



Universidad Nacional Autónoma de México

Facultad de Química

***“Producción de la proteína Tamapina
en el sistema de expresión pET40b”***

TESIS

***Para obtener el título de:
Químico Farmacéutico Biólogo***

Presenta:

Sergio Alland Colorado Cortes



Ciudad de México

2016



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Jurado Asignado

PRESIDENTE: M. en C. Rojo Callejas Francisco
VOCAL: Dra. Torres Márquez María Eugenia
SECRETARIO: Dra. Cano Sánchez Patricia
SUPLENTE 1: M. en C. Gama González Silvia Citlalli
SUPLENTE 2: Dr. del Rio Portilla José Federico

SITIO DONDE SE DESARROLLÓ EL TEMA:

Laboratorio 1 del departamento de Bioquímica y Laboratorio de Biología Molecular, Instituto de Química, UNAM

ASESOR DEL TEMA:

Dra. Patricia Cano Sánchez

SUPERVISOR TÉCNICO:

Dr. Federico del Río Portilla

SUSTENTANTE:

Sergio Alland Colorado Cortes

ÍNDICE GENERAL

ÍNDICE GENERAL.....	I
LISTA DE ABREVIATURAS	III
ÍNDICE DE FIGURAS	IV
1. INTRODUCCIÓN	1
2. HIPÓTESIS.....	2
3. OBJETIVO GENERAL.....	2
4. OBJETIVOS PARTICULARES	2
5. MARCO TEÓRICO.....	3
5.1 Toxinas de Alacrán.....	3
5.1.1 Tamapina	3
5.2.1 Aminoácidos	4
5.2.2 Proteínas	5
5.2.2.1 Estructura de proteínas.....	6
5.3 Síntesis de proteínas.....	7
5.3.1 DNA.....	7
5.3.2 Dogma Central de la Biología Molecular	8
5.3.2.1 Transcripción	8
5.3.2.2 Traducción	8
5.3.3 Tecnología del DNA recombinante	9
5.3.3.1 Plásmidos.....	9
5.3.3.1.1 Vector pET40b	10
5.3.3.2 Transformación	11
5.3.4 Expresión de Proteínas recombinantes.....	11
5.3.4.1 Regulación de la expresión	11
5.3.4.1.1 Producción de proteínas en periplasma	12
5.3.4.1.2 Sistema Dsb (Disulfide bond)	13
5.4 Purificación de proteínas.....	14
5.4.1 Cromatografía de afinidad metal quelato (CAMQ).....	15
5.4.2 Cromatografía de Líquidos de Alta Eficiencia (HPLC).....	16
5.5. Masas MALDI-TOF	17
5.6 Resonancia Magnética Nuclear (RMN)	17
5.6.1 Transiciones del espín nuclear	18
5.6.2 Desplazamiento Químico.....	19
5.7 Dicroísmo Circular.....	21

6. DISEÑO EXPERIMENTAL	23
6.1 Construcción del vector pET40b	23
6.2 Expresión en medio Luria Bertani (LB)	24
6.2.1 Inducción de la expresión de la proteína.....	24
6.2.2 Optimización de la producción de la proteína recombinante	25
6.2.2.1 Preparación de las muestras	26
6.3 Lisis selectiva de periplasma.....	26
6.4 Purificación de la proteína recombinante.....	27
6.4.1 Cromatografía de Afinidad Metal-Quelato.....	27
6.4.2 Corte con trombina	27
6.4.3 HPLC.....	28
6.5 Caracterización	28
6.5.1 Masas MALDI-TOF	28
6.5.2 Dicroísmo	29
6.5.3 RMN	29
7. RESULTADOS Y DISCUSIÓN	30
7.1 Construcción del vector.....	30
7.2 Optimización de la producción de la proteína recombinante	32
7.3 Producción de la proteína recombinante	35
7.4 Purificación por CAMQ.....	35
7.5 Corte con trombina.....	36
7.6 Purificación por HPLC	36
7.7Caracterización	37
7.7.1 Dicroísmo	37
7.7.2 Masas MALDI-TOF.....	39
7.7.3 RMN.....	40
8. CONCLUSIONES.....	42
9. BIBLIOGRAFÍA.....	43
Apéndice A. Abreviaturas de los 20 aminoácidos presentes en la naturaleza con su clasificación del grupo R.....	47
Apéndice C. Composición del medio LB.....	48
Apéndice D. Solución de lisis por sonicación 5X.	48
Apéndice E. Componentes de las soluciones para la lisis de periplasma.	48

LISTA DE ABREVIATURAS

ACN	Acetonitrilo
CAMQ	<i>Cromatografía de Afinidad Metal-Quelato</i>
DC	Dicroísmo Circular
DNA	Ácido Desoxirribonucleico
Dsb	<i>Disulfide bond</i>
HPLC	<i>Cromatografía de Líquidos de Alta Resolución</i>
IPTG	<i>Isopropil-β-D-1-tiogalactopiranosido</i>
KDa	<i>Kilo Dalton; 1DA = 1 g/mol</i>
MALDI-TOF	<i>Matrix-Assisted Laser Desorption Ionization-Time Of Flight</i>
PF	<i>Proteína de fusión</i>
RMN	<i>Resonancia Magnética Nuclear</i>
RNA_m	<i>Ácido Ribonucleico mensajero</i>
RNA_r	<i>Ácido Ribonucleico ribosomal</i>
RNA_t	<i>Ácido Ribonucleico de transferencia</i>
SDS PAGE	<i>Electroforesis en Gel de Poliacrilamida con DodecilSulfato de Sodio</i>
SK o K_{ca}	<i>Canales de potasio dependientes de calcio</i>
TFA	Ácido Trifluoroacético

ÍNDICE DE FIGURAS

Figura		Página
Figura 1.	<i>Secuencia de aminoácidos de la tamapina con las conectividades de puentes disulfuro.</i>	4
Figura 2.	<i>Estructura general de un aminoácido.</i>	5
Figura 3.	<i>Formación del enlace peptídico.</i>	6
Figura 4.	<i>Estructura de una proteína.</i>	7
Figura 5.	<i>Diseño del vector.</i>	10
Figura 6.	<i>Mecanismo de inducción de la PF.</i>	12
Figura 7.	<i>Sistema de plegamiento de proteínas en periplasma mediante las proteínas Dsb.</i>	14
Figura 8.	<i>Cromatografía de Afinidad Metal Quelato.</i>	15
Figura 9.	<i>Representación gráfica del proceso de separación de una mezcla a través de una columna.</i>	16
Figura 10.	<i>Efecto Zeeman de un núcleo.</i>	19
Figura 11.	<i>Efecto del Dicroísmo Circular.</i>	22
Figura 12.	<i>Representación esquemática de la construcción del vector pET40b con el inserto de la tamapina.</i>	24
Figura 13.	<i>Corte con la enzima BglI en el vector pET40-b con el inserto de la tamapina.</i>	30
Figura 14.	<i>Resultados de secuenciación.</i>	31
Figura 15.	<i>Secuenciación de la construcción del vector.</i>	31
Figura 16.	<i>Cinética de inducción.</i>	32
Figura 17.	<i>Fracción insoluble.</i>	33
Figura 18.	<i>Inducción a temperaturas variables y estudio de cepas.</i>	34
Figura 19.	<i>Experimento a diferentes concentraciones de IPTG.</i>	34
Figura 20.	<i>Cromatograma correspondiente a la purificación de la tamapina.</i>	37
Figura 21.	<i>Espectro de Dicroísmo Circular.</i>	38
Figura 22.	<i>Espectro de Masas MALDI-TOFF de la Tamapina.</i>	40
Figura 23.	<i>Espectro mononuclear de protón (¹H).</i>	41



1. INTRODUCCIÓN

El veneno de alacrán ha sido uno de los venenos más estudiados de los artrópodos. Está compuesto por diversas sustancias como hialuronidasa, fosfolipasas, proteasas, péptidos citolíticos, lípidos, carbohidratos entre otras sustancias desconocidas. La presencia de neurotoxinas en estos organismos puede afectar al hombre y a otros seres vivos, es por eso que es de vital importancia realizar investigación en este campo.¹

La tamapina es un péptido de 31 aminoácidos proveniente del veneno de alacrán, es conocida por su actividad bloqueadora y reguladora de canales iónicos. Pertenece a la subfamilia de toxinas α KTx5, su estructura terciaria contiene una α -hélice, dos hojas β y posee tres puentes disulfuro que le brindan estabilidad de tipo cisteínas α/β . La tamapina inhibe principalmente a canales KCa2 los cuales participan en la regulación de la excitabilidad sináptica. Dada a la distribución de estos canales en el sistema nervioso central y periférico ha sido propuesto como un novedoso blanco para tratar depresión, arritmias cardíacas, distrofia muscular y para el potenciamiento cognitivo.^{2,3}

Obtener de forma recombinante y bien plegada esta proteína nos permitirá comprobar y explorar la actividad de este péptido con posible potencial de aplicaciones biotecnológicas y de investigación.

En el presente trabajo se comparan dos sistemas de expresión bacteriano (rosseta+pET32a) con un nuevo vector de expresión (pET40b); esperando que la conectividad de los puentes disulfuro sea la nativa. En esta tesis se reportan los experimentos realizados para la expresión, purificación, caracterización y estudios preliminares de RMN.



2. HIPÓTESIS

La Tamapina se exporta a periplasma, donde la DsbC ejercerá su actividad de isomerasa de puentes disulfuro y habrá un aumento en el rendimiento de proteína con plegamiento nativo.

3. OBJETIVO GENERAL

Comparar el rendimiento y plegamiento de la proteína tamapina en dos sistemas de expresión bacterianos.

4. OBJETIVOS PARTICULARES

- Construir el vector pET40b con el inserto del gen de la tamapina.
- Determinar la cepa de expresión.
- Estandarizar el tiempo, temperatura y concentración de IPTG para realizar la inducción de la proteína recombinante.
- Determinar el rendimiento de GM-Tamapina.
- Realizar los experimentos de Dicroísmo Circular, Masas MALDI-TOF y RMN para caracterizar la proteína.



5. MARCO TEÓRICO

5.1 Toxinas de Alacrán

La palabra alacrán proviene del árabe y quiere decir escorpión. Pertenecen al grupo de los arácnidos y se consideran los animales terrestres más antiguos de todos por sus restos fósiles que corresponden al periodo silúrico. Actualmente, se les encuentra en todas las regiones del país, pero son más frecuentes en las zonas calientes y templadas no solo de México, sino de todo el mundo.⁴

La intoxicación por picadura de alacrán, es un problema de salud de notable importancia en México, que requiere de apoyo en materia de salud e investigación. La cifra de mortalidad anual y promedio en el país de 2006 a 2012 fue de 12.4 por 100,00 habitantes.⁵

Una toxina es una sustancia de origen biológico, natural, animal, vegetal o de microorganismos. Es una mezcla compleja de diferentes sustancias químicas que contienen docenas o incluso cientos de distintas proteínas que ejercen sus efectos en diferentes sistemas biológicos. Las toxinas de los alacranes actúan principalmente a nivel del sistema nervioso por lo que se le denominan neurotoxinas.^{3,6}

Por más de 50 años después de que se logró secuenciar la primera neurotóxina mediante purificación por cromatografías, se han utilizado como una herramienta para el estudio de la función de canales iónicos y es por eso que se ha investigado más a fondo los componentes particulares de los venenos de alacrán.⁷

5.1.1 Tamapina

La Tamapina es una toxina del veneno del alacrán rojo de la india (*Mesobuthus Tamulus*), aislada por Pedarzani.² Esta proteína se encuentra estabilizada por tres puentes disulfuro altamente conservados entre (C2-C5, C3-C6 y C1-C4)³. **(Fig.1)**



Fig. 1 Secuencia de aminoácidos de la tamapina con las conectividades de puentes disulfuro. En la imagen se puede observar la presencia de seis cisteínas que poseen el grupo tiol para la formación del enlace disulfuro. Adaptación de la referencia [2]

Esta Neurotóxina tiene la actividad de bloquear canales de potasio de baja conductancia dependientes de calcio (SK o K_{Ca}), principalmente el canal de tipo 2.1 (K_{Ca} 2.1).³

La Resonancia Magnética Nuclear (RMN) es la mejor técnica para el estudio estructural de péptidos que resultan complicado cristalizar y que regularmente contienen menos de 50 aminoácidos.³¹

Actualmente, se ha trabajado en la obtención de la tamapina por métodos recombinantes produciéndola como proteína de fusión con la tioredoxina y se ha elucidado la estructura terciaria de la misma mediante RMN⁸. Los rendimientos que se obtuvieron para la proteína de fusión en *E. coli* cepa rosseta gami fueron de 20 mg/L y 1.7 mg/L de GM-tamapina; la cual se obtiene después de cortar la proteína de fusión con trombina produciendo el péptido de interés con dos residuos adicionales en el amino terminal (Glicina y Metionina).⁸

5.2.1 Aminoácidos

Los aminoácidos son las unidades básicas de las proteínas, en la naturaleza existen 20 aminoácidos estándar y son α -aminoácidos. Los cuales tienen un grupo carboxilo y un grupo amino unido al mismo átomo de carbono (carbono α). Difieren unos de otros en sus cadenas laterales, o grupos R, que varían en estructura, tamaño y carga eléctrica (**Fig.2**).



A los aminoácidos se les han asignado abreviaturas de tres letras y símbolos de una sola letra (**Apéndice A**). El carbono α al poseer un carbón asimétrico o centro quiral genera dos tipos de isómeros los cuales son enantiómeros (imágenes especulares no superponibles). Siendo los L-aminoácidos los que se encuentran en las proteínas.⁹

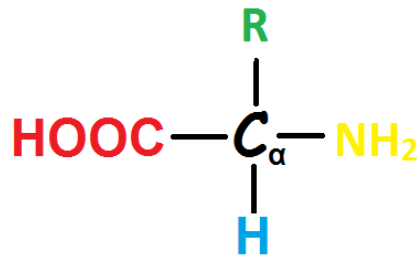


Fig.2 Estructura general de un aminoácido. Incluye a todos los α -aminoácidos con excepción de la prolina. En rojo se encuentra el grupo carboxilo, en amarillo el grupo amino y el grupo R es diferente para cada aminoácido.

5.2.2 Proteínas

Las proteínas son biomoléculas que varían en tamaño desde aquellas que contienen dos o tres aminoácidos (péptidos) hasta muy grandes que contienen cientos de ellos.

Para la formación de un enlace peptídico se requieren dos moléculas de aminoácidos que forman un enlace de tipo covalente a través de la deshidratación del grupo α carboxilo de un aminoácido y el grupo α amino de otro (enlace peptídico).⁹ **Fig.3**

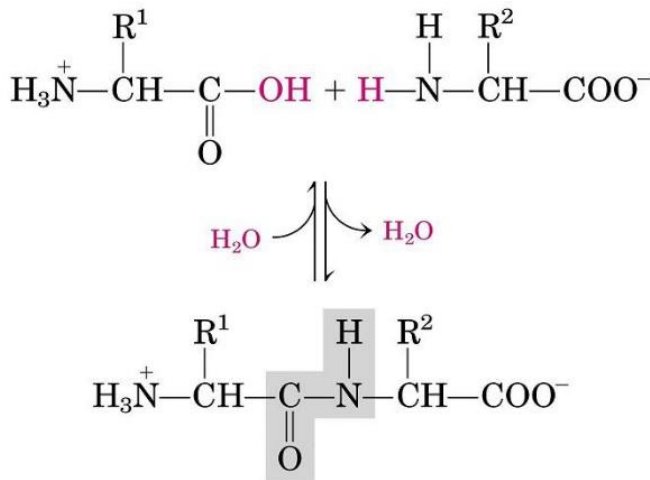


Fig. 3 Formación del enlace peptídico. Se observa la reacción reversible donde la pérdida de H_2O en la reacción forma el enlace peptídico (Sombreado). Recopilado de la referencia [9]

5.2.2.1 Estructura de proteínas

A finales de la década de 1930 Linus Paulin y Robert Corey iniciaron los estudios del enlace peptídico que sentaron las bases sobre la estructura proteica.⁹

Una vez que los aminoácidos se unen y forman una cadena lineal en un determinado orden constituyen la cadena polipeptídica conocida como estructura primaria. Toda proteína está definida por esta estructura; que es una única secuencia de aminoácidos. Se debe considerar que todos los niveles de organización (estructuras secundaria, terciaria y cuaternaria) subsisten en su estructura primaria.¹⁰

La estructura secundaria se forma cuando la cadena polipeptídica adquiere una conformación que origina un arreglo tridimensional.¹¹ Los principales tipos de estructuras son las *alfa* (α) Hélice y *Beta* (β) plegadas; ambas se encuentran estabilizadas por puentes de hidrógeno y a las partes de las proteínas que no tiene un patrón regular en su conformación se les llaman *giros* y *asas*.¹¹

En el siguiente nivel se encuentra la estructura terciaria que representa la cadena del polipéptido plegado y puede tener actividad biológica. La estabilización de esta estructura se logra mediante interacciones de los grupos R de los aminoácidos para evitar la repulsión de los mismos. Las principales interacciones que se



presentan son: enlaces disulfuro, puentes de hidrógeno, fuerzas de Van der Waals, interacciones hidrofóbicas y electroestáticas.¹⁰

Por último, se encuentra la estructura cuaternaria que se forma cuando dos o más polipéptidos se unen entre sí para formar una proteína; es decir que el arreglo espacial de los residuos de los aminoácidos está separado por su secuencia primaria. Esta estructura también posee una función biológica y se encuentra estabilizada por puentes de hidrógeno, interacciones hidrofóbicas, iónicas y/o enlaces disulfuro (S-S).¹⁰ **(Fig. 4)**

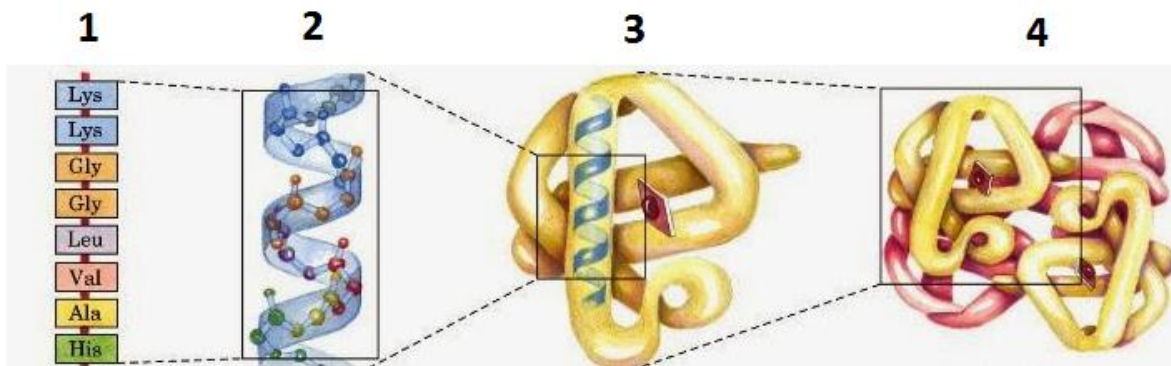


Fig. 4. Estructura de una proteína. La estructura primaria se muestra con su secuencia lineal de aminoácidos (1), la estructura secundaria se encuentra estabilizada por puentes de hidrógeno y forman α -hélice o β -plegadas (2), la estructura terciaria tiene plegamiento y puede tener actividad biológica (3) y la cuaternaria consta de más de una cadena polipeptídica. Recopilado de la referencia [9]

5.3 Síntesis de proteínas

5.3.1 DNA

Las moléculas de DNA (*Ácido Desoxirribonucleico*) contienen la información genética para la expresión de proteínas; las cuales realizan y regulan las actividades de las células para guiar su desarrollo y funcionamiento.¹² El DNA es una molécula de doble cadena unida por puentes de hidrógeno, cada cadena está conformada por nucleótidos que a su vez están compuestos por un azúcar



(Desoxirribosa), una base nitrogenada (Adenina, Timina, Guanina o Citocina).y un grupo fosfato, unidas por un enlace fosfodiéster.¹³

5.3.2 Dogma Central de la Biología Molecular

De acuerdo al dogma central de la biología molecular, la información genética fluye en los siguientes sentidos: (1) de DNA a DNA durante el proceso de transmisión de generación en generación y durante la replicación celular, y (2) de DNA a RNA y éste a la proteína durante la expresión fenotípica en un organismo. En el caso particular de algunos virus puede ocurrir el proceso de transmitir RNA a DNA o de RNA a RNA.¹²

5.3.2.1 Transcripción

Para que las células puedan codificar la información que contiene el DNA, primero se debe llevar a cabo la transcripción y consiste en la producción de RNA mensajero (RNA_m). El RNA_m es una copia complementaria de una molécula de DNA. Las principales diferencias entre estas moléculas son: (1) El azúcar que conforma el RNA es una Ribosa y en DNA es una desoxirribosa, (2) Es una molécula de cadena sencilla y (3) En el RNA la timina es remplazada por la base uracilo.¹⁴

Este proceso se lleva a cabo en tres etapas: La primera es la **iniciación** en la cual, la enzima RNA polimerasa se une a la región promotora de forma específica. La segunda etapa es la **elongación** que es la síntesis de la cadena de RNA mensajero. Y finalmente, se concluye el proceso con la **terminación** en la cual la cadena se libera de la enzima.¹⁵

5.3.2.2 Traducción

En todo este proceso la secuencia de un RNA_m se usa como un molde para unir aminoácidos a una cadena de polipéptido en un orden específico mediante codones (tres nucleótidos de RNA_m), los cuales de acuerdo al código genético codifican para un aminoácido (**Apéndice B**). Para que puedan ser leídos y



traducidos los codones, el RNA de transferencia (RNA_t) es el encargado de codificar y acarrear el aminoácido al péptido; Sin embargo, todo este proceso no se llevaría a cabo sin un último componente el cual es el RNA ribosomal (RNA_r); que está compuesto por una subunidad grande y pequeña, y tiene la función de asociar todo el complejo con los otros RNA's y otras proteínas como factores.¹⁶ De la misma forma, este proceso se divide en las etapas: **iniciación**, **elongación** y **terminación**.¹⁵

5.3.3 Tecnología del DNA recombinante

El DNA recombinante se construye mediante la incorporación de un fragmento de DNA foráneo en una molécula pequeña capaz de replicarse. Por ejemplo; un plásmido bacteriano. Las enzimas de restricción nos permiten cortar el DNA en sitios con secuencia específica, dependiendo el tipo de enzima, se producen dos fragmentos con extremos cohesivos; los cuales son apropiados para la inserción en un vector que haya sido cortado con la misma enzima.¹⁷

5.3.3.1 Plásmidos

Los plásmidos son moléculas circulares de DNA de doble cadena. Estos elementos genéticos son extracromosomales, es decir que son independientes al DNA genómico, se replican de manera independiente y codifican para una gran variedad de enzimas que le confieren a su hospedero resistencias a antibióticos y a metales pesados, degradan complejos orgánicos y/o producen toxinas. Dadas estas características se han utilizado como vectores de clonación de moléculas de DNA foráneas. Hoy en día, se dispone de una gran variedad de plásmidos sintéticos para la expresión de proteínas recombinantes cuyas modificaciones genéticas permiten controlar la producción de las mismas.¹⁸



5.3.3.1.1 Vector pET40b

El plásmido pET40b (Novagen) está diseñado para la expresión de la proteína de fusión con DsbC (**Fig.5**). Este vector contiene el sitio específico para transcribir la región codificante a través de la enzima T7 RNA polimerasa que es selectiva al promotor del plásmido. Y por último, permite que la bacteria adquiera la resistencia a kanamicina como marcador molecular para seleccionar a aquellas células competentes que hayan adquirido el material genético que le fue insertado mediante transformación.²¹

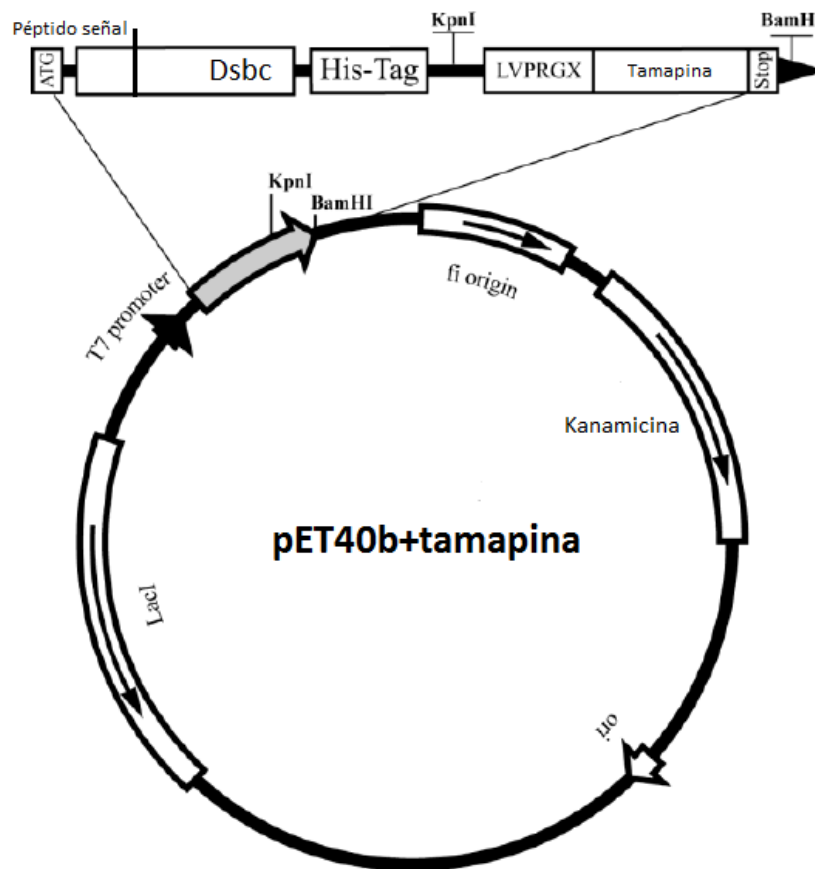


Fig. 5 Diseño del vector. La proteína de fusión para el vector pET40b; consta de la isomerasa de puentes disulfuro (DsbC), un péptido señal que permite translocar la proteína a través de la membrana interna a periplasma, una etiqueta de histidinas, la secuencia de aminoácidos para realizar el corte con trombina y por último, la secuencia del péptido de interés (tamapina). Adaptado de la referencia [8]



5.3.3.2 Transformación

La primera construcción y transformación de plásmidos fue lograda por Kohnert, permitiendo la expresión de proteínas recombinantes en *E. coli*.¹⁹ Al proceso mediante el cual se introduce genes exógenos a una bacteria se le conoce como transformación y puede realizarse de forma natural o artificial.²⁰ Para poder efectuar una transformación se requieren de células competentes (células capaces de aceptar material genético exógeno). Es importante que el vector de clonación o plásmido contenga un marcador que permita seleccionar a las bacterias transformadas.²⁰

5.3.4 Expresión de Proteínas recombinantes

Una proteína recombinante es aquella que se genera a partir de la tecnología del DNA recombinante.²² A pesar de los avances que se han tenido en la expresión de proteínas en sistemas de expresión Eucariontes (*S. cerevisiae*, animales transgénicos, líneas celulares, entre otros). *E. coli* sigue siendo el principal hospedero que se utiliza en investigación y en la industria biotecnológica ya que su producción es rápida, barata y fácil de obtener en comparación de otros organismos.²³

5.3.4.1 Regulación de la expresión

Para lograr la expresión selectiva de la proteína recombinante se utilizan cepas modificadas genéticamente, que utilizan el operón de la lactosa como elemento inducible para expresar la proteína T7 RNA polimerasa cuya secuencia proviene del bacteriófago λ DE3. De esta forma se logra la transcripción específica de la proteína de interés localizada en el vector dado que la transcripción del gen foráneo está dada por el promotor específico para la T7 RNA polimerasa y esta última como se mencionó antes se regula bajo el operador lac. Debido a esto se utiliza IPTG (isopropil- β -D-1-tiogalactopiranosido) como inductor; el cual es un análogo no hidrolizable de la galactosa.²¹

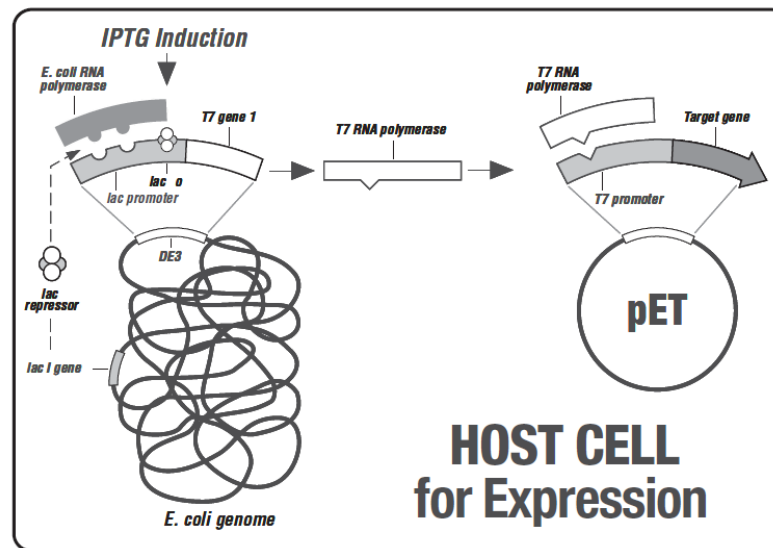


Fig. 6 Mecanismo de inducción de la PF. El vector posee un sitio de reconocimiento específico para transcribir la proteína de fusión a través de la T7 RNA polimerasa; que a su vez se encuentra regulada en el DNA genómico de la célula huésped por el operón de la lactosa. Recopilado de la referencia [21]

5.3.4.1.1 Producción de proteínas en periplasma

Los enlaces disulfuro se forman por la oxidación entre dos grupos tioles (SH), generando un enlace covalente entre dos cisteínas. Por lo tanto, para realizar la producción de proteínas que poseen puentes disulfuro se requiere un ambiente o agente oxidante. Sin embargo, dentro de una célula el ambiente es reductor. Y frecuentemente, si se producen proteínas de origen eucarionte en bacterias se enfrentan muchas problemáticas, como las modificaciones postraduccionales. Por ejemplo; los puentes disulfuro.²⁴

Para producir proteínas con puentes disulfuro en *E. coli* se puede diseñar la estrategia de translocar la PF a periplasma, es decir, dirigir el péptido al compartimiento entre las dos membranas de la bacteria. Este método nos permitiría catalizar la formación de los enlaces disulfuro; esto es porque en periplasma posee condiciones oxidantes en comparación al citoplasma.

Para efectuar este proceso se requiere que la proteína sea translocada con ayuda de un péptido señal a través de la membrana interna.²⁴ En el sistema pET la



proteína de fusión posee la secuencia de 20 aminoácidos que sirve como péptido señal, la principal proteína chaperona que lleva a cabo el proceso de translocación es la SecB. Esta debe reconocer a la proteína blanco en el citoplasma y una vez realizado este proceso se elimina el péptido señal.²³

5.3.4.1.2 Sistema Dsb (**Disulfide bond**)

En eucariontes la proteínas responsable de catalizar la formación y la isomerización de los puentes disulfuro es la **Proteína Isomerasa de Puentes Disulfuro (PDI)**.²³ Sin embargo, cuando se desea sintetizar en una célula procarionte una proteína recombinante que posee grupos tioles oxidados se pueden formar múltiples combinaciones de puentes disulfuro por lo que el rendimiento de la producción proteica disminuye debido al mal plegamiento y/o formación de agregados o cuerpos de inclusión.²³

En el periplasma de las bacterias se encuentra una serie de proteínas llamadas Dsb (**Disulfide bond**) que catalizan la formación y correcta conectividad de los enlaces disulfuro. Dentro de este sistema existen principalmente dos tipos de proteínas: La DsbA la cual posee la actividad catalítica de oxidar de forma inespecífica a los grupos tioles de las cisteínas para formar el correspondiente enlace covalente y la DsbC la cual tiene actividad de isomerasa de enlaces disulfuro permitiendo que la proteína adquiera su plegamiento con la conectividad nativa (**Fig. 7**).²³

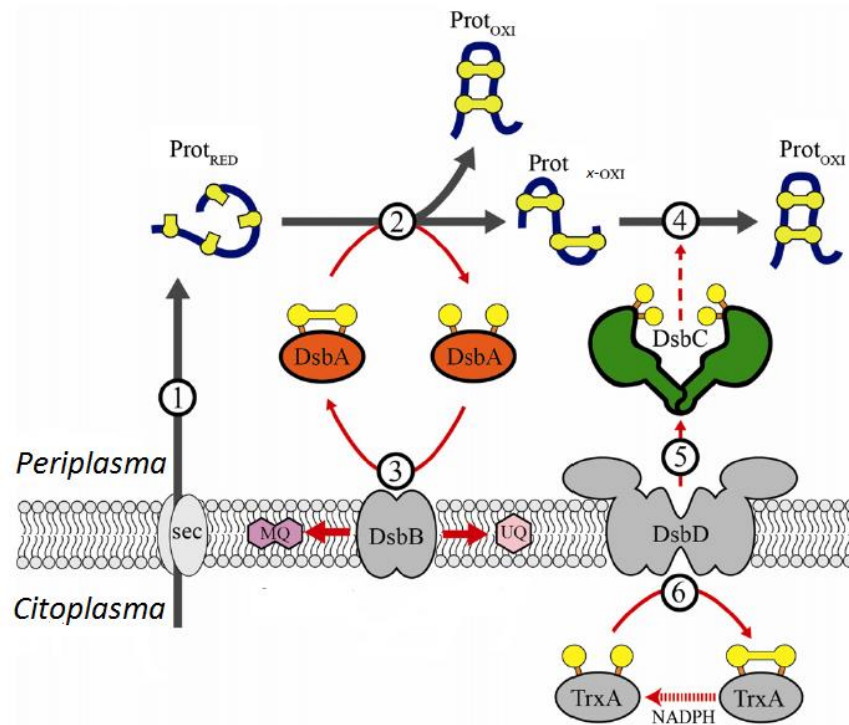


Fig. 7. Sistema de plegamiento de proteínas en periplasma mediante las proteínas Dsb. En la imagen, se observa lo que se espera que ocurra con la PF que es sintetizada en citoplasma; primero la proteína debe de ser translocada a periplasma mediante el péptido señal a través del sistema Sec (1), en seguida la proteína se oxida mediante los sitios activos de la DsbA (2), DsbB sigue proporcionando el poder oxidante a DsbA (3), si la proteína ha adquirido un plegamiento no nativo la DsbC se encarga de modificar la conectividad de puentes disulfuro (4) y quien mantiene a la DsbC en estado reducido es la proteína transmembranal DsbD (5) que a su vez en el citoplasma la tioreduxina junto con el NADPH mantienen el estado reducido de la misma (6). Recopilado de la referencia [23]

5.4 Purificación de proteínas

El objetivo de la purificación después de realizar la lisis celular es obtener una cantidad de proteína recombinante sin alguna otra proteína inherente a la lisis, el método de purificación más eficiente son las cromatografías como HPLC por sus siglas en inglés *High Performance Liquid Chromatography*. Sin embargo, en la mayoría de los pasos se utilizan pretratamientos para realizar estas técnicas. Por otra parte, es importante tomar en consideración que para obtener el mayor rendimiento de la proteína, la purificación debe contener el menor número posible de procesos.²⁵



5.4.1 Cromatografía de afinidad metal quelato (CAMQ)

Esta cromatografía consta en empacar o inmovilizar Ni^{2+} en la columna de sefrosa para que el ion pueda quelar ligandos, y como la proteína de fusión está diseñada con una secuencia consecutiva de 6 histidinas (HisTag o etiqueta de histidinas) se forma un enlace de coordinación con Ni^{2+} , confiriendo la afinidad necesaria para que se retenga en la columna y el resto de las proteínas resultantes de la lisis no se unan o lo hagan débilmente.

Para obtener la fracción de la proteína de fusión se utiliza Imidazol para despegar a la proteína de la columna debido a que este compuesto tiene una mayor afinidad por el Ni^{2+} y por lo tanto, al pasar un gradiente de concentración el imidazol desplaza a la proteína de fusión que se encuentra retenida en la columna.²⁶

(Fig. 8)

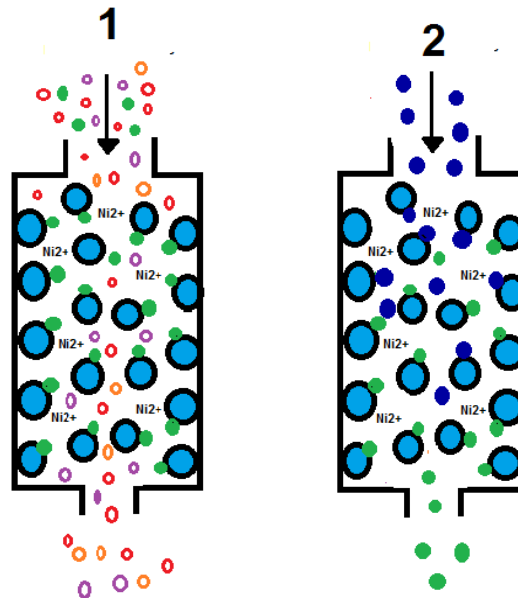


Fig. 8 Cromatografía de Afinidad Metal Quelato. Se observa que al pasar la proteína de fusión (verde) queda adherida a la columna debido a la formación de un enlace coordinado entre el Ni^{2+} y con las histidinas que posee la proteína de fusión. De esta forma los demás componentes de la mezcla son eludidos (1). Para recuperar la PF se realizan lavados con Imidazol (Azul rey) (2). Por tanto, hace un método eficiente y barato para la purificación de proteínas.



5.4.2 Cromatografía de líquidos de Alta Eficiencia (HPLC)

Debido a que esta cromatografía puede separar los compuestos de acuerdo a sus propiedades físicas con un alto poder de resolución, también se puede utilizar como método de purificación, caracterización e identificación para péptidos separando a la proteína de interés de otras impurezas provenientes de la lisis celular.²⁷ Esta técnica nos ofrece una purificación de moléculas que son solubles y que posean un peso molecular entre 1 a 1000 KDa.

Hoy en día, la industria biotecnológica y farmacéutica utiliza HPLC para investigación, desarrollo, control de calidad y manufactura de biomacromoléculas.

En esta cromatografía existen dos componentes; el primero es la **Fase Móvil** la cual es un líquido donde se encuentra la mezcla de compuestos y la segunda es la **Fase Estacionaria** que es una columna la cual esta empacada de un material de soporte. Para lograr separar algún compuesto de una mezcla se hace pasar la fase móvil a la columna donde existe una interacción de acuerdo al coeficiente de reparto de cada compuesto, de esta forma se logran separar los componentes en función a la afinidad que tengan por la Fase Estacionaria, y finalmente son analizados de forma individual por un detector, el más frecuente es el de UV-Visible.²⁸ **Fig.9**

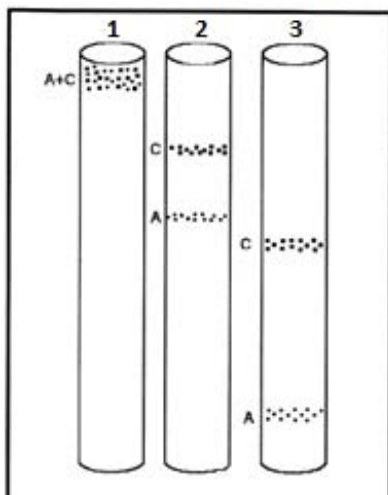


Fig.9. Representación gráfica del proceso de separación de una mezcla a través de una columna. En la imagen (1) se observa una mezcla de los componentes A y B que son pasados a través de la columna, al pasar más fase móvil por la fase estacionaria los compuestos se retienen dependiendo de sus coeficientes de reparto (2). De esta forma con esta técnica se logran separar mezclas complejas (3). Recopilado de la referencia [28]



5.5. Masas MALDI-TOF

La espectroscopia de masas es una técnica para analizar formas ionizadas de moléculas en fase gaseosa. Dado caso que los péptidos y proteínas nos son volátiles, se ha desarrollado la espectrometría de Desorción-ionización en matriz inducida por láser (MALDI, *Matrix-Assisted Laser Desorption-Ionization*) de esta forma un péptido o alguna macromolécula se coprecipita con un compuesto orgánico que absorbe luz láser de una longitud de onda apropiada. El destello de un láser sobre la preparación expulsa moléculas de la superficie que capturan electrones cuando salen de la matriz y por tanto, la dejan como iones cargados negativamente. En ESI (*ionización por electrospray*), las disoluciones que contienen la proteína fluyen a través de una punta fina metálica mantenida por un potencial eléctrico distinto de cero liberándose pequeñas partículas cargadas. Posteriormente, se realiza un análisis de tiempo de vuelo (TOF, *Time Of Flight*), los iones se aceleran en un campo eléctrico hacia un detector de tal forma que los iones más pequeños experimentan una mayor aceleración y llegan primero al detector. De esta forma la espectrometría de masas MALDI-TOF proporciona una herramienta poderosa para la caracterización e identificación de proteínas, péptidos, polímeros, fragmentos de DNA, y todos aquellos compuestos de alto peso molecular.^{29, 30}

5.6 Resonancia Magnética Nuclear (RMN)

La RMN es una poderosa técnica espectroscópica que se usada en investigación para la determinación de estructuras, cinéticas y dinámicas químicas de un amplio intervalo de sistemas químicos y bioquímicos.

Se puede determinar la estructura de proteínas a través de la RMN en estado líquido y sólido, siendo un método alterno para determinar la estructura de aquellas proteínas que no se pueden cristalizar. Sin embargo, esta técnica tiene la limitación de no poder procesar moléculas con peso molecular mayor de 50-60 KDa.^{31, 32}



5.6.1 Transiciones del espín nuclear

Para cualquier espectroscopía es necesario tener un sistema que posea dos o más niveles de energía. En el caso de la RMN estas interacciones de energía son debidas al *momento angular de espín* del núcleo. Cuando un núcleo con espín diferente de cero se coloca en un campo magnético, se generan dos niveles de energía. La transición entre estos niveles debido a la absorción o emisión de esta energía que se ubica en la región de las radiofrecuencias es lo que llamamos Resonancia Magnética Nuclear.

La energía de estos niveles es función de la componente en Z del momento angular de espín, de la intensidad de campo magnético aplicado y de la constante giromagnética, dada por la **Ecuación 1**.

$$E = -\gamma\hbar B m_z \quad (1)$$

Donde γ es la constante giromagnética, \hbar es la constante de Planck y m_z es el componente en z del momento angular de espín.

Esta ecuación hace referencia al efecto de Zeeman dada por el Hamiltoniano. Lo cual hace posible que se puedan representar los diagramas de energía para una partícula con espín de $\frac{1}{2}$ en función de la intensidad del campo magnético. **Fig. 10**

El estado basal o de menor energía se representa como α y corresponde a $m_z=+1/2$ y al estado excitado o de alta energía se representa con β y corresponde a $m_z=-1/2$. La energía de estos dos estados se calculada mediante la **Ecuación 2**:

$$E_\alpha = -\frac{\gamma\hbar B}{2} \quad E_\beta = +\frac{\gamma\hbar B}{2} \quad (2)$$

Por lo tanto, la diferencia de energía entre estos dos estados es:

$$\Delta E = E_\beta - E_\alpha = \gamma\hbar B \quad (3)$$

Considerando que la diferencia de población entre los dos estados de energía es muy pequeña; la energía de transición en RMN también lo es, por lo que requiere ondas de radio para excitar los espines del núcleo.

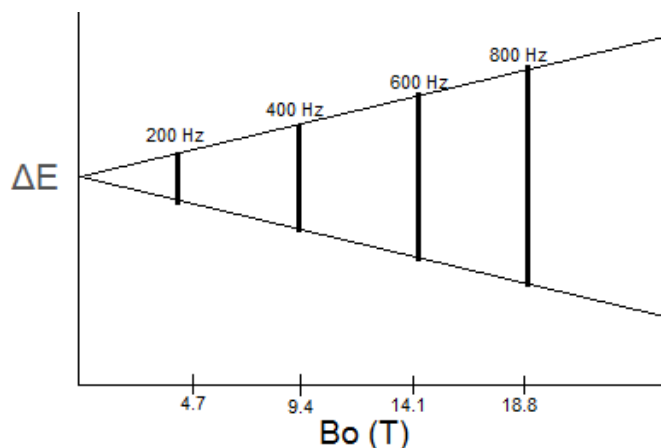


Fig.10. Efecto Zeeman de un núcleo. Cuando se alinea un núcleo en un campo magnético (B_0), los niveles energéticos de los posibles estados de espín nuclear se desdoblan proporcionalmente al campo aplicado; en consecuencia, las transiciones entre los niveles energéticos del espín nuclear no serán de mayor energía en comparación con el campo magnético aplicado.³³ Adaptado de la referencia [33]

La diferencia de población puede ser calculada mediante la distribución de Boltzman:

$$\frac{N_\beta}{N_\alpha} = e^{\frac{-\gamma h B}{kT}} \quad (4)$$

Tener una pequeña diferencia de población según la distribución de Boltzman, la espectroscopia de RMN es relativamente poco sensible. Para los experimentos de RMN con proteínas se requieren concentraciones de al menos 1 mM.³¹

El protón ^1H es el más sensible de todos los núcleos que pueden ser determinados mediante RMN. La sensibilidad del núcleo está dada por la constante giromagnética y en protón es la más grande de todas. Por otra parte el porcentaje de abundancia en la naturaleza del ^1H es del 99.98 % y, por tanto, existe una mayor diferencia poblacional.

5.6.2 Desplazamiento Químico

El entorno electrónico y las interacciones dinámicas que posee una estructura molecular afecta al núcleo. El ambiente magnético local está influenciado por la



capa de valencia, hibridación de orbitales y el tipo de enlace, carga, polaridad del enlace, y la electronegatividad de los átomos vecinos, así como el disolvente afectan la energía de transición de los núcleos. Todas estas pequeñas perturbaciones al campo magnético local producen pequeños cambios donde la señal de un núcleo en específico aparece en un espectro; a lo cual se le conoce como desplazamiento químico.³²

El campo magnético aplicado B_0 usualmente difiere significativamente cuando el núcleo está siendo *apantallado* por la densidad electrónica que rodea al núcleo. Este apantallamiento es debido a los electrones. La aplicación de la fuerza del campo magnético local en el núcleo está dada por la **Ecuación 5:**³⁴

$$B = B_0 - B' = B_0(1 - \sigma) \quad (5)$$

Donde σ representa el apantallamiento del espín nuclear. Como resultado de estas condiciones tenemos que para un núcleo de una molécula que se encuentre en resonancia:

$$E = \frac{\gamma B_0(1-\sigma)}{2\pi} \quad (6)$$

En RMN de transformada de Fourier medimos la fuerza electromotriz en función del tiempo, a lo cual le llamamos dominio del tiempo, conocido como el FID (Free Induce Decay). En contraste, en la mayoría de las otras espectroscopias se mide la absorción o emisión de energía en función de la frecuencia para generar un espectro. Es por eso que el FID que se obtiene después de aplicar un pulso en RMN se le aplica la transformada de Fourier (TF); que es una relación matemática que genera el espectro en función de la frecuencia e intensidad.

Para eliminar la dependencia de las frecuencias que son debidas al campo magnético aplicado (B_0), se construye una escala adimensional con unidades de ppm (parte por millón). **Ecuación 7**

$$\delta = \frac{\nu - \nu_0}{\nu_0} \times 10^6 \quad (7)$$



Donde V_0 es la frecuencia de referencia (tetrametilsilano para H1, C13 y Si19), y V es la frecuencia de la muestra. Como la diferencia de frecuencias es muy pequeña se compensa multiplicando por un millón para obtener la escala en número más grandes (ppm).³⁴

5.7 Dicroísmo Circular

En el dicroísmo circular (DC) se hace pasar un plano de luz polarizada de dos tipos; la primera rota en dirección contraria a las manecillas del reloj y se le llama L y la que rota en sentido de las manecillas del reloj es conocida como R. El DC se fundamenta en la absorción de estos dos tipos de luz, si después de pasar a través de una muestra que es analizada la luz L y R, no son absorbidas o son absorbidas en mismas cantidades la luz polarizada se dirige al plano original. Sin embargo, cuando la luz R y L son absorbidas en diferentes grados se genera una polarización elíptica.³⁹

En esta técnica se mide la diferencia de absorción entre la luz polarizada L y R pero también puede ser reportada en términos de grados de elipticidad (Θ). Las señales de absorción que se obtienen nos permiten determinar ciertas características estructurales de una molécula. En el caso particular de proteínas nos permite tener información que puede ser obtenida de un espectro que se genera por la absorción en regiones menores de 240 nm principalmente por los enlaces peptídicos y entre 260-320 nm por los aminoácidos aromáticos, puentes disulfuro o cofactores. De esta manera se puede determinar diferentes tipos de estructuras secundarias con los espectros que se generan con el DC.³⁹

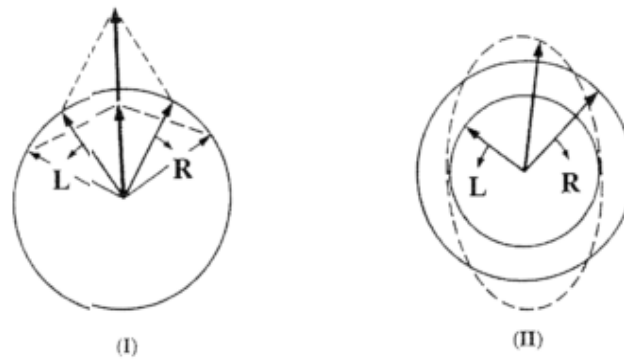


Fig.11 Efecto del Dicroísmo Circular. Se irradia la muestra con los componentes de la luz polarizada L y R. En (I) Los dos componentes tienen la misma amplitud y cuando se combinan generan un plano de radiación polarizada. En (II) Los componentes son de diferente magnitud y como resultado se genera la elipticidad. Recopilado de la referencia [39]



6. DISEÑO EXPERIMENTAL

El diseño experimental consiste en la construcción del plásmido pET40b con el inserto del gen de la tamapina para realizar la expresión en células de *E. coli*. Posteriormente, se establecerán las condiciones de tiempo de inducción, temperatura, cepa de expresión, concentración de IPTG y lisis selectiva de periplasma. Por último se purifica la proteína, se caracteriza y cuantifica.

6.1 Construcción del vector pET40b

Para realizar la subclonación del gen de la tamapina se utilizó el vector previamente construido en el laboratorio pET32a + tamapina (Novagen), ubicándose entre los sitios de restricción *KpnI* y *BamHI* por lo cual se realiza una reacción con ambas enzimas (New England Biolabs) y se purifica la banda del gen de acuerdo al protocolo gel Extraction Kit (QIAGEN). De igual forma se realiza el corte del vector pET40b (Novagen) con las mismas enzimas de restricción y finalmente, se realiza la reacción de ligación con la enzima T4 DNA ligase (New England Biolabs). **(Fig. 12)**

Una vez que se haya realizado la reacción de ligación, ésta se utiliza para transformar células DH5- α para poder seleccionar clonas positivas, se purifica el correspondiente DNA plasmídico y se determina la presencia del inserto del gen mediante un ensayo de restricción y posteriormente un análisis de secuenciación.

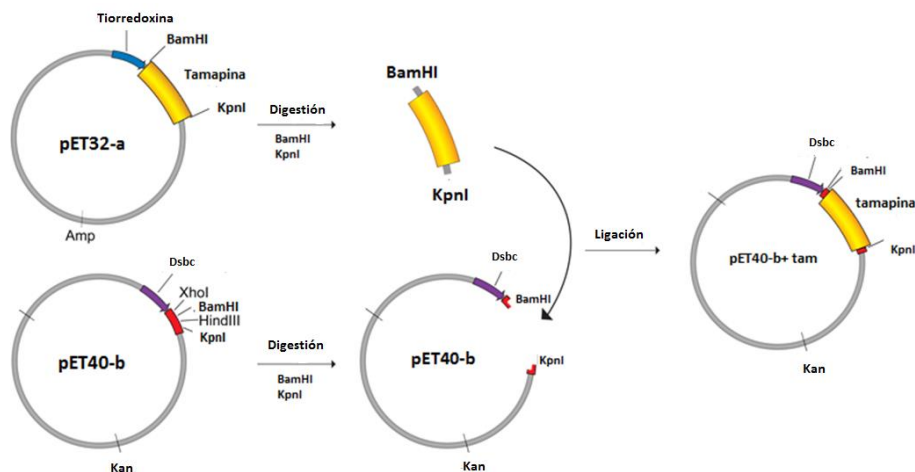


Fig.12 Representación esquemática de la construcción del vector pET40b con el inserto de la tamapina.

6.2 Expresión en medio Luria Bertani (LB)

Es importante para el sistema de expresión (Cepa + Vector) establecer las condiciones óptimas de inducción (tiempo, temperatura, concentración de IPTG) para tener buenos rendimientos, el medio de cultivo también es un factor esencial.

El medio LB (**Apéndice C**) está enriquecido, especialmente para *Escherichia coli* permite que la bacteria se encuentre con todas las condiciones nutrimentales necesarias como fuente de carbono y nitrógeno, aminoácidos, balance osmótico, etc. para que tenga los requerimientos energéticos y los elementos para realizar la síntesis de sus componentes celulares.³⁵

6.2.1 Inducción de la expresión de la proteína

Para realizar la inducción de la proteína se debe de picar una colonia transformada con el nuevo vector construido y colocar un preinóculo con medio LB y el antibiótico correspondiente. Se deja incubando con agitación constante a 37°C durante toda la noche (16 h). Posteriormente, se inocula un litro y medio de medio nuevo con 50 ml de preinóculo y se monitorea el crecimiento bacteriano a una



longitud de onda de 600 nm hasta llegar a una densidad óptica de 0.7 de absorbancia. Finalmente, se adiciona IPTG al 0.5 mM para inducir producción de la proteína.

6.2.2 Optimización de la producción de la proteína recombinante

La primera variable que se determina es el tiempo de inducción de la proteína, y para ello se realiza una cinética tomando los siguientes tiempos de muestreo: 0, 3, 6, 9, 12, 16, 20, 24 y 32 horas. Considerando las condiciones establecidas para el otro sistema de expresión (IPTG al 0.5mM y Temperatura a 30°C).

Una vez establecido el tiempo de inducción, se realiza un estudio de cepas de *E. coli* para realizar la expresión. Se compararon las siguientes cepas: Rosseta, Shuffle, C41 y BL21 para determinar qué sistema de expresión (Célula hospedera + vector) es el ideal.

BL21 (DE3) contiene el gen que codifica para la T7 RNA polimerasa proveniente del bacteriófago λ asociado a un promotor que es inducible en presencia de IPTG para realizar la expresión de cualquier plásmido que contenga el promotor específico T7.³⁶ Las otras cepas que se utilizaron son variantes de esta bacteria que fueron modificadas genéticamente.

La célula C41 posee una mutación que previene la muerte celular asociada a la expresión de proteínas recombinantes que le resulten tóxica a la bacteria. Por tanto, esta cepa es ideal para la expresión de toxinas provenientes de cualquier organismo como virus, plantas, levaduras o células animales.³⁶ La cepa Rosseta tiene la característica de contener un plásmido extra que produce más represor de la Lactosa, está diseñada para potenciar la expresión de proteínas de origen Eucarionte que contienen codones que raramente son usados de forma natural por la bacteria.³⁷ Y por último Shuffle es una variante modificada por ingeniería genética especialmente para favorecer la formación de puentes disulfuro en citoplasma.³⁸



Finalmente, para establecer todas las condiciones se determina temperatura óptima de inducción entre 20 ° C, 30 ° C o 37 ° C, y se determina la concentración de IPTG (0.25mM, 0.5mM, 0.75mM y 1mM).

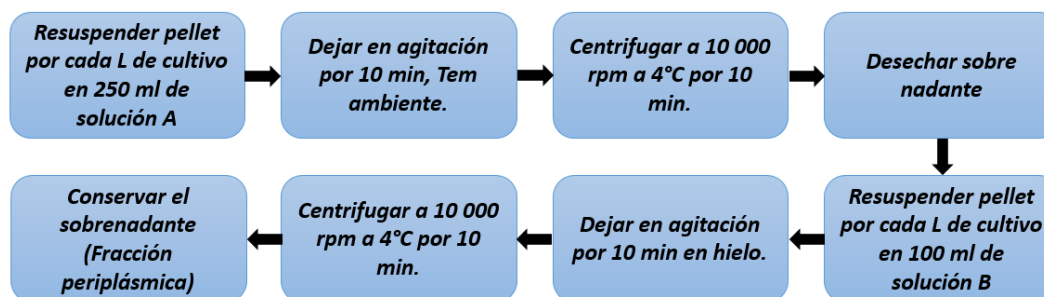
6.2.2.1 Preparación de las muestras

Las muestras se prepararon de la siguiente forma: Se lisaron todos los pellets con 100 µL de buffer de lisis (**Apéndice D**), se dieron choques sónicos durante 10 segundos hasta dejar de observar la muestra turbia y viscosa, se centrifugaron a 13 000 rpm durante 10 min y el sobrenadante fue colectado en tubos eppendorf nuevos. La observación de la proteína de fusión se realizó mediante electroforesis SDS-PAGE colocando la misma cantidad de proteína para cada muestra, lo cual se determinó por absorbancia a 280.

6.3 Lisis selectiva de periplasma

El método consiste en realizar un choque osmótico primero con una solución hipertónica (**Solución A**); seguida de una solución hipotónica (**Solución B**). En el **Apéndice E** se puede encontrar la composición de las soluciones.

Diagrama para realizar la lisis de periplasma.





6.4 Purificación de la proteína recombinante

6.4.1 Cromatografía de Afinidad Metal-Quelato

El sobrenadante se filtra con una membrana de tamaño de poro de 0.45 μm y se hace pasar por una columna HiTrap cargada con Ni^{2+} . Una vez que se hizo pasar el lisado se realiza una serie de lavados con cuatro volúmenes con agua destilada filtrada y con 10 volúmenes de columna con solución de Imidazol al 15 mM con el fin de eliminar sustancias inespecíficas que pudieron llegarse a retener en la columna. Finalmente, se eluye toda la PF con dos volúmenes de columna con la solución de Imidazol al 300 mM.

6.4.2 Corte con trombina

Antes de realizar la reacción de proteólisis de la proteína de fusión, las fracciones que fueron eluidas de la cromatografía de afinidad fueron dializadas con buffer de corte (Tres recambios con agitación después de 1, 3 y 6 horas) con una membrana de 8-10 KDa. De esta forma el Imidazol se elimina del medio y se puede proceder a la reacción de corte con trombina (Proteasa de corte específico) para obtener la GM-tamapina. Se coloca la fracción del paso anterior en un tubo falcón de 15 mL y se adiciona 60 μL de la suspensión de partículas con trombina (SIGMA KIT). La mezcla se coloca en agitación en un agitador de tipo carrusel el tiempo necesario para realizar el corte completo de la proteína de fusión (12 h). Finalmente, se recupera la trombina.

La fracción resultante (GM-Tamapina + proteína de fusión) es filtrada con una membrada de 0.45 μm y se hace pasar nuevamente por la columna HiTrap cargada con Ni^{2+} . En este proceso la fracción que nos interesa es la que no se adhiere a la columna puesto que la etiqueta de histidinas se queda en la proteína de fusión.



6.4.3 HPLC

Antes de iniciar con este proceso de purificación la muestra se hace pasar por una columna en fase reversa “Strata” (phenomenex, C18-T); este proceso es una extracción en fase sólido. Después realizar esta purificación se de liofiliza la muestra, se resuspende en ACN al 15% con TFA al 0.5% y se filtra a un tamaño de poro de 0.22 μm para ser inyectada al cromatógrafo.

Se utilizó un cromatógrafo de líquidos de alta eficiencia (HPLC) de Varian ProStar, con una columna de fase reversa C12 (phenomenex), con un gradiente creciente de hidrofobicidad del 15% al 37% de H_2O y ACN ambos con TFA al 0.5%v/v; con un flujo de 1 ml/min y fue analizado a una longitud de onda de 230 nm con detector UV-visible modelo PS325. Para obtener la proteína de interés se requiere de TFA y ACN para que el péptido presente mayores interacciones hidrofóbicas y pueda retenerse a la columna. Esta técnica es el último paso de la purificación. Sin embargo, debe de retirarse el ACN y TFA que se incorporaron en la fase móvil y lo cual se logra liofilizando la muestra.

6.5 Caracterización

6.5.1 Masas MALDI-TOF

Se determina el peso molecular mediante espectrometría de masas MALDI-TOF con el equipo de Bruker Daltonics. Con esta técnica nos podemos percatar si la proteína se encuentra en su estado oxidado, es decir, si tiene los puentes disulfuro mediante la perdida de los protones de los grupos tioles.

Se utilizó como matriz alfa-cianohidroxicinamico (CHCA) analizado en modo reflectron con una ventana de 700 a 5000 Da.



6.5.2 Dicroísmo

Mediante esta técnica se determina la estructura secundaria de nuestra proteína; la cual posee una alfa hélice y dos betas plegadas. Se realiza el experimento en el espectropolarímetro Jasco J-720; para realizar este experimento se requiere de una concentración de al menos 100 $\mu\text{g/ml}$ de proteína disuelta en agua milliQ. Se utilizó una celda para dicroísmo de 1 cm de largo; se inicia el experimento a una longitud de onda de 260 nm hasta 190 nm con una resolución de 1 nm, con una sensibilidad de 20 mdeg y con una velocidad de 20 nm/min.

6.5.3 RMN

La técnica de RMN es muy poco sensible, es por eso que al menos se requieren de concentraciones de 1 mM para determinación estructural. Se prepara la muestra disolviéndola en agua y agua deuterada (D_2O) al 5%. La muestra se coloca en un tubo para resonancia normal o “shigemi” para ser analizada en el equipo de RMN de 500 MHz Varian UnityPlus.



7. RESULTADOS Y DISCUSIÓN

7.1 Construcción del vector.

Después de transformar la reacción de ligación (pET40b con el inserto de la tamapina) en células DH5 alfa, se seleccionaron ocho colonias para realizar la extracción del plásmido y verificar que el vector poseía el gen de la tamapina. En el programa Serial Cloner se construyó el plásmido y se identificó que la enzima BglI tiene sitio de corte en el gen y en otra región del plásmido, por lo tanto, se utilizó esta enzima para corroborar las clonas positivas que deben de dar dos bandas con pesos aproximados de 3.5 y 2.8 kb. (**Fig.13**)

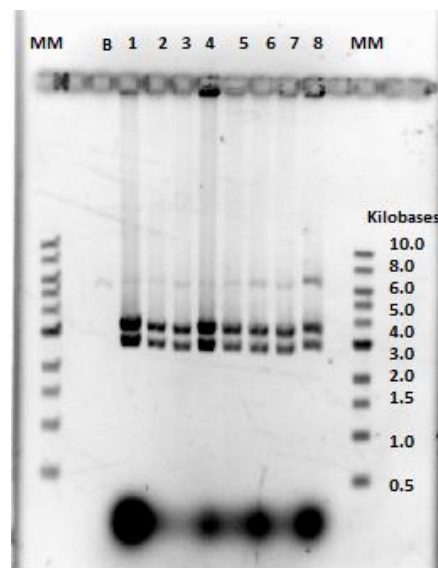


Fig.13 Corte con la enzima BglI en el vector pET40-b con el inserto de la tamapina. En la imagen se observan dos bandas que indican la presencia del gen de las muestras 1 a la 8, en el Blanco (B) se encuentra el vector sin inserto.

De las ocho muestras se seleccionaron cuatro (2022, 2023, 2026 y 2028), fueron secuenciadas por el método de Sanger (Laragen Inc.) con el primer T7 terminal (Dirección 3'-5') y se realizó un alineamiento en clustal donde se comprobó que las cuatro muestras poseen la secuencia del gen de la tamapina en el vector pET40-b. (**Fig.13**)

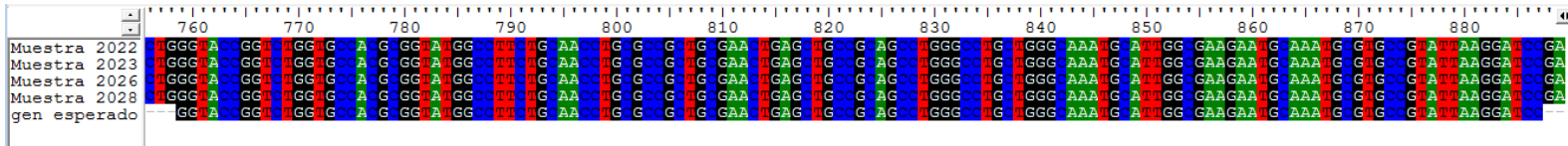


Fig. 14 Resultados de secuenciación. En la imagen se observa el alineamiento en clustal del primer T7 terminador donde todas las muestras poseen la secuencia deseada.

A continuación, se detalla la secuenciación de la muestra 2023 (Fig.15). Se verificó el marco de lectura que codifica para la proteína de fusión; la secuencia de nucleótidos que codifica para los 31 aminoácidos que conforman la tamapina (Cuadro rojo), el sitio de corte para la proteasa trombina (Cuadro verde), la etiqueta de histidinas (Cuadro Rojo con H rosas) y el codón de paro (Asterisco) sean las correctas.

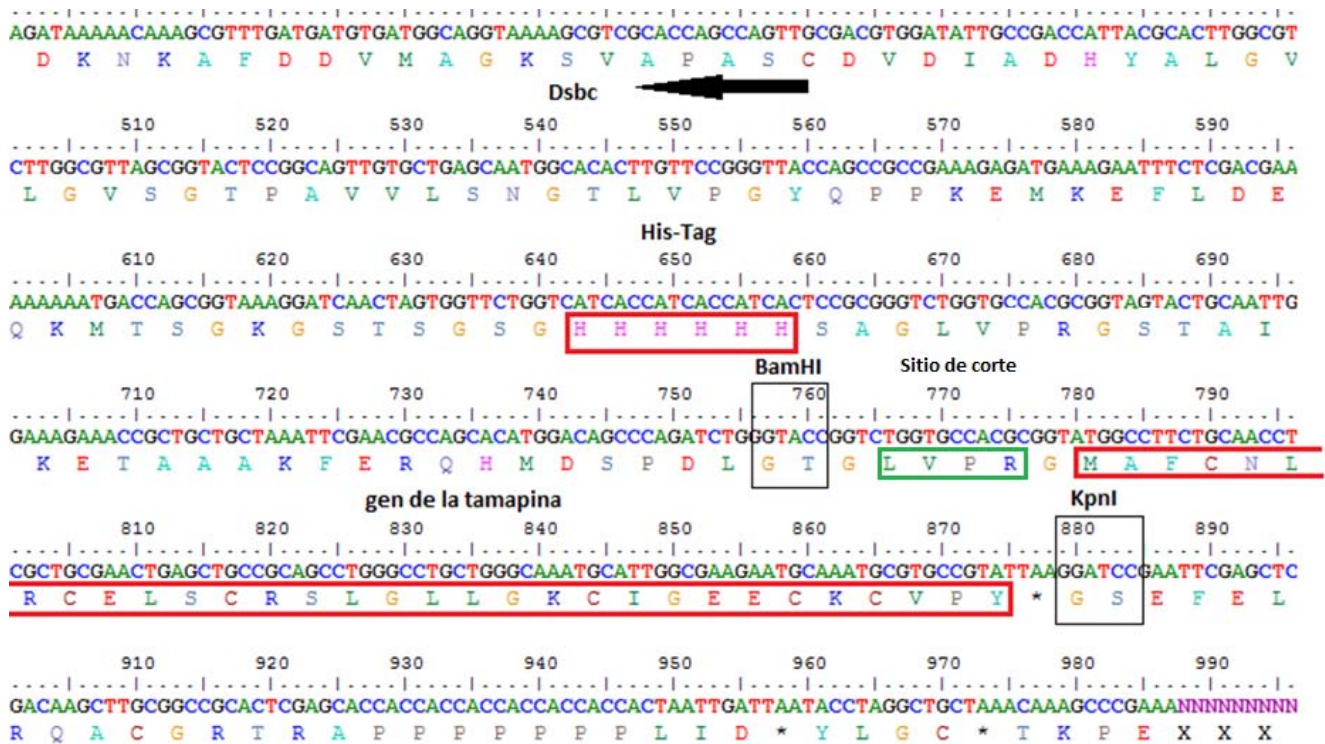


Fig.15 Secuenciación de la construcción del vector. Se observa que todos los componentes que codificaran para la proteína de fusión son correctos, así como el marco de lectura.



7.2 Optimización de la producción de la proteína recombinante

Se decidió iniciar los experimentos con la cepa C41 de *E. coli* debido a que posee una mutación fenotípica que le confiere tolerancia a las toxinas.³⁶ Se transformó el vector construido a la célula y se utilizó una electroforesis en gel de poliácridamida (SDS-PAGE) para monitorear la expresión, el peso de la proteína de fusión teórico es 32 kDa.

Se determinó que el tiempo óptimo de inducción esta entre las 12 y 16 horas (**Fig. 16**). Nos percatamos que la banda de la PF disminuye de intensidad después de las 24 horas de inducción; Esto puede deberse a que la bacteria comience a degradar la proteína o forme agregados. Cabe mencionar que en este gel se analiza la fracción soluble de proteínas tomando en cuenta el método de lisis (sonicación). Se puede observar que la perdida de la PF por la formación de cuerpos de inclusión es muy poca (**Fig.17**). Para realizar este experimento solo se resuspendieron los pellets del ensayo anterior adicionando 25 μ L de Buffer de carga, se disgregan por sonicación, se hierven y las muestras se cargan al gel.

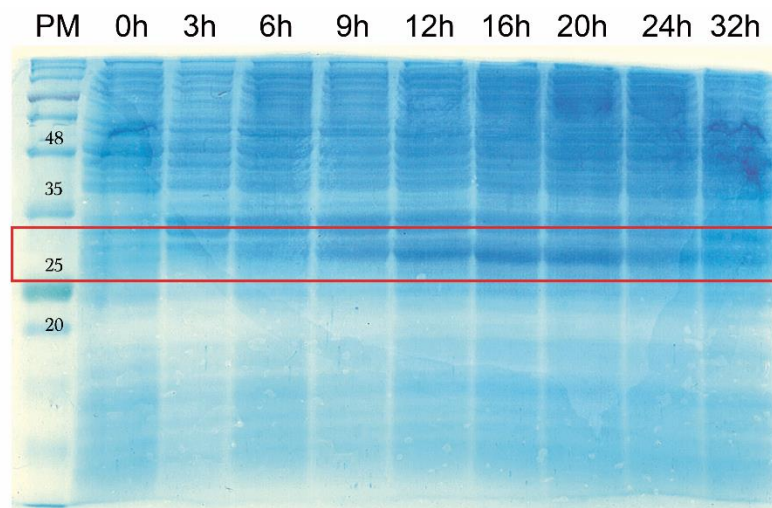


Fig. 16 Cinética de inducción. En el presente gel de poliácridamida al 12 % se analiza la fracción de proteína soluble, se observa que en el tiempo cero no se encuentra la PF porque aún no se inicia la inducción y por tanto, el tiempo óptimo de inducción se encuentra entre las 12 y 16 horas debido a la intensidad de la banda.

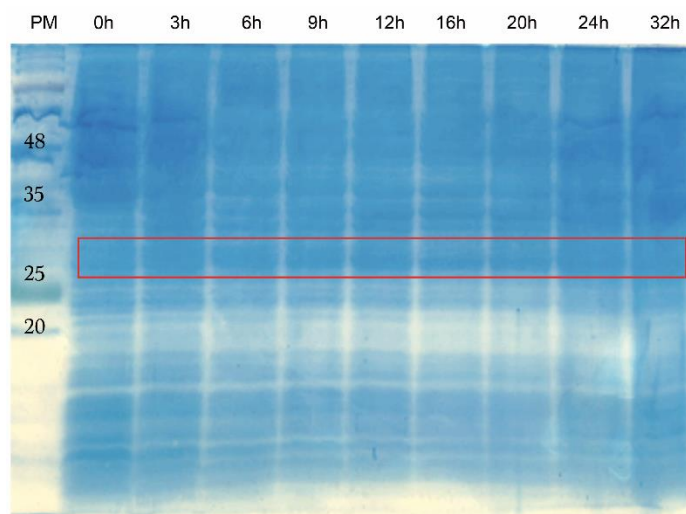


Fig. 17. Fracción insoluble. En este gel se analiza la fracción insoluble del experimento anterior, como se puede observar la PF forma una cantidad aceptable de cuerpos de inclusión o aglomerados.

Una vez establecido el tiempo de inducción se procede a determinar la temperatura óptima de inducción, para este experimento se utilizó la misma cepa del experimento anterior (C41) y se indujo la proteína a 20°C, 30°C y 37°C a una concentración de 0.5 mM de IPTG por 12 horas. (**Fig. 18**). Nos percatamos que la cantidad de proteína que se produce es mayor a 30 °C y 37°C. Sin embargo, se estableció la temperatura de inducción a 30 °C puesto que a mayor temperatura existe la posibilidad de generar agregados de proteína debido al incremento de las interacciones hidrofóbicas.³⁷

Por último, se realizó un estudio de cepas para determinar cual nos proporciona un mayor rendimiento. Se decidió iniciar la inducción con Rosseta ya que era la cepa que tenía mejores rendimientos respecto al vector pET32a + tamapina, Shuffle posee un ambiente oxidante ideal para que se oxiden las cisteínas y formen los puentes disulfuro en citoplasma, BL21 porque se han reportado en artículos que tiene buenos rendimientos de proteínas que se exportan a periplasma y por último C41 por la razón antes mencionada. (**Fig.18**). La cepa C41 sigue siendo la más efectiva para la producción de la proteína.

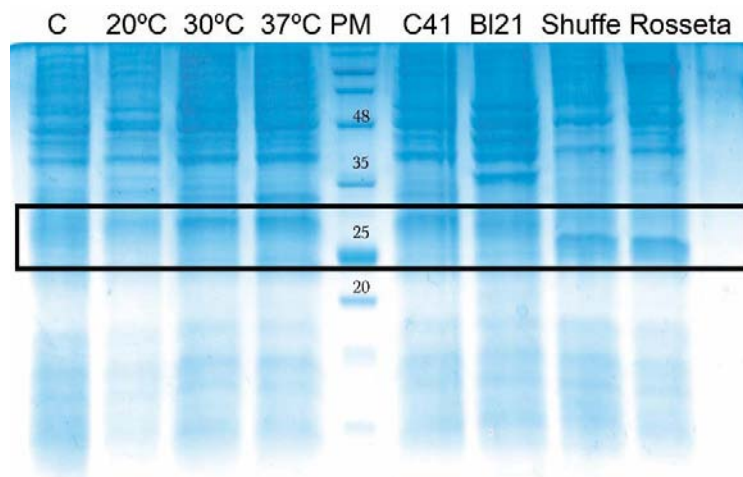


Fig.18. Inducción a temperaturas variables y estudio de cepas. En el presente gel de poliacrilamida al 12%, se establece que la temperatura óptima de inducción es de 30°C, en el control (negativo C) no se observa la banda de interés y la cepa con mayor producción de proteína sigue siendo C41.

Por último, se estableció la concentración de IPTG óptima para la inducción, en el gel (**Fig. 19**), no se observa diferencia con la variación de las concentraciones. Por tanto, se seguirá utilizando la concentración al 0.5 mM de IPTG que ha sido definida.

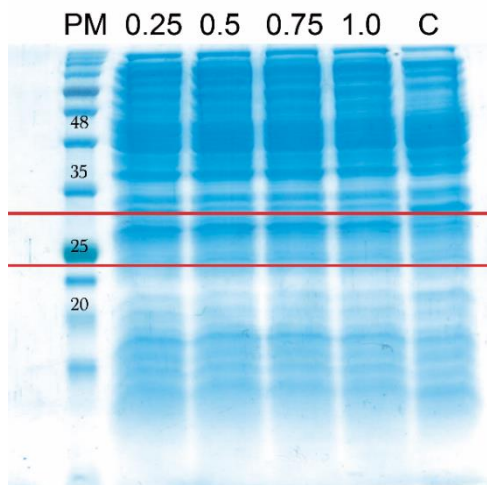


Fig. 19 Experimento a diferentes concentraciones de IPTG. En el gel no se observa diferencia en la concentración de la banda de la PF al variar la concentración.



7.3 Producción de la proteína recombinante

Una vez que se estandarizaron las condiciones de inducción, se produjeron tres litros de medio LB en dos matraces Fernbach. Antes de realizar la inducción, como se mencionó en el procedimiento experimental, se requiere de 50 ml de preinóculo por cada litro y medio de cultivo. En incubación a 37°C y con agitación, se monitoreó la densidad óptica a 600 nm hasta medir una absorbancia de 0.7 y se adicionó la cantidad necesaria de IPTG para obtener una concentración final de 0.5 mM. Finalmente, se indujo a una temperatura de 30°C con agitación durante 16 h.

Para recuperar el pellet, el cultivo se centrifugó a 6500 rpm durante 12 min y el medio fue colectado para analizar la presencia de proteína por SDS-PAGE. Se procedió a realizar la lisis selectiva de periplasma de acuerdo al procedimiento establecido.

7.4 Purificación por CAMQ

Después de la producción se procedió a realizar el primer proceso de purificación de la proteína. Para ello se utilizó la cromatografía de afinidad metal-quelato; donde se retuvo la PF debido al histag o etiqueta de histidinas. En la fracción de elución con 500 mM de Imidazol se obtuvo una banda que corresponde al peso molecular de la PF, además dicha fracción se observaba limpia de algún otro péptido; esto nos facilita tanto el corte con trombina y los siguientes procesos de purificación.

En el caso del medio que se colectó después de centrifugar se ajustó a una concentración de 50 mM de TRIS HCL pH=8, y se procedió a pasar todo medio por la columna. Sin embargo, no se obtuvo PF; lo que nos indica que la proteína no se está perdiendo en el medio.



7.5 Corte con trombina

Se procedió a realizar el corte con la proteasa trombina para obtener la GM-Tamapina de acuerdo al protocolo del diseño experimental, una vez que se terminó la reacción se colecta la trombina y se procede a purificar por CAMQ donde se separan las fracciones de la PF que contiene el Histag y la GM-Tamapina.

7.6 Purificación por HPLC

Antes de iniciar con este proceso de purificación se realizó una extracción en fase sólida (Strata C:18); Esta técnica tiene como objetivo limpiar la muestra de impurezas como sales, péptidos con peso molecular mayores a 75 kDa y concentra la muestra.

En la **Fig.20** se presenta el cromatograma correspondiente a la purificación de la tamapina por HPLC. En él se puede apreciar que el tiempo de retención de la GM-Tamapina expresada en el vector pET32a es 14.8 min (pico A) y se colectó el pico B con un tiempo de retención de 15.2 min que corresponde a la tamapina producida en el vector pET40b.

A partir de la intensidad de la señal cromatográfica, se puede observar que la producción de la proteína utilizando el mismo volumen de medio de cultivo es aproximadamente tres veces menor en comparación con el nuevo sistema de expresión.

A pesar de que se obtiene una menor cantidad de proteína en el nuevo sistema de expresión, se tiene la ventaja que la muestra contiene una menor cantidad de contaminantes.

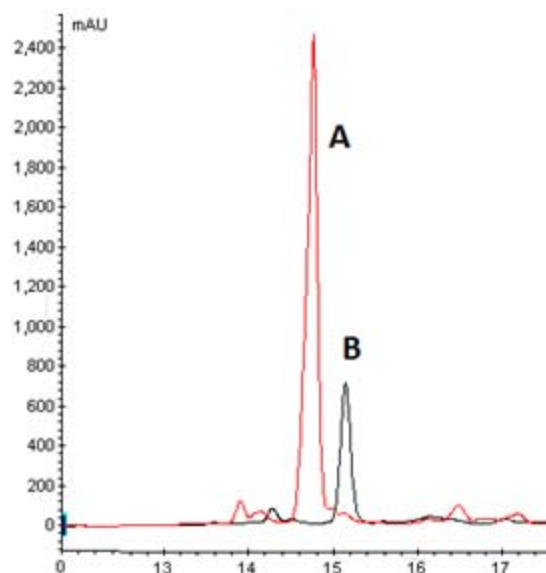


Fig. 20 Cromatograma correspondiente a la purificación de la Tamapina por HPLC

7.7 Caracterización

7.7.1 Dicroísmo

Una vez obtenidos los resultados, se realizó el cálculo para determinar la elipticidad molar con la siguiente fórmula:

$$\theta_{molar} = \frac{MRW \times \theta_{\lambda}}{10 \times d \times c}$$

Donde **MRW** es el peso del residuo medio que se calcula con el peso molecular de la proteína en daltones entre el número de aminoácidos de la cadena menos uno; se representa con la siguiente fórmula: $MRW = M(N-1)$, θ_{λ} es la elipticidad a una determinada longitud de onda, **d** es el largo de la celda en cm y **c** es la concentración en g/ml.³⁹ Finalmente, se gráfica la elipticidad molar en función de la longitud de onda a la cual se le incidió la luz polarizada a la muestra. **Fig.21** En este experimento se comparó el plegamiento que se obtuvo por la producción citoplasmática con la producción periplasma y como se observa ambas tienen el mismo patrón.



De acuerdo a la bibliografía la proteína tamapina posee una alfa hélice y dos betas plegadas, por tanto; Se observa el espectro de DC con estructuras proteicas que contienen α/β ; este tipo de espectro es muy similar a las muestras que contienen solo alfas hélices debido a que regularmente predominan sobre las láminas beta. Así es que, en ambos casos se presentan dos mínimos en 222 y entre 208-210 mn así como un máximo en 190 mn. En el espectro que fue obtenido se logran apreciar esos mínimos y el máximo, siendo los más notorios y se encuentran dentro del intervalo establecido de la bibliografía. ^{39,40}

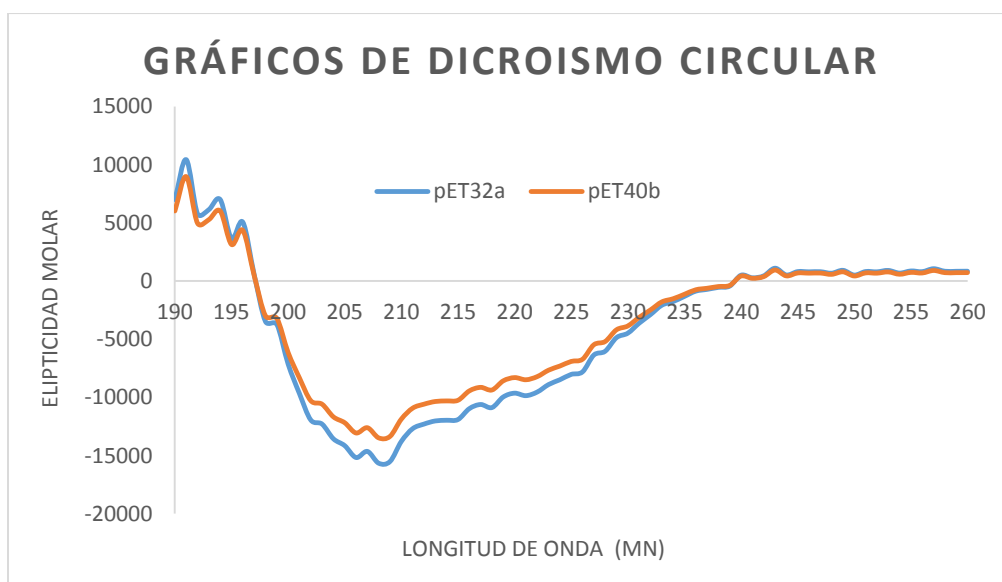


Fig. 21 Espectro de Dicroísmo Circular. En la presente grafica se compara el plegamiento de la tamapina obtenida por la producción en dos sistemas de expresión bacterianos y se observar que ambas poseen el mismo plegamiento. Se determinaron los mínimos experimentales de 208 y 222 mn así como el máximo de 190 mn característico de proteínas α/β . Por tanto, nos percatamos que ambas proteínas se encuentran plegadas.

Mediante esta técnica podemos determinar de forma cualitativa que nuestra proteína se encuentra plegada, para ello se debe de observar el máximo de 190 nm, por el contrario si la proteína se encuentra desplegada o desnaturalizada adquiere un comportamiento desarreglado y en el espectro de DC se observaría una banda intensa en la región negativa. Por último, el espectro nos permite saber



si la proteína tiene puentes disulfuro debido a la presencia de una banda muy pequeña en aproximadamente 228 mn.^{39,40}

7.7.2 Masas MALDI-TOF

La Tamapina posee una secuencia de 31 aminoácidos, sin embargo, como el diseño génico que se realiza para poder producirla de manera recombinante se generan dos residuos extra de aminoácidos (Glicina y Metionina) debido a la secuencia de corte con trombina. Mediante el programa de ExPASy protein parameters se calculó que el peso molecular del péptido obtenido es de 3653.4 Da.

Los resultados de masas MALDI-TOFF nos proporcionan el peso molecular de la proteína después de haber llevado acabo todo su proceso de purificación; el resultado de acuerdo a al espectrómetro de masas es de 3646.319 Da (**Fig.22**). No obstante, de acuerdo al peso molecular teórico existe una diferencia de siete Daltons y podemos inferir que la proteína se encuentra en su estado oxidado, es decir, que se encuentra formando tres puentes disulfuro que corresponden a los seis aminoácidos de cisteína que contiene el péptido y por tanto, se tiene una pérdida de un protón por cada grupo tiol de este aminoácido. Finalmente, se tiene un Dalton de sobra y puede ser justificado por un protón que se haya adicionado al péptido para que ionice la biomacromolécula y pueda ser leída por el equipo de masas. También mediante este resultado podemos justificar que no hay presencia de formación de dímeros o polímeros de nuestra proteína; si fuera el caso, se presentaría el doble o el equivalente del peso molecular en masas.

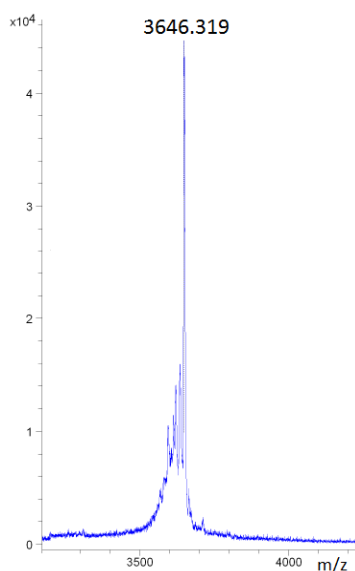


Fig.22. Espectro de Masas MALDI-TOFF de la Tamapina. En el espectro se observa el ion molecular que nos proporciona el peso molecular de la proteína; por tanto, con esta información podemos asegurar que se encuentra en su estado oxidado.

7.7.3 RMN

Se obtuvo el espectro de protón mediante RMN. En la **Fig. 23**, se observan señales dispersas bien definidas; lo que indica que la proteína se encuentra plegada. Esta dispersión de los protones NH se debe a los puentes de hidrógeno en su estructura secundaria.

A campo bajo entre 6.5 ppm y 9.5 ppm podemos observar las señales que corresponden a protones aromáticos y de amida. A campo alto, encontramos principalmente los protones de cadenas alifáticas. Por ejemplo: entre 0.5 a 1 ppm se observan los protones de metilo. Adquirir el espectro de protón es importante debido a que podemos determinar si la proteína se encuentra plegada. Por otra parte, se requiere este espectro como estudio preliminar para poder realizar los experimentos en 2D tales como: COSY, TOCSY y NOESY que son esenciales para obtener la elucidación estructural de la proteína.

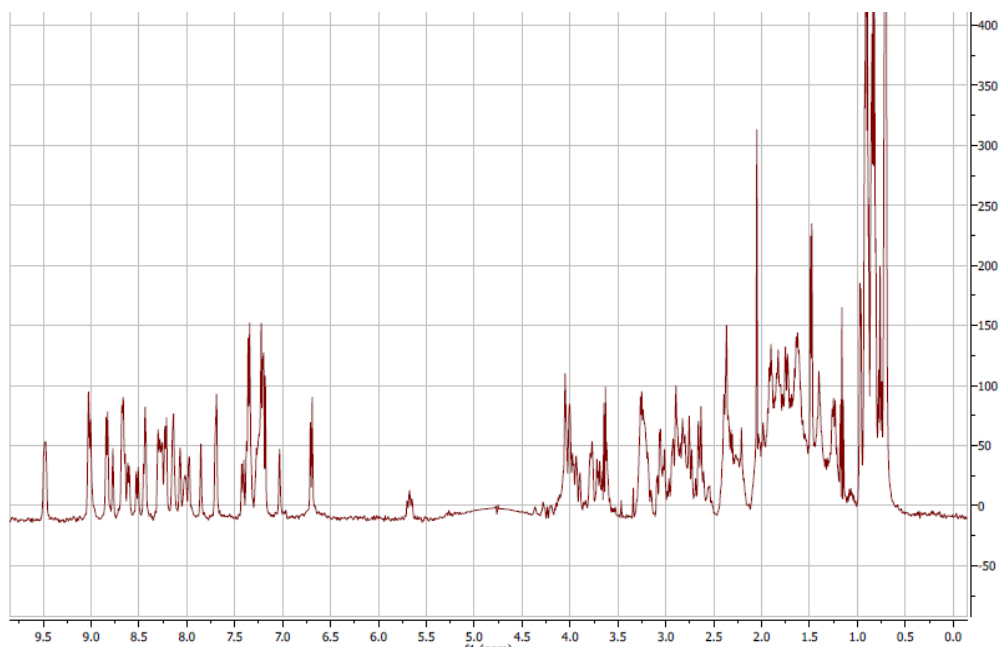


Fig. 23 Espectro mononucler de protón (^1H). En el presente espectro de RMN protónica fue obtenido a 25°C y se realizó supresión de agua. Se observan señales dispersas y definidas indicándonos que la proteína se encuentra plegada.

Se obtuvo un rendimiento final de 0.5 mg/L de cultivo, es decir, casi tres veces menor de acuerdo con el sistema establecido anteriormente. En el caso particular de la tamapina ha sido comprobado que su plegamiento de forma recombinante cumple con las conectividades nativas de puentes disulfuro. Por tanto, al comparar el espectro de protón de la misma proteína obtenida por el nuevo sistema de expresión; podemos concluir que la conectividad de puentes disulfuro y plegamiento idéntico al reportado.



8. CONCLUSIONES

Se logró la construcción del vector pET40b con el inserto del gen de la tamapina y se optimizó la producción de la proteína en medio LB con los parámetros encontrados.

De acuerdo a los resultados de caracterización de la GM-Tamapina se concluye que se encuentra en su estado oxidado. Se confirmó la presencia de hojas beta y hélices alfa; y finalmente, se determinó que la proteína se encuentra plegada.

En este proyecto se comprobó que la proteína que se exporta a periplasma adquiere el mismo plegamiento comparado con la producción que se lleva a cabo en citoplasma. Sin embargo, el rendimiento de la producción fue bajo.



9. BIBLIOGRAFÍA

1. Santibáñez C., Possani L. (2015) Overview of the Knottin scorpion toxin-like peptides in scorpion venoms: Insights on their classification and evolution. *Toxicon* 107, 317-326.
2. Pedarzani P., Dieter, D., Doorty, K. B., Wadsworth, J. D. F., Joseph, J. S., Jeyaseelan, K., Strong, P. N. (2002). Tamapin , a Venom Peptide from the Indian Red Scorpion (Mesobuthus tamulus) That Targets Small Conductance Ca^{2+} -activated K^+ Channels and Afterhyperpolarization Currents in Central Neurons, 277 (48), 46101–46109.
3. Ramírez, B., Toledano, Y., Cano, P., Hernández, R., Flores, D., Saucedo, A., Del Río, F. (2014). Cytotoxicity of recombinant tamapin and related toxin-like peptides on model cell lines. *Chemical Research in Toxicology*, 27(6), 960–967.
4. Hoffman A. (1993) El Maravilloso Mundo de los Arácnidos. Fondo de Cultura Económica: La Ciencia desde México. México D.F. III.
5. Manual de procedimientos Estandarizados para la vigilancia epidemiológica de la intoxicación por picadura de Alacrán, Dirección general de epidemiología de la secretaria de salud. (2012).
6. Norma Oficial Mexicana NOM-033-SSA2-2002, Para la vigilancia, prevención y control de la intoxicación por picadura de alacrán. <http://www.salud.gob.mx/unidades/cdi/nom/033ssa202.html>. Consultada el 16 de marzo de 2016.
7. Ricardo C., Rodríguez de la vega, Elisabeth F. Schwartz, Lourival D. Possani. Mining on scorpion venom biodiversity (2010). *Toxicon* 56 1155-1161.
8. Ramirez B. (2009) Expresión, Caracterización y Determinación Estructural por RMN de la GS-tamapina (Tesis de Maestría), Universidad Nacional Autónoma de México, Instituto de Química.
9. Nelson D., Cox M. (2009) Lehninger principios de Bioquímica Quinta edición, España. pp.
10. Whitford D. (2005) Proteins Structure and function, John Wiley and Sons, Ltd,



Spain

11. Liljas A., Liljas L., Piskur J., Lindblom J., Nissen P., Kjeldgaard M., (2011) Text book of structural biology. World Scientific Printers, Singapore.
12. Simons N. (2012) Principles of genetics. 5ta edición, editorial Wiley, Impreso en Estados Unidos de América.
13. Charlotte K., Paul F. (2004) Genes and DNA a beginner's guide to genetics and its applications. Impreso en Estados Unidos de America, Columbia University Press New York.
14. Pierce B. (2009) *Génética un enfoque conceptual*. Editorial Médica Panamericana.
15. Devlin T. (2011) Text book of Biochemistry with clinicar correlation. 7th ed. Joun Wilet and sons, Inc. Printed in the United Stats of America, Tercera edición.
16. Lodish H., Berk A., Kaiser C., Krieger M., Bretscher A., Ploegh H., Amon A., Scott M. (2013). *Molecular Cell Biology*. Primera edición, Editorial Freeman and Company, Impreso en Estados Unidos de América.
17. Griffiths A., Miller J., Suzuki D., Lewontin R., Gelbart W. (2002) *Genética*. Séptima Edición, Editorial McGraw-Hill, Impreso en España.
18. Brown T. (2008) *Genomas*. 3^{ra} Edición, Impreso en Argentina, Editorial Médica Panamericana.
19. Zhuo, X., Zhang, Y., Guan, Y., & Yao, S. (2014). Co-expression of disulfide oxidoreductases DsbA / DsbC markedly enhanced soluble and functional expression of reteplase in *Escherichia coli*. *Journal of Biotechnology*, 192, 197–203.
20. Dale J., Park S. (2004). *Molecular generics of bacteria*. 4ta edición, Editorial John Wiley and Sons, Ltd, Impreso en Gran Bretaña.
21. Novagen. pET System Manual. 11th Manual. 2005
22. Parnham M., Bruinvels J. (2001) *Recombinant protein drugs*. Springer Basel
23. Berkmen, M. (2012). Production of disulfide-bonded proteins in *Escherichia coli*. *Protein Expression and Purification*, 82(1), 240–251.



24. Malik, A., Rudolph, R., & Söhling, B. (2006). A novel fusion protein system for the production of native human pepsinogen in the bacterial periplasm, *47*, 662–671.
25. Ewelina, Ł., Kowalczyk, T., Olejniczak, S., & Sakowicz, T. (2016). Protein Expression and Purification Extraction and purification methods in downstream processing of plant-based recombinant proteins, *120*, 110–117.
26. Zhuo, X., Zhang, Y., Guan, Y., & Yao, S. (2014). Co-expression of disulfide oxidoreductases DsbA / DsbC markedly enhanced soluble and functional expression of reteplase in *Escherichia coli*. *Journal of Biotechnology*, *192*, 197–203.
27. Gross P, Burkart S., Muller R. (2014) Analytics of the therapeutic peptide aviptadil by sheathless CE-MS and comparison with nanoRP-HPLC-MS.
28. Cunico R., Gooding K., Wehr T. (1998) Basic HPLC and CE of Biomolecules. Bay Bioanalytical Laboratory Inc., Impreso Estados Unidos de América.
29. Berg J., Tymoczko J., Stryer L. (2007) Bioquímica. 6ta edición, editorial reverté, impreso en España.
30. López M., Dopico M., López J., González M. (2016) MALDI-TOF to compare polysaccharide profiles from commercial health supplements of different mushroom species. *Food Chemistry* 597-604.
31. Rule G., Hitchens T. (2006) Fundamental of Protein NMR Spectroscopy. Springer, Printed in the Netherlands.
32. Del Rio F., Ramirez B. Nuclear magnética resonance of proteins. Departamento de Bioquímica, Instituto de Química, UNAM. Research signpost.
33. Keeler, James. Understanding NMR spectroscopy. Cambridge: John Wiley and Sons, 2005
34. Hore P. (2005) Nuclear Magnetic Resonance, Universidad de Oxford, Impreso en Gran Bretaña.
35. Negroni M. Microbiología Estomatológica fundamentos y guía práctica (1999) Editorial Médica Panamericana
36. Zhuo, X., Zhang, Y., Guan, Y., & Yao, S. (2014). Co-expression of disulfide



- oxidoreductases DsbA / DsbC markedly enhanced soluble and functional expression of reteplase in *Escherichia coli*. *Journal of Biotechnology*, 192, 197–203.
37. [http://www.emdmillipore.com/US/en/product/Rosetta%E2%84%A2-2\(DE3\)-Competent-Cells](http://www.emdmillipore.com/US/en/product/Rosetta%E2%84%A2-2(DE3)-Competent-Cells). Consultada el 25 de Marzo de 2016.
38. <https://www.neb.com/products/c3026-shuffle-t7-competent-e-coli>. Consultada el 25 de Marzo de 2016.14
39. Kelly, S. M., Jess, T. J., & Price, N. C. (2005). How to study proteins by circular dichroism, 1751, 119–139.
40. Gerald D. (1996) *Circular Dichroism and Conformational Analysis of Biomolecules*. Editorial Plenum Press, Impreso en Estados Unidos de América.



Apéndice A. Abreviaturas de los 20 aminoácidos presentes en la naturaleza con su clasificación del grupo R

Nombre	Abreviaturas		Clasificación del grupo R
	Una letra	Tres letras	
Alanina	A	ALA	Alifático no polar
Glicina	G	GLY	Alifático no polar
Isoleucina	I	ILE	Alifático no polar
Leucina	L	LEU	Alifático no polar
Metionina	M	MET	Alifático no polar
Prolina	P	PRO	Alifático no polar
Valina	V	VAL	Alifático no polar
Fenilalanina	F	PHE	Aromático
Tirosina	Y	TYR	Aromático
Triptófano	W	TRP	Aromático
Aspartato	D	ASP	Cargado negativamente
Glutamato	E	GLU	Cargado negativamente
Arginina	R	ARG	Cargado positivamente
Histidina	H	HIS	Cargado positivamente
Lisina	K	LYS	Cargado positivamente
Asparagina	N	ASN	Polar sin carga
Cisteína	C	CYS	Polar sin carga
Glutamina	Q	GLN	Polar sin carga
Serina	S	SER	Polar sin carga
Treonina	T	THR	Polar sin carga

Apéndice B Código genético

		Segunda Base					
		U	C	A	G		
Primera Base	U	UUU } Fenil-alanina UUC } UUA } Leucina UUG }	UCU } Serina UCC } UCA } UCG }	UAU } Tirosina UAC } UAA } Stop codon UAG } Stop codon	UGU } Cysteine UGC } UGA } Stop codon UGG } Tryptophan	U C A G	
	C	CUU } Leucina CUC } CUA } CUG }	CCU } Prolina CCC } CCA } CCG }	CAU } Histidina CAC } CAA } Glutamina CAG }	CGU } Arginina CGC } CGA } CGG }	U C A G	
	A	AUU } Isoleucina AUC } AUA } AUG } Metionina start codon	ACU } Treonina ACC } ACA } ACG }	AAU } Asparagina AAC } AAA } Lisina AAG }	AGU } Serina AGC } AGA } Arginina AGG }	U C A G	
	G	GUU } Valina GUC } GUA } GUG }	GCU } Alanina GCC } GCA } GCG }	GAU } Ácido Aspártico GAC } GAA } Acido Glutámico GAG }	GGU } Glicina GGC } GGA } GGG }	U C A G	



Apéndice C. Composición del medio LB

Para un litro de cultivo:

<i>Componente</i>	<i>Cantidad</i>
Triptona	10 g
Extracto de levadura	5 g
NaCl	10 g

Apéndice D. Solución de lisis por sonicación 5X.

<i>Componente</i>	<i>Cantidad</i>
TRIS	50 mM
NaCl	300 mM
Imidazol	20 mM
pH	8

Apéndice E. Componentes de las soluciones para la lisis de periplasma.

Solución A (Buffer de sacarosa o solución hipertónica)

pH=8

<i>Componente</i>	<i>Concentración</i>
Sacarosa	20%
TRIS	50 mM
EDTA	1 mM

Solución B (hipotónica)

<i>Componente</i>	<i>Concentración</i>
MgCl₂	5 mM