



INSTITUTO POLITÉCNICO NACIONAL
CENTRO INTERDISCIPLINARIO DE CIENCIAS MARINAS



EVALUACIÓN DE ESTRATEGIAS
METODOLÓGICAS PARA LA EXPRESIÓN
RECOMBINANTE DE LA ADN POLIMERASA
(FAMILIA B) DE *Thermococcus gammatolerans*

TESIS

QUE PARA OBTENER EL GRADO DE
MAESTRÍA EN CIENCIAS EN MANEJO DE RECURSOS MARINOS

PRESENTA

VALERIA ALEXANDRA VILLEGAS SILVA

LA PAZ B.C.S., DICIEMBRE DEL 2018



INSTITUTO POLITÉCNICO NACIONAL
SECRETARIA DE INVESTIGACIÓN Y POSGRADO
ACTA DE REVISIÓN DE TESIS

En la Ciudad de La Paz, B.C.S., siendo las 12:00 horas del día 28 del mes de Noviembre del 2018 se reunieron los miembros de la Comisión Revisora de Tesis designada por el Colegio de Profesores de Estudios de Posgrado e Investigación de CICIMAR para examinar la tesis titulada:

“EVALUACIÓN DE ESTRATEGIAS METODOLÓGICAS PARA LA EXPRESIÓN RECOMBINANTE DE LA ADN POLIMERASA (FAMILIA B) DE Thermococcus gammatolerans”

Presentada por el alumno:

VILLEGAS

Apellido paterno

SILVA

materno

VALERIA ALEXANDRA

nombre(s)

Con registro: B 1 6 0 9 7 6

Aspirante de:

MAESTRIA EN CIENCIAS EN MANEJO DE RECURSOS MARINOS

Después de intercambiar opiniones los miembros de la Comisión manifestaron APROBAR LA DEFENSA DE LA TESIS, en virtud de que satisface los requisitos señalados por las disposiciones reglamentarias vigentes.

LA COMISION REVISORA

Directores de Tesis

[Signature of Dr. Cesar Salvador Cardona Félix]

DR. CESAR SALVADOR CARDONA FÉLIX
Director de Tesis

[Signature of Dra. Gracia Alicia Gomez Anduro]

DRA. GRACIA ALICIA GOMEZ ANDURO
2ª. Directora de Tesis

[Signature of Dra. Claudia Judith Hernández Guerrero]

DRA. CLAUDIA JUDITH HERNÁNDEZ GUERRERO

[Signature of Dr. Mauricio Muñoz Ochoa]

DR. MAURICIO MUÑOZ OCHOA

[Signature of Dra. Bárbara González Acosta]

DRA. BÁRBARA GONZÁLEZ ACOSTA

PRESIDENTE DEL COLEGIO DE PROFESORES

[Signature of Dr. Sergio Hernández Trujillo]

DR. SERGIO HERNÁNDEZ TRUJILLO



I.P.N.
CICIMAR
DIRECCIÓN



**INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA DE INVESTIGACIÓN Y POSGRADO**

CARTA CESIÓN DE DERECHOS

En la Ciudad de La Paz, B.C.S., el día 05 del mes de Diciembre del año 2018
El (la) que suscribe BM. VALERIA ALEXANDRA VILLEGAS SILVA Alumno (a) del Programa
MAESTRÍA EN CIENCIAS EN MANEJO DE RECURSOS MARINOS
con número de registro B160976 adscrito al CENTRO INTERDISCIPLINARIO DE CIENCIAS MARINAS
manifiesta que es autor(a) intelectual del presente trabajo de tesis, bajo la dirección de:
DR. CÉSAR SALVADOR CARDONA FÉLIX y DRA. GRACIA ALICIA GÓMEZ ANDURO
y cede los derechos del trabajo titulado:
"EVALUACIÓN DE ESTRATEGIAS METODOLÓGICAS PARA LA EXPRESIÓN RECOMBINANTE
DE LA ADN POLIMERASA (FAMILIA B) DE *Thermococcus gammatolerans*"
al Instituto Politécnico Nacional, para su difusión con fines académicos y de investigación.

Los usuarios de la información no deben reproducir el contenido textual, gráficas o datos del trabajo sin el permiso expreso del autor y/o director del trabajo. Éste, puede ser obtenido escribiendo a la siguiente dirección: vavs.villegas@gmail.com - ccardona.felix@gmail.com

Si el permiso se otorga, el usuario deberá dar el agradecimiento correspondiente y citar la fuente del mismo.


BM. VALERIA ALEXANDRA VILLEGAS SILVA

Nombre y firma del alumno

El presente trabajo de investigación se desarrolló en el laboratorio de Microbiología y Biología Molecular del Centro Interdisciplinario de Ciencias Marinas del Instituto Politécnico Nacional (CICIMAR-IPN) bajo la dirección del Dr. César Salvador Cardona Félix, Catedrático CONACyT comisionado al referido Centro de Investigación y de la Dra. Gracia Alicia Gómez Anduro, Profesora-Investigadora del Centro de Investigaciones Biológicas del Noroeste, con la asesoría de la Dra. Claudia Judith Hernández Guerrero, Dra. Bárbara González Acosta y Dr. Mauricio Muñoz Ochoa, todos Profesores-Investigadores del CICIMAR-IPN, en el período comprendido de agosto de 2016 a diciembre de 2018.

A mi mejor guía y maestro de vida.

T.A.S.M.M.

“...Antes de partir me gustaría tener la certeza de que sabes que todos los días de su vida este señor, tu padre, ha pensado en ti en cada amanecer, en cada anochecer, en los días más bellos y en los más nublados o lluviosos. Siempre visualizándote sana, alegre, positiva, triunfadora, con una energía desbordante y contagiosa, noble y fuerte como un roble y sobre todo creyendo con toda mi fe en ti, en tus metas en tus proyectos...”

Fragmento de una carta de un padre a su hija.

R.V.V (1960-2017)

AGRADECIMIENTOS

Al CONACyT por el financiamiento al Proyecto Número 247842 titulado: “*Endolisinas fágicas recombinante: Nueva estrategia para el control de bacterias patógenas en cultivos de importancia acuícola en México*”, de la convocatoria de proyectos de desarrollo científico para atender problemas nacionales 2014, así como por la beca para estudios de Maestría 612100.

Quiero agradecerles a mis directores de tesis: el Dr. César Salvador Cardona Félix y la Dra. Gracia Alicia Gómez Anduro por su apoyo y confianza en la realización del presente trabajo, incluso en los momentos más complicados. Así mismo, por su orientación y guía en laboratorio durante las pruebas realizadas.

A la Dra. Isabel Gómez Gómez del Instituto de Biotecnología de la UNAM por su apoyo en los análisis de Western-Blot.

A todos los miembros del comité por sus valiosas aportaciones y comentarios que ayudaron a dar un mejor enfoque al trabajo, así como por el tiempo y dedicación a la hora de aclarar dudas.

Al Dr. Mauricio Muñoz Ochoa, por su apoyo incondicional a través de los años. La paciencia, los consejos, la confianza y motivación en los días más difíciles, gracias por no dejar de creer en mí. De igual manera a la Dra. Claudia Judith Hernández Guerrero, gracias por su atención, comprensión y apoyo a lo largo del trabajo de tesis, el proyecto BEIFI SIP 20170434, la inscripción al quinto semestre, pero sobre todo por las oportunidades, la experiencia y la confianza de permitirme trabajar por proyecto a su lado. A la Dra. Bárbara González Acosta por su apoyo en tópicos de biología molecular, así como sus constantes observaciones y exigencias, que de manera directa me impulsaron a ser una mejor estudiante.

A mi “fan #1”, mi padre, por su eterno apoyo en todas mis decisiones (incluso las más cuestionables), su interés en todos y cada uno de mis proyectos, pero principalmente por todo el amor, los consejos y sus enseñanzas. Y es que, a pesar de la ausencia física, su cariño ha sido el motor principal que impulsa el avance de mi vida académica. Te Amo Siempre Mucho Más.

A mi familia, mamá, hermanos y amigos, quienes me han acompañado siempre con una sonrisa a lo largo de este proceso, haciendo de este tiempo algo más ameno, gracias por ser y estar. A Ariel, por su motivación, apoyo y paciencia especialmente en los momentos de frustración académica. De la manera más sincera, y cursi: Gracias amor.

A mis compañeros y amigos de laboratorio de Microbiología y Biología Molecular, particularmente Karla y Angélica por los momentos y las risas, porque los días en CICIMAR no habrían sido lo mismo sin ustedes y porque “hoy será un gran día”. Gracias por escucharme y ayudarme a replantear ideas. A Karen por su apoyo en laboratorio, así como la capacidad de adaptación a las noches de desvelo; y Antony por la ayuda en laboratorio.

A mis ex compañeros, amigos y nuevas amistades del laboratorio de Química de Algas Marinas, gracias por las anécdotas y las escapadas de des-estrés. Porque no importa el tiempo, el lugar, ni la circunstancia, el cariño es sincero y el apoyo incondicional.

Finalmente todas las personas que estuvieron de la mano conmigo a lo largo de los años: ¡Arriba y adelante!. Gracias.

ÍNDICE GENERAL

ÍNDICE DE FIGURAS.....	IX
ÍNDICE DE TABLAS.....	X
ABREVIATURAS.....	XI
GLOSARIO.....	XIII
RESUMEN.....	XV
ABSTRACT.....	XVI
1. INTRODUCCIÓN.....	1
1.1 Proteínas recombinantes.....	1
1.2 ADN polimerasas.....	1
1.2.1 Características de las ADN polimerasas de la familia B.....	4
1.2.2 Adaptación enzimática de ADN polimerasas de la familia B de extremófilos.....	4
1.3 <i>Thermococcus gammatolerans</i>	5
2. ANTECEDENTES.....	8
2.1 Producción de proteínas recombinantes.....	8
2.1.1 Cepa de expresión: <i>Escherichia coli</i>	8
2.1.2 Vector de expresión.....	8
2.2 Complicaciones en la expresión de proteínas recombinantes en <i>E. coli</i>	10
2.3 Trabajos previos en la expresión de ADN polimerasas de arqueas.....	11
3. JUSTIFICACIÓN.....	13
4. HIPÓTESIS.....	14
5. OBJETIVO GENERAL.....	14
6. OBJETIVOS ESPECÍFICOS.....	14
7. METODOLOGÍA.....	15
7.1 Obtención de construcciones.....	15
7.1.1 Diseño de oligonucleótidos para el gen que codifica TgPolB.....	15
7.1.2 Amplificación del gen a partir del ADN genómico.....	15
7.1.3 Clonación del gen que codifica TgPolB en pJET1.2 BluntCloning Vector.....	16
7.1.4 Subclonación en vectores de expresión.....	17
7.1.4.1 Confirmación de construcciones por medio de secuenciación del gen codificante de TgPolB.....	17
7.2 Expresión de TgPolB.....	18
7.2.1 Transformación en las cepas de expresión BL21 y Rosetta II (Ril).....	18
7.2.2 Cinética de inducción de TgPolB.....	18

7.2.3 Inducción	18
7.2.4 Relación medio de cultivo y expresión	19
7.3 Purificación de la enzima TgPolB.....	20
7.3.1 Purificación (Regulador con EDTA).....	20
7.3.2 Purificación (Regulador sin EDTA).....	20
7.3.3 Purificación con detergentes.....	21
7.4 Cuantificación de proteína	21
7.5 Análisis de actividad de TgPolB por medio de PCR.....	22
7.6 Análisis de Western-Blot	23
7.7 Subclonación en pET32a	23
7.8 Análisis bioinformáticos de TgPolB	24
7.8.1 Modelado estructural <i>in silico</i>	24
7.8.2 Análisis de identidad de secuencias.....	24
7.8.3 Ubicación de N-terminal.....	24
7.8.4 Sitios Activos	25
7.8.5 Índice de adaptación de codones.....	25
7.8.6 Enlaces disulfuro.....	25
8. RESULTADOS Y DISCUSIÓN.....	26
8.1 Obtención de construcciones.....	26
8.1.1 Diseño de oligonucleótidos para el gen que codifica TgPolB	26
8.1.2 Amplificación y clonación del gen que codifica TgPolB en pJET1.2 BluntCloning Vector	26
8.1.3 Subclonación en vectores de expresión.....	26
8.1.3.1 Confirmación de construcciones por medio de secuenciación.....	28
8.2 Expresión de TgPolB	28
8.2.1 Cinética de inducción de TgPolB en BL21(DE3)	28
8.2.2 Inducción	28
8.2.3 Relación del medio de cultivo y la expresión.....	31
8.3. Purificación	33
8.3.1 Purificación (Regulador con EDTA).....	34
8.3.2 Purificación (Regulador sin EDTA).....	34
8.3.3 Purificación con detergentes.....	36
8.4 Cuantificación de proteína	38
8.5 Análisis de actividad de TgPolB por medio de PCR.....	38

8.6	Análisis de Western-Blot	40
8.7	Análisis bioinformático de TgPolB.....	41
8.7.1	Modelo estructural <i>in silico</i>	41
8.7.2.	Análisis de identidad de secuencias.....	42
8.7.3	Ubicación del N-terminal.....	44
8.7.4	Sitios Activos	45
8.7.5	Índice de adaptación de codones.....	46
8.7.6	Enlaces disulfuro.....	47
8.8	Subclonación en pET32a	48
9.	CONCLUSIONES Y RECOMENDACIONES	51
11.	BIBLIOGRAFÍA.....	53
Anexo I.	Extracción del ADN plasmídico	60
Anexo II.	Electroforesis de proteínas en geles de poliacrilamida (SDS-PAGE)	61
Anexo III.	Composición de medios de cultivo	63
Anexo IV.	Preparación de soluciones y reguladores	64
Anexo V.	Secuenciación del inserto en los vectores de expresión pET19b y pColdI.....	68

ÍNDICE DE FIGURAS

Figura 1. Distribución de los diferentes tipos de ADN polimerasas en los tres dominios de la vida.....	3
Figura 2. Ubicación de los diferentes dominios.....	3
Figura 3. Clonación de TgPolB en pJET 1.2 Blunt Cloning Vector.....	27
Figura 4. Cinética de inducción de TgPolB.....	29
Figura 5. Inducción de TgPolB en cepas BL21 y Rosetta II (RII).....	30
Figura 6. Inducción TgPolB con diferentes medios de cultivo.....	32
Figura 7. Purificación de TgPolB.....	35
Figura 8. Purificaciones con detergentes.....	37
Figura 9. Cuantificación de proteínas en gel.....	38
Figura 10. Análisis de actividad enzimática de TgPolB.....	39
Figura 11. Comparación de secuencias de ADN polimerasas comerciales obtenidas a partir de <i>Thermococcales</i> y TgPolB.....	41
Figura 12. Análisis in silico de TgPolB.....	42
Figura 13. Comparación de modelos estructurales y dominios en ADN pol B de <i>Thermococcales</i>	43
Figura 14. Ubicaciones de la región N-terminal en los modelos estructurales de ADN pol B.....	44
Figura 15. Sitios activos de TgPolB.....	45
Figura 16. Análisis predictivo de enlaces disulfuro y predictor de conectividad elaborado en el servidor en línea DISULFIND.....	49
Figura 17. Ensayo de doble restricción de la construcción 4 en pColdI utilizando las enzimas de restricción BamHI y NdeI.....	50
Figura 18. Mutación en sitio de restricción.....	50

ÍNDICE DE TABLAS

Tabla I. Ejemplos de proteínas recombinantes y sus diferentes aplicaciones.	2
Tabla II. Principales estudios realizados con <i>T. gammatolerans</i> desde su descubrimiento a la actualidad.	6
Tabla III. Características principales y secuencia de aminoácidos de la ADN polimerasa (Familia B) de <i>Thermococcus gammatolerans</i> .	7
Tabla IV. Cepas de <i>E. coli</i> comúnmente utilizadas para la expresión de ADN polimerasas (Familia B) de arqueas.	9
Tabla V. Ejemplos de sistemas y vectores utilizados en la expresión de ADN polimerasas (Familia B) de arqueas.	10
Tabla VI. Comparación de las características de ADN polimerasas obtenidas de arqueas, utilizadas comercialmente en PCR.	12
Tabla VII. Mezcla de reacción utilizada para la amplificación del gen codificante de TgPolB.	15
Tabla VIII. Programa de temperaturas y tiempo para amplificación del gen codificante de TgPolB.	16
Tabla IX. Componentes de la reacción para el ensayo de doble restricción con las enzimas BamHI y NdeI.	17
Tabla X. Condiciones de inducción seleccionadas para las pruebas de purificación.	19
Tabla XI. Reguladores utilizados para el análisis de actividad enzimática.	23
Tabla XII. Programa control, utilizado para la reacción de PCR a 72°C y 68°C.	23
Tabla XIII. Concentración de ADN plasmídico de las construcciones de trabajo.	26
Tabla XIV. Porcentaje de identidad entre algunas ADN polimerasas (Familia B) de <i>Thermococcales</i> utilizadas en la actualidad de manera comercial y ADN polimerasa (Familia B) de <i>Thermococcus gammatolerans</i> (TgPolB).	42
Tabla XV. Análisis del uso preferencial de codones de TgPolB en <i>E. coli</i> .	46

ABREVIATURAS

ADN	Ácido desoxirribonucleico
ADN pol	ADN polimerasas
ADN pol B	ADN polimerasas de la familia B
AE	Adaptación enzimática
ATP	Adenosín trifosfato
EDTA	Ácido etilendiaminotetraacético
ER	Enzima Recombinante
ERO	Especies reactivas de oxígeno
DL50	Dosis letal media
DMSO	Dimetilsulfóxido
dNTP's	Deoxinucleósido trifosfato
FPLC	Fast Protein Liquid Chromatography/ Cromatografía líquida de proteínas a alta velocidad
Gy	Gray
IAC	Índice de adaptación de codones
IPTG	Isopropil- β -D-1 tiogalactopiranosido
kDa	Kilo daltons
LB	Luria Bertani
MPM	Marcador de peso molecular
nm	Nanómetros
pb	Pares de bases

PCR	´Polymerase Chain Reaction/ Reacción en Cadena de la Polimerasa
Pfu	ADN polimerasa comercial de <i>Pyrococcus furiosus</i>
PGBD	ADN polimerasa comercial Deep Vent
PMSF	Fluoruro de fenilmetilsulfonio
PR	Proteína recombinante
p/v	Peso sobre volumen
RI	Radiación ionizante
rpm	Revoluciones por minuto
SB	Super broth
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis/ Electroforesis en gel de poliacrilamida con dodecil sulfato sódico
TBE	Tris/Ácido Bórico/EDTA
Tgo	ADN polimerasa de <i>Thermococcus gorgonarius</i>
TgPolB	ADN polimerasa de <i>Thermococcus gammatolerans</i>
UNAM	Universidad Autónoma de México
UUSMB	Unidad Universitaria de Secuenciación Masiva y Bioinformática
VC	Volúmenes de columna
WB	Western-Blot
WGA	Whole Genome Amplification/Amplificación de genomas completos
2XYT	Yeast Extract Tryptone/Extracto de levadura Triptona

GLOSARIO

Actividad 3'-5' exonucleasa. Actividad que permite a la polimerasa corregir sus propios errores en sentido contrario a la síntesis de cadena (Sánchez, 2012).

Dominios. Regiones de aminoácidos con disposición común o función específica (Pierce, 2009).

Enzima. Proteínas que catalizan reacciones químicas (Velázquez Martínez, 2012).

Exonucleasa. Enzima que hidroliza los enlaces fosfodiéster de los nucleótidos a partir del extremo terminal (3' o 5') (Mukherjee *et al.*, 2004).

In silico. Uso de software informáticos para realizar simulaciones biológicas empleando modelos matemáticos (Fina *et al.*, 2013).

In vitro. Su traducción del latín significa “en vidrio”, es un término utilizado para describir un sistema de cultivo que involucra medio líquido o sólido en un contenedor de vidrio (Desmond, 1990).

Mutagénesis. Es un método propuesto por Michael Smith, que permite adaptar proteínas para ciertas aplicaciones alterando su secuencia de aminoácidos en sitios específicos (Voet & Voet, 2006).

Parálogos. Se le denomina así a los genes que teniendo un mismo origen y posterior a un proceso de duplicación, uno de ellos se adapta para una nueva función (Pierce, 2009).

Proteína. Polímeros lineales constituidos por más de 50 aminoácidos distintos, unidos por enlaces covalentes (McKee & McKee, 2014).

Punto isoeléctrico. Se le denomina así al pH en el que la carga eléctrica neta de una sustancia es igual a cero (Fuentes-Arderiu, 1997).

Tasa de error. Frecuencia de mutación por par de base por evento de replicación (Adams & Kelly, 2001)

Tasa de extensión. Velocidad en la que se incorporan los nucleótidos por molécula de ADN polimerasa. Puede expresarse en la cantidad de pares de bases incorporadas por el tiempo (minutos o segundos) (Zhang *et al.*, 2015)

Termoestabilidad. Capacidad de la proteína para mantener su estructura, función y las propiedades químicas de la cadena polipeptídica ante las fluctuaciones térmicas (Kumwenda *et al.*, 2013).

RESUMEN

La capacidad de *Thermococcus gammatolerans* para sobrellevar los daños provocados por las condiciones extremas en las que habita (elevadas temperaturas, niveles de presión, radiación, etc.), hacen de este microorganismo, un buen modelo de estudio para el análisis de sus enzimas. Por lo anterior, se seleccionó la ADN polimerasa B de *T. gammatolerans* (TgPolB) como modelo de trabajo para el análisis de su expresión. Dada la gran similitud estructural y el alto valor de identidad que comparte con otras ADN pol B, se buscó obtener resultados similares, bajo un sistema de expresión semejante al utilizado en otras enzimas homólogas. Para esto, se realizó una exhaustiva revisión bibliográfica, utilizando como referencia la información previamente publicada para modelos similares. El primer paso fue la amplificación del gen de TgPolB a partir del ADN genómico de *T. gammatolerans*, el cual fue clonado en el vector pJET 1.2 Bluntcloning, y subclonado en los vectores de expresión pColdI y pET19b. Estas construcciones fueron confirmadas por medio de PCR, ensayo de doble restricción y secuenciación. Los resultados obtenidos a partir de la expresión en las cepas de Rosetta II y BL21 (DE3) con diferentes condiciones de inducción, no mostraron expresión con ninguna de las construcciones analizadas independientemente del medio de cultivo (Luria Bertani, Super broth, 2XYT), temperatura (35°C y 16°C), tiempo de inducción (4 y 16 hrs) o concentraciones de inductor utilizados (0.25, 0.5, 1 y 1.25 mM). Por su parte, las pruebas de purificación no permitieron la obtención de TgPolB, por lo que es probable que presente cambios estructurales que impidan su identificación por el método de Western-Blot. Finalmente se propone el uso de 4 estrategias de trabajo derivadas de los resultados obtenidos: 1) uso de cepas de expresión codón-plus; 2) optimización del gen; 3) subclonación preferentemente simultánea en diferentes vectores de expresión; y 4) cambiar el sistema de expresión.

Palabras clave:

Proteínas recombinantes, arqueas, *Thermococcus gammatolerans*, ADN polimerasas de la familia B.

ABSTRACT

Thermococcus gammatolerans capability to survive the damage caused by the extreme conditions in which it lives (high temperatures, pressure levels, radiation, etc.), makes this microorganism a good study model for the analysis of its enzymes. Therefore, the DNA polymerase B from family B of *T. gammatolerans* (TgPolB), was selected as a working model for the analysis of its expression. Given the great structural similarity and the high identity value that this enzyme shares with other DNA polymerases from family B, we sought to obtain similar results by inducing the protein expression under an expression system similar to that used in other homologous enzymes. To achieve this, an exhaustive bibliographical review was carried out, using as reference the previously published information for similar models. The TgpolB gene was amplified from the genomic DNA of *T. gammatolerans*, which was cloned into pJET1.2 Blunt Cloning Vector and subcloned into the expression vectors pCold and pET19b. These constructions were confirmed by PCR, double restriction and sequencing. The results obtained from the expression in the strains of Rosetta II and BL21 (DE3) with different induction conditions, demonstrate that the absence of expression is not due to the culture medium (Luria-Bertani, SuperBroth, 2XYT), temperature (35°C y 16°C), induction time (4, 16 hrs) or inductor concentrations used (0.25, 0.5, 1 y 1.25 mM). On the other hand, purification tests didn't allow to obtain TgPolB, , so it is likely to present structural changes that prevent its identification by the Western-Blot method. Finally, the use of four work strategies derived from the obtained results is proposed in order to achieve protein expression: 1) use of codon-plus expression strains 2) gene optimization 3) subcloning preferably simultaneous in different expression vectors 4) change the expression system.

Key words:

Recombinant proteins, archaeae, *Thermococcus gammatolerans*, DNA polymerases family B.

1. INTRODUCCIÓN

1.1 Proteínas recombinantes

Una de las aportaciones más relevantes de la biotecnología moderna es la producción de proteínas recombinantes (PR), las cuales son un producto de la ingeniería genética en donde regiones génicas son insertadas en vectores que permiten la expresión de la proteína de interés para su producción. Cada vez son más las aplicaciones que requieren cantidades elevadas de proteína, desde las técnicas utilizadas para su análisis y caracterización en laboratorio, hasta su comercialización en industrias como la alimenticia y farmacéutica. A escala comercial se producen cientos de PR, no obstante, su producción aún sigue representando un reto en muchos casos debido a las diferentes problemáticas que suelen presentarse durante la estandarización y optimización del proceso de producción (Palomares *et al.*, 2004).

1.2 ADN polimerasas

El uso de las PR se ha extendido a lo largo de múltiples áreas de investigación con diversas aplicaciones (**Tabla I**), siendo una de ellas la obtención de enzimas recombinantes (ER) (Palomares *et al.*, 2004). Una de las ER más utilizadas como herramienta en técnicas de biología molecular y biotecnología para la manipulación *in vitro* del ADN, son las ADN polimerasas (ADN pol). Las ADN pol son enzimas ubicuas que sintetizan las hebras complementarias del ADN en todas las células vivas cumpliendo una función fundamental en el mantenimiento de la integridad del genoma durante los procesos de replicación y reparación del ADN (Gardner *et al.*, 2014; Ishino & Ishino, 2014).

Desde el descubrimiento y caracterización de la ADN pol I de *Escherichia coli* en la década de los cincuenta por Arthur Kornberg y colaboradores, se han aislado y caracterizado múltiples ADN pol a partir de diversos organismos (Kornberg *et al.*, 1956).

Tabla I. Ejemplos de proteínas recombinantes y sus diferentes aplicaciones.

Proteína Recombinante	Algunas aplicaciones	Área de investigación
ADN polimerasas	Procesos que requieren síntesis de ADN (PCR) y mutagénesis <i>in vitro</i>	Clonación genética
Quimosina	Manufacturación de queso	Industria alimenticia
Somatotropina bovina	Incrementa la producción de leche de vaca de alta calidad	
Insulina	Tratamiento de diabetes	Terapéutica
Hormona de crecimiento	Tratamiento de deficiencia de hormona de crecimiento	
Factor VIII	Tratamiento de hemofilia	
Proteasas	Remover suciedad, favorecer la penetración de ingredientes activos a la piel, ablandador de carne, alimentación de ganado	Varias
Lipasas	Manufactura de queso Roquefort, tratamiento de aguas ricas en lípidos, producción de cosméticos	Varias

*[Información obtenida de (Amid, 2015; Clark *et al.*, 2016; Nadeem *et al.*, 2018)].

Las ADN pol I se han clasificado de acuerdo a la comparación de las secuencias de aminoácidos que contienen, en siete diferentes familias: A, B, C, D, X, Y, y la transcriptasa inversa (**Figura 1**). Siendo las familias A y B (ambas presentes en bacterias, arqueas, eucariotas y virus/fagos) las más estudiadas y las únicas utilizadas como herramienta en ingeniería genética. Lo anterior, debido a la aplicación de sus propiedades en el desarrollo y perfeccionamiento de diversas técnicas biotecnológicas como los procesos de reacción en cadena de la polimerasa (PCR, por sus siglas en inglés), secuenciación, amplificación de genomas completos (WGA, por sus siglas en inglés), marcaje de ADN, mutagénesis y otros propósitos como biología sintética y diagnosis molecular (Ishino & Ishino, 2014; Zhang *et al.*, 2015).

	FAMILIA							
	A	B		C	D	E	X	Y
Bacteria	I ■	II ■		III ■				IV V ■ ■
Arquea		BI ■	B3 ■		D ■	E ■		Y ■
Eukarya	γ θ ■ ■	α δ ε ζ ■ ■ ■ ■					β λ μ σ ■ ■ ■ ■	η ι κ ■ ■ ■

Figura 1. Distribución de los diferentes tipos de ADN polimerasas en los tres dominios de la vida (Ishino y Ishino, 2014).

La habilidad fundamental de las ADN pol de sintetizar una cadena complementaria de ADN se mantiene en todas las enzimas de este tipo. No obstante, a pesar de presentar la misma función básica y una gran similitud estructural (dominios denominados “palma”, “pulgar” y “dedos”; **Figura 2**), las ADN pol difieren ampliamente entre sí en muchas de sus propiedades específicas como la procesividad, fidelidad (precisión de síntesis) y selectividad de nucleótidos (Beard *et al.*, 2003; Killelea *et al.*, 2014). Es por ello que, aunque se han obtenido múltiples ADN pol con el avance de las técnicas de ingeniería en proteínas, su estudio sigue siendo motivo de interés para los investigadores que buscan satisfacer la demanda de nuevas y diferentes ADN pol con propiedades dirigidas a propósitos específicos (Zhang *et al.*, 2015).

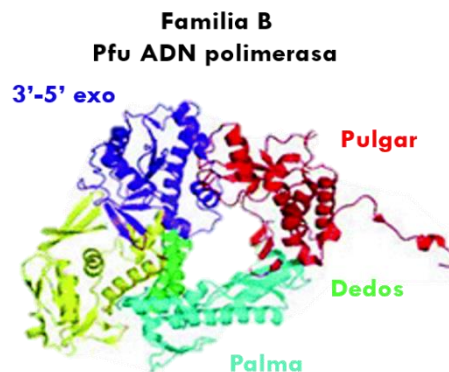


Figura 2. Ubicación de los diferentes dominios. dedos (verde), pulgar (rojo), palma (magenta) y exo (azul) en la ADN polimerasa (Familia B) de *Pyrococcus furiosus* (Pfu) (Ishino & Ishino, 2013).

1.2.1 Características de las ADN polimerasas de la familia B

Las ADN pol de la familia B (ADN pol B) son la familia de polimerasas más abundantes, con miembros en virus, bacterias, arqueas y eucariotas. Estas polimerasas en arqueas son representadas principalmente por enzimas replicativas que se encuentran en todos los linajes, llegando a presentar diferentes parálogos, algunos de los cuales se encuentran inactivos (Rogozin *et al.*,2008;Makarova *et al.*,2014). Además de la base principal de las ADN pol (dominio de la palma, dedos y pulgar), la mayoría de las ADN pol B contienen un dominio exonucleasa 3'-5' en el extremo N-terminal, así como un dominio de reconocimiento de uracilo (Rothwell *et al.*,2005;Makarova *et al.*,2014). Se caracterizan por presentar una muy alta fidelidad y notable estabilidad térmica. La alta fidelidad se debe a que las ADN pol B presentan esta actividad 3'-5' exonucleasa, la cual permite a la polimerasa corregir sus propios errores en sentido contrario a la síntesis de cadena. Esta actividad se realiza a través de la hidrólisis del nucleótido incorrecto desde el extremo 3' de la hebra en elongación antes de la incorporación del nucleótido siguiente (Sánchez, 2012).

1.2.2 Adaptación enzimática de ADN polimerasas de la familia B de extremófilos

Las arqueas son el principal grupo de microorganismos capaces de desarrollarse en condiciones ambientales consideradas extremas para los seres humanos (Bakermans, 2015). Estos microorganismos cuentan con mecanismos complejos de supervivencia para mantener sus componentes celulares estables y activos en condiciones como, altas temperaturas y elevados niveles de radiación, siendo uno de estos mecanismos la adaptación enzimática (AE). La AE se define como el cambio en las propiedades bioquímicas de ciertas enzimas en un organismo, inducido por la presencia de moléculas o condiciones específicas en el ambiente, presión selectiva y diferencias en la secuencia de aminoácidos. Lo anterior se traduce como variaciones en la estructura, flexibilidad, carga e hidrofobicidad de las enzimas. Estos cambios no siguen un patrón específico ya que

las proteínas de microorganismos extremófilos cuentan con una gran variabilidad de adaptaciones a condiciones físicas y químicas extremas.

Estudios comparativos de enzimas de mesófilos y termófilos han demostrado que en la mayoría de los casos, la secuencia de aminoácidos se conserva (entre 30-50%). No obstante, es posible que se presenten diferencias en su composición (Wintrode *et al.*, 2000; Reed *et al.*, 2013; Sarmiento *et al.*, 2015). Un ejemplo que describe esto, es el trabajo de Maras y colaboradores (1994) quienes compararon la enzima glutamato dehidrogenasa del termófilo *P. furiosus*, con versiones homologas obtenidas de organismos mesófilos, observando un mayor contenido de aminoácidos hidrofóbicos, una disminución en el contenido de aminoácidos polares y con cargas, así como disminución en los residuos de glicina.

1.3 *Thermococcus gammatolerans*

Son pocos los estudios realizados con *T. gammatolerans* (**Tabla II**), un arquea anaerobia, perteneciente al orden de los *Thermococcales* y al filo Euryarchaeota, que presenta un crecimiento óptimo a 88°C. Este microorganismo fue descubierto y aislado de una ventila hidrotermal en la cuenca de Guaymas (Golfo de California, México), a una profundidad de 2,600 m. Estudios realizados por Zivanovic y colaboradores en el 2009 han demostrado que *T. gammatolerans* es una especie capaz de soportar dosis de 5,000 Gy de radiación gamma sin presentar pérdida de viabilidad, y puede sobrevivir a dosis de hasta 30,000 Gy en cultivos enriquecidos (Zivanovic *et al.*, 2009). En otras palabras, esto quiere decir que es capaz de resistir dosis de radiación 12 veces mayor que la que resiste *Bacillus subtilis* (2,500 Gy), 142 veces mayor a *Escherichia coli* (210 Gy) y 6,000 veces mayor a la dosis letal media (DL50) reportada para humanos (3-5 Gy) (Powsner & Powsner, 2006; Moeller *et al.*, 2012). Lo anterior convierte a *T. gammatolerans* en el arquea más resistente a la radiación ionizante (RI) descrita en la actualidad (Zivanovic *et al.*, 2009).

Se sabe que la RI suele depositar energía en las biomoléculas de la célula durante su paso, lo cual tiene como consecuencia modificaciones en proteínas, ADN y básicamente cualquier macromolécula celular a través de la generación de

especies reactivas de oxígeno (ERO). En el caso del ADN, una fracción de los sitios de daño inducido por la RI puede provocar pérdida o modificación de bases, rupturas de cadena sencilla o doble y entrecruzamientos de tipo ADN-ADN o ADN-proteína. Lo anterior puede bloquear la replicación del genoma en caso de no ser reparado, resultando en una pérdida significativa de la información genética. Además de los daños producidos por la RI, el ADN también puede sufrir daños cuando está sujeto a altas temperaturas como las presentes en ambientes en los que las especies hipertermófilas habitan (Lomax *et al.*, 2013; Bakermans, 2015).

Tabla II. Principales estudios realizados con *T. gammatolerans* desde su descubrimiento a la actualidad.

Artículo	Referencia
<i>Thermococcus gammatolerans</i> a hyperthermophilic archaeon from a Deep-sea hydrothermal vent that resists ionizing radiation	(Jolivet <i>et al.</i> , 2003)
Recovery of ionizing-radiation damage after high doses of gamma ray in <i>T. gammatolerans</i>	(Tapias <i>et al.</i> , 2009)
Genome analysis and genome-wide proteomics of <i>T. gammatolerans</i>	(Zivanovic <i>et al.</i> , 2009)
Genome-Wide Transcriptional Response of the Archaeon <i>Thermococcus gammatolerans</i> to Cadmium	(Lagorce <i>et al.</i> , 2012)
Prioritizing targets for structural biology through the lens of proteomics: TGAM_1934	(Yang <i>et al.</i> , 2015)

La capacidad de *T. gammatolerans* para controlar este tipo de daños; dadas las condiciones en las que habita, ha hecho de este microorganismo un interesante modelo de estudio para el análisis de sus enzimas. Por esta razón, se seleccionó la ADN pol B de *T. gammatolerans* (TgPolB) como modelo de trabajo para el análisis de su expresión (**Tabla III**). Dada la gran similitud estructural y el alto valor de

identidad que comparte con otras ADN pol B, se busca producir una enzima de importancia biotecnológica obtenida a partir de germoplasma mexicano que permita desarrollos sin conflicto de patente, bajo un sistema de expresión semejante al utilizado en otras enzimas homólogas.

Tabla III. Características principales y secuencia de aminoácidos de la ADN polimerasa (Familia B) de *Thermococcus gammatolerans*.

Proteína: TgPolB	
Tamaño del gen (pb)	2,328
Peso molecular (kDa)	89.9
Punto isoeléctrico	8.66
Secuencia de aminoácidos	
MILDTDYITENGPVIRVFKKENGEFKIEYDRTFEPYFYALLKDDSAIEDVKKVTAKRHGAVVKV RAEKVQRKFLGRPIEVWKL YFTHPQDVP AIRDKIRAHPAVVDIYEYDIPFAKRYLIDKGLIPMEGD EELRMLAFDIETLYHEGEEFGTGPILMISYADENEARVITWKKIDLPYVDV VSTEKEMIKRFLRVVK EKDPDVLITYNGDNFDFAYLKKRCEKLGIFTLGRDGSEPKIQRMGDRFAVEVKGRIHFDLYPVIR RTINLPTYTLEAVYEAVFGKPKVKVYAAEITEAWESGEGLERVARYSMEDAKVTYELGREFFPM EAQLSRLIGQSLWDVSRSSSTGNLVEWFLLRKAYERNELAPNKPDERELARRRESYAGGYVKEP ERGLWDNIVYLDFRSLYPSIIITHNVSPDTLNREGCKEYDVAPEVGHKFKDFPGFIPSLLDLLE ERQKIKRKMKATIDPLEKKLLDYRQRAIKILANSFYGYGYAKARWYCKECAESVTAWGREYIET TIREIEEKFGFKVLYADTDGFFATIPGADAETVKKKAKEFLKYINAKLPGLLELEYEGFYVVRGFFVT KKKYAVIDEEGKITTRGLEIVRRDWSEIAKETQARVLEAILKHGDVEEAVRIVKEVTEKLSKYEVPP EKLVIHEQITRDLRDYKATGPHVAVAKRLAAKGVKIRPGTVISYIVLKGSGRIGDRAIPAEFDPTK HRYDAEYYIENQVLP AVERILKAFGYRKEDLRYQKTKQVGLGAWLKVKGKK	

2. ANTECEDENTES

2.1 Producción de proteínas recombinantes

La producción de proteínas recombinantes presenta tres componentes principales: el gen que codifica para la proteína de interés, la célula huésped que recibirá la modificación para la expresión de la proteína y un vector de expresión que transportará el material genético dentro de la célula huésped.

2.1.1 Cepa de expresión: *Escherichia coli*

Existen muchos sistemas de expresión basados en cultivos celulares de mamíferos, insectos, levaduras y bacterias (Guerrero-Olazarán *et al.*, 2004). De todos estos sistemas, la bacteria *E. coli* es el organismo más utilizado como huésped para la expresión de proteínas recombinantes (Rosano & Ceccarelli, 2014). Esto se debe a que, comparado con otros sistemas, *E. coli* presenta diversas ventajas: facilidad de manejo, una amplia gama de vectores compatibles, crecimiento con fuentes de carbono económicas, tiempo de generación corto y eficiente, escalamiento sencillo y rendimientos de 10 al 30% de la proteína total (Goulding & Perry, 2003). Aunado a lo anterior, *E. coli* cuenta con más de 40 años de conocimiento acumulado sobre su uso, por lo que se encuentra caracterizada a nivel molecular y se han diseñado diferentes cepas modificadas que permiten facilitar los procesos de clonación y expresión de proteínas en ellas (García *et al.*, 2013; Rosano & Ceccarelli, 2014). Algunas de estas cepas de *E. coli* modificadas son utilizadas para la expresión de ADN pol obtenidas de arqueas (**Tabla IV**). Mientras que existe sólo un reporte de la expresión de una ADN pol que utiliza la levadura *Pichia pastoris* como sistema de expresión (Wang *et al.*, 2015).

2.1.2 Vector de expresión

Los vectores de expresión son fragmentos de ADN que se utilizan como herramienta de apoyo en el proceso de transferencia de un gen exógeno al interior de la célula. Éstos agilizan la entrada del material genético al hospedero, permitiendo que se encuentre disponible intracelularmente. Los vectores más utilizados para la expresión de proteínas recombinantes son los plásmidos de tipo pET. Estos

vectores contienen un promotor del bacteriófago T7 y un operón lac sobre el gen de interés. Además un gen para la resistencia a antibiótico, sitios de restricción para la clonación y en ocasiones una etiqueta como la de histidinas (His-tag) que permiten la posterior purificación de la proteína de interés. Un buen vector de expresión debe presentar ciertas características, como la presencia de promotores fuertes con expresión basal baja, ser transferibles a otras cepas y ser inducidos por métodos simples y económicos. Los vectores más utilizados para la expresión de proteínas recombinantes son los plásmidos de tipo T7. La T7 ARN polimerasa se expresa en muchas de las cepas de *E. coli* y el operón lac permite una expresión inducible utilizando el análogo sintético de la lactosa Isopropil- β -D-1-tiogalactopiranosido (IPTG) como inductor (Hashemzadeh-Bonehi *et al.*, 1998; Duong-Ly & Gabelli, 2014).

Tabla IV. Cepas de *E. coli* comúnmente utilizadas para la expresión de ADN polimerasas (Familia B) de arqueas.

Cepa de expresión	Descripción
BL21(DE3)	Más utilizada, deficiente en proteasas Lon y OmpT
Origami(DE3)	Presenta mutaciones en genes que codifican para tiorredoxina reductasa y glutatión reductasa, permitiendo mejorar la formación de los enlaces disulfuro
BL21-CodonPlus	Contiene ARNt para codones raros
Rosetta (DE3)	Contiene ARNt para codones raros
Rosetta-gami(DE3)	Derivada de Origami, contiene ARNt para codones raros presentes en Rosetta
BL21pLyS	Produce lisozima del fago T7 para reducir el nivel basal de expresión del gen de interés
Rosetta pLyS	Produce lisozima del fago T7 para reducir el nivel basal de expresión del gen de interés. Tiene los ARNt raros de Rosetta

[*Tabla modificada de (Duong-Ly *et al.*, 2014)]

Debido a que cada ADN pol presenta características particulares a pesar de pertenecer a una misma familia, no existe una metodología uniforme para la producción de estas enzimas recombinantes. Sin embargo, existen diversos estudios que emplean los vectores T7 en combinación con el sistema de expresión en *E. coli* (Ej. Takagi *et al.*, 1997; Kim *et al.*, 2007; Lee *et al.*, 2009; Kim *et al.*, 2011) (Tabla V).

Tabla V. Ejemplos de sistemas y vectores utilizados en la expresión de ADN polimerasas (Familia B) de arqueas.

Microorganismo	Célula competente	Plásmido	Referencia
<i>Thermococcus guaymasensis</i>	BL21CodonPlus (DE3) RIL	pET22b(+)	(Lee <i>et al.</i> ,2009)
<i>Thermococcus sp.</i>	BL21CodonPlus (DE3) RIL	pET24a(+)	(Kim <i>et al.</i> ,2007)
<i>Thermococcus celericrescens</i>	Rosetta DE3pLysS	pET20b(+)	(Kim <i>et al.</i> ,2011)
<i>Pyrococcus sp.</i>	BL21(DE3)	pET8c	(Takagi <i>et al.</i> ,1997)

2.2 Complicaciones en la expresión de proteínas recombinantes en *E. coli*

A menudo suelen presentarse complicaciones durante la expresión de genes exógenos en *E. coli*, siendo uno de los problemas principales la formación de cuerpos de inclusión debido a la insolubilidad de la proteína de interés. Otras complicaciones que se pueden presentar son bajos niveles de expresión, degradación de la proteína, toxicidad y producción de proteína no funcional, siendo estos últimos relacionados principalmente a la falta de modificaciones postranscripcionales de *E. coli* (Sahdev *et al.*, 2008; Rosano & Ceccarelli, 2014). A pesar de que no existe una estrategia universal que permita resolver este tipo de problemáticas, se han elaborado diversos manuales, guías y revisiones (Makrides, 1996; Stevens, 2000; Baneyx & Mujacic, 2004; Jana *et al.*, 2005; Sørensen *et al.*, 2005; Duong-Ly & Gabelli., 2014;Rosano & Ceccarelli, 2014; Amid, 2015), que describen las estrategias a seguir para cada una de estas complicaciones. Estas estrategias se resumen en cinco pasos a seguir para mejorar los niveles de expresión y mejorar la solubilidad de la proteína a través de un plegamiento

correcto: 1) cambio de vector; 2) cambio de huésped (cepa de expresión); 3) cambio de los parámetros de cultivo del organismo huésped; 4) co-expresión con otros genes; y 5) optimización del gen a través de cambios en la secuencia. Por lo anterior es recomendable que si se trata de una proteína nueva que no requiera muchas modificaciones postranscripcionales y puede obtenerse de manera soluble, se inicie con *E. coli* como sistema de expresión, siendo la cepa BL21 (DE3) la opción más recomendada (Baneyx & Mujacic, 2004; Peleg *et al.*, 2012; Duong-Ly & Gabelli, 2014).

2.3 Trabajos previos en la expresión de ADN polimerasas de arqueas

La investigación de las ADN pol obtenidas de arqueas comenzó a principios de los años noventa, cuando P. Mattila y colaboradores caracterizaron la primer ADN pol a partir del Euryarchaeon *Thermococcus litoralis* (Mattila *et al.*, 1991). Desde entonces se han aislado, caracterizado y comercializado para su uso en laboratorio, una gran cantidad de ADN pol obtenidas de arqueas, principalmente del orden de los *Thermococcales* y de la familia B (**Tabla VI**) (Zhang *et al.*, 2015).

Tabla VI. Comparación de las características de ADN polimerasas obtenidas de arqueas, utilizadas comercialmente(Zhang *et al.*,2015).

Archaea	DNA pol	Familia	Tasa de error (x10⁻⁶)	Tasa de extensión (kbp/min)	Actividad 3'-5' exonucleasa	Temp. óptima (°C)	Termo-estabilidad	Referencias
<i>Crenarchaeota</i>								
<i>Sulfolobus solfataricus</i>	Dpo4	Y	8000-300	NP	No	NP	NP	(McDonald <i>et al.</i> ,2006)
<i>Euryarchaeota</i>								
<i>Pyrococcus furiosus</i>	Pfu	B	2.2 - 0.7	0.5-1.5	Sí	72-80	95°C/1h	(Cline,1996)
<i>Pyrococcus GB-D</i>	Deep Vent TM	B	12-2.7	1.4	Sí	72-75	95°C/23h	(Cline,1996)
<i>Pyrococcus abyssi</i>	Isis TM	B	6.7-0.6	NP	Sí	70-80	100°C/5h	(Dietrich <i>et al.</i> ,2002)
<i>Pyrococcus abyssi</i>	Pwo	B	NP	NP	Sí	72	100°C/2h	(Dabrowski <i>et al.</i> ,1998)
<i>Thermococcus kodakarensis</i>	KOD1	B	2.6	6.0-7.8	Sí	72-75	95°C/12h	(Takagi <i>et al.</i> ,1997)
<i>Thermococcus litoralis</i>	Vent TM	B	45-2.8	1	Sí	72-80	95°C/6.7h	(Cline,1996)
<i>Thermococcus gorganarius</i>	Tgo	B	5.6-3.5	1.5	Sí	72	95°C/2h	(Waltraud Ankenbauer <i>et al.</i> ,1997)

*NP No ha sido Publicado

3. JUSTIFICACIÓN

La obtención de proteínas recombinantes para uso en biotecnología, a partir de microorganismos aislados en México, que permita desarrollos sin conflicto de patente, conlleva distintos retos, uno de ellos es la estandarización del proceso de producción. Para la producción de PR generalmente se utiliza un organismo diferente al que pertenece la proteína de interés, lo anterior puede provocar; entre otras cosas, la producción de poca cantidad de proteína debido a un efecto tóxico de ésta en la célula huésped o un mal plegamiento de la misma que provoque su insolubilidad. La solución a este tipo de problemas suele representar un desafío complejo en el que no siempre existe una solución clara, por lo que en muchos casos es necesario recurrir a guías o manuales que permitan idear un mejor plan de trabajo (Duong-Ly & Gabelli, 2014).

Como se mencionó anteriormente, las ADN pol son una de las enzimas más utilizadas como herramienta de trabajo en biología molecular. Diversos estudios bioquímicos y estructurales revelan que cada ADN pol obtenida de arqueas a pesar de su alta similitud estructural y elevado porcentaje de identidad en secuencias, presenta diferentes características entre sí, con respecto a propiedades específicas como lo son la procesividad, fidelidad y termoestabilidad (**Tabla VI**). Debido a características como su alta fidelidad y fuerte termoestabilidad, las ADN pol de la familia B obtenidas a partir de arqueas son ampliamente utilizadas en técnicas como PCR de alta fidelidad, secuenciación de ADN y mutagénesis dirigida. Ya que la actividad de las polimerasas varía ampliamente entre sí incluso entre especies cercanas, su búsqueda y caracterización puede conducir al potencial desarrollo de nuevos reactivos con aplicaciones dirigidas al perfeccionamiento de este tipo de técnicas (Ishino & Ishino, 2014; Zhang *et al.*, 2015).

Existen múltiples guías que facilitan la toma de decisiones durante el proceso de obtención de proteínas recombinantes (Ej. Makrides, 1996; Stevens, 2000; Baneyx & Mujacic, 2004; Jana *et al.*, 2005; Sørensen *et al.*, 2005; Duong-Ly & Gabelli, 2014; Rosano & Ceccarelli, 2014; Amid, 2015). No obstante, no existen publicaciones que permitan observar las problemáticas relacionadas a la expresión

específica de ADN pol B de arqueas. Lo anterior, quizá se deba a la tendencia de publicaciones a enfocarse sólo en los resultados significativos para dar sentido a los experimentos, y no al proceso metodológico que le respalda.

4. HIPÓTESIS

Siendo *T. gammatolerans* una especie de arquea anaerobia hipertermófila, se espera que su ADN pol (B) pueda obtenerse bajo un sistema de expresión similar a los empleados para enzimas homólogas utilizadas en aplicaciones biotecnológicas.

5. OBJETIVO GENERAL

- Evaluar los sistemas de expresión utilizados en enzimas homólogas para la obtención de la proteína recombinante ADN polimerasa (Familia B) en *T. gammatolerans*.

6. OBJETIVOS ESPECÍFICOS

-Considerando los trabajos previos con enzimas homólogas de TgPolB, proponer las estrategias de expresión y purificación a utilizar.

- Probar las estrategias de expresión y purificación previamente seleccionadas según las características de los vectores seleccionados.

7. METODOLOGÍA

7.1 Obtención de construcciones

7.1.1 Diseño de oligonucleótidos para el gen que codifica TgPolB

El ADN genómico de *T. gammatolerans* fue donado por el Dr. Patrick Forterre del Instituto Pasteur. La secuencia del gen que codifica para TgPolB se consultó en las bases de datos de NCBI. Con ella se diseñaron los oligonucleótidos específicos utilizando el programa BLAST para el análisis de especificidad de los mismos.

7.1.2 Amplificación del gen a partir del ADN genómico

La mezcla de reacción, así como el programa de PCR utilizados para la amplificación del gen se describe a detalle en las (Tablas VII y VIII). Por su parte, la temperatura óptima de fusión de los oligonucleótidos se determinó mediante PCR de gradiente entre 56 - 72°C tomando en cuenta la temperatura sugerida por el proveedor.

Tabla VII. Mezcla de reacción utilizada para la amplificación del gen codificante de TgPolB.

Reactivo	Concentración final	Volumen por reacción
H ₂ O	--	9.25 µl
Regulador GC5X	1X	5 µl
dNTP's	200 µM	2.5 µl
Oligo Sentido	1mM	2.5 µl
OligoAnti-sentido	1mM	2.5 µl
DMSO	3%	0.75 µl
ADN genómico	98 ng	2 µl
Phusion polimerasa	1 U	0.5 µl
<i>Volumen final</i>		<i>25 µl</i>

Tabla VIII. Programa de temperaturas y tiempo para amplificación del gen codificante de TgPolB.

Proceso	Temperatura	Tiempo	Ciclos
Desnaturalización Inicial	98°C	3 min.	1 ciclo
Desnaturalización	98°C	10 seg.	35 ciclos
Alineación Extensión	72°C	70 seg.	
Extensión final	72°C	10 min.	1 ciclo
Almacenamiento	4°C		

7.1.3 Clonación del gen que codifica TgPolB en pJET1.2 BluntCloning Vector

El producto de PCR se clonó en el plásmido pJET1.2 BluntCloning Vector utilizando 1 µl de producto de PCR sin purificar en presencia de regulador de ligación 1X y 1µl de T4 ADN ligasa (Promega). Esta reacción se incubó a temperatura ambiente por 15 min y se almacenó a -20°C hasta el momento de la ligación. La ligación se transformó en células competentes DH5α y por medio de una reacción de PCR se confirmó la presencia del inserto en las colonias transformadas. Posteriormente se realizó extracción de ADN plasmídico (**Anexo I**) a las colonias que resultaron positivas y se les realizó una segunda verificación a través de un ensayo de doble restricción. Para esto, el ADN plasmídico se incubó a 37°C durante toda la noche en dos mezclas de reacción distintas (**Tabla IX**). La primera reacción, con las enzimas BamHI y NdeI, mientras que la segunda sólo con BamHI. Lo anterior, para verificar la liberación del inserto por el corte con ambas enzimas (Reacción 1) y la linealización del vector debida al corte único en el sitio de restricción de BamHI (Reacción 2).

Finalmente, las colonias que resultaron positivas para ambos ensayos se verificaron por medio de secuenciación del gen codificante para la proteína TgPolB.

Tabla IX. Componentes de la reacción para el ensayo de doble restricción con las enzimas BamHI y NdeI.

Reactivo	Reacción 1	Reacción 2
pJET1.2 Blunt Cloning Vector		
Enzima NdeI (NEB)	1 μ l	-
Enzima BamHI (NEB)	1 μ l	1 μ l
Regulador 3.1 (NEB)	1X	1X
Agua libre de nucleasas	a 10 μ l	a 10 μ l

7.1.4 Subclonación en vectores de expresión

Posteriormente las clonas positivas se subclonaron en los vectores de expresión pET19by pColdI (Takara Biolnc, 2017), siguiendo los pasos anteriormente descritos para la confirmación de las colonias. Estos vectores agregan una etiqueta de seis histidinas en el extremo N-terminal a las proteínas para su posterior purificación por cromatografía de afinidad a níquel. La extracción del ADN plasmídico se realizó de acuerdo a la metodología descrita en **Anexo II**, y se procedió a realizar la cuantificación en un espectrofotómetro ND-100 (NanoDrop®, Delaware, EE.UU.), estableciendo una relación de absorbancia a una longitud de 260 y 280 nm entre un rango de 1.8-2.0.

7.1.4.1 Confirmación de construcciones por medio de secuenciación del gen codificante de TgPoIB.

La secuenciación se realizó en la unidad universitaria de secuenciación masiva y bioinformática de la UNAM (UUSMB), así como en el Instituto de Biotecnología de Macrogen Inc.

7.2 Expresión de TgPolB

7.2.1 Transformación en las cepas de expresión BL21 y Rosetta II (RII)

Las diferentes construcciones con el gen de interés que codifica la proteína TgPolB, se transformaron en las cepas de expresión BL21 y RII. En tubos colocados previamente en hielo se agregaron 50 μ l de células competentes (BL21 o RII) y se agregó 1 μ l de las construcciones en pColdI y pET19b. Este tubo se dejó incubar en hielo por 15 min y se sometió a las células a un choque térmico durante 45 segundos a 42°C, dejando enfriar en hielo durante 2 min. Posterior a esto, se agregaron 250 μ l de medio de cultivo Luria Bertani (LB) y se dejaron incubar durante 1 hr a 37°C y aproximadamente 150 rpm. Finalmente se realizó un cultivo masivo agregando 100 μ l de células a una caja previamente preparada con LB y ampicilina a una concentración final de 100 μ g/ml para las transformaciones realizadas en BL21 y LB con ampicilina (100 μ g/ml) y cloranfenicol a una concentración final de 34 μ g/ml para las transformaciones en RII.

7.2.2 Cinética de inducción de TgPolB

Se realizó una cinética de inducción de la proteína TgPolB con las construcciones en pET19b y pColdI en las cepas de expresión BL21 colectando alícuotas de 1 ml cada 5 horas durante 35 horas para verificar el momento en que se obtienen los mejores rendimientos de expresión.

7.2.3 Inducción

Se preparó un preinóculo con una colonia de cada una de las cepas de expresión transformadas. La inducción se realizó añadiendo diferentes concentraciones de Isopropil- β -D-1-tiogalactopiranosido (IPTG) (0.25, 1 y 1.25 mM) a 1 ml de cultivo fresco de las transformaciones anteriores, previamente preparadas y ajustadas a una densidad óptica entre 0.6 y 0.8 a 600 nm. Las construcciones en pET19b, se dejaron incubar durante 4 hrs a 35°C, mientras que aquellas en pColdI se dejaron incubar por 16 hrs a una temperatura entre 16 y 18°C. Una vez

recolectada la alícuota se centrifugó a 10,000 rpm durante 2 min y se resuspendió en regulador de carga con colorante azul de bromofenol.

Los resultados de las inducciones (**Tabla X**) se analizaron por electroforesis en geles de poliacrilamida al 10% (**Anexo II**), utilizando como control negativo 1 ml de células no inducidas y como control positivo la expresión de una proteína bien caracterizada. Los geles fueron observados en un fotodocumentador (Bio-Rad). El control positivo se preparó en geles al 15% debido al peso molecular de la proteína control.

Tabla X. Condiciones de inducción analizadas.

Construcciones	Temperatura de inducción (°C)	Tiempo de inducción (Hrs)	Cepas de expresión	Concentración del Inductor
pET19C1 pET19C2 pET19C3	35	4-6	R11 y BL21	0.25, 0.5, 1 y 1.25 mM
pColdC4	16-18	16-20		

7.2.4 Relación medio de cultivo y expresión

Se realizó nuevamente la inducción utilizando como cepa de expresión BL21 (DE3) con las condiciones previamente descritas (**Tabla XI**) pero probando diferentes medios de cultivo (**Anexo III**):

- 1) Super broth (medio de autoinducción)
- 2) 2XTY enriquecido con 0.01% de glucosa (medio de autoinducción, la glucosa permite evitar la expresión basal de sistemas inducibles por IPTG).

7.3 Purificación de la enzima TgPolB

Basados en la bibliografía (**Tabla V y VI**) se seleccionó la cepa de expresión BL21(DE3), el vector pColdI, la temperatura de inducción de 16 °C, y el tiempo de inducción de 16 horas como condiciones de inducción de la enzima TgPolB para el proceso de purificación.

La purificación de la enzima se realizó utilizando la técnica de cromatografía por afinidad a níquel, por el método manual y con el sistema de FPLC (Fast Protein Liquid Chromatography) en el equipo ÄKTA Pure (GE) con una columna HisTrap (GE Healthcare) de 1 ml de resina.

7.3.1 Purificación (Regulador con EDTA)

Para esta purificación se utilizó regulador de lisis (50 mM Tris pH 8.0, 100 mM NaCl y 1mM EDTA). Se resuspendió el paquete celular en regulador de lisis y se añadió lisozima a una concentración final de 1 mg/ml. Éste se dejó incubar en hielo durante 30 minutos con agitación constante y se sometió a dos ciclos de congelación (4 min) con nitrógeno líquido. Posteriormente se agregó inhibidor de proteasas (PSMF) a una concentración final de 1 mM y se procedió a sonicar la muestra durante 3 ciclos de 20s de pulso y 20s de reposo. La fracción no soluble se obtuvo por medio de centrifugación a 4,000 rpm durante 30 minutos. Mientras que la fracción soluble se filtró (0.45 µm) y se inyectó a través de una columna de afinidad a níquel (GE Healthcare) utilizando un gradiente escalonado con lavados de 15 volúmenes de columna (VC) con 20 y 40 mM de imidazol. La elución de esta purificación se realizó de manera manual y se obtuvo con 3 VC de regulador de elución (50 mM Tris-HCl, 100 mM NaCl y 500 mM imidazol). El resultado se verificó con un gel de poliacrilamida al 8%.

7.3.2 Purificación (Regulador sin EDTA)

Para esta purificación se utilizó regulador de lisis (50 mM Tris pH 8.0 y 100 mM NaCl). Se resuspendió el paquete celular en regulador de lisis y se añadió lisozima a la muestra. Este se dejó incubar en hielo durante 30 min con agitación

constante y se sometió a dos ciclos de congelación (4 min cada uno) con nitrógeno líquido. Una vez descongelado se procedió a sonicar la muestra durante 3 ciclos de 20s de pulso y 20s de reposo. La fracción no soluble se obtuvo por medio de centrifugación a 4,000 rpm durante 30 minutos. Mientras que la fracción soluble se filtró (0.45 μ m) y se inyectó a través de una columna de afinidad a níquel (GE Healthcare).

El lavado y elusión de la columna se realizó de manera manual con un gradiente escalonado de imidazol: 20 mM, 40 mM y 500mM. También se llevó a cabo una segunda purificación cuya elución de la columna se realizó en el equipo ÄKTA Pure con un gradiente escalonado de imidazol: 20 mM, 40 mM, 100 mM, 150 mM, 175 mM, 190 mM, 200 mM y 500 mM. Finalmente se analizaron las muestras en geles de poliacrilamida al 10% (**Anexo II**) y se seleccionaron para su análisis las fracciones donde se observó la proteína de interés.

7.3.3 Purificación con detergentes

Se llevó a cabo el uso de dos detergentes distintos para descartar que TgPolB se encontrara en cuerpos de inclusión. Sarcosyl al 5% y TritonX-100 al 0.1%. Las purificaciones se realizaron de manera similar a las purificaciones anteriores, variando únicamente en la conformación del regulador de lisis. Se utilizó TrisHCl 50 mM, NaCl 100 mM pH:7.9 y Sarcosyl 5% p/v para la purificación con Sarcosyl y TrisHCl 50 mM, NaCl 100 mM pH: 7.9 con TritonX-100 0.1% para la purificación con TritonX-100. Los lavados y elusiones de la columna se realizaron de manera manual con un gradiente escalonado de imidazol (20 mM, 40 mM y 500 mM), y se analizaron las muestras en geles de poliacrilamida al 10%.

7.4 Cuantificación de proteína

Se preparó una curva con concentraciones conocidas de albúmina (8 mg, 4 mg, 2 mg, 1 mg y 0.5 mg) y se cargó junto con las fracciones seleccionadas de las eluciones de la purificación manual y en ÄKTA Pure en un gel de poliacrilamida al 10%. La cuantificación se realizó por medio de densitometría comparando la

intensidad de color por azul de Comassie, entre las muestras de la purificación y las concentraciones de albúmina conocidas.

7.5 Análisis de actividad de TgPolB por medio de PCR

Las fracciones seleccionadas a partir de las eluciones de la purificación manual y por ÄKTA Pure, con y sin EDTA se emplearon para el análisis de actividad por medio de PCR. Para esto se realizó la amplificación de un fragmento de 336 pb, cuyo programa de amplificación está bien establecido (Almanza, 2018). Se utilizó como control positivo la enzima Taq pol de Vivantis. Se probaron concentraciones de 0.064 mM, 0.16 mM y 0.32 mM para las fracciones 1 y 3, y concentraciones de 0.032 mM, 0.8 mM y 0.16 Mm para la elución 2. El ADN se estableció a una concentración de 100 ng/μl, y para la preparación de las muestras se utilizó regulador de la Taqpol de Vivantis, dNTP's 10 mM y oligonucleótidos específicos (sentido y anti-sentido) a 10 μM. Se preparó el gel de agarosa en el cual se añadieron 3 μl del marcador de peso molecular [2μl del fluorocromo GelRed (Biotium) con 1 μl de marcador] y 4 μl de cada muestra (2μl de muestra con 2μl de GelRed).

Se agregó además en el mismo gel las fracciones con ADN para el análisis de actividad exonucleasa, utilizando como control el ADN resuspendido en la misma agua en la que se llevan a cabo las reacciones.

Se analizaron además diferentes reguladores (**Tabla XI**) y temperaturas de reacción de 72°C y 68°C (**Tabla XII**).

Tabla XI. Reguladores utilizados para el análisis de actividad enzimática.

Clave	Regulador
A	S- Taq Vivantis
B	5X Phusion GC
C	5X Phusion HF
D	10X Pfu- Altaenzymes (Sin Mg ⁺²)
E	10X A- Taq Vivantis (Sin Mg ⁺²)

Tabla XII. Programa control, utilizado para la reacción de PCR a 72°C y 68°C.

Pasos	Temperatura	Temperatura	Tiempo
1	94°C	94°C	5 min.
2	94°C	94°C	30 seg.
3	53°C	53°C	30 seg.
4	72°C	68°C	40 seg.
5	Regresar al paso 2	Regresar al paso 2	35 veces
6	72°C	72°C	10 min.

7.6 Análisis de Western-Blot

Las fracciones obtenidas de las purificaciones 1 y 3, se enviaron al Laboratorio de Microbiología Molecular del Instituto de Biotecnología de la UNAM para su posterior análisis por Western-Blot (WB). Además se repitió la purificación por el método manual y se enviaron para su análisis las alícuotas colectadas durante diferentes puntos del proceso para dar seguimiento a la expresión de la proteína TgPolB y detectar si existía degradación durante su manipulación en alguno de los pasos.

7.7 Subclonación en pET32a

Para la obtención del inserto se realizó la digestión de la construcción C4 en el vector pCold-I con las respectivas enzimas de restricción BamHI y NdeI. Debido a los resultados obtenidos a partir de este ensayo, se procedió a repetir la prueba siguiendo el procedimiento descrito previamente en el apartado 7.1.3, referente al ensayo de doble restricción, utilizando el resto de las construcciones.

7.8 Análisis bioinformáticos de TgPolB

7.8.1 Modelado estructural *in silico*

Se utilizó SWISS-MODEL (<https://swissmodel.expasy.org/>) (Waterhouse *et al.*, 2018), para la elaboración de un modelo estructural de la proteína TgPolB. Posteriormente las secuencias de diferentes ADN pol B pertenecientes al orden de los *Thermococcales* (TgPolB: *T. gammatolerans*; PGBD: *Pyrococcus furiosus* DeepVent; Tgo: *T. gorgonarius*; Pfu: *Pyrococcus furiosus*) se analizaron en InterPro (Hunter *et al.*, 2009), para la identificación de sus dominios: Exo y multifuncional. Además se realizó el alineamiento de las secuencias y se elaboró una imagen en donde con los reportes de la patente WO2007043769A1-2007M y los trabajos de Argos, 1991 y Zhang, 2015, se identificaron los aminoácidos característicos que distinguen los dominios de estas ADN pol (Bordo *et al.*, 1991; Zhang *et al.*, 2015).

Una vez identificados los aminoácidos característicos, se procedió a ubicarlos dentro del modelo estructural previamente diseñado con SWISS-MODEL y se elaboró una imagen utilizando el software Chimera (Pettersen *et al.*, 2004).

7.8.2 Análisis de identidad de secuencias

Utilizando el programa BLAST se realizó un análisis de identidad de las principales secuencias de ADN pol B utilizadas actualmente de manera comercial, obtenidas a partir de *Thermococcales*.

7.8.3 Ubicación de N-terminal

La ubicación del N-terminal dentro de la secuencia se ubicó comparando el alineamiento de las secuencias de las ADN pol B con los reportes de la patente WO2007043769A1-2007M y los trabajos de Argos, 1991 y Zhang, 2015 (Bordo *et al.*, 1991; Zhang *et al.*, 2015). Posteriormente se utilizó el software CCP4MG en donde se visualizó el modelo estructural previamente elaborado en SWISS-MODEL y dentro del cual se ubicaron los aminoácidos correspondientes al N-terminal.

Finalmente se obtuvo también un modelo con el potencial electrostático para la identificación de las cargas en este punto.

7.8.4 Sitios Activos

Para la identificación de los sitios activos en la ADN pol B de *T. gammatolerans*, se ingresó el modelo estructural previamente elaborado por SWISS-MODEL en la página de Supercomputing Facility for Bioinformatics & Computational Biology, IIT Delhi (SCFBio) http://www.scfbio-iitd.res.in/dock/ActiveSite_new.jsp. Los resultados obtenidos se compararon con los sitios de unión a ADN reportados ya en un modelo cristalográfico para *T. gorgonarius*, (Tgo) y se superpusieron en el modelo estructural de TgPolB previamente elaborado por SWISS-MODEL. Se consideró como potenciales sitios activos aquellas cavidades que concordaban con los sitios de unión a ADN de la enzima homóloga Tgo.

7.8.5 Índice de adaptación de codones

Para verificar si el uso preferencial de codones representa una limitante en la producción de la proteína recombinante se ingresó la secuencia en el servidor en línea Genscript (<https://www.genscript.com/tools/rare-codon-analysis>) para el análisis del Índice de Adaptación de Codones (IAC) con respecto al huésped *E. coli*. Posteriormente se realizó una simulación de optimización en el programa Codon Optimization tool (<https://www.idtdna.com/codonopt>), la cual fue analizada nuevamente en Genscript y comparada con la secuencia original de TgPolB. Para este análisis los “codones raros” se definieron como los codones que son empleados por *E. coli* en una frecuencia de <1% (Kane, 1995).

7.8.6 Enlaces disulfuro

Para el cálculo de la potencial formación de enlaces disulfuro, se ingresó la secuencia en los servidores de DISULFIND (Ceroni *et al.*, 2006) y DiANNA (Ferrè *et al.*, 2006), ambos disponibles en línea:

DISULFIND: <http://disulfind.dsi.unifi.it/monitor.php?query=UB7Z9A>

DiANNA1.1: <http://clavius.bc.edu/~clotelab/DiANNA/>

8. RESULTADOS Y DISCUSIÓN

8.1 Obtención de construcciones

8.1.1 Diseño de oligonucleótidos para el gen que codifica TgPolB

Se diseñaron los siguientes oligonucleótidos específicos:

Oligo Sentido: 5'-GGTCGGAATTCCATATGATTCTCGATACCGACTACATCACC-3'

Oligo Anti-sentido: 5'-GGTTGGGGATCCTCACTTCTTCCCTTTCACC-3'

8.1.2 Amplificación y clonación del gen que codifica TgPolB en pJET1.2 BluntCloning Vector

El producto de amplificación obtenido por PCR a partir del ADN genómico (**Figura 3A**) se subclonó en el vector pJET 1.2 BluntCloning Vector de manera exitosa obteniendo cinco colonias positivas. Éstas fueron cultivadas para extracción de ADN plasmídico, confirmando la presencia del inserto por medio de PCR (**Figura 3B**). El ensayo de doble restricción realizado a partir de las clonas positivas confirmó la liberación del inserto (2.32 Kb) en presencia de ambas enzimas de restricción (BamHI y NdeI) y la linealización del vector (5.32Kb) en la reacción incubada sólo con BamHI (**Figura 3C**). Lo anterior fue visualizado en un gel de agarosa al 1% en TBE (Tris/Ácido Bórico/EDTA) y teñido con el fluorocromo GelRed.

8.1.3 Subclonación en vectores de expresión

Se lograron confirmar 4 construcciones positivas, tres en pET19b y una en pColdI (**Tabla XIII**), las cuales fueron analizadas en el presente estudio.

Tabla XIII. Concentración de ADN plasmídico de las construcciones de trabajo.

Clave	Construcciones	Concentración [ng/μl]	Absorbancia			
			260	280	260/280	260/230
C1	pET19b clona 1	248.9	4.97	2.459	2.02	2.44
C2	pET19b clona 2	297.6	5.952	2.90	2.04	2.26
C3	pET19b clona 12	238	4.760	2.32	2.04	2.26
C4	pColdI clona 12	338	6.765	3.30	2.08	2.23

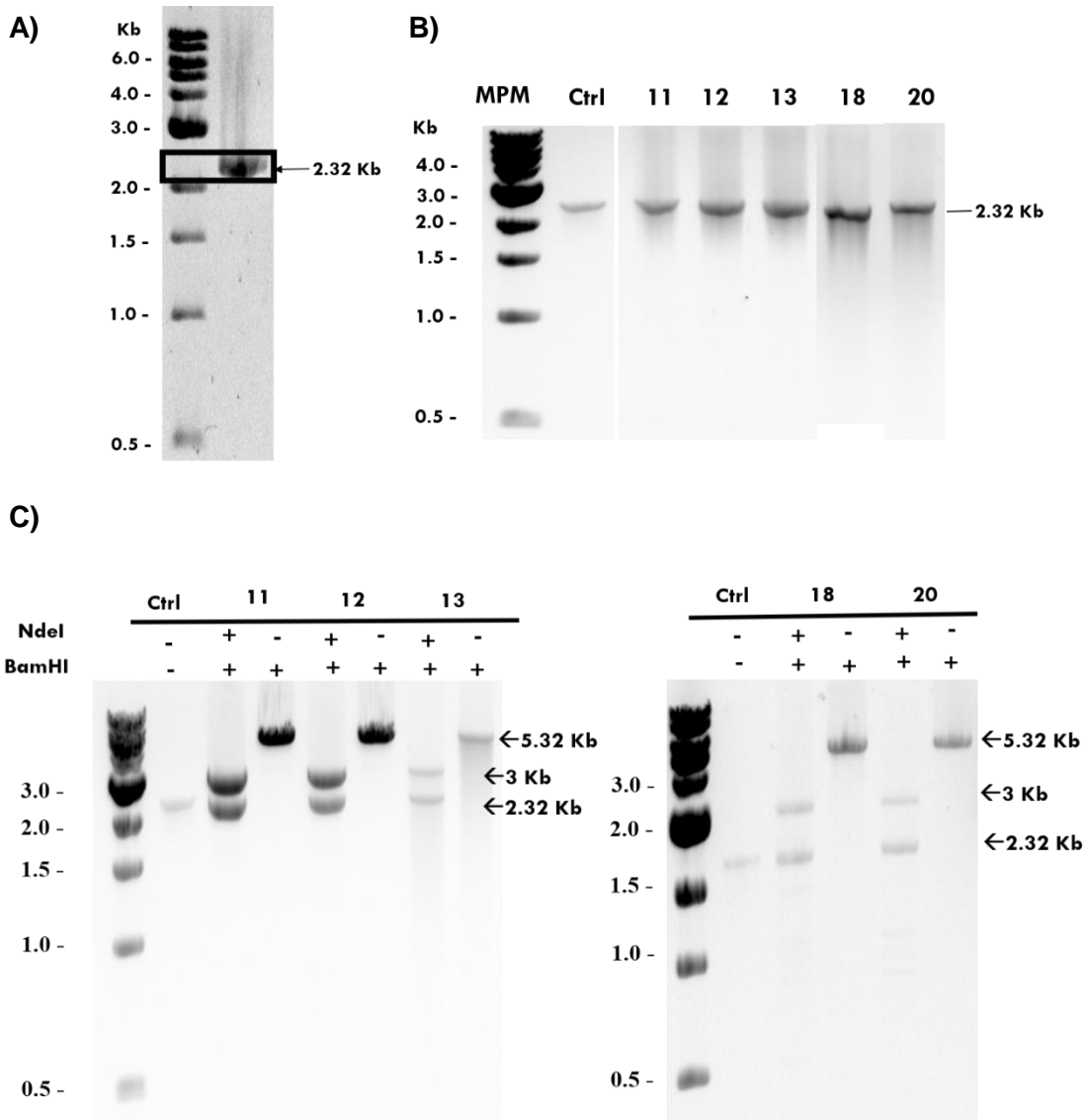


Figura 3. Clonación de TgPolB en pJET 1.2 Blunt Cloning Vector. A) Amplificación de TgPolB a partir del ADN genómico. Amplificación realizada a partir del ADN genómico con los oligos específicos para el gen de interés. El producto de amplificación de 2.32Kb se indica con una flecha. **B)** Confirmación de las colonias positivas en pJET1.2 BluntCloning Vector por medio de PCR. **C)** Ensayo de doble restricción con enzimas NdeI y BamHI para la confirmación de las colonias positivas en pJET1.2 BluntCloning Vector. += presencia; -= ausencia; Ctrl= amplificación del gen; MPM= marcador de peso molecular.

8.1.3.1 Confirmación de construcciones por medio de secuenciación

La secuenciación de las construcciones de trabajo, demostró que ninguna de ellas presentó mutaciones dentro de la región codificante para la proteína de interés (**Anexo V**).

8.2 Expresión de TgPolB

8.2.1 Cinética de inducción de TgPolB en BL21(DE3)

La mayoría de las ADN pol B que han sido expresadas, purificadas y caracterizadas a partir de arqueas, han utilizado *E. coli* como sistema de expresión de manera exitosa, destacando el uso de cepas como BL21(DE3), BL21-CodonPlus, B834 (DE3), Rosetta(DE3), y Rosetta pLyS (Takagi *et al.*, 1997; Hopfner *et al.*, 1999; Lee *et al.*, 2009; Kim *et al.*, 2011). Al no existir reportes que sugieran el uso particular de alguna de estas cepas para la expresión de ADN polimerasas, se seleccionó la cepa BL21(DE3) como cepa de expresión para la producción recombinante de TgPolB, debido a que es la más recomendada para los análisis exploratorios cuando se comienza con una proteína nueva (Baneyx & Mujacic, 2004; Peleg *et al.*, 2012; Duong-Ly & Gabelli., 2014). No obstante, la cinética de inducción realizada con la cepa BL21(DE3), inducida con 0.5 mM IPTG no mostró la sobreexpresión aparente de TgPolB (89.9 kDa) en ningún punto de la expresión (**Figura 4**). Por lo que fue necesario analizar otras variables diferentes del tiempo, tales como la cepa de expresión y la concentración el inductor.

8.2.2 Inducción

Las condiciones del cultivo son de gran importancia para la expresión de PR, ya que distintas cepas o vectores de expresión presentarán diferencias en los niveles de producción de una misma PR bajo condiciones de cultivo similares. Por lo anterior, trabajos como los de Gopal & Kumar (2013) y Gräslund (2008) proponen realizar el análisis conjunto de diferentes condiciones de cultivo para una mejor estandarización y optimización del proceso (Gräslund *et al.*, 2008; Gopal *et al.*, 2013). Entre las condiciones de cultivo analizadas se encuentran la cepa de expresión Rosetta II (RII), la cual contiene ARN de transferencia para codones

diferentes a los de *E. coli*, favoreciendo así la expresión de proteínas ricas en codones raros (como es el caso de TgPoIB), lo cual se describirá a detalle más adelante.

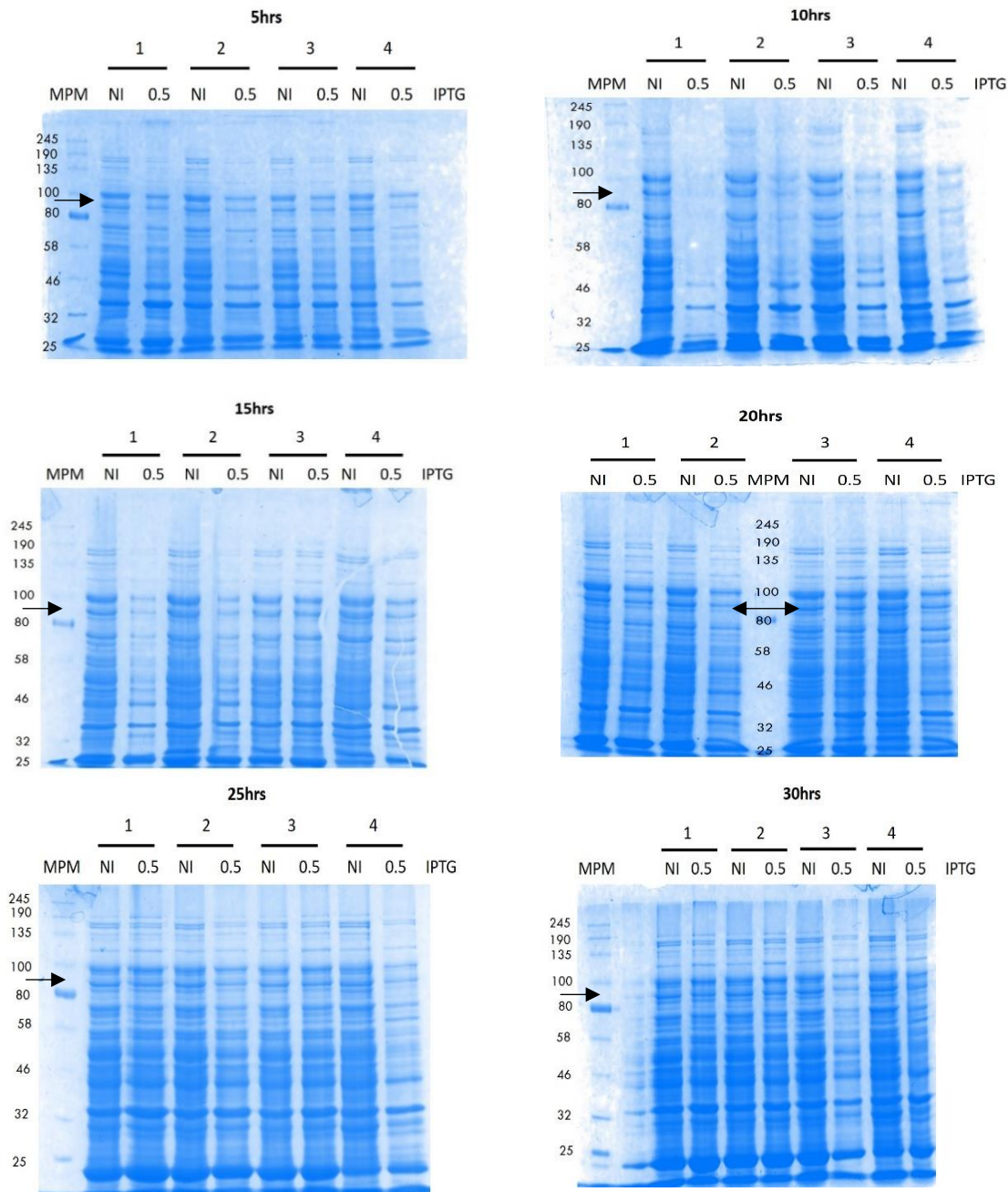


Figura 4. Cinética de inducción de TgPoIB. SDS-PAGE 10% de Inducción a DO 0.600 con 0.5 mM IPTG en BL21. Los números del 1 al 4 representan la construcción analizada. 1-3: pET19b y 4: pColdI. NI= control no inducido. MPM= marcador de peso molecular.

Por su parte, la concentración del inductor se ha descrito como factor clave en la solubilidad de la proteína de interés (Duong-Ly & Gabelli, 2014)., reportando que a concentraciones más bajas de IPTG (0.1mM) se favorece una expresión más lenta de la proteína, permitiendo el plegamiento correcto de proteínas complejas.

En este estudio no fue posible apreciar diferencias significativas en los niveles de expresión con respecto a la cepa de expresión, vector de expresión y concentración de inductor (IPTG) utilizados (**Figura 5**), ya que, independientemente de las condiciones analizadas, no se distingue la sobreexpresión de TgPolB a la altura del peso molecular esperado de la proteína (89.9 kDa). Cabe destacar, que este resultado no indica la ausencia total de la expresión de TgPolB. Existe la posibilidad de que la proteína pueda expresarse en niveles muy bajos e indetectables para la técnica de tinción utilizada (azul Comassie: 30 ng mínimo), en los geles de poliacrilamida.

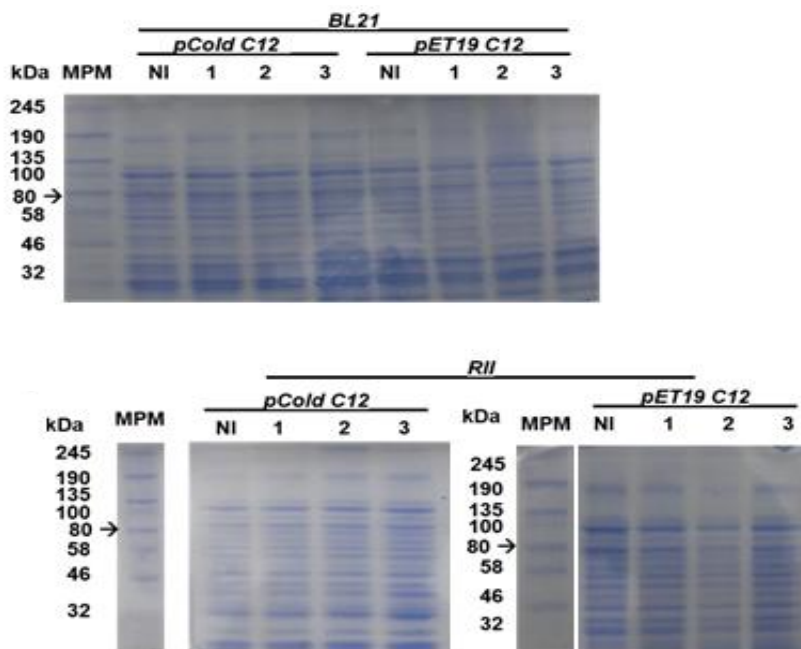


Figura 5. Inducción de TgPolB en cepas BL21 y Rosetta II (RII). SDS-PAGE 10% a diferentes concentraciones de inductor (IPTG): 1= 0.25 mM; 2= 1mM; 3= 1.25 mM; NI= control no inducido. Las cepas en el plásmido pColdI se incubaron durante 16 hrs a una temperatura entre 16 y 18°C, mientras que las cepas en pET19b se incubaron durante 4 hrs a 35°C.

8.2.3 Relación del medio de cultivo y la expresión

Luria-Bertani (LB) es el medio más utilizado de manera inicial para la obtención de proteína recombinante, no obstante, a pesar de tratarse de un medio enriquecido, éste no tolera el crecimiento celular hasta muy altas densidades, particularmente porque contiene bajas cantidades de fuente de carbono (Lara, 2011). Autores como Rosano & Ceccarelli (2014) sugieren, como estrategia ante niveles muy bajos de expresión, incrementar la cantidad de proteína recombinante a través de un aumento en la densidad del cultivo. Por lo anterior, y considerando los resultados de las pruebas de inducción previas, se cambió la composición del medio, utilizando medios de autoinducción: 1) 2XYT enriquecido con glucosa (0.01%) y 2) Super Broth (SB). En los medios de autoinducción enriquecidos con glucosa (0.01%), la glucosa es la fuente de carbono preferida y se metaboliza preferentemente durante el crecimiento, previniendo así el consumo de lactosa hasta que la glucosa se agota. Posteriormente se consume el glicerol y la lactosa, siendo esta última el inductor de las proteínas controladas por el operón lac. De este modo, el monitoreo de la biomasa para la adición del inductor puede evitarse, así como la manipulación del cultivo (Gräslund *et al.*, 2008).

Contrario a lo esperado, el uso de 2XYT enriquecido con 0.01% de glucosa y Super Broth como medios de cultivo alternativos para la expresión de TgPolB, no representó diferencia significativa en los resultados previamente observados, ya que en ninguno de los geles con las diferentes condiciones analizadas fue posible observar la sobreexpresión aparente de la proteína de interés (89.9 kDa). No obstante, sí se detecta un aumento en la densidad celular, el cual puede observarse a través de la saturación de color en los geles (**Figura 6**).

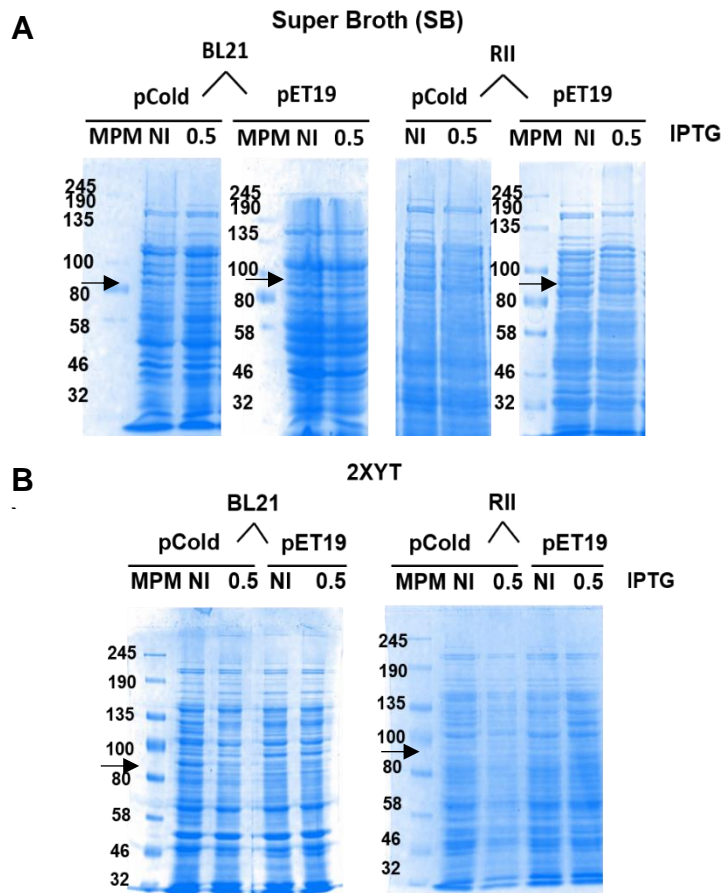


Figura 6. Inducción TgPolB con diferentes medios de cultivo. SDS-PAGE 10% de inducción de TgPolB (0.5 mM de IPTG) en cepas BL21 y Rosetta II (R11) con diferentes medios de cultivo. MPM= marcador de peso molecular; NI= no inducido; 0.5= inducido con IPTG 0.5 mM. * Las cepas en el plásmido pColdI se incubaron durante 16 hrs a una temperatura entre 16°C, mientras que las cepas en pET19b se incubaron durante 4 hrs a 35°C. **A)** Inducción con Superbroth; **B)** Inducción con 2XYT.

En ese sentido, Gopal & Kumar (2013) y Duong-Ly & Gabelli (2014), sugieren que ante problemáticas de poca o nula expresión de PR se realice la co-expresión de la proteína de interés con otras proteínas de interacción, como lo son las chaperonas moleculares (Gopal *et al.*, 2013; Duong-Ly & Gabelli, 2014). Las cuales son proteínas que se asocian con polipéptidos que no se han plegado, suelen prevenir la asociación inespecífica de las cadenas polipeptídicas y con ello la pérdida de actividad biológica, promoviendo el plegamiento correcto a través de una reacción dependiente de ATP (Walter *et al.*, 2002). Considerando lo anterior y a pesar de no haber observado expresión aparente de TgPolB en éste vector, se seleccionó la construcción 4 (pColdI) para continuar con los ensayos de expresión y purificación de TgPolB, ya que este vector cuenta con el promotor del gen de choque en frío *cspA*, el cual se activa al disminuir la temperatura del cultivo de *E. coli*. Una vez activado, comienza a producir proteínas que ayudan a mejorar el plegamiento correcto de la PR, lo que le permite que se mantenga en forma soluble, por lo que es recomendado para proteínas difíciles de expresar en los vectores convencionales que utilizan un sistema T7. Además, aumenta la pureza de la proteína de interés ya que la expresión en frío reprime la expresión de proteínas basales del huésped (Takara BioInc, 2017). Debido a las características de la construcción en el vector de expresión pColdI previamente descritas, se esperaba que esta construcción favoreciera la sobreexpresión de TgPolB tanto en el proceso de inducción como en el de purificación.

8.3. Purificación

Basados en la bibliografía (**Tabla V**), se optó por trabajar con la construcción en el vector pColdI y la cepa BL21(DE3) (cepa más recomendada para pruebas iniciales) (Baneyx & Mujacic, 2004; Peleg *et al.*, 2012; Duong-Ly & Gabelli, 2014). Para continuar con los análisis de purificación correspondientes esperando obtener una mayor cantidad de proteína concentrada y a su vez, detectable para el método.

8.3.1 Purificación (Regulador con EDTA)

En esta ocasión fue posible observar una banda a la altura del peso molecular esperado de la proteína. No obstante, una vez calentada la muestra a 85°C durante 15 min en un segundo paso de purificación, la proteína se degradó. Esto puede explicarse debido a que el EDTA es un agente quelante; es decir, un secuestrador de metales pesados, que puede interferir con la actividad de la enzima al remover el agente cofactor, en este caso el magnesio (Bornhorst *et al.*, 2000). La falta de su cofactor, podría desestabilizar la enzima, haciéndola más susceptible a la degradación térmica. Por lo anterior, se optó por realizar nuevamente las purificaciones utilizando un nuevo regulador de lisis sin EDTA (50 mM Tris pH 8.0 y 100 mM NaCl).

8.3.2 Purificación (Regulador sin EDTA)

El análisis del gel de poliacrilamida al 10% con las muestras obtenidas de la purificación en ÄKTA con regulador sin EDTA, permitió observar un patrón electroforético similar al de la proteína de interés en la elución a 100 mM de imidazol (**Figura 7A**), por lo que se procedió a concentrar la muestra en filtros de 10 kDa, para su análisis en Western-Blot. De la misma manera el análisis del gel de poliacrilamida 10% elaborado a partir de la purificación manual con regulador sin EDTA permitió distinguir claramente las bandas casi puras en las eluciones 1 y 2 (E1, y E2, respectivamente) con 500 mM de imidazol (**Figura 7B**).

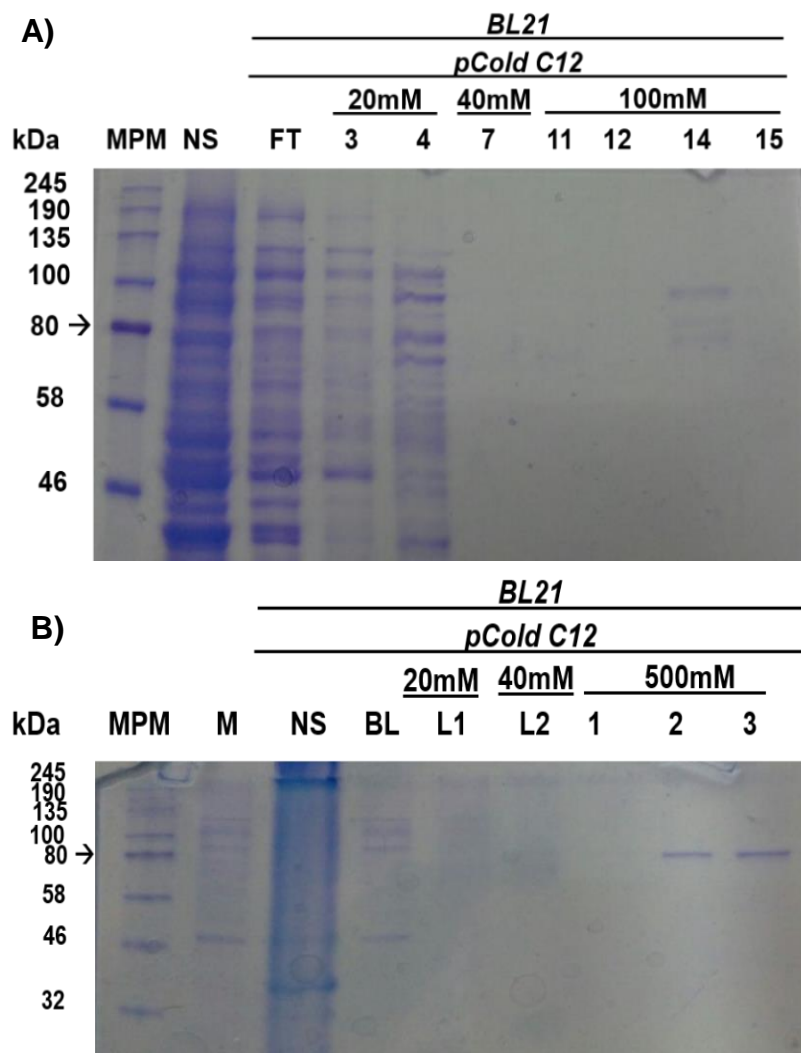


Figura 7. Purificación de TgPolB. A) Gel de poliacrilamida 10% con las muestras obtenidas de la purificación en ÄKTA (50 mM Tris pH 8.0 y 100 mM NaCl) de la construcción 4 en BL21. MPM= marcador de peso molecular; NS= fracción no soluble; FT= paso a través de la columna; 3,4= 20 mM imidazol; 7= 40 mM imidazol; 11-15= 100 mM imidazol. **B)** Gel de acrilamida al 10% con las muestras obtenidas de la purificación manual (50 mM Tris pH 8.0 y 100 mM NaCl) de C12 en pColdI en BL21. MPM= marcador de peso molecular; NS= fracción no soluble; FT= paso a través de la columna; BL= lavado con regulador de lisis, L1= lavado con 20 mM imidazol, L2= 40mM imidazol; 1, 2 y 3= 500 mM imidazol.

8.3.3 Purificación con detergentes

Uno de los pasos críticos para la obtención de grandes cantidades de proteína correctamente plegada, es el uso un sistema de expresión eficiente. Como se ha mencionado anteriormente, el sistema de expresión en *E. coli* ofrece una gran cantidad de ventajas. No obstante, el plegamiento de proteínas de gran tamaño no siempre resulta correcto, ya que grandes cantidades de proteína inactiva e insoluble suelen acumularse en agregados de partículas que se denominan cuerpos de inclusión (Rosano & Ceccarelli, 2014). Se ha descrito que sólo el 50% de las proteínas de eubacterias o arqueas y un 10% de las de eucariotas pueden ser expresadas de manera soluble en *E. coli*, disminuyendo considerablemente la probabilidad de éxito para la expresión de proteína soluble a pesos moleculares por encima de 80 kDa (Gräslund *et al.*, 2008). Dado que TgPolB tiene un peso molecular esperado de 89.9 kDa, existe la posibilidad de que la proteína no se esté expresando de forma soluble debido a un plegamiento incorrecto y por lo tanto podría estarse precipitando dentro de *E. coli* a través de la formación de cuerpos de inclusión. La formación de los cuerpos de inclusión depende de diversos factores como la secuencia de la proteína, la temperatura o la velocidad de crecimiento. A pesar de que no es posible predecirles, estos pueden removerse a través de una purificación selectiva con el uso de detergentes (Lara, 2011). Por lo anterior, y debido a la ausencia de TgPolB en las condiciones de expresión y purificación previamente analizadas, se optó por probar con el uso de detergentes iónicos (Sarcocyl 5%) y aniónicos (Tritón X-100) sin éxito aparente.

Como se puede observar en la **Figura 8**, el uso de Sarcosyl durante el proceso de purificación, permitió extraer una mayor cantidad de proteínas contaminantes en comparación a Tritón X-100. No obstante, en ninguna de las purificaciones realizadas con el uso de detergentes fue posible detectar la expresión de TgPolB.

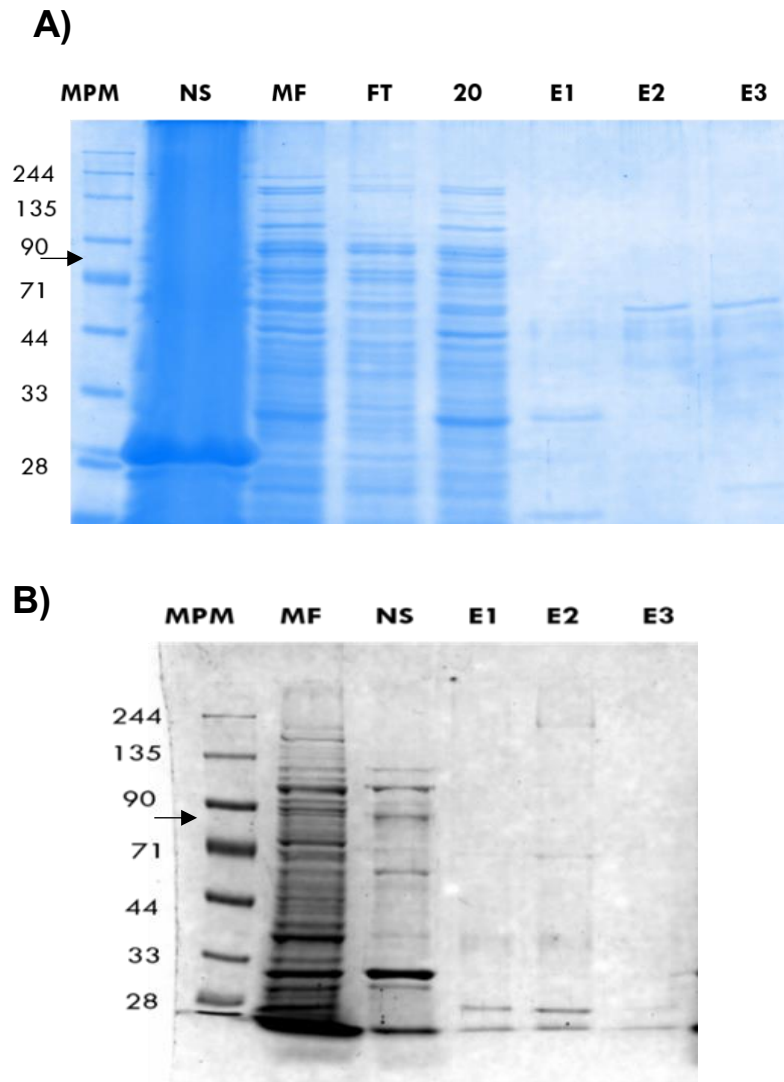


Figura 8. Purificaciones con detergentes. A) Purificación manual utilizando Tritón X-100 0.1% como detergente. MPM= marcador de peso molecular; NS= fracción no soluble; MF= muestra filtrada; FT= paso a través de la columna; 20= lavado con 20 mM imidazol; E1, 2 y 3= eluciones con 500 mM imidazol. **B)** Purificación manual utilizando Sarcosyl 5% como detergente. MPM= marcador de peso molecular; NS= fracción no soluble; MF= muestra filtrada; FT= paso a través de la columna; E1, 2 y 3= eluciones con 500 mM de imidazol.

8.4 Cuantificación de proteína

La intensidad de color observada en los geles de poliacrilamida de las muestras comparada con las obtenidas a partir de las concentraciones de albúmina conocidas (**Figura 9**), indicó una concentración aproximada de 1 mg/μl y 0.5 mg/μl para las fracciones E1 y E2 respectivamente, obtenidas a partir de la elución de la purificación manual sin EDTA; así como una concentración de 1 mg/μl para la fracción E3 obtenida a partir de la elución de la purificación en ÄKTA. Con los resultados obtenidos se realizó la estimación de las concentraciones molares de las fracciones, correspondiendo a 1.6 mM y 0.8 mM para las fracciones E1 y E2 respectivamente, y de 1.6 mM para la fracción E3.

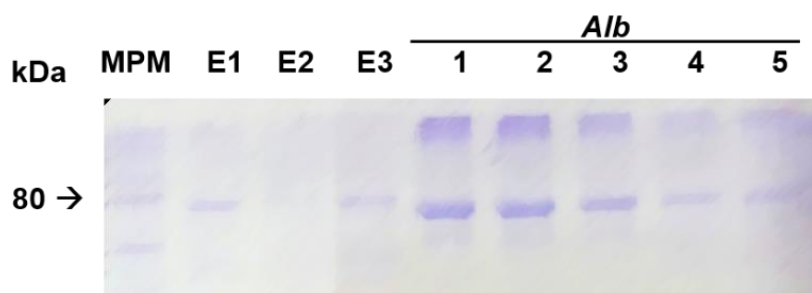


Figura 9. Cuantificación de proteínas en gel. Gel de cuantificación. PM= marcador de peso molecular; E1 y E2= eluciones de purificación manual con 500 mM imidazol; E3= elución de purificación en ÄKTA con 100 mM imidazol; concentraciones de albúmina: 1= 8 mg; 2= 4 mg; 3= 2 mg; 4= 1 mg; 5= 0.5 mg; 6=0.25 mg y 7=0.125mg.

8.5 Análisis de actividad de TgPolB por medio de PCR

Como puede observarse en las **Figuras 10**, no fue posible detectar actividad enzimática de tipo polimerasa ni exonucleasa a partir de las fracciones obtenidas de la purificación de TgPolB. No obstante, la falta de actividad podría deberse a que la proteína TgPolB se encuentre de forma inactiva debido a plegamientos inespecíficos o una síntesis incorrecta (Rosano & Ceccarelli, 2014). Por lo que se decidió realizar un análisis de tipo Western-Blot de las muestras para determinar la presencia de la proteína de interés.

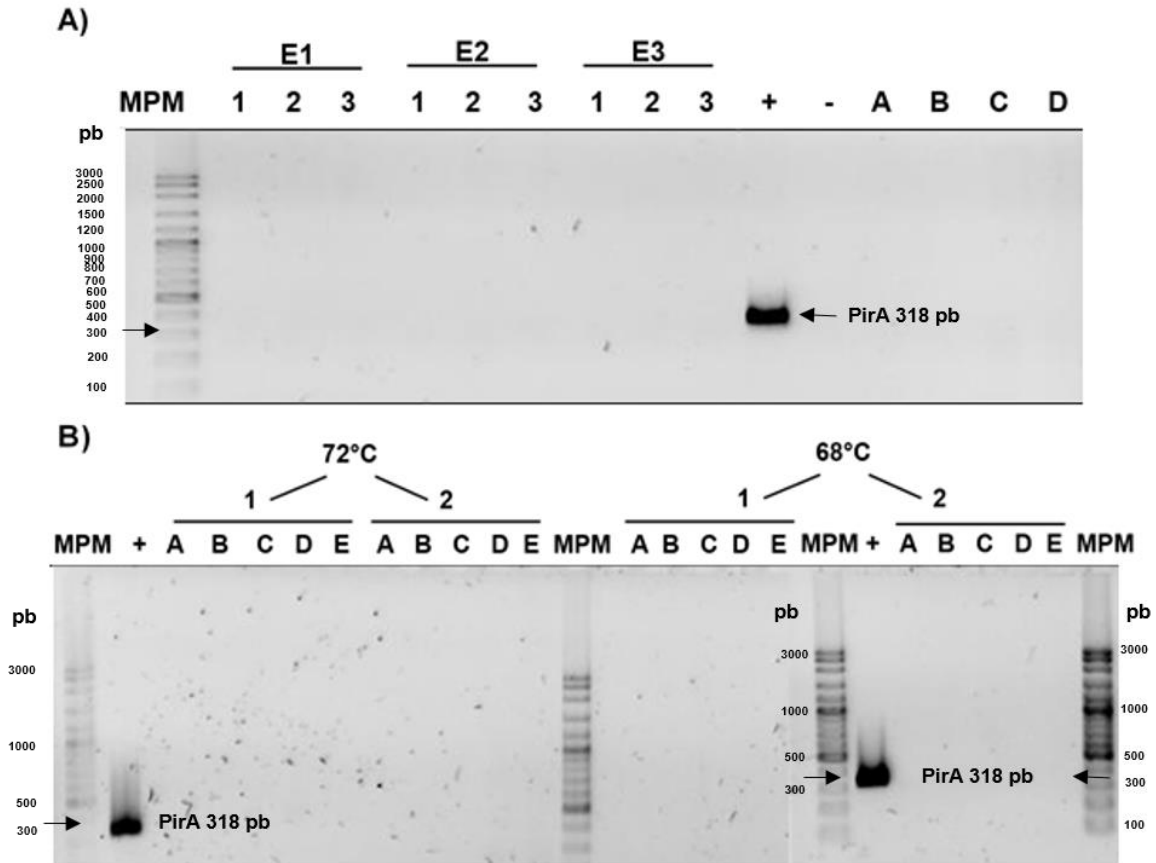


Figura 10. Análisis de actividad enzimática de TgPolB. **A)** Gel de agarosa al 1%. E1 y E2 = eluciones de C4 pColdI obtenidas por purificación manual. E3= elución de C4 pColdI obtenida de purificación con ÄKTA; += control positivo; -= control negativo. Se probaron las muestras con las siguientes concentraciones: E1 y E3, 1= 0.064 mM; 2= 0.16 mM; 3 = 0.32 mM. E2, 1= 0.032 mM; 2= 0.08 mM; 3 =0.16 mM. Además se probó la actividad exonucleasa incubando la enzima TgPolB de ambas purificaciones con una concentración conocida de ADN: A= E1; B= E2; C= E3; D= Control. **B)** Gel de agarosa 1%. 1 = Concentrado de eluciones de C12 pColdI obtenidas de la purificación manual; 2 = Elución de C4p ColdI obtenida de purificación en ÄKTA. Reguladores utilizados: A= S-TaqVivantis; B= PhusionGC; C= Phusion HF; D= Pfu-Altanzymes (Sin Mg⁺²)*; E= A-TaqVivantis (Sin Mg⁺²)*; += control positivo de Taq DNA pol. *A los reguladores sin magnesio se les añadió el magnesio durante la preparación del PCR.

8.6 Análisis de Western-Blot

Los resultados del análisis por Western-Blot de las fracciones obtenidas de la purificación manual y con ÄKTA (1 y 2 respectivamente), no demostraron la presencia de la etiqueta de histidinas. Por lo anterior, se colectaron alícuotas de 1 ml a lo largo de los diferentes puntos clave del proceso de purificación con la finalidad de detectar si la proteína se estaba degradando en alguno de estos pasos. Los resultados obtenidos de este último análisis detectan una banda a la altura entre 28 y 36 kDa, no obstante, este peso molecular no corresponde al de la proteína de interés (89.9 kDa) por lo que ésta no fue identificada en ninguno de los análisis.

Esta ausencia de señal podría deberse a un plegamiento incorrecto de la proteína que altere su estructura terciaria de manera que la etiqueta de histidinas se oculte, lo que impediría el reconocimiento de ésta por parte del anticuerpo (Bornhorst *et al.*, 2000). Otra posible explicación, podría ser la degradación de la proteína debido a un mal manejo de las muestras, no obstante, esta posibilidad se descartó al analizar paso a paso el proceso de purificación sin detectar la presencia de TgPolB en ninguno de ellos, incluido la expresión.

8.7 Análisis bioinformático de TgPoIB

8.7.1 Modelo estructural *in silico*

Dentro de la secuencia fue posible ubicar los dominios EXO I, EXO II y EXO III. Dominios de la Palma A, Motivos A, B, C y V, Motivo Pol III, la región N-terminal, y el sitio de unión a ADN (Figura 11).

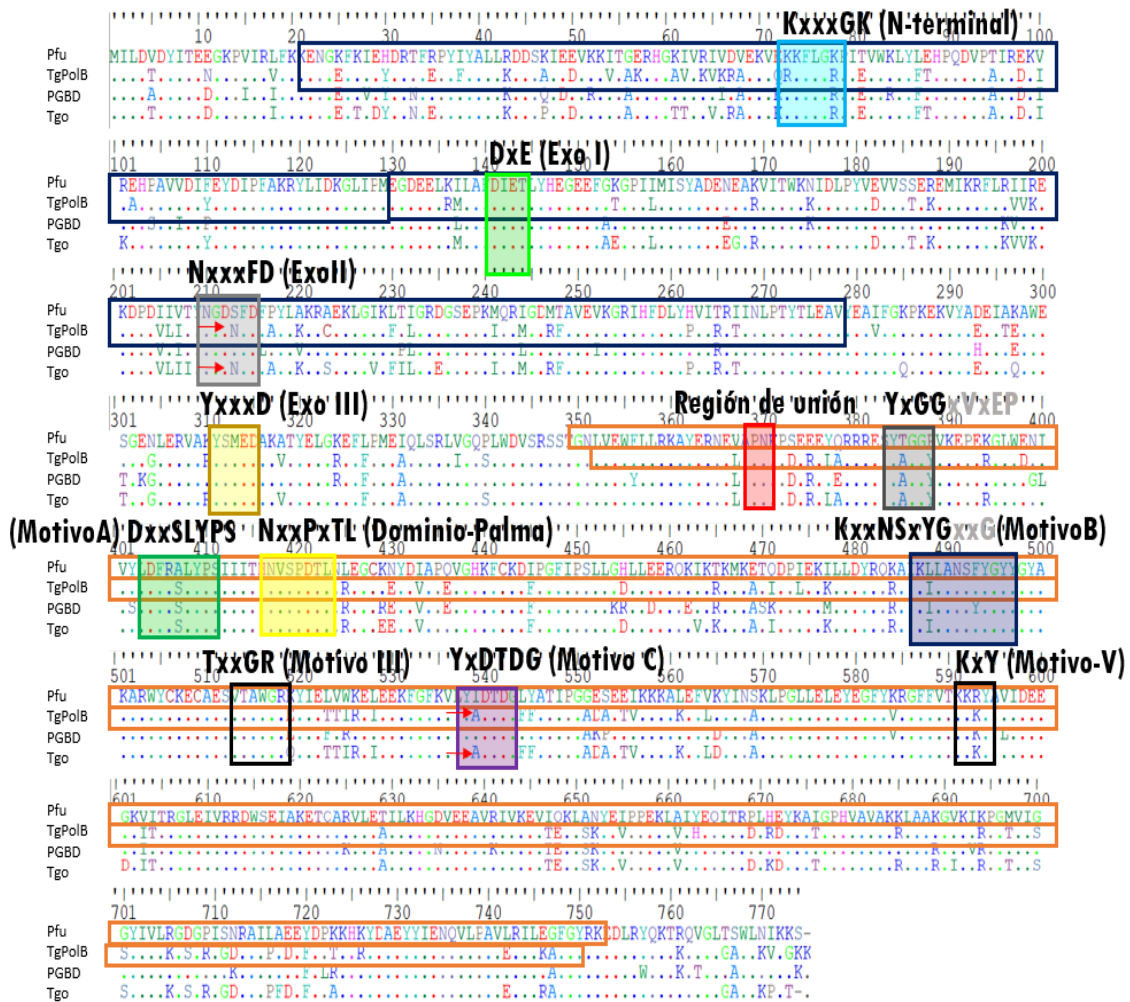


Figura 11. Comparación de secuencias de ADN polimeras comerciales obtenidas a partir de *Thermococcales* y TgPoIB. En recuadros azules se marca el dominio Exo y en naranja el multifuncional. En colores se indican los principales dominios y motivos identificados dentro de la secuencia. Pfu = *Pyrococcus furiosus*; PGBD = *Pyrococcus furiosus* Deep Vent; Tgo = *Thermococcus gorgonarius*; TgPoIB = *Thermococcus gammatolerans*. Información obtenida de patente WO2007043769A1-2007M (Bordo et al., 1991; Zhang et al., 2015).

Los aminoácidos identificados como característicos de estos sitios se ubicaron en regiones que concuerdan estructuralmente con los reportados para otras ADN pol B (**Figura 12**).

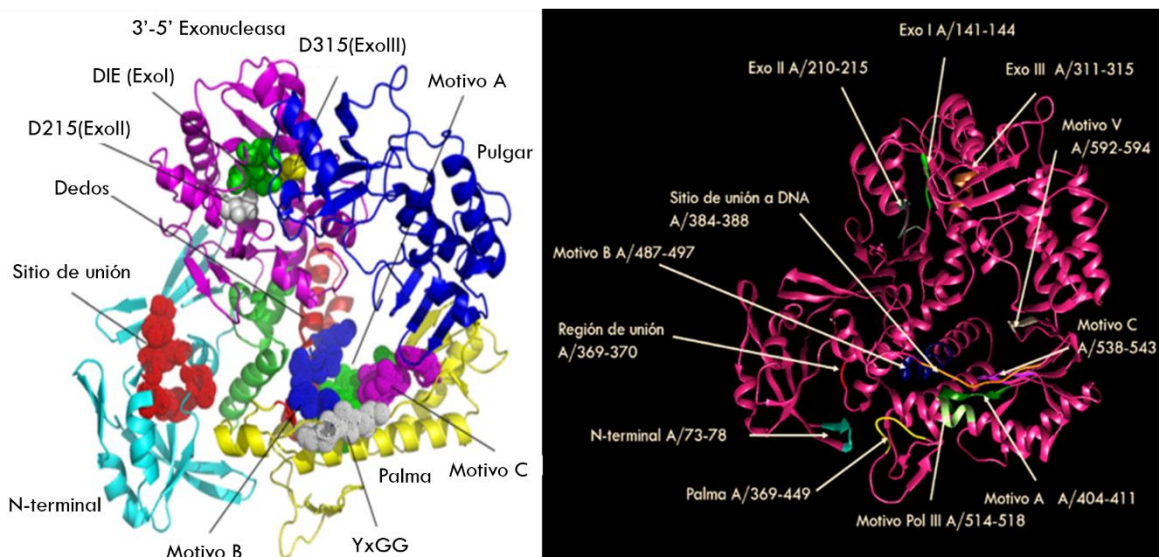


Figura 12. Análisis in silico de TgPolB. A) Estructura cristalográfica de la ADN pol B de Tgo: *Thermococcus gorgonarius* (Zhang *et al.*,2015). B) Modelo estructural de TgPolB elaborado en el software Chimera (Pettersen *et al.*,2004).

8.7.2. Análisis de identidad de secuencias

Se observó que todas presentan un elevado porcentaje de identidad entre 79.72 y 92.37%, encontrando los valores más bajos entre las especies de distinto género (**Tabla XIV**).

Tabla XIV. Porcentaje de identidad entre ADN polB de *Thermococcales* utilizadas en la actualidad de manera comercial.

Especie	Nombre	KOD1	Tgo	TgPolB	Pfu	PGBD
<i>T. kodakarensis</i>	KOD1	100				
<i>T. gorgonarius</i>	Tgo	92.24	100			
<i>T. gammatolerans</i>	TgPolB*	92.12	92.37	100		
<i>P. furiosus</i>	Pfu	79.72	80.08	80.49	100	
<i>Pyrococcus GBD</i>	PGBD	81.40	83.05	82.95	85.53	100

*Enzima objeto de estudio del presente trabajo.

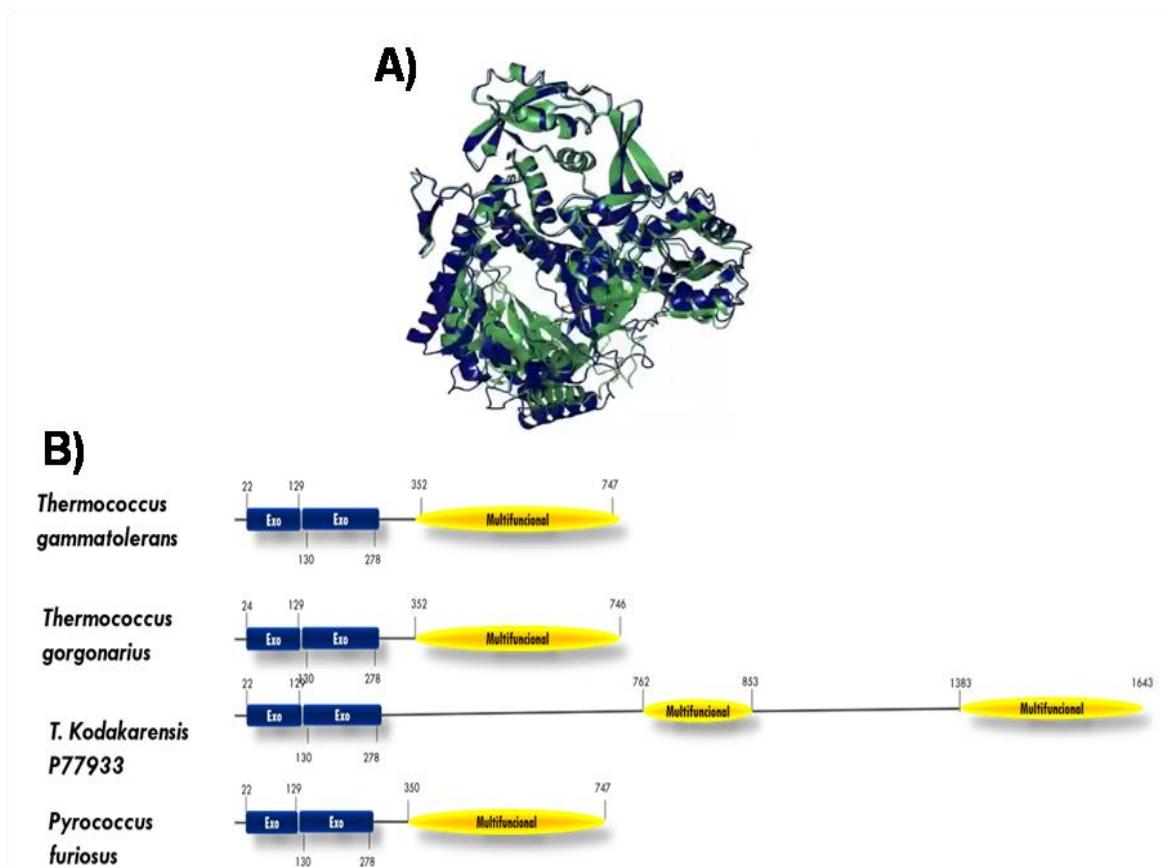


Figura 13. Comparación de modelos estructurales y dominios en ADN pol B de Thermococcales. **A)** Comparación de modelos estructurales de la ADN pol de *T. kodakarensis* (azul) y *T. gorgonarius* (verde), obtenidos utilizando el programa CCP4MG. **B)** Ubicación de los dominios exo y multifuncional en las ADN pol B de Thermococcales. Ubicación de los dominios exo (azul) y multifuncional (amarillo) de las ADN polimerasas (Familia B) de Thermococcales utilizadas de manera comercial, en comparación con la ADN polimerasa (Familia B) de *T. gammatolerans*. Se observa que todas presentan una ubicación similar a excepción de la ADN polimerasa de *T. kodakarensis*. Estas diferencias en la ubicación del dominio multifuncional, se explican por la presencia de intrones dentro de la secuencia. Toda la información se obtuvo por medio de un análisis de las secuencias en InterPro (Hunter *et al.*, 2009).

Del mismo modo, estas ADN pol cuentan con una alta similitud estructural (Beard y Wilson, 2003) (**Figura 13A**) y todas presentan los dominios principales que les caracterizan (**Figura 13B**). No obstante, como es posible observar en la **Tabla VI**, pese al alto nivel de identidad (Ej. Tgo y KOD1 92.24% de identidad), la actividad de estas enzimas es muy diferente. Por lo anterior, se considera que las ADN pol derivadas de este grupo de microorganismos extremófilos, suelen presentar

propiedades particulares dentro de esas diferencias en secuencia, que les permiten catalizar reacciones químicas en condiciones rigurosas como las que se requieren en algunas técnicas de biología molecular (Sarmiento *et al.*,2015), despertando así el interés de los investigadores por las arqueas como recurso potencial para la obtención de nuevas ADN pol.

8.7.3 Ubicación del N-terminal

El N-terminal de las enzimas ADN pol B de *T. gammatolerans* (TgPolB) y *T. gorgonarius* (Tgo), se ubicó entre los aminoácidos 74-79. Al comparar su ubicación con el potencial electrostático del modelo estructural se observó esta región expuesta en ambas enzimas analizadas, indicando así accesibilidad al sitio catalítico (**Figura 14**).

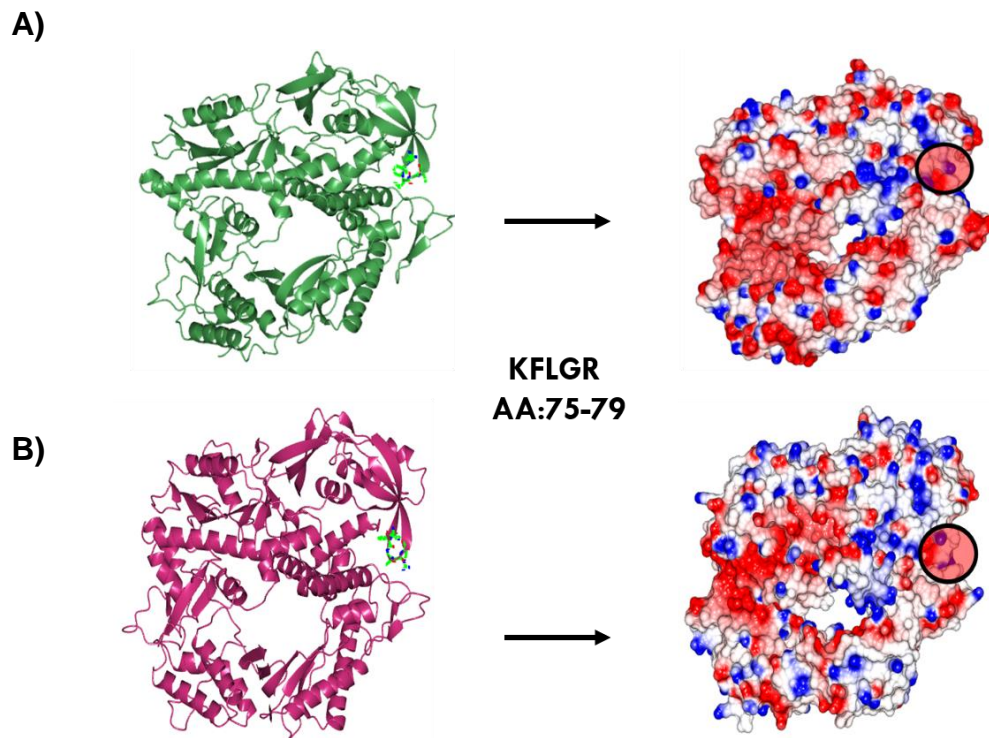


Figura 14. Ubicaciones de la región N-terminal en los modelos estructurales de ADN pol B de, A) Tgo: *T. gorgonarius* y B) TgPolB: *T. gammatolerans*. Los modelos estructurales así como las imágenes del modelo de potencial electrostático se obtuvieron con el software CCP4MG (McNicholas *et al.*,2011).

8.7.4 Sitios Activos

El programa en que se analizaron los sitios activos calculó 46 posibles cavidades de las cuales; posterior al análisis comparativo con el modelo estructural de la ADN pol B de *T. gorgonarius* (Tgo) y su sitio de unión al ADN, cinco resultaron positivas (**Figura 15**). Las cinco cavidades positivas se unieron para conformar una sola cavidad, ya que se considera como sitio activo no sólo el lugar de unión sino también las cadenas laterales de los aminoácidos que podrían encontrarse participando de forma activa en el proceso catalítico. Los resultados obtenidos de esta comparación estructural, demostraron una alta similitud entre el sitio activo predicho para TgPolB y el sitio activo descrito previamente para Tgo (Firbank *et al.*, 2008). La estructura del sitio activo suele utilizarse para orientar de forma óptima al sustrato, por lo que la alta similitud entre el sitio activo de TgPolB y Tgo predice que el complejo enzima-sustrato puede formarse de manera correcta (McKee & McKee, 2014).

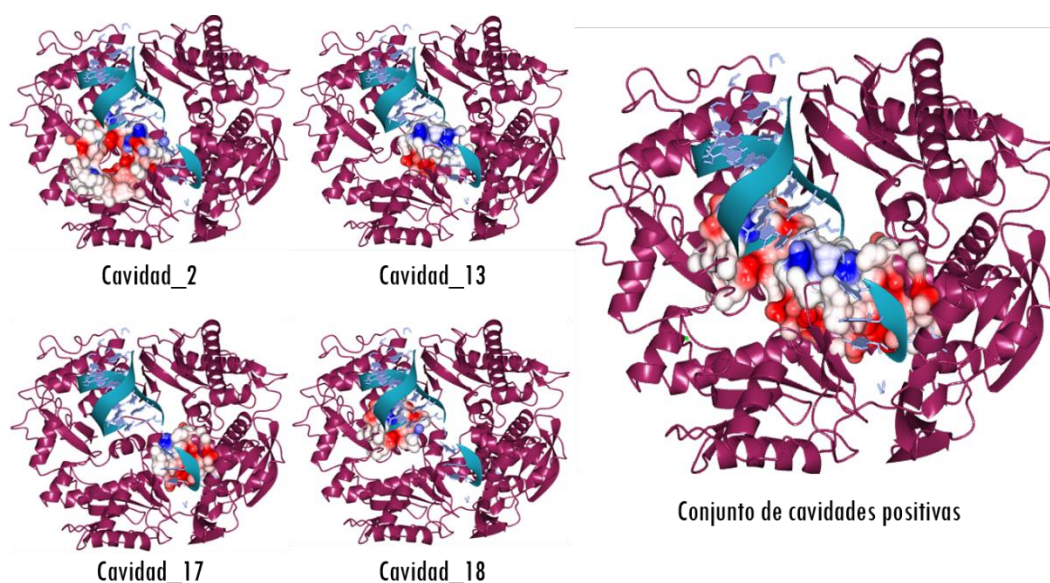


Figura 15. Sitios activos de TgPolB. Cavidades que resultaron positivas al análisis comparativo de los sitios activos de TgPolB con los sitios de unión al ADN de la ADN pol de *T. gorgonarius* (Tgo). Azul: cadena de ADN ubicada en el sitio de unión de la enzima Tgo. Morado: Enzima TgPolB. Potencial electrostático: Corresponde a los sitios activos de las diferentes cavidades que podrían encontrarse interactuando con el ADN y por lo tanto ser sitios activos.

8.7.5 Índice de adaptación de codones

El análisis del uso preferencial de codones en *E. coli* dio como resultado un valor de IAC de 0.57% para la secuencia no optimizada, y un valor de 0.69% para la secuencia optimizada. Cuando la frecuencia de ocurrencia de los codones sinonimos del ADN codificante es significativamente distinta a la del hospedero se dice que existen diferencias en el uso preferencial de codones. Entre mayor son las diferencias, incrementa la probabilidad de obtener una baja expresión o pérdida de actividad debido al agotamiento de los ARNt de baja abundancia, lo cual provoca la incorporación incorrecta y/o truncamiento de los polipéptidos (Rosano & Ceccarelli, 2014). Por lo anterior, el valor obtenido del IAC de la secuencia no optimizada que se utilizó en el presente trabajo (0.57), indica que ésta podría ser una de las causas de la baja o nula expresión de TgPolB en *E. coli*.

Tabla XV. Análisis del uso preferencial de codones de TgPolB en *E. coli*

	Análisis de uso preferencial de codones en <i>E.coli</i>					
	IAC		Contenido de GC		Elementos negativos de CIS	Elementos negativos repetidos
	Valor	Valor ideal	Valor	Valor ideal		
No optimizada	0.57	0.8-1.0	55.49	30-70%	9	0
Optimizada	0.69	0.8-1.0	46.8	30-70%	8	0

Existen dos estrategias que permiten resolver esta situación: 1) la optimización de la secuencia y 2) incrementar la disponibilidad de los ARNt a través de una modificación del hospedero usando cepas ricas en “codones raros”. La optimización de la secuencia consiste en modificar los codones sinónimos para ajustarlos al uso preferencial de codones del huésped de manera que se incremente el IAC, requiriendo la selección de la mejor secuencia, entre un gran número de combinaciones posibles; Por su parte, el uso de cepas ricas en codones raros permite incrementar la disponibilidad de estos codones a través de modificaciones en el huésped (Sorensen & Mortensen, 2005).

Si bien, la secuencia utilizada no fue optimizada, el uso preferencial de codones se intentó favorecer con el implemento de la cepa RII, la cual se caracteriza por presentar “codones raros”. No obstante, tampoco fue posible la expresión de TgPolB en esta cepa. Por lo anterior, se propone el uso de cepas que presenten una mayor variabilidad en el uso de codones, tales como las líneas celulares de “Codon-Plus” (Gopal & Kumar, 2013).

8.7.6 Enlaces disulfuro

El análisis predictivo de la formación de enlaces disulfuro y su conectividad obtenido por medio del servidor en línea DISULFIND (Ceroni *et al.*,2006),detectó cinco residuos de cisteína (aminoácidos 221, 426, 440, 504 y 507), que presentan un alto potencial de conformar enlaces disulfuro (**Figura 16**). Por su parte, el análisis elaborado en DiANNA 1.1 (Ferrè *et al.*,2006) detectó una elevada probabilidad de enlaces disulfuro entre los residuos de cisteína de Cis223-Cis508 y Cis442-Cis509.

Estos enlaces disulfuro concuerdan parcialmente con los reportados en la bibliografía, lo cual sugiere que son fundamentales para la estabilidad de la proteína (Hopfner *et al.*,1999). Además, se detectó otro residuo de cisteína Cis221 con un elevado potencial para la formación de enlace disulfuro. Para muchas proteínas recombinantes, la formación correcta de estos enlaces es vital para la conformación tridimensional de su forma biológicamente activa (Messens & Collet, 2016). La incapacidad de *E. coli* para la formación correcta de enlaces disulfuro podría ser un impedimento para la conformación tridimensional de TgPolB ya que existe la posibilidad de plegamientos incorrectos, así como de agregación en cuerpos de inclusión que imposibiliten la detección de la proteína (Baneyx & Mujacic, 2004).

Una de las razones por la cual resulta complicada la expresión de proteínas ricas en enlaces disulfuro en *E. coli* es debido su formación se da en el periplasma, mientras que el citoplasma presenta potencial redox más negativo y se mantiene en ambiente reductor por el sistema tiorredoxina-tiorredoxina reductasa (trxB) y el sistema glutarredoxin-glutarredoxin reductasa (gor) (Stewart *et al.*, 1998). Para solucionar este problema Rosano & Ceccarelli (2014), proponen dirigir la proteína

al periplasma, a través de la fusión de la proteína recombinante a un péptido líder. También sugieren continuar la expresión en el citoplasma haciendo uso de las cepas modificadas de *E. coli*, las cuales contienen un ambiente citoplasmático oxidativo que favorece la formación de los enlaces disulfuro (Ej. cepas Origami (Novagen) y SHuffle (NEB)).

Por lo anterior y considerando los resultados obtenidos a lo largo del presente estudio, se eligió el vector pET32a como alternativa para la expresión de TgPolB, esperando que el dominio de tiorredoxina presente en este vector, permita mejorar el plegamiento de la proteína.

8.8 Subclonación en pET32a

Se intentó realizar la subclonación del inserto en un nuevo vector de expresión pET32a, para mejorar la formación de enlaces disulfuro y el plegamiento correcto de la proteína de interés, sin embargo, ésta no pudo realizarse de manera exitosa. Lo anterior debido a que mientras se realizaba la restricción con las enzimas de corte correspondientes (BamHI y NdeI) para la liberación del inserto, se observó un vector linealizado en presencia de ambas enzimas. Del mismo modo, la reacción que presentaba sólo NdeI mostró la formación de un vector linealizado; mientras que en presencia de BamHI el vector no presentó ningún corte, observándose de manera similar al control (vector sin cortar). Indicando así que la enzima BamHI no fue capaz de reconocer el sitio de corte (**Figura 17**). Este problema se observó en todas las construcciones independientemente del vector, por lo que se procedió a secuenciar nuevamente el gen, lo que permitió detectar la mutación en el sitio de restricción de BamHI (**Figura 18**).

```

.....10.....20.....30.....40.....50.....60.....70.....
AA      MILDTDYITENKPKVIRVFKKENGEFKIEYDRTFEPYFYALLKDDSAIEDVKKVTAKRHGAVVKKRAEKVQRKFLGRP
DB_state
DB_conf

80.....90.....100.....110.....120.....130.....140.....150.....
AA      IEVWKLYFTHPQDVPAIRDKIRAHPAVVDIYEYDIPFAKRYLIDKGLIPMEGDEELRMLAFDIETLYHEGEEFGTGPIL
DB_state
DB_conf

.160.....170.....180.....190.....200.....210.....220.....230.....
AA      MISYADENEARVITWKKIDLPYVDVSTEKEMIKRFLRVVKEKDPDVLITYNGDNFDFAYLKKRCEKLGKFTLGRDGS
DB_state                                     0
DB_conf                                     9

..240.....250.....260.....270.....280.....290.....300.....310....
AA      EPKIQRMGDRFAVEVKGRIFDLYPVIRRTINLPYTL EAVVEAVFGPKPEKVYAE EITEAWESGEGLERVARYSMEDA
DB_state
DB_conf

...320.....330.....340.....350.....360.....370.....380.....390...
AA      KVTYELGREFFPMEAQLSRLIGQSLWDVSRSTGNLVEVFLLRKAYERNELAPNKPDERELARRRESYAGGVYKEPERG
DB_state
DB_conf

...400.....410.....420.....430.....440.....450.....460.....470..
AA      LWDNIVYLDFRSLYPSIIITHNVSPDTLNREGCKEYDVAPEVGHKFKCKDFPGFIPSELLGDLEERQKIKRKMKATIDPL
DB_state                                     0 0
DB_conf                                     9 9

.....480.....490.....500.....510.....520.....530.....540.....550.
AA      EKKLLDYRQRAIKILANSFYGYGYAKARWYCKECAESVTANGREYIETTIREIEEKFGFKVLYADTDGFFATIPGADA
DB_state                                     0 0
DB_conf                                     9 9

.....560.....570.....580.....590.....600.....610.....620.....630
AA      ETVKKKAKEFLKYINAKLPGLLELEYEGFYVRGFFVTKKKYAVIDEEGKITTRGLEIVRRDWSEIAKETQARVLEAILK
DB_state
DB_conf

.....640.....650.....660.....670.....680.....690.....700.....710
AA      HGDVEEAVRIVKEVTEKLSKYEVPEKLVIIHEQITRDLRDKATGPHVAVAKRLAAKGVKIRPGTVISYIVLKGSGRIG
DB_state
DB_conf

.....720.....730.....740.....750.....760.....770...
AA      DRAIPADEFDPTKHRYDAEYYIENQVLP AVERILKAFGYRKEDLRYQTKQVGLGAWLKVKGKK
DB_state
DB_conf

```

Figura 16. Análisis predictivo de enlaces disulfuro y predictor de conectividad elaborado en el servidor en línea DISULFIND. AA: Secuencia de aminoácidos; DB_state: Estado predictivo del enlace disulfuro (1= Con enlace disulfuro, 0= Sin enlace disulfuro); DB_conf: confianza de la predicción (0= baja, 9= alta).

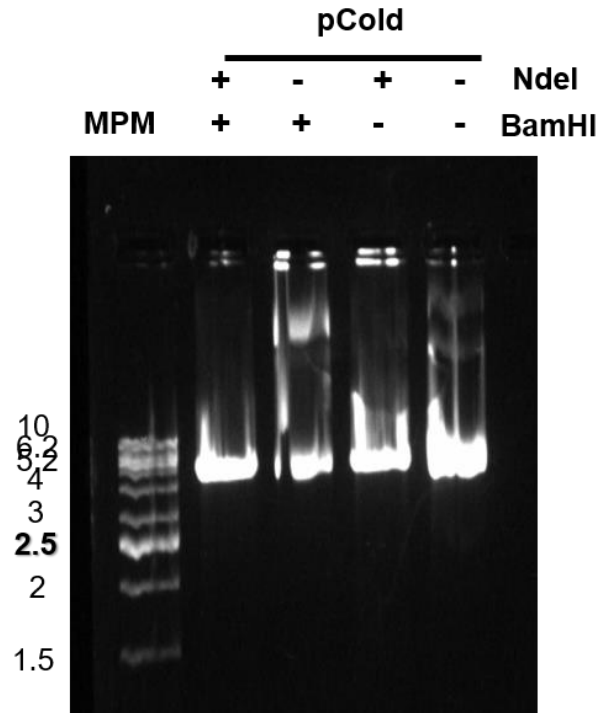


Figura 17. Ensayo de doble restricción de la construcción 4 en pColdI utilizando las enzimas de restricción BamHI y NdeI. Gel de agarosa 1%. += presencia; -= ausencia; MPM= marcador de peso molecular.

```

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
          2260      2270      2280      2290      2300
TgPolB_EJ3  CGGAAGGAGG ATTGAGGTA CCAGAAGACG AAGCAGGTCG GTTTGGGCGC
C4pCold    CGGAAGGAGG ATTGAGGTA CCAGAAGACG AAGCAGGTCG GTTTGGGCGC

.....|.....| .....|.....| .....|.....| .....|
          2310      2320      2330
TgPolB_EJ3  GTGGCTGAAG GTGAAAGGGA AGAAGTGAGG ATCC
C4pCold    GTGGCTGAAG GTGAAAGGGA AGAAGTGAGG ATCG

```

Figura 18. Mutación en sitio de restricción. TgPolB_EJ3: Secuencia de TgPolB obtenida de gene bank. C4pCold: Resultado de la secuenciación de la construcción en pCold Azul. Sitio de restricción de la enzima. Amarillo: mutación.

9. CONCLUSIONES Y RECOMENDACIONES

Los resultados obtenidos a partir de la expresión en las cepas de Rosetta II y BL21 (DE3) con diferentes condiciones de inducción, no muestran expresión con ninguna de las construcciones utilizadas, independientemente del medio de cultivo (Luria-Bertani, Super broth, 2XYT), temperatura (35°C y 16°C), tiempo de inducción (4, 16 hrs) o concentraciones de inductor utilizados (0.25, 0.5, 1 y 1.25 mM).

Por su parte, los ensayos de purificación realizados para incrementar la concentración de la proteína expresada no permitieron la obtención de TgPolB. Del mismo modo, las purificaciones con uso de detergentes permitieron extraer proteínas contaminantes de las muestras, sin embargo, en ninguna fue posible detectar la presencia de TgPolB.

Debido a lo anterior, se concluye que las estrategias de expresión y purificación propuestas para la producción recombinante de TgPolB no resultaron efectivas para esta enzima bajo estas condiciones, a pesar de tratarse de una proteína con un elevado porcentaje de identidad de secuencias y estructuralmente muy similar a enzimas homólogas. Por consiguiente, se rechaza la hipótesis; ya que, el uso de sistemas de expresión similares a los empleados para la producción recombinante de enzimas homólogas no es recomendable como principio para la selección de estrategias a utilizar en la producción recombinante de TgPolB.

El análisis bioinformático de TgPolB predice a través del modelo *in silico* que la enzima cuenta con los dominios característicos de las ADN polimerasas B, necesarios para su actividad. No obstante, la ausencia de expresión de la proteína con las diferentes estrategias utilizadas, así como la elevada probabilidad de enlaces disulfuro entre los residuos de cisteína de Cis223-Cis508 y Cis442-Cis509, y el residuo Cis221, sugieren como una de las posibles causas la formación incorrecta de estos enlaces.

Por su parte, el bajo índice de adaptación de codones de TgPolB en *E. coli* y el tamaño de la proteína (89.9 kDa), sugieren como otra posible explicación el alto requerimiento energético que implica para *E. coli* la producción de esta proteína en

ausencia de codones “raros”. Por lo anterior, se propone el uso de nuevas cepas de expresión “codon plus” enriquecidas con este tipo de codones.

La mutación en el sitio de restricción de la enzima no explica la ausencia de expresión, no obstante, es probable que existan otras mutaciones que aún no han sido detectadas que pudieran estar afectando la expresión (alteraciones en el promotor o en los reguladores de expresión). Por lo anterior, se recomienda realizar una secuenciación completa de las construcciones y no sólo del gen a expresar.

Finalmente, para los futuros estudios de expresión de TgPolB se sugieren los siguientes pasos: 1) Hacer uso de otras cepas de expresión “codón plus”; 2) Optimización del gen para el ajuste del uso preferencial de codones en *E. coli*; 3) Subclonación preferentemente simultánea en diferentes vectores de expresión que favorezcan el plegamiento correcto de la proteína de interés y eviten la formación de cuerpos de inclusión; y 4) Siendo TgPolB una proteína de elevado peso molecular (89.9 kDa) y con un índice preferencial de codones muy distinto al de *E. coli*, se recomienda como alternativa la expresión de TgPolB en otro sistema como la levadura *Pichia pastori*.

11. BIBLIOGRAFÍA

- Adams, Michael W.,W; KELLY, Robert M. Finding and using hyperthermophilic enzymes. *Trends in Biotechnology*, 1998, **vol.** 16, no 8, 526 pp.
- Almanza, M., N., 2018: Interacción de la toxina binaria PirAB de *Vibrio parahaemolyticus* con proteínas hepatopancreáticas de camarón blanco (*Litopenaeus vannamei*). Tesis de maestría. Instituto Politécnico Nacional. Centro Interdisciplinario de Ciencias Marinas, La Paz B.C.S.
- Amid, A., 2015: Introduction to Recombinant Enzyme Pre-commercialization. *Recombinant Enzymes - From Basic Science to Commercialization*. Springer International Publishing, Cham, pp. 1–11.
- Bakermans, 2015. Microbial evolution under extreme conditions. Walter de Gruyter GmbH & Co KG, Corien (ed.).
- Baneyx, F.; Mujacic, M., 2004, November: Recombinant protein folding and misfolding in *Escherichia coli*. *Nature Biotechnology*.
- Beard, W. A.; Wilson, S. H. Structural insights into the origins of DNA polymerase fidelity. *Structure (London, England: 1993)*, 2003, vol. **11**, no 5, p. 489
- Bordo, D.; Argos, P., 1991: Suggestions for 'safe' residue substitutions in site-directed mutagenesis. *Journal of Molecular Biology.*, **217**, 721–729.
- Bornhorst, J. A.; Falke, J. J., 2000: Purification of proteins using polyhistidine affinity tags. *Methods in enzymology.*, **326**, 245–254.
- Ceroni, A.; Passerini, A.; Vullo, A.; Frasconi, P., 2006: Disulfind: A disulfide bonding state and cysteine connectivity prediction server. *Nucleic Acids Research.*, **34**, W177–W181.
- Clark, D. P.; Pazdernik, N. J., 2016: Recombinant DNA Technology. *Biotechnology (Second Edition)*. Academic Cell, pp. 63–95.
- Cline, J., 1996: PCR fidelity of pfu DNA polymerase and other thermostable DNA polymerases. *Nucleic Acids Research.*, **24**, 3546–3551.
- Dabrowski, S. S.; Kur, J., 1998: Cloning and expression in *Escherichia coli* of the recombinant His-tagged DNA polymerases from *Pyrococcus furiosus* and *Pyrococcus woesei*. *Protein Expression and Purification.*, **14**, 131–138.

- Dietrich, J.; Schmitt, P.; Zieger, M.; Preve, B.; Rolland, J. L.; Chaabihi, H.; Gueguen, Y., 2002: PCR performance of the highly thermostable proof-reading B-type DNA polymerase from *Pyrococcus abyssi*. *FEMS Microbiology Letters.*, **217**, 89–94.
- Duong-Ly, K. C.; Gabelli, S. B., 2014: Explanatory chapter: Troubleshooting recombinant protein expression: General. *Methods in Enzymology.*, **541**, 209–229.
- Ferrè, F.; Clote, P., 2006: DiANNA 1.1: An extension of the DiANNA web server for ternary cysteine classification. *Nucleic Acids Research.*, **34**, W182–W185.
- Fina, B. L.; Lombarte, M.; Rigalli, A., 2013: Investigación de un fenómeno natural: ¿Estudios *in vivo*, *in vitro* o *in silico*? *Actualizaciones en Osteología.*, **9**, 294–299.
- Gardner F., A.; Kelman, Z., 2014: DNA polymerases in biotechnology. *Frontiers in microbiology.*, **5**, 3.
- Gopal, G. J.; Kumar, A., 2013: Strategies for the Production of Recombinant Protein in *Escherichia coli*. *The Protein Journal.*, **32**, 419–425.
- Goulding, C. W.; Perry, L. J., 2003: Protein production in *Escherichia coli* for structural studies by X-ray crystallography. *Journal of Structural Biology.*, **142**, 133–143.
- Gräslund, S.; Nordlund, P.; Weigelt, J.; Hallberg, B. M.; Bray, J.; Gileadi, O.; Knapp, S.; Oppermann, U.; Arrowsmith, C.; Hui, R.; Ming, J.; dhe-Paganon, S.; Park, H. won; Savchenko, A.; Yee, A.; Edwards, A.; Vincentelli, R.; Cambillau, C.; Kim, R. et al., 2008: Protein production and purification. *Nature Methods.*, **5**, 135–146.
- Hashemzadeh-Bonehi, L.; Mehraein-Ghomi, F.; Mitsopoulos, C.; Jacob, J. P.; Hennessey, E. S.; Broome-Smith, J. K., 1998, November: Importance of using lac rather than ara promoter vectors for modulating the levels of toxic gene products in *Escherichia coli*. *Molecular Microbiology.*

- Hopfner, K. P.; Eichinger, A.; Engh, R. A.; Laue, F.; Ankenbauer, W.; Huber, R.; Angerer, B., 1999: Crystal structure of a thermostable type B DNA polymerase from *Thermococcus gorgonarius*. *Proceedings of the National Academy of Sciences of the United States of America.*, **96**, 3600–3605.
- Hunter, S.; Apweiler, R.; Attwood, T. K.; Bairoch, A.; Bateman, A.; Binns, D.; Bork, P.; Das, U.; Daugherty, L.; Duquenne, L.; Finn, R. D.; Gough, J.; Haft, D.; Hulo, N.; Kahn, D.; Kelly, E.; Laugraud, A.; Letunic, I.; Lonsdale, D. et al., 2009: InterPro: The integrative protein signature database. *Nucleic Acids Research.*, **37**, D211–D215.
- Ishino, S.; Ishino, Y., 2013: DNA polymerases and DNA ligases. *Thermophilic Microbes in Environmental and Industrial Biotechnology: Biotechnology of Thermophiles*. Springer Netherlands, Dordrecht, pp. 429–457.
- Ishino, S.; Ishino, Y., 2014: DNA polymerases as useful reagents for biotechnology - the history of developmental research in the field. *Frontiers in microbiology.*, **5**, 465.
- Jana, S.; Deb, J. K., 2005: Strategies for efficient production of heterologous proteins in *Escherichia coli*. *Applied microbiology and biotechnology*, vol. 67, no 3, p. 289-298
- Jolivet, E.; L'Haridon, S.; Corre, E.; Forterre, P.; Prieur, D., 2003: *Thermococcus gammatolerans* sp. nov., a hyperthermophilic archaeon from a deep-sea hydrothermal vent that resists ionizing radiation. *International journal of systematic and evolutionary microbiology.*, **53**, 847–851.
- Kane, J. F. 1995. Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. *Curr. Opin. Biotechnol.* 6, 494–500
- Killelea, T.; Ralec, C.; Bossé, A.; Henneke, G., 2014: PCR performance of a thermostable heterodimeric archaeal DNA polymerase. *Frontiers in Microbiology.*, **5**, 1–11.

- Kim, K. P.; Cho, S. S.; Lee, K. K.; Youn, M. H.; Kwon, S. T., 2011: Improved thermostability and PCR efficiency of *Thermococcus celericrescens* DNA polymerase via site-directed mutagenesis. *Journal of Biotechnology.*, **155**, 156–163.
- Kim, Y. J.; Lee, H. S.; Bae, S. S.; Jeon, J. H.; Lim, J. K.; Cho, Y.; Nam, K. H.; Kang, S. G.; Kim, S. J.; Kwon, S. T.; Lee, J. H., 2007: Cloning, purification, and characterization of a new DNA polymerase from a hyperthermophilic archaeon, *Thermococcus* sp. NA1. *Journal of Microbiology and Biotechnology.*, **17**, 1090–1097.
- Lagorce, Arnaud, et al. Genome-wide transcriptional response of the archaeon *Thermococcus gammatolerans* to cadmium. *PloS one*, 2012, vol. 7, no 7, p. e41935.
- Lee, J. Il; Kil, E. J.; Song, J. G.; Kim, Y. J.; Choi, J. J.; Shim, H.; Kwon, S. T., 2009: Characterization and PCR optimization of the thermostable family B DNA polymerase from *Thermococcus guaymasensis*. *Enzyme and Microbial Technology.*, **45**, 103–111.
- Lomax, M. E.; Folkes, L. K.; O'Neill, P., 2013: Biological consequences of radiation-induced DNA damage: Relevance to radiotherapy. *Clinical Oncology.*, **25**, 578–585.
- Makarova, K. S.; Krupovic, M.; Koonin, E. V., 2014: Evolution of replicative DNA polymerases in archaea and their contributions to the eukaryotic replication machinery. *Frontiers in Microbiology.*, **5**, 1–10.
- Makrides, S. C., 1996: Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol Rev.*, **60**, 512–538.
- Mattila, P.; Korpela, J.; Tenkanen, T.; Pitkänen, K., 1991: Fidelity of DNA synthesis by the *Thermococcus litoralis* DNA polymerase--an extremely heat stable enzyme with proofreading activity. *Nucleic acids research.*, **19**, 4967–4973.
- McDonald, J. P.; Hall, A.; Gasparutto, D.; Cadet, J.; Ballantyne, J.; Woodgate, R., 2006: Novel thermostable Y-family polymerases: Applications for the PCR amplification of damaged or ancient DNAs. *Nucleic Acids Research.*, **34**, 1102–1111.

- McKee, Trudy; McKee, James Robert. *Biochemistry: the molecular basis of life*. New York: McGraw-Hill, 2003.
- McNicholas, S.; Potterton, E.; Wilson, K. S.; Noble, M. E. M., 2011: Presenting your structures: The CCP4mg molecular-graphics software. *Acta Crystallographica Section D: Biological Crystallography.*, **67**, 386–394.
- Messens, J., and Collet, J. F. (2006). Pathways of disulfide bond formation in *Escherichia coli*. *Int. J. Biochem. Cell Biol.* 38, 1050–1062.
- Moeller, R.; Reitz, G.; Li, Z.; Klein, S.; Nicholson, W. L., 2012: Multifactorial Resistance of *Bacillus subtilis* Spores to High-Energy Proton Radiation: Role of Spore Structural Components and the Homologous Recombination and Non-Homologous End Joining DNA Repair Pathways. *Astrobiology.*, **12**, 1069–1077.
- Nadeem, Tariq, et al. Glycosylation of Recombinant Anticancer Therapeutics in Different Expression Systems with Emerging Technologies. *Cancer research*, 2018, **vol. 78**, no 11, p. 2787-2798..
- Palomares, L. A.; Estrada-Mondaca, S.; Ramírez, O. T., 2004: Production of Recombinant Proteins: *Challenges and Solutions*. *Recombinant Gene Expression*. Humana Press, New Jersey, **Vol. 267**pp. 15–52.
- Peleg, Y.; Unger, T., 2012: Resolving bottlenecks for recombinant protein expression in *E. coli*. *Methods in Molecular Biology.*, **800**, 173–186.
- Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., 2004: UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry.*, **25**, 1605–1612.
- Powsner, R. A.; Powsner, E. R. *Essential nuclear medicine physics* 2. ed. *Massachussets, Estados Unidos: Blackwell*, 2006
- Lara, R. A., 2011: Recombinant protein production in *Escherichia coli*. *Revista Mexicana de Ingeniería Química.*, **10**, 209–223.
- Reed, C. J.; Lewis, H.; Trejo, E.; Winston, V.; Evilia, C., 2013: Protein adaptations in archaeal extremophiles. *Archaea.*, **2013**, 14.

- Rogozin, I. B.; Makarova, K. S.; Pavlov, Y. I.; Koonin, E. V., 2008: A highly conserved family of inactivated archaeal B family DNA polymerases. *Biology Direct.*, **3**.
- Rosano, G. L.; Ceccarelli, E. A., 2014: Recombinant protein expression in *Escherichia coli*: Advances and challenges. *Frontiers in Microbiology.*, **5**, 1–17.
- Rothwell, B. P. J.; Waksman, G., 2005: Structure and mechanism of DNA polymerase. *Advances.*, **71**, 401–440.
- Sahdev, Sudhir; Khattar, Sunil K.; SAINI, Kulvinder Singh. Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Molecular and cellular biochemistry*, 2008, **vol. 307**, no 1-2,
- Sánchez, Ángel Herráez. Texto ilustrado e interactivo de biología molecular e ingeniería genética StudentConsult en español: Conceptos, técnicas y aplicaciones en ciencias de la salud. Elsevier Health Sciences, 2012.
- Sarmiento, F.; Peralta, R.; Blamey, J. M., 2015: Cold and Hot Extremozymes: Industrial Relevance and Current Trends. *Frontiers in bioengineering and biotechnology.*, **3**, 148.
- Sorensen, H. P.; Mortensen, K. K. 2005. Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *Journal of biotechnology.* **vol. 115**, no 2
- Stevens, R. C., 2000, September: Design of high-throughput methods of protein production for structural biology. *Structure*.
- Stewart, E. J., Aslund, F., and Beckwith, J. (1998). Disulfide bond formation in the *Escherichia coli* cytoplasm: an in vivo role reversal for the thioredoxins. *EMBO J.* **17**, 5543–5550.
- Takagi, M.; Nishioka, M.; Kakiyama, H.; Kitabayashi, M.; Inoue, H.; Kawakami, B.; Oka, M.; Imanaka, T., 1997: Characterization of DNA polymerase from *Pyrococcus sp.* strain KOD1 and its application to PCR. *Applied and Environmental Microbiology.*, **63**, 4504–4510.
- Takara BioInc, 2017: pCold DNA cold-shock expression system.

- Tapias, A.; Leplat, C.; Confalonieri, F., 2009: Recovery of ionizing-radiation damage after high doses of gamma ray in the hyperthermophilic archaeon *Thermococcus gammatolerans*. *Extremophiles.*, **13**, 333–343.
- Velázquez Martínez, I., 2012: Aminoácido péptidos y proteínas. *Bioquímica: Las bases moleculares de la vida*. McGraw-Hil., 1–55.
- Walter, S.; Buchner, J., 2002: Molecular Chaperones—Cellular Machines for Protein Folding. *Angewandte Chemie International Edition.*, **41**, 1098–1113.
- Waltraud Ankenbauer; Svetlichny, V.; Bonch-Osmolovskaya, E.; Ebenbichler, C.; Angerer, B.; Schmitz-Agheguian, G.; Laue, F., 1997, October: Thermostable nucleic acid polymerase from *Thermococcus gorgonarius*.
- Wang, F.; Li, S.; Zhao, H.; Bian, L.; Chen, L.; Zhang, Z.; Zhong, X.; Ma, L.; Yu, X., 2015: Expression and characterization of the RKOD DNA polymerase in *Pichia pastoris*. *PLoS ONE.*, **10**, 1–12.
- Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; De Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T., 2018: SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research.*, **46**, W296–W303.
- Wintrode, P. L.; Arnold, F. H., 2000: Temperature adaptation of enzymes: lessons from laboratory evolution. *Advances in protein chemistry*.
- Yang, Y. S.; Fernandez, B.; Lagorce, A.; Aloin, V.; De Guillen, K. M.; Boyer, J. B.; Dedieu, A.; Confalonieri, F.; Armengaud, J.; Roumestand, C., 2015: Prioritizing targets for structural biology through the lens of proteomics: The archaeal protein TGAM_1934 from *Thermococcus gammatolerans*. *Proteomics.*, **15**, 114–123.
- Zhang, L.; Kang, M.; Xu, J.; Huang, Y., 2015: Archaeal DNA polymerases in biotechnology. *Applied Microbiology and Biotechnology.*, **99**, 6585–6597.
- Zivanovic, Y.; Armengaud, J.; Lagorce, A.; Leplat, C.; Guérin, P.; Dutertre, M.; Anthouard, V.; Forterre, P.; Wincker, P.; Confalonieri, F., 2009: Genome analysis and genome-wide proteomics of *Thermococcus gammatolerans*, the most radioresistant organism known amongst the Archaea. *Genome biology.*, **10**, R70.

Anexo I. Extracción del ADN plasmídico

- 1) Concentrar el paquete celular en un tubo eppendorf centrifugando todo el volumen a 10000 rpm.
- 2) El paquete celular de cada muestra se resuspendió en 500µL de solución STET y se agregaron 5µL de una solución de lisozima (100mg/mL)
- 3) Incubar por 10 min. A 35°C (5 min. A 37°C)
- 4) Incubar 1 min. a 95°C (Entre 90-95)
- 5) Centrifugar a 13000 rpm y 15 min. y retirar el paquete celular con un palillo estéril.
- 6) Agregar 20µL de CTAB(5%) y 5µL de ARNasa (10µg/mL), se agita suavemente y se deja incubar por 3 min. a temperatura ambiente.
- 7) Centrifugar las muestras por 15 min. a 13000 rpm para eliminar el sobrenadante.
- 8) Disolver el pellet en 300µL de solución de NaCl (1.2 M)
- 9) Precipitar el ADN con 750µL de etanol absoluto.
- 10) Centrifugar por 20 min. a 13000 rpm y se eliminar el sobrenadante.
- 11) Agregar 500µL de etanol 70% y agitar suavemente.
- 12) Centrifugar por 5 min. a 13000 rpm
- 13) Eliminar sobrenadante con pipeta y dejar secar en la incubadora a 35°C durante toda la noche.
- 14) Agregar 50µL de agua grado biología molecular para resuspender el pellet y analizar en nanodrop.

Anexo II. Electroforesis de proteínas en geles de poliacrilamida (SDS-PAGE)

Para el presente trabajo se utilizaron geles de poliacrilamida al 10%, para su preparación se hizo uso de la **Tabla I**. La preparación de las soluciones utilizadas se describe más adelante.

Tabla I. Preparación de geles a partir de 40% Acrilamida/Bis (BioRad).

Componentes	Gel		Gel separador			
	concentrador	4%	7.5%	10%	12%	15%
H ₂ O Destilada*	1270	2735	2425	2175	1800	3.675 - (A) ^{&} = mL H ₂ O
40% Acrilamida/Bis *	200	940	1250	1500	1875	(2.5*X%)/20= (A) ^{&} mL
0.5 M Tris-HCL pH 6.8 *	500	-	-	-	-	-
1.5 M Tris-HCl pH 8.8 *	-	1250	1250	1250	1250	1250
10% SDS *	20	50	50	50	50	50
10% APS *	10	25	25	25	25	25
TEMED *	4	5	5	5	5	5
Volumen final (mL)	2	5	5	5	5	5

[&]La letra (A) se refiere al volumen de Acrilamida/Bis al 40% necesaria para la preparación de un gel a X%.

* Volúmenes en μ L

La Acrilamida/Bis al 40% indicada en esta Tabla se encuentra en una proporción 37.5:1.

Regulador de corrida de proteínas

1. Añadir las sales indicadas en la Tabla de Preparación (**Tabla II**), a un vaso de precipitados adecuado para el volumen de la solución
2. Agregar el 80% del volumen requerido de H₂O destilada (Arrowhead)
3. Mezclar en una plancha de agitación magnética hasta diluir las sales
4. En el caso de soluciones TAE 1X, ajuste el pH a 7.6 con 1 M HCl o 10 M NaOH (según sea necesario)
5. Aforar la solución con H₂O destilada (Arrowhead) al volumen final requerido.

Tabla II.- Preparación de reguladores de corrida de proteínas Tris-glicina

Regulador	Solución de trabajo	Solución Stock/Litro
Tris-glicina	1X	5X
	25 mM Tris-HCl	15.1 g Tris
	250 mM Glicina	94 g Glicina
	0.1% SDS	50 mL SDS 10%

Persulfato de amonio (APS) 10% (w/v)

Peso molecular: 228.18 g/mol

Ubicación: Refrigerador a 4°C

Preparación: 10 mL - 10%

Disolver 1 g de persulfato de amonio en 10 mL de H₂O destilada (Arrowhead) y almacenar a 4°C.

Solución de tinción con Azul de Coomassie

Peso molecular: 854.03 g/mol

Ubicación: Temperatura Ambiente

Preparación: 100 mL

Disolver 0.25 g de Azul Brillante de Coomassie R-250 en 90 mL de metanol:agua (1:1, v/v) y 10 mL de Ácido Acético Glacial. Filtrar la solución por papel Whatman No.1. Este paso de filtración es para remover cualquier partícula. Almacenar a temperatura ambiente.

DTT (Dithiothreitol)

Peso molecular: 154.253 g/mol

Ubicación: Congelador

Preparación: 10 mL - 1 M

Disolver 1.54 g de DTT en 10 mL de 0.01 M de acetato de sodio (pH 5.2). Esterilizar por filtración. Repartir en alícuotas de 1 mL y almacenar a -20°C. Baje estas condiciones, el DTT es estable a la oxidación por el aire.

Anexo III. Composición de medios de cultivo

Tabla III. Formulación del medio de cultivo Luria Bertani.

SuperBroth	Formulación g/L
Triptona	10 g
Extracto de levadura	5 g
NaCl	10 g

Tabla IV. Formulación del medio de cultivo SuperBroth (SB).

SuperBroth	Formulación g/L
Triptona	35 g
Extracto de levadura	20 g
NaCl	5 g

pH final (25°C) 7.0 ± 0.2

Tabla V. Formulación del medio de cultivo 2XYT.

SuperBroth	Formulación g/L
Triptona	16 g
Extracto de levadura	10 g
NaCl	5 g
Glucosa	0.01%

Anexo IV. Preparación de soluciones y reguladores

NaCl (Cloruro de Sodio, 5 M)

Peso molecular: 58.44 g/mol

Preparación: 1 L - 5 M

Disolver 292 g de NaCl en 800 mL de H₂O destilada (Arrowhead). Ajustar el volumen a 1 litro con H₂O destilada (Arrowhead). Si es necesario, distribuir en alícuotas y esterilizar con autoclave. Almacenar a temperatura ambiente.

IPTG (isopropil-β-D-1-tiogalactopiranosido)

Peso molecular: 238.30 g/mol

Ubicación: Congelador

Descripción: El IPTG libre de dioxano es un análogo de lactosa que inactiva al represor *lac*. El represor *lac* es una proteína de unión al ADN que inhibe la expresión de genes. Por lo tanto, al ser inhibida por el IPTG se permite la expresión de genes de interés y la expresión de proteínas. IPTG no puede ser hidrolizado o descompuestos por *E. coli*, por tanto, la concentración se mantiene constante durante la replicación celular. El IPTG se utiliza a menudo a una concentración final que va desde 0.5-1.0 mM. *Preparación:* 10 mL - 1 M

1. Disolver 2.38 g de IPTG en 8 mL de H₂O destilada (Biología Molecular)
2. Ajustar a un volumen final de 10 mL con H₂O destilada (Biología Molecular)
3. Esterilizar por filtración con un filtro de jeringa de 0.22 μm
4. Distribuir en alícuotas de 1 mL y almacenar a -20°C

NOTA.- El IPTG es sensible a la temperatura. Mantener en hielo y evitar repetidos ciclos de congelado/descongelado.

Regulador STET (Sacarosa/Triton/EDTA/Tris)

Composición:

8% (w/v) de sacarosa

5% de Tritón X-100

50 mM EDTA

50 mM Tris-HCl pH 8

Preparación: 100 mL

1. Medir 50 mL de H₂O destilada (Arrowhead) en un vaso de precipitado de 100 mL
2. Añadir 8 g de sacarosa
3. Añadir 5 mL de Tris HCl 1 M (pH 8.0)
4. Añadir 10 mL de 0.5 M EDTA (pH 8.0)
5. Añadir 5 mL de Triton X-100
6. Llevar a un volumen de 100 mL con H₂O destilada (Arrowhead)
7. Esterilizar por filtración con membrana de 0.22 µm

Regulador de carga (Loading Buffer) para ADN 6X

Preparación: 10 mL - 6X

1. 25 mg Azul de Bromofenol
2. 3 mL Glicerol al 100%
3. 4 g Sacarosa
4. Ajuste el volumen a 10 mL con H₂O destilada (Biología Molecular) *NOTA.- La migración nominal del Azul de Bromofenol es de 300 pb*

Tris-HCl pH 8.0 (1 M)

Peso molecular: 121.14 g/mol

Preparación: 1 L - 1 M

Primera parte

Preparar 100 mL de HCl 1 M. Agregar 8.62 mL de HCl concentrado a 91 mL de H₂O destilada (Arrowhead) previamente colocados en un vaso de precipitados de 250 mL. ¡NUNCA agregar el agua al ácido! Mezcle en una plancha agitadora magnética durante 5 minutos. Aforar a 100 mL de H₂O destilada (Arrowhead).

Segunda parte

1. Disolver 121 g de Tris Base en 800 mL de H₂O destilada (Arrowhead). Ajuste la solución a un pH de 8.0 con HCl 1 M agregándolo gota a gota con una pipeta Pasteur limpia
2. Aforar a 1 litro con H₂O destilada (Arrowhead)
3. Esta solución se puede esterilizar por filtración a través de unidades de filtración equipadas con membranas de 0.2 µm.
4. Conservar a temperatura ambiente o en refrigeración

EDTA (Ácido etilendaminotetraacético) pH 8.0 (0.5 M)

Peso molecular: 292.24 g/mol

Preparación: 1 L - 0.5 M

Primera parte

Preparar 100 mL de Hidróxido de Sodio 10 M (10 M NaOH) agregando 40 g de NaOH a 40 mL de H₂O destilada (Arrowhead) previamente colocados en un vaso de precipitados de 250 mL. Mezclar con una barra magnética en una plancha agitadora hasta que el NaOH se haya disuelto por completo. Aforar a 100 mL con H₂O destilada (Arrowhead). ¡Precaución, esta reacción es exotérmica!

Segunda parte

1. Disolver 146.12 g de EDTA en 700 mL de H₂O destilada (Arrowhead). Ajuste el pH a 8.0 con 10 M NaOH (~50 mL)
2. Aforar a 1 litro con H₂O destilada (Arrowhead)
3. Esta solución se puede esterilizar por filtración a través de unidades de filtración equipadas con membranas de 0.2 µm o con Autoclave
4. Conservar a temperatura ambiente o en refrigeración

Reguladores de electroforesis de ADN (TBE-Tris, Borato, EDTA)

Preparación:

1. Añadir las sales indicadas en la Tabla de Preparación (**Tabla VII**), a un vaso de precipitados adecuado para el volumen de la solución
2. Agregar el 80% del volumen requerido de H₂O destilada (Arrowhead)
3. Mezclar en una plancha de agitación magnética hasta diluir las sales
Amacene a temperatura ambiente.
4. Estas soluciones se pueden esterilizar por filtración a través de unidades de filtración equipadas con membranas de 0.2 µm o con Autoclave

Tabla VI. Preparación de regulador TBE

Regulador	Solución de trabajo	Solución stock/Litro
TBE	0.5X	5X
	45 mM Tris-borate	54g Tris Base
		27.5 g Ácido Bórico
	1mM EDTA	20 mL EDTA 0.5 M (pH 8.0)

Anexo V. Secuenciación del inserto en los vectores de expresión pET19b y pColdI

```

.....| .....| .....| .....| .....| .....| .....| .....| .....| .....|
10 20 30 40 50 60 70 80 90 100
TgPo1B      ATGATTCTCG ATACCGACTA CATCACCAG AACGGGAAG CCGTGATAAG GGTCTTCRAG AAGGAGAACG GCGAGTTTAA AATCGAGTAC GACAGAACCT
pET19b_C1   -----| -----| -----| -----| -----| -----| -----| -----| -----|
pET19b_C2   -----| -----| -----| -----| -----| -----| -----| -----| -----|
pET19b_C3   -----| -----| -----| -----| -----| -----| -----| -----| -----|
pColdI_C4   -----| -----| -----| -----| -----| -----| -----| -----| -----|

.....| .....| .....| .....| .....| .....| .....| .....| .....| .....|
110 120 130 140 150 160 170 180 190 200
TgPo1B      TCGAGCCATA CTTCTACGCC CTCCTGAAGG ATGATTCCGC AATAGAAGAC GTGAAGAAAG TAACCCGCCAA AAGGCACGGG GCGGTCGTCA AGGTGAAGCG
pET19b_C1   TCGAGCCATA CTTCTACGCC CTCCTGAAGG ATGATTCCGC AATAGAAGAC GTGAAGAAAG TAACCCGCCAA AAGGCACGGG GCGGTCGTCA AGGTGAAGCG
pET19b_C2   TCGAGCCATA CTTCTACGCC CTCCTGAAGG ATGATTCCGC AATAGAAGAC GTGAAGAAAG TAACCCGCCAA AAGGCACGGG GCGGTCGTCA AGGTGAAGCG
pET19b_C3   TCGAGCCATA CTTCTACGCC CTCCTGAAGG ATGATTCCGC AATAGAAGAC GTGAAGAAAG TAACCCGCCAA AAGGCACGGG GCGGTCGTCA AGGTGAAGCG
pColdI_C4   TCGAGCCATA CTTCTACGCC CTCCTGAAGG ATGATTCCGC AATAGAAGAC GTGAAGAAAG TAACCCGCCAA AAGGCACGGG GCGGTCGTCA AGGTGAAGCG

.....| .....| .....| .....| .....| .....| .....| .....| .....| .....|
210 220 230 240 250 260 270 280 290 300
TgPo1B      CGCCGAGAAG GTGCAGAGGA AGTTCCCTCG CAGGCCGATA GAGGTCTGGA AGCTCTACTT CACCCATCCT CAGGACGTC CCGCGATTCC AGACAAGATA
pET19b_C1   CGCCGAGAAG GTGCAGAGGA AGTTCCCTCG CAGGCCGATA GAGGTCTGGA AGCTCTACTT CACCCATCCT CAGGACGTC CCGCGATTCC AGACAAGATA
pET19b_C2   CGCCGAGAAG GTGCAGAGGA AGTTCCCTCG CAGGCCGATA GAGGTCTGGA AGCTCTACTT CACCCATCCT CAGGACGTC CCGCGATTCC AGACAAGATA
pET19b_C3   CGCCGAGAAG GTGCAGAGGA AGTTCCCTCG CAGGCCGATA GAGGTCTGGA AGCTCTACTT CACCCATCCT CAGGACGTC CCGCGATTCC AGACAAGATA
pColdI_C4   CGCCGAGAAG GTGCAGAGGA AGTTCCCTCG CAGGCCGATA GAGGTCTGGA AGCTCTACTT CACCCATCCT CAGGACGTC CCGCGATTCC AGACAAGATA

.....| .....| .....| .....| .....| .....| .....| .....| .....| .....|
310 320 330 340 350 360 370 380 390 400
TgPo1B      CGCGCCCATC CAGCAGTTGT GGACATCTAC GAGTACGACA TACCCTTCGC CAAGCGCTAC CTCATCGACA AGGGCCTGAT TCCGATGGAG GGGGACGAAG
pET19b_C1   CGCGCCCATC CAGCAGTTGT GGACATCTAC GAGTACGACA TACCCTTCGC CAAGCGCTAC CTCATCGAGC AGGGCCTGAT TCCGATGGAG GGGGACGAAG
pET19b_C2   CGCGCCCATC CAGCAGTTGT GGACATCTAC GAGTACGACA TACCCTTCGC CAAGCGCTAC CTCATCGACA AGGGCCTGAT TCCGATGGAG GGGGACGAAG
pET19b_C3   CGCGCCCATC CAGCAGTTGT GGACATCTAC GAGTACGACA TACCCTTCGC CAAGCGCTAC CTCATCGACA AGGGCCTGAT TCCGATGGAG GGGGACGAAG
pColdI_C4   CGCGCCCATC CAGCAGTTGT GGACATCTAC GAGTACGACA TACCCTTCGC CAAGCGCTAC CTCATCGACA AGGGCCTGAT TCCGATGGAG GGGGACGAAG

.....| .....| .....| .....| .....| .....| .....| .....| .....| .....|
410 420 430 440 450 460 470 480 490 500
TgPo1B      AGCTCAGGAT GCTCGCCTTT GACATCGAGA CTCTCTACCA TGAAGGCCGAG GAGTTCGGAA CCGGGCCGAT TCTCATGATA AGCTACGCCG ACGAGAACGA
pET19b_C1   AGCTCAGGAT GCTCGCCTTT GACATCGAGA CTCTCTACCA TGAAGGCCGAG GAGTTCGGAA CCGGGCCGAT TCTCATGATA AGCTACGCCG ACGAGAACGA
pET19b_C2   AGCTCAGGAT GCTCGCCTTT GACATCGAGA CTCTCTACCA TGAAGGCCGAG GAGTTCGGAA CCGGGCCGAT TCTCATGATA AGCTACGCCG ACGAGAACGA
pET19b_C3   AGCTCAGGAT GCTCGCCTTT GACATCGAGA CTCTCTACCA TGAAGGCCGAG GAGTTCGGAA CCGGGCCGAT TCTCATGATA AGCTACGCCG ACGAGAACGA
pColdI_C4   AGCTCAGGAT GCTCGCCTTT GACATCGAGA CTCTCTACCA TGAAGGCCGAG GAGTTCGGAA CCGGGCCGAT TCTCATGATA AGCTACGCCG ACGAGAACGA

.....| .....| .....| .....| .....| .....| .....| .....| .....| .....|
510 520 530 540 550 560 570 580 590 600
TgPo1B      GGCRAAGGTTG ATAACCTGGA AGAAGATTGA CCTTCCGTA CTTGACGTCG TTTCCGACCGA GAAGGAGATG ATTAAGCGCT TCCTCCGCGT CGTCAAGGAG
pET19b_C1   GGCRAAGGTTG ATAACCTGGA AGAAGATTGA CCTTCCGTA CTTGACGTCG TTTCCGACCGA GAAGGAGATG ATTAAGCGCT TCCTCCGCGT CGTCAAGGAG
pET19b_C2   GGCRAAGGTTG ATAACCTGGA AGAAGATTGA CCTTCCGTA CTTGACGTCG TTTCCGACCGA GAAGGAGATG ATTAAGCGCT TCCTCCGCGT CGTCAAGGAG
pET19b_C3   GGCRAAGGTTG ATAACCTGGA AGAAGATTGA CCTTCCGTA CTTGACGTCG TTTCCGACCGA GAAGGAGATG ATTAAGCGCT TCCTCCGCGT CGTCAAGGAG
pColdI_C4   GGCRAAGGTTG ATAACCTGGA AGAAGATTGA CCTTCCGTA CTTGACGTCG TTTCCGACCGA GAAGGAGATG ATTAAGCGCT TCCTCCGCGT CGTCAAGGAG

.....| .....| .....| .....| .....| .....| .....| .....| .....| .....|
610 620 630 640 650 660 670 680 690 700
TgPo1B      AAGGACCCCG ACGTGCTCAT CACCTACAAC GCGCAACA CTTGACTTCG CTACCTGAAA AAGCGCTGTG AGAAACTCGG GATAAAGTTC ACGCTCGGGA
pET19b_C1   AAGGACCCCG ACGTGCTCAT CACCTACAAC GCGCAACA CTTGACTTCG CTACCTGAAA AAGCGCTGTG AGAAACTCGG GATAAAGTTC ACGCTCGGGA
pET19b_C2   AAGGACCCCG ACGTGCTCAT CACCTACAAC GCGCAACA CTTGACTTCG CTACCTGAAA AAGCGCTGTG AGAAACTCGG GATAAAGTTC ACGCTCGGGA
pET19b_C3   AAGGACCCCG ACGTGCTCAT CACCTACAAC GCGCAACA CTTGACTTCG CTACCTGAAA AAGCGCTGTG AGAAACTCGG GATAAAGTTC ACGCTCGGGA
pColdI_C4   AAGGACCCCG ACGTGCTCAT CACCTACAAC GCGCAACA CTTGACTTCG CTACCTGAAA AAGCGCTGTG AGAAACTCGG GATAAAGTTC ACGCTCGGGA

.....| .....| .....| .....| .....| .....| .....| .....| .....| .....|
710 720 730 740 750 760 770 780 790 800
TgPo1B      GAGACGGGAG CGAGCCGGAAG ATCCAGCGCA TGGCGGACCG CTTCCCGCTG GAGGTAAGG GTAGGATACA CTTCCGATCTC TATCCAGTCA TAAGGCCGAC
pET19b_C1   GAGACGGGAG CGAGCCGGAAG ATCCAGCGCA TGGCGGACCG CTTCCCGCTG GAGGTAAGG GTAGGATACA CTTCCGATCTC TATCCAGTCA TAAGGCCGAC
pET19b_C2   GAGACGGGAG CGAGCCGGAAG ATCCAGCGCA TGGCGGACCG CTTCCCGCTG GAGGTAAGG GTAGGATACA CTTCCGATCTC TATCCAGTCA TAAGGCCGAC
pET19b_C3   GAGACGGGAG CGAGCCGGAAG ATCCAGCGCA TGGCGGACCG CTTCCCGCTG GAGGTAAGG GTAGGATACA CTTCCGATCTC TATCCAGTCA TAAGGCCGAC
pColdI_C4   GAGACGGGAG CGAGCCGGAAG ATCCAGCGCA TGGCGGACCG CTTCCCGCTG GAGGTAAGG GTAGGATACA CTTCCGATCTC TATCCAGTCA TAAGGCCGAC

.....| .....| .....| .....| .....| .....| .....| .....| .....| .....|
810 820 830 840 850 860 870 880 890 900
TgPo1B      GATAAACCTC CCGACCTACA CCCTTGAGGC CGTTTACGAG GCGGTCTTCG GGAAGCCRAAG GAGAAGGTTT ACGCCGAGGA GATAACAGAG CCCTGGGAGA
pET19b_C1   GATAAACCTC CCGACCTACA CCCTTGAGGC CGTTTACGAG GCGGTCTTCG GGAAGCCRAAG GAGAAGGTTT ACGCCGAGGA GATAACAGAG CCCTGGGAGA
pET19b_C2   GATAAACCTC CCGACCTACA CCCTTGAGGC CGTTTACGAG GCGGTCTTCG GGAAGCCRAAG GAGAAGGTTT ACGCCGAGGA GATAACAGAG CCCTGGGAGA
pET19b_C3   GATAAACCTC CCGACCTACA CCCTTGAGGC CGTTTACGAG GCGGTCTTCG GGAAGCCRAAG GAGAAGGTTT ACGCCGAGGA GATAACAGAG CCCTGGGAGA
pColdI_C4   GATAAACCTC CCGACCTACA CCCTTGAGGC CGTTTACGAG GCGGTCTTCG GGAAGCCRAAG GAGAAGGTTT ACGCCGAGGA GATAACAGAG CCCTGGGAGA

.....| .....| .....| .....| .....| .....| .....| .....| .....| .....|
910 920 930 940 950 960 970 980 990 1000
TgPo1B      GCGGGGAGGG CCTTGAAAGG GTTGCCCGCT ACTCGATGGA GGACGCGAAG GTGACCTACG AGCTGGGAAG GGAGTTCITTT CCGATGGAGG CCCAGCTTTT
pET19b_C1   GCGGGGAGGG CCTTGAAAGG GTTGCCCGCT ACTCGATGGA GGACGCGAAG GTGACCTACG AGCTGGGAAG GGAGTTCITTT CCGATGGAGG CCCAGCTTTT
pET19b_C2   GCGGGGAGGG CCTTGAAAGG GTTGCCCGCT ACTCGATGGA GGACGCGAAG GTGACCTACG AGCTGGGAAG GGAGTTCITTT CCGATGGAGG CCCAGCTTTT
pET19b_C3   GCGGGGAGGG CCTTGAAAGG GTTGCCCGCT ACTCGATGGA GGACGCGAAG GTGACCTACG AGCTGGGAAG GGAGTTCITTT CCGATGGAGG CCCAGCTTTT
pColdI_C4   GCGGGGAGGG CCTTGAAAGG GTTGCCCGCT ACTCGATGGA GGACGCGAAG GTGACCTACG AGCTGGGAAG GGAGTTCITTT CCGATGGAGG CCCAGCTTTT

```



```

.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
 2110      2120      2130      2140      2150      2160      2170      2180      2190      2200
TgPolB   ACATCGTCCT AAAGGGCTCT GGCAGGATAG GTGACAGGGC GATTCCAGCC GATGAATTCG ACCCGACGAA GCACCCGCTAC GATGCGGAAT ATTACATCGA
pET19b_C1 ACATCGTCCT AAAGGGCTCT GGCAGGATAG GTGACAGGGC GATTCCAGCC GATGAATTCG ACCCGACGAA GCACCCGCTAC GATGCGGAAT ATTACATCGA
pET19b_C2 ACATCGTCCT AAAGGGCTCT GGCAGGATAG GTGACAGGGC GATTCCAGCC GATGAATTCG ACCCGACGAA GCACCCGCTAC GATGCGGAAT ATTACATCGA
pET19b_C3 ACATCGTCCT AAAGGGCTCT GGCAGGATAG GTGACAGGGC GATTCCAGCC GATGAATTCG ACCCGACGAA GCACCCGCTAC GATGCGGAAT ATTACATCGA
pColdI_C4 ACATCGTCCT AAAGGGCTCT GGCAGGATAG GTGACAGGGC GATTCCAGCC GATGAATTCG ACCCGACGAA GCACCCGCTAC GATGCGGAAT ATTACATCGA

```

```

.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
 2210      2220      2230      2240      2250      2260      2270      2280      2290      2300
TgPolB   GAACCCAGGTT CTGCCGGCAG TGGAGAGGAT TCTAAAAGCC TTCGGCTACC GGAAGGAGGA TTTGAGGTAC CAGAAGACGA AGCAGGTCGG TTTGGGCGCG
pET19b_C1 GAACCCAGGTT CTGCCGGCAG TGGAGAGGAT TCTAAAAGCC TTCGGCTACC GGAAGGAGGA TTTGAGGTAC CAGAAGACGA AGCAGGTCGG TTTGGGCGCG
pET19b_C2 GAACCCAGGTT CTGCCGGCAG TGGAGAGGAT TCTAAAAGCC TTCGGCTACC GGAAGGAGGA TTTGAGGTAC CAGAAGACGA AGCAGGTCGG TTTGGGCGCG
pET19b_C3 GAACCCAGGTT CTGCCGGCAG TGGAGAGGAT TCTAAAAGCC TTCGGCTACC GGAAGGAGGA TTTGAGGTAC CAGAAGACGA AGCAGGTCGG TTTGGGCGCG
pColdI_C4 GAACCCAGGTT CTGCCGGCAG TGGAGAGGAT TCTAAAAGCC TTCGGCTACC GGAAGGAGGA TTTGAGGTAC CAGAAGACGA AGCAGGTCGG TTTGGGCGCG

```

```

.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
 2310      2320
TgPolB   TGGCTGAAGG TGAAGGGGAA GAAGTGA
pET19b_C1 TGGCTGAAGG TGAAGGGGAA GAAGTGA
pET19b_C2 TGGCTGAAGG TGAAGGGGAA GAAGTGA
pET19b_C3 TGGCTGAAGG TGAAGGGGAA GAAGTGA
pColdI_C4 TGGCTGAAGG TGAAGGGGAA GAAGTGA

```