



**UNIVERSIDAD NACIONAL AUTÓNOMA DE  
MÉXICO**

**FACULTAD DE QUÍMICA**

**“Expresión y caracterización por métodos  
espectroscópicos de un novedoso dominio  
bacterial de reconocimiento de  
peptidoglicano”**

**T E S I S**

QUE PARA OBTENER EL TÍTULO DE:

**Q U Í M I C A**

**F A R M A C É U T I C A**

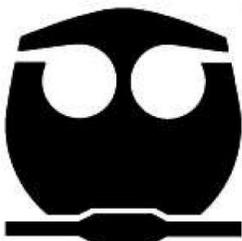
**B I Ó L O G A**

P R E S E N T A:

**EIRA JASIVE GARCÍA PÉREZ**

MÉXICO, D. F.

2013



## **JURADO ASIGNADO:**

**PRESIDENTE:** Profesor: Homero Hernández y Montes

**VOCAL:** Profesor: José Luz González Chávez

**SECRETARIO:** Profesor: José Federico del Río Portilla

**1er. SUPLENTE:** Profesor: Silvia Citlalli Gama González

**2° SUPLENTE:** Profesor: Perla Deyanira Maldonado Jiménez

**SITIO DONDE SE DESARROLLÓ EL TEMA: LABORATORIO 1 DEL DEPARTAMENTO DE QUÍMICA DE BIOMACROMOLÉCULAS Y LABORATORIO DE BIOLOGÍA MOLECULAR, INSTITUTO DE QUÍMICA UNAM**

**ASESOR DEL TEMA: DR. JOSÉ FEDERICO DEL RÍO PORTILLA**

**SUPERVISOR TÉCNICO: DRA. ALICIA ACELY GARZA GARCÍA**



**SUSTENTANTE: EIRA JASIVE GARCÍA PÉREZ**

## ÍNDICE

CONTENIDO	PÁGINA
<b>ABREVIATURAS</b> .....	i
<b>1. INTRODUCCIÓN</b> .....	1
<b>1.1 HIPÓTESIS</b> .....	3
<b>1.2 OBJETIVOS</b> .....	3
<b>2. ANTECEDENTES</b> .....	4
<b>2.1 PÉPTIDOS Y PROTEÍNAS</b> .....	5
2.1.1 Arquitectura proteica.....	5
2.1.2 Síntesis de proteínas (Replicación, Transcripción y Traducción).....	7
<b>2.2 SÍNTESIS DE PROTEÍNAS RECOMBINANTES</b> .....	8
2.2.1 Vectores de clonación (plásmidos).....	8
2.2.2 Transformación genética.....	9
2.2.3 Sistema T7 para la expresión de proteínas recombinantes.....	10
2.2.4 Proteína de fusión.....	10
2.2.5 Proteasas.....	11
2.2.6 Enriquecimiento isotópico.....	14
<b>2.3 PURIFICACIÓN DE PROTEÍNAS</b> .....	15
2.3.1 Cromatografía de afinidad metal- quelato.....	15
2.3.2 Cromatografía Líquida de Alta Presión (HPLC).....	16
<b>2.4 RESONANCIA MAGNÉTICA NUCLEAR DE PROTEÍNAS</b> .....	17
<b>2.5 NAMLA AMIDASAS</b> .....	20
2.5.1 Peptidoglicano.....	21
2.5.2 Clasificación de las NAMLA amidasas.....	22
2.5.3 NAMLA amidasas como blancos para combatir enfermedades bacterianas.....	22
2.5.4 Dominios de reconocimiento de carbohidratos relacionados con la familia Amidasa 3.....	24
2.5.5 Dominios de reconocimiento de carbohidratos.....	24
2.5.6 Dominio Amidasa 2-C de una NAMLA amidasa presente en <i>Bacillus</i> .....	25
2.5.7 Estudios bioinformáticos del dominio Amidasa 2-C de <i>Bacillus</i> .....	26

<b>3. METODOLOGÍA</b> .....	30
<b>3.1 EXPRESIÓN DE LA PROTEÍNA DE FUSIÓN GB1- AMIDASA 2C EN <i>E.coli</i> Rosetta</b>	31
3.1.1 Expresión en medio Lysogeny Broth (LB).....	32
3.1.2 Expresión en medio mínimo PG.....	33
<b>3.2 PURIFICACIÓN DE LA PROTEÍNA RECOMBINANTE</b> .....	34
3.2.1 Cromatografía de afinidad metal quelato de la proteína de fisión.....	34
3.2.2Corte de la proteína de fusión.....	34
3.2.3 Cromatografía Líquida de Alta Presión (HPLC).....	35
<b>3.3 CARACTERIZACIÓN DEL DOMINIO AMIDASA 2-C</b> .....	35
3.3.1 Espectrometría de masas.....	35
3.3.2 MALDI-TOF.....	36
<b>3.4 ANÁLISIS POR RESONANCIA MAGNÉTICA NUCLEAR (RMN)</b> .....	36
<b>4. RESULTADOS Y DISCUSIÓN</b> .....	38
<b>4.1 EXPRESIÓN DE LA PROTEÍNA DE FUSIÓN GB1- AMIDASA 2C</b> .....	39
4.1.2 Monitoreo de las expresiones.....	39
4.1.3 Condiciones para la expresión en medio Lysogeny Broth (LB).....	39
4.1.4 Condiciones para la expresión en medio mínimo PG.....	40
4.1.5 Lisis celular.....	43
<b>4.2 PURIFICACIÓN DE LA PROTEÍNA DE FUSIÓN</b> .....	44
4.2.1 Purificación por la columna de afinidad metal – quelato.....	44
4.2.2 Corte de la proteína de fusión.....	45
4.2.3 Cromatografía Líquida de Alta Presión (HPLC).....	46
<b>4.3 CARACTERIZACIÓN DEL DOMINIO AMIDASA 2-C</b> .....	48
4.3.1 Espectrometría de masas.....	48
4.3.2 Máximo de absorción en la región de UV.....	51
<b>4.4 RESONANCIA MAGNÉTICA NUCLEAR</b> .....	52
<b>5. CONCLUSIONES</b> .....	58
<b>6. BIBLIOGRAFÍA</b> .....	61

<b>7. APÉNDICES.....</b>	<b>67</b>
<b>APÉNDICE A. PROTOCOLO DE CÉLULAS COMPETENTES Y TRANSFORMACIÓN.....</b>	<b>68</b>
<b>APÉNDICE B. COMPOSICIÓN DE LOS MEDIOS DE CULTIVO.....</b>	<b>70</b>
<b>APÉNDICE C. PROTOCOLOS DE EXPRESIÓN PARA 1 L DE MEDIO MÍNIMO.....</b>	<b>72</b>
<b>APÉNDICE D. DISOLUCIONES AMORTIGUADORAS.....</b>	<b>76</b>
<b>APÉNDICE E. PROTOCOLO PARA PURIFICAR LA PROTEÍNA DE FUSIÓN POR LA COLUMNA DE AFINIDAD METAL QUELATO.....</b>	<b>77</b>
<b>APÉNDICE F. CUANTIFICACIÓN DE PROTEÍNAS POR EL MÉTODO DE BRADFORD.....</b>	<b>78</b>

---

# *Abreviaturas*

ABREVIATURA	SIGNIFICADO
NAMLA	N-acetilmuramoil-L-alanina amidasas
ADN	Ácido Desoxirribonucleico
ARN	Ácido Ribonucleico
tARN	ARN de transferencia
mARN	ARN mensajero
GB1	Dominio B1 de la proteína G de <i>Streptococcus sp.</i>
TEV	Proteasa Tobacco Etch Viruse
IPTG	Isopropil $\beta$ -tiogalactopiranosido
PMSF	Fluoruro de fenilmetilsulfonilo
KDa	Kilo Dalton; Da = 1g/mol
MALDI- TOF	MALDI, matrix-assisted laser desorption ionization TOF, time of flight
HPLC	Cromatografía líquida de alta presión
RP-HPLC	Cromatografía líquida de alta presión en fase reversa
DEUA	Dólares de Estados Unidos de América
COSY	Correlation Spectroscopy
TOCSY	Total Correlation Spectroscopy
NOESY	Nuclear Overhauser Effect Spectroscopy
HSQC	Heteronuclear Single-Quantum Correlation
<sup>15</sup> N- HSQC	Heteronuclear Single-Quantum Correlation <sup>15</sup> N y <sup>1</sup> H
<sup>13</sup> C- HSQC	Heteronuclear Single-Quantum Correlation <sup>13</sup> C y <sup>1</sup> H
TFA	Ácido trifluoroacético
DTT	Ditiotreitól

**MS**

Eespectrometría de masas

---

# *Introducción*

## 1. INTRODUCCIÓN

Las bacterias, a diferencia de las células eucariotas, tienen una pared celular que las hace más rígidas y que está formada de carbohidratos unidos covalentemente a pequeños péptidos, lo que se conoce como peptidoglicano.<sup>(1)(2)</sup>

Las N-acetilmuramoyl-L-alanina (NAMLA) amidasas son proteínas que reconocen la pared celular y catalizan la hidrólisis del enlace entre el aminoácido alanina y el monosacárido N-acetilmurámico del peptidoglicano.<sup>(3)</sup> Estas enzimas son utilizadas por las bacterias en algunas etapas de su ciclo celular ya que durante la citocinesis la pared celular tiene que ser hidrolizada para que las células se puedan separar.<sup>(3)(4)</sup> También hay amidasas NAMLA en el sistema inmune innato de la gran mayoría de los animales,<sup>(3)</sup> donde tienen una función importante en la defensa contra las bacterias patógenas y son expresadas también por los bacteriófagos donde son necesarias para la infección y el metabolismo viral.<sup>(1)(3)(5)</sup> Las bacterias con actividad NAMLA amidasa anormal presentan defectos citológicos drásticos que implican el crecimiento bacteriano en cadenas. Esto sugiere que las NAMLA amidasas pueden ser nuevos blancos para fármacos antibacterianos ahora que los antibióticos están presentando problemas de salud pública como la resistencia que adquieren las bacterias debido a que los fármacos utilizados actualmente suelen ser variantes de un fármaco original.<sup>(4)(6)</sup>

Las NAMLA amidasas son proteínas modulares. Además de poseer el dominio enzimático, cada NAMLA amidasa tiene uno o más dominios que reconocen carbohidratos y/o peptidoglicano u otros componentes de la pared celular y de esta manera guían al dominio enzimático hacia su sustrato. Existen muchas clases de dominios de reconocimiento y se cree que son éstos los que determinan la función específica de cada una de las muchas NAMLA amidasas que expresan las bacterias. Las NAMLA amidasas se dividen en dos familias: Amidasa2 y Amidasa3, de acuerdo con la secuencia de aminoácidos del dominio enzimático.<sup>(7)</sup> La NAMLA amidasa que se ha estudiado en esta investigación es una NAMLA de la familia Amidasa3, la cual es producida por *Bacillus cereus* y se encuentra clasificada en Uniprot con el número de acceso Q736Q1\_BACC1.<sup>(8)</sup> Está formada por el dominio enzimático Amidasa 3 en el extremo N-terminal y en el extremo C-terminal lo que

se presume es un dominio de reconocimiento de carbohidrato y/o peptidoglicano denominado Amidasa 2-C.

El dominio Amidasa 2-C está presente tanto en bacterias como en virus, y ha sido muy poco estudiado<sup>(9)</sup>, actualmente no existe ninguna información sobre la función que realizan las proteínas que lo contienen o la naturaleza del carbohidrato que reconoce.

En esta tesis se reportan los experimentos realizados para la identificar la expresión, el marcaje isotópico, la purificación, la caracterización y estudios preliminares de Resonancia Magnética Nuclear (RMN) del dominio Amidasa 2-C.

Este trabajo representa el preámbulo del estudio estructural del dominio Amidasa 2-C, así como de las interacciones de este dominio con el peptidoglicano bacteriano mediante Resonancia Magnética Nuclear.

### 1.1 HIPÓTESIS

El dominio Amidasa 2-C podrá ser sobre expresado en la bacteria *Escherichia coli*, enriquecido con los isótopos  $^{15}\text{N}$  y  $^{13}\text{C}$ , aislado de la bacteria y caracterizado por espectroscopía.

### 1.2 OBJETIVOS

- Expresar el dominio Amidasa 2-C recombinante en *E. Coli*.
  - Establecer las condiciones experimentales para la expresión del dominio Amidasa 2-C en medio Lysogeny Broth (LB) y medio mínimo PG.
  - Identificar los beneficios de los métodos utilizados para incorporar isótopos activos en RMN.
  - Comparar el grado de expresión al utilizar el método de reducción de volumen.
- Aislar y purificar el dominio Amidasa 2-C del sistema productor.
  - Establecer las condiciones óptimas para la purificación de la proteína Amidasa 2-C
- Caracterizar el dominio Amidasa 2-C mediante métodos espectroscópicos (UV-VIS, espectrometría de masas y RMN).

---

# *Antecedentes*

## 2. ANTECEDENTES

### 2.1 PÉPTIDOS Y PROTEÍNAS

#### 2.1.1 Arquitectura proteica <sup>(10)</sup>

Las proteínas desempeñan una enorme variedad de funciones: unas transportan y almacenan moléculas pequeñas, otras son estructurales y algunas son enzimas. Dependiendo del papel que representan dentro de un organismo es el grado de complejidad o estructuración de cada una.

Todas las proteínas son polímeros y los  $\alpha$ -aminoácidos son los monómeros que se combinan para formarlas. La estructura general de éstos se muestra en la Figura 2.1, en donde R representa la cadena lateral perteneciente a cada uno, el resto de los átomos que no son la cadena lateral constituyen la cadena principal o “columna vertebral” de la proteína. Los distintos  $\alpha$ -aminoácidos se diferencian por sus cadenas laterales, éstas puede contener cadenas alifáticas, grupos aromáticos o grupos que proporcionan características ácidas, básicas o polares a una proteína. En los genes de todos los organismos están codificados veinte aminoácidos diferentes que se incorporan a las proteínas, todos éstos a excepción de la glicina pueden existir en la forma D o L, ya que en todos los casos el carbono  $\alpha$  es quiral. El hecho importante es que todos los aminoácidos que incorporan los organismos a las proteínas son de la forma L.

Los aminoácidos pueden unirse entre ellos en forma covalente por formación de un enlace de amida entre el grupo  $\alpha$ - carboxilo de uno y el grupo  $\alpha$ -amino de otro, éste enlace se denomina enlace peptídico y los productos formados de esta unión se llaman péptidos, si la cadena de aminoácidos es muy larga se llama polipéptido. Las proteínas son polipéptidos de secuencia definida, dicha secuencia es la que compone la *estructura primaria* de una proteína.

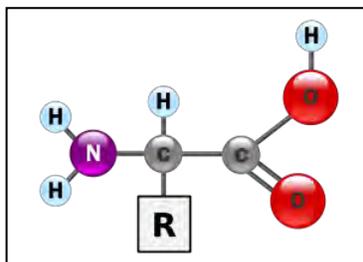


Figura 2.1. Estructura general de los aminoácidos.

La *estructura secundaria* de una proteína consiste en el tipo de organización local que adquieren los segmentos del polipéptido mediante diferentes tipos de interacciones establecidas por la cadena principal. Hay tres tipos de estructuras secundarias regulares: hélice, hoja plegada  $\beta$ , y

vuelta. En una hélice la cadena de aminoácidos parece encontrarse parcialmente enrollada en regiones de estructura helicoidal, hay más de un tipo de hélice dependiendo de la longitud de cada giro y de la dirección a la cual la cadena principal se enrolla, pero la hélice más común es la hélice  $\alpha$ , que se caracteriza por estar enrollada hacia la derecha y tener 3.6 aminoácidos en cada giro de la hélice. La hoja plegada  $\beta$  es una estructura en forma de lámina extendida formada de dos o más segmentos de la cadena principal que se acomodan uno al lado del otro, ya sea en forma paralela o anti-paralela, cada uno de los segmentos que forman una hoja  $\beta$  se denomina hebra  $\beta$ . Las vueltas son estructuras que permiten cambiar la dirección de la cadena principal, aunque las vueltas generalmente no tienen una estructura regular periódica como las hélices y las hojas  $\beta$ , son comúnmente rígidas y de estructura definida.<sup>(11)</sup>

La *estructura terciaria* de las proteínas se adquiere cuando las formas de estructura secundaria interactúan entre sí originando estructuras globulares y compactas, este nivel de estructuración se estabiliza por interacciones entre los grupos de las cadenas laterales (interacciones de van der Waals, interacciones carga-carga, dipolo-dipolo, puentes de hidrógeno) El efecto más común que resulta en interacciones entre los elementos de estructura secundaria es el efecto hidrófobo, en el que los residuos de éstas características se sitúan hacia el interior de la estructura y aquellos con características hidrofílicas quedan en el exterior en contacto con el agua. Otra manera de estabilizar estructuras terciarias es la formación de puentes disulfuro por los residuos de cisteína.

La mayoría de las proteínas está formada por más de un *dominio*. Un dominio es una región compacta de la estructura terciaria que en principio puede plegarse independientemente del resto del polipéptido. Los dominios están conectados entre sí por la cadena polipeptídica formando una estructura que semeja las cuentas en un collar, aunque es común que los dominios interactúen entre ellos.

Puede existir un cuarto nivel de estructuración, la *estructura cuaternaria*, ésta se da por la interacción entre dos o más cadenas polipeptídicas plegadas. Las subunidades que interactúan para dar lugar a este nivel de estructura pueden ser idénticas o casi idénticas (homotípicas) o diferentes (heterotípicas) Este nivel de estructura se estabiliza de la misma manera que la estructura terciaria.

### 2.1.2 Síntesis de proteínas <sup>(10)</sup>

Cada organismo lleva en cada una de sus células al menos una copia de toda la información genética que necesita (genoma) Normalmente esta información genética está codificada en el ADN de doble cadena a excepción de algunos virus. En cada organismo una fracción importante del ADN contiene información para la síntesis de proteínas. Los segmentos de ADN que codifican para una proteína (aunque también aquellos que codifican para la producción de moléculas de ARN funcionales para la célula como los ARNs de transferencia (tARN) o ribosomales) se denominan *genes*.

El dogma central de la biología molecular establece que la información para la síntesis de proteínas está codificada en el ADN y que esta información primero se copia a moléculas ARN mensajero (mARN) mediante un proceso llamado transcripción. Las moléculas de ARN son entonces leídas por el ribosoma, el cual forma la cadena polipéptica mediante el ensamblado de los aminoácidos individuales, a éste proceso se le denomina traducción.

### **Replicación**

En la replicación se construye una copia complementaria de cada una de las dos cadenas de ADN, lo que normalmente da lugar a dos copias idénticas de la original. La replicación se produce mediante un complejo de enzimas que está concentrado sobre una proteína denominada ADN polimerasa, éste complejo permite abrir el ADN y guiar el apareamiento de cada desoxirribonucleósido trifosfato, así como catalizar la formación del enlace fosfodiéster.

### **Transcripción**

De igual manera que el ADN puede dirigir la replicación, también guía la transcripción que es la formación de una cadena de ARN complementaria. Naturalmente, los monómeros empleados para la transcripción son los ribonucleósidos trifosfato. La transcripción requiere un conjunto de enzimas denominadas ARN polimerasas.

### **Traducción**

A partir del ADN se transcriben copias complementarias de los genes que es necesario traducir en forma de mRNA que transporta la información desde el ADN hasta la maquinaria de síntesis proteica.

Las moléculas de mRNA se unen a un ribosoma y las moléculas de tARN llevan a los aminoácidos al ribosoma de uno en uno. Cada tARN identifica el codón apropiado en el mRNA y añade éste aminoácido a la cadena proteica en crecimiento. El ribosoma se desplaza a lo largo del mRNA.

Cuando se libera la cadena polipeptídica de un ribosoma, no está necesariamente acabada, debe plegarse en su estructura tridimensional correcta y en algunos casos es preciso que se formen enlaces disulfuro, éste proceso generalmente se lleva a cabo por un conjunto de proteínas y enzimas denominadas chaperonas. Cabe mencionar que después de la síntesis proteica, algunas de las proteínas sufren diferentes modificaciones postraduccionales tales como glicosilación, acilación y proteólisis limitada.

## **2.2 SÍNTESIS DE PROTEÍNAS RECOMBINANTES**

Los conocimientos actuales del proceso de síntesis de proteínas han proporcionado la posibilidad de introducir modificaciones genéticas en los organismos vivos mediante técnicas de ADN recombinante.<sup>(10)</sup>

### **2.2.1 Vectores de clonación**

Un vector de clonación es una molécula de ADN pequeño, fácil de aislar y caracterizar, con una zona de secuencias de restricción conocida (sitios que pueden ser cortados por enzimas de restricción concretas), de fácil introducción a la célula anfitriona y una vez allí con capacidad de replicación autónoma. Es conveniente que el vector posea un gen marcador (por ejemplo, de resistencia antibióticos) que permita identificar a la células que llevan el ADN recombinante.

La misión del vector es funcionar como unidad replicadora del fragmento de ADN que se quiere introducir a la célula anfitriona. La introducción del dicho fragmento (el inserto) en un vector se conoce como clonación. En éstos días existen diferentes técnicas para unir un inserto con un vector, pero la técnica tradicional consiste en digerir tanto el inserto como el

vector con una (o dos) enzimas de restricción que generen extremos de secuencia complementaria conocidos como cohesivos o “pegajosos”. Cuando el vector y el inserto son mezclados éstos se unen con el fragmento respectivo, como la región complementaria es de tan sólo unas cuantas bases y no es muy estable, es necesario añadir una ligasa de ADN que forma los enlaces fosfodiéster entre el inserto y el vector dando lugar a una sola molécula de ADN de doble hebra llamada ADN recombinante. Los plásmidos, virus, BACs, PACs, YACs, cósmidos y fagémidos son ejemplos de vectores de clonación.<sup>(12)</sup> Las proteínas que se obtienen de la expresión de genes del ADN recombinante se denominan proteínas recombinantes.<sup>(13)</sup>

### **Plásmido**

Los plásmidos son moléculas de ADN de doble hebra de pequeño tamaño (2 a 5 Kpb) y de forma circular. Permiten la incorporación de insertos de un tamaño máximo de 10 Kpb. Para su replicación autónoma se requiere de una secuencia origen de replicación. El plásmido sufre replications múltiples y origina varias copias por célula (desde unos cuantos hasta varias decenas) Los plásmidos se mantienen estables como círculos extracromosómicos.<sup>(12)</sup>

### **2.2.2 Transformación genética**

La transformación es el proceso por medio del cual una célula bacteriana incorpora ADN extraño o foráneo, de esta manera los plásmidos utilizados como vector de clonación pueden incorporarse al interior de la célula anfitriona. Para esto, la bacteria es tratada con cationes divalentes (generalmente  $\text{Ca}^{2+}$  y/o  $\text{Mg}^{2+}$ ) que provocan la permeabilidad temporal a moléculas pequeñas de ADN, se dice que la célula se hace competente. Aún bajo las mejores condiciones, los plásmidos se establecen solamente en una pequeña población bacteriana, por lo que las células transformantes deben de ser seleccionadas mediante marcadores de selección, ejemplo de esto es la resistencia a antibióticos.<sup>(12)</sup> <sup>(14)</sup> Existen diversos protocolos para producir células competentes y transformación genética, un ejemplo de éstos es el protocolo 25 “Preparation and transformation of competent *E. coli* using Calcium Chloride” del manual de laboratorio Molecular Cloning Volume 1 de Sambrook and Russell.

### 2.2.3 Sistema T7 para la expresión de proteínas recombinantes.

Éste sistema es ampliamente utilizado para la expresión de proteínas recombinantes en *E. coli*. El sistema se basa en el uso de la ARN polimerasa del bacteriófago T7 para la transcripción de la proteína de interés, La ARN polimerasa T7 es altamente selectiva para el promotor T7 y en principio no transcribe ninguna de las proteínas codificadas en el cromosoma bacteriano.

Existen cepas de *E. coli* modificadas que contienen el lisógeno del fago  $\lambda$  DE3 que expresa el gen de la ARN polimerasa T7. Las cepas que han sido modificadas de esta manera son denominadas con las siglas (DE3) y son utilizadas para la expresión de proteínas recombinantes mediante éste sistema.

Para la expresión de la proteína recombinante debe introducirse a una cepa DE3 el plásmido que contiene el gen que codifica para la proteína que se desea producir. La transcripción del gen contenido en el plásmido debe estar controlada por el promotor T7.<sup>(15)</sup>

La adición al medio de lactosa o de un análogo como el IPTG, libera al represor *lac* del promotor *lac*, permitiendo la transcripción del gen de la ARN polimerasa T7. Una vez expresada, la ARN polimerasa T7 procede a su vez a reconocer al promotor T7 en el plásmido y a transcribir exclusivamente al gen recombinante. A diferencia de la lactosa, el IPTG contiene un átomo de azufre que crea un enlace covalente no hidrolizable por la bacteria lo que evita que éste sea degradado progresivamente tras su adición.<sup>(16)</sup>

La Figura 2.2 muestra la regulación de la transcripción mediante el uso del operador *lac*. La Figura 2.3 esquematiza el mecanismo de expresión de una proteína recombinante mediante el sistema T7.

### 2.2.4 Proteína de fusión

Una proteína de fusión es aquella que resulta de expresar un gen artificial híbrido que combina dos o más fragmentos polipeptídicos que no existen juntos en la naturaleza; la construcción resultante se transcribe y traduce como si se tratase de un sólo gen, produciendo una única proteína. Se utilizan para distintos fines que incluyen: 1) añadir una etiqueta a una proteína para facilitar su purificación y/o incrementar su nivel de expresión; 2) producir una proteína que combine las características de dos proteínas naturales; 3) producir una proteína en la que dos actividades diferentes estén físicamente ligadas.<sup>(17)</sup>

Es común expresar una proteína recombinante de interés como proteína de fusión al añadirle proteínas auxiliares (etiquetas de fusión) que proporcionan ventajas en el proceso de obtención de la proteína blanco. Las etiquetas de fusión pueden mejorar la expresión de la proteína, promover el plegamiento, aumentar la solubilidad, y facilitar los procesos posteriores de purificación.<sup>(18)</sup> Existen proteínas que son comúnmente utilizadas como etiquetas de fusión, ejemplos de éstas se muestran en la Tabla 2.1

### 2.2.5 Proteasas

Las proteasas o peptidasas son enzimas que rompen los enlaces peptídicos de las proteínas adicionando una molécula de agua para hacerlo, por lo tanto se clasifican como hidrolasas.<sup>(19)</sup> Las proteasas pueden clasificarse como endopeptidasas cuando actúan en medio de una cadena polipeptídica o exopeptidasas cuando actúan en dipéptidos o tripeptidos situados al final de la cadena, éstas pueden actuar en el extremo amino terminal (aminopeptidasas) o en el extremo carboxilo (carboxipeptidasas).<sup>(20)</sup>

Las proteasas participan en innumerables procesos biológicos, entre otros, la coagulación sanguínea, algunas partes de la resistencia inmunitaria, la apoptosis e incluso la replicación viral.<sup>(20)</sup>

En técnicas de biología molecular, se adiciona dentro de la secuencia de la proteína de fusión un sitio de corte (aproximadamente 7 aminoácidos) para una proteasa en específico. Éste sitio se encuentra entre las secuencias auxiliares y la proteína de interés, lo que permite separar la proteína blanco cuando actúa la proteasa específica para ese sitio. Esta estrategia se utiliza debido a que las proteínas auxiliares pueden influir en el comportamiento de la proteína blanco, por lo tanto, es deseable retirarlas en algún paso del proceso de purificación.<sup>(18)</sup>

Existen varias proteasas que se utilizan para remover las proteínas auxiliares, las más comunes se muestran en la Tabla 2.2.

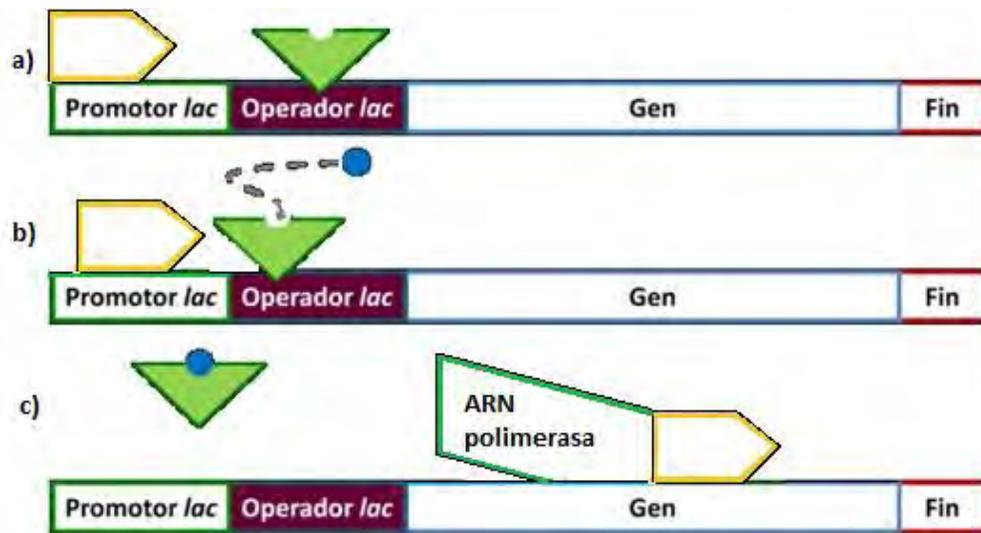


Figura 2.2. Modelo general de regulación de transcripción mediante el uso del operador *lac* e IPTG. El inciso a) muestra la acción de la proteína represora. En ausencia de lactosa, o moléculas análogas como el IPTG, el represor se encuentra fuertemente unido al operador *lac* evitando la transcripción del gen por la ARN polimerasa. b) El IPTG se une al represor ocasionando un cambio de conformación. c) Una vez unido al IPTG el represor pierde su afinidad por el ADN y se desprende, permitiendo a la ARN polimerasa iniciar la transcripción del gen.<sup>(21)</sup>

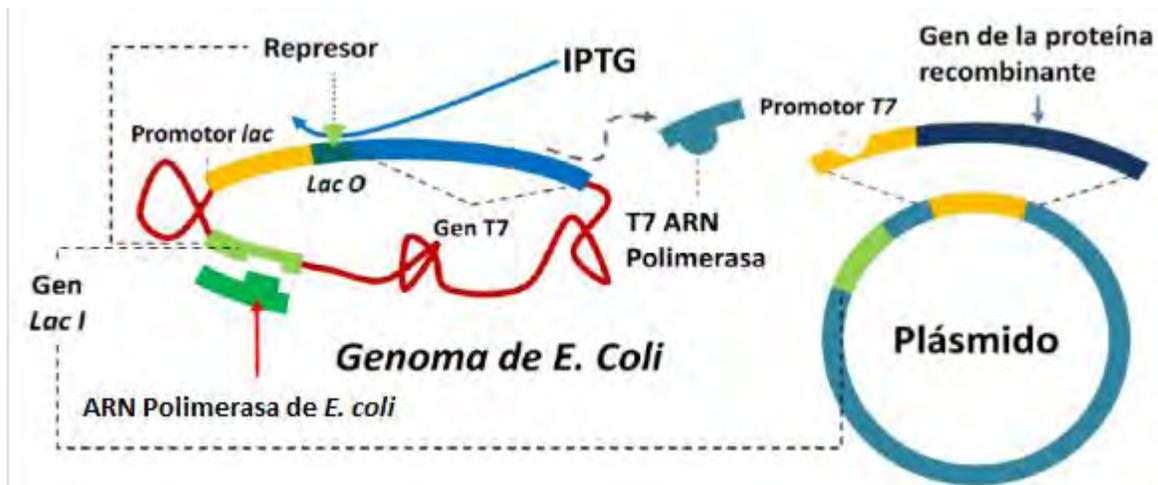


Figura 2.3. Mecanismo de expresión de una proteína recombinante mediante el sistema T7. El gen *lac I* codifica al represor *lac*, el cual es una proteína que se une al operador de lactosa (*lac o*) impidiendo la transcripción del gen de la ARN polimerasa T7. El IPTG se une al represor *lac* induciendo un cambio de estructura que lo hace desprenderse del *lac o*, una vez libre del represor la transcripción de la ARN polimerasa T7 puede llevarse a cabo. A su vez la ARN polimerasa T7 encuentra a su respectivo promotor en el plásmido y comienza la transcripción de la proteína recombinante de interés.<sup>(21)</sup>

ETIQUETA DE FUSIÓN	NÚMERO DE AMINOÁCIDOS	CARACTERÍSTICAS
Tiorredoxina (Trx)	135	Facilita tanto la expresión como la solubilidad cuando se fusiona en el extremo N- terminal de la proteína blanco. Es utilizada para la expresión de proteínas con puentes disulfuro ya que cataliza la reacción ditiol- disulfuro.
Proteína de unión a maltosa (MBP)	390	Facilita la expresión, la solubilidad y la purificación al fusionarse ya sea en el extremo N-terminal o C-terminal de la proteína blanco. Favorece la expresión de la proteína blanco tanto en el citosol como en periplasma. Favorece el plegamiento. Facilita la purificación ya que puede unirse fuertemente a azúcares tales como amilosa y dextrina.
Glutación S- Transferasa (GST)	243	Facilita la expresión de la proteína blanco al catalizar el ataque nucleofílico del glutatión sobre sustratos electrofílicos con el fin de reducir su reactividad con otras biomoléculas. Forma dímeros por lo que podría conducir a la precipitación de la proteína de fusión si la proteína blanco tiende a formar oligómeros. Facilita la purificación al unirse fuertemente al glutatión.
Dominio B1 de la proteína G (GB1) de <i>Streptococcus</i> sp.	56	Es una proteína pequeña y altamente soluble en comparación con otras comúnmente utilizadas como etiquetas de fusión. Por lo tanto, al utilizarla como etiqueta de fusión se obtiene una proteína de fusión altamente soluble. Puede mejorar los rendimientos de expresión, la estabilidad y la solubilidad de la proteína blanco sin afectar su estructura y función. <sup>(22)</sup>
Hexahistidina (His <sub>6</sub> )	6	Consta de una secuencia de 6 histidinas. Es utilizada para facilitar la purificación de la proteína blanco ya cada histidina puede coordinarse con metales como el Ni <sup>2+</sup> , el cual se encuentra unido a una resina de cromatografía de afinidad a metal. Al ser una etiqueta de fusión pequeña, suele utilizarse en combinación con otras etiquetas que promuevan la solubilidad y la expresión de la proteína blanco.

Tabla 2.1. Etiquetas de fusión comúnmente utilizadas para expresar proteínas recombinantes como proteínas de fusión.<sup>(18)</sup>

ENZIMA	DESCRIPCIÓN	SITIO DE RECONOCIMIENTO
<b>Enterocinasa</b>	Subunidad catalítica de enterocinasa bovina.	DDDDK-
<b>Trombina</b>	Serin proteasa que convierte el fibrinógeno en fibrina.	LVPR-GS
<b>Factor Xa</b>	Serin proteasa que convierte la protrombina a trombina.	LVPR-GS I(E/N)GR- No se unirá en el sitio seguido por P ó R
<b>Proteasa TEV</b>	Dominio catalítico de una proteína de inclusión nuclear (NIa) Cistein proteasa del virus del tabaco (TEV)	EXXYXQ-(G/S) Más común: ENLYFQ-X
<b>3C</b>	Forma recombinante de la proteasa 3C de rinovirus humano tipo 14.	EVLFQ-GP

**Tabla 2.2. Proteasas comúnmente usadas para remover proteínas auxiliares de la proteína de fusión.**<sup>(18)(23)</sup>

Se han mencionado las proteasas que se utilizan más comúnmente (Tabla 2.2), sin embargo, enzimas como la trombina, el factor Xa y al enterocinasa pueden cortar de forma promiscua y no específica. Hoy en día una de las proteasas más utilizadas es la proteasa TEV ya que el uso de ésta ofrece diferentes ventajas. Corta eficientemente, es muy específica y es activa en diversas condiciones experimentales, valores de pH, disoluciones amortiguadoras y temperaturas.<sup>(18)</sup> También puede ser producida muy económicamente en bacteria, a diferencia de otras proteasas clásicas como trombina y el Factor Xa que son purificadas de sus fuentes naturales.

### 2.2.6 Enriquecimiento isotópico.

Las señales en RMN de un espectro de <sup>1</sup>H de una proteína se traslapan demasiado, por lo que es necesario incrementar la dispersión de las señales en experimentos de 2 y 3 dimensiones. Cuando una proteína está constituida por más de 50 aminoácidos se requiere que los núcleos atómicos presentes en las proteínas sean activos en RMN (es necesario que posean espín diferente a 0) Sin embargo, la abundancia natural de núcleos activos como <sup>15</sup>N y <sup>13</sup>C presentes en las proteínas es muy baja y por esta razón se recurre a técnicas de biología molecular para obtener proteínas enriquecidas con éstos isótopos.<sup>(24)</sup>

Para obtener proteínas marcadas isotópicamente se hace crecer a la cepa de expresión en un medio mínimo, el cual proporciona los nutrientes mínimos necesarios para el crecimiento y desarrollo bacteriano, de esta manera se obliga a la célula a incorporar las únicas fuentes de nitrógeno y carbono presentes en el medio. Para éste fin se utilizan sales que contienen  $^{15}\text{N}$ , por ejemplo;  $^{15}\text{NH}_4\text{Cl}$  o  $(^{15}\text{NH}_4)_2\text{SO}_4$  y una fuente de carbono  $^{13}\text{C}$ , la cual generalmente es  $^{13}\text{C}$ -Glucosa y en algunos casos suele utilizarse  $^{13}\text{C}$ -Glicerol.

Generalmente las proteínas marcadas isotópicamente se obtienen haciendo crecer a la cepa de expresión en un medio mínimo que contiene una solución amortiguadora que prevenga la acidificación excesiva del medio de cultivo, metales necesarios para el metabolismo bacteriano tales como  $\text{Mg}^{2+}$  y  $\text{Ca}^{2+}$  y algunos metales de transición en pequeñas cantidades, así como las fuentes de nitrógeno y carbono ya mencionadas.<sup>(25)</sup> El medio mínimo PG ha sido específicamente diseñado para el óptimo aprovechamiento de reactivos enriquecidos que permitan el enriquecimiento isotópico de las proteínas expresadas.<sup>(25)</sup>

El crecimiento bacteriano en un medio mínimo generalmente es lento, lo cual puede resultar en una disminución en la expresión de la proteína dada por producción de sustancias que pueden ser citotóxicas para la célula. Por éste y otros motivos se ha diseñado una estrategia que ofrece diversas ventajas en la producción de proteínas marcadas isotópicamente.<sup>(26)</sup> Esta estrategia consiste en generar la biomasa celular para la expresión de la proteína en medio de cultivo rico (medio LB), posteriormente obtener el paquete celular y transferirlo a  $\frac{1}{4}$  de volumen de medio mínimo el cual contiene los reactivos enriquecidos y en el cual se hace la inducción para la producción de la proteína. Esta estrategia ofrece la obtención de altos rendimientos de proteína marcada isotópicamente así como una reducción en los tiempos de producción y en los costos.<sup>(26)</sup>

## 2.3 PURIFICACIÓN DE PROTEÍNAS

### 2.3.1 Cromatografía de afinidad metal-quelato<sup>(27)</sup>

Las proteínas recombinantes pueden ser adicionadas con una secuencia de 6 histidinas que permite la purificación mediante cromatografía de afinidad metal-quelato. Los grupos imidazólicos presentes en las histidinas forman complejos con iones metálicos tales como el níquel II y el cobalto II.

Los iones metálicos se fijan a la fase estacionaria de una columna cromatográfica por medio de grupos quelantes que a su vez están unidos covalentemente a perlas de algún carbohidrato insoluble. Al pasar una mezcla de proteínas a través de la columna cromatográfica, quedarán retenidas aquellas que puedan formar complejos con los iones presentes en la columna.

Algunas proteínas presentes en la naturaleza pueden quedar retenidas en la columna, al igual que las proteínas recombinantes que presentan una cola de histidinas. La elución de las proteínas que han sido retenidas en la columna se hace utilizando una disolución amortiguadora que permita modular la afinidad de las proteínas por el níquel, generalmente se utilizan disoluciones amortiguadoras con un agente quelante como el imidazol. Se pueden aplicar gradientes de concentración de imidazol en los que las proteínas con baja afinidad por el níquel eluyen en concentraciones bajas de imidazol y la proteína recombinante que presenta gran afinidad por el níquel eluye en una concentración mayor de imidazol. De esta manera, se realizan lavados para separar la proteína recombinante del resto de las proteínas que pudieron ser retenidas.

Éste tipo de cromatografía se sugiere como primer paso de purificación y puede ser utilizado para la purificación de proteínas en su forma nativa, así como en condiciones desnaturalizantes.

### **2.3.2 Cromatografía Líquida de Alta Presión (HPLC) <sup>(28)</sup>**

La cromatografía líquida de alta presión (HPLC) es una técnica de separación utilizada para determinar o purificar componentes de una mezcla compleja basándose en interacciones químicas entre la fase estacionaria de la columna cromatográfica y las sustancias analizadas. En éste tipo de cromatografía, la fase móvil es bombeada a alta presión. La utilización de presión en éste tipo de cromatografías incrementa la velocidad lineal de los compuestos dentro la columna y reduce así su difusión dentro de la columna, mejorando la resolución de la cromatografía.

La cromatografía en fase reversa (RP-HPLC) es una de las herramientas que se ha utilizado extensamente en separaciones de muestras bioquímicas. Éste tipo de cromatografía aprovecha las propiedades hidrofóbicas que generalmente están presentes en las proteínas. Explota el aspecto único en la estructura de péptidos y proteínas, de modo que tanto la

unión de éstos como la elución de la matriz son muy selectivas y están dadas principalmente por interacciones débiles de Van der Waals. El poder de resolución de esta técnica permite en la mayoría de los casos una separación total de moléculas casi idénticas permitiendo la separación de péptidos y proteínas que difieren únicamente en un aminoácido.

En cromatografía de fase reversa la fase estacionaria es no polar e hidrofóbica, la fase móvil inicial es polar, por ejemplo agua, con la cual eluyen de la columna las moléculas que casi no interactúan con la columna. Posteriormente, para eluir las moléculas que han interactuado con mayor fuerza, se utiliza una fase móvil menos polar; por ejemplo, una fase que incluya metanol, isopropanol o acetonitrilo. Los solutos eluyen en orden creciente de hidrofobicidad o en orden decreciente de la carga neta, grado de ionización o capacidad de formación de puentes de hidrógeno.

Generalmente para la elución de péptidos y proteínas se realiza una elución en gradiente y no una elución isocrática. Es común utilizar un gradiente de agua: acetonitrilo comenzando la purificación en una concentración baja de acetonitrilo. Generalmente se adiciona ácido trifluoroacético (TFA) a la fase móvil, esta adición produce un decremento del pH en la fase móvil ( $\text{pH} < 4$ ) con el cual se evita la interacción de residuos cargados en la proteína con grupos silanos libres presentes en la matriz cromatográfica, ya que a valores de pH menores a 4 los grupos silanos no se encuentran cargados. El TFA también interactúa con las proteínas acrecentando su carácter hidrofóbico y aumenta la interacción entre las proteínas y la fase estacionaria permitiendo una mejor separación de la mezcla de péptidos.

### **2.4 RESONANCIA MAGNÉTICA NUCLEAR DE PROTEÍNAS <sup>(24)</sup>**

La RMN es un fenómeno que se observa cuando ciertos núcleos atómicos en presencia de un campo magnético estable son expuestos a otro campo magnético variable. La espectroscopía de RMN utiliza éste fenómeno para obtener información del entorno químico que rodea a éstos núcleos atómicos y por lo tanto resulta muy útil para la determinación de estructuras de moléculas sencillas y complejas, así como para el estudio de interacciones y dinámica a nivel molecular.

Para experimentar el fenómeno de RMN un núcleo tiene que poseer espín. El espín es una propiedad fundamental de las partículas, así como la carga y la masa, y puede ser interpretado como un pequeño campo magnético. Las partículas subatómicas tienen espín así como ciertos núcleos atómicos tales como el  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ,  $^{19}\text{F}$  y  $^{31}\text{P}$ .

Al momento de introducir átomos con espín nuclear en un campo magnético externo fijo, el núcleo se comporta como un pequeño imán y tiende a orientarse, ya sea a favor (baja energía) o en contra (alta energía) de campo magnético externo, con una ligera preferencia a alinearse a favor. El núcleo puede pasar del estado de baja energía al de alta energía mediante la absorción de un fotón con la energía exacta entre los dos niveles; de la misma manera un núcleo puede pasar del nivel de alta energía al de baja energía mediante la emisión de un fotón. La diferencia de energía entre éstos dos niveles ( $\Delta E$ ) corresponde a radiación electromagnética de la región de las radiofrecuencias y está dada por la relación:

$$\Delta E = h \gamma B_0 / 2\pi$$

donde  $h$  es la constante de Planck,  $\gamma$  es la constante giromagnética que depende de cada núcleo y  $B_0$  es el campo magnético externo fijo. Siendo  $h$  y  $2\pi$  constantes universales, esta relación indica que la diferencia de energía entre los dos niveles depende únicamente del valor del  $B_0$  y del núcleo del que se trate, asimismo dado un mismo campo magnético, pero diferentes núcleos, la frecuencia de la radiación cambia. Ahora, si se aplica un pulso electromagnético al sistema y la energía del pulso es igual a  $\Delta E$  se dice que el sistema entra en resonancia.

En un campo magnético determinado se esperaría que todos los  $^1\text{H}$  absorban a la misma frecuencia de radiación; sin embargo, existen pequeñas variaciones (del orden de las centenas o millares de Hz) en las posiciones exactas de excitación para cada núcleo de la misma especie química. Estas variaciones son provocadas por los campos magnéticos locales y específicos generados por la densidad electrónica local que rodea a cada núcleo. A la posición en donde se observa la excitación de cada núcleo se le conoce como desplazamiento químico, el cual es empleado para identificar cada tipo de núcleo en una molécula y se mide en partes por millón.

En moléculas pequeñas, la determinación del desplazamiento químico puede ser suficiente para la determinación estructural de la molécula; sin embargo, a medida que se incrementa el número de núcleos, la sobreposición de señales aumenta. Esto es lo que ocurre en el caso

de las proteínas, en las que el número de núcleos es muy grande y las señales se traslapan impidiendo la interpretación de los espectros.

Además del desplazamiento químico, existen dos propiedades espectroscópicas fundamentales para la interpretación de los espectros, el acoplamiento escalar y el acoplamiento dipolar.

El acoplamiento escalar se representa como la influencia que ejerce un núcleo activo en RMN sobre otro activo a través de los enlaces. Al tener dos núcleos acoplados indirectamente, uno de ellos siente a través de los enlaces químicos un incremento o decremento del campo magnético debido al espín a favor o en contra del núcleo vecino. El acoplamiento escalar se observa generalmente de 1 a 3 enlaces de distancia y no se manifiesta entre núcleos equivalentes. Éste tipo de acoplamiento indica la presencia del número de núcleos vecinos. Éste fenómeno origina que las señales en los espectros se desdoblen y aparenten ser más complejas.

El acoplamiento que se presenta a través del espacio se le llama dipolar, esto es, sólo se requiere de su proximidad espacial y depende de la orientación con respecto al campo magnético. En disolución éste acoplamiento no se observa, dado que el promedio de los diferentes acoplamientos en moléculas que se encuentran en movimiento aleatorio con respecto al campo magnético se promedian obteniendo un valor de cero. Pese a ello, éste acoplamiento se manifiesta a través del efecto nuclear Overhauser (NOE por sus siglas en inglés)

Debido al traslape de las señales en los espectros monodimensionales de una proteína, se requiere obtener espectros de RMN en dos y en tres dimensiones para la interpretación de los espectros y la determinación de la estructura. Los experimentos de RMN en dos dimensiones se clasifican como homonucleares y heteronucleares. Los experimentos homonucleares muestran la correlación entre núcleos  $^1\text{H}$ . Ejemplo de éste tipo de experimentos son los experimentos COSY, TOCSY y NOESY. En cada uno de los experimentos se muestran diferentes tipos de correlación entre los núcleos.

Los experimentos heteronucleares muestran la correlación entre núcleos distintos. Frecuentemente los núcleos que se correlacionan en una muestra proteica son los  $^1\text{H}$  con los  $^{15}\text{N}$  y/o con los  $^{13}\text{C}$ . Cuando la correlación es a un enlace de distancia, los experimentos se llaman HSQC o HMQC, si son a dos o más enlaces de distancia el experimento se llama

HMBC. Éste tipo de espectros ayudan a identificar cada una de las señales de protón. Sin embargo, en la naturaleza los isótopos  $^{15}\text{N}$  y  $^{13}\text{C}$  no son abundantes por lo que se requiere la obtención de muestras de proteínas marcadas isotópicamente mediante el uso de biología molecular. La RMN ha sido empleada mediante la modalidad del experimento HSQC para corroborar la expresión, rendimiento, solubilidad y plegamiento en cuestión de horas, desde que se inserta el gen, hasta que se obtiene el experimento de RMN. <sup>(29)</sup>

### **2.5 NAMLA AMIDASAS.**

La pared celular bacteriana es una estructura presente en la mayoría de las bacterias. Es una capa semirrígida responsable de la configuración de la célula, se localiza en el exterior de la membrana plasmática, protegiendo a esta y al interior de la célula de los cambios adversos del medio externo. Su principal función es evitar la ruptura de la célula bacteriana. Químicamente la pared celular se compone principalmente de peptidoglicano, la disposición de este y la proporción presente en la pared celular permite establecer diferencias entre los principales tipos bacterianos. Clínicamente la pared celular reviste importancia porque contribuye a la virulencia de algunas especies bacterianas y es el sitio de acción de algunos antibióticos. <sup>(30)</sup>

La mayoría de las bacterias producen enzimas que hidrolizan el peptidoglicano bacteriano. Existen diferentes hidrolasas de peptidoglicano que químicamente se clasifican como glicosidasas y amidasas. <sup>(7)</sup>

Las hidrolasas de peptidoglicano actúan específicamente y tienen diversas funciones, por ejemplo, son útiles para el metabolismo bacteriano, regulan el crecimiento de la pared celular, regulan la separación de las células hijas durante la división celular, participan en la formación de los poros en los que se encuentran los flagelos y el pili, participan en la ruptura del peptidoglicano durante la formación y germinación de las esporas bacterianas. <sup>(3)</sup>

Las N-acetilmuramoyl-L-alanina (NAMLA) amidasas son proteínas modulares que degradan el peptidoglicano bacteriano actuando entre el extremo amino terminal del residuo de alanina y el ácido N-acetilmurámico <sup>(3)</sup> de la estructura del peptidoglicano (Figura 2.4), al ser hidrolasas de peptidoglicano participan en los procesos antes mencionados. Las NAMLA amidasas también se encuentran presentes en animales, donde

constituyen parte importante del sistema de defensa antibacterial <sup>(3)</sup> y también en algunos virus que atacan a las bacterias. <sup>(1)(3)(5)</sup>

### 2.5.1 Peptidoglicano

El peptidoglicano (PGN) es el componente mayoritario de la pared celular bacteriana. Tanto en bacterias gran positivas como gran negativas (Figura 2.4). Químicamente, el peptidoglicano está constituido por secuencias repetidas de N-acetilglucosamina (Glc Nac) y ácido  $\beta$ -(1-4)- N-acilmurámico (Mur Nac) entrecruzados por cadenas de aminoácidos atadas a los residuos ácido N- acilmurámico. El primero y segundo aminoácidos de la cadena suelen ser L-alanina y D-glutamina o isoglutamina, mientras que el último residuo suele ser D-alanina. En cocos gram positivos (ej. especies de *Staphylococcus* y *Streptococcus*) el tercer residuo de la cadena suele ser una lisina (PGN tipo Lys) y en bacterias gram negativas, así como en algunas especies gram positivas el tercer residuo es un *meso*-diaminopimelato (mDAP). (ej. Especies de *Bacillus* y *Listeria*). <sup>(1)</sup>

Sin embargo, la estructura del peptidoglicano puede modificarse de una especie a otra <sup>(1)(2)</sup> y puede verse modificada en respuesta a cambios en las condiciones del crecimiento bacteriano como puede ser la presencia de antibióticos en el medio. <sup>(1)</sup>

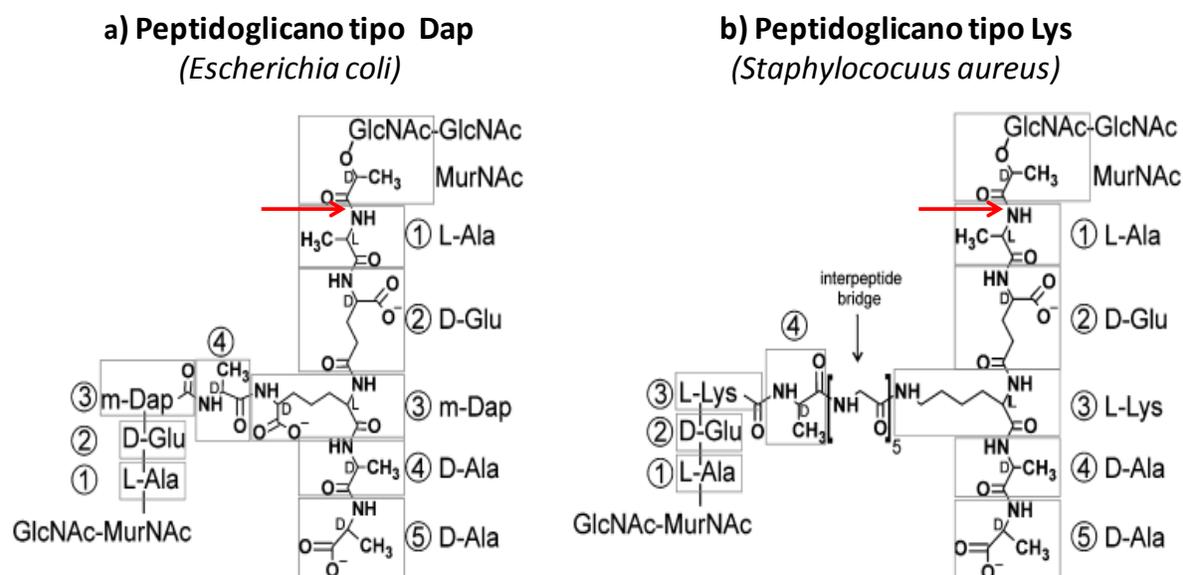


Figura 2.4. Estructura del peptidoglicano tipo Dap (a) y tipo Lys (b) <sup>(7)</sup>. La flecha roja muestra el sitio en el cual actúan las NAMLA amidasas.

### 2.5.2 Clasificación de las NAMLA amidasas <sup>(7)</sup>

Las amidasas de peptidoglicano se clasifican por la similitud de sus secuencias en familias y por criterios estructurales en grupos de plegamiento. En esta clasificación, las amidasas de un mismo grupo de plegamiento pertenecen al mismo tipo catalítico; pero, cada tipo catalítico puede comprender diferentes grupos de plegamiento. Las amidasas de peptidoglicano se clasifican en metalopeptidasas, cistein peptidasas y serin peptidasas.

Las NAMLA amidasas están clasificadas dentro de las metalopeptidasas. De acuerdo con la secuencia de aminoácidos del dominio enzimático se dividen en dos familias, familia Amidasa 2 (T7 lisozima) y familia Amidasa 3 (Carboxipeptidasa T).

La familia Amidasa 2 es la mejor estudiada de las dos porque se encuentra presente en todos los animales, además de bacterias y bacteriófagos; existen más de 20 estructuras resueltas en la base de datos PDB (Protein Data Bank). El dominio enzimático Amidasa 2 consiste en una lámina  $\beta$  formada de cinco hebras y rodeada de tres  $\alpha$ - hélices.

La familia Amidasa 3 sólo está presente en bacterias y por lo tanto ha despertado menos interés, sólo existen 5 estructuras (Códigos PDB 1JWQ, 1XOV, 3CZX, 3NE8 y 3QAY) y tres de ellas fueron resueltas automáticamente en iniciativas de genómica estructural. La topología del dominio enzimático Amidasa 3 consiste en una lámina  $\beta$  formada de seis hebras y rodeada de cinco o seis  $\alpha$ - hélices.

La Figura 2.5 muestra un ejemplo de la estructura secundaria del dominio enzimático tipo Amidasa 2 y un ejemplo del dominio enzimático tipo Amidasa 3.

En ambas familias el sitio activo incluye un catión  $Zn^{2+}$  coordinado por dos histidinas. Otros residuos menos conservados que también se encuentran comúnmente coordinando al  $Zn^{2+}$  incluyen aspartato, glutamato y en la familia Amidasa 3 también cisteína y lisina.

### 2.5.3 NAMLA amidasas como blancos para combatir enfermedades bacterianas

Actualmente, la resistencia a antibióticos es un problema que va en aumento. Una posible causa de la resistencia de las bacterias se debe al uso indiscriminado de los antibióticos, por lo tanto las bacterias pueden adquirir rápidamente resistencia a esas drogas. Por esta razón, en la actualidad se ha dado prioridad a otras áreas de investigación; por ejemplo, al desarrollo de fármacos cardiovasculares, los cuales se administran de por vida, en comparación con los antibióticos que se utilizan de manera transitoria.<sup>(4)</sup>

Actualmente, los antibióticos que existen en el mercado intervienen en cuatro procesos bacterianos esenciales, los cuales se encuentran bien identificados; estos procesos son: la síntesis de proteínas, la síntesis de ácidos nucleicos, la síntesis de la pared celular y la síntesis de los folatos. El proceso de división celular bacteriana es esencial; sin embargo, no ha sido blanco para los antibióticos probados clínicamente. <sup>(4)</sup>

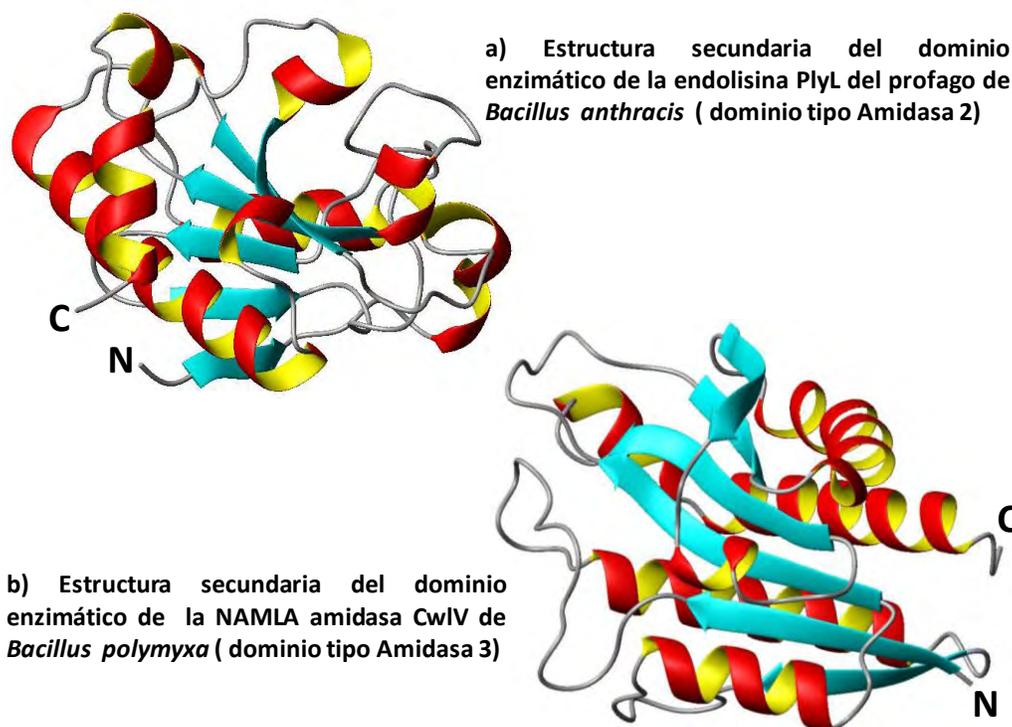


Figura 2.5. a) Estructura secundaria del dominio enzimático de la endolisina PlyL del profago de *Bacillus anthracis* con código PDB 1YB0 .Es un ejemplo de la estructura secundaria de un dominio enzimático de la familia Amidasa 2; se aprecia una lamina  $\beta$  formada por 5 hebras rodeada por  $\alpha$ - hélices. b) Estructura secundaria del dominio enzimático de la NAMLA amidasa CwIV de *Bacillus polymyxa* con código PDB 1JWQ. Es un ejemplo de la estructura secundaria de un dominio enzimático de la familia Amidasa 3 ,se aprecia una lamina  $\beta$  formada por 6 hebras rodeada por 6  $\alpha$  hélices. En ambas estructuras se muestran los extremos amino y carboxilo terminal.

Las hidrolasas de peptidoglicano intervienen en el proceso de división celular, en la ruptura del peptidoglicano para la separación de las células hijas durante la mitosis. Éstas, al igual que otras proteínas que participan en este proceso, son blancos terapéuticos para combatir enfermedades bacterianas. Particularmente, existe una NAMLA amidasa de *E.coli* llamada Ami C, la cual se sabe que es fundamental para la división del peptidoglicano durante el proceso de división celular<sup>(6)</sup>; adicionalmente se sabe que Ami C es una amidasa que se

encuentra presente en un gran número de bacterias, incluyendo bacterias patógenas de importancia médica<sup>(4)</sup>

Se piensa que el proceso de división celular puede inhibirse con el uso de inhibidores que actúen sobre las proteínas que participan en dicho proceso.<sup>(4)</sup>

Por otra parte, se ha propuesto el uso de hidrolasas que degradan la pared celular bacteriana como posibles antibacterianos para combatir infecciones causadas por bacterias patógenas que presentan resistencia a los antibióticos actuales.<sup>(31)</sup>

### **2.5.4 Dominios de reconocimiento de carbohidratos relacionados con la familia Amidasa 3.**

La familia Amidasa 3 en PFAM incluye miles de proteínas distribuidas en 173 arquitecturas, la gran mayoría de ellas compuestas por un dominio enzimático Amidasa 3 y uno o más dominios de reconocimiento de carbohidratos, peptidoglicano y/o de otros componentes de la pared celular. Entre los dominios de reconocimiento de carbohidratos y/o peptidoglicano mejor caracterizados se encuentran: SPOR, SLH, PG-binding-1, PG-binding-2, PG-binding-3 (miembros del clan PGBD), Lys M, CW-binding-1, CW-binding-2 y el dominio Amidasa 2-C.<sup>(32)</sup>

### **2.5.5 Dominios de reconocimiento de carbohidratos.<sup>(33)</sup>**

Muchas enzimas activas frente a carbohidratos son proteínas modulares que están compuestas de dominios catalíticos unidos a uno o más dominios no catalíticos de unión a carbohidratos (CBMs). Los CBMs se unen a una amplia gama de polisacáridos, tanto solubles como insolubles; su función principal es dirigir al dominio catalítico a su sustrato específico. Los CBMs se unen a sus ligandos con alta especificidad y afinidad, por lo que resultan ser excelentes modelos para estudiar los mecanismos de interacción proteína-carbohidrato.

Estudios estructurales demuestran que el plegamiento encontrado comúnmente para los CBMs es de  $\beta$ -sándwich. En dichas proteínas se encuentran iones metálicos que contribuyen a la estabilidad de la estructura.

Se ha propuesto una clasificación de tres categorías para los CBMs que se basan en la unión al sustrato: Tipo A o CBMs de unión a superficie, tipo B o CBMs de unión a cadenas de glicano y tipo C o de unión a azúcares pequeños.

Los CBMs tipo A incluye familias de proteínas que se unen altamente a la celulosa cristalina y a la quitina. El sitio de unión está compuesto por residuos aromáticos, muestra una arquitectura plana que se piensa es complementaria a la superficie de los cristales de celulosa y quitina. Este tipo de CBMs muestra poca o ninguna afinidad por carbohidratos solubles. En los CBMs tipo B los sitios de unión se describen como hendiduras y comprenden varios subsitios capaces de adaptarse a las unidades individuales de azúcar del ligando polimérico. Al igual que en el tipo A los ligandos son reconocidos por residuos aromáticos. Las cadenas laterales de tirosina, triptófano y fenilalanina pueden formar plataformas planas, espirales o sándwich para la unión del ligando. Finalmente, los CBMs tipo C tienen la propiedad, al igual que la lectina, de unirse óptimamente a mono, di ó trisacáridos y por lo tanto carecen de sitios de unión extendidos tipo hendidura como los observados en los CBMs tipo B.

### **2.5.6 Dominio Amidasa 2-C de una NAMLA amidasa presente en *Bacillus***

El dominio Amidasa 2-C es un dominio de reconocimiento de carbohidratos, tiene una longitud aproximada de 50 aminoácidos y se encuentra presente en bacterias y virus. Hasta el momento se conocen cuatro arquitecturas proteicas en las que se encuentra presente: en forma aislada, o en combinación con tres dominios de hidrolasas (glicosil-hidrolasa 25, Amidasa 2 y Amidasa 3).<sup>(9)</sup> Actualmente, el dominio Amidasa 2-C ha sido muy pobremente estudiado y su blanco de reconocimiento no es conocido, razón por la cual es importante y novedoso caracterizarlo.

El dominio Amidasa 2-C que se caracterizó en este trabajo pertenece a una NAMLA amidasa 3 proveniente de *B. cereus* (Uniprot Q736Q1\_BACC1).<sup>(8)</sup> Esta proteína está constituida por 244 residuos, con un dominio Amidasa 3 en el extremo N-terminal y un el dominio Amidasa 2-C en el extremo C-terminal (Figura 2.6).

MARYSLHGGHNSIVQGANFGNRKEHVLDRQVKDAVATK  
 LRSLGHTVYDDTDEVGTTQSQNLNNIIRNSNSHAVDLVIS  
 FHLNASDGNGQGVEVLYDQKDLAAKISAQLAKDIGWR  
 DRGAKQRDLAVLNGTKAPAILIELGFIDNESDMAKWNV  
 DKIAN SIVFALTGQTGGGATDLLKVKTTGGVAFSNLQALAQ  
 AMVDAGIDGQIVVQKDGIGYAITNGYPSGNIDKFTAWLD  
 ARKWYYEYLR

Figura 2.6. Estructura primaria de la N- acetilmuramoil L- alanina amidasa de *Bacillus cereus* (Q736Q1\_BACC1). La secuencia representada con letras rojas corresponde a la secuencia de aminoácidos del dominio enzimático Amidasa 3, la secuencia representada en letras azules corresponde a la secuencia de aminoácidos del dominio de reconocimiento de peptidoglicano Amidasa 2-C y la secuencia representada en letras negras es la secuencia de aminoácidos que une al dominio enzimático Amidasa 3 con el dominio de reconocimiento de peptidoglicano Amidasa 2-C.

### 2.5.7 Estudios bioinformáticos del dominio Amidasa 2-C de *Bacillus*

Desde décadas anteriores se han desarrollado una serie estrategias *in silico* que permiten hacer un modelo de la estructura terciaria de una estructura primaria extrayendo información de proteínas homólogas de estructura tridimensional conocida. Una de estas metodologías se conoce como enhebrado o “threading” y consiste primero en detectar una o más estructuras homólogas a la secuencia de interés utilizando información sobre similitudes en la secuencia de aminoácidos y predicciones de la estructura secundaria. Una vez seleccionada una o más proteínas homólogas, a las que se les denomina estructuras molde, el modelo tridimensional de la secuencia blanco se construye transfiriendo las coordenadas de las regiones equivalentes entre las dos proteínas y después construyendo las secciones disimilares por métodos de dinámica molecular.

Se obtuvo un modelo por homología del dominio Amidasa 2-C. El modelo se hizo en el servidor I- TASSER ONLINE Protein Structure & Function Predictions<sup>(34)</sup>. El algoritmo utilizado por este servidor tomó 10 proteínas homólogas como moldes para estimar la estructura tridimensional.

En la Figura 2.7 se muestra el modelo por homología del dominio Amidasa 2-C obtenido por el servidor. La estructura molde detectada por el servidor tiene el número de acceso en PDB 2L48, pertenece al dominio de reconocimiento de carbohidratos de una NAMLA

Amidasa 2 de bacteriófago conocida como PlyG (Uniprot Q8LTE6), desafortunadamente no existe ningún artículo asociado con esta estructura. El porcentaje de identidad entre el dominio Amidasa 2-C y la estructura 2L48 es de tan solo 22%, por lo que es de esperarse de que la estructura tridimensional experimental del dominio Amidasa 2-C de *Bacillus* difiera en algunos aspectos.

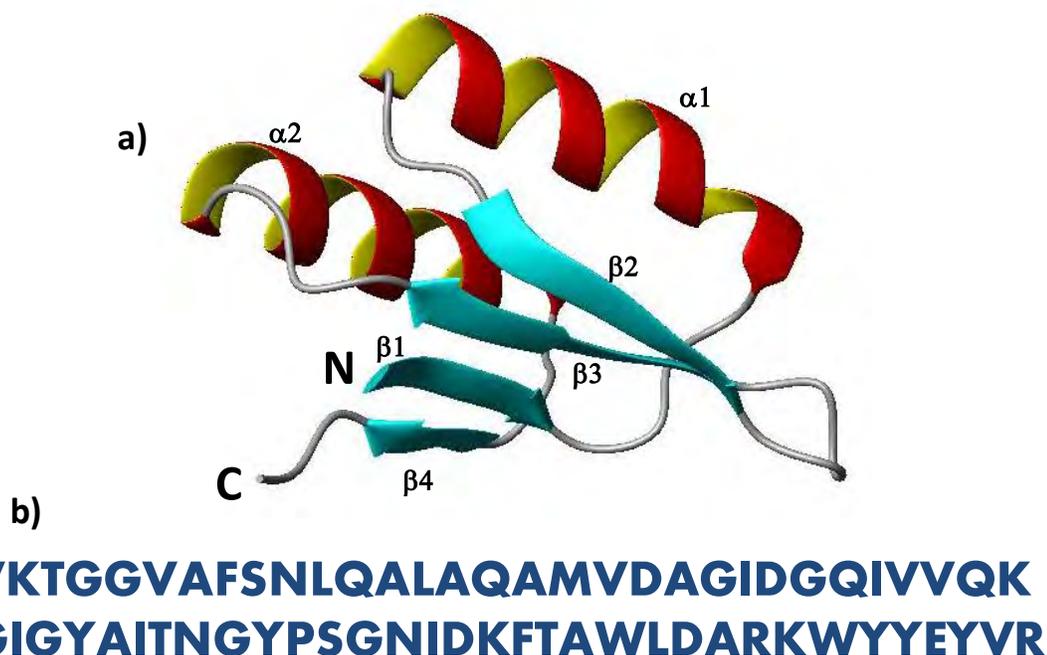


Figura 2.7. a) Estructura secundaria teórica del dominio Amidasa 2-C obtenida en el I- TASSER ONLINE Protein Structure & Function Predictions; b) Secuencia primaria del dominio Amidasa 2-C.

Se muestra una secuencia de elementos de estructura secundaria (hebra  $\beta 1$  - hélice  $\alpha 1$  - hebra  $\beta 2$  - hebra  $\beta 3$  - hélice  $\alpha 2$  - hebra  $\beta 4$ , en donde todas las hebras forman una lámina beta antiparalela  $\beta 2\beta 3\beta 1\beta 4$ . En el modelo de la estructura secundaria se muestran los extremos amino y carboxilo terminal (N y C respectivamente)

La secuencia de los elementos de estructura secundaria que presentan los dominios Amidasa 2-C (hebra  $\beta 1$  - hélice  $\alpha 1$  - hebra  $\beta 2$  - hebra  $\beta 3$  - hélice  $\alpha 2$  - hebra  $\beta 4$ , en donde todas las hebras forman una lámina beta antiparalela  $\beta 2\beta 3\beta 1\beta 4$ ) es semejante a la presentada por otro dominio de reconocimiento de carbohidratos conocido como SPOR (Mark A Williams, Birkbeck, University of London, comunicado personal), el cual forma parte de proteínas implicadas en la esporulación y la división celular<sup>(35)</sup>. En la Figura 2.8 se muestran las dos estructuras conocidas de dominios tipo SPOR (con códigos PDB 1UTA y

1X60) y el modelo obtenido para el dominio Amidasa 2-C. En esta se observa la similitud estructural que pudiera existir entre estos dominios.

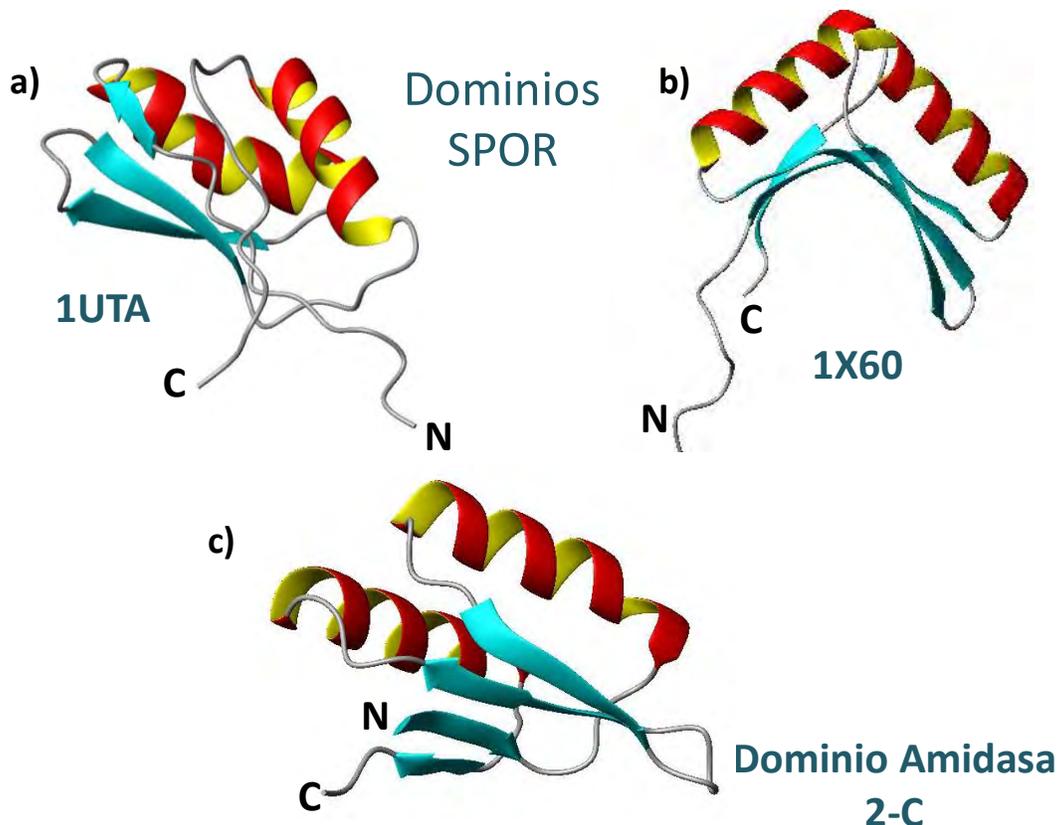


Figura 2.8. a) Estructura tridimensional con código PDB 1UTA, b) Estructura tridimensional con código PDB 1X60. Ambas estructuras se encuentran clasificadas dentro del dominio tipo SPOR. c) Estructura tridimensional teórica del dominio Amidasa 2-C. Se aprecia una secuencia de elementos en la estructura secundaria (hebra  $\beta$ - hélice  $\alpha$  - hebra  $\beta$ - hebra  $\beta$ - hélice  $\alpha$  - hebra  $\beta$ ) muy similar en las tres proteínas.

El alineamiento de la secuencia de aminoácidos de las proteínas tipo SPOR (Acely Garza García, MRC National Institute for Medical Research, Reino Unido, comunicado personal) con la secuencia de aminoácidos del dominio Amidasa 2-C (Figura 2.9) muestra que la cantidad de aminoácidos iguales entre las secuencias es baja; sin embargo, existe una conservación razonable de las características químicas de la posición en la secuencia, lo que sugiere una similitud entre estas proteínas. Existe entonces la posibilidad de que el universo de dominios de reconocimiento de carbohidratos sea menos diverso de lo que sugiere una simple comparación de secuencias de aminoácidos y que las relaciones que existen entre las diferentes familias sean sólo evidentes al nivel de estructura terciaria. Sea este el caso o no

sólo resultará evidente a medida de que más estructuras tridimensionales de reconocimiento de carbohidratos sean resueltas.



Figura 2.9 Alineamiento de la secuencia de aminoácidos de las proteínas tipo SPOR con la secuencia de aminoácidos del dominio Amidasa 2-C. Los aminoácidos polares se encuentran en color verde; los hidrofóbicos en azul cielo; los ácidos en color morado; los básicos en color rojo; la histidina y tirosina se muestran en color cian; la glicina en color naranja y la prolina en color amarillo. Se muestra una conservación de las características químicas de la posición en la secuencia.

---

# Metodología

### 3. METODOLOGÍA

#### 3.1 EXPRESIÓN DE LA PROTEÍNA DE FUSIÓN GB1-AMIDASA 2-C EN *E. coli*

##### *Rosetta*

Para poder obtener el dominio Amidasa 2-C recombinante se cuenta con el gen que codifica para dicha proteína, el gen está inserto en un plásmido modificado a partir del plásmido comercial pET24d (Novagen). Este derivado fue generado en el Laboratorio Europeo de Biología Molecular en Heidelberg, Alemania por Gunter Stier. La construcción del plásmido consiste en el gen fusionado a una serie de secuencias auxiliares que facilitan la expresión y la purificación de la proteína. Una vez aprovechadas las propiedades de las secuencias auxiliares deben retirarse por lo que la proteína de fusión incluye un sitio de reconocimiento para una proteasa específica que corta a la proteína de fusión liberando a la proteína objetivo de las secuencias auxiliares.

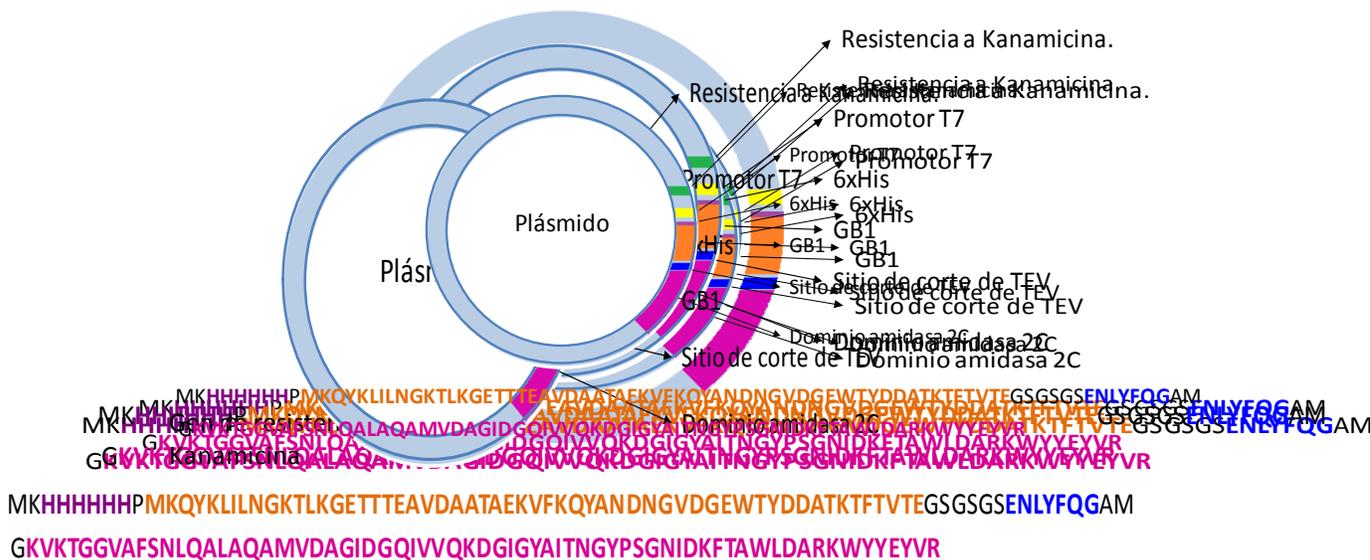


Figura 3.1 Se muestra la construcción del plásmido utilizado para la expresión de la proteína de fusión, así como la secuencia de aminoácidos. El plásmido cuenta con una secuencia de DNA que le confiere a la cepa de expresión resistencia a kanamicina, una secuencia codificante para el promotor T7 y la secuencia que codifica para la proteína de fusión.

La proteína de fusión consta de la secuencia de aminoácidos mostrada en la Figura 3.1. El color morado representa la secuencia de 6 histidinas consecutivas (6xHis) que facilitan la purificación de la proteína de fusión por la columna de afinidad metal-quelato, en naranja se

muestra la secuencia de aminoácidos de la proteína GB1, la cual favorece la solubilidad y estabilidad del dominio Amidasa 2-C, en color azul se muestra la secuencia de aminoácidos que son reconocidos como sitio de corte para la proteasa TEV y la secuencia de aminoácidos del dominio Amidasa 2-C está representada en color rosa.

Existe una amplia variedad de cepas aptas para la expresión de proteínas utilizando el sistema T7. En este caso se probó la expresión de la proteína de fusión en *E. coli Rosetta* (DE3) y se observó un buen nivel de expresión.

Para lograr la expresión de la proteína de fusión por la bacteria es necesario introducir el plásmido mostrado en la Figura 3.1 en células competentes de *E. coli Rosetta* (DE3) mediante el proceso de transformación bacteriana (APÉNDICE A). Dicha cepa de expresión presenta resistencia al cloranfenicol, por lo tanto, después del proceso de transformación se seleccionan las colonias bacterianas que se desarrollan en presencia de kanamicina y cloranfenicol ya que dichas colonias son las que han adquirido el plásmido que codifica para la proteína de fusión.

La expresión de las células transformadas es inducida mediante la adición de isopropil- $\beta$ -D-1-tiogalactopiranosido (IPTG) al medio de cultivo. Esta inducción es la señal para la producción paulatina de la proteína de fusión, llegando a un máximo en un tiempo desconocido. Para saber el tiempo en el cual la producción de proteína es máxima, se fija la temperatura de crecimiento e inducción y se realiza una cinética de expresión por 24 horas tomando muestras a tiempos determinados y analizando éstas mediante electroforesis en geles de poliacrilamida en presencia de dodecilsulfato de sodio (SDS-PAGE).

### **3.1.1 Expresión en medio Lysogeny Broth (LB)**

El medio LB (APÉNDICE B) contiene aminoácidos como fuente de nitrógeno y de energía así como otras moléculas complejas tales como nucleósidos y vitaminas, por lo tanto es considerado un medio rico en donde las bacterias no tienen que sintetizar *de novo* todas las biomoléculas necesarias para su metabolismo. Esto se refleja en mayores velocidades de crecimiento y generalmente mejor expresión de proteínas recombinantes. En este medio se realiza la expresión de la proteína sin enriquecimiento isotópico para hacer la caracterización de la proteína.

Se hace crecer a las células las células hasta que alcanzan un valor de densidad óptica igual a 0.7 a 600 nm, en este momento se hace la inducción añadiendo IPTG y se incuba a la temperatura óptima de crecimiento bacteriano 37 °C, la cual puede o no coincidir con la

temperatura óptima de expresión

### 3.1.2 Expresión en medio mínimo PG

En este medio de cultivo se realiza la expresión de la proteína doblemente marcada ( $^{15}\text{N}$ ,  $^{13}\text{C}$ -Amidasa 2-C) debido a que este medio de cultivo contiene los requerimientos mínimos nutricionales para el crecimiento y desarrollo celular, de esta manera se obliga a la célula a utilizar moléculas simples como  $\text{NH}_4^+$  y glucosa para sintetizar todas las biomoléculas necesarias para su metabolismo, y de esta manera incorporar los núcleos marcados en la proteína que se desea estudiar. En el medio PG están optimizadas las concentraciones de reactivos enriquecidos en el medio de cultivo para su óptimo aprovechamiento en el enriquecimiento isotópico.<sup>(25)</sup>

Se probó la expresión de la proteína de fusión utilizando dos metodologías diferentes (Expresión en medio PG y expresión en medio PG con reducción de volumen), esto con la finalidad de comparar las diferencias en la obtención de la proteína de fusión y así hacer la elección de la técnica más conveniente. Se realizó un monitoreo de expresión con ambas técnicas a 30 °C y 37 °C, los resultados se evaluaron mediante SDS-PAGE. Se aisló la proteína de fusión, se purificó por columna de afinidad metal-quelato y se cuantificó mediante la técnica de Bradford.

La metodología de expresión en medio PG consiste en preparar un cultivo celular crecido en medio mínimo PG por una noche previa a la inducción, esto para obtener una masa celular suficiente y adaptada el medio en el cual se hará la inducción, de modo que la densidad de inducción se alcance en un tiempo razonable. Posteriormente el cultivo de toda la noche se pone en un volumen mayor de medio PG hasta alcanzar una densidad óptica de 0.7 a 600 nm, se hace la inducción con IPTG y se incuba a la temperatura a la cual se probará la expresión.

La expresión en medio PG con reducción de volumen consiste en crecer a las células en medio LB hasta un valor de densidad óptica igual a 0.7 a 600 nm, cuando se alcanza este valor, se centrifuga el cultivo celular para quitar los desechos metabólicos que pudieran resultar tóxicos para las células y que acidifican el medio se obtiene el paquete celular y se lava con disolución amortiguadora de fosfatos del medio mínimo, posteriormente, se resuspende el paquete celular en  $\frac{1}{4}$  de volumen inicial con disolución amortiguadora de medio mínimo sin fuentes de carbono y nitrógeno y se deja incubando por 1 hora para que las células se adapten al medio, transcurrido este tiempo se centrifugan nuevamente, se lavan con disolución amortiguadora del

medio mínimo y se resuspenden nuevamente en ¼ de volumen inicial con medio mínimo, en este momento se hace la inducción con IPTG y se incuban a la temperatura a la cual se probará la expresión.<sup>(26)</sup>

El protocolo detallado de ambas metodologías se muestra en el APÉNDICE C.

## **3.2 PURIFICACIÓN DE LA PROTEÍNA RECOMBINANTE**

Ya que la cepa bacteriana ha producido la proteína de fusión en las condiciones óptimas para la expresión, se debe extraer de la célula y separar de otras proteínas y componentes celulares. Para esto se hace una lisis celular mediante sonicación para liberar a la proteína de la célula. Si la proteína de fusión se encuentra soluble, se trabaja con la parte soluble del lisado bacteriano y se comienza con los pasos de purificación.

### **3.2.1 Cromatografía de afinidad metal quelato de la proteína de fusión.**

La proteína de fusión soluble puede separarse de otras proteínas cuando se hace pasar la parte soluble del lisado celular a través de una columna de afinidad metal quelato. La columna se encuentra cargada con iones de  $\text{Ni}^{2+}$ , las seis histidinas que forman parte de la proteína de fusión se coordinan con el níquel y la proteína es retenida en la columna. La retención de proteína de fusión puede modificarse utilizando disoluciones con diferentes concentraciones de imidazol. Puede aplicarse un gradiente de imidazol, en donde las disoluciones poco concentradas sirven para eluir proteínas con poca o mediana afinidad por el níquel y una disolución más concentrada de imidazol permite eluir a la proteína de fusión.

### **3.2.2 Corte de la proteína de fusión<sup>(36)</sup>**

Se debe separar el dominio Amidasa 2-C de las secuencias auxiliares. La proteína de fusión cuenta con una secuencia de aminoácidos que son reconocidos como sitio de corte de la proteasa TEV. La secuencia general de reconocimiento para esta enzima es EXXYXQ(G/S), donde X es cualquier aminoácido e hidroliza en el enlace peptídico entre Q (glutamina) y G (glicina) o entre Q y S (serina). Se utiliza comúnmente la secuencia ENLYFQG.

La proteasa TEV es recombinante y cuenta también con una secuencia de 6 histidinas, de modo que una vez realizado el corte se puede utilizar otra vez una columna de afinidad metal quelato para separar a la proteína que ha sido exitosamente cortada, la cual ya no posee la etiqueta de polihistidina de la proteasa, la proteína sin cortar y la etiqueta His<sub>6</sub>GB1 resultante del corte.

La proteasa TEV es activa entre valores de pH de 4 a 9, su máxima actividad la alcanza a 34 °C aunque es recomendable hacer el corte a temperatura ambiente o a 4 °C. Se requiere de un medio reductor para que la enzima sea activa, comúnmente se utiliza ditioneitol (DTT) en una concentración de 1 mM, si la proteína blanco forma puentes disulfuro, el DTT debe ser remplazado por alguna disolución amortiguadora que permita mantener las condiciones reductoras, esto con la finalidad de preservar los puentes disulfuro.

### **3.2.3 Cromatografía Líquida de Alta Presión (HPLC)**

Después de realizar el corte de la proteína de fusión, se debe separar el dominio Amidasa 2-C de otras proteínas de la lisis celular que aún pudieran estar presentes y otros contaminantes. Esto se hace con HPLC utilizando columnas C12 o C18 de fase reversa. Para lograr la separación y purificación de los residuos de corte, se utiliza un gradiente de acetonitrilo y agua con 0.05 % de ácido trifluoroacético (TFA). Este es el paso final de la purificación. La proteína de interés se encuentra con cierta cantidad de acetonitrilo y TFA que deben ser retirados para hacer la caracterización de la proteína y obtener los espectros de Resonancia Magnética Nuclear, esto se logra liofilizando la muestra. La liofilización de la muestra permite preservarla por un tiempo prolongado.

## **3.3 CARACTERIZACIÓN DEL DOMINIO AMIDASA 2-C**

Después de purificar una proteína es importante obtener datos que corroboren que se ha obtenido la proteína correcta.

### **3.3.1 Espectrometría de masas**

El peso molecular es una característica altamente específica de una molécula dada y con frecuencia es la primera propiedad físico-química que se determina para la caracterización de una proteína. Tradicionalmente, la electroforesis de gel de poliacrilamida, la ultracentrifugación analítica o la cromatografía de permeación en gel han sido las técnicas utilizadas para realizar la determinación del peso molecular. Recientemente, el desarrollo de las técnicas de espectrometría de masas, tanto en la preparación de la muestra como en la instrumentación, ha permitido realizar la determinación de la masa de proteínas intactas con una precisión muy superior a los métodos mencionados anteriormente. De esta manera, la espectrometría de masas (MS) es utilizada para analizar la estructura primaria de las proteínas en las últimas etapas de la

purificación, así como también antes de realizar estudios para la determinación de la estructura, tales como el dicroísmo circular, la cristalografía de rayos X y la resonancia magnética nuclear. Un espectrómetro de masa genera iones en fase gaseosa a partir de una muestra, separa los iones de acuerdo con su relación masa carga y registra su abundancia relativa para generar un espectro que será característico de las condiciones de análisis empleadas.

### **3.3.2 MALDI- TOF <sup>(37)</sup> <sup>(38)</sup>**

Una de las técnicas más relevantes para la aplicación en trabajos de biomoléculas es la ionización por desorción láser asistida por una matriz (MALDI, matrix-assisted laser desorption ionization) combinada con el análisis de espectrometría de masas (MS) empleando un detector de tiempo de vuelo (TOF, time of flight). La técnica de MALDI-TOF tiene la ventaja de ser muy sensible, tal que permite utilizar concentraciones de análisis en las escalas femto y picomolar. Otras técnicas para análisis de MS son la ionización por electrospray y el bombardeo rápido de átomos (FAB, fast atom bombardment) empleando detección con cuadrupolo.

Para el análisis por MS es necesario generar iones intactos en fase gaseosa que serán generados por la fuente de iones del espectrómetro. El paso crítico para el análisis por MALDI es la preparación de la muestra. El analito es mezclado con un compuesto que es empleado como matriz capaz de absorber la radiación UV, para formar microcristales de proteína embebidos en la matriz. La sublimación de los microcristales producida por el efecto combinado del alto vacío y la irradiación láser, conduce a la formación de la fase gaseosa de iones moleculares protonados. Los iones son entonces acelerados por un campo electrostático de modo que adquieran la misma energía cinética, en estas condiciones los iones ligeros vuelan más rápido que los pesados y llegan antes al detector, de modo que el tiempo de vuelo es directamente proporcional a la relación de la masa dividida entre la carga.

### **3.4 ANÁLISIS POR RESONANCIA MAGNÉTICA NUCLEAR (RMN)**

La RMN es una técnica que permite determinar la estructura secundaria, terciaria y hasta cuaternaria de las proteínas y proporciona información acerca de procesos dinámicos que pudiera estar experimentando la proteína. Esto es posible porque la frecuencia de resonancia de cada núcleo es extremadamente sensible tanto a la magnitud como a la dirección de la densidad electrónica que lo rodea, la cual está dada por la naturaleza y la distancia de los átomos que se

encuentran cercanos a cada núcleo.

En una proteína existe una cantidad significativa de  $^1\text{H}$  activos en RMN, lo cual origina la sobreposición de las señales y el traslape de las mismas impidiendo de esta manera la interpretación de los espectros monodimensionales. Es por esta razón que la determinación de la estructura de una proteína requiere la obtención de espectros en 2 y 3 dimensiones. Con estos experimentos es posible determinar qué núcleos dentro de una proteína se encuentran acoplados entre sí.

Para esta proteína se obtiene un espectro de protón monodimensional, un espectro en dos dimensiones homonuclear (TOCSY) que permite encontrar las correlaciones de todos los núcleos del sistema de espín y experimentos en dos dimensiones heteronucleares ( $^{15}\text{N}$ - HSQC y  $^{13}\text{C}$ - HSQC) que facilitan la determinación de la estructura en proteínas mayores a 50 aminoácidos. <sup>(24)</sup>

---

# Resultados y Discusión

### 4. RESULTADOS Y DISCUSIÓN

#### 4.1 EXPRESIÓN DE LA PROTEÍNA DE FUSIÓN GB1- AMIDASA 2-C

La inducción de la expresión de la proteína en medio LB, así como de la proteína expresada en medio mínimo por ambas metodologías mencionadas anteriormente se hizo con una concentración de IPTG de 0.5 mM a una velocidad de agitación de 200 rpm.

##### 4.1.2 Monitoreo de las expresiones

Para monitorear la expresión de la proteína de fusión expresada en el medio LB y en el medio mínimo, se hace la expresión en volúmenes pequeños de medio de cultivo correspondiente (20-50 mL). El monitoreo se hace tomando una alícuota de 1 mL del cultivo en tiempos determinados considerando el tiempo cero de inducción y tiempos posteriores a ésta. Las muestras tomadas en ambos casos se colocaron en tubos de microfuga y fueron centrifugadas (13,000 x g por 3 min.) para obtener el paquete celular y desechar el sobrenadante. Las muestras fueron congeladas a  $-20\text{ }^{\circ}\text{C}$  para su posterior tratamiento.

Para poder ver la expresión por SDS-PAGE, las muestras fueron resuspendidas en 50  $\mu\text{L}$  de disolución amortiguadora de carga 2X para SDS-PAGE y 50  $\mu\text{L}$  de disolución amortiguadora A (APÉNDICE D), estas se calentaron a ebullición por 5 min en un baño de agua y posteriormente se tomó una alícuota de 8  $\mu\text{L}$  de cada muestra para cargar en el gel.

##### 4.1.3 Condiciones para la expresión en medio Lysogeny Broth (LB)

La expresión de la proteína de fusión en medio LB se realizó a una temperatura de  $37\text{ }^{\circ}\text{C}$ . Los resultados del monitoreo de expresión muestran (Figura 4.1) que en 6 horas posteriores a la inducción se obtiene el máximo de expresión de la proteína.

Considerando los resultados del monitoreo de expresión, se decide hacer la expresión de la proteína de fusión en las condiciones establecidas ya que los resultados muestran que a esta temperatura se obtiene un buen nivel de expresión en un tiempo corto.

Se hace la inducción en 1.5 L de medio LB, se incuba durante 6 horas y se obtiene el paquete celular centrifugando a 5500 x g. Posteriormente se realiza la lisis celular y los

pasos de purificación.

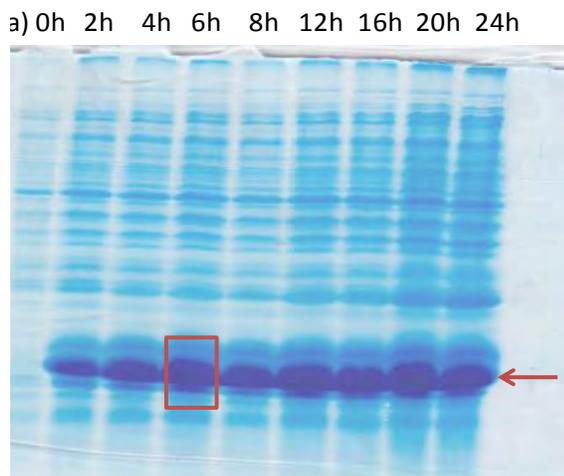


Figura 4.1. Gel de SDS-PAGE del monitoreo de expresión en medio LB. Los resultados muestran la expresión de la proteína de fusión en muestras a diferentes tiempos posteriores a la inducción. La flecha roja indica la banda que corresponde a la proteína de fusión en el gel de poliacrilamida. Se aprecia la proteína de fusión en el peso esperado (~16.3 kDa) observándose el máximo de expresión en 6 horas posteriores a la inducción (recuadro rojo)

#### 4.1.4 Condiciones para la expresión en medio mínimo PG

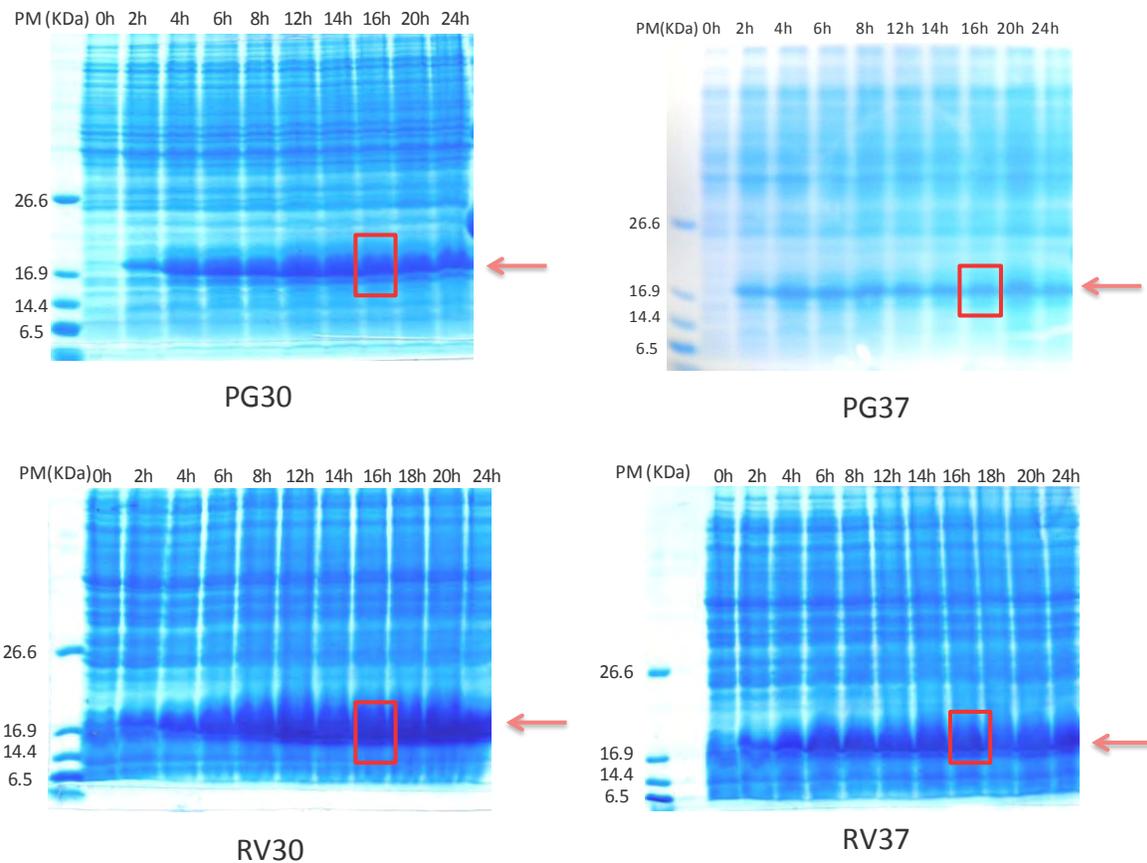
La finalidad de expresar la proteína en un medio mínimo es obtener la proteína marcada isotópicamente. Se decide hacer la expresión por el método tradicional y probando la técnica descrita de reducción de volumen ya que se ha mencionado que ofrece la obtención de buenos rendimientos así como una disminución en los costos.<sup>(26)</sup>

Se realiza el monitoreo de expresión a 30 y 37 °C utilizando ambas técnicas descritas en la metodología. El monitoreo de expresión se hace sin enriquecimiento isotópico. Los resultados se muestran en geles de SDS- PAGE (Figura 4.2) observándose que en todos los casos se mantiene el máximo de expresión a las 16 horas posteriores a la inducción. Los resultados mostrados por SDS-PAGE son cualitativos debido a que la concentración bacteriana de las muestras tomadas en el monitoreo es mayor cuando se hace la reducción de volumen.

Se decidió hacer una prueba de expresión para determinar cuantitativamente el método que proporciona mayores beneficios para la expresión de proteína enriquecida. Para esta prueba se realiza la expresión por ambas técnicas durante 16 horas ya que el máximo de expresión de la proteína se mantiene durante este tiempo. La expresión se realiza en las dos

## Resultados y Discusión

temperaturas antes mencionadas. Posteriormente se aísla la proteína de fusión, se cuantifica y se comparan los resultados. Se probó la expresión en medio PG haciendo la inducción en 125 mL (matraz preparativo) del medio PG utilizando las proporciones equivalentes de nutrimentos con respecto al protocolo descrito en el APÉNDICE C. Para la técnica de reducción de volumen en el medio PG el crecimiento de las células se inició en un volumen de 100 mL de medio LB y haciendo la reducción de volumen a 25 mL de medio PG. Después de 16 horas tras la inducción en las dos temperaturas (30 y 37 °C) se obtiene el paquete celular centrifugando a 5500 x g durante 20 min.



**Figura 4.2. Geles de SDS-PAGE del monitoreo de expresión en medio PG. 1) PG30: Monitoreo de expresión en medio PG a una temperatura de 30 °C; 2) PG37: Monitoreo de expresión en medio PG a una temperatura de 37 °C; 3) RV30: Monitoreo de expresión en medio PG con reducción de volumen a una temperatura de 30 °C; 4) RV37: Monitoreo de expresión en medio PG con reducción de volumen a una temperatura de 37 °C. Las flechas rojas señalan la banda correspondiente a la proteína de fusión GB1-Amidasa 2-C (~16.3 KDa). Los recuadros rojos muestran que a las 16 horas posteriores a la inducción se mantiene el máximo de expresión de la proteína de fusión. En cada uno de los geles se muestran las bandas correspondientes al marcador de pesos moleculares y los tiempos posteriores a la inducción en los que se monitorea la expresión de proteína.**

## Resultados y Discusión

Se pesa el paquete celular y se hace la lisis de la forma que se indica en la sección 4.1.5; se realiza la purificación por columna de níquel siguiendo los pasos indicados en el APÉNDICE E, se realiza un cambio de disolución amortiguadora de la Fracción E1 (sección 4.2.1) obtenida de la purificación de la columna de níquel a disolución amortiguadora D (APÉNDICE D) y se cuantifica la cantidad de proteína mediante el método de Bradford (APÉNDICE F). Se compara la cantidad de proteína obtenida así como el costo que tiene la producción en cada caso.

Los resultados obtenidos se muestran en la Tabla 4.1.

Muestra	Vol. de inducción (mL)	Cantidad utilizada de <sup>13</sup> C-Glucosa (g)	Cantidad utilizada de <sup>15</sup> N-(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> (g)	Cantidad obtenida de proteína de fusión (mg)	Cantidad gastada en <sup>13</sup> C-Glucosa (DEUA)	Cantidad gastada en <sup>15</sup> N-(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> (DEUA)	Costo de la producción (DEUA/mg)
<b>PG30</b>	<b>125</b>	<b>0.625</b>	<b>0.438</b>	<b>8.98</b>	<b>121.25</b>	<b>21.02</b>	<b>15.84</b>
<b>PG37</b>	125	0.625	0.438	2.66	121.25	21.02	53.48
<b>RV30</b>	<b>25</b>	<b>0.125</b>	<b>0.088</b>	<b>5.38</b>	<b>24.25</b>	<b>4.20</b>	<b>5.29</b>
<b>RV37</b>	25	0.125	0.088	2.91	24.25	4.20	9.77

**Tabla 4.1** La tabla muestra la cantidad de proteína obtenida, así como la cantidad utilizada de reactivos enriquecidos y el costo de producción en cada caso. Se consideró el costo por gramo de <sup>13</sup>C-Glucosa (DEUA 194) y el costo de <sup>15</sup>N-(NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (DEUA 48).

DEUA= Dólares de Estados Unidos de América.

Los resultados de esta prueba muestran que la temperatura óptima de producción es de 30 °C. A esta temperatura de expresión se compara la cantidad obtenida de proteína mediante el método tradicional de expresión en medio mínimo y la cantidad obtenida al hacer la reducción de volumen. Se obtienen 3.6 mg menos de proteína al hacer la reducción de volumen. Esta diferencia no es relevante al considerar el costo de la producción por el método tradicional, el cual es tres veces mayor en comparación con el costo al hacer la reducción de volumen.

Los resultados de la Tabla 4.1 muestran que la técnica más conveniente para hacer la expresión de la proteína marcada isotópicamente es haciendo la reducción de volumen, ya que permite la obtención de buenos rendimientos, lo cual puede explicarse porque en el crecimiento bacteriano se generan sustancias que reducen el pH del medio y/o son citotóxicas o inhibitorias para el crecimiento bacteriano y la expresión de la proteína. (25; 26)

En la técnica de reducción de volumen estas sustancias son retiradas al realizar un cambio de medio a medio fresco (medio mínimo) en el cual se hace la inducción para la expresión de la proteína. Dicha técnica ofrece una reducción significativa en los costos de producción debida a una disminución en la cantidad de reactivos enriquecidos. Además acorta el tiempo en el que se genera la biomasa celular para inducir la expresión de la proteína ya que la biomasa es generada en medio LB. Esto es válido ya que se ha demostrado que no se requiere tener masa celular marcada con  $^{13}\text{C}$  y  $^{15}\text{N}$  para tener una buena incorporación de los isótopos después de hacer la inducción.<sup>(26)</sup>

Se decide hacer la expresión de la proteína marcada isotópicamente siguiendo la técnica de reducción de volumen a una temperatura de 30 °C. La masa celular se genera en 4 L de medio LB a 37 °C hasta alcanzar una OD a 600 nm de 0.7, posteriormente se obtuvo el paquete celular, se resuspendió en 1 L de disolución amortiguadora de medio mínimo sin fuentes de carbono y nitrógeno para permitir que las células se adapten al medio. Transcurrida 1 hora se centrifugó nuevamente para obtener el paquete celular, el cual se resuspendió en 1 L de medio mínimo PG que contenía los reactivos enriquecidos. Se hizo la inducción con isopropil- $\beta$ -D-1-tiogalactopiranosido (IPTG) y se dejó incubando por 12 horas a 30 °C. Posteriormente, se obtuvo el paquete celular por centrifugación y se congeló a -20 °C para su posterior tratamiento.

### 4.1.5 Lisis celular

De las expresiones realizadas tanto en medio LB, así como en medio mínimo, se obtuvo el paquete celular centrifugando a 5500 x g por 20 min. El paquete celular se pesa y se resuspende en la disolución amortiguadora A (APÉNDICE D), adicionando 20 mL de disolución amortiguadora por cada gramo del paquete celular. Se adicionó fluoruro de fenilmetilsulfonilo (PMSF) en una concentración de 1 mM como inhibidor de proteasas. La lisis celular se realiza sonicando el paquete celular resuspendido (20 golpes sónicos de 39 watts de potencia cada uno, con 1 min de incubación en hielo entre cada uno).

El lisado celular fue centrifugado a 35000 x g durante 40 min. Se separó el sobrenadante y se filtró a través de una membrana de 0.22  $\mu\text{M}$  para purificarla por una columna de afinidad metal - quelato.

## 4.2 PURIFICACIÓN DE LA PROTEÍNA DE FUSIÓN

### 4.2.1 Purificación por la columna de afinidad metal - quelato

La fracción soluble del lisado celular contiene la proteína de fusión que debe ser separada del resto de las proteínas presentes en la mezcla. La proteína de fusión contiene una secuencia de 6 histidinas que pueden coordinarse con el  $\text{Ni}^{2+}$ , el cual se encuentra fijo a la fase estacionaria de la columna especializada para la separación de mezclas de proteínas Hi-Trap Chelating®. Las proteínas que pueden formar enlaces de coordinación con el ión metálico se quedan retenidas al hacer pasar la mezcla de proteínas por la columna. Para eluir posteriormente las proteínas retenidas en la columna se realizaron lavados aplicando un gradiente de imidazol, el primer lavado se hace con la disolución amortiguadora A para arrastrar proteínas que probablemente se encuentran retenidas en los espacios de la resina y proteínas de baja afinidad por el níquel. El segundo lavado se hizo utilizando la disolución amortiguadora B para eluir de la columna las proteínas con mediana afinidad por el níquel. Finalmente la proteína de fusión que presenta gran afinidad por el níquel es eluída utilizando la disolución amortiguadora C la cual tiene una concentración de imidazol de 300 mM (APÉNDICE D y E). Se toman muestras de los pasos de purificación y se analizan las diferentes fracciones por SDS-PAGE (Figura 4.3)

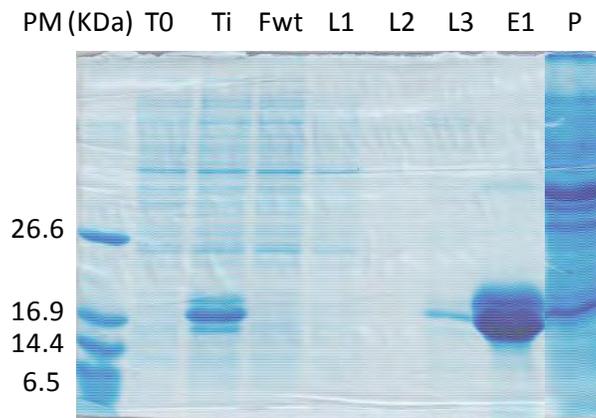


Figura 4.3. Gel de SDS-PAGE que muestra los resultados del proceso de purificación de la proteína de fusión al pasar la fracción soluble del lisado celular a través de una columna de afinidad metal-quelato. T0: Tiempo cero de inducción; Ti: Fracción soluble del lisado celular después del tiempo de inducción y antes de hacer pasar la muestra por la columna; Fwt: Fracción no retenida por el níquel después de pasar la muestra a través de la columna; L1: Primer lavado con disolución A; L2: Segundo lavado con disolución A; L3: Lavado con disolución amortiguadora B; E1: Elución con disolución amortiguadora C de la proteína de fusión retenida por el níquel ; P: Muestra del paquete celular resultante después de la lisis

### 4.2.2 Corte de la proteína de fusión

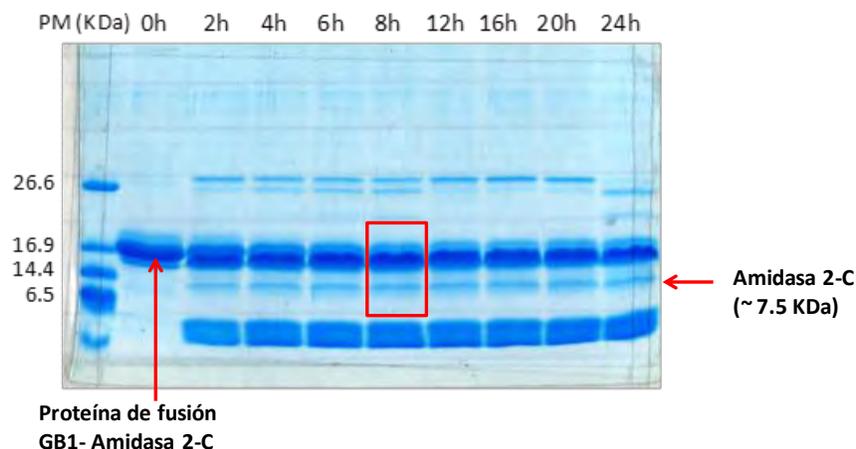
Después de la purificación de la proteína por la columna de afinidad metal-quelato es necesario realizar el corte de la proteína de fusión con la proteasa TEV para separar el dominio Amidasa 2-C de las secuencias auxiliares.

La fracción E1 eluída con la disolución amortiguadora C contiene la proteína de fusión. El gel de electroforesis (Figura 4.3) muestra que aproximadamente más del 90 % de la proteína presente en ésta fracción corresponde a la proteína de fusión por lo que se decide hacer el corte con la proteasa TEV.

El imidazol puede inhibir el corte por TEV, por lo tanto, éste debe realizarse en una disolución que no lo contenga. Para realizar el corte de la proteína de fusión se hizo un cambio de disolución amortiguadora. La proteína de fusión se cambia de la disolución amortiguadora C a la disolución D (APÉNDICE D). Este cambio se hace mediante diálisis utilizando una membrana de tamaño de poro de 8000 Da (el peso de la proteína de fusión es de 16.3 kDa).

Se realizaron pruebas para determinar las condiciones de corte, éstas se hicieron con proteína de fusión expresada en medio LB sin enriquecimiento isotópico. En dichas pruebas se observó que el corte no era eficiente ya que la proteína de fusión no se cortaba al 100 %. Sin embargo, las mejores condiciones encontradas para el corte fueron las siguientes: 1) Mantener la concentración de la proteína de fusión en 1.3 mg/mL; 2) Adicionar 1 mol de TEV por cada 50 moles de proteína de fusión; 3) Tener presente una concentración de 1 mM de DTT en la reacción de corte y 4) Cortar a una temperatura de 30 °C. Dadas estas condiciones se realizó una cinética de corte en la que se encontró que a las 8 horas se obtiene el máximo corte de la proteína de fusión (Figura 4.4).

Después de realizar el corte es necesario separar el dominio Amidasa 2-C de los fragmentos resultantes. Para poder realizar el siguiente paso de purificación es necesario retirar el DTT que puede originar reducción en la fase estacionaria de la columna de níquel, lo cual se hizo mediante diálisis utilizando una membrana de tamaño de poro de 3000 Da, al mismo tiempo se aumentó la concentración de NaCl a 500 mM (disolución amortiguadora E) para favorecer que el dominio Amidasa 2-C no quedara retenido en la columna mediante interacciones iónicas.



**Figura 4.4.** Gel de SDS-PAGE de la Cinética de corte de la proteína de fusión en las siguientes condiciones: 1) Concentración de la proteína de fusión en 1.3 mg/mL; 2) 1 mol de TEV por cada 50 moles de proteína de fusión; 3) Concentración de 1 mM de DTT en la reacción de corte y 4) Temperatura de 30 °C. A las cero hora de corte se muestra la proteína de fusión. El gel muestra la banda correspondiente a la proteína de fusión (~16.3 KDa) antes del corte con la proteasa TEV y la banda correspondiente al dominio Amidasa 2-C obtenido después del corte. El recuadro rojo muestra que en 8 horas posteriores a la adición de la proteasa TEV se obtiene el máximo corte de la proteína de fusión.

Después del corte el dominio Amidasa 2-C, que ya no cuenta con la secuencia de 6 histidinas a diferencia del fragmento restante de la proteína de fusión; se hace pasar la reacción de corte a través de la columna de afinidad metal-quelato Hi-Trap Chelating® en la que quedarán retenidos tanto la proteasa TEV como los fragmentos de corte que contienen la cola de histidinas. Por lo tanto, en este paso de purificación la fracción que pasa a través de la columna y que contiene las proteínas no retenidas es en la que se encuentra el dominio Amidasa 2-C

### 4.2.3 Cromatografía Líquida de Alta Presión (HPLC)

El paso final de purificación del dominio Amidasa 2-C se realizó mediante HPLC de fase reversa (RP-HPLC). La purificación se hizo utilizando como fase móvil acetonitrilo y agua, ambos conteniendo ácido trifluoroacético (TFA) al 0.05 %. La purificación fue monitoreada utilizando una lámpara de deuterio calibrada a 230 nm, longitud de onda a la cual es detectado el enlace peptídico. Se realizaron pruebas de gradientes para determinar el gradiente que sería utilizado para la purificación, estas pruebas se hicieron con proteína producida en medio LB. El dominio Amidasa 2-C enriquecido y no enriquecido fueron

purificados con el mismo gradiente obteniendo el mismo patrón cromatográfico en ambos casos.

La Figura 4.5 muestra el gradiente optimizado. Éste comienza en 15 % de acetonitrilo el cual se mantiene durante 5 minutos; en este momento las proteínas son retenidas en la columna acorde a su naturaleza hidrofóbica. Posteriormente se aumenta gradualmente la concentración de acetonitrilo en la fase móvil llevándola hasta 50 % en 15 minutos para eluir las proteínas que han sido retenidas en la columna. El siguiente paso fue aumentar la concentración de acetonitrilo hasta un 100 % y mantenerla así durante 10 minutos, esto con la finalidad de eluir cualquier proteína o compuesto que pudiera estar muy retenida en la columna con características no polares. Finalmente, se pasa de 100 % de acetonitrilo a 15 %; se mantiene en esa proporción para estabilizar la columna a las condiciones iniciales y realizar la siguiente corrida cromatográfica.

El cromatograma (Figura 4.5) muestra que el pico principal tiene un tiempo de retención de 19.8 minutos, el cual eluye con una concentración de acetonitrilo del 50 %. Esta fracción fue colectada y analizada por espectrometría de masas para comparar con la masa esperada para el dominio Amidasa 2-C. En un ambiente acuoso, las proteínas plegadas tienden a orientar las cadenas de los aminoácidos hidrofóbicos hacia el centro de la molécula, siendo pocos los que están expuestos al exterior. Por lo tanto, las proteínas plegadas tienen un tiempo de retención corto en comparación con las proteínas que se encuentran desnaturalizadas, las cuales tienen mayor superficie de contacto con la columna cromatográfica. Es muy probable que la proteína aislada en el pico principal del cromatograma se encuentre plegada ya que no se muestra muy retenida en la columna.

El ácido trifluoroacético (TFA) se usa en HPLC como agente apareador de iones mejorando la separación de proteínas ya que interacciona con los grupos cargados presentes en la fase estacionaria anulando los efectos que éstos tienen en la separación, interactúa con las proteínas acrecentando su carácter hidrofóbico y aumenta la interacción entre las proteínas y la fase estacionaria permitiendo una mejor separación de las mezcla de péptidos.<sup>(28)</sup>

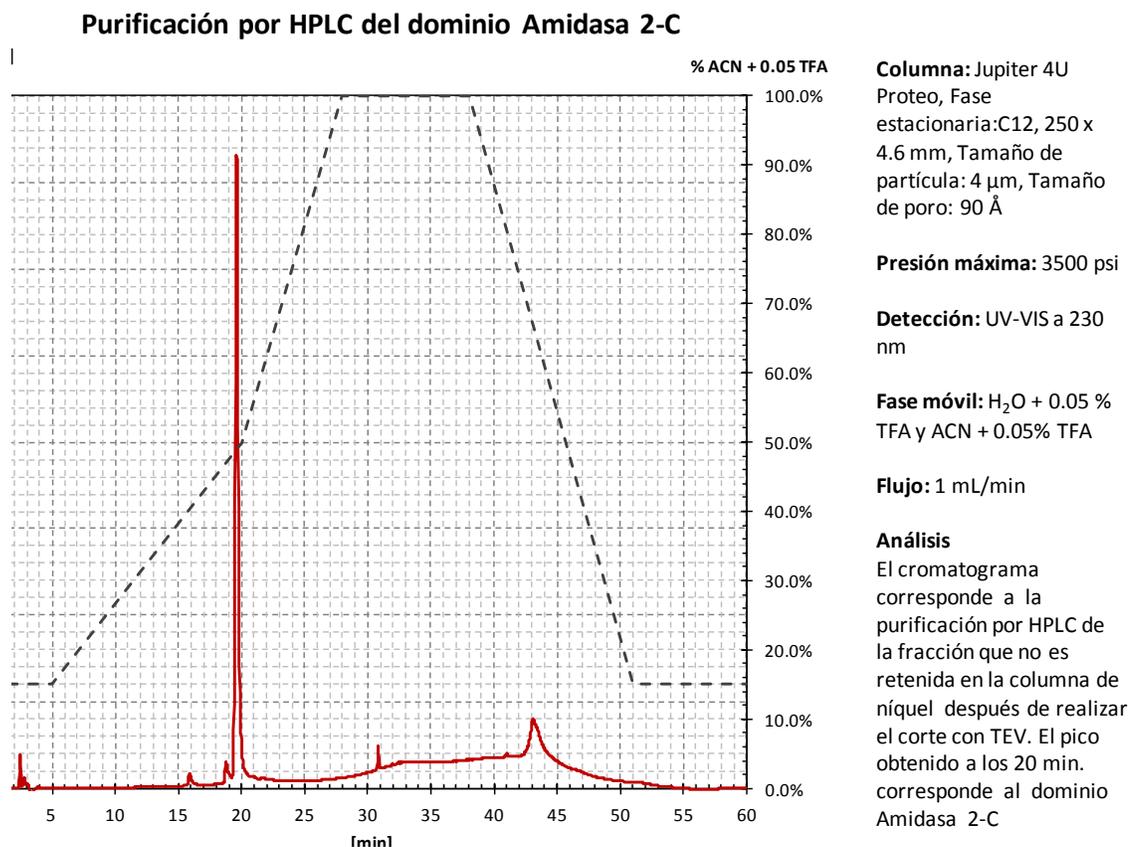


Figura 4.5. Se muestra el cromatograma obtenido de la purificación por HPLC del dominio Amidasa 2-C así como el gradiente utilizado y las condiciones de la corrida cromatográfica. La purificación se hizo en una columna C12 con un gradiente de ACN+ 0.05% de TFA que inicia en 15%. El pico principal tiene un tiempo de retención de 19.8 min.

## 4.3 CARACTERIZACIÓN DEL DOMINIO AMIDASAS 2-C

### 4.3.1 Espectrometría de masas

Para confirmar la masa de la proteína de fusión así como la masa del dominio Amidasa 2-C del pico correspondiente al tiempo de retención de 19.8 min. aislado por HPLC, se realizó un análisis mediante la técnica de MALDI-TOF (espectrometría de masas).

Se contaba con la masa promedio teórica de la proteína de fusión y del dominio Amidasa 2-C, en la que se considera la abundancia natural de los isótopos.<sup>(39)</sup>

Se realizó el cálculo para saber la masa teórica del dominio Amidasa 2-C marcada isotópicamente con <sup>15</sup>N y <sup>13</sup>C obteniéndose un valor de 7989.396 g/mol considerando un enriquecimiento isotópico del 100%.<sup>(40)</sup>

## Resultados y Discusión

Los resultados en los espectros de masas MALDI-TOF están dados por la relación masa/carga y la intensidad absoluta, es decir, el número de iones de cada especie que llegan al detector. Las moléculas de proteína están contenidas en una matriz ácida que además de protegerlas para evitar que se fragmenten al ser irradiadas, permite la ionización. Cuando las moléculas de proteína son irradiadas con rayo láser se ionizan tomando generalmente sólo un protón de la matriz ácida. Así, son impulsadas por un acelerador hasta que adquieren un potencial eléctrico elevado y después son separadas en función del tiempo que les toma llegar al detector, este tiempo es directamente proporcional a la masa de la molécula dividida entre su carga. Si la molécula de proteína se ha ionizado solo con un protón de la matriz, el resultado representa la masa de la molécula más un protón ( $MH^+$ ). Algunas veces, la molécula de proteína se ioniza tomando dos protones de la matriz, cuando esto ocurre la molécula viaja más rápido por tener un mayor impulso debido a un incremento en la carga; por tanto, llega más rápido al detector y la masa se representa como el cociente de la masa de la proteína entre 2, ya que está doblemente cargada. El resultado se observa como un pico en el espectro con la mitad de la masa esperada para la proteína ( $MH^{2+}$ ).

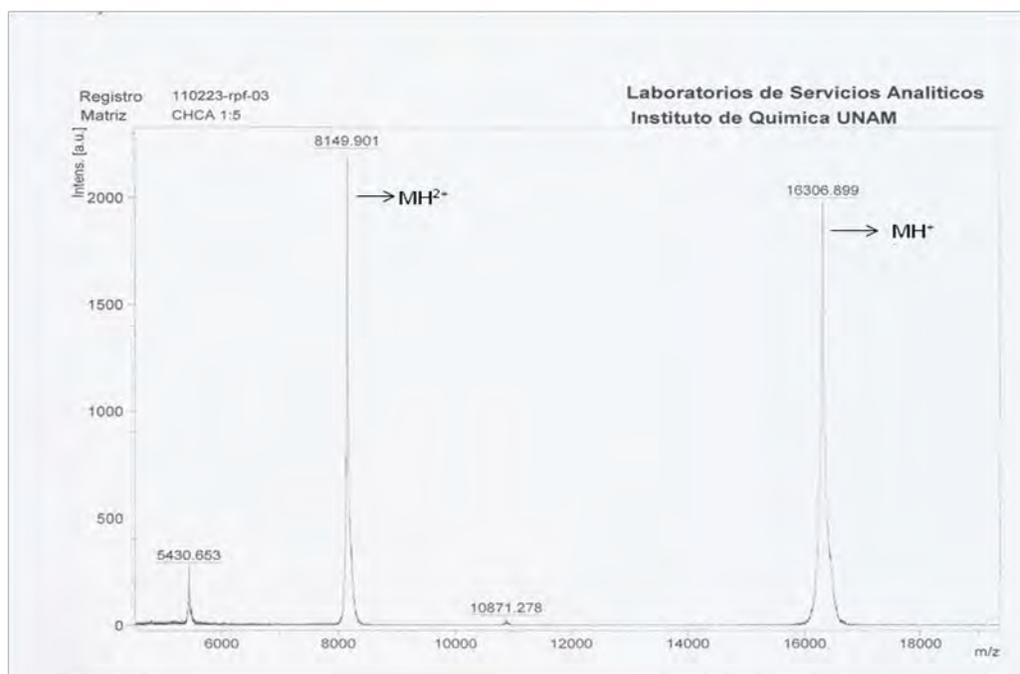


Figura 4.6 Espectro de masas MALDI-TOF de la proteína de fusión GB1- Amidasa 2-C.

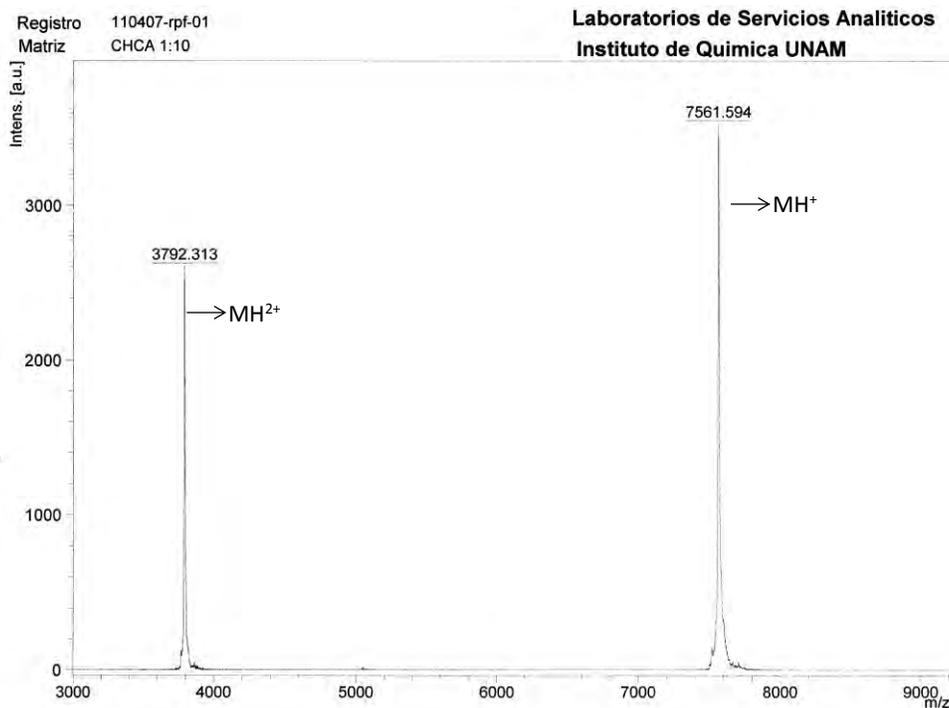


Figura. 4.7. Espectro de masas MALDI-TOF obtenido de la fracción aislada por HPLC de la purificación del dominio Amidasa 2-C.

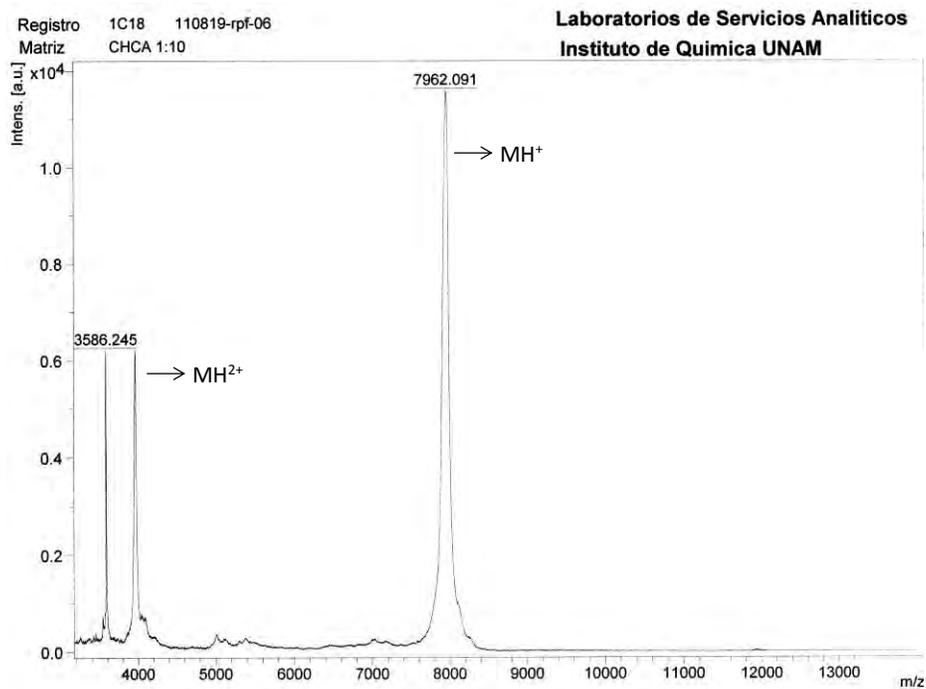


Figura. 4.8. Espectro de masas MALDI-TOF obtenido de la fracción aislada por HPLC de la purificación del dominio Amidasa 2-C marcado isotópicamente ( $^{15}N$ ,  $^{13}C$ -dominio Amidasa 2-C).

## Resultados y Discusión

Muestra	Masa calculada MH <sup>+</sup> (g/mol)	Masa experimental MH <sup>+</sup> (g/mol)	Masa calculada MH <sup>2+</sup> (g/mol)	Masa experimental MH <sup>2+</sup> (g/mol)
Proteína de fusión	16301.1	16306.899	8150.55	8149.901
Dominio Amidasa 2C	7561.1	7561.594	3780.55	3792.313
<sup>15</sup> N, <sup>13</sup> C-Dominio Amidasa 2C	7990.396	7962.091	3995.69	3981.045

Tabla 4.2 Se muestran los cálculos teóricos para espectrometría de masas MALDI TOF considerando la masa promedio así como los resultados obtenidos experimentalmente. Los cálculos teóricos hechos para el dominio Amidasa 2-C marcado isotópicamente consideran el 100 % de enriquecimiento isotópico.

Los espectros de masas MALDI- TOF corroboraron la masa esperada tanto para la proteína de fusión como para el dominio Amidasa 2-C no enriquecido. La diferencia de unidades que se observa entre los cálculos teóricos y los resultados experimentales (Tabla 4.2) se debe al error experimental, ya que la relación masa carga determinada depende de la también de la distribución natural de los isótopos de cada elemento que han sido incorporados en la proteína. Con estos resultados y todo el proceso de biología molecular podemos concluir sin ambigüedad que tenemos la proteína deseada.

La Figura 4.8 muestra una masa experimental menor a la masa calculada para el dominio Amidasa 2-C enriquecido al 100 %; esto indica un enriquecimiento isotópico menor al 100 %. El enriquecimiento isotópico del dominio Amidasa 2-C fue del 99.65 %.

### 4.3.2 Máximo de absorción en la región del UV

El espectro de absorción de una muestra representa la fracción de la radiación electromagnética incidente que absorbe dentro de un intervalo de longitudes de onda. <sup>(41)</sup> En la región lejana del UV (típicamente de 240 nm a 190 o 180 nm) hay absorción por parte del enlace peptídico. Existe una débil transición  $n \rightarrow \pi^*$  alrededor de 220 nm y una transición intensa  $\pi \rightarrow \pi^*$  alrededor de 190 nm <sup>(42)</sup>. Cada uno de los aminoácidos tiene un espectro característico. El enlace peptídico absorbe a una longitud de onda y los aminoácidos aromáticos a otra provocando que cada aminoácido adicione señales particulares a la proteína en general, si no hay aminoácidos aromáticos la proteína no absorbe a  $\sim 285$  nm.

En el UV cercano las cadenas de los aminoácidos aromáticos absorben en un intervalo de 250 a 290 nm. <sup>(42)</sup>. El triptófano (Trp) muestra absorción entre 290 y 305 nm; la tirosina

(Tyr) tiene un pico entre 275 y 282 nm y la fenilalanina (Phe) muestra una absorción débil entre 255 y 270 nm.<sup>(43)</sup>

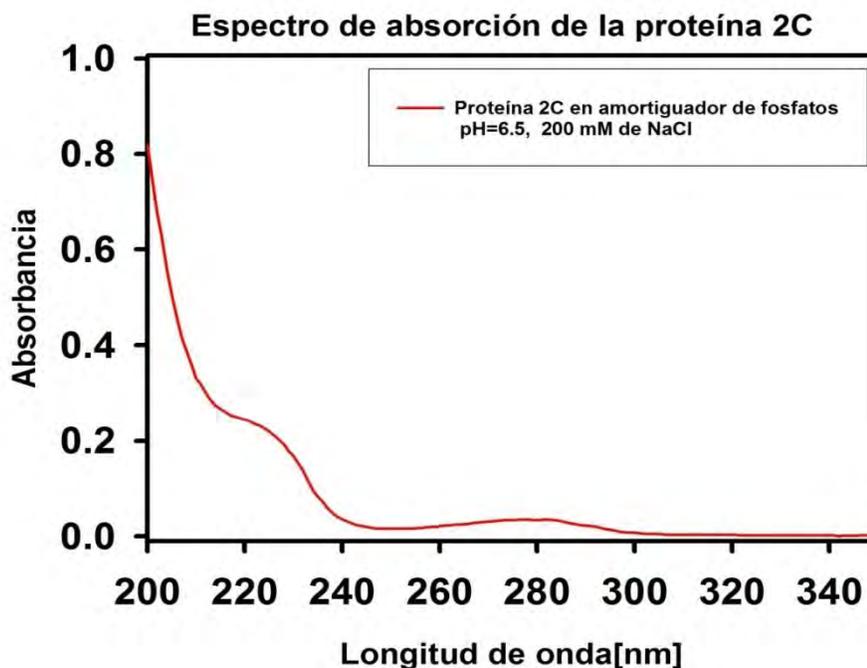


Figura 4.9. Espectro de absorción del dominio Amidasa 2-C. El espectro muestra absorción correspondiente al enlace peptídico entre 210 y 240 nm y un segundo pico correspondiente a los aminoácidos aromáticos entre 260 y 300 nm.

En la Figura 4.9 se muestra el espectro de absorción obtenido para el dominio Amidasa 2-C. La absorción entre 210 y 240 nm corresponde al enlace peptídico y la absorción a 280 nm está dada por los aminoácidos aromáticos presentes en la estructura de la proteína.

#### 4.4 RESONANCIA MAGNÉTICA NUCLEAR (RMN)

La RMN es una técnica espectroscópica que permite obtener información sobre el ambiente electrónico en el que se encuentran inmersos los núcleos con espín. Los núcleos con espín 1/2 y que se encuentran en las moléculas biológicas son el  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{31}\text{P}$  y  $^{15}\text{N}$ . Todos ellos pueden ser observados por esta técnica. Sin embargo, otros núcleos más abundantes como el  $^{12}\text{C}$  y el  $^{16}\text{O}$  no poseen espín y por tanto no pueden observarse por RMN.

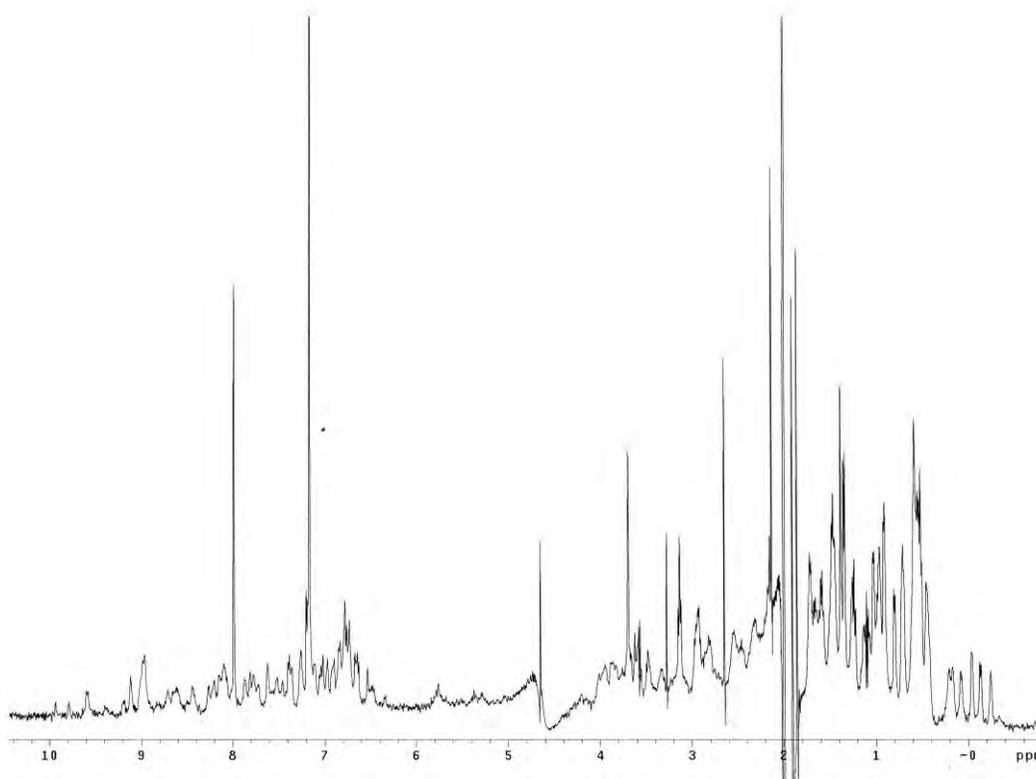
## Resultados y Discusión

---

Los parámetros más usados para la caracterización de biomoléculas (proteínas, ADN y ARN) por RMN son el desplazamiento químico, los acoplamiento escalar y dipolo-dipolo, los mecanismos de relajación de los núcleos y la integración de las señales. Con estos parámetros generalmente es posible determinar la estructura tridimensional, dinámica e interacción inter e intramolecular de las biomoléculas.

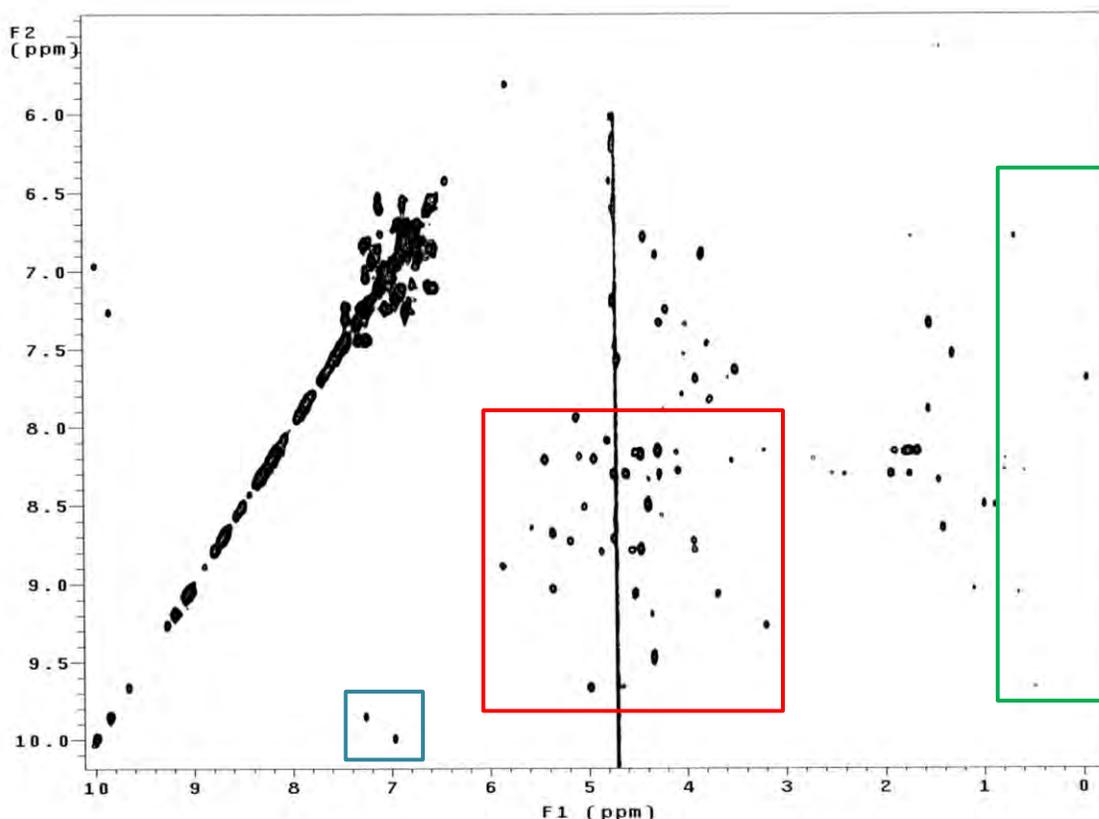
En la Figura 4.10 se muestra el espectro de protón del dominio Amidasa 2-C. En él se puede observar que existen señales de -0.03 hasta 10.0 ppm. Normalmente, las señales de muestras no plegadas se encuentran de 0.7 hasta 9.0 ppm. Esta amplia distribución de señales indica que los diversos protones en la proteína se encuentran inmersos en una gran variedad de ambientes químicos estables, lo cual generalmente indica que la proteína tiene una estructura tridimensional definida. Este simple resultado que se puede obtener en unos cuantos minutos es condición suficiente para establecer que la molécula se encuentra plegada; pero, este resultado no es suficiente para determinar su estructura tridimensional. Para poder establecer el tipo de plegamiento, tanto su estructura secundaria como terciaria, es necesario contar con espectros en dos y en tres dimensiones que permitan la asignación de las señales al átomo que las origina.

Los espectros de RMN en dos dimensiones permiten dispersar las señales en una dimensión adicional resolviendo en alguna medida la sobreposición de señales que se observa en los espectros de una dimensión. Los espectros en dos dimensiones se pueden clasificar en homonucleares y heteronucleares. Los espectros homonucleares más comunes son el COSY, TOCSY y NOESY. El experimento COSY permite encontrar la correlación entre dos protones a dos y a tres enlaces de distancia. El experimento TOCSY permite encontrar a los sistemas de espín que pertenecen al mismo aminoácido; sin embargo, no es posible diferenciar a aminoácidos que tengan el mismo sistema de espín como, por ejemplo, los aminoácidos aromáticos. El NOESY permite encontrar correlación a través del espacio para posteriormente determinar distancias entre los núcleos.



**Figura 4.10.** Espectro de protón del dominio Amidasasa 2-C. Existen señales de -0.03 hasta 10.0 ppm. Esta amplia distribución de señales indica que los diversos protones en la proteína se encuentran inmersos en una gran variedad de ambientes químicos estables, lo cual generalmente indica que la proteína tiene una estructura tridimensional definida.

En la Figura 4.11 se muestra un experimento TOCSY. Este experimento en dos dimensiones muestra las correlaciones de las señales que pertenecen al mismo aminoácido, o comúnmente llamado como el sistema de espín. En él se observa las correlaciones en las coordenadas ( $\sim 7.0$  ppm,  $\sim 10.0$  ppm) las señales del HN y CH del grupo indol de los dos triptófanos presentes en la proteína, recuadro azul en la Figura 4.11. Se observan las señales de  $-\text{CH}_3$  que se encuentran a alto campo, recuadro verde en la Figura 4.11, indicando frecuentemente que algunos metilos entran en la región de protección de anillos aromáticos. Estos resultados sólo son posibles si la molécula adopta una conformación particular. Así mismo se observa en el recuadro rojo a las correlaciones provenientes de los HN y los  $\text{H}\alpha$ , en donde se observa que el desplazamiento químico de los  $\text{H}\alpha$  se encuentra a mayores partes por millón, lo cual es indicativo de la presencia de hebras beta.



**Figura 4.11** Espectro TOCSY del dominio Amidasas 2-C. El recuadro azul muestra las señales del HN y CH del grupo indol de los dos triptófanos presentes en la proteína; en el recuadro verde se observan las señales de  $-CH_3$  que se encuentran a alto campo y en el recuadro rojo se observan las correlaciones provenientes de los HN y los  $H\alpha$ , en donde se observa que el desplazamiento químico de los  $H\alpha$  se encuentra a mayores partes por millón, lo cual es indicativo de la presencia de hojas beta.

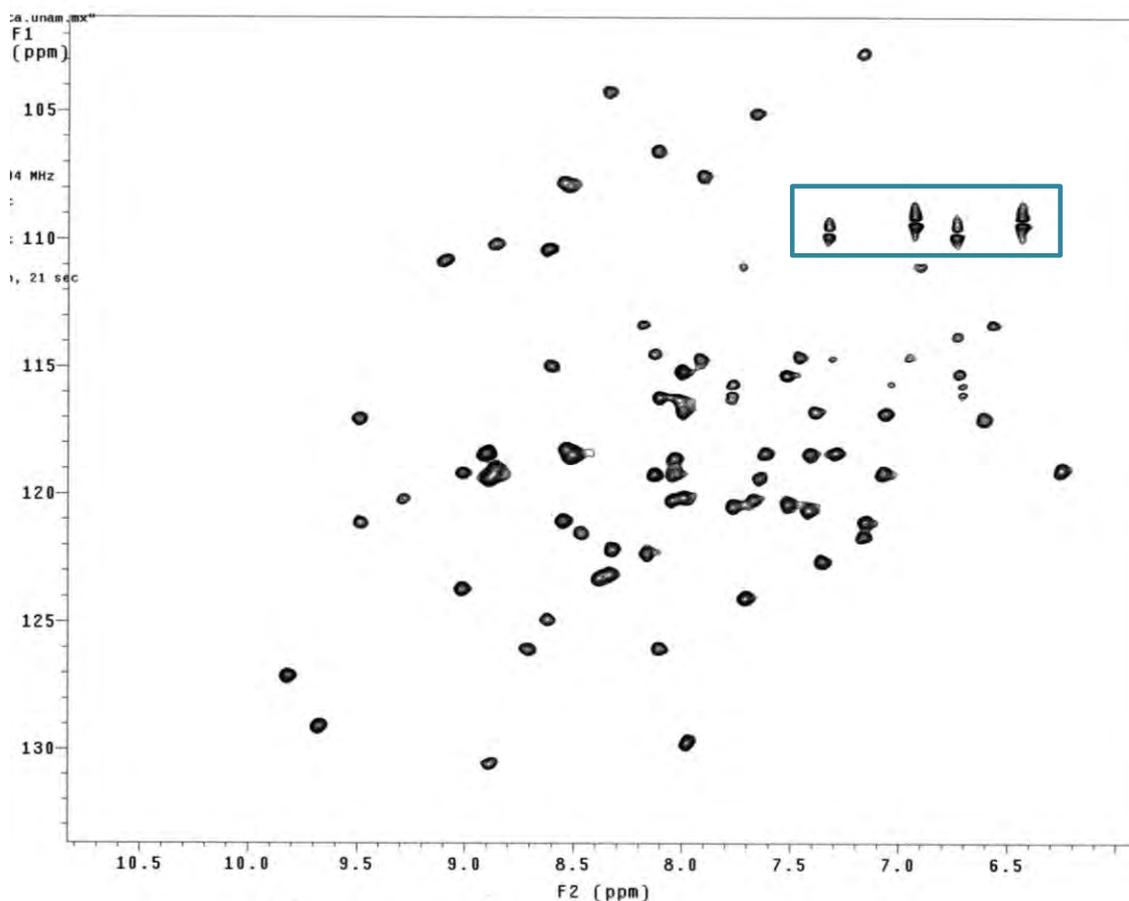
Dado que el dominio Amidasas 2-C tiene más de 50 aminoácidos, la interpretación de los espectros se facilita si se puede contar con la proteína marcada con  $^{15}N$  y  $^{13}C$ . Por tal motivo, se optimizaron los procedimientos para obtener proteína enriquecida con estos isótopos mediante expresión en medio mínimo. Cuando la proteína se expresa en medio donde la única fuente de  $^{15}N$  es  $(^{15}NH_4)_2SO_4$ , y la única fuente de  $^{13}C$  es glucosa marcada completamente con  $^{13}C$ , es posible obtener espectros de correlación HN,  $^{15}N$  y  $^{13}C$ . Con ellos la interpretación de los espectros se facilita; dado que ahora es posible encontrar correlaciones no sólo entre protones sino también entre  $^1H$  y  $^{15}N$  y entre  $^1H$  y  $^{13}C$ .

El espectro heteronuclear en dos dimensiones más común es el HSQC, que permite detectar interacciones a un enlace de distancia entre  $^1H$  y  $^{13}C$  o entre  $^1H$  y  $^{15}N$ . Este experimento en sus dos variantes, es el experimento preliminar para iniciar con los espectros en tres

## Resultados y Discusión

dimensiones que permitirán la interpretación de los espectros en tres dimensiones.

En la Figura 4.12 se muestra el experimento de correlación a un enlace de distancia HN- $^{15}\text{N}$ ,  $^{15}\text{N}$ -HSQC, en el cual se observa que el enriquecimiento con  $^{15}\text{N}$  fue exitoso, que la Amidasa 2-C se encuentra plegada y la intensidad y distribución de las señales sugiere que no será complejo determinar su estructura por RMN. Las señales en antifase, recuadro azul, correspondientes a las cadenas laterales de argininas, asparaginas y glutaminas.



**Figura 4.12** Espectro  $^{15}\text{N}$ - HSQC del dominio Amidasa 2-C. En el recuadro azul se muestran las señales en antifase correspondientes a las argininas, asparaginas y glutaminas.

En la Figura 4.13 se muestra el experimento de correlación a un enlace de distancia H- $^{13}\text{C}$ ,  $^{13}\text{C}$ -HSQC, al igual que el experimento anterior se muestra el enriquecimiento con  $^{13}\text{C}$ . Las señales que se muestran dentro del recuadro rojo pertenecen a los carbonos  $\alpha$  de la proteína, las señales dentro del recuadro azul corresponden a los carbonos de los aminoácidos aromáticos y las señales dentro del recuadro verde corresponden a los metilos

de la proteína.

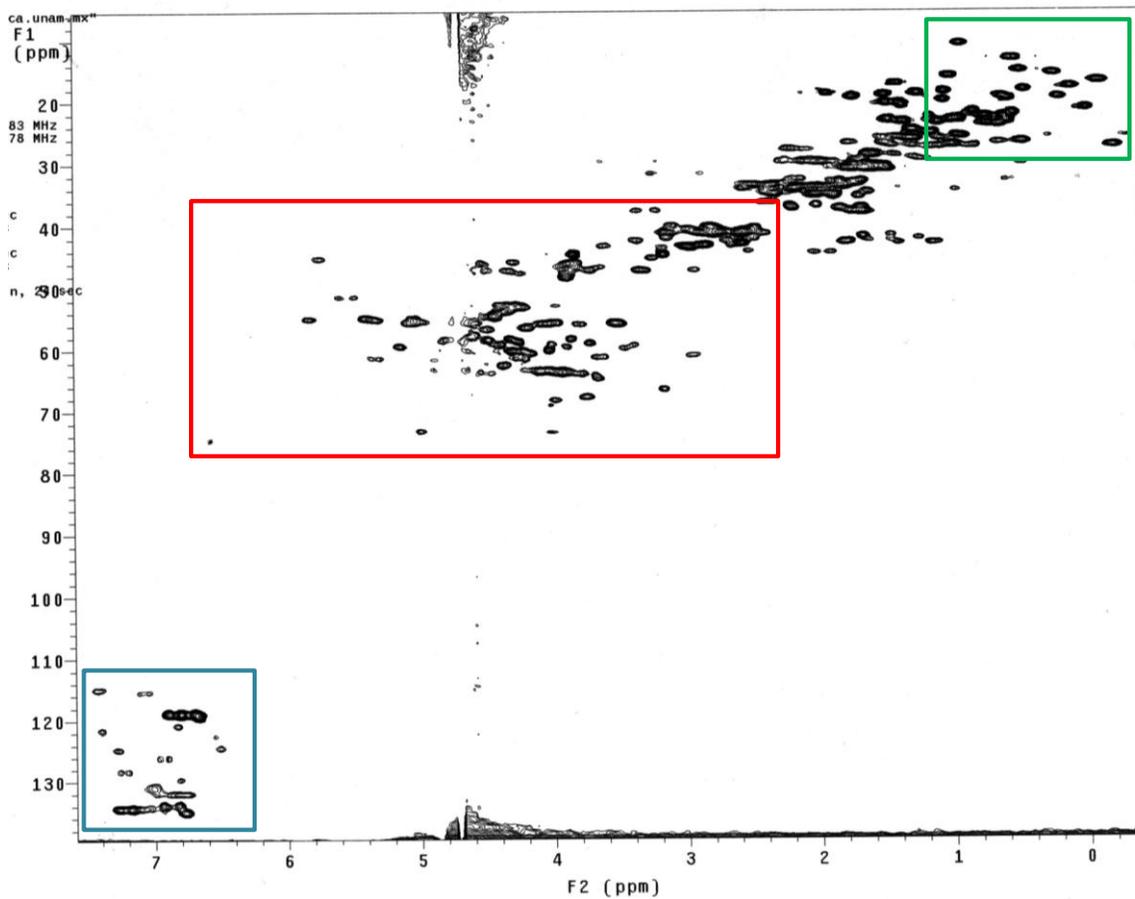


Figura 4.13 Espectro  $^{13}\text{C}$ -HSQC del dominio Amidasin 2-C. Las señales que se muestran dentro del recuadro rojo pertenecen a los carbonos  $\alpha$  de la proteína, las señales dentro del recuadro azul corresponden a los carbonos de los aminoácidos aromáticos y las señales dentro del recuadro verde corresponden a los metilos de la proteína.

---

# *Conclusiones*

### 5. CONCLUSIONES

El dominio de reconocimiento de carbohidratos Amidasa 2-C de la N-acetilmuramoil-L-alanina amidasa 3 proveniente de la bacteria patógena *Bacillus cereus* (Uniprot Q736Q1\_BACC1), se pudo expresar en *Escherichia coli* cultivada en medio rico en nutrientes (LB) y medio mínimo (PG).

Debido a que los isótopos más abundantes de carbono y nitrógeno ( $^{12}\text{C}$  y  $^{14}\text{N}$ ) son invisibles para la técnica de RMN, fue necesario optimizar las condiciones de expresión en medio con  $^{15}\text{NH}_4^+$  y  $^{13}\text{C}_6$ -glucosa como únicas fuentes de nitrógeno y carbono respectivamente; y de este modo expresar proteína marcada con estos isótopos. Dado el elevado costo de estos compuestos, fue de particular importancia encontrar condiciones que maximizarán rendimiento y al mismo tiempo minimizaran la cantidad de  $^{15}\text{NH}_4^+$  y  $^{13}\text{C}_6$ - glucosa utilizados. Esto se logró implementando la técnica de reducción de volumen. En este método, las bacterias son crecidas en medio LB hasta que llegan a la densidad adecuada para la inducción de expresión y sólo entonces se transfieren a tan solo una cuarta parte del volumen inicial de medio mínimo marcado.

El dominio Amidasa 2-C pudo ser aislado del lisado celular utilizando cromatografía de afinidad metal-quelato, tratado con la proteasa TEV para separarlo de las secuencias auxiliares, y finalmente purificado mediante cromatografía líquida de alta eficiencia en fase reversa (RP-HPLC).

El dominio Amidasa 2-C así purificado fue caracterizado por espectrometría de masas utilizando la técnica MALDI-TOF, así como espectroscopía UV-visible y RMN. Los espectros obtenidos por RMN confirman que la proteína se encuentra formando una estructura tridimensional definida y estable. La calidad de los espectros respecto a la distribución de las señales, la mínima superposición y la homogeneidad en su intensidad sugiere que esta proteína es apta para estudios estructurales y de unión de posibles ligandos. Es poco lo que se sabe de la interacción entre los dominios bacterianos de reconocimiento de carbohidratos y sus blancos. Además, es posible que la inhibición de estas interacciones sea una estrategia terapéutica efectiva para el tratamiento de infecciones bacterianas. El trabajo desarrollado aquí ha abierto las puertas para subsecuentes investigaciones que tienen el potencial de contribuir en el conocimiento y caracterización de estos sistemas.

## Conclusiones

---

El modelo por homología del dominio Amidasa 2-C obtenido en el servidor I- TASSER ONLINE Protein Structure & Function Predictions sugiere una estructura secundaria compuesta por una secuencia de elementos hebra  $\beta 1$  - hélice  $\alpha 1$  - hebra  $\beta 2$  - hebra  $\beta 3$  - hélice  $\alpha 2$  - hebra  $\beta 4$ , en donde todas las hebras forman una lámina beta antiparalela  $\beta 2 \beta 3 \beta 1 \beta 4$ . Dicha estructura es semejante a la estructura secundaria de otro dominio de reconocimiento a carbohidratos denominado SPOR. Se requiere obtener la estructura de forma experimental para confirmar este resultado.

---

# Bibliografía

## 6. BIBLIOGRAFÍA

1. **Jessica Humann and Laurel L. Lenz**, *Bacterial peptidoglycan degrading enzymes and their impact on host muropeptide detection*, *J. Innate. Immun.*, **1** (2009) 88-97.
2. **Waldemar Vollmer, Ute Bertsche**, *Murein (peptidoglycan) structure, architecture and biosynthesis in Escherichia coli*, *B. B. A. Biomembranes.*, **1778** (2008) 1714–1734.
3. **Waldemar Vollmer, Bernard Joris, Paulette Charlier, Simon Foster**, *Bacterial peptidoglycan (murein) hydrolases*, *FEMS Microbiol Rev*, **32** (2008) 259–286.
4. **Rowena L. Lock and Elizabeth J. Harry** *Cell-division inhibitors: new insights for future antibiotics*, *Nat. Rev. Drug Discov.*, **7**, 2008, 324-338.
5. **Mark Fenton, Paul Ross, Olivia McAuliffe, Jim O'Mahony and Aidan Coffey**, *Recombinant bacteriophage lysins as antibacterials*, *Bioeng Bugs*, **1** (2010) 9-16.
6. **Christoph Heidrich, Markus F. Templin, Astrid Ursinus, Melisa Merdanovic, Jurgen Berger, Heinz Schwarz, Miguel A. de Pedro and Joachim-Volker Holtje**, *Involvement of N-acetylmuramyl-L-alanine amidases in cell separation and antibiotic induced autolysis of Escherichia coli*, *Mol Microbiol.*, **41** (2001) 167–178.
7. **Malgorzata Firczuk and Matthias Bochtler**, *Folds and activities of peptidoglycan amidases*, *FEMS Microbiol Rev.*, **31** (2007) 676–691
8. <http://www.uniprot.org/uniprot/Q736Q1>
9. <http://pfam.sanger.ac.uk/family/PF12123.3#tabview=tab1>
10. **Christopher K. Mathews, Kevin G. Ahern, K. E. Van Holde**, *Bioquímica*, 3ª Edición, Pearson Education, Madrid (2002).
11. **Thomas E. Creighton**, *Proteins: structures and molecular properties*, W.H. Freeman and Company, New York, (1993) Capítulo 5, 171-198.

12. **José Luque Cabrera, Ángel Herráez Sánchez**, *Texto ilustrado de Biología Molecular e Ingeniería Genética*, Elsevier Health Sciences (2001)
13. **Bruce Alberts, Dennis Bray, Julian Lewis** *Biología Molecular de la Célula*, 3ª Edición, Ediciones Omega (1999)
14. **Claudia Andrea Senegal Kischinevzky y Jesús Ortega Lule**, *Manual de prácticas Biología Molecular de la Célula I*, Publidisa, Facultad de Ciencias UNAM, México (2005) 77-78.
15. **F. William Studier, Alan H. Rosenberg, John J. Dunn, John W. Dubendorff**, *Use of T7 RNA Polymerase to Direct Expression of Cloned Genes*, *Methods Enzymol.*, **185** (1990) 60 - 89.
16. **Hansen LH, Knudsen S**, *The effect of the lacY gene on the induction of IPTG inducible promoters, studied in Escherichia coli and Pseudomonas fluorescens*, *Curr. Microbiol.* **36** (1998) 341-347.
17. **Juan Jáuregui Rincón y Norma Angélica Chávez Vela**, *Glosario de Biotecnología*, 1ª Edición, Universidad Autónoma de Aguascalientes, México (2006) 165.
18. **Dana M. Francis and Rebecca Page** *Strategies to Optimize Protein Expression in E.Coli* , *Curr. Protoc. Protein*, Chapter 5 (2010) 29.
19. **Jeremy Mark Berg, Lubert Stryer y John L. Tymoczko**, *Bioquímica*, Editorial Reverté, España (2008) 243.
20. **Werner-Müller Esterl, Ulrich Brand, Oliver Anderka, Stefan KieB, Katrin Ridinger y Michael Plenikowski**, *Bioquímica: Fundamentos para Medicina y ciencias de la vida.*, Editorial Reverté, España (2008) 168.
21. **Flores Solís David**, *Expresión y caracterización por métodos espectroscópicos (RMN) de la Defensina 4*, Tesis de Licenciatura, Facultad de Química UNAM, México D.F. (2010).

22. **Wen-Jing Bao, Yong-Guang Gao**, *Highly efficient expression and purification system of small-size protein domains in Escherichia coli for biochemical characterization*, Protein Expres. Purif., **47** (2006) 599-606.
23. **David S. Waugh**, *An overview of enzymatic reagents for the removal of affinity tags*, Protein Expres. Purif. **80** (2011) 283–293.
24. **José Federico del Río Portilla**, *Determinación de la estructura de proteínas por Resonancia Magnética Nuclear*, Mensaje Bioquímico, Vol. XXVII. Dpto. de Bioquímica, Facultad de Medicina, UNAM, Ciudad Universitaria, México D.F. (2003).
25. **Studier, F. William**. *Protein production by auto-induction in high-density shaking cultures*, Protein Expres. Purif. **41** (2005) 207 - 234.
26. **Marley, Jonathan, Lu, Min y Bracken, Clay**, *A method for efficient isotopic labeling of recombinant proteins*, J.Biomol. NMR, **20** (2001) 71 - 75.
27. **Joseph Sambrook, David Willian Rissell**, *Molecular Cloning: a laboratory manual*, Volume 3, 3ª Edición, CSHL Press, New York (2001) 15.44-15.48.
28. **Robert L. Cunico, Karen M. Gooding y Tim Wehr**, *Basic HPLC and CE of Biomolecules*, 1ª Edición, Bay Bioanalytical Laboratory, Richmond California (1998).
29. **Shiranthi Keppetipola, Wieslaw Kudlicki, Bao D. Nguyen, Xi Meng, Kevin J. Donovan, and A. J. Shaka**, *From Gene to HSQC in under Five Hours: High-Throughput NMR Proteomics*, JACS, **128** (2006) 4508- 4509.
30. **Gerard J. Tortora, Berdell R. Funke, Christine L. Case**, *Introducción a la microbiología*, 9ª edición, Editorial Médica Panamericana, Argentina, (2007) 85-89.
31. **Albert Parisien and Christopher Q. Lan**, *Classification of bacterial cell wall hydrolases and their potentials as novel alternatives to antibiotics – a response to the letter of Biziulevicius and Kazlauskaite*, J. Appl. Microbiol., **106** (2009) 1754–1759.
32. <http://smart.embl-heidelberg.de/>

33. **H. Hashimoto**, *Recent structural studies of carbohydrate-binding modules*, Cell. Mol. Life Sci., **63** (2006) 2954- 2967.
34. <http://zhanglab.ccmb.med.umich.edu/I-TASSER/>
35. <http://pfam.sanger.ac.uk/family/PF05036#tabview=tab0>
36. **Davis S. Waugh**, *Laboratorio de cristalografía macromolecular, sección de ingeniería de proteínas, TEV portease FAQ*, Instituto Nacional de Cáncer, septiembre 2010 [mcl1.ncifcrf.gov/waugh\\_tech/faq/tev.pdf](http://mcl1.ncifcrf.gov/waugh_tech/faq/tev.pdf)
37. **Hillenkamp, F., Karas, M., Beavis, R. C., Chait B. T.**, *Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers*, Anal Chem., **63** (1991) 1193A-1203A.
38. **U. Bahr, M. Karas, F. Hillenkamp**, *Analysis of biopolymer by Matrix-Assisted Laser Desorption/Ionization (MALDI) mass spectrometry*, J. Anal Chem., **348** (1994) 783-791.
39. **Bioinformatics, Swiss Institute of.**, *ExPASy (Expert Protein Analysis System)*. [En línea] <http://www.expasy.ch/>
40. *Protein calculator*, Putnam Chris [En línea].
41. **Douglas A. Skoog, Donald M. West**, *Introducción a la Química Analítica*, Editorial Reverté, Barcelona (1986) Capítulo 20.
42. **Sharon M. Kelly, Nicholas C. Price**, *The application of circular dichroism to studies of protein folding and unfolding*, B.B.A. Biomembranes, **1338** (1997). 161–185.
43. **Sharon M. Kelly, Thomas J. Jess, Nicholas C. Price**, *How to study proteins by circular dichroism*, B.B.A Biomembranes, **1751** (2005) 119 – 139.
44. **John M. Walker**, *The protein protocols Handbook*, Humana Press, Totowa New Jersey (2002) 15-20.

45. **Tsaffrir Zor and Zvi Selinger**, *Linearization of the Bradford Protein Assay Increases Its Sensitivity: Theoretical and Experimental Studies*, *Anal. Biochem.*, **236** (1996) 302 – 308.

46. <http://www.piercenet.com/browse.cfm?fldID=02020104>

47. **Thermo Scientific**, nota científica, Pierce Biotechnology, *Coomassie Plus (Bradford) Assay Reagent*, ficha técnica 23238.

---

# Apéndices

**APÉNDICE A. PROTOCOLOS DE CÉLULAS COMPETENTES Y TRANSFORMACIÓN. <sup>(21)</sup>**

- I. Seleccionar una colonia aislada e inocular 100 mL de medio LB en un matraz de 250 mL. Incubar con agitación a 37 °C hasta alcanzar un valor de densidad óptica de 0.4 medido a 600 nm.
- II. Transferir el cultivo a dos tubos estériles de 50 mL e incubar en hielo por 10 min.
- III. Centrifugar durante 10 minutos a 4,000 x g a 4 °C.
- IV. Decantar y eliminar el medio restante invirtiendo los tubos por 1 min.
- V. Resuspender cada paquete celular en 10 mL de CaCl<sub>2</sub> 0.1 M (estéril y previamente incubado en hielo) e incubar en hielo durante 10 min.
- VI. Centrifugar a 2,500 x g durante 5 minutos a 4 °C. Desechar el sobrenadante.
- VII. Repetir el paso V y VI para un total de 3 incubaciones en cloruro de calcio.
- VIII. Resuspender cada paquete celular en 2 mL de CaCl<sub>2</sub> 0.1 M (frío y estéril).
- IX. Agregar 0.5 mL de glicerol al 50 % w/v a cada 2 mL de suspensión del paso anterior.
- X. Preparar alícuotas de 200 µL en tubos estériles con tapa de rosca y guardarlos a -80 °C.

**TRANSFORMACIÓN**

- I. Tomar un tubo de células competentes e incubarlo en hielo hasta que se derrita.
- II. Transferir 50 µL de células competentes a un tubo frío de microfuga de 1.5 mL.
- III. Añadir 2-5 µL del DNA plasmídico que se desea transformar, mezclar cuidadosamente con el tubo tapado.
- IV. Mantener en hielo durante 30 minutos.
- V. Transferir el tubo a un baño de agua a 42 °C durante 45 segundos.

**VI.** Regresar el tubo al hielo por 2 minutos.

**VII.** Añadir 200  $\mu$ L de medio LB o SOC e incubar las células durante 1 hora con agitación ocasional.

**VIII.** Tomar 50 – 100  $\mu$ L del cultivo y sembrarlo en una caja con medio LB sólido con los antibióticos apropiados.

**APÉNDICE B. COMPOSICIÓN DE LOS MEDIOS DE CULTIVO <sup>(21)</sup>**

**MEDIO LURIA BERTANI (LB)**

• Tripsina	10 g/L
• Extracto de levadura	5 g/L
• NaCl	10 g/L
• pH	7.0

Mezclar los primeros tres componentes en 1 L de agua desionizada o destilada y ajustar el pH de la disolución con NaOH o HCl según sea necesario. Esterilizar en el autoclave.

**MEDIO MÍNIMO PG**

Condiciones diseñadas para preparar un litro de cultivo.

• Na <sub>2</sub> HPO <sub>4</sub>	50 mM
• KH <sub>2</sub> PO <sub>4</sub>	50 mM
• (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	25 mM
• MgSO <sub>4</sub>	2 mM
• “Stock” de Metales traza	1 mL
• “Stock” de Vitaminas	1 mL
• Glucosa	5 g/L

**Disolución “stock” de disolución amortiguadora (10X)**

• Na <sub>2</sub> HPO <sub>4</sub>	0.5 M
• KH <sub>2</sub> PO <sub>4</sub>	0.5 M

Esterilizar por autoclave

**Disolución “stock” de metales traza (1000X)**

• FeCl <sub>3</sub> (en ~ 0.3 % HCl)	0.1 M	5.0 mL
• CaCl <sub>2</sub>	0.2 M	1.0 mL
• MnCl <sub>2</sub>	0.4 M	0.25 mL
• ZnSO <sub>4</sub>	0.2 M	0.5 mL

---

• CoCl <sub>2</sub>	0.1 M	0.2 mL
• CuSO <sub>4</sub>	0.1 M	0.2 mL
• NiSO <sub>4</sub>	0.1 M	0.2 mL
• H <sub>3</sub> BO <sub>3</sub>	0.1 M	*0.2 mL

\*Disolver en H<sub>2</sub>O caliente

Agregar H<sub>2</sub>O hasta aforar 10 mL y esterilizar por filtración.

(1 mL de “stock” de metales por litro de cultivo)

**“Stock” de magnesio (1000x)**

• MgSO <sub>4</sub>	2 M
---------------------	-----

(1 mL de “stock” de magnesio por litro de cultivo)

**Disolución “stock” de Vitaminas (1000X)**

• Cloruro de Colina	1.0 g/L
• Ácido Fólico	1.0 g/L
• Ácido D-Pantoténico	1.0 g/L
• Mio-inositol	2.0 g/L
• Nicotinamida (Niacinamide)	1.0 g/L
• Piridoxal HCl	1.0 g/L
• Tiamine HCl	1.0 g/L
• Riboflavin	0.1 g/L
• D-Biotina	1.0 g/L

Esterilizar por filtración

Guardar en alícuotas de 1 mL a -20°C

(1 mL de “stock” de vitaminas por litro de cultivo)

---

**APÉNDICE C. PROTOCOLOS DE EXPRESIÓN PARA 1 L DE MEDIO MÍNIMO**  
(21)

**EXPRESIÓN EN MEDIO PG CON REDUCCIÓN DE VOLUMEN**

**I.** Seleccionar una colonia aislada de una transformación reciente o desprender una hojuela de un stock de glicerol. Inocular 10 ml de medio LB. Crecer el cultivo incubando las células por 12-16 horas a 30 °C ó 37 °C con agitación. Este cultivo se puede dejar creciendo durante una noche. Al día siguiente se inocula un litro de medio LB.

**II.** Crecer el cultivo incubándolo a 37 °C con agitación hasta que alcance un valor de densidad óptica de 0.7 a 600 nm. Alcanzado el valor de densidad óptica, centrifugar el cultivo (3000 x g por 5-10 min.), lavar con la disolución amortiguadora de medio mínimo resuspendiendo el paquete celular con 100 mL de disolución amortiguadora, centrifugar de nuevo. Lavar nuevamente con disolución amortiguadora.

**III.** Resuspender el paquete celular en ¼ del volumen inicial (250 mL) con disolución amortiguadora de medio mínimo sin fuentes de carbono y nitrógeno. Dejar incubando por 1 hora a 37 °C con agitación, para que el sistema de expresión recupere el crecimiento y produzca metabolitos sin enriquecimiento isotópico.

**IV.** Después de la hora de incubación centrifugar a 3000 x g. Resuspender el paquete celular nuevamente en ¼ del volumen inicial de medio mínimo con los nutrientes señalados, si el marcaje es deseado, se añade los reactivos enriquecidos isotópicamente.

**V.** Inducir la expresión con la adición de IPTG hasta una concentración de 0.5 mM en el cultivo (para el monitoreo de la expresión **NO** usar reactivos enriquecidos).

**VI.** Incubar a la temperatura óptima hasta que la producción de proteína sea máxima (es necesario el monitoreo de la expresión para determinar las condiciones óptimas de producción: tiempo y temperatura de inducción)

**VII.** Transcurrido el tiempo óptimo de máxima expresión, centrifugar (5500 x g por 20 minutos) para obtener el paquete celular. Si se comienza con el proceso de purificación de la proteína de fusión se resuspende el paquete celular con el buffer que se usará para la lisis, de no ser así, el paquete celular debe congelarse a -20°C para la posterior purificación de la proteína de fusión.

En todo momento, el medio de cultivo, así como la disolución amortiguadora deben contar con la cantidad de antibióticos adecuados para la cepa de expresión y el plásmido con el cual fue transformada.

### **EXPRESIÓN CON MEDIO PG**

#### **PREPARACIÓN DE LA DISOLUCIÓN AMORTIGUADORA**

Para llevarlo a cabo el enriquecimiento isotópico, como primer paso se prepara un litro de disolución amortiguadora con 50 mM  $\text{Na}_2\text{HPO}_4$  y 50 mM  $\text{KH}_2\text{PO}_4$ . Este litro se divide en 2 partes iguales y se coloca cada una de ellas en un matraz para cultivo de 2 L. Estos serán los **matraces preparativos** donde se expresará la proteína de interés. Es necesario preparar 100 mL adicionales de la disolución amortiguadora anterior para colocarla en un matraz de 250 o 500 mL. Esta disolución es la de los **cultivos semilla**. Estas disoluciones se esterilizan por autoclave.

#### **MEZCLA DE NUTRIENTES PREPARATIVOS (PARA LOS DOS MATRACES GRANDES; CON REACTIVOS ENRIQUECIDOS ISOTÓPICAMENTE O NO, DEPENDIENDO DEL EXPERIMENTO)**

Estas disoluciones contienen las fuentes de carbono y nitrógeno que servirán como materia prima a la célula para la producción de la proteína de interés. Dichas fuentes pueden ser o no contener nutrientes enriquecidos con  $^{15}\text{N}$  y/o  $^{13}\text{C}$ , dependiendo si se quiere enriquecer en uno o ambos isótopos o el cultivo es un ensayo de la expresión u optimización de las condiciones para expresar proteína. Dichas disoluciones se preparan como a continuación se describe:

1. Pesar el equivalente a 0.025 moles de  $(\text{NH}_4)_2\text{SO}_4$  en un vaso de precipitado de 50 mL conteniendo un agitador magnético, añadir 1 mL del “stock de magnesio”, 1 mL del “stock” de vitaminas, 1 mL del “stock” de metales. Finalmente añadir agua hasta alcanzar el volumen de 15 mL. Agitar hasta que el  $(\text{NH}_4)_2\text{SO}_4$  se haya disuelto.

2. Pesar 5 g de glucosa, añadir al vaso de precipitado y añadir el volumen de agua restante para alcanzar un volumen de 40 mL, agitar hasta que la glucosa se haya disuelto.

3. Esterilizar por filtración, guardar a 4° C (si será utilizado al siguiente día) o congelar.

4. No añadir a los matraces hasta que sea tiempo de inocularlos con el cultivo semilla.

**MEZCLA DE NUTRIENTES SEMILLA (PARA LOS CULTIVOS SEMILLA; SIEMPRE SIN REACTIVOS ENRIQUECIDOS)**

Al igual que la mezcla nutritiva preparativa, esta mezcla contiene las fuentes de carbono y nitrógeno que servirán como materia prima para que la célula produzca la proteína de nuestro interés. La disolución se prepara de forma idéntica con la excepción que esta mezcla jamás llevará reactivos enriquecidos con  $^{15}\text{N}$  y/o  $^{13}\text{C}$ . La cantidad preparada de mezcla de nutrientes semilla es suficiente para 10 cultivos semilla. No se debe añadir a los matraces hasta que sea tiempo de inocularlos con las células transformadas.

**PROTOCOLO PG**

*Primer día*

1. Se selecciona una colonia bien aislada de una transformación fresca o una hojuela de un stock de glicerol. Se inoculan 10 mL de medio LB. Se crece el cultivo mediante la incubación de las células por 7-10 horas a 37 °C con agitación.

2. Mientras el cultivo del paso 1 crece, se preparan los matraces semilla y preparativos, sin añadir los nutrientes.

3. Completado el paso 1 se centrifugan las células a 3'000 x g por 5-15 minutos, se decanta el sobrenadante.

4. Se deben añadir los antibióticos apropiados y la mezcla de nutrientes al (los) matraz (ces) semilla.

5. El paquete celular obtenido en el paso 3 se resuspende en un poco de medio del matraz semilla a inocular, después se inocula el resto del medio con las células resuspendidas.

6. El cultivo se deja crecer por 12-16 horas (toda una noche) incubando a 30 o 37 °C con agitación constante.

*Segundo día*

7. Al siguiente día se deben poner 45 mL del cultivo dentro de cada uno de dos tubos falcon de 50 mL estériles para recolectar las células centrifugando los tubos falcon a  $3'000 \times g$  por 15 minutos, después se decanta el sobrenadante.

8. Mientras se completa el paso 7 se añaden los antibióticos adecuados y la mezcla nutritiva a los matraces preparativos.

9. Completado el paso 7 se resuspende el paquete celular de uno de los tubos falcon en un poco de volumen del medio preparativo, una vez resuspendido el paquete celular, se inocula el resto del medio de uno de los matraces preparativos. Repetir con el otro falcon y el otro matraz preparativo.

10. Las células crecen incubando el cultivo a  $37\text{ }^{\circ}\text{C}$  con agitación a 200 rpm hasta que la densidad óptica (DO) a 600 nm sea de 0.7 - 0.8 (usualmente 1 - 2 h).

11. Inmediatamente se induce la expresión añadiendo IPTG hasta alcanzar una concentración 0.5 mM en el medio de cultivo. Se incuba a la temperatura optima de expresión y se deja transcurrir el tiempo hasta que la expresión de proteína sea máxima (se requiere hacer un monitoreo de expresión para fijar estas condiciones).

### *Segundo/Tercer día*

12. Terminado el tiempo de expresión óptimo, se recolectan las células del cultivo centrifugándolo a  $5500 \times g$  por 20 minutos.

13. Se decanta el medio y el paquete celular debe congelarse a  $-20\text{ }^{\circ}\text{C}$  o proceder con la lisis celular.

**APÉNDICE D. DISOLUCIONES AMORTIGUADORAS**

**Disoluciones amortiguadoras utilizadas para la lisis celular y la purificación por columna de afinidad metal-quelato**

Componente	Disolución amortiguadora		
	A	B	C
Fosfatos	50 mM		
NaCl	300 mM		
Imidazol	10 mM	50 mM	300 mM
pH	8.0		

**Disoluciones amortiguadoras utilizadas para realizar el corte de la proteína de fusión y la purificación de los fragmentos de corte.**

Componente	Disolución amortiguadora	
	D	E
Fosfatos	50 mM	
NaCl	100 mM	500 mM
pH	8.0	

**Disolución amortiguadora de carga para SDS-PAGE (2X)**

- Agua 4.2 mL
- Tris-HCl 0.5M pH= 6.8 1.0 mL
- Glicerol 800 µL
- SDS al 10% ( w/v) 1.6 mL
- β-mercaptoetanol 400 µL
- Azul de bromofenol al 10% (w/v) 20 µL

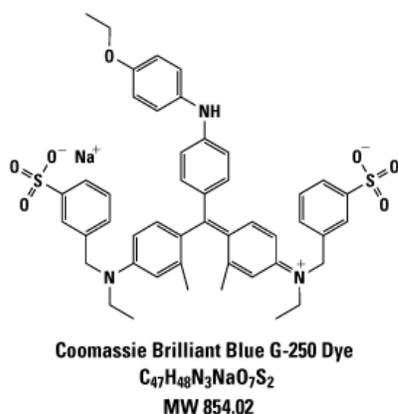
**APÉNDICE E. PROTOCOLO PARA PURIFICAR LA PROTEÍNA DE FUSIÓN  
POR LA COLUMNA DE AFINIDAD METAL- QUELATO**

- I. La columna Hi-Trap Chelating® debe encontrarse debidamente cargada con níquel divalente.
- II. Equilibrar la columna con 5-10 volúmenes de columna (existen columnas de 1 mL y de 5 mL) con disolución amortiguadora con la cual se deja fija la proteína de fusión a la columna (disolución amortiguadora A).
- III. Se hace pasar la parte soluble del lisado celular previamente filtrada en poro de 0.22  $\mu\text{m}$  considerando que la capacidad máxima de fijación de proteína por la columna es de alrededor de 12 mg de proteína/mL.
- IV. Se lava la columna con 10 volúmenes de columna de disolución amortiguadora de fijación. Los primeros dos volúmenes de lavado con esta disolución se colectaron en una sola fracción denominada L1. Los 8 volúmenes restantes se colectaron como una fracción diferente y se denominó L2, esto con la finalidad de monitorear la fijación de la proteína de fusión a la columna, así como de otras proteínas presentes en el lisado celular.
- V. Se realiza un lavado de 5 volúmenes de columna con una disolución que contenga mayor cantidad de imidazol que la disolución de fijación. En este caso se hizo con disolución amortiguadora B. Se colectó esta fracción y se denominó L3.
- VI. Se eluye la proteína de fusión con 3 volúmenes de columna de disolución amortiguadora C (300 mM de imidazol). La fracción colectada se denomina E1.
- VII. Finalmente se pasan 10 volúmenes más de disolución de fijación para retirar de la columna cualquier proteína que aún pudiera estar presente.
- VIII. Se hacen pasar de 5-10 volúmenes de columna de H<sub>2</sub>O destilada.
- IX. Se pasan 5 volúmenes de columna de etanol al 20 % y en estas condiciones se guarda la columna.
- X. Se toman muestras de las diferentes fracciones de purificación y se analizan por SDS-PAGE.

**APENDICE F. CUANTIFICACIÓN DE PROTEÍNAS POR EL MÉTODO DE BRADFORD** <sup>(44)</sup> <sup>(45)</sup> <sup>(46)</sup>

El Coomassie G-250 es un reactivo colorimétrico que se utiliza para la detección y cuantificación de proteínas totales. Fue descrito por primera vez por el Dr. Marion Bradford in 1976.

El ensayo de Bradford se basa en la unión del colorante azul de Coomasie G-250 a las proteínas. El ensayo se realiza a pH ácido. En estas condiciones existen tres formas



Estructura química del colorante Coomassie. Formulado en un amortiguador de ácido fosfórico a pH bajo. Esta forma es básica para el ensayo de proteínas .

cargadas del colorante. Las formas catiónicas roja y verde tienen su máximo de absorbancia a 470 nm y 650 nm respectivamente. La forma aniónica azul es la que se une a las proteínas y tiene una absorbancia máxima a 590 nm. El complejo colorante – proteína absorbe luz a 594 nm, por lo cual es común determinar la cantidad de proteína haciendo la medición de la absorbancia a 595 nm.

En la unión del colorante a la proteína participan fuerzas de Van der Waals e interacciones hidrofóbicas. El número de ligandos que se unen a

cada molécula de proteína es aproximadamente proporcional al número de cargas positivas presentes en la proteína; por lo tanto, el colorante se une más fácilmente a los aminoácidos básicos como lisina, arginina e histidina presentes en la proteína. Esto puede conducir a una variación en la respuesta a diferentes proteínas en el ensayo.

Aminoácidos libres, péptidos y proteínas de bajo peso molecular no producen respuesta a este ensayo. En general la masa de la proteína debe ser igual o mayor a 3KDa.

La cuantificación de la proteína de fusión GB1- Amidasa 2C se realizó utilizando el reactivo comercial de Thermo Scientific Coomasie Plus (Bradford) Protein Assay (con número de producto 23238) el cual contiene aditivos que retardan la formación de agregados colorante – colorante y colorante - proteína que tienden a formarse comúnmente en todos los reactivos a base de Coomasie.

PROTOCOLO PARA LA CUANTIFICACIÓN DE PROTEÍNA. <sup>(47)</sup>

1. Se preparan los estándares de BSA (albumina sérica bovina) que se utilizarán para construir una curva estándar con las siguientes concentraciones:

Estándar	Concentración de BSA
A	2000 $\mu\text{g/ mL}$
B	1500 $\mu\text{g/ mL}$
C	1000 $\mu\text{g/ mL}$
D	750 $\mu\text{g/ mL}$
E	500 $\mu\text{g/ mL}$
F	250 $\mu\text{g/ mL}$
G	125 $\mu\text{g/ mL}$
H	25 $\mu\text{g/ mL}$

El rango de lineal de trabajo con BSA es de 125-1000  $\mu\text{g/ mL}$

- Colocar 5 $\mu\text{L}$  de cada estándar de BSA y de las muestras de concentración desconocida en tubos de microfuga de 2mL. Identificar adecuadamente los tubos. Cada muestra desconocida debe hacerse por triplicado.
- Adicionar a cada tubo 1mL del reactivo de Coomasie y mezclar bien.
- Para tener mayor consistencia en los resultados incubar las muestra durante 10 min. a temperatura ambiente.
- Con una longitud de onda de 595 nm ajustar el espectrofotómetro en un valor de absorbancia igual a cero con 1 mL de reactivo de Coomasie (Blanco). Subsecuentemente medir la absorbancia de todas las muestras y registrar la lectura.
- Construir una curva estándar con la medida de absorbancia obtenida para cada estándar contra la concentración de cada uno.
- Utilizar la ecuación de la curva estándar para determinar la concentración de proteína de cada muestra desconocida con en valor promedio de absorbancia obtenido para cada una.