plásmidos derivados de pz8-ptac</i>	73
5.6	<i>Expresión de proteínas en C. glutamicum</i>	78
5.7	<i>Purificación de proteínas secretadas de C. glutamicum</i>	82
5.8	<i>Construcción de pTCueO y pTYahl</i>	86
5.9	<i>Edición genómica de E. coli MG1655 y BL21(DE3)</i>	88
CAPÍTULO 6	91
6.	DISCUSIÓN DE RESULTADOS	91
CAPÍTULO 7	104
7.	CONCLUSIONES	104
CAPÍTULO 8	106
8.	SEGURIDAD EN EL TRABAJO DE DISPOSICIÓN DE RESIDUOS	106

LISTA DE FIGURAS

<i>Secuencias de TorA y CgR_0949</i>	40
<i>Proteínas de fusión SmbP y CgR marcadas con TorA y CgR</i>	40
<i>Diseño del nuevo sitio de clonación múltiple y localización de la proteína de fusión y la proteína reportera</i>	41
<i>Diseño de plásmidos para la serie de TorA a partir del pz8-ptac y serie de CgR a partir del plásmido pz8-ptac</i>	42
<i>Edición genómica empleando CRISPR/Cas9 facilitada por recombinación homologa</i>	48
<i>Inserto diseñado para los plásmidos derivados del pTarget</i>	48
<i>pTCueO y pTYahl, plásmidos derivados de pTarget</i>	49
<i>Amplificación de CusF y sus variantes.</i>	59
<i>Amplificación de CusF3H+.</i>	60
<i>Amplificación de GFP marcado con CusF3H+</i>	61
<i>Amplificación de CusF3H+-G-CSF.</i>	62
<i>Expresión piloto de GFP marcada con CusF3H+</i>	62
<i>Expresión piloto de GFP marcada con CusF3H+</i>	63
<i>Expresión piloto de CusF3H+-G-CSF y His-G-CSF.</i>	64
<i>Expresión piloto de CusF3H+_AXR3e</i>	65
<i>Expresión piloto de IMB1 marcada con CusF3H+ y His,</i>	66
<i>Expresión de PIF3 marcada con CusF3H+ y His.</i>	67
<i>Expresión piloto de PixE marcada con CusF3H+ y His.</i>	68
<i>Cromatograma obtenido de la purificación de CusFc-GFP con FPLC ÄKTA prime plus usando Ni(II).</i>	69
<i>SDS-PAGE 12% de purificación de CusFc-GFP</i>	69
<i>SDS-PAGE 12% de purificación de CusFc-GFP 2</i>	70
<i>Cromatograma obtenido de la purificación de CusF3H+_GFP con el FPLC ÄKTA prime plus empleando una columna tipo IMAC cargada con Ni(II)</i>	71
<i>SDS-PAGE 12% de purificación de CusF3H+_GFP</i>	71
<i>SDS-PAGE de fracciones recolectadas de la purificación de His_GFP</i>	72
<i>SDS-PAGE 15% mostrando el corte de la proteína de fusión empleando enteroquinasa de cadena ligera.</i>	73
<i>Digestión de plásmidos pUC57 para la purificación de los genes +</i>	74
<i>Amplificación de los genes de CusF3H+ marcados con los dos diferentes péptidos señal.</i>	75
<i>Análisis de restricción para los genes de SmbP marcados con los dos diferentes péptido señal.</i>	76
<i>Amplificación de clonas para tamizaje de GFP.</i>	76
<i>Amplificación de GFP para tamizaje de clonas positivas</i>	77
<i>Amplificación del ADN plasmídico extraído de una clona de TorA-SmbP-GFP</i>	78
<i>Análisis del lisado celular y medio de cultivo después de la inducción de TorA-SmbP-GFP.</i>	79
<i>Análisis de la expresión de CgR-SmbP-GF.</i>	80
<i>Análisis de la expresión de Tora-CusF3H+-GFP.</i>	81
<i>Análisis de expresión de CgR-CusF3H+-GFP.</i>	82
<i>Purificación de TorA-SmbP-GFP por FPLC ÄKTA prime plus con Cu(II).</i>	83
<i>Purificación de TorA-CusF3H+-GFP con FPLC ÄKTA prime plus con Ni(II). a) cromatograma, b) SDS-PAGE de las fracciones recuperadas</i>	84

<i>Purificación de CgR-SmbP-GFP por medio del FPLC ÄKTA prime plus usando Ni(II).</i>	85
<i>Purificación de CgR-CusF3H+-GFP con FPLC ÄKTA prime plus empleando Ni (II).</i>	86
<i>Reacciones de amplificación para los fragmentos de los insertos</i>	87
<i>a) Amplificación de alineamiento para Yahl y CueO, b) repetición de amplificación para Yahl</i>	87
<i>a) Análisis de restricción para pTCueO con enzimas BamHI y Sall, b) análisis de restricción para pTYahl con enzimas BamHI y SacI</i>	88
<i>Tamizaje con primer set de iniciadores con productos esperados de 1092 bp, Ctrl1: MG1655+pCassac, Ctrl2: pTCueO, Ctrl3: BL21(DE3)+pCassac</i>	89
<i>Segundo tamizaje con el segundo set de iniciadores con productos esperados en 964 bp, Ctrl1: MG1655+pCassac, Ctrl2: pTCueO, Ctrl3: BL21(DE3)+pCassac</i>	89
<i>Amplificación realizada con los tres conjuntos de iniciadores de las cepas mutantes MG1655ΔcueO::smbp-gfp y BL21(DE3)ΔcueO::smbp-gfp</i>	90
<i>Secuenciación de pET30a-CusF3H+</i>	91

LISTA DE TABLAS

<i>Tabla 1. Iniciadores para amplificación de CusF para la mutagénesis dirigida</i>	32
<i>Tabla 2. Composición de reacción 1x para la amplificación de CusF3H+</i>	32
<i>Tabla 3. Construcciones obtenidas y sus pesos moleculares</i>	34
<i>Tabla 4. Proteínas empleadas para pruebas comparativas junto con su origen</i>	34
<i>Tabla 5. Secuencias de aminoácidos para cada una de las proteínas de fusión marcadas</i>	40
<i>Tabla 6. Composición 1 x para la reacción de amplificación de CusF3H+</i>	43
<i>Tabla 7. Iniciadores empleados para la amplificación de GFP</i>	43
<i>Tabla 8. Composición 1x para amplificación del gen de GFP</i>	44
<i>Tabla 9. Genes seleccionados para corte con Cas9</i>	47
<i>Tabla 10. Iniciadores diseñados para la PCR de alineamiento para la construcción del inserto</i>	50
<i>Tabla 11. Constitución 1 x para amplificación de fragmentos para los insertos</i>	51
<i>Tabla 12. Pesos moleculares y relación molar de fragmentos para el inserto de pTCueO</i>	51
<i>Tabla 13. Pesos moleculares y relación molar de fragmentos para el inserto de pTYahl</i>	52
<i>Tabla 14. Constitución 1 x para PCR de alineamiento para el inserto de pTCueO</i>	52
<i>Tabla 15. Constitución 1 x para PCR de alineamiento para el inserto de pTYahl</i>	53
<i>Tabla 16. Pesos moleculares y relación molar de fragmentos para la segunda reacción de alineamiento para pTYahl</i>	54
<i>Tabla 17. Constitución 1 x para reacción de alineamiento para el inserto de pTYahl</i>	54
<i>Tabla 18. Iniciadores delanteros y reverso para confirmación de edición genómica en E. coli</i>	57
<i>Tabla 19. Reacción 1 x para tamizaje de mutantes</i>	57
<i>Tabla 20. Secuencias de aminoácidos de CusF y el mutante CusF3H+.</i>	59

RESUMEN

Teresa Vargas Cortez

Fecha de graduación: Diciembre, 2017

Universidad Autónoma de Nuevo León

Facultad de Ciencias Químicas

Título de la tesis: EXPRESIÓN Y PURIFICACIÓN DE PROTEÍNAS RECOMBINANTES EN *CORYNEBACTERIUM GLUTAMICUM* Y *ESCHERICHIA COLI* UTILIZANDO SMBP Y UN MUTANTE DE CUSF COMO PROTEÍNAS DE FUSIÓN

Número de páginas:

Candidato para el grado de Doctor en Ciencias con orientación en Microbiología aplicada

Área de estudio: Microbiología molecular

Propósito y Método del Estudio: El objetivo del presente estudio fue el de ampliar la aplicación de dos sistemas de expresión/purificación nobel, CusF/*E. coli* y SmbP/*E. coli*. Una serie de estrategias avanzadas en biología molecular fueron empleadas para dicho fin, incluyendo la modificación del gen de *cusF* generando un mutante con mayor afinidad a Ni(II), el uso de péptidos señal para señalización Tat y la modificación directa del genoma de *E. coli* para generar un sistema de expresión libre de plásmido acoplado a SmbP.

Contribuciones y conclusiones: Se obtuvo la proteína de fusión CusF3H+, la cual es una variante mejorada de CusF permitiendo obtener niveles de pureza mayores al 70% en un solo paso, estando a la par de otras proteínas de fusión en el mercado. Se obtuvieron cuatro plásmidos modificados a partir del p_{z8}-ptac para la obtención de proteínas marcadas con SmbP y CusF3H+ en *C. glutamicum* utilizando dos péptidos señales de naturaleza Tat, TorA de *E. coli* y CgR_0949 de *C. glutamicum*. Con ambos péptidos señal, la proteína GFP marcada con SmbP y CusF fue secretada al medio sobrenadante y purificada mediante IMAC. Mediante la técnica de CRISPR/Cas9 se logró realizar la delección del gen *cueO* para la inserción de un casete con el conjunto de SmbP-GFP, así como el promotor, 5'-UTR y 3'-UTR de la proteína de choque frío CspA (para realizar la inducción por cambios de temperatura) y los dos brazos de homología río arriba y río abajo a *cueO* en el genoma de *E. coli*. Obteniendo así el mutante de *E. coli* MG1655 Δ *cueO*::*smbp-gfp*.

FIRMA DEL ASESOR:

CAPÍTULO 1

1. Introducción

1.1 Producción de proteínas recombinantes

El campo de las proteínas recombinantes se ha expandido a lo largo de las últimas décadas y por el mundo. Conforme pasa el tiempo nuevas proteínas recombinantes surgen para cubrir los crecientes requerimientos de una amplia gama de áreas.¹ De hecho, la aparición de las proteínas recombinantes surgió por la urgente necesidad de obtener biofármacos de manera práctica, eficiente y segura para el paciente. Siendo la insulina el primer biofármaco obtenido de manera recombinante, adquiriendo todas las ventajas de ser producido de esta manera: eficiente, bajo costo y disminución de efectos adversos. De esta manera se abrieron las puertas al uso de la tecnología de ADN recombinante para la producción de proteínas con diversos fines.² Las proteínas recombinantes poseen las mismas características que sus contrapartes de origen natural e incluso esta tecnología ha avanzado tanto que hoy en día las proteínas de origen natural se pueden modificar para mejorar características en específico o incluso diseñar nuevas proteínas sin previa precedencia en la naturaleza.¹

Cada año son publicadas nuevas investigaciones alrededor sobre la producción de las proteínas recombinantes. A grandes rasgos, la producción de proteínas recombinantes se realiza por los siguientes pasos: 1) la selección del gen de interés (aquel que codifica para la proteína deseada), 2) inserción del gen en un vector específico, 3) introducción del vector en un organismo hospedero de expresión, 4) expresión de la proteína, 5) purificación de la proteína recombinante.¹ Dependiendo de una amplia variedad de factores, cada uno de estos pasos puede implicar cierto nivel de dificultad.

Los elementos a tomar en cuenta para la producción de proteínas son: la naturaleza de la proteína, origen de la misma, aplicación final, pureza necesaria, cantidad requerida, costo e infraestructura.^{1,3} Una mala planeación en cualquiera de los pasos mencionados puede derivar, en el peor de los casos, a la nula obtención de la proteína deseada.⁴ Las dificultades que llegan a surgir representan un enorme reto con el cual lidiar. A veces, estas dificultades no llegan a ser superadas y algunas proteínas recombinantes se quedan solo en el protocolo.

Baja o nula expresión, poco rendimiento de biomasa, formación de cuerpos de inclusión, proteína inactiva, baja pureza, por mencionar algunos, son los principales problemas que se presentan al momento de producir una proteína recombinante. Es importante recalcar que cada proteína es diferente entre si, un sistema de expresión-purificación puede funcionar bien con una proteína, pero ser deficiente para otra. Por lo que el sistema a implementar debe ser cuidadosamente seleccionado para disminuir la incidencia de cualquier posible riesgo.⁵

1.2 Organismos hospederos de expresión

Dentro de la producción de proteínas recombinantes, una de las elecciones más importantes es la selección del organismo hospedero. El hospedero de expresión es el encargado de producir la proteína recombinante, por lo que existe una amplia variedad de organismos usados con este fin.⁶ Como células procariotas, células eucariotas (levaduras, animales y de plantas) hasta organismos superiores como plantas, insectos o mamíferos.⁷ De entre todas estas opciones, las células procariotas son las más usadas por su facilidad de uso y bajo costo de producción.⁸ El organismo hospedero de expresión más usado en el mundo, pertenece a dicha categoría y se trata de *Escherichia coli*.⁷

E. coli es un organismo procariota gram negativo cuya genética y bioquímica ha sido estudiada extensivamente.^{9,5} Dicha bacteria ha sido empleada para la producción de proteínas recombinantes desde hace ya varios años mostrando varias ventajas sobre otros organismos hospederos. La producción en *E. coli* es fácilmente escalable a nivel industrial, es un organismo de crecimiento rápido y su mantenimiento es sencillo además de económico.⁴ Todas estas ventajas lo hacen un excelente organismo hospedero de expresión que permite altos rendimientos y un costo de producción relativamente bajo.⁵

Debido a su condición de organismo procariota, *E. coli* puede llegar a presentar algunas desventajas en el sistema de expresión. Principalmente relacionadas con la falta de modificaciones post-traduccionales, posible toxicidad de la proteína expresada y formación de cuerpos de inclusión cuando se expresan genes heterólogos (regularmente de origen eucariótico).⁸ Para hacer frente a dichas complicaciones se debe poner atención especial al diseño de vectores de expresión así como a las condiciones de crecimiento y expresión entre otras cosas.⁴ Una parte importante de los esfuerzos por mejorar la producción de proteínas recombinantes se centra en buscar la optimización en la producción de diversas proteínas en *E. coli* y establecer protocolos que puedan servir de base.

Dentro de la gama de opciones que las células de origen procariota nos ofrecen, se encuentran también *Corynebacterium glutamicum*. Se trata de un organismo gram positivo no esporulado cuyo uso para la producción de proteínas recombinantes ha aumentado en los últimos años.¹⁰ Originalmente, *C. glutamicum* se empleaba principalmente para la producción de aminoácidos de interés comercial (como el propio nombre lo sugiere).⁶ Sin embargo, diversos estudios han mostrado la capacidad de *C. glutamicum* para producir de manera eficiente proteínas recombinantes heterólogas. La

ventaja que proporciona *C. glutamicum* radica en su habilidad para transportar eficientemente proteínas recombinantes plegadas correctamente hacia el medio de cultivo.¹¹ Esta característica permite la recuperación de la proteína de manera directa del medio obteniendo una mayor pureza en menos pasos, ya que en otros sistemas de expresión, la proteína debe ser extraída del citoplasma o el periplasma de la célula.⁷

C. glutamicum ha mostrado ser capaz de expresar proteínas recombinantes funcionales como biofármacos o enzimas de uso industrial, un ejemplo bien documentado es la producción de la proteína verde fluorescente (GFP por sus siglas en inglés).¹² Su principal desventaja es que su uso como organismo hospedero de expresión es de reciente explotación (siendo el primer artículo reportado sobre el tema de 1992 por Liebln y col.).⁶ Por lo tanto la información existente al respecto es relativamente escasa lo que lleva a una falta de opciones comerciales de vectores de expresión disponibles para *C. glutamicum*.¹² Además, la falta de información puede llevar a una planeación inapropiada en la producción de proteínas recombinantes ya que los principales retos que podrían presentarse aún no se han explorado del todo.

Aún con la mejor elección posible de hospedero y vector de expresión, sin duda, la purificación de proteínas siempre será una parte desafiante del proceso de obtención de proteínas. Decidir la mejor opción para purificar una proteína de interés (tanto recombinantes como de origen natural) implica tomar en cuenta una serie de factores determinantes. Entre ellos, la naturaleza física y bioquímica de la proteína, el compartimento celular de expresión, el organismo hospedero de expresión y las posibles interferencias que pudiesen haber.⁵

1.3 Proteínas de fusión

Una vez que se comenzaron a producir las proteínas recombinantes de manera exponencial comenzaron a surgir nuevos retos, entre ellos encontrar una forma de purificar proteínas recombinantes de manera práctica y obteniendo la mayor pureza posible en pocos pasos. Las proteínas de fusión surgieron como una respuesta al nuevo reto planteado en base a la purificación.⁵ Las proteínas de fusión pueden ser pequeñas cadenas peptídicas hasta proteínas completas, su principal característica es su elevada afinidad hacia un ligando específico. Esta afinidad se puede adaptar a un sistema de cromatografía por afinidad, donde el ligando se emplea unido a una fase estacionaria.¹³ La proteína de fusión se une al ligando de manera específica y se separa del resto, finalmente la proteína es eluida empleando un buffer de elución que permite romper la unión de la proteína de interés y el ligando unido a la fase estacionaria.³ Cuando una proteína es purificada mediante cromatografía de afinidad es posible obtener una alta pureza en tan solo un paso.

Para purificar una proteína recombinante empleando una proteína de fusión, es necesario clonar el gen de la proteína de fusión junto al de la proteína de interés, de forma que ambas queden bajo el marco de lectura.⁴ Así, al expresar la proteína de interés, la proteína de fusión se expresará en conjunto formando una proteína quimérica, es importante evitar agregar un codón de terminación entre ambas proteínas para evitar la terminación temprana de la traducción. Esta proteína quimérica poseerá la proteína de fusión en alguno de sus extremos (por lo regular en el N-terminal) permitiendo su unión a un ligando específico y su posterior purificación mediante cromatografía de afinidad obteniendo un alto grado de pureza.³

Las proteínas de fusión han mostrado no solo ser excelentes para la purificación de proteínas de interés, sino que también aumentan los niveles de expresión de la proteína

en sistemas heterólogos. Un ejemplo muy típico es la proteína de unión a maltosa (MBP de 42 kDa) y ha sido usada por su habilidad para elevar la expresión y solubilidad de la proteína de interés.^{9,13} Otra opción comercial es la proteína glutatión S-transferasa (GST) de 26 kDa, que al igual que MBP ha mostrado elevar los niveles de expresión de la proteína de interés. GST eleva también la solubilidad de la proteína de interés, aunque no al mismo grado que MBP.⁴

La principal desventaja que llegan a presentar las proteínas de fusión MBP y GST se hace visible al momento de remover la proteína de fusión de interés.¹⁴ En la mayoría de los casos para poder estudiar o emplear la proteína que se está produciendo, es necesario remover la proteína de fusión. Esto generalmente se logra cortando con alguna proteasa que reconozca algún sitio específico entre la proteína de interés y de fusión.¹⁴ Con proteínas de pesos moleculares elevados, el rendimiento final de la proteína de interés obtenido después del corte suele ser mucho menor al esperado. Considerando una proteína de interés de 20 kDa, unido a la proteína GST de 26 kDa se obtiene una proteína quimérica de 46 kDa. Es decir, el peso total de la proteína quimérica está constituido en un 56.5% por la proteína de fusión. Por lo que, entre menor sea el peso de la proteína de fusión, mayores rendimientos finales se obtendrán.

La remoción de la proteína de fusión puede significar también en algunos casos la pérdida de la estructura terciaria de la proteína de interés.¹⁵ Donde la actividad deseada se pierde y es necesario involucrar pasos adicionales para el repliegamiento de la misma. Esta situación suele observarse con cierta regularidad en el caso de MBP.¹⁴ Por lo tanto, la generación de proteína soluble, no siempre es en todos los casos un sinónimo de actividad en la proteína de interés obtenida, sobre todo después del corte.

Debido a estas razones, se puede concluir que una proteína de fusión ideal deberá poseer las siguientes características: tener bajo peso molecular, aumentar los niveles

de expresión de la proteína de interés, así como su solubilidad, permitir la purificación mediante cromatografía de afinidad y no afectar la estructura terciaria de la proteína de interés antes y después de su remoción y deben poder aplicarse a una amplia variedad de proteínas de diversas características y naturaleza. Por tanto, el desarrollo y búsqueda de nuevas proteínas de fusión y el diseño de un sistema adecuado de expresión-purificación para su uso son importantes áreas de oportunidad.¹³

1.4 Uso de SmbP y CusF como proteínas de fusión

SmbP (del inglés: *Small metal-binding Protein*) es una proteína de bajo peso molecular aislada de *Nitrosomonas europaea*. Tiene la capacidad unir iones metálicos divalentes como el Ni(II) y el Cu(II) y algunos trivalentes como el Fe(III). SmbP posee un importante rol en la regulación de metales en *N. europaea* con el exterior. En estudios previos se ha producido SmbP de manera recombinante y para su purificación se ha empleado cromatografía tipo IMAC (*immobilized-metal affinity chromatography*) aprovechando su capacidad de unión a metales sin necesidad de unirla a otra proteína de fusión.¹⁶

Recientemente se ha comprobado el uso eficaz de SmbP como proteína de fusión para la purificación de proteínas de interés empleando cromatografía tipo IMAC. La expresión de la proteína de interés en *E. coli* empleando a SmbP como proteína de fusión se ve elevada aun siendo SmbP una proteína heteróloga para dicho organismo hospedero. Así mismo se ha comprobado un aumento en la solubilidad de la proteína de interés comparable al de otras proteínas de fusión como MBP y GST. Estas propiedades en conjunto con su bajo peso molecular la han convertido en una proteína de fusión idónea.¹⁷

La proteína CusF ha sido objeto de estudio por varios años, pero recientemente se ha considerado su aplicación como proteína de fusión. Esta proteína es parte del complejo

de flujo CusCFBA de *E. coli*. Al igual que SmbP participa en la regulación de iones metálicos con el exterior; posee una alta afinidad por iones metálicos monovalentes como Ag(I) y Cu(I). CusF es también una proteína de bajo peso molecular que ha mostrado aumentar la expresión y solubilidad de diversas proteínas de interés como la proteína GFP en comparación con otras proteínas de fusión comerciales como MBP y GST. Así mismo, su afinidad por el ión Cu(II) hace viable la purificación de la proteína de interés unida a CusF por cromatografía tipo IMAC.¹⁸

CusF y SmbP han sido probadas con proteínas de diferente naturaleza y peso molecular. Su habilidad como proteínas potenciadoras de la expresión y solubilidad ha sido ya verificada. Sin embargo, es necesario ampliar la perspectiva de ambas proteínas y mejorar las condiciones del sistema de expresión-purificación con el fin de desarrollar la propuesta de uso más adecuada. En el caso de CusF se propone realizar modificaciones en su estructura, propiamente añadir tres histidinas en el N-terminal con el fin de aumentar su afinidad a iones de Ni(II) y permitir la purificación con IMAC funcionalizada con dichos iones. Se sabe que el uso de Ni(II) en cromatografía tipo IMAC eleva en gran medida el grado de pureza de la proteína a purificar en comparación al Cu(II).¹⁸

Para la mejora de SmbP se generarán nuevos vectores de expresión que amplíen sus términos de uso aplicándolo a nuevos hospederos de expresión. En este caso, se pretende adaptar el sistema para su uso en *C. glutamicum* y aprovechar las ventajas inherentes de dicho organismo. De esta manera, se simplificará el método de purificación y se elevará la pureza de las proteínas obtenidas anteriormente en *E. coli* empleando a SmbP y cromatografía tipo IMAC.

1.5 ¿Cómo mejorar los sistemas de expresión y purificación existentes?

Existen diversos ángulos aplicables a cada uno de los pasos que compete la producción y purificación de las proteínas recombinantes. Algunas de las aproximaciones que se han desarrollado para el mejoramiento de los sistemas ya existentes y en el desarrollo de nuevos es la elección del promotor, el uso de péptidos señal, las condiciones de cultivo e inducción e incluso desarrollo de sistemas libres de plásmidos en bacterias.

1.5.1 Promotores

Los promotores consisten en secuencias de ADN reconocibles de manera específica por la ARN polimerasa participando así dentro de la regulación transcripcional. Estas secuencias son cortas y se posicionan río arriba al gen que regulan. Existen diferentes tipos de promotores y cada uno varía también dependiendo de la especie de origen, es obvio decir que los promotores de origen procarionta solo funcionarán en organismos de dicho origen y no funcionarían en organismos de origen eucariota. Dependiendo del diseño del sistema de expresión se selecciona el promotor que mejor se adapte a nuestros requerimientos. En la mayoría de los casos lo ideal sería usar un promotor fuerte e inducible que nos permita la generación de una alta cantidad de proteína.⁴ Sin embargo, a veces es necesario mantener la expresión baja en situaciones donde la proteína puede presentar toxicidad al organismo hospedero, por tanto se deberá seleccionar un promotor débil. Así mismo, se puede decidir entre seleccionar un promotor constitutivo en vez de uno inducible si la situación sugiere una expresión constante de la proteína.

Los promotores inducibles permiten controlar la expresión al modificar las condiciones que dan lugar a la activación o inactivación del mismo. Por ejemplo, para la expresión de proteínas en organismos como *E. coli* uno de los promotores más empleados es el

promotor *lac*, proveniente del operón *lac* del mismo organismo. Este promotor puede ser inducido mediante IPTG o lactosa teniendo así cierto control de la expresión (por lo regular siempre existe una cantidad de expresión basal constante). Con el avance de la biología molecular, incluso han surgido combinaciones como el *T7lac* que se trata de un promotor originado de la combinación del promotor del fago *T7* y del operón *lac*. Gracias a esta combinación la expresión basal es prácticamente nula (a comparación de la que el promotor *lac* presenta) debido a que necesita específicamente de la ARN polimerasa *T7*, por lo que el uso de una cepa especial es estricto. Al igual que el promotor *lac*, *T7lac* puede ser inducido con IPTG o lactosa.

Además de los promotores controlados por inductores o represores, existen otros que se activan por cambios físicos en el medio, por ejemplo, en la concentración de sales, ambientes altamente reductores u oxidantes, condiciones de anaerobiosis o cambios en la temperatura.⁹ La identificación de los promotores inducibles ante tales condiciones abre una completa gama de posibilidades para su explotación. Considerando que una de las condiciones que favorecen el correcto plegamiento de una proteína es la expresión a bajas temperaturas sonaría lógico la explotación de promotores con esta característica para dicho fin.

1.5.1.1 Promotores regulados por cambios de temperatura

Diferentes situaciones de estrés ambientales fomentan la expresión de diferentes proteínas que le permiten al organismo sobrellevar las condiciones a la que es sometido. Los cambios de temperatura no son la excepción a esto. En *E. coli* se comenzaron a estudiar las proteínas que se llegaban a producir en estos cambios de temperatura, regularmente a bajas temperaturas. Se descubrió así que cambios por debajo de la temperatura óptima de crecimiento (37 °C) derivaba en la producción de proteínas de choque frío.¹⁹

El interés de por la regulación de las proteínas de choque frío ha derivado de su estudio la identificación de varios promotores a través del tiempo y su aplicación para regular la expresión de proteínas en sistemas heterólogos. Uno de estos promotores es el *pL* originario del phago lambda, el cual presenta una actividad a temperaturas de 20 °C la cual declinaba al aumentar la temperatura.²⁰

Una de las proteínas de choque frío más estudiada es la CspA, y su regulación ha sido objeto de estudio. En 1996, Vasina y col, reportaron el uso de un sistema para la expresión que empleaba el promotor *cspA* para la producción de proteínas recombinantes a bajas temperaturas. En dicho artículo produjeron la proteína recombinante β -galactosidasa bajo el control del promotor *cspA* creando un sistema en el que puntúan tres mayores ventajas: una represión relativamente eficiente de la proteína a temperaturas fisiológicas, una rápida inducción de la expresión después de bajar la temperatura y una inducción que no requiere el uso de reactivos químicos.²¹ El éxito de dicho promotor se revela en subsecuentes artículos donde se dieron a la tarea de emplearlo para la producción de proteínas recombinantes, así mismo una serie de plásmidos comerciales, conocidos como pCOLD se encuentran el mercado.^{4,22}

1.5.2 Péptidos señal

En un sistema procarionta, las proteínas que son expresadas se pueden secretar y/o acumular en diferentes áreas, es decir, en el caso de *E. coli*, las proteínas pueden ser expresadas en el citoplasma, el periplasma, como proteínas transmembranales o incluso ser secretadas al medio. En general, los sistemas de exportación de las proteínas en los sistemas procariontas son similares entre ellos. Entre los sistemas más estudiados se encuentra el sistema Tat (Twin arginine translocation) y el sistema Sec (general secretion pathway).

Para que una proteína tenga un fin diferente al de acumularse en el citoplasma, es decir, para que termine en el periplasma, como proteína transmembranal o incluso ser secretada, debe pasar por un determinado medio de transporte. Es así como se puede manipular de cierta manera el recorrido que tendrá una proteína después de ser sintetizada. El camino y el fin que seguirá dependerá mucho de la naturaleza de la proteína y del propio sistema hospedero de expresión. Una vez que estos dos factores se han seleccionado, las proteínas pueden ser marcadas con un péptido señal adecuado que permita a la proteína ser transportada de la manera deseada.

El sistema Sec tiene como fin el transporte de proteínas al periplasma de la bacteria y existe una vasta cantidad de péptidos señal identificados que permiten la exportación eficiente de las proteínas marcadas. El sistema Tat funciona también como un método de transporte para las proteínas, sin embargo es menos usada que el sistema Sec.²³ El sistema de translocalización Tat se distingue primordialmente del Sec en el que el primer requisito para que estas sean translocalizadas es que deben estar correctamente plegadas.⁵ Las proteínas que siguen esta vía deben estar marcadas con un péptido señal especial que contiene el motif de doble arginina (*twin arginine* en inglés, de allí su nombre), se podría decir que dichas proteínas tienen una “señal de evasión sec” (sec avoidance signal) que evita su tránsito por la vía Sec.²⁴ Los péptidos señal en general poseen tres partes, la región n (dominio N-terminal), la región h (dominio hidrofóbico) y la región c (dominio C-terminal), en el caso de los péptidos señal de Tat, el motif doble arginina se encuentra en la región N y por lo regular sigue un patrón SRRxFLK (donde “X” señala a un aminoácido polar).^{23,25} Los péptidos señales para la vía de Tat suelen ser más largos que los de Sec con una región h menos hidrofóbica y una región c con mayor cantidad de residuos básicos.²³

El sistema de translocasa de Tat se divide en varias proteínas transmembranales, en el caso de *E. coli* se identifican las proteínas TatA, TatB y TatC las cuales son expresadas en un solo operón conocido como *tatABC*. Algunas bacterias gram negativas contienen también genes para proteínas TatD, una proteína citosólica, y TatE que se considera es un duplicado de TatA, este último es un gen monocistronico.²⁵ El sistema Tat en bacterias gram positivas es bastante similar al de las bacterias gram negativas, pero cuenta con algunas diferencias importantes. Las bacterias gram positivas expresan en el mayor de los casos solo variantes de las proteínas TatA y TatC, cuyos genes se encuentran concentrados también en un solo operón produciendo translocasas Tat mínimas.²⁵ La expresión de las translocasas en gram positivas puede variar ligeramente entre ellas, los sistemas más estudiados son los del género estafilococo, micobacterias y las especies *B. subtilis* y *C. glutamicum*. Esto indica que aunque el sistema conceso está compuesto por TatA, TatB y TatC, el centro esencial del sistema lo constituyen el binomio de TatA-TatC.²⁶

El sistema Tat es una peculiar área de oportunidad debido a la capacidad de exportar proteínas al periplasma y secretarlas de manera plegada. El sistema Tat de *C. glutamicum* se ha convertido de particular interés y ha sido estudiado en los últimos diez años.²⁷ En diferentes artículos, se ha mostrado la capacidad que tiene esta bacteria para secretar proteínas heterólogas recombinantes al medio de cultivo. Más de una proteína ha sido secretada con éxito y se han podido intercalar diversos péptidos señal del sistema Tat propios de otras bacterias con particular éxito.^{27,28,29} Manipular el sistema de translocalización para la secreción/expresión de cierta proteína de interés es posible empleando péptidos señales adecuados para dicho fin.

1.5.3 Inducción y condiciones de cultivo

En el caso de sistemas basados en expresión en células procariotas es importante tomar en cuenta las condiciones de cultivo con detalle. Estas incluyen: temperatura de crecimiento, agitación, tipo de reactor y componentes del medio de cultivo las cuales se deben adaptar a dos cuestiones, el organismo hospedero y el método de inducción. En el caso de *E. coli* existen varias cepas que han sido modificadas genéticamente para favorecer la expresión de proteínas, por ejemplo, la línea de BL21 (incluyendo DE3 y DE3plys y Star), Rosetta y Origami por mencionar algunas. Algunas de estas requieren de vectores especiales que se complementan con ellas para permitir la expresión de la proteína de interés.⁵

Una extensa variedad de vectores para la expresión de proteínas se encuentra presente en el mercado haciendo un excepcional dúo con las cepas modificadas. Cada uno de estos vectores posee dentro de su secuencia genes de resistencia a antibióticos para la selección de clonas positivas, genes de proteínas o péptidos etiquetas y promotores/terminadores que permiten tener un control de la expresión. De manera que la inducción es posible al agregar un inductor específico, como es el caso del promotor *lac* cuya inducción se cataliza por la lactosa o un similar, como el IPTG.

Para seleccionar el método de inducción se debe tomar en cuenta el origen y naturaleza de la proteína a expresar. Elegir un promotor fuerte como el *lac* asegura la expresión de una gran cantidad de proteína, sin embargo, esta podría resultar tóxica para la bacteria, por lo que sería conveniente la elección de un promotor débil. O incluso, la proteína podría ser termolábil, en donde el uso de un promotor inducido a bajas temperaturas podría ser preferible. Por lo regular, las condiciones óptimas de expresión se dan por prueba y error y es por mucho dependiente de la proteína de interés.

1.5.4 Expresión libre de plásmidos

Hoy en día la expresión de proteínas recombinantes se realiza en su mayoría a través de sistemas dependientes de plásmidos. Dichos plásmidos poseen características especiales como ya se ha mencionado anteriormente que permiten la expresión de manera regulable de grandes cantidades proteína recombinante. Sin embargo, se ha demostrado varias veces que la expresión de proteínas con estos sistemas puede llegar a ser mermada por el organismo hospedero de expresión al experimentar una alta carga metabólica.³⁰ Se ha planteado que la integración del compendio genético de un plásmido de expresión dentro del cromosoma del hospedero de expresión podría aliviar la carga metabólica que experimenta al realizar la expresión en base a plásmidos eliminando también algunas otras desventajas como la inestabilidad que pueden llegar a presentar los plásmidos incluyendo la segregación del mismo de la cepa.^{30,31}

1.6 Técnicas de edición

El constante avance en técnicas de ingeniería genética nos permite modificar lo ya existente, con el fin de mejorar alguna característica que nos sea de especial interés. Dicha aplicación es vasta, desde la modificación de rutas metabólicas a la modificación de las propias proteínas. Con el uso de técnicas avanzadas de edición, es posible maximizar la eficiencia de los sistemas de proteínas de fusión novel.

1.6.1 CRISPR/Cas9

Una de las técnicas más novedosas para la edición genómica consiste en el sistema CRISPR (del inglés: *Clustered, Regularly, Interspaced, Short Palindromic Repeats*). CRISPR fue descubierto como un sistema de defensa contra virus y es usado hoy en día como una técnica eficiente para la edición genómica en células eucariotas y procariontas. Los loci de CRISPR están compuestos de una serie de repeticiones

separadas por un “espaciador”, las cuales son secuencias que son complementarias a genomas de bacteriófagos y otros elementos genéticos móviles.³²

El acomodo entre las repeticiones y el espaciador generan un ARNcr (ARN guía de CRISPR), además para el corte es necesaria la presencia de un motif conocido como PAM (*Protospacer-Adjacent Motif*). El sistema emplea la enzima Cas9 asociada a CRISPR, la cual se trata de una nucleasa que emplea el ARNcr como guía para el área específica de corte, permitiendo la edición dirigida de un sitio fijo dentro del genoma. Es necesario también un ARNtr (ARN transactivador del ARNcr) que forma junto con el ARNcr el complejo de activación sitio-específico para el corte de Cas9.^{32,33}

Con el sistema de CRISPR se puede editar el genoma, insertando el promotor, resistencia a antibiótico, gen de proteína de fusión (SmbP y CusF) y de interés. El sistema de expresión se puede ver favorecido mediante la inserción de promotores termorregulables. Se han descrito varios promotores que regulan la expresión de proteínas en dependencia de la temperatura. Ejemplos de estos son el promotor *pL* y el *cspA* que se usan en vectores comerciales como alternativa para la expresión de proteínas.³⁴

El promotor *cspA* es altamente inducible a bajas temperaturas (15 °C – 23 °C) y es reprimido a mayores temperaturas (37 °C) lo cual lo hace de particular interés ya que la expresión de proteínas a bajas temperaturas es siempre favorable.³⁵ Las bajas temperaturas pueden permitir un adecuado plegamiento de las proteínas lo que permite obtener niveles altos de proteína soluble y activa.

CAPÍTULO 2

2. Antecedentes

La cromatografía tipo IMAC fue uno de los primeros sistemas de purificación de proteínas recombinantes empleados. Este sistema fue desarrollado en 1975 por Jerker Porath y está basado en la interacción entre iones metálicos [Co(II), Ni(II), Cu(II) y Zn(II), siendo el Cu(II) el que presenta mayor afinidad] en una fase estacionaria con residuos específicos de la proteína a purificar (ya sea nativa o quimérica).³ Dichos iones metálicos son considerados como aceptores de pares de electrones, mientras que los residuos de aminoácidos actúan como donadores gracias a los heteroátomos en su estructura. Los residuos específicos son: histidina, cisteína y triptófano, conocidos comúnmente como triada de Porath.³⁶

El aminoácido más importante para dicha interacción es la histidina. Es por esto que se desarrolló junto con la técnica de IMAC un marcador afín, en este caso los His-tag, que pueden variar en su contenido de histidinas, siendo el His6-tag el más usado. Este marcador, que es de apenas unos cuantos aminoácidos, tiene la capacidad de permitir la purificación de una proteína de interés mediante IMAC.^{3,36}

La GST es una de las proteínas de fusión de mayor uso hoy día. GST de aproximadamente 26 kDa fue usada por Taylor y col., en 1994 como proteína de fusión para la purificación de proteínas de interés.³ MBP es por mucho una de las proteínas más usadas para la expresión y purificación de proteínas recombinantes, ya que ha mostrado aumentar la solubilidad de la proteína de interés.^{14,37} El uso de MBP, de aproximadamente 42kDa, como proteína de fusión fue descrito por primera vez por Di Guan y col., en 1988 empleando amilosa como ligando de unión en la fase estacionaria para su purificación.^{3,13}

La proteína SmbP fue aislada y caracterizada en el 2004 por Barney y col. Debido a la afinidad de la proteína por metales fue purificada mediante IMAC con Ni(II). Considerando dicha cualidad y que se trata de una proteína pequeña, fue probada como proteína de fusión con éxito en el 2015 por Vargas-Cortez y col., en *E. coli*. Diversas proteínas de diferente peso molecular y naturaleza fueron expresadas con dicha proteína, observando una clara potenciación de la solubilidad de las proteínas de interés comparándolas con MBP y GST. Así mismo se purificó la proteína GFP marcada con SmbP mediante IMAC con Ni(II) y Cu(II).¹⁷ Por su parte, la proteína CusF fue estudiada en el 2016 por Cantu y col., empleando una aproximación similar a la empleada con SmbP. Las pruebas de expresión piloto mostraban un incremento en la expresión y solubilidad de las proteínas de interés marcadas con CusF, también equiparable al presentado con otras proteínas de fusión como GST y MBP.¹⁸ Cabe resaltar, que en ambos casos, las proteínas fueron también exportadas al periplasma de la bacteria empleando los péptidos señal propios de cada proteína. A pesar de que las proteínas marcadas con SmbP podían ser fácilmente purificadas con IMAC Ni(II), las proteínas marcadas con CusF no mostraban dicha posibilidad y la purificación debía ser realizada empleando IMAC Cu(II). El uso de Cu(II) como ligando implicaba comprometer en gran medida la obtención de una elevada pureza en un solo paso de purificación como se esperaba.

C. glutamicum ha sido considerada como una opción para la producción de proteínas recombinantes para hacer frente a las desventajas que puede llegar a presentar *E. coli*. *C. glutamicum* ha sido utilizada durante muchos años para la producción de aminoácidos de interés industrial, además no es patógena y no genera toxinas a diferencia de *E. coli* por lo que es considerado como un organismo GRAS.³⁸ *C. glutamicum* presenta un muy eficiente sistema de translocalización Tat que permite la

secreción de proteínas, cabe mencionar que *C. glutamicum* presenta una mínima actividad de proteasas, aumentando la eficiencia del uso de esta vía.³⁹ En el año 1992, Liebl y col., desarrollaron un sistema para la producción de la nucleasa *SNase* (originaria de *S. aureus*) usando como hospedero *C. glutamicum*.⁴⁰ A partir de entonces, varias proteínas han sido producidas y exportadas con éxito en dicha bacteria. Tanto que un sistema conocido como CORYNEX™ ha sido desarrollado en Japón por la compañía Ajimoto para la expresión y secreción de proteínas de interés en *C. glutamicum*.⁴¹ Diferentes péptidos señal de *C. glutamicum* han sido caracterizados ya y usados para la secreción de proteínas recombinantes, además, se han podido utilizar péptidos señal Tat heterólogos, como el TorA perteneciente de *E. coli*, donde se ha mostrado que el sistema Tst de *C. glutamicum* tiene afinidad también por dichos péptidos señal y es capaz de usarlos para la secreción de proteínas por dicha ruta.²⁹

En la mayoría de los casos, la producción de proteínas en hospederos de origen procarionta se realiza mediante la transformación de la cepa con un plásmido diseñado para dicho fin. Este plásmido está bajo un promotor regulable y confiere resistencia a antibióticos. Sin embargo, se han observado diversas desventajas en el uso de plásmidos, por ejemplo, la generación de copias y la propia resistencia a antibióticos generan una pesada carga en el metabolismo de la bacteria, esto puede derivar a inestabilidad y/o mutaciones en el plásmido así como una disminución en la velocidad de crecimiento.^{42,43} Podría entonces suponerse que modificar cierta cepa directamente en su genoma para la producción de una proteína en específico podría ser una buena estrategia introduciendo dentro del genoma las mismas características necesarias para la inducción encontradas en un plásmido destinado a ello.^{31,30} Pero, la literatura reporta poca investigación realizada con el fin de generar cepas estables para la producción de proteínas recombinantes, al contrario de los realizados con plásmidos. Si bien, la

mayoría de los reportes de modificaciones a nivel genómico son para la producción de metabolitos menos complejos (aminoácidos, vitaminas, metabolitos secundarios heterólogos, etc) aun así la modificación de una ruta o introducción de otra pueda significar un enorme reto.⁴⁴ La modificación genómica de *E. coli* tiene como principal fin, la generación de cepas que puedan mejorar la producción de proteínas, es decir, conferirles una o varias características que faciliten ya sea la expresión, la inducción o el plegamiento de las proteínas recombinantes en general.^{45,46} En el 2010, Striedner y col., producen un sistema de expresión libre de plásmido empleando el promotor *T7* en *E. coli*. En dicha investigación se expresó la proteína GFP tanto en un sistema con plásmido pET11a y la cepa generada con el gen GFP bajo la influencia del promotor *T7*, la expresión para el sistema libre fue de casi el doble ya que permitió el crecimiento continuo de la cepa, contrario al sistema con plásmido donde la producción y crecimiento de la bacteria se detuvo a las 6.5 h.³¹ En el 2012 Li y col., introdujeron en el genoma de *E. coli* promotores del tipo *tac* en tándem y emplearon también la proteína GFP como reportera. Li y col., encontraron que en las construcciones en las que se tenía mayor cantidad de promotores, la cantidad de proteína producida era mayor.⁴²

Hasta el momento existen varios métodos para la integración específica de genes en el cromosoma de bacterias como *E. coli*. Entre estas técnicas se encuentran MAGE (Multiplex automated genome engineering), recombinación λ Red (recombineering), TRMR (trackable multiplex recombineering), ZFNs (zinc finger nucleases), TALEN'S (transcription activator-like effector nucleases) y CRISPR acoplada tanto a Cas9 o a Cpf1.^{47,48} Este último posee una alta especificidad y una practicidad superior a sus antecesores ya que la endonucleasa usada (Cas9 o Cpf1) solo requiere de un pequeño ARN guía siendo muy poco probable la interacción de la enzima fuera del ADN objetivo.⁴⁹ La eficiencia llega a aumentarse cuando es combinada con otra técnica, por

ejemplo, recombinación con λ Red.⁵⁰ El sistema CRISPR fue descrito por primera vez en 1987 por Nakata y col., cuando estudiando la enzima *iap* encontraron dichas secuencias.^{33,51} Posteriormente fueron nombrados CRISPR (*Clustered Regular Interspaced Short Palindromic Repeats*) por Jansen y col. En el 2005 Mojica y col., descubrieron que dichas secuencias contenían ADN de bacteriófagos.³³ En ese mismo periodo, Bolotin y col., encontraron genes tipo *cas* que codificaban para una ADN endonucleasa. Estos hallazgos permitieron comprender que la función primaria del sistema CRISPR/Cas era la degradación de ADN exógeno.³³

A partir de entonces, el sistema CRISPR/Cas9 fue usado y adaptado con el fin de permitir la edición del genoma de diferentes líneas celulares. Siendo posible la inserción, delección, edición o mutación de sitios específicos guiados mediante este sistema.^{32,33} Este sistema sigue evolucionando y la aplicación en sistemas bacterianos es ya posible con métodos cada vez más estandarizados.³² Sin embargo, a pesar del enorme crecimiento, optimización y uso de la técnica de CRISPR, su aplicación para edición en genomas de origen procarionta se ha visto mermado por la baja tasa de éxito, esto debido a que la edición mediante CRISPR se realiza mediante el sistema de HRH, cuya eficiencia en organismos procariontas es menor que en eucariotas y los cortes que realiza la enzima Cas9 o Cpf1 en el genoma de la bacteria resultan ser mortales.⁵² A pesar de dicho obstáculo, en el 2013, Jiang y col., emplearon la técnica de CRISPR dual-RNA:Cas9 logrando la edición genómica de *E. coli*.⁴⁹ Posteriormente, en el mismo año, 2015, Li y col., así como Jiang y col., lograron también mejorar la eficacia de CRISPR-Cas9 para la edición genómica de *E. coli*, ambos aplicando un sistema de doble plásmidos apoyándose en recombinación con λ Red.^{47,53}

Una de las principales aproximaciones en el diseño de estrategias para expresar proteínas en *E. coli* tanto en sistemas con plásmidos y libres de ellos es en el uso

racional de promotores para modificar y controlar la expresión de la proteína.^{31,42} El uso del promotor y terminador de la proteína de choque frío CspA fue empleada en 1996 por Vasina y col., para la producción de proteínas recombinantes en *E. coli* generando una serie de plásmidos a partir de dichos hallazgos conocidos como pCOLD, cuya propiedad radica en la capacidad de inducir la expresión de proteínas recombinantes al cambiar la temperatura de 37°C a 25°C.^{4,21}

En cuanto a la producción de G-CSF, Choe, H. y col., en el 2014 publicaron una serie de comparaciones de expresión y solubilidad de G-CSF con diferentes proteínas de fusión y marcadores. Dichas proteínas incluían a la isomerasa de unión a disulfuro (PDI), a la proteína de N-utilización de sustancia (NusA), a MBP, al dominio b de PDI (PDib´a´) a GST y a la proteína tioredoxina (Trx). También se empleó, como marcador, a His6-tag. Todas las proteínas quiméricas fueron expresadas en el citoplasma de *E. coli*. Cada proteína de fusión afectó de manera diferente la expresión y solubilidad del G-CSF a cada una de las condiciones de expresión sometidas. En el caso de la proteína marcada con His6-tag, esta se expresó en forma insoluble bajo las condiciones probadas.⁵⁴

CAPÍTULO 3

3. Objetivos e hipótesis

3.1 Objetivo general

Desarrollar metodologías para la expresión y purificación de proteínas recombinantes heterólogas en los microorganismos genéticamente modificados *Escherichia coli* y *Corynebacterium glutamicum*.

3.2 Objetivos específicos

- Mutar el gen CusF para obtener dos variantes mutantes por medio de mutagénesis dirigida.
- Caracterizar el mutante y optimizar el método de purificación de proteínas recombinantes por medio de cromatografía de afinidad.
- Producir proteínas recombinantes marcadas con SmbP y las variantes mutantes de CusF en *C. glutamicum* como microorganismo hospedero.
- Modificar el genoma de *E.coli* por medio del sistema CRISPR para la producción de proteínas recombinantes marcadas con SmbP y las variantes mutantes de CusF, así como inducir la expresión de proteínas recombinantes por medio de cambios de temperatura.

3.3 Hipótesis

A través de la ingeniería genética es posible mejorar el sistema de expresión/purificación de las proteínas SmbP y CusF, modificando la afinidad a ligandos de IMAC, aprovechamiento del sistema TAT en un diferente hospedero y generando un sistema libre de plásmidos.

CAPÍTULO 4

4. METODOLOGÍA

4.1 Material, equipo y ubicación del equipo

Los equipos que se mencionaran a continuación para el desarrollo del presente proyecto se encuentran en el Laboratorio de Expresión y Purificación de Proteínas, ubicado en la División de Estudios de Posgrado de la Facultad de Ciencias Químicas de la UANL. El desarrollo de la metodología correspondiente al punto 3.5. se realizó en el laboratorio del Institute of Plant Physiology & Ecology de Shanghai Institutes for Biological Sciences (CAS).

Construcción de plásmidos (técnicas clásicas de clonación): Microcentrífuga, cámaras de electroforesis vertical y horizontal, fuentes de poder, termociclador, incubadora, plancha de calentamiento/agitación, espectrofotómetro Nanodrop™, autoclave.

Expresión y purificación de proteínas: Microcentrífuga, incubadora con agitación, campana de flujo laminar, espectrofotómetro UV-Vis, planchas de calentamiento/agitación, homogeneizador para lisis con perlas de vidrio, ÄKTA prime plus™ (FPLC), columnas HisTrap HPTM, columnas HiTrap Q HPTM.

Edición genómica con CRISPR/Cas9: Microcentrífuga, cámara de electroforesis horizontal, fuente de poder, termociclador, incubador, espectrofotómetro Nanodrop™, autoclave, electroporador.

4.2 Estrategia general experimental

El desarrollo del proyecto se realizó en tres etapas: 1) Mutación del gen de CusF y desarrollo del nuevo sistema de expresión/purificación, 2) Desarrollo del sistema de expresión, secreción y purificación de proteínas recombinantes en *C. glutamicum*, 3)

Desarrollo de un sistema de expresión de proteínas recombinantes libre de plásmido en *E. coli* mediante la técnica CRISPR/Cas9. Para cada una de las etapas se realizaron técnicas clásicas de clonación, así como técnicas avanzadas de edición genómica.

4.3 Mejoramiento del sistema CusF/*E. coli* para la expresión y purificación de proteínas recombinantes heterólogas

4.3.1 Mutagénesis sitio dirigida del gen CusF

La mutagénesis dirigida tendrá el objetivo de insertar tres residuos de histidina en el N-terminal de CusF. Esto se realizará mediante el diseño de iniciadores que permitan la inserción de los codones codificantes para histidina. Mediante la técnica de PCR (reacción en cadena de la polimerasa) se amplificará el ADN empleando como molde CusF. Las amplificaciones serán purificadas y usadas posteriormente para la construcción de vectores. CusF se une de manera natural a Cu(I) y Ag(I) por interacciones catión- π con el residuo Trp44, en el mismo sitio de unión existe una coordinación con los residuos Met47, Met49 y His36.⁵⁵ CusF posee cinco histidinas, un triptófano y ningún residuo de cisteína, por lo que se cree que la unión de CusF específicamente a Cu(II) puede darse por medio de los residuos de histidina y metionina.⁵⁶ Los residuos His25, His26 y His27 se encuentran después del corte del péptido señal y antes del inicio de la estructura secundaria de CusF dejándolos disponibles para la interacción con metales divalentes como el Cu(II) o el Ni(II). Para aumentar la afinidad de CusF a Ni(II) se modificó este sitio agregando tres histidinas más en el N-terminal de CusF formando el mutante CusF3H+. El péptido maduro posee ocho residuos de histidinas, seis de ellos entre las posiciones 2 a 7. La inserción de estos nueve codones extras se realizó mediante mutagénesis sitio dirigida con reacción en cadena de la polimerasa para modificaciones en ADN plasmídico. Se diseñó el iniciador delantero agregando los tres codones extra y el sitio de restricción *NdeI*. Para

el iniciador reverso se agregó el sitio de corte *KpnI* con el fin de dejar el sitio de reconocimiento de corte de enteroquinasa del plásmido pET30a. En la tabla 1 se muestran los iniciadores empleados, es importante señalar que se agregaron los codones CAC para permitir el alineamiento apropiado del iniciador al gen haciendo distinción entre los codones originales que son CAT.

Iniciadores	
Delantero	5' – AGTCAGTCAC CATATG CACCACCACCATCATCATGAAACCATGAGCGAAGTAC – 3'
Reverso	5' – ATGCATGCAG GTACC CTGGCTGACTTTAATATCCTGTAA – 3'

Tabla 1. Iniciadores para amplificación de CusF para la mutagénesis dirigida

Para la amplificación se realizó a un volumen final de 50 µL en microtubos de PCR de 200 µL. La reacción 1x se muestra en la siguiente tabla 2:

Reacción de amplificación para CusF3H+	
ADN patrón	1 µL (10 ng)
Iniciador delantero	3 µL (60 pmoles)
Iniciador reverso	3 µL (60 pmoles)
dNTP's	1 µL (10 mM)
Vent DNA polimerasa	0.5 µL (2 unidades)
Buffer de reacción	5 µL (10x)
Agua mQ	36.5 µL
Volumen final	50 µL

Tabla 2. Composición de reacción 1x para la amplificación de CusF3H+

El producto de amplificación fue purificado empleando el Kit Gel Extraction de QIAGEN siguiendo indicaciones de la casa comercial.

4.3.2 Construcción de vectores con CusF3H+

El producto amplificación originado de la mutagénesis sitio-dirigida vía reacción en cadena de la polimerasa fue digerido con las enzimas de restricción *NdeI* y *KpnI* (New England Biolabs). El plásmido comercial pET30a(+) fue linearizado con las mismas enzimas y tratado con SAP después del corte para evitar la recirculización del mismo.

El nuevo gen, CusF3H+ fue ligado en el vector linearizado obteniendo el plásmido pET30a-CusF3H+ el cual fue seleccionado vía resistencia a kanamicina.

Se seleccionaron cinco diferentes genes de diferente origen y peso molecular para realizar las pruebas de expresión y solubilidad. Entre ellos se incluyó la proteína reportera GFP. Los genes de las proteínas de interés fueron amplificados con iniciadores específicos para introducir los sitios de corte *NcoI* y *XhoI*, en el caso de GFP se amplifico para incluir los sitios de *NcoI* y *XhoI* y mantener intacto el sitio de corte de enteroquinasa. Los genes fueron digeridos con las enzimas de restricción señaladas para cada caso, así mismo el vector pET30a-CusF3H+ fue linearizado con las enzimas necesarias para cada inserto agregando al final de cada reacción SAP para evitar la recirculización del mismo. Cada uno de los genes de interés digeridos fueron insertados dentro del vector linearizado obteniéndose cinco nuevos vectores. También se realizaron construcciones en el mismo vector y los mismos genes de interés, pero sin la proteína de fusión CusF dejando el sitio de etiqueta de histidina presente previamente en el vector comercial para fines comparativos. En la tabla 3 se muestran las construcciones realizadas, junto con su peso molecular. En la tabla 4 se señala el nombre completo y naturaleza de los genes seleccionados. La construcción de los vectores se realizó en la cepa especial de *E. coli* DH5 α empleando la técnica de transformación por choque térmico empleando células rubidio competentes.

CusF3H+	Peso (bp)	Peso (kDa)	His-tag	Peso (bp)	Peso (kDa)
GFP	981	36.7	GFP	714	26.0
G-CSF	789	32.0	G-CSF	522	22.0
IMB1	1442	53.3	IMB1	1175	43.4
PixE	1407	54.0	PixE	1140	43.1
AXR3	977	35.2	AXR3	780	25.3
PiF3	1847	66.9	PiF3	1580	57.0

Tabla 3. Construcciones obtenidas y sus pesos moleculares

Proteína	Nombre completo	Origen
GFP S65T	Green Fluorescent Protein	<i>Aequorea Victoria</i>
PixE	PatA subfamily protein	<i>Synechocystis sp</i>
AXR3	Auxin Response 3	<i>Arabidopsis thaliana</i>
PiF3	Phytochrome-interacting factor 3	<i>Arabidopsis thaliana</i>
IMB1	Imbibition-inducible 1	<i>Arabidopsis thaliana</i>

Tabla 4. Proteínas empleadas para pruebas comparativas junto con su origen

4.3.3 Expresión de proteínas de interés y GFP

Las construcciones obtenidas con la proteína de fusión CusF3H+ y la etiqueta de histidina fueron transformadas en la cepa de *E. coli* BL21(DE3) mediante choque térmico empleando células rubidio-competentes.

4.3.3.1 Expresión piloto de proteínas de interés y GFP

Se seleccionaron varias clonas resultado de la transformación de las construcciones obtenidas en la cepa de *E. coli* BL21(DE3). Para la inducción de la expresión piloto para cada una de ellas se inocularon las clonas seleccionadas en 2 mL de caldo LB (Luria Bertani) con kanamicina a una concentración final de 30 µg/mL. El cultivo se incubó a 37 °C a 200 rpm hasta un O.D. 600 de entre 0.4 – 0.6, al llegar a dicha densidad los cultivos se retiraban de la incubadora hasta alcanzar temperatura ambiente y fueron

inducidos con IPTG a una concentración final de 0.1 mM. Las células fueron incubadas nuevamente, pero ahora a 25 °C en agitación a 200 rpm durante 16 h. El sedimento celular fue recolectado mediante centrifugación a 13000 rpm por 1 minuto y se descartó el LB.

4.3.3.2 Fracción soluble e insoluble

Para la obtención de la fracción soluble de la expresión, las células recolectadas se resuspendieron en 120 µL de agua mQ y se agregó 40 µL de buffer de carga 6x. La mezcla fue llevada a ebullición durante 10 minutos para permitir la lisis celular y después centrifugada a 13000 rpm por 10 minutos. El sobrenadante fue recuperado y tratado como la fracción soluble de la expresión. El sedimento remanente fue lavado con agua mQ para eliminar cualquier remanente de fracción soluble que pudiera interferir. El sedimento fue resuspendido en un buffer de urea 8 M y hervido durante otros 10 minutos y después centrifugado por durante 10 minutos a 13000 rpm. El sobrenadante obtenido fue guardado y tratado como la fracción insoluble de la expresión. Ambas fracciones fueron visualizadas mediante electroforesis en geles de poliacrilamida (SDS-PAGE) en fase discontinua con el gel de resolución al 10% y el gel de apilamiento a 4%. Los geles fueron teñidos empleando con azul de Commassie.

4.3.3.3 Análisis de fracciones de expresión

Todas las expresiones obtenidas fueron analizadas con el programa ImageJ (versión 1.49) del National Institute of Health. El programa permite el análisis de las bandas de electroforesis de una manera semi-cuantitativa. Cada banda que aparece dentro del mismo carril es analizada por el programa arrojando para cada una un pico con un área proporcional a la intensidad mostrada en la electroforesis. El pico de cada área es sumado para obtener un cien por ciento y poder relacionar el área de cada pico obtenido con un porcentaje del total. De esta manera se obtiene un estimado de la representación

de la banda de expresión en relación con el total de proteínas obtenidas del lisado celular.

4.3.3.4 Escalamiento de la expresión para CusF_GFP, CusF3H+_GFP y His_GFP

Se escaló la expresión para la proteína reportera GFP marcada con las proteínas CusF, CusF3H+ y la etiqueta de histidina. Para esto se preparó un pre-inóculo en 5 mL de LB con kanamicina (30 µg/mL) el cual se incubó a 37 °C a 200 rpm durante 16 h. Se prepararon matraces triptinizados de 500 mL con 125 mL de LB con kanamicina (30 µg/mL) y se inocularon con 125 uL del pre-inoculo. El cultivo se incubó a 37 °C a 220 rpm hasta que se alcanzó un O.D.600 de entre 0.4 – 0.6. Una vez alcanzado el O.D.600 necesario, los matraces se retiraban de la incubadora hasta alcanzar temperatura ambiente y se agregó IPTG a una concentración de 0.1 mM, los matraces se incubaron nuevamente en agitación a 220 rpm pero a una temperatura de 25 °C durante 16 h.

4.3.4 Purificación de GFP marcada con las diferentes proteínas de fusión

4.3.4.1 Purificación de CusF_GFP

Con el fin de entender el comportamiento de CusF original con Ni(II) se realizó primero la purificación de GFP marcada con CusF. La purificación de GFP marcada con CusF, CusF3H+ y la etiqueta de histidina se realizó con el FPLC ÄKTA prime plus (GE Healthcare). La expresión de CusF_GFP se escaló a un total de 2 L, el sedimento celular se recolectó mediante centrifugación a 12000 rpm durante 15 minutos a 4 °C. El sedimento obtenido se resuspendió en la solución amortiguadora de lisis 50 mM Tris, 500 mM NaCl y se lisó utilizando perlas de vidrio y una cámara de lisado apropiada en ciclos de 15 segundos con 30 segundos de descanso entre cada pulso. Las células fueron lisadas mediante fricción y el contenido celular es liberado a la matriz. Para separar el contenido citoplásmico del resto celular y las perlas de vidrio, se centrifugó

durante 20 minutos a 15000 rpm a 4 °C. El contenido citoplásmico fue separado y se agregaron 10 µL de coctel de inhibidor de proteasas. El lisado clarificado se cargó en una columna HisTrap FF de 1 mL previamente funcionalizada con Ni(II) y equilibrada con la solución amortiguadora 50 mM Tris, 500 mM NaCl pH 8 (la misma solución amortiguadora de lisis). Una vez cargado el clarificado, la columna se lavó con 10 volúmenes (10 mL) empleando la misma solución amortiguadora de equilibrio y lisis y la proteína unida fue eluida con la solución amortiguadora de elución (50 mM Tris, 500 mM NaCl, 50 mM de imidazol) en fracciones de 0.5 mL, se mantuvo una presión de 0.6 MPa (según las especificaciones de la columna) y un flujo de 0.5 mL/min. Las fracciones donde se observaba proteína fueron recolectadas y analizadas por SDS-PAGE.

4.3.4.2 Purificación de CusF3H+_GFP

Las condiciones de expresión para GFP marcada con CusF3H+ fueron iguales a las usadas con su contraparte original. El sedimento celular se resuspendió en una solución amortiguadora con baja concentración de imidazol (50 mM Tris, 500 mM NaCl, 5 mM imidazol) y se lisó con perlas de vidrio en ciclos de 15 segundos con 30 segundos de descanso entre cada pulso de rotación. El clarificado obtenido fue cargado en una columna HisTrap FF de 1mL, se pasaron 10 volúmenes de solución amortiguadora de lavado (50 mM Tris, 500 mM NaCl, 10 mM imidazol) con el fin de disminuir la unión de proteínas inespecíficas. La proteína fue eluida con la solución amortiguadora de elución (50 mM Tris, 500 mM NaCl, 250 mM de imidazol) en gradiente de 20 volúmenes y se recolectaron fracciones de 0.5 mL. La purificación se realizó a 0.6 MPa y a un flujo constante de 0.5 mL/min. Las fracciones obtenidas con proteína presente visible fueron analizadas mediante SDS-PAGE.

4.3.4.3 Purificación de His_GFP

Para fines comparativos se realizó una expresión y purificación de His_GFP bajo las mismas condiciones que CusF_GFP y CusF3H+_GFP. El sedimento celular obtenido de 1 L de cultivo fue resuspendido en solución amortiguadora de equilibrio con imidazol (50 mM Tris, 500 mM NaCl, 5 mM imidazol) y lisado con perlas de vidrio. Al clarificado obtenido de la lisis se le agregó 10 µL de un cóctel de proteasas para evitar la degradación de la proteína a purificar. El clarificado se cargó en una columna HisTrap FF de 1 mL previamente funcionalizada con Ni(II). Una vez pasado todo el producto de clarificado, se lavó la columna con 10 volúmenes (10 mL) con la solución amortiguadora de lavado (50 mM Tris, 500 mM NaCl y 10 mM imidazol). La elución se realizó en gradiente de 20 volúmenes empleando la solución amortiguadora de elución (50 mM Tris, 500 mM NaCl, 250 mM imidazol) y se recolectaron fracciones de 0.5 mL. La presión se mantuvo durante toda la corrida en 0.6 MPa de acuerdo a las especificaciones de la columna y con un flujo bajo de 0.5 mL/min. Las fracciones de elución obtenida con proteína visible fueron analizadas mediante SDS-PAGE. En los tres casos, la cuantificación de proteína obtenida se realizó mediante Bradford.

4.3.5 Remoción de la proteína de fusión CusF3H+_GFP con enteroquinasa

Se recolectaron las fracciones con mayor pureza y se realizó una cuantificación con Bradford. Se preparó una alícuota con 200 µg de proteína la cual fue cortada con 12.5 U de enteroquinasa (la cantidad necesaria se calculó en base a las especificaciones de la enzima dada por NEB). La reacción se incubó 16 h a temperatura ambiente y se analizó en SDS-PAGE para corroborar el completo corte de la proteína. Una vez confirmado, se separó la proteína CusF del resto mediante una segunda cromatografía de afinidad empleando la resina Profinity IMAC (BioRad) cargada con Ni(II). La reacción de corte se incubó con la resina durante 1 h a 4 °C. La mezcla se centrifugó a 5000 rpm

por un minuto, la proteína de interés se recuperó en el sobrenadante. La resina se resuspendió en una solución amortiguadora de elución con 250 mM de imidazol y se volvió a centrifugar a 5000 rpm por un minuto para separar la proteína de fusión. Todas las fracciones obtenidas fueron analizadas por SDS-PAGE en fase discontinua, empleando un gel de resolución al 15% y uno de apilamiento al 4%.

4.4 Desarrollo de los sistemas CusF3H+/C. *glutamicum* y SmbP/C. *glutamicum* para la expresión, secreción y purificación de proteínas recombinantes heterólogas

4.4.1 Diseño de plásmidos para secreción de proteínas heterólogas en C. *glutamicum*

Los plásmidos para la secreción de proteínas expresadas en *Corynebacterium glutamicum* se diseñaron a partir del plásmido pz8-tac (Addgene plasmid #74062). El pz8-tac es un tipo de plásmido conocido como *shuttle vector*, este tipo de plásmidos se pueden usar en dos o más organismos diferentes. De manera que la construcción de los plásmidos derivados se realizó en la cepa de *E. coli* DH5 α . El plásmido pz8-tac tiene el replicón de pCG1 (plásmido de origen natural de *C. glutamicum*), además está bajo el promotor ptac permitiendo la inducción mediante IPTG.

Para secretar una proteína por medio del sistema de Tat (Twin arginine translocation) de *C. glutamicum* es necesario marcar las proteínas con un péptido señal que sea reconocido por esta vía. En este caso se seleccionaron dos péptidos señal, TorA de *E. coli* y Cg_0949 de *C. glutamicum*. Ambos péptidos señal han demostrado previamente ser reconocidos por el sistema Tat de *C. glutamicum* y secretar las proteínas marcadas. Las secuencias de aminoácidos de ambos péptidos señal se presentan en la figura 1 a continuación:

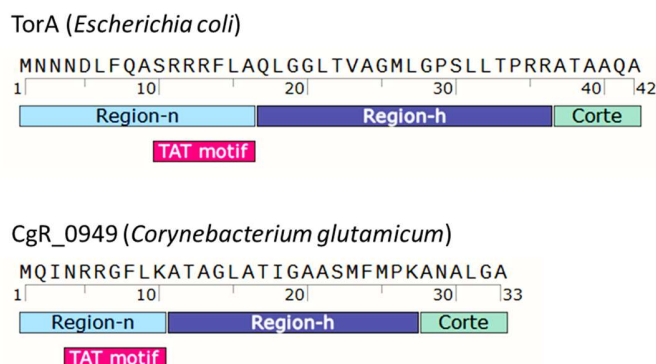


Figura 1. Secuencias de TorA y CgR_0949

Los dos péptidos señal fueron fusionados con las proteínas de fusión en estudio: CusF3H+ y SmbP con el fin de desarrollar dos versiones de cada una. Para CusF3H+ se generaron TorA-CusF3H+ y CgR-CusF3H+, para SmbP fueron TorA-SmbP y CgR-SmbP como se muestra en la figura 2, en la tabla 5 se presentan las secuencias de aminoácidos para cada una de estas proteínas de fusión.



Figura 2. Proteínas de fusión SmbP y CgR marcadas con TorA y CgR

	Péptido señal	Proteína de fusión
TorA-CusF3H+	MNNNDLFQASRRRFLAQL GGLTVAGMLGPSLLTPRRA TAAQA	MHHHHHHHETMSEAQPVISATGVVKGIDLESKITIHHDPAAVNWPE MTMRFTITPQTKMSEIKTGDKVAFNFVQQGNLSLLQDIKVSQGS
TorA-SmbP	MNNNDLFQASRRRFLAQL GGLTVAGMLGPSLLTPRRA TAAQA	SGHTAHVDEAVKHAEEVAHGKEGHTDQLLEHAKESLTHAKAASEA GGNTHVGHGIKHLEDAIKHGEEGHVGVATKHAQEAIHLRASEHK
CgR-CusF3H+	MQINRRGFLKATAGLATIG AASMFMPKANALGA	MHHHHHHHETMSEAQPVISATGVVKGIDLESKITIHHDPAAVNWPE MTMRFTITPQTKMSEIKTGDKVAFNFVQQGNLSLLQDIKVSQGS
CgR-SmbP	MQINRRGFLKATAGLATIG AASMFMPKANALGA	SGHTAHVDEAVKHAEEVAHGKEGHTDQLLEHAKESLTHAKAASEA GGNTHVGHGIKHLEDAIKHGEEGHVGVATKHAQEAIHLRASEHK

Tabla 5. Secuencias de aminoácidos para cada una de las proteínas de fusión marcadas

El plásmido pz8-ptac fue modificado para permitir la inserción de los genes deseados. El sitio de clonación múltiple (MCS) de pz8-ptac era bastante limitado, por lo que se diseñó y añadió un nuevo sitio de clonación múltiple entre los sitios *EcoRI* y *Sall* que ya

contenía el vector. El nuevo sitio se diseñó para incluir los sitios *NdeI*, *KpnI*, *BamHI* y *NotI* entre los sitios mencionados, además se agregó un sitio de reconocimiento para enteroquinasa entre *NdeI* y *KpnI*. La proteína de fusión quedaría entre los sitios de *EcoRI-KpnI* o *NdeI-KpnI*, mientras que la proteína de interés quedaría entre *BamHI-NotI* o *BamHI-Sall*. El diseño del sitio de clonación múltiple para *pz8* se muestra en la figura 3.

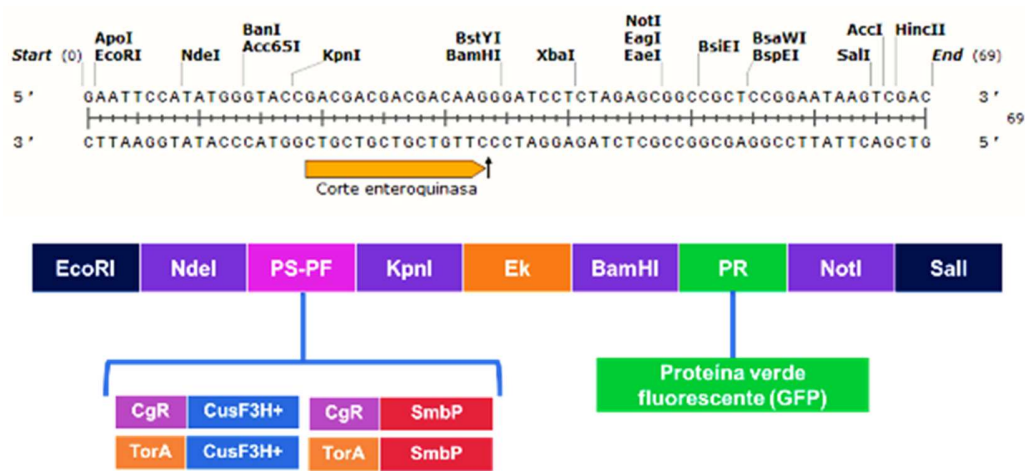


Figura 3. Diseño del nuevo sitio de clonación múltiple y localización de la proteína de fusión y la proteína reportera

Se diseñaron cuatro genes de cada una de las combinaciones de péptidos señal y proteínas de fusión con sus respectivos nuevos sitios de restricción y reconocimiento para enteroquinasa. Cada uno de los genes fueron optimizados para su expresión en *C. glutamicum* de acuerdo al algoritmo de GeneScript donde fueron sintetizados. En la figura 4a) se muestran los plásmidos derivados de *pz8-ptac* con las proteínas de fusión marcadas con el péptido señal *TorA*, mientras que los plásmidos generados con el péptido señal *CgR_0949* se presentan en la figura 4b).

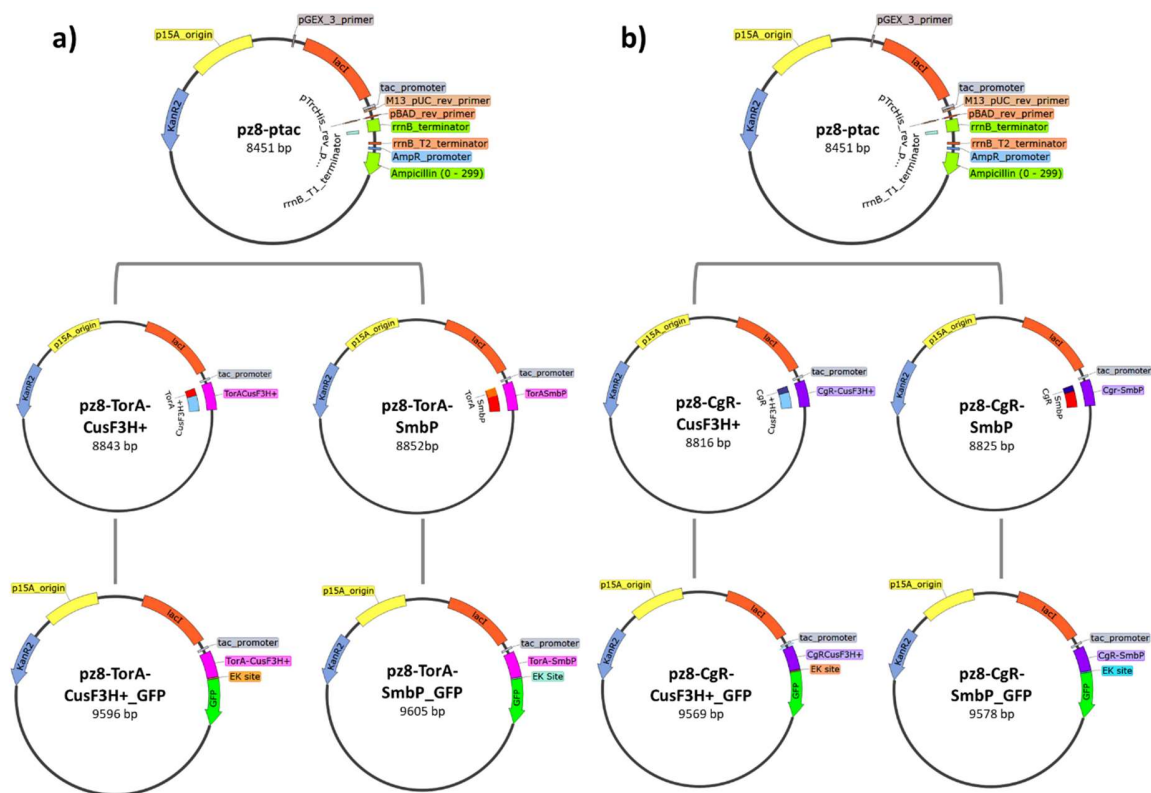


Figura 4. a) Diseño de plásmidos para la serie de TorA a partir del pz8-ptac, b) Diseño de plásmidos para la serie de CgR a partir del plásmido pz8-ptac

4.4.2 Construcción de plásmidos derivados de pz8-ptac

Los plásmidos de origen pUC57 con los cuatro genes diseñados fueron transformados en *E. coli* DH5 α para propagación del ADN plasmídico. El ADN plasmídico obtenido fue digerido con las enzimas *EcoRI* y *Sall*, el producto de digestión fue corrido en un gel de agarosa para recuperar los insertos y purificarlos con el kit QIAquick Gel Extraction de QIAGEN. El plásmido pz8-ptac fue linearizado empleando también las enzimas *EcoRI* y *Sall*, agregando al final SAP para evitar la re-circulización del plásmido. La ligación se realizó con la T4 ligasa de BioBasic empleando las indicaciones de la casa comercial y una relación de 1:10 y 1:20 (plásmido:inserto). El producto de ligación fue transformado en DH5 α y el cribado en el caso de las construcciones de CusF3H+ se realizó mediante PCR. La reacción de amplificación para los insertos de TorA-CusF3H+ y CgR-CusF3H+ se llevó a cabo en 50 μ L de reacción, la composición 1 x se muestra

en la tabla 6, los iniciadores empleados para la reacción fueron los mismos mostrados en la tabla 2.

Reacción de amplificación para CusF3H+	
ADN patrón	1 µL (10 ng)
Iniciador delantero	3 µL (60 pmoles)
Iniciador reverso	3 µL (60 pmoles)
dNTP's	1 µL (10 mM)
Vent DNA polimerasa	0.5 µL (2 unidades)
Buffer de reacción	5 µL (10x)
Agua mQ	36.5 µL
Volumen final	50 µL

Tabla 6. Composición 1 x para la reacción de amplificación de CusF3H+

El producto de amplificación fue analizado mediante un gel de agarosa y teñido con bromuro de etidio. Para el cribado de las clonas con el inserto TorA-SmbP y CgR-SmbP se analizó el ADN plasmídico mediante restricción empleando las enzimas *EcoRI* y *Sall*. Para dicha reacción se cortaron 2µg de cada plásmido y se incubó durante 2 h a 37 °C. El resultado de la digestión se cargó en un gel de agarosa al 0.7% para su posterior análisis.

Los cuatro plásmidos resultantes fueron linearizados con las enzimas *BamHI* y *NotI* y se agregó SAP para evitar la re-circulización del plásmido. Los sitios de restricción *BamHI* y *NotI* se agregaron al gen de GFP mediante amplificación empleando iniciadores diseñados con dichos sitios, en la tabla 7 se muestran los iniciadores empleados.

Iniciadores	
Delantero	5' – AGTCAGTCAGGATCCATGGTGAGCAAGGGCGAGG – 3'
Reverso	5' – ATGCATGCAGCGGCCGCTTGTACAGCTCGTCCATG – 3'

Tabla 7. Iniciadores empleados para la amplificación de GFP

La reacción de amplificación se realizó a un volumen total de 50 µL, el contenido de la reacción a 1x se muestra en la tabla 8.

Reacción de amplificación para GFP	
ADN patrón	1 μ L (10 ng)
Iniciador delantero	3 μ L (60 pmoles)
Iniciador reverso	3 μ L (60 pmoles)
dNTP's	1 μ L (10 mM)
Vent DNA polimerasa	0.5 μ L (2 unidades)
Buffer de reacción	5 μ L (10x)
Agua mQ	36.5 μ L
Volumen final	50 μ L

Tabla 8. Composición 1x para amplificación del gen de GFP

El producto de amplificación fue purificado con el kit QIAquick gel extraction de QIAGEN y digerido empleando las enzimas *BamHI* y *NotI*. La reacción de ligación entre el inserto (GFP) y los cuatro plásmidos linearizados se realizó a una proporción de vector:inserto de 1:10 y 1:20. El producto de ligación se transformó mediante choque térmico empleando células rubidio-competentes de *E. coli* DH5 α . El cribado de las clonas transformadas se realizó mediante amplificación del inserto a partir del ADN plasmídico extraído de las mismas empleando las mismas condiciones usadas para la obtención del producto.

4.4.3 Expresión y secreción de proteínas en *C. glutamicum*

4.4.3.1 Preparación de células competentes para electroporación de *C. glutamicum*

Para la transformación de *C. glutamicum* se preparó un preinóculo de una colonia aislada e inoculada en 4 mL de LB durante 16 h a 30 C°. Se inocularon dos matraces triptinizados con 100 mL de medio MB (5g de extracto de levadura, 15g de triptona, 5g de hidrolizado de soya, 5g de NaCl y 35g de glicina por litro) y se incubaron a 30 °C en 200 rpm de agitación hasta un O.D.600 de entre 0.2 y 0.25. Una vez alcanzada la densidad deseada se suplemento el medio con ampicilina a 100 ug/mL de concentración final y se volvió a incubar por otra hora y media más. Las células fueron recuperadas por centrifugación a 5000 rpm por 10 minutos a 4 °C y después resuspendidas en 30 mL de una solución al 5% de glicerol a un pH 7.2 repitiéndose la centrifugación, se repitió

el proceso dos veces más. El sedimento final fue resuspendido en 1.5 mL de solución al 15% de glicerol a un pH de 7.2 y se transfirieron en alícuotas de 150 µL. Las células se mantuvieron en hielo para seguir con el protocolo de transformación.

4.4.3.2 Transformación de *C. glutamicum*

A una alícuota de 150 µL de células competentes se agregaron 50 ng de cada uno de los plásmidos de pz8-ptac obtenidos. La mezcla se incubó en hielo durante 5 minutos y se transfirió a celdas para electroporación de 2mm de previamente enfriadas en hielo. Para la electroporación se empleó el electroporador Eppendorf Eporator, las celdas se colocaron en el electroporador a 2.5 kV, 200 Ω, 25 µF y se inició el pulso. Directamente a la celda se agregaron 200 µL de LB y se transfirió a un microtubo de 1.5 mL y se incubó durante 2 h a 30 °C. El contenido se pasó a placas de LB adicionadas con kanamicina a concentración de 25 µg/mL y se incubó durante 48 h a 30 °C.

4.4.3.3 Expresión de proteínas en *C. glutamicum*

La expresión de las dos diferentes proteínas marcadas con los diferentes péptidos señales se realizó empleando un medio definido modificado de Jae Woong⁵⁷ el cual contiene 20 g/L de dextrosa por litro, 3 g de K₂HPO₄, 1 g de KH₂PO₄, 2 g de urea, 2 g de MgSO₄, 1 mg de ZnSO₄, 200 µg de CuSO₄, 10 mg de CaCl₂, 2 g de levadura y 7 g de hidrolizado de caseína por litro. En matraces triptinizados de 500 mL se inocularon en 100 mL del medio definido adicionado con kanamicina a una concentración de 25 µg/mL y se incubaron a 200 rpm a 30 °C. El cultivo se incubó hasta un O.D.600 de entre 0.8 – 1.0 y se agregó IPTG para la inducción a una concentración final de 1 mM. Después de la inducción se incubó durante 48 h a 30 °C.

4.4.3.4 Purificación de proteínas secretadas por *C. glutamicum*

Para la purificación de CusF3H+_GFP y SmbP-GFP marcadas con TorA y CgR, se recolectaron 600 mL de medio de cultivo de una expresión de 48 h a 30 °C a 200 rpm de agitación. El cultivo se centrifugó a 10000 rpm durante 10 minutos a 4 °C para separar el sedimento del medio de cultivo. El medio recuperado se cargó en una columna HisTrapFF de 5 mL cargada con Ni(II) para GFP marcada con TorA-CusF3H+, CgR-CusF3H+ y CgR-SmbP, y con Cu(II) para GFP marcada con TorA-SmbP. La columna se lavó con 10 volúmenes de solución amortiguadora de lavado (50 mM Tris, 500 mM NaCl). Finalmente, la proteína fue eluida con 10 volúmenes de solución amortiguadora de elución (50 mM Tris, 500 mM NaCl, 500 mM imidazol) en un solo paso. La purificación se realizó bajo 0.6 mPa de presión con un flujo constante de 1 mL/min. Las fracciones obtenidas fueron analizadas mediante gel de poliacrilamida SDS-PAGE.

4.5 Diseño de sistema de expresión de proteínas recombinantes libre de plásmidos usando la técnica CRISPR/Cas9

Para la modificación a nivel genómico del organismo hospedero se empleará una técnica de edición asistida con CRISPR/Cas9 diseñada por Yang y col. de doble plásmido. Dicho sistema consiste en la transformación de la cepa de *E. coli* con el plásmido pCassac* que incluye el gen de la proteína Cas9 responsable del corte y un segundo plásmido que incluye el sgARN que permite guiar la enzima y el ADN donador junto con los brazos de homología necesarios para permitir la recombinación homóloga y la posterior reparación.⁵³

4.5.1 Diseño de pTCueo y pTYahl

Los genes *cueO* y *yahl* fueron seleccionados como flancos para la edición genómica al ser no esenciales para *E. coli* de acuerdo a la base de datos: Pec Profiling of *E. coli* Chromosome Ver.4 del Japan National Institute of Genetics

(<https://shigen.nig.ac.jp/ecoli/pec/index.jsp>). El gen *cueO* (Gene ID: 947736, 137083..138633) codifica para la proteína laccasa (multicopper oxidase) y tiene un tamaño de 1551 bp. Mientras que *yahl* es un gen de 951 bp (Gene ID: 944984, 340165..341115) que codifica para la proteína “parecida” a carbamato quinasa (carbamate kinase-like protein). Ambos genes fueron analizados con el algoritmo del programa en línea Benchling (www.benchling.com) arrojando 223 y 135 PAM para *cueO* y *yahl* respectivamente. La elección del N20 para cada uno fueron en cada caso aquellos que presentaron la mayor puntuación en sitio y fuera de sitio, dicha puntuación fue también calculada con el programa en línea Benchling, mismo que se basa en la investigación de Doench, Fusi y col., en el 2016.⁵⁸ En la siguiente tabla se muestran los sitios seleccionados junto con sus puntuaciones, localización y PAM.

Gen	Secuencia	PAM	Puntuación “en sitio”	Puntuación “fuera de sitio”	Localización
cueO	TTAATGCTGCCAAAACAGTG	GGG	75.3*	99.3	137590 – 137609
yahl	TGTGTGGCAGATACGCAGGG	CGG	85.4*	99.8	340405 – 340424

Tabla 9. Genes seleccionados para corte con Cas9

*Calculado en base a la investigación de Doench, Fusi y col., en el 2016

Los genes *cueO* y *yahl* se reemplazarán en su totalidad por el ADN donador, el cual consiste en el gen de la proteína reportera GFP fusionada al gen de la proteína de fusión SmbP bajo la inducción y control del promotor, 5’-UTR y 3’-UTR del gen de la proteína de choque frío *cspA*. Permitiendo a inducción de la proteína al disminuir la temperatura. Para hacer el reemplazo de los genes se debe adicionar en 5’ y 3’ un brazo homólogo fuera de los genes de seleccionados que permitan realizar la recombinación homóloga como respuesta de reparación al corte de Cas9. La figura 5 muestra el proceso de edición genómica empleando el sistema CRISPR/Cas9 en conjunto con la recombinación homóloga.

* En la investigación de Yang y col del 2016 se emplea el plásmido pCas9, sin embargo, el mismo Dr. Sheng Yang sugirió el uso de un segundo plásmido que recién habían desarrollado, el pCassac.

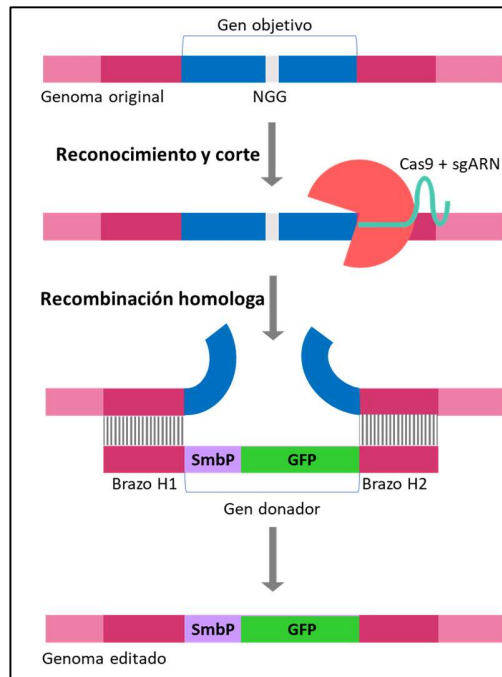


Figura 5. Edición genómica empleando CRISPR/Cas9 facilitada por recombinação homóloga

En la figura 6 se muestra el gen donador junto con los brazos de homología en 5' y 3' así mismo en 5' del brazo de homología 1 se agregó el N20 y el sgARN que permitirán guiar a Cas9 al sitio de corte. Este gen tiene los sitios de cortes XbaI en 5' y Sall en 3' para permitir su introducción dentro del plásmido pTarget. Yang y col., en el 2016 demostraron que la tasa de eficiencia de edición aumentaba cuando el ADN donador se encontraba dentro del mismo plásmido que lleva el N20 y el sgARN.⁵³ Debido al tamaño del gen donador los brazos de homología consistieron de 350 bp como sugerencia del equipo de trabajo del profesor Sheng Yang. En la figura 7 se muestran los plásmidos diseñados derivados del plásmido pTarget.

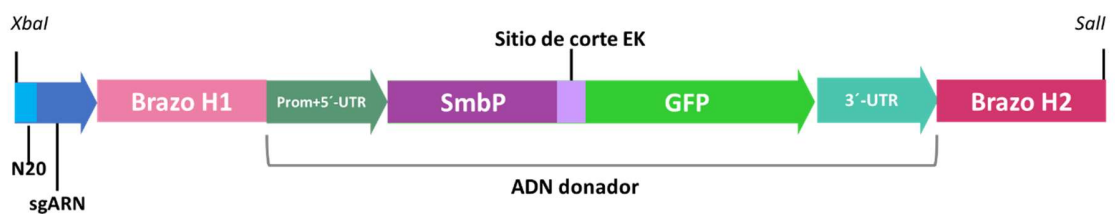


Figura 6. Inserto diseñado para los plásmidos derivados del pTarget

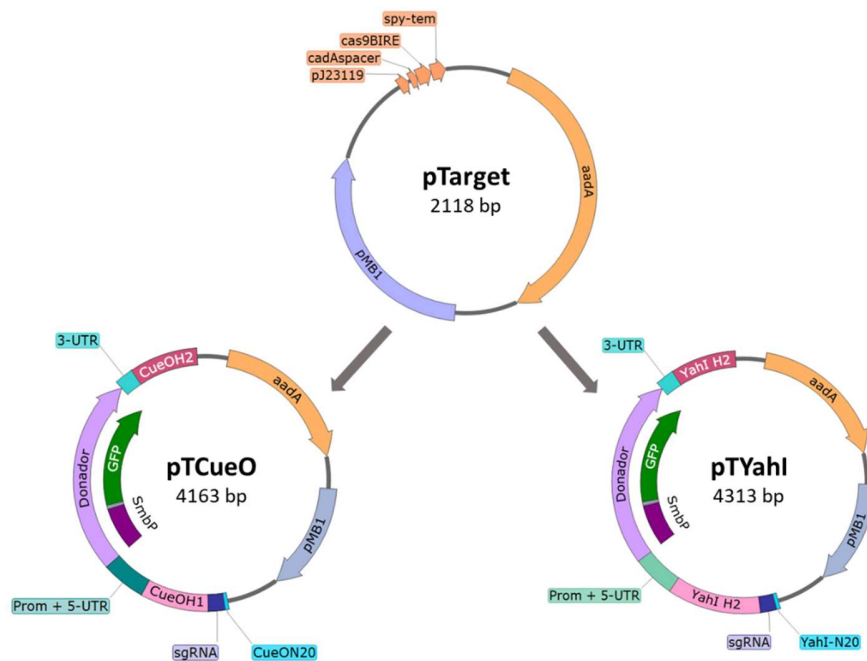


Figura 7. pTCueO y pTYahl, plásmidos derivados de pTarget

4.5.2 Construcción de plásmidos pTCueo y pTYahl

Para la construcción de los plásmidos con el sgARN, el N20 y el ADN donador se siguió la técnica de alineamiento por PCR. Para ello se diseñaron iniciadores especiales para cada una de las partes, cada uno de estos contenía un excedente de 20 bp en el 5' que coincidían con el fragmento inmediato rio arriba para permitir su alineamiento. El iniciador delantero para el fragmento N20+sgARN posee la secuencia del N20 y el sitio de corte *XbaI* para su posterior ligación, el ultimo iniciador reverso que corresponde al fragmento de brazo homologo 2 posee la secuencia de corte para *Sall*. Los iniciadores diseñados se muestran en la tabla 10:

Fragmento		Secuencia
Donador	Delantero	5'- ATTATTAAGGTAATACACTATGTCCGGTCACACCGCAC - 3'
	Reverso	5'- TCTGTGCTTTTAAGCAGAGATTACTTGTACAGCTCGTCCATGC - 3'
Serie CueO		
CueoN20+sgARN	Delantero	5'- TGCTCTAGATTAATGCTGCCAAAACAGTGGTTTTAGAGCTAGAAATAGCAAGTTAAAAT - 3'
	Reverso	5'- CGTCGCGCACTGGTCTTTTTCTCAAAAAAGCACCGACTCG - 3'
CueOH1	Delantero	5'- GAGTCGGTGCTTTTTTTGAGAAAAAGACCAGTGC CGA - 3'
	Reverso	5'- CATATTTATGATTAATCGGTAGTTATTTCTTATTCTTAAGCAGGCG - 3'
Prom-CspA-CueO	Delantero	5'- TTAAGAATAAGGAAATAACTACCGATTAATCATAAATAGAAAAATAATTGTTGC - 3'
	Reverso	5'- TGTGCGGTGTGACCGGACATAGTGATTACCTTTAATAATTAAGTGTGCCT - 3'
Term-CspA-CueO	Delantero	5'- TGGACGAGCTGTACAAGTAATCTCTGCTTAAAAGCACAGAATCTAA - 3'
	Reverso	5'- TATTTCCGAATACGGTCTTTAGATTTTATTACGCGATCGATTATTTATTTC - 3'
CueOH2	Delantero	5'- TCGATCGCGTAATAAAATCTAAAGACCGTATTCGGAATATGCC - 3'
	Reverso	5'- ACGCGT CGAC GTTTACCAGCGGGTGGTC - 3'
Serie Yahl		
YahlN20+sgARN	Delantero	5'- TGCTCTAGATGTGTGGCAGATACGCAGGGTTTTAGAGCTAGAAATAGCAAGTTAAAAT - 3'
	Reverso	5'- CTGATGCGCTACGCTTATCACTCAAAAAAGCACCGACTCG - 3'
Yahl H1	Delantero	5'- GAGTCGGTGCTTTTTTTGAGTGATAAGCGTAGCGCATCAG - 3'
	Reverso	5'- CATATTTATGATTAATCGGTGTTGTTATGTCCGCTACAGAGG - 3'
Prom-CspA-Yahl	Delantero	5'- TCTGTAGCGGACATAACAACACCGATTAATCATAAATAGAAAAATAATTGTTGC - 3'
	Reverso	5'- TGTGCGGTGTGACCGGACATAGTGATTACCTTTAATAATTAAGTGTGCCT - 3'
Term-CspA-Yahl	Delantero	5'- TGGACGAGCTGTACAAGTAATCTCTGCTTAAAAGCACAGAATCTAA - 3'
	Reverso	5'- CTTTCTTTCATCTTACGTCCAGATTTTATTACGCGATCGATTATTTATTTC - 3'
Yahl H2	Delantero	5'- GAGTCGGTGCTTTTTTTGAGTGGCGGGCATTGTGCATC - 3'
	Reverso	5'- ACGCGT CGAC AGCATCAGCTTACCGCC - 3'

Tabla 10. Iniciadores diseñados para la PCR de alineamiento para la construcción del inserto

Los fragmentos se amplificaron con la polimerasa KOD (Toyobo). La composición de la reacción se muestra en la tabla 11. La amplificación se realizó con un primer paso de 95 °C por 10 minutos, seguido de 30 ciclos de 98 °C por 10 segundos, 57 °C por 45 segundos y 68 °C por 30 segundos, con un paso final de 68 °C por 10 minutos. Para la amplificación de los fragmentos con el N20 y el sgARN se empleó como patrón el

plásmido pCB003 amablemente proporcionado por el equipo de trabajo del profesor Yang, ya que este poseía la secuencia del sgARN (plásmido actualizado del pCB001). Para el inserto se empleó el plásmido pZ8-Tora-SmbP-GFP obtenido de este mismo proyecto. Para el resto de los fragmentos, el ADN patrón empleado para la amplificación fue ADN genómico de *E. coli* BL21(DE3) extraído con el kit AxyPrep Bacterial Genomic DNA. La reacción se realizó en el termociclador Thermo Professional Trio de Biometra.

Patrón	Plásmido	ADN genómico*
ADN patrón	1 µL (10 ng)	1 µL (100 ng)
Iniciador delantero	1.5 µL (30 pmoles)	1.5 µL (30 pmoles)
Iniciador reverso	1.5 µL (30 pmoles)	1.5 µL (30 pmoles)
dNTP's	5 µL (0.2 mM)	5 µL (0.2 mM)
KOD DNA polimerasa	1 µL (1 U)	1 µL (1 U)
Buffer de reacción	5 µL (10x)	5 µL (10x)
MgSO ₄	3 µL (1.5 mM)	3 µL (1.5 mM)
Agua mQ	32 µL	32 µL
Volumen final	50 µL	50 µL

Tabla 11. Constitución 1 x para amplificación de fragmentos para los insertos

Los fragmentos fueron purificados por gel de agarosa empleando el kit de Axygen AxyPrep DNA Gel Extraction kit. A partir de los pesos moleculares de cada fragmento se realizó una correlación molar tomando en cuenta al fragmento con menor peso molecular como el limitante, siendo en ambos casos el fragmento de N20+sgARN con 126 bp. Las tablas 12 y 13 muestran los pesos moleculares de cada fragmento y los moles necesarios por cada fragmento como ADN patrón para el alineamiento.

Fragmento	Peso molecular	Relación molar
N20+sgARN-CueO	126	1
CueOH1	390	3.1
Prom/CspA-CueO	267	2.1
Donador	1069	8.5
3'-UTR/CspA-CueO	138	1
CueOH2	390	3

Tabla 12. Pesos moleculares y relación molar de fragmentos para el inserto de pTCueO

Fragmento	Peso molecular	Relación molar
N20+sgARN-Yahl	126	1
YahlH1	540	4.2
Prom/CspA-Yahl	267	2.1
Donador	1069	8.5
3'-UTR/CspA-Yahl	138	1
YahlH2	390	3

Tabla 13. Pesos moleculares y relación molar de fragmentos para el inserto de pTYahl

Para ambas construcciones se realizaron tres diferentes alineaciones, la primera con las tres primeras partes, la segunda con las tres partes restantes y una tercera con los seis fragmentos. La cantidad de ADN patrón para cada una de las reacciones dependió de la cantidad requerida en base al fragmento limitante, la constitución 1 x de cada una de las reacciones para los fragmentos pTCueO y pTYahl se muestran en las tablas 14 y 15 respectivamente.

Reactivo		Reacción N20 – Prom/CspA-CueO	Reacción Donador – CueOH2	Reacción N20 – CueOH2
ADN patrón	N20+sgARN-CueO	1 mol	-	1 mol
	CueOH1	3.1 mol	-	3.1 mol
	Prom/CspA-CueO	2.1 mol	-	2.1 mol
	Donador	-	8.5 mol	8.5 mol
	3'-UTR/CspA-CueO	-	1 mol	1 mol
	CueOH2	-	3 mol	3 mol
Iniciador delantero		1.5 µL (30 pmoles)	1.5 µL (30 pmoles)	1.5 µL (30 pmoles)
Iniciador reverso		1.5 µL (30 pmoles)	1.5 µL (30 pmoles)	1.5 µL (30 pmoles)
dNTP's		5 µL (0.2 mM)	5 µL (0.2 mM)	5 µL (0.2 mM)
KOD DNA polimerasa		1 µL (1 U)	1 µL (1 U)	1 µL (1 U)
Buffer de reacción		5 µL (10x)	5 µL (10x)	5 µL (10x)
MgSO4		3 µL (1.5 mM)	3 µL (1.5 mM)	3 µL (1.5 mM)
Agua mQ		x µL	x µL	x µL
Volumen final		50 µL	50 µL	50 µL

Tabla 14. Constitución 1 x para PCR de alineamiento para el inserto de pTCueO

Reactivo		Reacción	Reacción	Reacción
		N20 – Prom/CspA-Yahl	Donador – YahlH2	N20 – YahlH2
ADN patrón	N20+sgARN-Yahl	1 mol	-	1 mol
	YahlH1	4.2 mol	-	4.2 mol
	Prom/CspA-Yahl	2.1 mol	-	2.1 mol
	Donador	-	8.5 mol	8.5 mol
	3'-UTR/CspA-Yahl	-	1 mol	1 mol
	YahlH2	-	3 mol	3 mol
Iniciador delantero		1.5 µL (30 pmoles)	1.5 µL (30 pmoles)	1.5 µL (30 pmoles)
Iniciador reverso		1.5 µL (30 pmoles)	1.5 µL (30 pmoles)	1.5 µL (30 pmoles)
dNTP's		5 µL (0.2 mM)	5 µL (0.2 mM)	5 µL (0.2 mM)
KOD DNA polimerasa		1 µL (1 U)	1 µL (1 U)	1 µL (1 U)
Buffer de reacción		5 µL (10x)	5 µL (10x)	5 µL (10x)
MgSO4		3 µL (1.5 mM)	3 µL (1.5 mM)	3 µL (1.5 mM)
Agua mQ		x µL	x µL	x µL
Volumen final		50 µL	50 µL	50 µL

Tabla 15. Constitución 1 x para PCR de alineamiento para el inserto de pTYahl

La programación de la amplificación consistió en un primer paso de 95 °C por 10 minutos, seguido de 5 ciclos de 98 °C por 10 segundos, 57 °C por 45 segundos y 68 °C por 30 segundos, con un paso final de 68 °C por 10 minutos, estos 5 primeros ciclos se llevaron a cabo sin iniciadores para inducir el alineamiento de los fragmentos. Al término de los primeros ciclos se agregaron los iniciadores y se repitió el programa, pero con otros 20 ciclos. Los productos de amplificación se analizaron mediante electroforesis con gel de agarosa. En el caso de los fragmentos Yahl se observó baja especificidad y calidad en las amplificaciones. Por lo que se decidió disminuir la temperatura de alineamiento de 57 °C a 54 °C. Se repitieron las tres amplificaciones para el inserto de Yahl, sin embargo, para mejorar la calidad del inserto se emplearon los dos productos de las amplificaciones de solo tres fragmentos teniendo solo dos fragmentos en vez de seis en una segunda reacción de alineamiento. En la tabla 16 se muestra el tamaño de los fragmentos y el limitante y la tabla 17 la constitución de la reacción de amplificación de alineamiento con los dos fragmentos. Para el segundo alineamiento la programación se definió de la siguiente manera: un primer paso de 95 °C por 10 minutos, seguido de 30 ciclos de 98 °C por 10 segundos, 58 °C por 45 segundos y 68 °C por 30 segundos,

con un paso final de 68 °C por 10 minutos, omitiendo los iniciadores en los primeros 10 ciclos, al término de estos se agregaron los iniciadores correspondientes. La reacción se analizó por gel de agarosa y se purificó con el kit Axygen AxyPrep DNA Gel Extraction. El inserto para el plásmido de pTCueO obtuvo en la amplificación de alineamiento directa de los seis fragmentos una alta calidad, por lo que su purificó directamente sin necesidad de una segunda amplificación de alineamiento. Los fragmentos fueron purificados por gel de agarosa usando el kit Axygen AxyPrep DNA Gel Extraction.

Fragmento	Peso molecular	Relación molar
N20 – Donador	853	1
Terminador – Brazo H2	1503	1.7

Tabla 16. Pesos moleculares y relación molar de fragmentos para la segunda reacción de alineamiento para pTYahI

Reactivo	Reacción N20 – YahIH2
ADN patrón	N20 – Prom/CspA-YahI 1 mol Donador – YahIH2 1.7 mol
Iniciador delantero	1.5 µL (30 pmoles)
Iniciador reverso	1.5 µL (30 pmoles)
dNTP's	5 µL (0.2 mM)
KOD DNA polimerasa	1 µL (1 U)
Buffer de reacción	5 µL (10x)
MgSO4	3 µL (1.5 mM)
Agua mQ	x µL
Volumen final	50 µL

Tabla 17. Constitución 1 x para reacción de alineamiento para el inserto de pTYahI

Una vez obtenido los dos insertos conformados por los seis fragmentos para los dos plásmidos derivados de pTarget fueron digeridos con las enzimas de restricción *XbaI* y *Sall* (Fast Digestion de Thermofisher). El plásmido original, pTarget fue digerido con las

enzimas de restricción *BcuI* y *Sall* adicionando al final FastAP (Fast Digestion de Thermofisher). El plásmido y los insertos fueron purificados mediante gel de agarosa para su posterior uso. La ligación de los insertos con el plásmido se realizó con la enzima T4 DNA ligasa (Takara) en relación plásmido:inserto de 1:5 y 1:10 a un volumen final de 8 μ L, la reacción se incubó durante 16 h a 16 °C . Todo el producto de la ligación se transformó en 100 μ L de células calcio-competentes *E. coli* TopTen. Después del choque térmico se inocularon en agar LB con espectinomicina a concentración final de 100 μ g/mL y se incubaron a 37 °C durante 16 h. Después de la incubación se almacenaron a 4 °C hasta que la proteína verde fluorescente fuese expresada. Se seleccionaron las clonas que presentaron la coloración verde típica y se sometieron a digestión. En el caso de pTCueO se realizó una digestión con *BamHI* y *Sall* (Fast digestión Thermofisher) y para pTYahl con *BamHI* y *SacI*. Además, se secuenciaron para asegurar la integridad de cada parte que completaba el inserto, en especial el N20 fundamental para el corte con Cas9.

4.5.3 Edición genómica de *E. coli* MG1655 y BL21(DE3)

La edición genómica con CRISPR/Cas9 se realizó con un sistema de doble plásmido en sinergia con recombinación λ *red*, el primer plásmido contiene el gen de la proteína Cas9 bajo su promotor nativo constitutivo y del sistema de λ *red* bajo la inducción del promotor ARA, en este caso el pCassac (plásmido actualizado del pCas9). Un segundo plásmido, el pTarget (y plásmidos derivados, en este trabajo pTCueO y pTYahl) posee el sgARN para direccionar el corte junto con el ADN donador y los brazos de homología bajo el promotor pJ23119 el cual es un promotor sintético constitutivo. Se inocularon 5 mL de LB con una clona de cada una de las líneas a editar y se incubó a 37 °C por 16 h con agitación a 200 rpm. El cultivo se centrifugó a 4000 rpm durante 10 minutos a 4 °C para recolectar el sedimento celular, las células se hicieron calcio-competentes lavando el

sedimento dos veces con solución de CaCl_2 0.1M centrifugando durante 10 minutos a 4000 rpm a 4 °C y con una resuspensión final en 100 μL de CaCl_2 0.1M con glicerol al 10%. Mediante choque térmico se transformaron los 100 μL obtenidos de cada línea con 50 ng del plásmido pCassac y se inocularon en agar LB con kanamicina a concentración final de 30 $\mu\text{g}/\text{mL}$. Se tomo una clona transformada con pCassac de cada línea para inocular 5 mL de LB con kanamicina a concentración de 30 $\mu\text{g}/\text{mL}$ y se añadió L-arabinosa (Sigma Aldrich) a una concentración final de 0.3 M para inducir las proteínas λ red y se incubo a 37 °C en agitación a 200 rpm durante 3 horas. Al término de la incubación se recolectó el sedimento a 4000 rpm durante 10 minutos a 4 °C y se prepararon las células para electroporación lavando el sedimento tres veces con solución fría de glicerol al 10% resuspendiendo a un volumen final de 100 μL . Las células se transformaron con 100 ng de pTCueo y pTYahl con electroporación empleando celdas de 2mm a 200 Ω , 25 μF y 2.5 kV y el electroporador de Biorad Gene Pulser Xcell. Se añadió inmediatamente 1 mL de LB frio a las células tratadas y se incubaron durante 1 h a 37 °C con agitación a 200 rpm. Las células fueron transferidas a agar LB con kanamicina y espectinomicina a concentración final de 30 $\mu\text{g}/\text{mL}$ y 100 $\mu\text{g}/\text{mL}$ respectivamente e incubadas a 37 °C durante 16 h. Se seleccionaron diez clonas de la transformación con pTCueo y se realizó una reacción de amplificación empleando tres iniciadores delanteros rio arriba la parte complementara del brazo de homología 1 (44bp, 159 bp y 287 bp rio arriba respectivamente) y un iniciador reverso dentro del gen de SmbP con el objetivo de amplificar un fragmento intermedio entre el genoma y la edición dentro de un gen heterólogo de *E. coli* como lo es SmbP. Los iniciadores usados se muestran en la tabla 18. La reacción de amplificación se realizó con la polimerasa EasyTaq (Transgen), la composición 1x se muestra en la tabla 19.

Iniciador		
Delantero	CueOH1 + 44bp	5'– CGGTCAATATTTCTACAAAAGGTAT – 3'
	CueOH1 + 159 bp	5'– CTGGTCGTGTAGCGTTTCAT – 3'
	CueOH1 + 287 bp	5'– TCGGTGGTTTGTACTACGCCAT – 3'
Reverso	SmbP	5'– CAGGGACTCCTTTCGCGTGT – 3'

Tabla 18. Iniciadores delanteros y reverso para confirmación de edición genómica en *E. coli*

Reacción 1 x para tamizaje de mutantes	
Patrón	Clona seleccionada
Iniciador delantero	0.6 µL (12 pmoles)
Iniciador reverso	0.6 µL (12 pmoles)
dNTP's	2 µL (0.2 mM)
Easy taq ADN polimerasa	0.5 µL (0.5 U)
Buffer de reacción	2 µL (10x)
Agua mQ	14.3 µL
Volumen final	20 µL

Tabla 19. Reacción 1 x para tamizaje de mutantes

Se seleccionaron dos clones de cada una en base a los resultados obtenidos de la amplificación y de la posterior expresión de la proteína verde fluorescente al ser almacenada a 4 °C. Se inocularon 5 mL de LB con kanamicina a concentración final de 30 µg/mL y se añadió L-Rhamnosa a una concentración final de 8 mM y se incubaron a 37 °C a 200 rpm durante 16 h para el curado de pTCueO y pTYahl. Después de la incubación, se tomaron 10 µL y se inocularon en placas de LB más kanamicina y espectinomicina a concentración de 30 µg/mL y 100 µg/mL respectivamente y en una segunda placa de LB más kanamicina a 30 µg/mL de concentración para prueba de sensibilidad a espectinomicina. De la segunda placa (kanamicina) se tomó una colonia y se inoculó en 5 mL de LB con 250 mg de glucosa y se incubó a 37 °C con agitación

de 200 rpm durante 16 h. Se tomaron 10 μ L y se inocular en una placa de agar LB con concentración de 5% de glucosa y 5% de sacarosa, la placa se inoculó a 37 °C. Se tomó una de las colonias obtenidas y se inocularon en placas de kanamicina a concentración de 30 μ g/mL para confirmar la sensibilidad a dicho antibiótico al perder el pCassac en el segundo curado.

CAPÍTULO 5

5. Resultados

5.1 Mutación y amplificación del gen CusF

Mediante mutagénesis dirigida empleando cebadores especiales para dicho propósito, se añadieron tres histidinas en el N terminal del gen de CusF. En la siguiente tabla se muestra el gen original de CusF y el gen con la modificación, el cual será llamado CusF3H+.

	Secuencia de aminoácidos
CusF	ANEHHH ETMSEAQPQVISATGVVKGIDLESKKITIHHDPIAAVNWPEMTMRFTITPQTKM SEIKTGDVKAFNFVQQGNLSLLQDIKVSQ
CusF3H+	MHHHHH ETMSEAQPQVISATGVVKGIDLESKKITIHHDPIAAVNWPEMTMRFTITPQTK MSEIKTGDVKAFNFVQQGNLSLLQDIKVSQ

Tabla 20. Secuencias de aminoácidos de CusF y el mutante CusF3H+.

En la figura 8 se puede observar la amplificación de cuatro variantes del gen CusF. En el primer carril se observa el marcador de peso molecular, en el segundo la amplificación de CusFc (variante citoplásmica de 303 bp), en el tercero CusFp (variante periplásmica de 330 bp), en el cuarto CusFsh (variante mutante de 285 bp) y finalmente en el quinto, CusF3H+ (variante mutante de 306 bp).

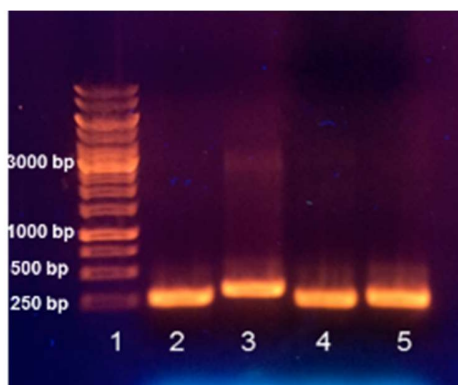


Figura 8. Amplificación de CusF y sus variantes. 1) marcador de peso molecular, 2) amplificación de CusFc, 3) CusFp, 4) CusFsh, 5) CusF3H+.

5.2 Construcción de plásmidos empleando el vector pET30a(+) y el gen de CusF3H+

5.2.1 Construcción de pET30a-CusF3H+

Los genes que fueron amplificados fueron introducidos al vector de expresión para *E. coli* seleccionado, en este caso, pET30a(+). En la figura 9 se puede observar la amplificación realizada a partir de cinco clonas obtenidas de la ligación de CusF3H+ a pET30a(+). Dichas amplificaciones se realizaron con los iniciadores T7. En la imagen del gel de agarosa se aprecia en el primer carril el marcador de peso molecular, en el segundo se observa un control negativo amplificando el espacio entre los T7 de pET30a que corresponde a ~375 bp. En el tercer y quinto carril se observa la amplificación positiva para dos clonas, la 1 y 3 respectivamente, mostrando el peso esperado de ~575 bp. Mientras que el carril 4 y 6 muestran resultados negativos.

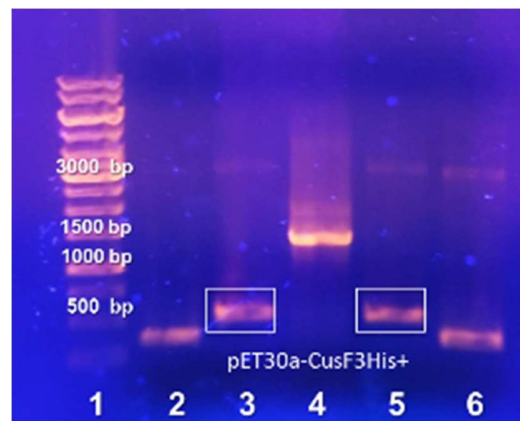


Figura 9. Amplificación de CusF3H+. 1) marcador de peso molecular, 2) control negativo, 3,5) amplificación para clonas positivas, 4,6) clonas negativas.

5.2.2 Construcción de pET30a-CusF3H+ con proteínas de interés

En la figura 10 se observa la amplificación realizada para los genes de CusF3H+-GFP con el iniciador delantero de CusF3H+ y el reverso de GFP. En el primer carril se observa el marcador de peso molecular, mientras que en los carriles segundo, tercero y cuarto se puede apreciar la banda de amplificación correspondiente a la suma de los pesos de CusF3H+ y GFP que es de aproximadamente 1020 bp.

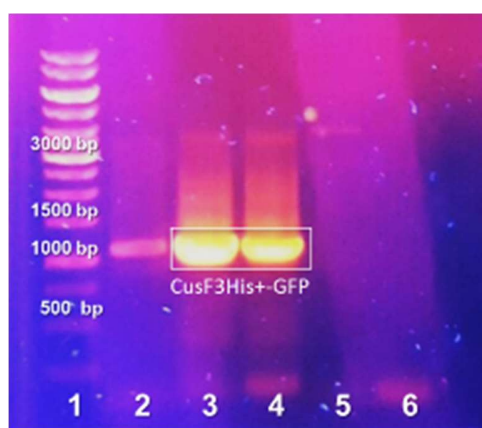


Figura 10. Amplificación de GFP marcado con CusF3H+, 1) marcador de peso molecular, 2 – 4) clonas positivas, 5,6) clonas negativas.

En la electroforesis presentada en la figura 11 se observa la amplificación realizada a partir del ADN plasmídico de las clonas obtenidas de la ligación de pET30a-CusF3H+ con el gen G-CSF. La reacción fue positiva para cuatro de las colonias probadas y el peso es el esperado situándose en aproximadamente las 789 bp. En este caso se empleó el iniciador delantero de CusF3H+ y el reverso para G-CSF.

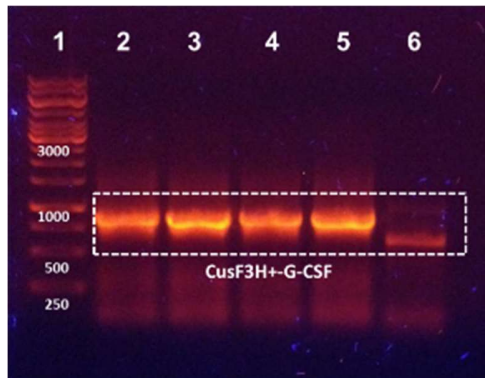


Figura 11. Amplificación de CusF3H+-G-CSF. 1) marcador de peso molecular, 2-5) clonas positivas, 6) clona negativa.

5.3 Expresión piloto de proteínas marcadas con CusF3H+ y His-tag

5.3.1 Expresión piloto de CusF3H+-GFP

En la figura 12 se puede observar el sedimento celular de la expresión piloto de la construcción CusF3H+-GFP en *E. coli*. Se aprecia claramente el color característico de la expresión de proteína verde fluorescente.



Figura 12. Expresión piloto de GFP marcada con CusF3H+

En la figura 13 se observa la expresión piloto de la proteína reportera, GFP, marcada con CusF3H+ y His-Tag. En la comparación se observa que la proteína reportera se expresó con las dos etiquetas, dando en el caso de CusF3H+ una proteína 37 kDa y para His_GFP una proteína de 26 kDa. El rendimiento fue calculado con el programa libre de ImageJ (versión 1.49), siendo para CusF3H+_GFP del 36% y para His_GFP del 31%.

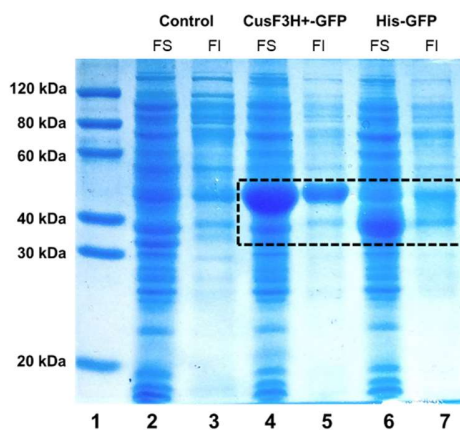


Figura 13. Expresión piloto de GFP marcada con CusF3H+ (carriles 4 y 5) y con His-tag (carriles 6 y 7), el marcador de peso molecular se muestra en el carril 1 y las fracciones solubles e insoluble se marcan con FS y FI respectivamente

5.3.2 Expresión piloto de CusF3H+_G-CSF

En la figura 14 se observan las expresiones piloto para dos diferentes proteínas quiméricas. El marcador de peso molecular se presenta en el primer carril, mientras que los carriles 2 al 5 muestran la expresión piloto para CuSF3H+-G-CSF y los carriles del 6 al 9 para His-G-CSF. Además, se muestran las fracciones solubles (FS) e insoluble (FI) para cada una de ellas. Para la proteína quimérica CusF3H+-G-CSF el peso molecular esperado es de 29.5 kDa y para His-G-CSF es de 22 kDa aproximadamente.

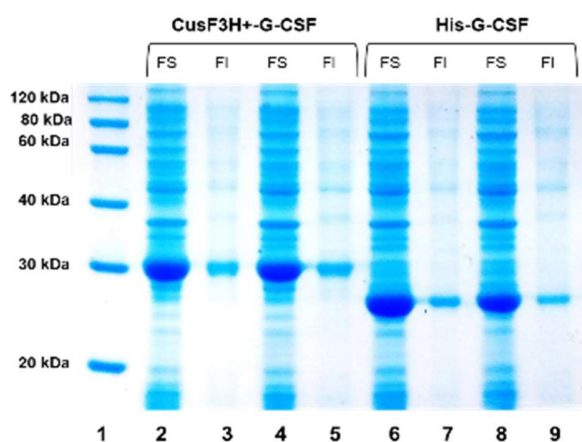


Figura 14. Expresión piloto de CusF3H+-G-CSF y His-G-CSF. 1) marcador de peso molecular, 2,4) fracción soluble de CusF3H+-G-CSF (rendimiento de expresión: ~ 38.3%), 3,5) fracción insoluble de CusF3H+-G-CSF, 6,8) fracción soluble de His-G-CSF (rendimiento de expresión: ~ 36.2%), 7,9) fracción insoluble de His-G-CSF

5.3.3 Expresión piloto de CusF3H+_AXR3

La figura 15 muestra la expresión piloto de la expresión de la proteína AXR3 marcada con CusF3H+. En el carril 1 se muestra el marcador de peso molecular, en el carril 2 y 3 aparece el control correspondiente al lisado celular (FS) de una cepa de *E. coli* BL21(DE3) y la fracción insoluble obtenida (FI). En el carril 4 y 5 se observa la expresión piloto de AXR3, correspondiendo el carril 4 a su fracción soluble y el cinco a la insoluble, la banda correspondiente a la proteína recombinante se encuentra dentro del

rectángulo. El rendimiento calculado para AXR3 marcado con CusF3H+ es del 44% del contenido total de proteínas del sedimento celular.

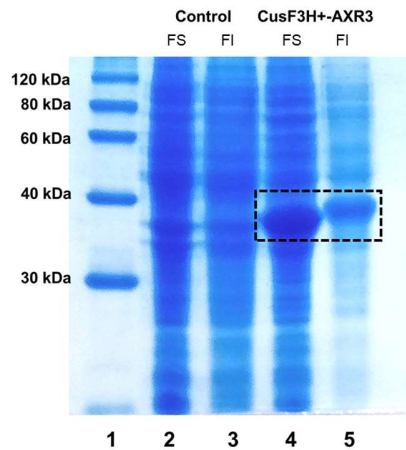


Figura 15. Expresión piloto de CusF3H+_AXR3, 1) marcador de peso molecular, 2, 3) control BL21(DE3) sin transformar, fracción soluble e insoluble, 4, 5) expresión de CusF3H+_AXR3 mostrando la fracción soluble e insoluble

5.3.4 Expresión piloto de CuF3H+_IMB1 y His_IMB1

En la imagen 16 se presenta el SDS-PAGE de la expresión piloto de IMB1 marcada con CusF3H+ y His-tag para comparación. En el carril 1 se muestra el marcador de peso molecular, en los carriles 2 y 3 se observa el control negativo correspondiente a una cepa sin transformar de *E. coli* BL21(DE3), correspondiendo el carril 2 a la fracción soluble y el 3 a la fracción insoluble, en los carriles 4 y 5 se muestra la expresión de IMB1 marcada con CusF3H+, en el 4 se presenta la fracción soluble y en el 5 la fracción insoluble. En los siguientes carriles se muestra la expresión de la proteína marcada únicamente con el His_tag, con su respectiva fracción soluble (carril 6) y la insoluble (carril 7). Los rendimientos calculados para IMB1 marcada con CusF3H+ y His-tag son de 27% y 26% respectivamente.

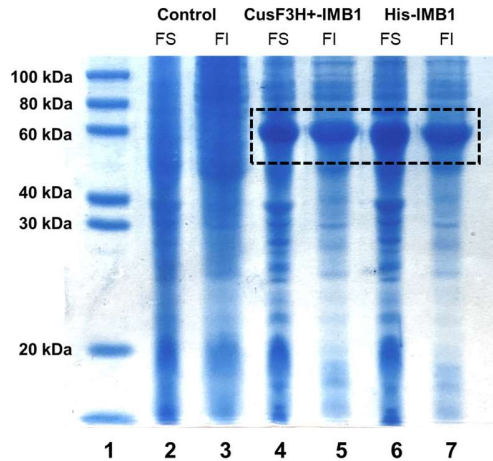


Figura 16. Expresión piloto de IMB1 marcada con CusF3H+ y His, 1) marcador de peso molecular, 2, 3) control BL21(DE3) fracción soluble e insoluble, 4, 5) fracción soluble e insoluble de IMB1 marcada con CusF3H+, 6,7) fracción soluble e insoluble de IMB1 marcada con His

5.3.5 Expresión piloto de CusF3H+_PiF3 y His_PiF3

La expresión correspondiente a PiF3 con las tres etiquetas se presenta en el SDS-PAGE de la figura 17. El marcador de peso molecular se muestra en el carril 1, los carriles 2 y 3 presentan el control correspondiente a una cepa de *E. coli* BL21(DE3) sin transformar, el carril 2 muestra la fracción soluble y el carril 3 la fracción insoluble. La expresión de PIF3 marcada con CusF3H+ se observa en los carriles 4 (fracción soluble) y 5 (fracción insoluble). Mientras que la proteína marcada únicamente con el His-tag se muestra en los carriles 6 (fracción soluble) y en el 7 (fracción insoluble). El rendimiento para PiF3 con CusF3H+ es del 21% y con His-tag es del 19%.

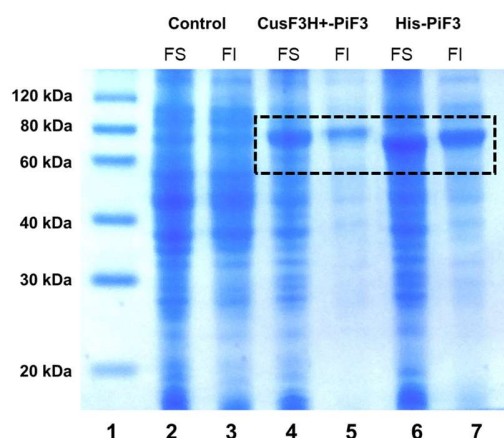


Figura 17. Expresión de PiF3 marcada con CusF3H+ y His. 1) marcador de peso molecular; 2, 3) control BL21(DE3), fracción soluble e insoluble; 4, 5) fracción soluble e insoluble de PiF3 marcada con CusF3H+; 6,7) fracción soluble e insoluble de PiF3 marcada con His

5.3.6 Expresión piloto de CusF3H+_PixE y His_PixE

El SDS-PAGE de la expresión piloto de PixE se muestra en la figura 18. El marcador de peso molecular se aprecia en el carril 1, en los carriles 2 y 3 las fracciones soluble e insoluble del control que corresponde a *E. coli* BL21(DE3) sin transformar. Los carriles 4 y 5 muestran las fracciones soluble e insoluble de la expresión de la proteína CusF3H+_PixE, mientras que los siguientes carriles 6 y 7 se observa las fracciones soluble e insoluble de la expresión de His_PixE. Los rendimientos calculados para PixE son de 18% con CusF3H+ y de 16% con His-tag.

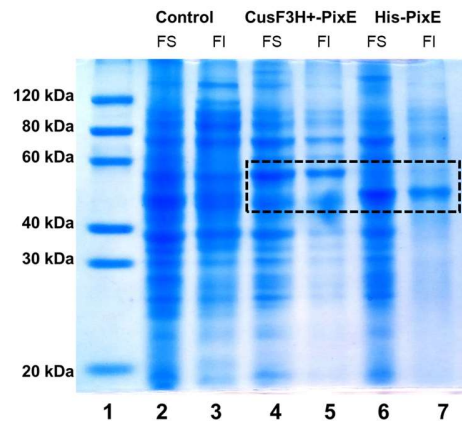


Figura 18. Expresión piloto de PixE marcada con CusF3H+ y His. 1) marcador de peso molecular; 2, 3) control BL21(DE3) fracción soluble e insoluble; 4, 5) fracción soluble e insoluble de PixE marcada con CusF3H+; 6, 7) fracción soluble e insoluble de PixE marcada con His

5.4 Purificación de proteínas

5.4.1 Purificación de CusFc-GFP

La figura 19 muestra el cromatograma obtenido de la purificación de CusFc-GFP, donde se aprecia el comportamiento típico de este tipo de cromatografía. En dicho cromatograma se observa a partir de los 100 mL el pico correspondiente a la elución de la proteína quimérica CusFc-GFP. Durante la purificación se observó visualmente una pérdida de la proteína durante el paso de la muestra por la columna cargada con Ni(II). En la figura 20 se observa el SDS-PAGE realizado a partir de las fracciones recuperadas del percolado mostrando un peso aproximado de 37.5 kDa. En el gel presentado se observa claramente la banda correspondiente a CusFc-GFP demostrando la pérdida de la misma. En el gel mostrado en la figura 21 se observan las fracciones recuperadas de la elución de la proteína con 50 mM de imidazol. Se observa la banda característica de la proteína quimérica junto con otras impurezas.

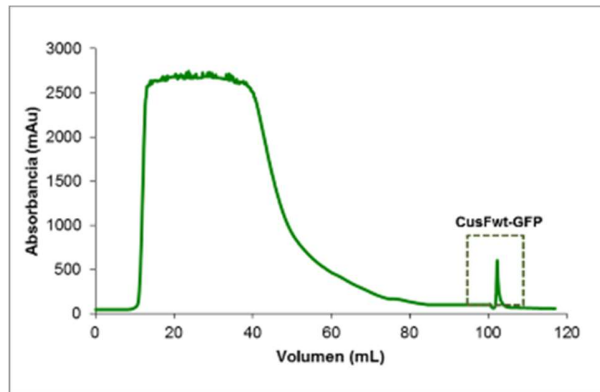


Figura 19. Cromatograma obtenido de la purificación de CusFc-GFP con FPLC ÄKTA prime plus usando Ni(II).

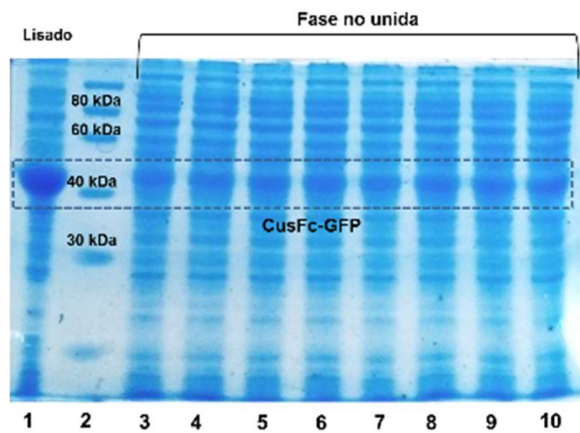


Figura 20. SDS-PAGE 12% donde el primer carril corresponde al lisado obtenido del sedimento celular, en el segundo se muestra el marcador de peso molecular, los carriles del 3 al 10 corresponden a las fracciones recuperadas a partir de la fase no unida a la columna con Ni(II) mediante el FPLC ÄKTA prime plus. Dichas fracciones corresponden a la "meseta" en el cromatograma de la figura 8.

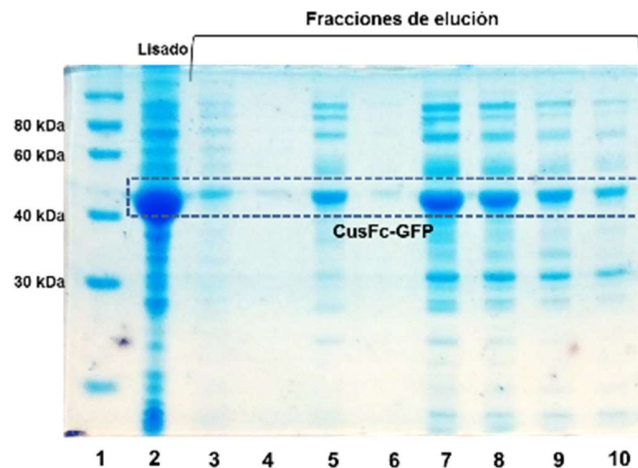


Figura 21. SDS-PAGE 12% en el carril 1 se muestra el marcador de peso molecular, en el segundo el lisado antes de la purificación, del carril 3 al 10 se muestran las fracciones recuperadas de la elución de CusFc-GFP remanente en la columna con 50 mM de imidazol. Dichas fracciones corresponden al "pico" observado en el cromatograma

5.4.2 Purificación de CusF3H+_GFP

En la figura 22 se presenta el cromatograma de la purificación de CusF3H+-GFP mostrando el comportamiento típico de este tipo de cromatografía. Aproximadamente a los 95 mL se observa el pico de elución que corresponde a CusF3H+-GFP. La purificación se realizó con las condiciones establecidas después de la purificación de CusFc-GFP. En la figura 23 se muestra el SDS-PAGE correspondiente a las fracciones de elución recuperadas. Se puede observar la presencia de la proteína quimérica en el lisado empleado para comparación. En los carriles donde se muestran las fracciones de elución recuperadas se aprecia la banda correspondiente a CusF3H+-GFP. El rendimiento para la proteína quimérica CusF3H+-GFP se calculó en aproximadamente 6 mg.

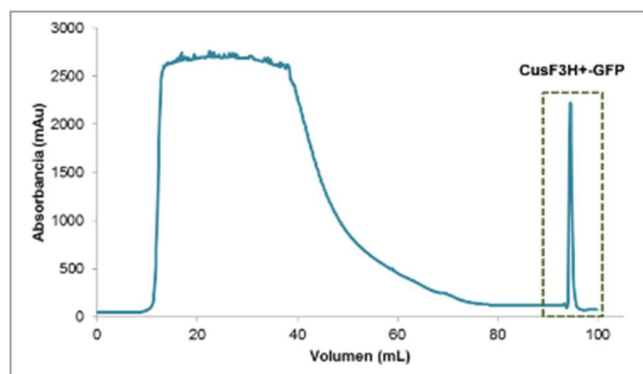


Figura 22. Cromatograma obtenido de la purificación de CusF3H+GFP con el FPLC ÄKTA prime plus empleando una columna tipo IMAC cargada con Ni(II)

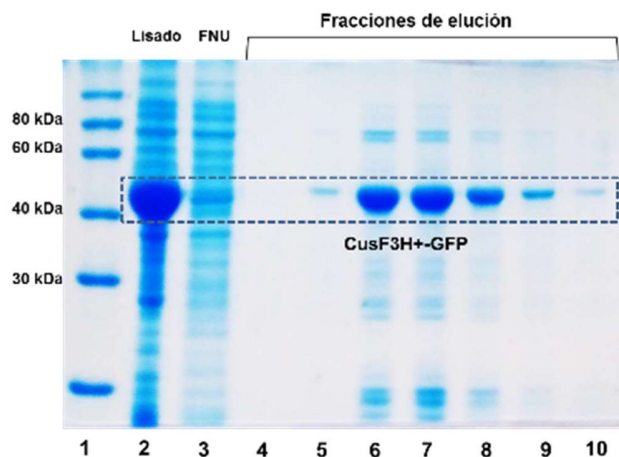


Figura 23. SDS-PAGE 12% mostrando en el carril 1 el marcador de peso molecular, en el carril 2 el lisado obtenido a partir del sedimento celular, en el carril 3 la fase no unida (“meseta” en el cromatograma) y del carril 4 al 10 se aprecian las fracciones recuperadas a partir de la elución de CusF3H+GFP con 250 mM de imidazol (“pico” en el cromatograma)

5.4.3 Purificación de His_GFP

En la figura 24 se muestra el SDS-PAGE de las fracciones obtenidas de la purificación de la etiqueta His-tag con la proteína reportera. La purificación se realizó con las condiciones establecidas por el proveedor de las columnas comerciales HiTrap FF (GE Healthcare Life Sciences) y presentó el comportamiento típico de la cromatografía IMAC.

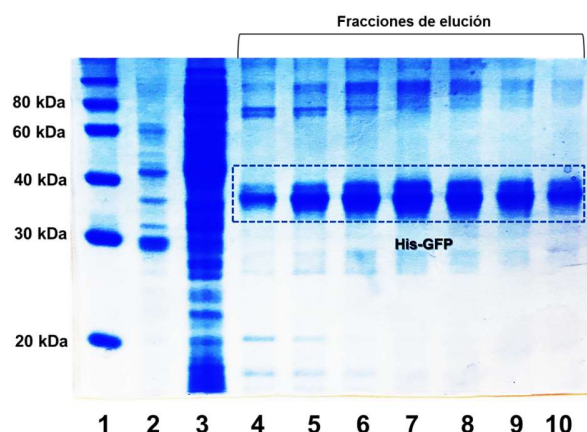


Figura 24. SDS-PAGE de fracciones recolectadas de la purificación de His_GFP

5.4.4 Remoción de CusF3H+ y obtención de GFP pura

En la figura 25 se muestra el SDS-PAGE realizado a partir del corte realizado con la enzima enteroquinasa. En el primer carril se observa la proteína quimérica obtenida de la purificación con IMAC con iones Ni(II), en el carril 2 se observa el corte realizado por la enzima, mostrando ambas proteínas separadas y diferenciadas por su peso molecular (CusF3H+ ~10 kDa, y GFP ~26.7 kDa). La proteína quimérica fue cortada en su totalidad a las 24 h. El tercer carril muestra el marcador de peso molecular. Los carriles 4 y 5 muestran a GFP obtenida después de una segunda purificación empleando una resina funcionalizada con Ni(II). En el carril 6 se observa la proteína CusF3H+ después de su elución con imidazol. Considerando que la proteína de fusión corresponde apenas a un 27% de la proteína quimérica se puede considerar que en la segunda purificación se pueden obtener hasta ~ 4.4 mg de GFP pura a partir de los 6 mg obtenidos por L. Para cortar 25 µg de proteína se requiere de 1 U de enteroquinasa ($1.6 \mu\text{g}^{-04}$), para cortar 200 µg de proteína se empleó 1 µL de enzima, que corresponden a 12.5 U.

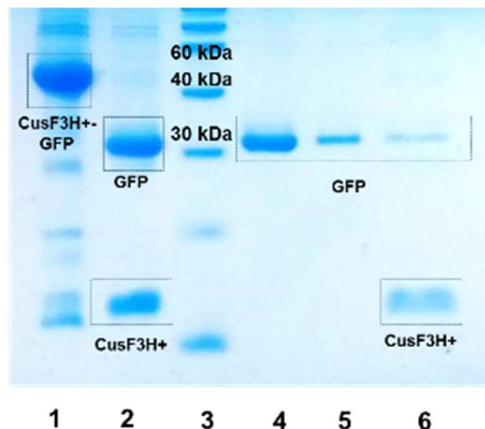


Figura 25. SDS-PAGE 15% mostrando el corte de la proteína de fusión empleando enteroquinasa de cadena ligera. El primer carril muestra la proteína recuperada de la purificación con Ni(II) con la técnica IMAC. En el carril 2 se muestra la reacción de corte a las 24 h de incubación, donde se aprecia que la proteína quimérica fue cortada en su totalidad. En el carril 3 se muestra el marcador de peso molecular, los carriles 4 y 5 muestran el sobrenadante recuperado después de la incubación de la reacción con una resina cargada con Ni(II), el carril 6 muestra la proteína de fusión, CusF3H+, recuperada con 250 mM de imidazol a partir de la resina.

5.5 Construcción de plásmidos derivados de pZ8-ptac

Los genes sintetizados fueron recibidos en el plásmido pUC57, así que los genes fueron obtenidos mediante una digestión empleando las enzimas *EcoRI* y *Sall*. Los productos de la digestión para cada uno de los plásmidos se muestran a continuación. En la figura 26 se muestra la digestión para la obtención de CgR-Smbp de 444 bp en el primer carril combinado, en el segundo muestra el marcador de peso molecular, en el tercer carril combinado se observa CgR-CusF3H+ con un peso de 435 bp. En la figura 26 se expone la digestión para la obtención del siguiente par de genes, en el primer carril combinado se observa el producto de TorA-SmbP con un peso de 471 bp, en el segundo se muestra el marcador de peso molecular y en el tercer carril combinado a TorA-CusF3H+ con 462 bp. Cabe mencionar que el peso de los cuatro genes es ligeramente mayor a la simple combinación de ambos por los sitios de restricción y de reconocimiento añadidos que suman 69 bp en cada uno.

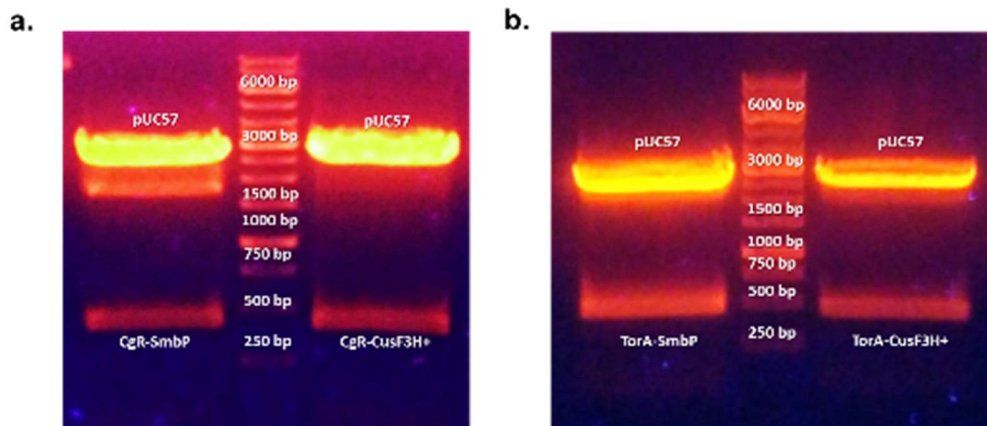


Figura 26. Digestión de plásmidos pUC57 para la purificación de los genes: a) CgR-SmbP, CgR-CusF3H+; b) TorA-SmbP, TorA-CusF3H+

Los genes digeridos fueron insertados dentro del plásmido pz8-ptac generando cuatro diferentes plásmidos. Para la confirmación de las clonas positivas, en el caso de los insertos de CusF3H+ se realizó una amplificación empleando los iniciadores específicos para dicho gen. En la electroforesis de la figura 27 se aprecia la amplificación para tres clonas, en el carril 1 y 2 se observa la amplificación de CusF3H+ a partir del plásmido pz8-CgR-CusF3H+, el peso del producto concuerda con el de CusF3H+ de 273 bp. En el carril se muestra el marcador de peso molecular, en el carril 4 se observa la amplificación realizada sobre el plásmido pz8-TorA-CusF3H+, al igual que en los carriles 1 y 2, el producto tiene el peso molecular esperado concordando con el de CusF3H+.

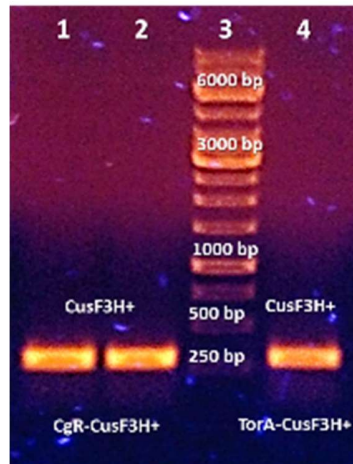


Figura 27. Amplificación de los genes de CusF3H⁺ marcados con los dos diferentes péptidos señal. En el caso de las construcciones con SmbP no fue posible realizar la confirmación de clonas positivas debido a que los iniciadores con los que se contaban diferían de la secuencia sintetizada debido a que fue optimizada para su uso en *C. glutamicum* por lo que las reacciones resultaron negativas (electroforesis de amplificación no mostrada). Debido a esto se confirmaron mediante análisis de restricción, empleando las enzimas *EcoRI* y *Sall*. En la siguiente imagen (figura 28) se muestra el análisis de restricción realizado sobre dos clonas. En el carril 1 se aprecia el marcador de peso molecular, en el carril 2 el análisis de una clona con el plásmido pZ8-CgR-Smbp, se aprecia el gen digerido a la altura esperada de 444 bp, en el carril 3 se muestra un control positivo, el gen CgR-SmbP obtenido originalmente de pUC57, en el carril 4 se encuentra la segunda digestión en este caso para el plásmido pZ8-TorA-SmbP donde es posible apreciar la banda a la altura esperada de 471 bp concordando con el control positivo del carril 5, donde se muestra el gen TorA-SmbP obtenido del pUC57 anteriormente.

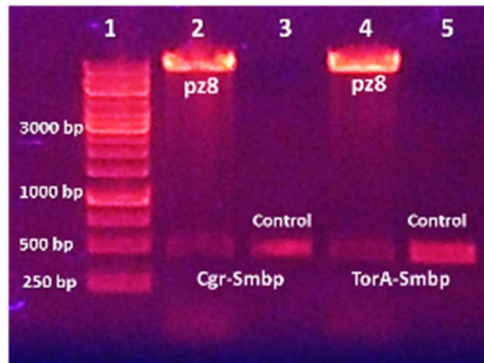


Figura 28. Análisis de restricción para los genes de SmbP marcados con los dos diferentes péptido señal. Una vez obtenidos los plásmidos de pz8 con los insertos de las cuatro combinaciones de péptido señal y proteínas de fusión se agregó la proteína reportera, en este caso GFP. En la figura 29 se muestra el producto de amplificación de cuatro diferentes clonas probadas, la reacción se realizó con los iniciadores específicos para GFP. El marcador de peso molecular se aprecia en el carril 1, en el carril 2 y 3 se observa la amplificación para las clonas de pz8-CgR-CusF3H+_GFP donde ambas muestran la banda correspondiente a GFP con 717 bp. En los carriles 4 y 5 se aprecia la amplificación para las clonas de pz8-TorA-CusF3H+_GFP, en donde la primera clona (carril 4) es positiva mientras que la segunda resultó negativa (carril 5), en el carril 6 se muestra un control positivo mostrando la banda de GFP en el peso molecular esperado.

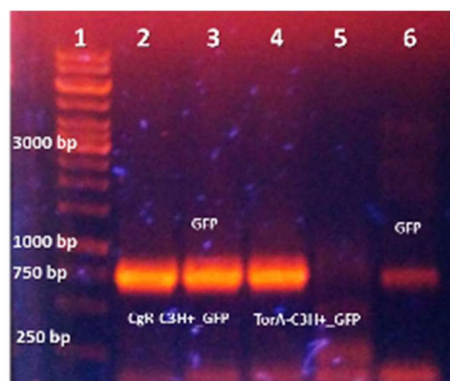


Figura 29. Amplificación de clonas para tamizaje de GFP, se observan ambas clonas de CgR-CusF3H+_GFP positivas, mientras que para TorA-CusF3H+_GFP solo una es positiva.

Para el tamizaje de las construcciones de pz8-CgR-SmbP-GFP se amplificó el gen de la proteína reportera GFP empleando los iniciadores específicos. En la figura 30 se observa la electroforesis de dicha amplificación, en el carril se muestra el marcador de peso molecular, en los carriles 2 y 3 se observa el producto de la amplificación en el peso molecular esperado de 717 bp, en el carril 4 un control positivo y en el carril 5 un control negativo (sin ADN patrón).

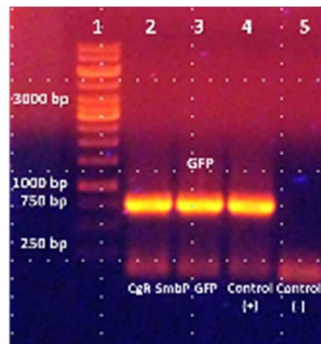


Figura 30. Amplificación de GFP para tamizaje de clonas positivas, las dos colonias de CgR-SmbP-GFP resultaron positivas.

Para pz8-TorA-SmbP-GFP se realizó la amplificación del ADN plasmídico de una sola clona. El producto se observa en la figura 31 en donde en el carril se muestra el marcador de peso molecular, en el carril 2 el producto de la amplificación de la clona mostrando una banda en el peso molecular esperado (717 bp), en el carril 3 se muestra un control negativo (sin ADN patrón) y en el carril 4 el control positivo mostrando la banda de GFP que coincide con la del producto del plásmido analizado.

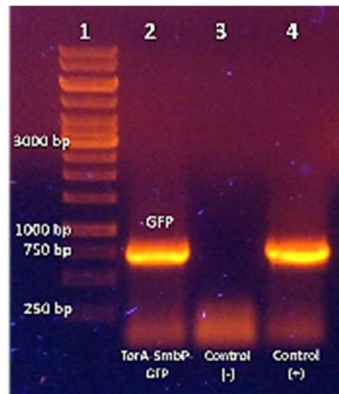


Figura 31. Amplificación del ADN plasmídico extraído de una clona de TorA-SmbP-GFP; la amplificación de GFP muestra que la clona es positiva para dicho inserto

5.6 Expresión de proteínas en *C. glutamicum*

Una vez que se obtuvieron las cuatro construcciones en pz8 se procedió con la expresión de proteínas en la cepa de *C. glutamicum* (ATCC 13032). Se empleó un medio definido modificado en base a lo publicado por Jae Woong Choi⁵⁷. En la figura 32 se muestran los resultados de la expresión de la proteína SmbP-GFP marcada con TorA. En el primer carril se muestra el marcador de peso molecular, en el carril 2 se muestra el lisado celular de 2 mL un cultivo de *C. glutamicum* sin transformar. En el carril 3 se muestra el lisado celular de 2 mL de un cultivo de *C. glutamicum* transformado con el plásmido pz8-TorA-SmbP-GFP sin inducir, en el carril 4 se muestra el lisado del mismo cultivo, pero inducido a un O.D.600 de 0.6, y el carril 5 inducido a un O.D.600 de 1.0. El carril 6 muestra el medio de cultivo recuperado de una cepa de *C. glutamicum* sin transformar, el carril 7 es el medio recuperado de la cepa de *C. glutamicum* transformada sin inducir, en el carril 8 se muestra el medio recuperado después de la inducción a 0.6 en el carril 9 el medio recuperado de la inducción a 1.0. El patrón observado en los carriles de lisado celular es similar en los cuatro, no existe una diferencia notoria, tanto de los sedimentos inducidos como el del control y la cepa sin inducir. Analizar el patrón de bandas de los lisados celulares permite observar si existe proteína aún recluida en

el citoplasma, ya sea por una identificación ineficiente del sistema Tat o una sobresaturación del mismo. La diferencia se hace presente en el análisis de los medios de cultivo, en donde tanto en la cepa sin transformar, como en la cepa sin inducir la presencia de bandas es bastante baja. La inducción sin embargo parece no solo permitir la expresión de nuestra proteína deseada, sino que también de otras más como se observa en el gel. La banda marcada en el recuadro corresponde a la proteína SmbP-GFP, dicha banda no se aprecia en los controles (carril 6 y 7). Cabe mencionar que los medios recuperados tienen una fluorescencia verde (aunque baja) al ser observados con luz ultravioleta, y tanto la cepa sin transformar y la cepa sin inducir carecen de esta fluorescencia.

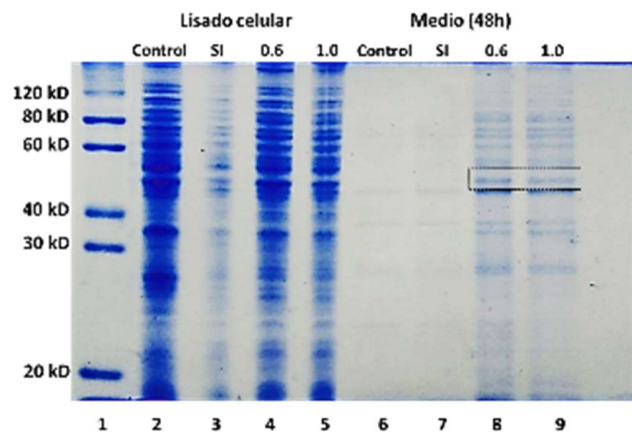


Figura 32. Análisis del lisado celular y medio de cultivo después de la inducción de TorA-SmbP-GFP. El carril 1 muestra el marcador de peso molecular, los carriles 2 y 3 corresponden al lisado celular de los dos controles empleados, correspondiendo el primero a la cepa de *C. glutamicum* sin transformar, y el 3 a una cepa transformada sin inducir, el carril 4 muestra el lisado celular completo después de la inducción a 0.6 de O.D.600, el carril 5 muestra el lisado a la expresión a 1.0 de O.D.600. Los carriles 6 y 7 muestran el medio de cultivo recuperado de los dos controles usados antes mencionados, el carril 8 muestra el sobrenadante obtenido después de la expresión a un O.D.600 de 0.6 y el 9 a un O.D.600 de 1.0

En la figura 33 se muestran los resultados de la expresión de la proteína SmbP-GFP marcada con CgR. En el carril 1 se muestra el marcador de peso molecular, en el carril 2 se muestra el control de lisado de celular de una cepa de *C. glutamicum* sin

transformar. En el carril 3 el lisado celular de una cepa de *C. glutamicum* transformada con el plásmido pz8-CgR-SmbP-GFP. En el carril 4 se muestra el lisado celular de la cepa transformada inducida a un O.D.600 a 0.6 y en el 5 inducida a un O.D.600 de 1.0. En el carril 6 se muestra el medio obtenido de la cepa de *C. glutamicum* sin transformar, en el carril 7 el medio obtenido de la cepa transformada sin inducir, en el carril 8 se muestra el medio obtenido de la inducción de la cepa a un O.D. 600 de 0.6 y en el 9 inducido a un O.D.600 de 1.0. El patrón se repite como en la figura 32, en donde el sedimento celular, el patrón de bandas en los cuatro carriles es bastante similar sin alguna diferencia característica. En el caso del análisis del medio, se observa una diferencia clara entre el medio recuperado de la inducción. Sin embargo, se observa que la inducción a 0.6 no es tan eficiente como a 1.0, en donde en esta última se observa la banda esperada (marcada en el recuadro).

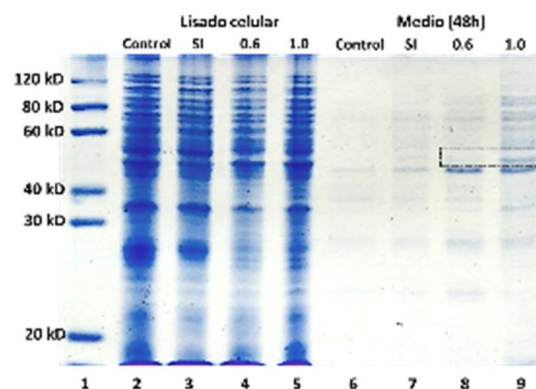


Figura 33. Análisis de la expresión de CgR-SmbP-GF. El carril 1 muestra el marcador de peso molecular, los carriles 2 y 3 corresponden al lisado celular de los dos controles empleados, correspondiendo el primero a la cepa de *C. glutamicum* sin transformar, y el 3 a una cepa transformada sin inducir, el carril 4 muestra el lisado celular completo después de la inducción a 0.6 de O.D.600, el carril 5 muestra el lisado a la expresión a 1.0 de O.D.600. Los carriles 6 y 7 muestran el medio de cultivo recuperado de los dos controles usados antes mencionados, el carril 8 muestra el sobrenadante obtenido después de la expresión a un O.D.600 de 0.6 y el 9 a un O.D.600 de 1.0

En la figura 34 se observa la expresión de la proteína CusF3H+-GFP marcada con TorA.

En el carril 1 se observa el marcador de peso molecular, en el carril 2 se observa el

control que corresponde al lisado celular de una cepa de *C. glutamicum* sin transformar. En el carril 3 se observa el lisado celular de una cepa de *C. glutamicum* transformada con el plásmido pz8-TorA-CusF3H+-GFP sin inducir, en el carril 4 se muestra el lisado celular de la inducción a 0.6 y a 1.0. En el carril 6 se muestra el medio obtenido de la cepa control sin transformar, en el carril 7 se observa el medio recolectado de la cepa transformada sin inducir, en el carril 8 se muestra el medio después de la inducción a 0.6 y en el carril 9 la inducción a 1.0. En la imagen se aprecia como el patrón se repite y no existe diferencia en los lisados celulares, la diferencia se marca en el medio recolectado. Después de la inducción se observa también la presencia de más bandas además de la esperada y la inducción a un O.D.600 de 0.6 parece ser más eficiente a diferencia de las otras 2.

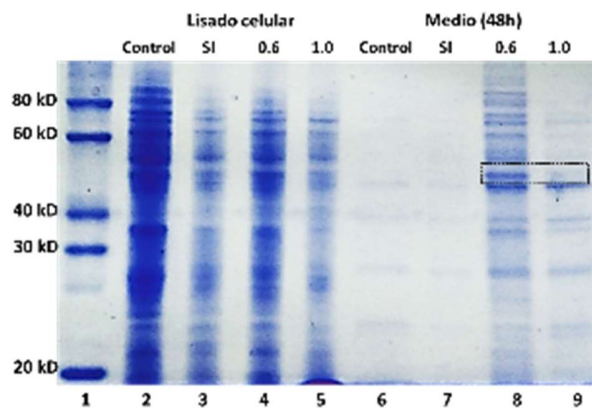


Figura 34. Análisis de la expresión de Tora-CusF3H+-GFP. El carril 1 muestra el marcador de peso molecular, los carriles 2 y 3 corresponden al lisado celular de los dos controles empleados, correspondiendo el primero a la cepa de *C. glutamicum* sin transformar, y el 3 a una cepa transformada sin inducir, el carril 4 muestra el lisado celular completo después de la inducción a 0.6 de O.D.600, el carril 5 muestra el lisado a la expresión a 1.0 de O.D.600. Los carriles 6 y 7 muestran el medio de cultivo recuperado de los dos controles usados antes mencionados, el carril 8 muestra el sobrenadante obtenido después de la expresión a un O.D.600 de 0.6 y el 9 a un O.D.600 de 1.0

En la figura 35 se aprecia la expresión de la proteína GFP marcada por el dúo de CusF3H+-CgR. En el carril 1 se observa el marcador de peso molecular, en el carril número 2 se cargó la cepa de *C. glutamicum* sin transformar como control y un segundo

control, de *C. glutamicum* transformada con el plásmido, pero sin inducir se agregó en el carril 3. Los carriles 4 y 5 muestran el sedimento celular lisado después de las inducciones a dos diferentes O.D.₆₀₀, 0.6 y 1.0. En el carril 6 se observa el sobrenadante obtenido del control de *C. glutamicum* sin transformar, mientras que en el carril 7 se muestra el segundo control correspondiente al control de *C. glutamicum* transformada, pero sin inducción con IPTG. Los carriles 8 y 9 muestran el sobrenadante obtenido de la inducción de cultivo a los dos diferentes O.D.₆₀₀ mencionados anteriormente.

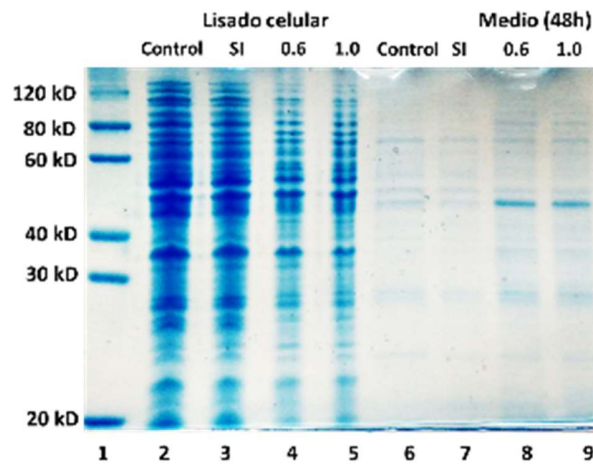


Figura 35. Análisis de expresión de CgR-CusF3H+-GFP. El carril 1 muestra el marcador de peso molecular, los carriles 2 y 3 corresponden al lisado celular de los dos controles empleados, correspondiendo el primero a la cepa de *C. glutamicum* sin transformar, y el 3 a una cepa transformada sin inducir, el carril 4 muestra el lisado celular completo después de la inducción a 0.6 de O.D.₆₀₀, el carril 5 muestra el lisado a la expresión a 1.0 de O.D.₆₀₀. Los carriles 6 y 7 muestran el medio de cultivo recuperado de los dos controles usados antes mencionados, el carril 8 muestra el sobrenadante obtenido después de la expresión a un O.D.₆₀₀ de 0.6 y el 9 a un O.D.₆₀₀ de 1.0

5.7 Purificación de proteínas secretadas de *C. glutamicum*

Del medio obtenido de la expresión de las cuatro distintas clonas transformadas de *C. glutamicum* se procedió la purificación para cada una. En un principio, la proteína TorA-SmbP-GFP expresada en el medio de *C. glutamicum* se empleó para pruebas de purificación preliminares con la resina profinity IMAC de BIORAD empleando iones de Cu(II). Sin embargo, los resultados obtenidos no fueron los esperados. La cantidad que

de proteína que se unía a la resina fue mínima, y la mayoría de las proteínas secretadas se unían también a la resina, y aun cuando se cambió a iones Ni(II) se repetía la situación. Pensando que quizás el CaCl₂ presente en exceso en el medio de cultivo podría haber interferido en la unión de la proteína a la resina. Se realizó una diálisis de 100 mL del medio obtenido de una expresión de TorA-SmbP-GFP contra una solución amortiguadora de 50 mM Tris, 500 mM NaCl a pH 8.0. Sin embargo, aún después de este tratamiento se presentó la misma situación aún con ambos iones metálicos. Considerando la situación, se prefirió emplear una técnica de FPLC para la purificación empleando las columnas especiales para cromatografía tipo IMAC.

Para la purificación de TorA-SmbP-GFP se decidió cargar la columna con Cu(II), ya que SmbP ha mostrado tener una gran afinidad a dicho metal y la presencia de pocas proteínas en el medio (a comparación de un lisado total o una fracción periplásmica) debería permitir una rápida purificación con dicho metal, asegurando la unión completa de la proteína de interés a la fase estacionaria y con una alta recuperación. El cromatograma obtenido se muestra en la figura 36a) en donde se observa el patrón típico que sigue una cromatografía de afinidad.

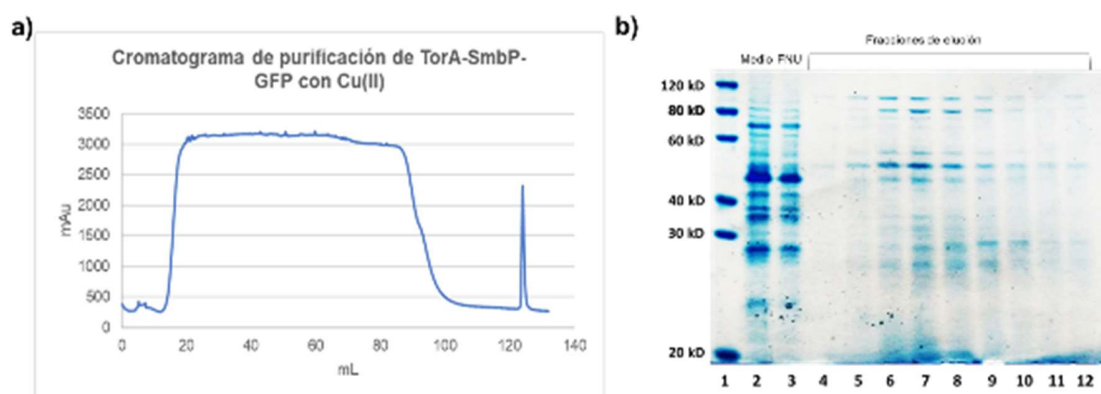


Figura 36. Purificación de TorA-SmbP-GFP por FPLC ÄKTA prime plus con Cu(II). a) cromatograma, b) análisis SDS-PAGE de las fracciones recuperadas

Las fracciones obtenidas que comprendían el pico de elución fueron analizadas por SDS-PAGE que se puede observar en la figura 36b). El gel muestra que, aunque haya habido una mayor afinidad de nuestra proteína por la fase estacionaria sin duda existe una importante presencia de bandas de las proteínas cuya secreción fue también inducida (no es posible decir si en si su producción fue también inducida).

La proteína TorA-CusF3H+-GFP fue purificada en cambio, utilizando los iones Ni(II) debido a que era la finalidad de la mutación de CusF. La figura 37 a) muestra el cromatograma de la purificación de esta proteína, en donde se observa el patrón característico que presentan las cromatografías de dicho tipo.

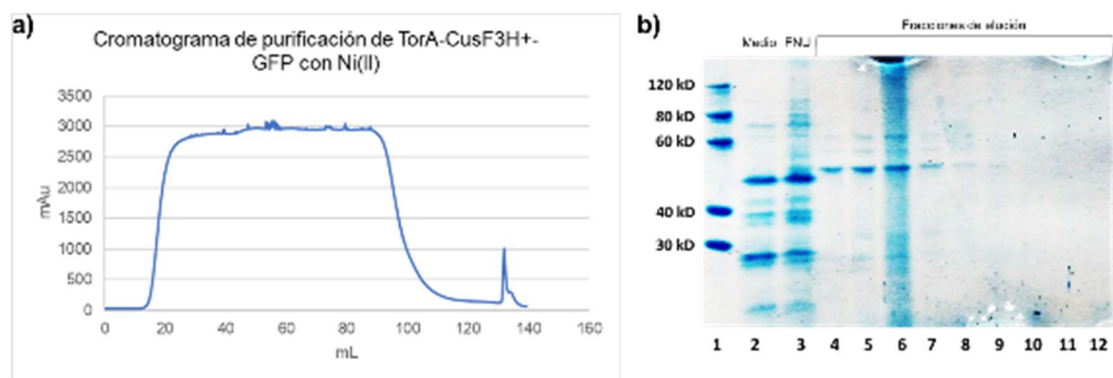


Figura 37. Purificación de TorA-CusF3H+-GFP con FPLC ÄKTA prime plus con Ni(II). a) cromatograma, b) SDS-PAGE de las fracciones recuperadas

La figura 37 b) muestra el SDS-PAGE que se corrió de las muestras obtenidas de la elución en donde se observó absorbancia. En esta purificación, el pico de elución fue mucho más bajo que en la previa y en el gel se puede constatar que hay una menor presencia de bandas correspondientes a impurezas constatando la diferencia en la absorbancia de los picos de elución. Así mismo, nuestra proteína parece tener una buena afinidad a la fase estacionaria con Ni(II) y no parece haber pérdida alguna en la fase no unida.

Debido a los resultados obtenidos en la primera ronda de purificación de las proteínas con el péptido señal de TorA, para la proteína CgR-SmbP-GFP se emplearon iones Ni(II) en vez de Cu(II) para la purificación IMAC en el FPLC. En la figura 38 a) se muestra el cromatograma obtenido de la purificación de dicha proteína.

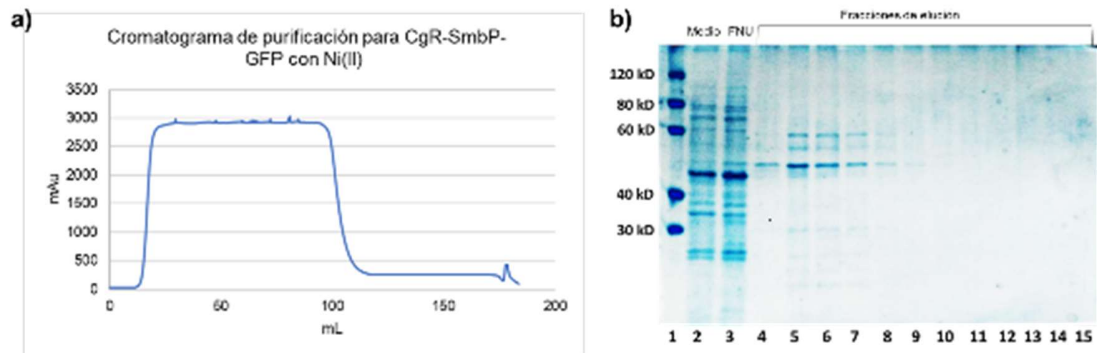


Figura 38. Purificación de CgR-SmbP-GFP por medio del FPLC ÄKTA prime plus usando Ni(II). a) cromatograma, b) SDS-PAGE de las fracciones de elución recolectadas.

En la figura 38 b) se muestra el SDS-PAGE de las fracciones obtenidas de la elución. A diferencia de su contra parte, TorA-SmbP-GFP, el pico de elución dio una absorbancia bastante baja. En los geles se observan por ende las bandas con baja intensidad. Aunque la pureza se ha mejorado, en comparación de la purificación realizada con Cu(II) la cantidad de proteína es bastante baja.

De la misma manera en que TorA-CusF3H+-GFP fue purificada, se purificó también CgR-CusF3H+-GFP. La figura 39 a) muestra el cromatograma obtenido de la purificación mostrando un pico de absorbancia ligeramente más alto que CgR-SmbP-GFP, pero más bajo que su contraparte etiquetada con TorA.

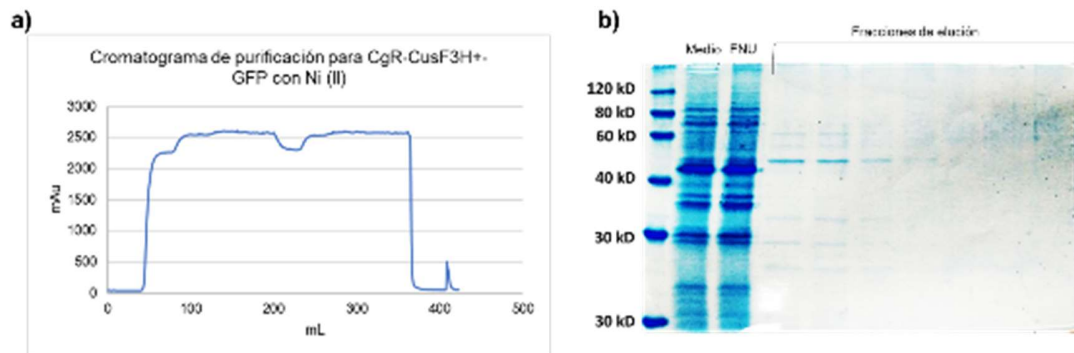


Figura 39. Purificación de CgR-CusF3H+-GFP con FPLC ÄKTA prime plus empleando Ni (II). a) cromatograma, b) SDS-PAGE de las fracciones de elución recolectadas

La figura 39 b) muestra el SDS-PAGE de la cromatografía realizada. Cabe señalar que, debido a la mayor definición y calidad del pico, la cantidad de fracciones obtenida fue mucho menor. El gel muestra las bandas que aparecieron en cada una de las fracciones obtenidas, donde se aprecia que la banda de interés posee una mayor intensidad mostrando que la proteína se unió de manera eficiente a la columna, aunque se existe una ligera presencia de impurezas en las fracciones recolectadas. Cabe señalar que, en las cuatro purificaciones, la proteína, cuya secreción es inducida en conjunto no parece tener una alta afinidad por ninguno de los dos metales.

5.8 Construcción de pTCueO y pTYahl

En la figura 40 se observa la reacción de amplificación obtenida para cada uno de los fragmentos que componen los dos diferentes insertos para los plásmidos derivados del pTarget. Cada uno de los fragmentos corresponde a los pesos moleculares esperados, siendo para los dos N20-sgARN el peso de 126 bp, para los brazos de homología 390 bp, (excepto para el brazo H1 de Yahl, el cual se tuvo que alargar a 540 bp debido a que la amplificación mostraba baja especificidad), para los 3'-UTR de 138 bp y para el donador (SmbP-GFP) de 1069 bp.

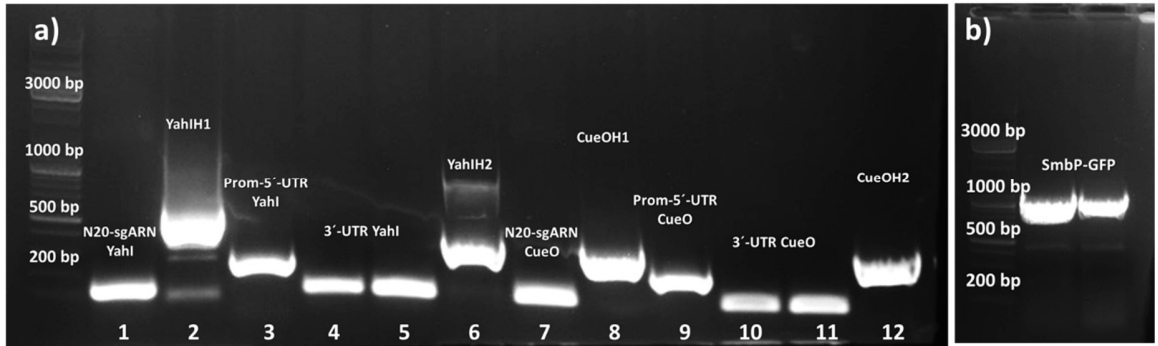


Figura 40. Reacciones de amplificación para los fragmentos de los insertos a) fragmentos específicos para Yahl y CueO, b) amplificación de SmbP-GFP

A partir de los productos de amplificación purificados de cada fragmento se realizó la PCR de alineamiento para obtener dos fragmentos conformados de tres partes y el inserto completo. En la figura 41 a) se observa la amplificación obtenida para los fragmentos de Yahl y CueO, sin embargo, debido a la baja calidad de los productos obtenidos para Yahl se cambiaron las condiciones de amplificación, el resultado se puede observar en la figura 41 b). El producto de los seis fragmentos de CueO de la primera amplificación, y de Yahl de la segunda amplificación se purificaron por gel y se insertaron en el pTarget.

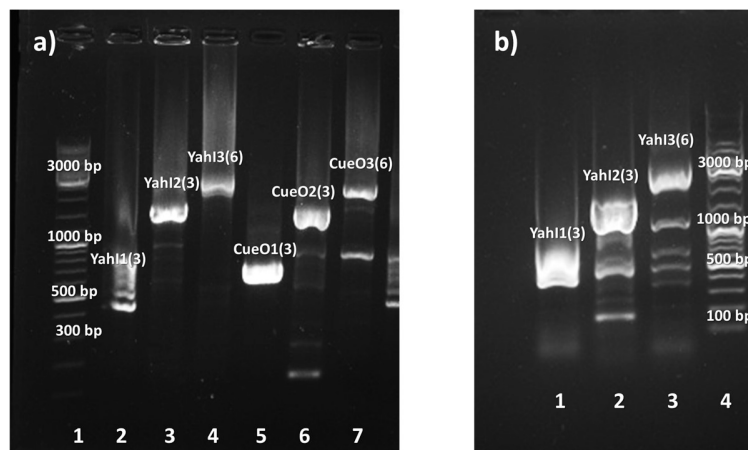


Figura 41. a) Amplificación de alineamiento para Yahl y CueO, b) repetición de amplificación para Yahl

Para realizar un tamizaje de las clonas positivas se realizó un análisis de restricción preliminar a partir de la selección de tres clonas que presentasen la coloración

característica de GFP. En la figura 42 a) se observa el análisis de restricción para el plásmido pTCueO (carriles 3,4 y 5) digerido con las enzimas *Bam*HI y *Sall*, con tres fragmentos esperados en 1962 bp, 1174 bp y 1027 bp. En la figura 42 b) se presenta el patrón de fragmentos de la digestión de pTYahl con las enzimas *Bam*HI y *Sac*I (carriles 3, 4 y5), los pesos esperados son de 2161 bp, 1177 bp y 975 bp. El ADN plasmídico de las clonas seleccionadas fue secuenciado para confirmar la clonación y asegurar la integridad del N20 de cada uno.

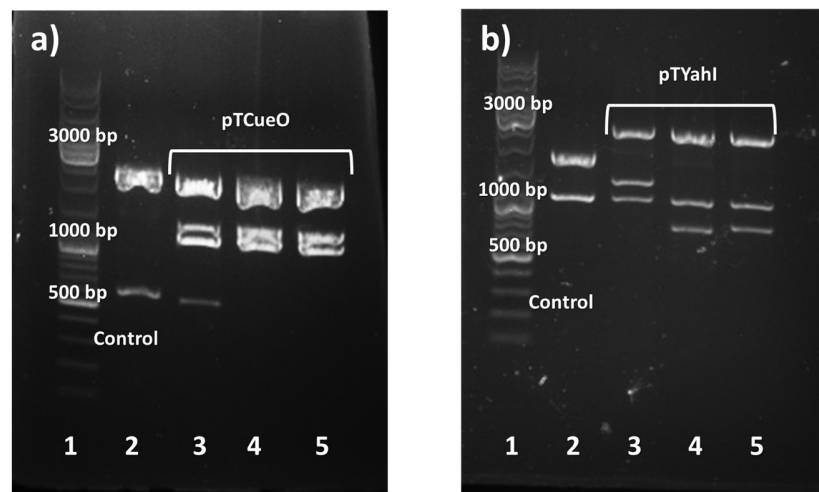


Figura 42. a) Análisis de restricción para pTCueO con enzimas *Bam*HI y *Sall*, b) análisis de restricción para pTYahl con enzimas *Bam*HI y *Sac*I

5.9 Edición genómica de *E. coli* MG1655 y BL21(DE3)

Para la confirmación de edición genómica en ambas líneas se realizó un tamizaje preliminar con los iniciadores especiales diseñados para amplificar río arriba al primer brazo de homología. En la imagen 43 se observa el resultado de la amplificación de diez clonas tomadas para cada línea, con pesos esperados para esta amplificación de 1092 bp. Una segunda amplificación fue realizada con un segundo iniciador delantero, en este caso debiendo resultar productos de 964 bp, en la imagen 44 se observa el resultado de este segundo tamizaje. Para ambas amplificaciones se corrieron tres controles, el

primer control fue una clona de la línea MG1655 transformada con pCassac (Ctrl1), el segundo consistió en la amplificación de 10 ng del plásmido pTCueO (Ctrl 2), el tercero fue la amplificación de una clona de BL21(DE3) transformada con pCassac (Ctrl3).

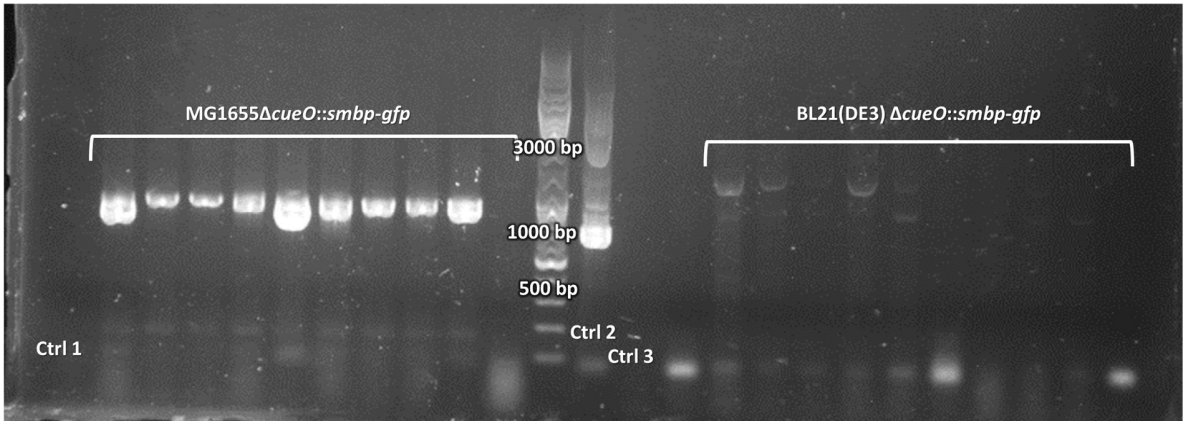


Figura 43. Tamizaje con primer set de iniciadores con productos esperados de 1092 bp, Ctrl1: MG1655+pCassac, Ctrl2: pTCueO, Ctrl3: BL21(DE3)+pCassac

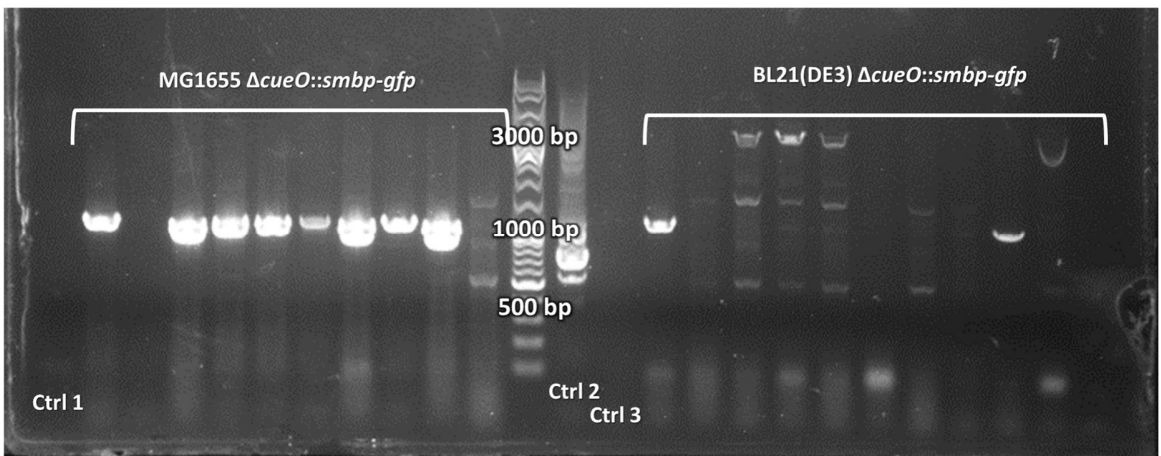


Figura 44. Segundo tamizaje con el segundo set de iniciadores con productos esperados en 964 bp, Ctrl1: MG1655+pCassac, Ctrl2: pTCueO, Ctrl3: BL21(DE3)+pCassac

En la figura 45 se observan las amplificaciones realizadas de dos clonas seleccionadas para MG1655 y cuatro de BL21(DE3) a partir de los resultados previos. En la figura 45 a) se muestra la amplificación con el primer set de iniciadores con pesos moleculares

esperados de 1092 bp, en la 45 b) con el segundo set con esos esperados de 964 bp, en la 45 c) con un tercer set de iniciadores con productos esperados en 849 bp. El control 1 (Ctrl1) corresponde a la amplificación hecha con los iniciadores respectivos a 10 ng de pTCueO, en el control 2 (Ctrl2) y el control 3 (Ctrl3) se emplearon para la amplificación las líneas MG1566 y BL21(DE3) respectivamente, transformadas con pCassac.

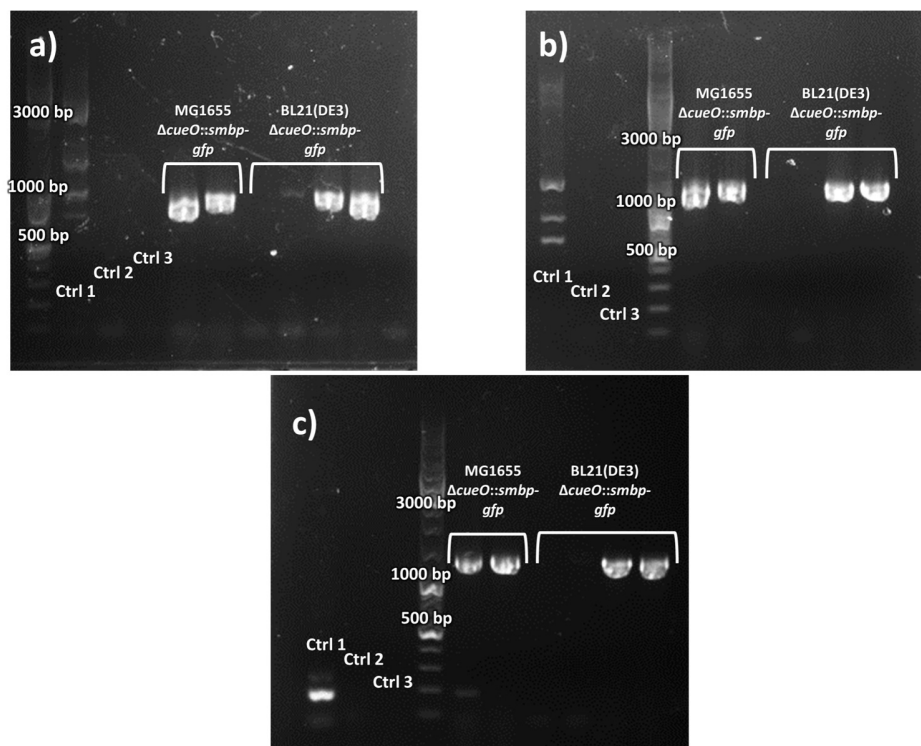


Figura 45. Amplificación realizada con los tres conjuntos de iniciadores de las cepas mutantes MG1655 $\Delta cueO::smbp-gfp$ y BL21(DE3) $\Delta cueO::smbp-gfp$ a) amplificación con primer conjunto, peso esperado de 1092 bp, b) amplificación con segundo conjunto, peso esperado de 964 bp, c) amplificación con tercer conjunto, peso esperado de 849 bp

CAPÍTULO 6

6. Discusión de resultados

La generación del mutante se realizó mediante mutagénesis sitio dirigida asistida por reacción en cadena de la polimerasa. El diseño de los iniciadores incluyó los codones CAC para permitir una distinción de la polimerasa entre los CAT originales del gen. En la figura 35 se muestra la primera parte de la secuenciación que se realizó de los plásmidos pET30a-CusF3H+ empleando los iniciadores del plásmido pET30 T7. Se observa claramente como la amplificación del gen con los iniciadores diseñados se alinearon de la manera esperada generando estos tres codones extras en 5' del gen además de la metionina. El diseño de los iniciadores también permitió la remoción de los residuos Ala, Asn y Glu al inicio del péptido maduro después del péptido señal. Quedando de esta manera modificado el extremo N terminal de la proteína, este cambio no compromete el sitio de unión de Cu(I) y Ag(I) y tampoco ninguna parte de la estructura de CusF3H+ ya que se encuentra fuera de la estructura terciaria de la misma.⁵⁹

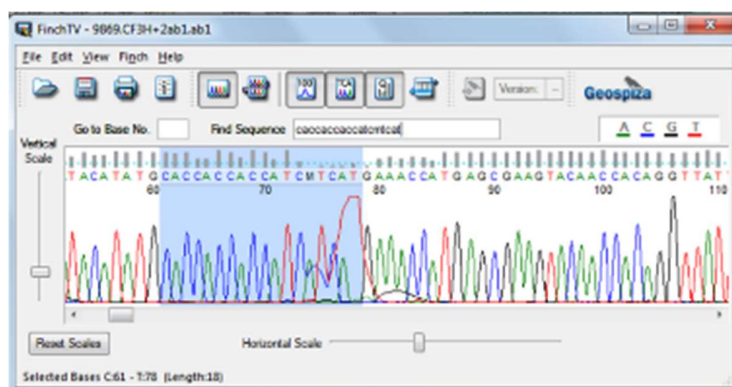


Figura 46. Secuenciación de pET30a-CusF3H+, la parte sombreada muestra los codones originales de CusF para His y los que han sido añadidos mediante la mutagénesis dirigida.

Con el gen mutante obtenido se realizaron las construcciones necesarias para probar la afinidad de la nueva proteína a iones de Ni(II) para su uso en cromatografía tipo IMAC. La primera evaluación de la capacidad de expresión de proteínas recombinantes

marcadas con CusF3H+ fue con la proteína GFP, se realizó una expresión piloto empleando IPTG 1 mM como inductor. En la figura 6 se observa la expresión de la proteína quimérica CusF3H+-GFP, el color característico de la proteína es claramente apreciable. Esto nos indica que la proteína se está expresando adecuadamente y el plegamiento de la misma es el correcto, ya que en caso de haber una falla en el plegamiento la proteína carecería de coloración.⁶⁰ Así mismo, la mutación aplicada no interfiere en la capacidad de CusF como proteína potenciadora de la expresión característica de la misma encontrada en estudios previos.¹⁸

La expresión de G-CSF (inducida también con IPTG 1 mM) fue visualizada mediante SDS-PAGE. G-CSF marcada con His-tag y CusF3H+ fue expresada a escala piloto con fines comparativos. En la figura 7 se puede apreciar el SDS-PAGE mencionado. En los cuatro carriles se aprecia claramente la banda de expresión correspondiente a la proteína quimérica, mostrando un alto rendimiento en los cuatro casos. Cabe mencionar que el gen empleado de G-CSF es un gen optimizado para la expresión en *E. coli*, por lo que en el caso de CusF3H+ existe una sinergia entre la capacidad de potenciar la expresión de la proteína de fusión y como la capacidad inherente de G-CSF optimizada. En los cuatro casos se puede observar también una muy baja formación de cuerpos de inclusión, siendo la proteína soluble recombinante la que aporta en mayor proporción al rendimiento final.

Debido a que el gen G-CSF fue optimizado para su expresión en *E. coli* se puede observar un excelente rendimiento y solubilidad. Estos resultados difieren con los obtenidos por Do, B. y col., en el 2014 donde G-CSF marcado con His6 se expresó de manera insoluble bajo todas las condiciones probadas.⁵⁴ Estudios donde se produce G-CSF con diferentes proteínas de fusión y condiciones, por lo regular tienen la dificultad de que se forman cuerpos de inclusión. Muchos protocolos incluyen un paso extra que

consiste en la resolubilización de G-CSF.⁶¹ Marcar G-CSF con la proteína de fusión CusF3H+ es una buena manera de hacer frente a dicho problema. Ya que como se observó en la expresión piloto, CusF3H+ produce G-CSF de forma soluble.

La principal desventaja que poseía el antecesor de CusF3H+, CusF era su baja afinidad a Ni(II) mostrando una mayor afinidad por Cu(II).¹⁸ Sin embargo, al haber un mayor rango de proteínas en la matriz citoplásmica de *E. coli* afines también a Cu(II) la pureza se veía considerablemente disminuida.⁶² La mutación dirigida sobre CusF para agregar las tres histidinas adicionales fue con dicho fin, es decir, aumentar sustancialmente la afinidad de CusF3H+ por iones Ni(II). Se sabe que la histidina juega un papel fundamental en dicha unión y es necesario que se encuentren disponibles en su posición para permitirla.³⁶

Comprobada la unión de CusF3H+-GFP a la fase estacionaria unida a iones Ni(II), se realizó previamente una purificación empleando la proteína quimérica CusFc-GFP. Esto con el fin de establecer las condiciones de purificación adecuadas para CusF3H+ considerando que esta proteína tiene una afinidad prácticamente nula a Ni(II). En la purificación se pudo comprobar que CusFc tiene una afinidad mínima a Ni(II) y la mayoría de la proteína fue perdida en la fase no unida (la cual se recolectó en fracciones), gracias al color de GFP la pérdida fue visible y corroborada después en gel de poliacrilamida (figura 9). En la figura 8 se observa el cromatograma obtenido de la misma, donde el máximo del pico de elución fue de apenas 600 mAu a diferencia de la pasada donde supero los 1000 mAu. En las fracciones de elución observadas en la figura 10 se aprecia la banda correspondiente a la proteína quimérica, sin embargo, no se observa una mejoría concreta en la pureza de la proteína obtenida, de hecho, disminuyó a ~28.5%. Esto comprueba la imposibilidad de purificar proteínas marcadas con CusFc empleando Ni(II).

Obtenidos los resultados anteriores, se estableció el uso de una concentración de 10 mM de imidazol para la solución amortiguadora de equilibrio y de lavado. Ya que a partir de esa concentración la proteína CusFc-GFP era separada de la fase estacionaria. Se realizó una posterior purificación de CusF3H+-GFP usando la concentración establecida para la solución de lisis, equilibrio y lavado no se observó pérdida alguna de la proteína CusF3H+-GFP. En la figura 12 en el gel de poliacrilamida se observa la banda correspondiente a la proteína quimérica en las fracciones de elución. La pureza se elevó a ~ 74.0% mejorando la misma empleando CusF3H+ y las condiciones de concentración de imidazol anteriormente determinadas. El cromatograma mostrado en la figura 11 es particularmente interesante. Ya que tanto para CusFc-GFP y CusF3H+ se lisó el total de 2 L de cultivo total. En el caso de CusFc-GFP el pico solo alcanzó hasta los 600 mAu, mientras que para CusF3H+-GFP el pico máximo excede los 2200 mAu. Mostrando que no se perdió una cantidad considerable de proteína durante la carga del lisado y los lavados, por tanto, el rendimiento de purificación es mayor.

Este aumento en la pureza obtenida de la proteína quimérica en un solo paso es considerable en comparación con los resultados previos donde CusFc-GFP se purificó con Cu(II) consiguiendo ~ 50% de pureza.¹⁸ Además, el porcentaje de pureza obtenido del 74% es comparable con sistemas comerciales que usan His6. Donde en el caso de GFP llegan a presentar rendimientos de pureza de alrededor de 75% según reportan. En pocos casos se reportan purezas arriba del 80% para His6. Para fines comparativos se purificó GFP marcada también con una simple etiqueta de seis histidinas. Los resultados obtenidos de esta purificación son bastante comparables con los reportados incluso por las mismas casas comerciales, empleando las condiciones de purificación sugeridas se observa una pureza de ~ 70%. Este resultado obtenido de la purificación de His-tag comparable con lo reportado implica que la reproducibilidad de las técnicas

de purificación no se ven limitadas ni existe una variación importante, por lo que se espera lo mismo de las purificaciones realizadas con la proteína CusF3H+, aunque es necesario probar su funcionalidad para purificar en más variedad de proteínas. CusF3H+ no solo aumenta la expresión y solubilidad de la proteína de interés, sino que también permite la purificación de la misma en pocos pasos y con purezas equiparables a las de uso comercial como His6. Considerando que dichos sistemas llevan varios años de perfección y optimización, los resultados obtenidos con CusF3H+ son de alto impacto.

Para usar cualquier proteína recombinante en la mayoría de los casos es necesario remover la proteína de fusión con la que ha sido marcada. En este caso, la proteína quimérica CusF3H+-GFP obtenida de la última purificación fue cortada con la proteasa enteroquinasa que reconoce el sitio específico de Asp-Asp-Asp-Asp-Lys que se encuentra entre CusF3H+-GFP.¹⁵ Como se puede observar en la figura 13 en el carril 1 se observa la proteína quimérica, en el segundo la proteína cortada, mostrando a CusF3H+ y GFP en su peso indicado. No se observó ningún remanente de la proteína quimérica por lo que la cantidad de unidades y el tiempo empleado fue el adecuado para la cantidad de proteína usada.

Posteriormente para separar ambas proteínas de su matriz se sometieron a una segunda purificación empleando una resina tipo IMAC cargada con iones Ni(II). Como fue de esperarse, GFP no se unió a la resina y después de la incubación por una hora permaneció en el sobrenadante. Las fracciones recuperadas se aprecian en los carriles 4 y 5 donde se observa la proteína obtenida con una elevada pureza ya que la única proteína que no era afín a Ni(II) en la matriz era la proteína de interés. Se eluyó finalmente con imidazol para liberar a CusF3H+ de la resina y en el carril 6 se observa claramente la proteína CusF3H+ recuperada y coincidiendo en peso molecular.¹⁵ Es

importante remarcar que en las fracciones de la proteína de interés (GFP) que fueron recuperadas, el color típico de la proteína se mantuvo. Demostrando que el corte y remoción de la proteína de fusión no afecta la estabilidad de la proteína de interés.

La cantidad obtenida por L de cultivo, para el caso de CusF3H+-GFP fue de 6 mg. Considerando el porcentaje que representa en la proteína quimérica CusF3H+ se puede calcular que aproximadamente 1.6 mg corresponden únicamente a nuestra proteína de fusión, dejando 4.4 mg restantes que corresponden a la proteína de interés. Comparando los resultados obtenidos con otras proteínas de interés con pesos moleculares más altos como MBP y GST, se puede observar como el pequeño peso molecular de la proteína CusF3H+ ayuda a mejorar enormemente los rendimientos finalmente obtenidos para la proteína de interés.⁴

Para la expresión de proteínas en *C. glutamicum* se empleó un medio definido modificado de Jae Woong⁵⁷. Este medio es bastante parecido en su composición con otros reportados en la literatura y es bastante importante recalcar el uso de sales en dicho medio. Al principio, después de la transformación de las colonias se realizaron algunas expresiones de prueba en el medio definido y en LB. La diferencia entre ambos cultivos fue bastante explícita, con el medio definido se obtenía una alta cantidad de biomasa a las 48 h de cultivo a 30 °C y 200 rpm, mientras que con LB, bajo las mismas condiciones la cantidad de biomasa era mínima.

Al observar que la cantidad de biomasa era mayor con el medio definido se realizaron las primeras expresiones en matraces de 500 mL con 100 mL de dicho medio y suplementados con kanamicina (25 µg/mL). Sin embargo, después de las 48 h de expresión no se observaba diferencia visual significativa entre ambos cultivos, incluso entre el medio de cultivo y la biomasa. Al analizar el medio después de las 48 h en un transiluminador UV no se observaba fluorescencia alguna, tanto en los cultivos de las

cepas transformadas como en el del control (una cepa de *C. glutamicum* sin transformar). En los geles de SDS-PAGE tanto de lisado celular como de medio de cultivo tampoco se observaba una diferencia significativa con los controles.

Teramoto en el 2011 probó el péptido señal CgR_0949 para la secreción de proteínas en *C. glutamicum* mediante la vía Tat.²⁸ En dicha publicación se empleó la proteína reportera GFP y el medio empleado es bastante similar en composición al usado en el presente proyecto exceptuando por las vitaminas. Aunque, Teramoto describe en el mismo la importancia de emplear Ca(II) para la secreción de GFP. Por lo que se modificó el medio para incluir una alta cantidad de Ca(II), Teramoto describe una concentración final de CaCl₂ de 2.0 g/L. Sin embargo, en la preparación de medios, existía una precipitación de la sal por interacción del MgSO₄ el cual se encontraba también en una concentración de 2 g/L. La cantidad máxima de CaCl₂ que se pudo añadir al medio de cultivo sin precipitación fue hasta una concentración de 200 mg/L, apenas una décima parte de lo que describe Teramoto.

La primera proteína en ser expresada con el nuevo medio fue TorA-SmbP-GFP, la cual se expresó a las mismas condiciones antes probadas, pero ahora con matraces triptinizados. Después de las 48 h de expresión se recolectaron alícuotas de 2 mL de cultivo y se centrifugaron. Al examinarlas tras el transiluminador se observó una ligera fluorescencia verde en las alícuotas con inducción por IPTG. Se observó una clara diferencia con los controles, tanto de *C. glutamicum* sin inducir como de *C. glutamicum* transformada, pero sin inducir, ya que estos dos controles no poseían fluorescencia alguna. El experimento se repitió, empleando el medio sin sales y 1 g/L de CaCl₂ y por otra parte el medio definido sin CaCl₂ (como se hacía en un principio). En el primer caso la biomasa obtenida fue bastante baja y en el segundo la biomasa no se comprometió, pero no existía fluorescencia como se observó en los primeros experimentos. Estos

resultados indican la importancia del Ca(II) para la inducción de la expresión y secreción de GFP. Cabe mencionar que la biomasa no presentaba fluorescencia bajo el transiluminador.

El experimento se repitió con las cuatro construcciones y se probaron dos diferentes densidades ópticas (0.6 y 1.0) para establecer la óptima condición para la inducción. En el caso de las proteínas marcadas con SmbP, se observó para la proteína con el péptido de TorA que la inducción tanto a 0.6 como a 1.0 no parecía marcar una importante diferencia entre ambas como se observa en la figura. En el caso de la marcada con CgR se aprecia una intensidad de la banda esperada de mayor intensidad a la densidad óptica de 1.0. Para GFP marcada con CusF3H+ se observó en el caso del péptido señal de TorA una producción mayor de la expresión a una densidad de 0.6 mostrando una clara diferencia con su contraparte inducida a 1.0. Es importante añadir que la inducción de la expresión con IPTG no solo desencadena la expresión de la proteína deseada, sino que también una importante presencia de bandas extra se observa. Es posible que la inducción promueve la secreción de otras proteínas inespecíficas al medio de cultivo. Sin embargo, la cantidad de dichas bandas es relativamente baja en comparación con las presentadas en el lisado celular, considerando también que en los carriles del lisado se agregaron únicamente 4 µL de muestra, mientras que en las de medio se cargaron 15 µL para *C. glutamicum* sin transformar y 10 µL para cada una de las expresiones. Las bandas obtenidas del lisado celular completo después de la inducción no presentan diferencia significativa aun en comparación con ambos controles. En los resultados de Teramoto también se presentaron las proteínas extras en el sobrenadante que aparentemente su secreción es potenciada por la presencia de Ca(II) así que esto puede presentar cierta desventaja.

En el caso de las proteínas marcadas con CusF, se observaron comportamientos similares a las marcadas con SmbP, parece ser que la proteína de fusión empleada no parece influir en la cantidad de proteína que es secretada al medio. En el caso de la cepa transformada con el plásmido pz8-SmbP-CgR, se observa que la secreción de proteínas se ve favorecida cuando se induce a 1.0 de O.D.₆₀₀, aunque en ambas O.D.₆₀₀ se aprecia la secreción de estas. En los lisados celulares no se observa diferencia, tanto en los sedimentos de los cultivos inducidos y los controles empleados. En el caso de la secreción de CusF3H+-CgR-GFP, se observa de una manera más clara la tendencia al patrón de proteínas secretadas reportadas por Teramoto, en donde la presencia de la proteína de aproximadamente 50 kDa sobresale de entre las demás y la banda correspondiente a nuestra proteína es bastante tenue a comparación de las pasadas secreciones, además, no se observa diferencia alguna entre la inducción a 0.6 y a 1.0 de O.D.₆₀₀.

Una vez recolectado el sobrenadante de cada una se procedió a la purificación. Se recolectaron diversos volúmenes de expresión de cada una de las proteínas recombinantes. Para la purificación de TorA-SmbP-GFP se empleó una columna cargada con Cu(II), el cromatograma mostró un patrón típico con un pico de elución bastante alto superando los 2000 mAu. Sin embargo, las fracciones recolectadas mostraban que, aunque nuestra proteína se había unido a la resina, también otras proteínas contaminantes. Tora-CusF3H+-GFP se purificó empleando una columna cargada con Ni(II), a diferencia de la primera, el pico de elución solo alcanzó los 1000 mAu, sin embargo, el análisis con SDS-PAGE reveló una mayor pureza, por tanto, la cantidad de proteína recuperada fue menor (por eso se presentó menos absorbancia) pero a su vez, era más pura.

Debido a los resultados obtenidos con TorA-SmbP-GFP, CgR-SmbP-GFP se purificó empleando una columna cargada con Ni(II). El cromatograma también presentó la forma característica esperada, aunque el pico de elución fue bastante bajo alcanzando apenas los 500 mAu y con una pobre resolución. Los geles de SDS-PAGE revelan la baja concentración de proteínas presentes en las fracciones de elución, aunque con una pureza considerable. Con la misma metodología se purificó CgR-CusF3H+-GFP, el patrón fue el esperado, con un pico de elución también bajo, aunque con una buena resolución. El SDS-PAGE de las fracciones de elución revela también la baja concentración de proteínas recolectadas. En las cuatro purificaciones, la proteína GFP marcada tanto con SmbP y CusF3H+ se unió a la fase estacionaria, además de que se logró concentrar y no se observa presencia alguna de la proteína co-secretada con el Ca(II). Lo que resulta también evidente es que cuando las proteínas son marcadas con el péptido señal de TorA parece favorecer la secreción de proteínas al medio.

Sin embargo, es necesario realizar la cuantificación final después de la purificación con el fin de establecer una clara diferencia en el uso de TorA y CgR_0949 que podría ser una realidad, como en los estudios de secreción de GFP también en *C. glutamicum* de Meissner D, y col., en el 2007, donde encontraron una clara diferencia entre el tipo de péptido señal que usaron para la secreción.²⁹ Hasta el momento de búsqueda bibliográfica, solo Meissner y Teramoto han exportado la proteína de GFP usando *C. glutamicum*. En ambos casos se reporta la exportación de la proteína de manera activa midiendo la fluorescencia de la misma, curiosamente en ninguno de ellos se hace mención sobre la presencia de la coloración de la proteína y la cantidad de la proteína obtenida es también baja.^{28,29} Probablemente, esta sea la principal desventaja de los sistemas de expresión/secreción en *C. glutamicum*, la baja cantidad de proteína obtenida, dato que es recalado en los estudios de Date, M. y col., en el 2006 e Itaya,

H. y col., en el 2008 al exportar la proteína (hFCE) (factor de crecimiento epidérmico humano) y una pro-transglutaminasa respectivamente. En dichos estudios la cantidad de proteína obtenida fue baja, siendo en el primer caso de apenas 6.5 mg/L por h de fermentación para el caso del hFCE y de 25 g por L durante siete días de fermentación para la pro-transglutaminasa.^{10,63}

La edición genómica de *E. coli* mediante la técnica de CRISPR/Cas9 desarrollada por Yang y col., en el 2016⁵³ propone el uso de dos plásmidos para realizar el corte y edición. Uno con el gen de Cas9 y la maquinaria de λ Red y un segundo plásmido con el ADN donador, los brazos de homología y sgARN-N20, siendo derivados del plásmido pTarget. Se construyeron dos plásmidos para la delección de los genes *cueO* y *yahI* con reemplazo del ADN donador que consistía en *smbp-gfp* junto con el promotor y 5'-UTR/3'-UTR de la proteína de choque frío CspA. La construcción se realizó mediante reacción de amplificación de alineamiento generando en la primera amplificación los fragmentos deseados junto con +20 bp río arriba y río abajo para permitir el ensamble en la reacción de alineamiento. En dicha reacción se realizan de 5 a 10 ciclos sin iniciadores con el fin de permitir el alineamiento de los fragmentos mediante los +20 bp en 5' y 3' de cada fragmento. Los productos de amplificación de dicha reacción se pueden observar en las figuras 41 a) y b) con los pesos esperados para cada uno. El fragmento final con el ensamble de las seis diferentes partes fue ligado dentro del plásmido pCB003 (pTarget) para originar los plásmidos pTCueO y pTYahI. La edición genómica se realizó induciendo los genes del sistema λ Red del plásmido pCassac mediante arabinosa, el gen de Cas9 se encuentra bajo un promotor sintético constitutivo, por lo cual la expresión de Cas9 es constante. Después de la activación es necesario transformar la línea con el segundo plásmido, en este caso el pTCueO y pTYahI que poseen el sgARN junto con el N20 necesario para direccionar el corte de Cas9, por

tanto, la generación de DSB (*double stranded breaks*) no se dará hasta que este segundo plásmido se encuentre dentro de la línea.⁵³ Mediante incubación se permite que todo el sistema en conjunto de lugar a la edición genómica del sitio elegido y se lleve a cabo la recombinación homóloga para la reparación del DSB generado permitiendo así la integración del ADN donador. Después de la clonación del segundo plásmido, las cepas se guardaron durante 16 h a 4 °C, después de este periodo la mayoría de las clonas obtenidas (+100) se habían tornado de color verde en el caso de la línea MG1655. En el caso de BL21(DE3) se obtuvieron también una alta cantidad de clonas, sin embargo, era notable que se habían generado dos distintas colonias con tamaños diferentes. Después de las 16 h a 4 °C fue interesante notar que las colonias más pequeñas fueron las únicas que se tornaron verdes. Incluso, después del cribado de amplificación, solamente las amplificaciones de las colonias pequeñas verdes dieron positivo en los tres conjuntos de iniciadores probados a diferencia de los resultados obtenidos de las colonias de mayor tamaño sin coloración, en donde solo dos dieron un resultado positivo en el cribado preliminar con el segundo conjunto de iniciadores, sin embargo, en la segunda amplificación con los tres conjuntos ninguna dio resultado positivo. La coloración verde de las colonias puede deberse tanto por la presencia del pTCueO como por la edición genómica realizada, por lo que es necesario realizar el cribado. Cabe señalar que el lento crecimiento y pequeño tamaño de las colonias de *E. coli* obtenidas pueden ser señal de que el DSB y edición se ha llevado a cabo, ya que esto puede estresar a la bacteria al tener que reponerse primero del DSB para asegurar su supervivencia comprometiendo así su velocidad de crecimiento.

La edición en el sitio del gen *yahI* se encuentra pendiente, ya que el plásmido generado poseía una mutación en el N20 impidiendo así el posible direccionamiento de Cas9. Se

realizó una mutagénesis dirigida amplificando todo el plásmido y digiriendo el original con *DpnI* para la generación de un nuevo plásmido corregido.

CAPÍTULO 7

7. Conclusiones

- Por medio de mutagénesis dirigida se obtuvo el mutante deseado de CusF al que se añadieron tres histidinas en el N-terminal, dicho mutante se nombró CusF3H+, este mutante es capaz de aumentar los niveles de expresión de la proteína de interés además presenta una baja formación de cuerpos de inclusión como se observa en la fracción insoluble.
- CusF3H+ a diferencia de CusF permite una purificación con IMAC usando columnas con Ni(II), se obtiene una mayor pureza como se puede observar en los geles de SDS-PAGE de las fracciones obtenidas.
- La capacidad de CusF de obtener buenos resultados de expresión y solubilidad no se ve comprometida en su variante CusF3H+ y la pureza final se ve altamente mejorada.
- La proteína quimérica puede separarse fácilmente empleando la enzima enteroquinasa de cadena ligera sin afectar el plegamiento de la proteína de interés, la cual se puede recuperar con una alta pureza (>~90%) aprovechando su baja afinidad a Ni(II) empleando una segunda purificación con cromatografía tipo IMAC.
- Tanto CgR como TorA son capaces de permitir la secreción de proteínas al medio, sin embargo, se requieren de condiciones específicas para que la secreción se pueda llevar a cabo, por ejemplo, en el caso de GFP el uso de calcio juega un papel importante, la presencia de sales en el medio definido es fundamental para la producción eficiente de biomasa.

- La construcción de los plásmidos pTCueO y pTYahl fue posible mediante reacción en cadena de la polimerasa con alineamiento de los fragmentos.
- El plásmido pTCueO al contener el gen de SmbP-GFP regulado bajo el promotor y 5'-UTR/3'-UTR de la proteína de choque frío CspA permite la expresión de la proteína al ser sometida a temperaturas menores de 25 °C.
- El plásmido pTCueO permite el direccionamiento de la proteína Cas9 en el sitio elegido de corte, así mismo gracias a la maquinaria λ Red del pCassac es posible permitir la recombinación después del DSB generado.
- La supervivencia de *E. coli* después de la edición es alta, debido al uso del plásmido pCassac (que es una versión mejorada de pCas9 pendiente de publicación) aunque la velocidad de crecimiento de la línea después de la edición parece estar comprometida.

CAPÍTULO 8

8. Seguridad en el trabajo de disposición de residuos

Los residuos generados durante la realización de este proyecto se dispusieron de acuerdo al Procedimiento para la Disposición Final de Residuos Peligrosos de la FCQ-UANL.

Referencias

1. Baneyx F. Recombinant protein expression in *Escherichia coli*. *Curr Opin Biotechnol*. 1999;10(5):411-421. doi:S0958-1669(99)00003-8 [pii].
2. Quianzon CC, Cheikh I. History of insulin. *J community Hosp Intern Med Perspect*. 2012;2(2). doi:10.3402/jchimp.v2i2.18701.
3. Terpe K. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol*. 2003;60(5):523-533. doi:10.1007/s00253-002-1158-6.
4. Rosano GL, Ceccarelli EA. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front Microbiol*. 2014;5:172. doi:10.3389/fmicb.2014.00172.
5. Gopal GJ, Kumar A. Strategies for the Production of Recombinant Protein in *Escherichia coli*. *Protein J*. 2013;32(6):419-425. doi:10.1007/s10930-013-9502-5.
6. Adrio J-L, Demain AL. Recombinant organisms for production of industrial products. *Bioeng Bugs*. 2010;1(2):116-131. doi:10.4161/bbug.1.2.10484.
7. Chen R. Bacterial expression systems for recombinant protein production: *E. coli* and beyond. *Biotechnol Adv*. 2012;30(5):1102-1107. doi:10.1016/j.biotechadv.2011.09.013.
8. Demain AL, Vaishnav P. Production of recombinant proteins by microbes and higher organisms. *Biotechnol Adv*. 2009;27(3):297-306. doi:10.1016/j.biotechadv.2009.01.008.
9. Schumann W, Ferreira LCS. Production of recombinant proteins in *Escherichia coli*. *Genet Mol Biol*. 2004;27(3):442-453. doi:10.1590/S1415-47572004000300022.
10. Itaya H, Kikuchi Y. Secretion of *Streptomyces mobaraensis* pro-transglutaminase by coryneform bacteria. *Appl Microbiol Biotechnol*. 2008;78(4):621-625. doi:10.1007/s00253-007-1340-y.
11. Kikuchi Y, Itaya H, Date M, Matsui K, Wu LF. TatABC overexpression improves *Corynebacterium glutamicum* Tat-dependent protein secretion. *Appl Environ Microbiol*. 2009;75(3):603-607. doi:10.1128/AEM.01874-08.
12. Lee J. Development and characterization of expression vectors for *Corynebacterium glutamicum*. *J Microbiol Biotechnol*. 2014;24(1):70-79. doi:10.4014/jmb.1310.10032.
13. Kimple ME, Brill AL, Pasker RL. Overview of affinity tags for protein purification. *Curr Protoc Protein Sci*. 2013;(SUPPL.73):608-616. doi:10.1002/0471140864.ps0909s73.

14. Zhao X, Li G, Liang S. Several affinity tags commonly used in chromatographic purification. *J Anal Methods Chem*. 2013;2013(Table 1):1-8. doi:10.1155/2013/581093.
15. Waugh DS. An Overview of Enzymatic Reagents for the Removal of Affinity Tags. *Protein Expr Purif*. 2011;72(2):181-204. doi:10.1038/nature13314.A.
16. Barney BM, LoBrutto R, Francisco W a. Characterization of a small metal binding protein from *Nitrosomonas europaea*. *Biochemistry*. 2004;43(35):11206-11213. doi:10.1021/bi049318k.
17. Vargas-Cortez T, Morones-Ramirez JR, Balderas-Renteria I, Zarate X. Expression and purification of recombinant proteins in *Escherichia coli* tagged with a small metal-binding protein from *Nitrosomonas europaea*. *Protein Expr Purif*. 2016;118:49-54. doi:10.1016/j.pep.2015.10.009.
18. Cantu-Bustos JE, Vargas-Cortez T, Morones-Ramirez JR, et al. Expression and purification of recombinant proteins in *Escherichia coli* tagged with the metal-binding protein CusF. *Protein Expr Purif*. 2016;121:61-65. doi:10.1016/j.pep.2016.01.007.
19. Goldstein J, Stephen Pollitt N, Inouyet M. Major cold shock protein of *Escherichia coli*. *Biochemistry*. 1990;87:283-287.
20. Makrides SC. Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol Rev*. 1996;60(3):512-538. <http://www.ncbi.nlm.nih.gov/pubmed/8840785>. Accessed February 9, 2017.
21. Vasina JA, Franç A, Baneyx O. Recombinant Protein Expression at Low Temperatures under the Transcriptional Control of the Major *Escherichia coli* Cold Shock Promoter *cspA*. *Appl Environ Microbiol*. 1996;62(4):1444-1447.
22. Vasina JA, Peterson MS, Baneyx F. Scale-up and optimization of the low-temperature inducible *cspA* promoter system. *Biotechnol Prog*. 1998;14(5):714-721. doi:10.1021/bp980061p.
23. Lee PA, Tullman-Ercek D, Georgiou G. The bacterial twin-arginine translocation pathway. *Annu Rev Microbiol*. 2006;60:373-395. doi:10.1146/annurev.micro.60.080805.142212.
24. Cristó S, De Gier J-W, Nielsen H, Von Heijne G. Competition between Sec- and TAT-dependent protein translocation in *Escherichia coli*. *EMBO J*. 1999;18(11):2982-2990.
25. Fröbel J, Rose P, Müller M. Twin-arginine-dependent translocation of folded proteins.

- Philos Trans R Soc Lond B Biol Sci.* 2012;367(1592):1029-1046.
doi:10.1098/rstb.2011.0202.
26. Goosens VJ, Monteferrante CG, Maarten J, Dijn V. The Tat system of Gram-positive bacteria. *BBA - Mol Cell Res.* 2014;1843:1698-1706. doi:10.1016/j.bbamcr.2013.10.008.
 27. Kikuchi Y, Date M, Itaya H, Matsui K, Wu L-F. Functional analysis of the twin-arginine translocation pathway in *Corynebacterium glutamicum* ATCC 13869. *Appl Environ Microbiol.* 2006;72(11):7183-7192. doi:10.1128/AEM.01528-06.
 28. Teramoto H, Watanabe K, Suzuki N, Inui M, Yukawa H. High yield secretion of heterologous proteins in *Corynebacterium glutamicum* using its own Tat-type signal sequence. *Appl Genet Mol Biotechnol.* 2011;91:667-687. doi:10.1007/s00253-011-3281-8.
 29. Meissner D, Vollstedt A, Maarten Van Dijn J, Freudl R. Comparative analysis of twin-arginine (Tat)-dependent protein secretion of a heterologous model protein (GFP) in three different Gram-positive bacteria. *Appl Genet Mol Biotechnol.* 2007;76:633-642.
 30. Tyo KEJ, Ajikumar PK, Stephanopoulos G. Stabilized gene duplication enables long-term selection-free heterologous pathway expression. *Nat Biotechnol.* 2009;27(8):760-765. doi:10.1038/nbt.1555.
 31. Striedner G, Pfaffenzeller I, Markus L, Nemecek S, Grabherr R, Bayer K. Plasmid-free T7-based *Escherichia coli* expression systems. *Biotechnol Bioeng.* 2010;105(4):786-794. doi:10.1002/bit.22598.
 32. Wenyan Jiang, David Bikard, David Cox, Feng Zhang and LAM. CRISPR-assisted editing of bacterial genomes. *Nat biotechnol.* 2013;31(3):233-239. doi:10.1038/nbt.2508.CRISPR-assisted.
 33. Hsu PD, Lander ES, Zhang F. Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell.* 2014;157(6):1262-1278. doi:10.1016/j.cell.2014.05.010.
 34. Eriksson S, Hurme R, Rhen M. Low-temperature sensors in bacteria. *Philos Trans R Soc Lond B Biol Sci.* 2002;357(1423):887-893. doi:10.1098/rstb.2002.1077.
 35. Ivancic T, Jamnik P, Stopar D. Cold shock CspA and CspB protein production during periodic temperature cycling in *Escherichia coli*. *BMC Res Notes.* 2013;6(1):248. doi:10.1186/1756-0500-6-248.
 36. Prasanna RR, Vijayalakshmi MA. Immobilized metal-ion affinity systems for recovery

- and structure–function studies of proteins at molecular, supramolecular, and cellular levels. *Pure Appl Chem*. 2010;82(1):39-55. doi:10.1351/PAC-CON-09-01-18.
37. Sommer B, Friehs K, Flaschel E, Reck M, Stahl F, Scheper T. Extracellular production and affinity purification of recombinant proteins with *Escherichia coli* using the versatility of the maltose binding protein. *J Biotechnol*. 2009;140(3-4):194-202. doi:10.1016/j.jbiotec.2009.01.010.
 38. Srivastava P, Deb JK. Gene expression systems in corynebacteria. *Protein Expr Purif*. 2005;40(2):221-229. doi:10.1016/j.pep.2004.06.017.
 39. Liu X, Yang Y, Zhang W, et al. Expression of recombinant protein using *Corynebacterium Glutamicum*: progress, challenges and applications. *Crit Rev Biotechnol*. 2015;8551(April 2016):1-13. doi:10.3109/07388551.2015.1004519.
 40. Liebl W, Sinskey AJ, Schleifer KH. Expression, secretion, and processing of staphylococcal nuclease by *Corynebacterium glutamicum*. *J Bacteriol*. 1992;174(6):1854-1861. <http://www.ncbi.nlm.nih.gov/pubmed/1548234>. Accessed June 1, 2017.
 41. Freudl R. Beyond amino acids: Use of the *Corynebacterium glutamicum* cell factory for the secretion of heterologous proteins. *J Biotechnol*. 2017. doi:10.1016/j.jbiotec.2017.02.023.
 42. Li M, Wang J, Geng Y, et al. A strategy of gene overexpression based on tandem repetitive promoters in *Escherichia coli*. *Microb Cell Fact*. 2012;11(1):19. doi:10.1186/1475-2859-11-19.
 43. Zwick F, Lale R, Valla S, et al. Strong stimulation of recombinant protein production in *Escherichia coli* by combining stimulatory control elements in an expression cassette. *Microb Cell Factories* 2012 111. 2012;25(2):402-408. doi:10.1006/meth.2001.1262.
 44. Jeong J, Cho N, Jung D, Bang D. Genome-scale genetic engineering in *Escherichia coli*. *Biotechnol Adv*. 2013;31(6):804-810. doi:10.1016/j.biotechadv.2013.04.003.
 45. Waegeman H, Soetaert W. Increasing recombinant protein production in *Escherichia coli* through metabolic and genetic engineering. *J Ind Microbiol Biotechnol*. 2011;38(12):1891-1910. doi:10.1007/s10295-011-1034-4.
 46. Mahalik S, Sharma AK, Mukherjee KJ. Genome engineering for improved recombinant protein expression in *Escherichia coli*. *Microb Cell Fact*. 2014;13(1):177. doi:10.1186/s12934-014-0177-1.

47. Li Y, Lin Z, Huang C, et al. Metabolic engineering of *Escherichia coli* using CRISPR–Cas9 mediated genome editing. *Metab Eng.* 2015;31:13-21. doi:10.1016/j.ymben.2015.06.006.
48. Guha TK, Wai A, Hausner G. Programmable Genome Editing Tools and their Regulation for Efficient Genome Engineering. *Comput Struct Biotechnol J.* 2017;15:146-160. doi:10.1016/j.csbj.2016.12.006.
49. Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol.* 2013;31(3):233-239. doi:10.1038/nbt.2508.
50. Peters JM, Silvis MR, Zhao D, Hawkins JS, Gross CA, Qi LS. Bacterial CRISPR: accomplishments and prospects. *Curr Opin Microbiol.* 2015;27:121-126. doi:10.1016/j.mib.2015.08.007.
51. Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J Bacteriol.* 1987;169(12):5429-5433. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=213968&tool=pmcentrez&rendertype=abstract>. Accessed July 6, 2015.
52. Luo ML, Leenay RT, Beisel CL. Current and future prospects for CRISPR-based tools in bacteria. *Biotechnol Bioeng.* 2016;113(5):930-943. doi:10.1002/bit.25851.
53. Jiang Y, Chen B, Duan C, Sun B, Yang J, Yang S. Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl Environ Microbiol.* 2015;81(7):2506-2514. doi:10.1128/AEM.04023-14.
54. Do BH, Ryu H-B, Hoang P, Koo B-K, Choe H. Soluble prokaryotic overexpression and purification of bioactive human granulocyte colony-stimulating factor by maltose binding protein and protein disulfide isomerase. *PLoS One.* 2014;9(3):e89906. doi:10.1371/journal.pone.0089906.
55. Xue Y, Davis A V, Balakrishnan G, et al. Cu(I) recognition via cation- π and methionine interactions in CusF. *Nat Chem Biol.* 2008;4(2):107-109. doi:10.1038/nchembio.2007.57.
56. Franke S, Grass G, Rensing C, Nies DH. Molecular Analysis of the Copper-Transporting Efflux System CusCFBA of *Escherichia coli*. *J Bacteriol.* 2003;185(13):3804-3812. doi:10.1128/JB.185.13.3804-3812.2003.
57. Choi JW, Yim SS, Kim MJ, Jeong KJ. Enhanced production of recombinant proteins with

- Corynebacterium glutamicum by deletion of insertion sequences (IS elements). *Microb Cell Fact.* 2015;14:207. doi:10.1186/s12934-015-0401-7.
58. Doench JG, Fusi N, Sullender M, et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol.* 2016;34(2):184-191. doi:10.1038/nbt.3437.
 59. Loftin IR, Franke S, Blackburn NJ, Mcevoy MM. Unusual Cu(I)/Ag(I) coordination of Escherichia coli CusF as revealed by atomic resolution crystallography and X-ray absorption spectroscopy. doi:10.1110/ps.073021307.
 60. Feilmeier BJ, Iseminger G, Schroeder D, Webber H, Phillips GJ. Green fluorescent protein functions as a reporter for protein localization in Escherichia coli. *J Bacteriol.* 2000;182(14):4068-4076.
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=94594&tool=pmcentrez&rendertype=abstract>. Accessed March 21, 2016.
 61. Vanz AL, Renard G, Palma MS, et al. Human granulocyte colony stimulating factor (hG-CSF): cloning, overexpression, purification and characterization. *Microb Cell Fact.* 2008;7:13. doi:10.1186/1475-2859-7-13.
 62. Bolanos-Garcia VM, Davies OR. Structural analysis and classification of native proteins from E. coli commonly co-purified by immobilised metal affinity chromatography. *Biochim Biophys Acta.* 2006;1760(9):1304-1313. doi:10.1016/j.bbagen.2006.03.027.
 63. Date M, Itaya H, Matsui H, Kikuchi Y. Secretion of human epidermal growth factor by Corynebacterium glutamicum. *Lett Appl Microbiol.* 2006;42(1):66-70. doi:10.1111/j.1472-765X.2005.01802.x.