

UNIVERSIDAD PERUANA CAYETANO HEREDIA
FACULTAD DE CIENCIAS Y FILOSOFÍA
“Alberto Cazorla Talleri”



Evaluación y comparación de la concentración y solubilidad de la proteína ClpC1 de *Mycobacterium tuberculosis* producida de manera recombinante mediante el uso de los plásmidos pET28a(+) y pMAL-cHT en tres cepas distintas de *Escherichia coli*.

Tesis para optar el título de Licenciado en Biología

Angela Anaid Rios Angulo

Lima-Perú

2021

Asesor (a)

PhD. Mirko Zimic Peralta

Co-asesor (a)

PhD. Patricia Sheen Cortavarria

AGRADECIMIENTOS

Agradezco primeramente a Dios por su guía y a mis padres Julia y Arnaldo por apoyarme constantemente en cada uno de los pasos que doy en la vida.

A mi asesor de Tesis el Dr. Mirko Zimic, co-asesora y a mis profesores Dra. Patricia Sheen y MSc. Eduardo Gushiken por sus consejos para el desarrollo del presente trabajo.

A Stefany Quiñonez por aconsejarme y guiarme en los diferentes ensayos y experimentos en el laboratorio, por sus ideas y aportes en los protocolos realizados; así como su ayuda emocional al estar pendiente de mi avance diario y amistad absoluta.

A Alonso Vigil por su constante apoyo y optimismo.

A mis compañeros de laboratorio por su ayuda continua en los procedimientos experimentales, especialmente a Erika, Ana Paula, Gabriel y Gustavo.

A mis amigos de la universidad por su motivación y preocupación en el desarrollo de la tesis, especialmente Fiorella, Angélica, Ana Flavia, Julio, Iván y Juan.

Por último, a mis amigos Seichi, Pedrito y María Paz por su creatividad en la elaboración de los proyectos.

RESUMEN

Las proteínas Clp son un grupo de proteasas caseolíticas del grupo serina dependiente de ATP. Su función principal es el control de calidad de las proteínas. En *Mycobacterium tuberculosis*, ClpC1 participa del mecanismo de acción de fármacos antituberculosos como pirazinamida (PZA) por ello es importante estudios sobre su función y estructura. Asimismo, monómero de ClpC1 de *Bacillus subtilis* se ha cristalizado completamente por ello se utiliza como plantilla para modelar ClpC1 de otras bacterias. ClpC1 es una proteína citosólica por lo que su expresión en forma soluble ha sido confirmada en estudios previos. El sistema de expresión en *Escherichia coli* usando el plásmido pET28a(+) como vector de expresión, ha sido utilizado para producir ClpC1; sin embargo, no hay información acerca del porcentaje de solubilidad y concentración de proteína producida. Por otro lado, el plásmido pMAL-cHT es ampliamente utilizado para aumentar la solubilidad de proteínas recombinantes en *E. coli*. pMAL-cHT adiciona el gen *male*, que codifica la proteína de unión a maltosa (MBP), en la proteína a expresar y por mecanismos aún no confirmados aumenta la solubilidad y cantidad de las proteínas recombinantes.

En este estudio se estimó la concentración y porcentaje de solubilidad de ClpC1 de *M. tuberculosis* expresada en diferentes sistemas de expresión usando vectores que incluían o no la proteína de fusión MBP, diferentes concentraciones de inductor y tiempos de inducción.

Las mejores condiciones para obtener mayor concentración y porcentaje de solubilidad de ClpC1 fueron con la cepa BL21(DE3)pLysS, el plásmido pMAL-cHT-clpC1 inducido con 0.5 mM de IPTG por 10 horas a 16°C. En este sistema se obtuvo 45,54 mg/ml (IC 95% 44,68 - 46,40) de ClpC1 con un porcentaje de solubilidad de 86.93% (IC 95% 80,41 - 93,45). De la misma manera, el modelo bioinformático de ClpC1 con MBP presentó mayor solubilidad global comparación de ClpC1.

Palabras clave:

Escherichia coli, *Mycobacterium tuberculosis*, ClpC1, porcentaje de solubilidad, modelamiento bioinformático.

ABSTRACT

Clp proteins are a group of caseolytic proteases of the ATP-dependent serine group. Its main function is the quality control of proteins. In *Mycobacterium tuberculosis*, ClpC1 participates in the mechanism of action of antituberculous drugs such as pyrazinamide (PZA), which is why studies on its function and structure are important. Also, ClpC1 monomer from *Bacillus subtilis* has been completely crystallized so it is used as a template to model ClpC1 from other bacteria. ClpC1 is a cytosolic protein so its expression in soluble form has been confirmed in previous studies. The expression system in *Escherichia coli* using the plasmid pET28a (+) as an expression vector has been used to produce ClpC1; however, there is no information about the percentage of solubility and concentration of protein produced. On the other hand, the plasmid pMAL-CHT is widely used to increase the solubility of recombinant proteins in *E. coli*. pMAL-CHT adds the *maltE* gene, which encodes the maltose-binding protein (MBP), to the protein to be expressed and, by mechanisms not yet confirmed, increases the solubility and quantity of recombinant proteins.

In this study, the concentration and percentage of solubility of *M. tuberculosis* ClpC1 expressed in different expression systems were estimated using vectors that included or did not include the MBP fusion protein, different concentrations of inducer and induction times.

The best conditions to obtain a higher concentration and percentage of solubility of ClpC1 were with the strain BL21 (DE3) pLysS, the plasmid pMAL-CHT-clpC1 induced with 0.5 mM of IPTG for 10 hours at 16 ° C. In this system, 45.54 mg / ml (95% CI 44.68 - 46.40) of ClpC1 was obtained with a solubility percentage of 86.93% (95% CI 80.41 - 93.45). In the same way, the bioinformatic model of ClpC1 with MBP presented greater global solubility compared to ClpC1.

Keywords:

Escherichia coli, *Mycobacterium tuberculosis*, ClpC1, percent solubility, bioinformatics modeling.

ÍNDICE

RESUMEN

ABSTRACT

1. INTRODUCCIÓN	1
2. MARCO TEÓRICO	4
2.1. Chaperonas: Definición y función	4
2.1.1. Proteína ClpC1: Características y función	4
2.1.1.1. ClpC1 en bacterias	5
2.1.1.2. ClpC1 en plantas	5
2.1.2. ClpC1 en <i>M. tuberculosis</i>	6
2.1.2.1. ClpC1 como blanco terapéutico de drogas antituberculosas	7
2.1.2.2. ClpC1 como blanco del ácido pirazinoico	8
2.2. Cristal de ClpC1	9
2.3. Expresión de ClpC1 de <i>M.tuberculosis</i> en <i>E.coli</i>	9
2.4. Problema de la expresión soluble de proteínas recombinantes en <i>E.coli</i>	10
2.4.1. Efecto de la solubilidad según el tipo de plásmido	11
2.4.1.1. Plásmido pET28a(+)	11
2.4.1.2. Plásmido pMAL-cHT	12
2.4.1.2.1. Maltose-binding protein (MBP)	12
2.4.2. Efecto de la solubilidad según el tipo de cepa de <i>E.coli</i>	12
2.4.2.1. Cepa BL21(DE3)pLySs	13
2.4.2.2. Cepa Lemo(DE3)pLySs	13
2.4.2.3. Cepa Rosetta(DE3)pLySs	13
2.4.3. Efecto de la solubilidad según la concentración de IPTG	14
2.4.4. Efecto de la solubilidad según el tiempo y temperatura de expresión	14
2.5. Modelamiento estructural de proteínas	15
2.5.1. Parámetros de evaluación del modelo	15
2.5.1.1. GMQE	15
2.5.1.2. QMEAN	15

2.5.1.3.	RMSD	16
2.5.2.	Parámetros de solubilidad	16
3.	PLANTEAMIENTO DEL PROBLEMA Y OBJETIVOS	16
3.1.	PLANTEAMIENTO DEL PROBLEMA	16
3.2.	OBJETIVO PRINCIPAL	18
3.3.	OBJETIVOS ESPECÍFICOS	18
4.	METODOLOGÍA	18
4.1.	Diseño de primers para el gen <i>Rv3596c</i>	18
4.2.	Amplificación del gen <i>Rv3596c</i>	19
4.3.	Clonación con <i>In-fusion</i> del gen <i>Rv3596c</i>	19
4.4.	Secuenciación de <i>Rv3596c</i>	19
4.5.	Expresión de ClpC1 recombinante	20
4.6.	Estimación de la concentración total de ClpC1 recombinante en los sistemas de expresión	22
4.7.	Estimación del porcentaje de solubilidad de ClpC1 recombinante en los sistemas de expresión	23
4.8.	Análisis Estadístico	25
4.9.	Purificación de ClpC1-MBP y remoción de MBP de ClpC1	25
4.10.	Modelamiento estructural de ClpC1 con y sin MBP	25
4.11.	Estimación de los parámetros de solubilidad de los modelos de Clpc1 con y sin MBP	26
5.	RESULTADOS	26
5.1.	Diseño de primers para el gen <i>Rv3596c</i>	26
5.2.	Amplificación del gen <i>Rv3596c</i>	27
5.3.	Clonación con <i>In-fusion</i> del gen <i>Rv3596c</i>	28
5.4.	Expresión de Clpc1 recombinante	29
5.5.	Comparación de la concentración y porcentaje de solubilidad de ClpC1 recombinante en los sistemas de expresión	30
5.6.	Purificación de ClpC1-MBP y remoción de MBP de ClpC1	38
5.7.	Modelamiento estructural de ClpC1 con y sin MBP	39
5.8.	Estimación de los parámetros de solubilidad de los modelos de Clpc1	

	con y sin MBP	40
6.	DISCUSIÓN	42
7.	CONCLUSIONES	45
8.	RECOMENDACIONES	46
9.	REFERENCIAS BIBLIOGRÁFICAS	47
10.	ANEXO	56

ÍNDICE DE TABLAS

Tabla 1. Expresión de ClpC1 de <i>M.tuberculosis</i> a diferentes condiciones en <i>E.coli</i> .	10
Tabla 2. Lista de primers utilizados para la secuenciación de <i>Rv3596c</i> en pMAL- <i>cHT</i> y pET28a(+).	20
Tabla 3. Tabla resumen de las diferentes condiciones probadas en cada expresión.	21
Tabla 4. Características de los primers diseñados para la clonación de <i>Rv3596c</i> en los plásmidos pET28a(+) y pMAL- <i>cHT</i> .	26
Tabla 5. Resumen de los parámetros de solubilidad de ClpC1 sin y con MBP.	42

ÍNDICE DE FIGURAS

Figura 1. Proteasas AAA + y sus respectivas proteínas adaptadoras de <i>B. subtilis</i> .	6
Figura 2. Mecanismo de regulación de la homeostasis proteica de ClpC1/P1/P2 en <i>M.tuberculosis</i> .	7
Figura 3. Mecanismo de acción propuesto de PZA/POA sobre PanD y ClpC1.	9
Figura 4. Procedimiento de las fracciones no inducidas, inducidas y extracto crudo.	23
Figura 5. Procedimiento de la fracción del extracto crudo y fracción soluble.	24
Figura 6. Amplificación del gen Rv3596c con los primers diseñados para el plásmido pET28a(+)	27
Figura 7. Amplificación del gen Rv3596c con los primers diseñados para el plásmido pMAL- cHT.	27
Figura 8. Digestión simple (SD) con <i>Hind</i> III y digestión doble (DD) con <i>Hind</i> III y <i>Nco</i> I del plásmido pET28a(+)-ClpC1.	28
Figura 9. Digestión simple (SD) con <i>Hind</i> III y digestión doble (DD) con <i>Hind</i> III y <i>Eco</i> RI del plásmido pMAL-cHT- ClpC1.	29
Figura 10. Expresión de ClpC1 recombinante en la cepa BL21(DE3)pLysS con el plásmido pET28a(+).	30
Figura 11. Expresión de ClpC1 recombinante en la cepa BL21(DE3)pLySs clonada con el plásmido pMAL- cHT.	30
Figura 12. Estimación de la concentración total promedio y porcentaje de solubilidad de ClpC1 entre los plásmidos pET28a(+) y pMAL- cHT.	32
Figura 13. Estimación de la concentración total promedio y porcentaje de solubilidad de ClpC1 recombinante en las cepas BL21(DE3)pLysS, Lemo(DE3)pLysS y Rosetta(DE3)pLysS con los plásmidos pET28a (+) y pMAL-cHT.	33
Figura 14. Estimación de la concentración total promedio y porcentaje de solubilidad de ClpC1 referente a las diferentes concentraciones de IPTG en cepas clonadas con pET28a(+) y pMALcHT.	35
Figura 15. Estimación de la concentración total promedio y porcentaje de solubilidad de ClpC1 referente a las diferentes concentraciones de IPTG en cepas clonadas con pET28a(+) y pMAL-cHT.	37
Figura 16. Purificación y corte de ClpC1-MBP.	38
Figura 17. Estructura tridimensional de ClpC1 sin MBP.	39
Figura 18. Estructura tridimensional de ClpC1 con MBP.	40

Figura 19. Porcentaje de residuos hidrofóbicos e hidrofílicos de ClpC1 sin MBP.	41
Figura 20. Porcentaje de residuos hidrofóbicos e hidrofílicos de ClpC1 con MBP.	41

1. INTRODUCCIÓN

Las proteínas Clp (proteínas caseolíticas) o de shock térmico *HSP 100* son una familia de chaperonas altamente conservadas entre especies y facilitan el correcto plegamiento de las proteínas (Feldman & Frydman, 2000). Su expresión aumenta significativamente en condiciones de estrés biológico (Lindquist & Craig, 1988). Dentro de esta familia se encuentra la proteína ClpC1, la cual es una proteína caseolítica del grupo serina dependiente de ATP. ClpC1 despliega proteínas mal plegadas mediante la hidrólisis de ATP y las transloca en hacia su degradación en los complejos ClpP (Gottesman *et al*, 1990). ClpC1 de *M. tuberculosis* es esencial tanto en su estado activo como en su estado de latencia (Kar *et al*, 2008; Leodolter *et al*, 2015). Por esta razón, es blanco de fármacos antituberculosos como lassomicina, cyclomarina A, ecumicina y rufomicina (Choules *et al*, 2019; Gao *et al*, 2015; Gavrish *et al*, 2014; Vasudevan *et al*, 2013; Weinhäupl *et al*, 2018; Wolf *et al*, 2019). De la misma manera, se ha demostrado que está involucrada en el mecanismo de acción de pirazinamida (PZA); en presencia del ácido pirazinoico (POA), compuesto activo de PZA, ClpC1 degrada la proteína PanD, la cual es una descarboxilasa que genera beta alanina (Gopal *et al*, 2020a). Sin embargo, hasta la fecha no se sabe si ClpC1 es blanco de POA. Para entender mejor la relación de PanD con el mecanismo de PZA así como en otras rutas metabólicas es importante primero obtener ClpC1 recombinante en forma soluble y activa para dar respuesta a dicha pregunta.

Estudiar la función y estructura de una proteína recombinante, su obtención de manera soluble y en una concentración adecuada es fundamental. La solubilidad depende del sistema de expresión, la temperatura, el tiempo de expresión, la naturaleza y concentración del inductor, así como la secuencia de aminoácidos y modificaciones de la misma proteína. Asimismo, las cepas más utilizadas para aumentar la solubilidad de proteínas recombinantes en *E. coli* son BL21(DE3)pLysS, Lemo(DE3)pLysS y Rosetta(DE3)pLysS. Entre las características que comparten se encuentra la denominación (DE3) y pLysS. DE3 se refiere a que la cepa presenta el fago lisogénico DE3 (λ DE3), por ello existe el gen que codifica la ARN polimerasa T7 dentro del cromosoma de la bacteria (Studier *et al*, 1990). De la misma manera, estas cepas presentan el plásmido pLysS, el cual presenta el gen que codifica la lisozima T7 y el marcador de resistencia a cloranfenicol (Studier, 1991).

La cepa BL21(DE3)pLysS se utiliza para producir de forma soluble proteínas insolubles mientras que Lemo(DE3)pLysS se utiliza para expresar proteínas de membrana, secretoras e insolubles, y Rosetta(DE3)pLysS expresa proteínas eucariotas (Angius *et al*, 2018; Schlegel *et al*, 2012).

Por otro lado, el inductor isopropil- β -D-1-tiogalactopiranosido (IPTG) es un compuesto químico que se utiliza para inducir la expresión de proteínas recombinantes en *E. coli*. IPTG es utilizado en sistemas de expresión basados en el operón lac, debido a que se une al represor lacI para expresar el ARN polimerasa T7. Este sistema se complementa al utilizar el vector pET, el cual presenta el promotor T7 (en donde se une el ARN polimerasa T7) antes del gen de la proteína a expresar (William Studier *et al*, 1990).

ClpC1 es una proteína citosólica por lo que su expresión en forma soluble es factible o altamente probable. Por lo cual, clonar y producir ClpC1 en forma soluble, a través de un sistema de expresión en *E. coli*, es fundamental para estudiar su posible implicancia dentro del mecanismo de PZA como blanco de POA u otras drogas antituberculosas.

El sistema de expresión en *E. coli* con el plásmido pET28a(+) ha sido utilizado para producir ClpC1 (Kandror *et al*, 2014), sin embargo, no hay información acerca del porcentaje de solubilidad obtenido. Por otro lado, el uso del plásmido pMAL-cHT permite que se adicione una proteína de unión a maltosa (MBP) en la proteína recombinante, lo cual aumenta la solubilidad y la cantidad de expresión (Sun *et al*, 2011; Walls & Loughran, 2011). En adición, MBP es removida por la proteasa TEV, la cual deriva del virus del grabado de tabaco (Raran-Kurussi *et al*, 2017).

Por otro lado, en *B. subtilis* se ha cristalizado ClpC1 y debido a ello, es utilizado como plantilla o base para modelar ClpC1 de otras bacterias (Wang *et al*, 2011). En *M. tuberculosis* se ha logrado cristalizar solo el dominio N-terminal de ClpC1 (Vasudevan *et al*, 2013). Por ello, modelar bioinformáticamente ClpC1 de *M. tuberculosis* usando como molde ClpC1 de *B. subtilis* es posible.

Asimismo, para hallar *in silico* los parámetros de solubilidad de una proteína, se utilizan el área superficial accesible al solvente o SASA (*Solvent-accessible surface area*), el porcentaje de residuos hidrofóbicos e hidrofílicos y el predictor de solubilidad protein-Sol (Ali *et al*, 2014; Hebditch *et al*, 2017).

En el presente trabajo se estimó la concentración y solubilidad de la proteína ClpC1 de *M. tuberculosis* producida en los plásmidos pET28a(+) y pMAL-cHT en las cepas BL21(DE3)pLysS, Lemo(DE3)pLysS y Rosetta(DE3)pLysS de *E. coli* a diferentes

concentraciones de IPTG y en distintas temperaturas y tiempos de expresión. Asimismo, se realizó el modelamiento estructural de ClpC1 con/sin MBP para poder determinar los cambios de los parámetros de solubilidad que ocasiona la adición de MBP en la estructura de ClpC1.

2. MARCO TEÓRICO

2.1. Chaperonas: Definición y función

Las chaperonas moleculares cumplen funciones como el ensamblaje, transporte, plegamiento y disgregación de proteínas dentro de una célula durante su crecimiento (Gottesman *et al*, 1997). Se sabe que, en condiciones de estrés biológico, la expresión de chaperonas aumenta significativamente (Lindquist & Craig, 1988).

Las proteínas de shock térmico (Hsps) actúan como chaperonas al facilitar el correcto plegamiento de proteínas debido a que interactúan con las regiones hidrofóbicas en polipéptidos desplegados, y de esta manera previenen la agregación de proteínas (Feldman & Frydman, 2000). Existen cinco familias principales de Hsps, Hsp100, Hsp90, Hsp70, Hsp60 y las proteínas pequeñas de shock térmico (sHsps) (Richter *et al*, 2010). Las proteínas Hsp100 o caseolíticas (Clp) son una familia altamente conservada entre organismos, desde bacterias hasta mamíferos (Gottesman *et al*, 1990; Kang *et al*, 2002; Squires *et al*, 1991; Squires & Squires, 1992). ClpC presenta actividad ATPasa, la cual es necesaria para traslocar proteínas mal plegadas o agregadas hacia su degradación en ClpP, proteasa caseolítica dependiente de ATP (Horwich *et al*, 1999).

2.1.1. Proteína ClpC1: Características y función

ClpC1 es una proteína que se encuentra en múltiples organismos como bacterias, cianobacterias fotosintéticas, cloroplastos de algas y plantas superiores (Clarke & Eriksson, 1996; Gottesman *et al*, 1990; Msadek *et al*, 1994). ClpC1 pertenece a la clase I de Clp debido a que presenta dos dominios ATPasas asociados con diversas actividades (AAA +) y un dominio N-terminal, el cual es homólogo a ClpA/B (Kirstein *et al*, 2006). ClpC es un homohexámero mientras que ClpP es un heptámero que se unen formando un complejo para degradar proteínas mal plegadas, la actividad ATPasa de ClpC facilita el despliegue de la proteína hacia la proteasa ClpP. ClpP no presenta actividad ATPasa y cuando no forma un complejo con ClpC, solo presenta actividad peptidasa. Sin embargo, cuando ClpP se une a ClpC actúa como una serina proteasa (Maurizi *et al*, 1990).

La arquitectura general de este complejo es altamente conservada en organismos, aunque pueden existir homólogos y parálogos tanto para ClpC y ClpP. Por ejemplo, en

plantas existen dos tipos de ClpC (ClpC1 y ClpC2); mientras que en las micobacterias poseen dos homólogos de ClpP (ClpP1 y ClpP2) (Leodolter *et al*, 2015).

2.1.1.1. ClpC1 en plantas

En cianobacterias, se ha demostrado que ClpC está relacionada a la viabilidad celular y al crecimiento fototrófico (Clarke & Eriksson, 1996). En las plantas, ClpC se localiza en los cloroplastos y existen 2 parálogos ClpC1 y ClpC2. ClpC1 se encarga de la degradación de proteínas mal plegadas que normalmente se dirigen a las membranas de los tilacoides (Halperin *et al*, 2001). En *Nicotiana benthamiana*, la supresión simultánea de ClpC1 y ClpC2 produce cambios morfológicos en los cloroplastos, los cuales conducen a un desequilibrio de las vías fotosintéticas y biosintéticas de la planta (Ali & Baek, 2020).

En *Arabidopsis thaliana*, ClpC1 está involucrada en la biogénesis del fotosistema I y II, además de su papel general de chaperona, cumple un rol vital en la homeostasis del cloroplasto (Sjögren *et al*, 2004). Igualmente, ClpC1 podría controlar la transcriptómica de *A.thaliana* mediante el mantenimiento de la homeostasis del proteoma (Shoudong Zhang *et al*, 2018).

2.1.1.2. ClpC1 en bacterias

A comparación de otros organismos, ClpC1 en las bacterias presenta múltiples funciones. Se ha demostrado que en *Listeria monocytogenes*, ClpC1 desempeña un papel fundamental en la supervivencia de la bacteria dentro de los macrófagos, así como en su virulencia (Rouquette *et al*, 1998). En *Staphylococcus aureus* ClpC1 regula la transcripción del gen *codY*, el cual es el principal factor de virulencia (Luong *et al*, 2011). En *B. subtilis* ClpC1 controla la expresión de genes de competencia, esporulación, motilidad y formación de biofilm, y supervivencia en condiciones estresantes (Luong *et al*, 2011). Sin embargo, a diferencia de otras bacterias, en *B.subtilis* ClpC1 necesita de proteínas adaptadoras como MecA, YpbH y McsB para poder reconocer qué sustrato se degradará durante la proteólisis (Elsholz *et al*, 2017). Figura 1.

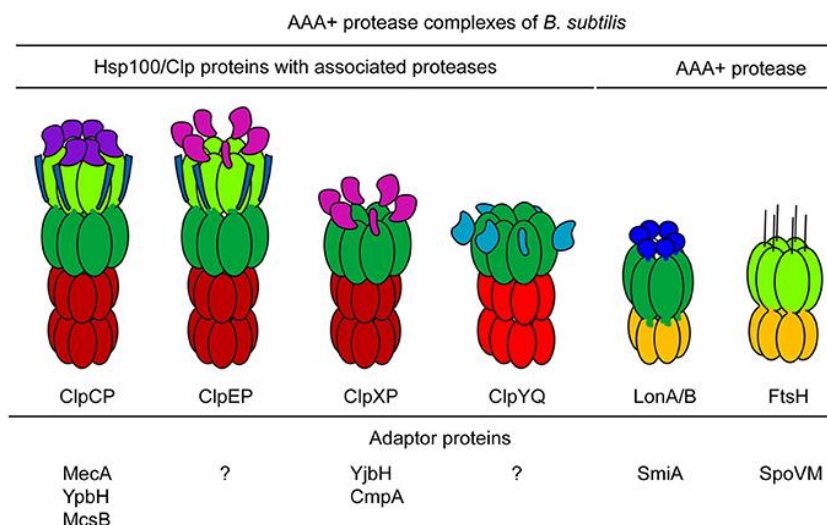


Figura 1. Proteasas AAA + y sus respectivas proteínas adaptadoras en *B. subtilis*.

2.1.2. ClpC1 en *M.tuberculosis*

En *M.tuberculosis*, ClpC1 es codificada por el gen *Rv3596c*, posee 848 aminoácidos y presenta una función conjunta con el complejo de proteasas ClpP1/P2 (Kar *et al*, 2008). ClpC1 reconoce las proteínas a degradar, y debido a su actividad ATPasa, transloca a las proteínas en el complejo proteolítico ClpP1/P2 en donde la proteína es degradada (Lee & Suh, 2016; Leodolter *et al*, 2015) (Figura 2). La función principal de este complejo de proteasas es el control de calidad de las proteínas debido a que degrada proteínas de vida corta, mal plegadas y dañadas ya sea por un estrés químico o térmico, la degradación de dichas proteínas es indispensable para mantener la homeostasis y las actividades metabólicas óptimas de la micobacteria (Lee & Suh, 2016). La mala función de este complejo genera altos niveles de proteínas mal plegadas lo que es tóxico para la micobacteria y conlleva a la muerte celular (McGillivray *et al*, 2015). El sistema de degradación y reciclaje de proteínas es importante durante el estado de latencia debido a que es una de las principales redes regulatorias en *M.tuberculosis* para responder a factores de estrés y con ello promover su supervivencia en condiciones específicas de estrés (McGillivray *et al*, 2015).

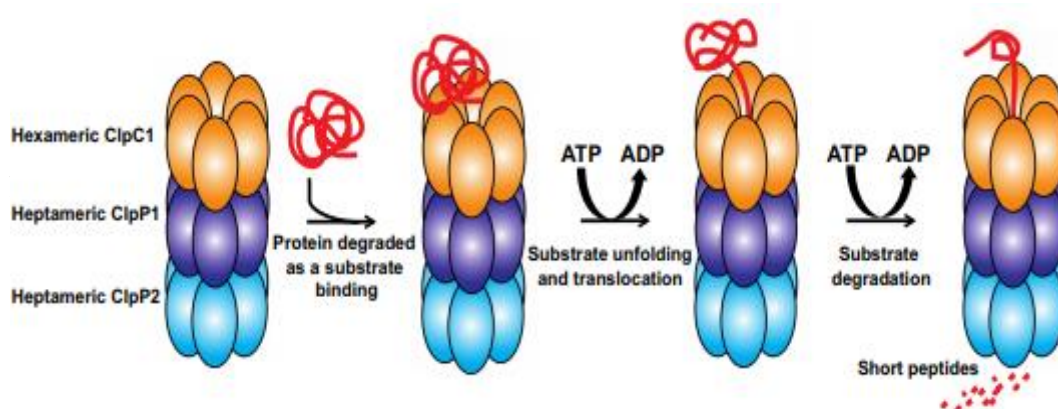


Figura 2. Mecanismo de regulación de la homeostasis proteica de ClpC1/P1/P2 en *M.tuberculosis*. (Lee & Suh, 2016).

Por otro lado, se ha demostrado que el *knockout* de ClpC1 en *M.tuberculosis* daña severamente el crecimiento bacteriano mientras que al introducir un plásmido que contiene ClpC1 el crecimiento se restaura, por ello se considera a esta proteína esencial en *M.tuberculosis* tanto *in vitro* como *in vivo* (McGillivray *et al*, 2015; Sassetti *et al*, 2003).

Con respecto a la estructura de ClpC1, no se ha cristalizado la proteína en su totalidad, pero se ha logrado cristalizar el dominio N- terminal en alta resolución en unión con la droga cyclomarina A (Vasudevan *et al*, 2013). Asimismo, se ha cristalizado el dominio N-terminal de ClpC1 con el fármaco rufomicina (Wolf *et al*, 2019).

2.1.2.1. ClpC1 como blanco terapéutico de drogas antituberculosas

ClpC1 es un blanco atractivo de fármacos antituberculosos, como es el caso de lassomicina, cyclomarina A, ecumicina y rufomicina (Choules *et al*, 2019; Gao *et al*, 2015; Gavrish *et al*, 2014; Vasudevan *et al*, 2013; Weinhäupl *et al*, 2018; Wolf *et al*, 2019). Se postula que la unión de lassomicina a ClpC1 aumenta su actividad ATPasa; sin embargo, desacopla la transferencia de energía de dicha hidrólisis a la actividad proteolítica de la proteína (Gavrish *et al*, 2014), de manera que no solo afecta el sistema de degradación de proteínas, sino también genera un mayor gasto innecesario de ATP. Cyclomarina A conduce a una proteólisis descontrolada en el complejo ClpC1/P1/P2 (Vasudevan *et al*, 2013). Eumicina estimula la actividad ATPasa de ClpC1, pero bloquea la degradación de proteínas en el complejo ClpC1/ P1/ P2 (Gao *et al*, 2015). Mientras que

rufomicina disminuye la función proteolítica en el complejo ClpC1/ P1/ P2, pero no tiene ningún efecto en la actividad ATPasa de ClpC1 (Choules *et al*, 2019; Wolf *et al*, 2019).

2.1.2.2. ClpC1 como posible blanco del ácido pirazinoico

Se ha observado que la proteína ClpC1 está asociada a la resistencia a PZA, debido a que al inducir resistencia a POA en cepas de *M.tuberculosis in vitro* con *pncA*- gen que codifica la proteína pirazinamidasa- wild type (WT), estas seleccionan mutaciones en el gen *Rv3596c*; por lo cual se postula que puede estar involucrado directamente en el mecanismo de acción de POA (Yee *et al*, 2016; Shuo Zhang *et al*, 2017). Igualmente, se ha reportado *in vivo*, mutaciones tanto en *Rv3596c* y *panD* (gen que codifica la proteína aspartato descarboxilasa PanD) que causan resistencia a POA (Gopal *et al*, 2017). La mayoría de las mutaciones en *Rv3596c*, en cepas resistentes a PZA y con *pncA* WT, se encuentran en el dominio N-terminal de ClpC1, el cual estaría encargado de reconocer proteínas mal plegadas, iniciando así el proceso de degradación (Schmitz & Sauer, 2014); sin embargo, no se ha demostrado si tales mutaciones ocasionan una actividad disfuncional en ClpC1 o en el complejo ClpC1/ P1/P2 o si poseen relación con la interacción PanD-ClpC1. El resto de las mutaciones se encuentran en posibles sitios de interacción para la formación de su estructura cuaternaria (Yee *et al*, 2016). Del mismo modo, no se ha demostrado si existe interacción directa entre ClpC1 y POA. Estudios recientes sugieren fuertemente que POA inhibe la función de PanD (Gopal *et al*, 2017; Sun *et al*, 2020). Recientemente, Gopal *et al* (2020) postularon un mecanismo de acción en donde demuestran que POA bloquea la biosíntesis de CoA al promover la degradación de la proteína PanD por ClpC1/P, esta aproximación indica que ClpC1 contribuye de manera indirecta en el mecanismo de la droga (Gopal *et al*, 2019, 2020a) (Figura 3). Sin embargo, este mecanismo no invalida que POA tenga otros blancos de unión o que POA tenga un efecto directo sobre ClpC1. Hasta el momento, no existen reportes funcionales sobre el efecto directo de POA sobre ClpC1.

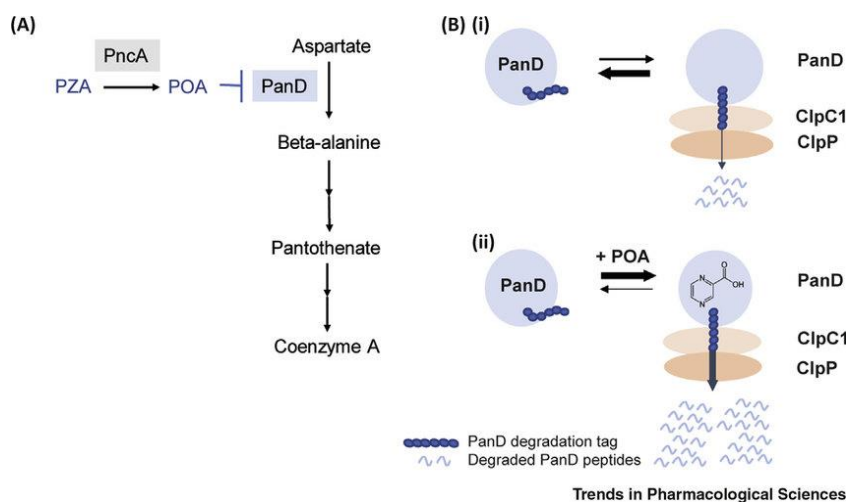


Figura 3. Mecanismo de acción propuesto de PZA/POA sobre PanD y ClpC1 (Gopal *et al*, 2019).

2.2 Cristal de ClpC1

ClpC1 de *B. subtilis* ha sido cristalizada por la metodología de “hanging-drop vapour diffusion” y presenta un 63% de identidad con el dominio N-terminal de ClpC1 de las micobacterias, en *M. tuberculosis* la identidad es del 60% (Vasudevan *et al*, 2013; Wang *et al*, 2011). La estructura cristalizada presenta la proteína adaptadora MecA (Wang *et al*, 2011). A diferencia de *B. subtilis* y la mayoría de bacterias gram-positivas, *M. tuberculosis* no requiere una proteína adaptadora para el correcto funcionamiento de ClpC1 (Vasudevan *et al*, 2013). No obstante, se sugiere que existe un tipo de degradación de proteínas mediada por una proteína adaptadora en ClpC1 (Marsee *et al*, 2018). De la misma manera, se ha logrado cristalizar solo el dominio N-terminal de ClpC1 de *M. tuberculosis* en unión con ecumicina y rufomicina (Vasudevan *et al*, 2013; Wolf *et al*, 2019).

2.3 Expresión de ClpC1 de *M. tuberculosis* en *E. coli*

ClpC1 de *M. tuberculosis* se ha expresado de manera recombinante en múltiples sistemas de expresión de *E. coli* en diferentes cepas, plásmidos, concentraciones de IPTG y condiciones de tiempo y temperatura. pET30a(+) y pET28a(+) son los plásmidos más utilizados para su expresión, mientras que la cepa más utilizada es BL21(DE3) (Tabla 1).

Tabla 1. Expresión de ClpC1 de *M. tuberculosis* a diferentes condiciones en *E. coli*.

Condiciones de expresión de ClpC1 de <i>M. tuberculosis</i>				
Plásmido	Célula	Temperatura y tiempo	[IPTG]	Referencia
p7XC3H FX	Rosetta(DE3)	20°C/ON	0.1 mM	(Leodolter <i>et al</i> , 2015)
pET-28a(+)	BL21-Gold (DE3)	37°C/4h	0.5 mM	(Kandror <i>et al</i> , 2014)
pET30(a)	BL21-Gold (DE3)	37°C/4h	0.25 mM	(Choules <i>et al</i> , 2019)
pET30(a)	BL21(DE3)	18°C/ON	1 mM	(Vasudevan <i>et al</i> , 2013)
pET-24a(+)	BL21(DE3)	18°C/ON	0.5 mM	(Marsee <i>et al</i> , 2018)
pET28b	BL21(DE3)	18°C/ON	0.5 mM	(Lunge <i>et al</i> , 2020)
pET30(a)	BL21C+	37°C/4h	1 mM	(Wolf <i>et al</i> , 2019)

2.4 Problema de la expresión soluble de proteínas recombinantes en *E. coli*.

La producción de proteínas recombinantes en *E.coli* es ampliamente utilizada en la industria como la investigación, sin embargo; existe un problema en común en el estudio de las proteínas debido en *E. coli* debido a que estas tienden a precipitar, en otras palabras, en acumularse en su forma insoluble e inactiva (Singh *et al*, 2015). Entre los factores que ocasionan la insolubilidad de las proteínas en *E.coli* se encuentra el tipo de sistema de expresión utilizado; naturaleza de la secuencia de la proteína, naturaleza del inductor, tiempo y temperatura de expresión (Wilkinson & Harrison, 1991); la alta concentración de la proteína sobre expresada (Makrides, 1996); altos niveles de glutatión en el citoplasma lo que previene la formación de enlaces disulfuro en las proteínas de interés (Lilie *et al*, 1998); falta de accesibilidad de chaperonas para el correcto plegamiento de las proteínas (Mogk *et al*, 2010); ineficiente traducción de proteínas que contienen en su secuencia

codones raros para *E.coli* (Komar *et al*, 1999); y falta de proteínas fusión que aumentan la solubilidad (Stevens, 2000).

Asimismo, se sabe que la solubilidad de las proteínas recombinantes depende principalmente de su interacción con el solvente (Schein, 1990). El aumento de las interacciones hidrofóbicas en las proteínas interrumpe las interacciones intermoleculares con su medio, lo que ocasiona la separación de las proteínas con la solución y la desnaturalización de estas (Schein, 1990). Se sabe que proteínas con un alto porcentaje de regiones hidrofóbicas en su superficie al sobre expresarlas facilitan la formación de agregados proteicos debido a que las interacciones hidrofóbicas entre ellas son mayores que las interacciones electrostáticas (interacciones con el solvente) (Idicula-Thomas & Balaji, 2005). Por ende, para producir proteínas de forma soluble estas deben interactuar tanto como sea posible con el solvente.

Es posible revertir este problema utilizando agentes denaturantes como la úrea y convertir una proteína insoluble en una proteína nativa biológicamente activa; sin embargo, estos procedimientos requieren un tiempo, costo adicional y solo un bajo porcentaje de proteína solubilizan (Drew *et al*, 2003).

2.4.1. Efecto de la solubilidad según el tipo de plásmido

2.4.1.1. Plásmido pET28a(+)

pET28a(+) es un vector de clonación y expresión de proteínas recombinantes, contiene al promotor T7 y operador lac, los cuales controlan la expresión de la proteína de interés mediante la adición de IPTG. Al añadir IPTG, este bloquea al represor del operon lac, *lacI* lo que ocasiona la expresión de la ARN polimerasa T7, esta se une al promotor T7 y se induce la expresión de la proteína recombinante.

En adición, los vectores pET-28a-c (+) llevan una configuración His-Tag N-terminal/trombina/T7•Tag más una secuencia Tag His-terminal opcional C-terminal. Se ha demostrado que el uso de pET28a(+) aumenta significativamente la producción de una proteína recombinante, sin embargo no posee un efecto en la cantidad de proteína soluble que se puede obtener (Shilling *et al*, 2020).

2.4.1.2. Plásmido pMAL-cHT

pMAL-cHT es ampliamente utilizado para mejorar la cantidad y solubilidad de las proteínas recombinantes. Esto se debe a que adiciona en el gen de interés el gen *malE* que codifica la proteína de unión a maltosa (MBP) en el dominio C-terminal mediante una secuencia de aminoácidos que generalmente es una secuencia de corte (Sun *et al*, 2011; Walls & Loughran, 2011). La expresión de la proteína de interés en pMAL-cHT está controlada por el promotor sintético *tac*, el cual es un promotor fusión entre el promotor *lac* y *trp*. Por ello, es inducible por IPTG (isopropil- β -D-1-tiogalactopiranosido); se ha demostrado que el promotor *tac* permite una expresión más alta de la proteína de interés a comparación de los promotores *lac* y *trp*. Sin embargo, estudios reportan que el promotor *tac* es menos eficiente que el promotor T7, presente en pET28a(+) (de Boer *et al*, 1983).

2.4.1.2.1. Maltose-binding protein (MBP)

MBP es una proteína periplásmica que forma parte de una vía metabólica de *E. coli*, la cual es responsable de la absorción y el catabolismo de las maltodextrinas (Nikaido, 1994). El mecanismo por el cual MBP favorece la expresión soluble de proteínas aún no está dilucidado, pero se postula que la proteína fusión (proteína de interés + MBP) es más afín a las chaperonas GroELS y con ello se facilita su correcto plegamiento y evita la formación de agregados proteicos (Bernier *et al*, 2018; Kapust & Waugh, 1999; Raran-Kurussi & Waugh, 2012). Múltiples estudios demuestran su efectividad en aumentar la solubilidad de proteínas recombinantes en *E.coli* (Bernier *et al*, 2018; Kapust & Waugh, 1999; Raran-Kurussi & Waugh, 2012). Sin embargo, es necesario cortar MBP de la proteína de interés para futuros estudios funcionales o estructurales. La proteasa TEV, del virus del grabado de tabaco, es una cisteína proteasa que se utiliza para separar MBP de proteínas al reconocer la secuencia de aminoácidos Glu- Asn- Leu- Tyr- Phe- Gln- (Gly/Ser) y corta entre Gln- (Gly/Ser) (Raran-Kurussi *et al*, 2017).

2.4.2. Efecto de la solubilidad según el tipo de cepa de *E. coli*.

BL21(DE3)pLysS al igual que las cepas Rosetta(DE3)pLysS y Lemo(DE3)pLysS derivan de la cepa BL21(DE3) por ello comparten características similares como la

denominación DE3, pLysS y la deficiencia de las proteasas OmpT y lon; estas proteasas reducen la producción de las proteínas recombinantes, por ende su deficiencia resulta en la acumulación de proteínas recombinantes dentro de la célula de expresión (Pan & Malcolm, 2000). La denominación DE3, significa que la cepa presenta al λ DE3 lisogénico y en consecuencia presenta una copia cromosómica del gen de la ARN polimerasa T7, la cual está bajo control del promotor lacUV5 (Studier *et al*, 1990). Asimismo, las cepas que presentan el plásmido pLysS expresan la lisozima T7, la cual suprime la expresión basal de la ARN polimerasa T7 antes de la inducción y de esta manera se controla la expresión de la proteína de interés (Studier, 1991); en adición, el plásmido confiere resistencia a cloranfenicol. El sistema de expresión se completa al utilizar vectores pET, los cuales presentan el promotor T7 antes de la secuencia de interés, en donde se une la ARN polimerasa T7 al añadir IPTG (inhibidor del represor del promotor lacUV5, lacI) (Studier *et al*, 1990).

2.4.2.1. Cepa BL21(DE3)pLysS

Las cepas BL21(DE3)pLysS se utilizan para expresar proteínas tóxicas, particularmente las proteínas insolubles que forman cuerpos de inclusión, para el hospedero debido a que mejoran la solubilidad de la proteína de interés al tener una expresión controlada por el promotor y la lisozima T7 (Angius *et al*, 2018).

2.4.2.2. Cepa Rosetta (DE3)pLysS

Las cepas Rosetta(DE3)pLysS se utilizan para aumentar la expresión de proteínas eucariotas que presentan en su mRNA codones que *E.coli* no utiliza; por lo tanto estas cepas presentan en su citoplasma tARN codones inusuales para *E.coli* que asegura la expresión exitosa de la proteína de interés.

2.4.2.3. Cepa Lemo (DE3)pLySs

Las cepas Lemo(DE3)pLysS - cuyo nombre significa “less is more”- una menor inducción de la expresión resulta en una mayor cantidad de proteína producida y se utilizan para expresar proteínas de membrana, secretoras e insolubles (Schlegel *et al*,

2012). Su principal característica es que utiliza dos plásmidos para expresar proteínas: pLemo y pET. El vector de expresión pET - en este trabajo se utilizó pET28a(+)- se basa en el promotor T7 mientras que pLemo se utiliza para la expresión basada en el promotor PrhaBAD, el cual es inducible por L-ramnosa, para producir la lisozima T7 (Kuipers *et al*, 2017). El mecanismo por el cual se expresan las proteínas se fundamenta en la actividad de la ARN polimerasa T7, la cual es modulada por la expresión de la lisozima T7, el gen que codifica esta lisozima se encuentra en pLemo- y de esta manera se regula la expresión de la proteína de interés.

2.4.3. Efecto de la solubilidad según la concentración de IPTG.

Un factor importante que mejora la solubilidad de las proteínas es la concentración de IPTG. Este inductor se agrega al medio de cultivo cuando se ha alcanzado una densidad óptica de 0.4- 0.6 (OD 600 nm: 0.4- 0.6), lo que indica un crecimiento exponencial en consecuencia, expresión de la proteína de interés (Donovan *et al*, 1996). El IPTG controla al represor del promotor *lac*, lo cual induce a la expresión del ARN polimerasa T7, luego ésta se une al promotor T7 que se encuentra en los plásmidos pET y de esta manera se comienza la expresión de la proteína de interés. Sin embargo, existe un nivel basal de expresión de la proteína sin la adición de IPTG, lo que significa un problema para proteínas tóxicas debido a que se evita que el cultivo celular alcance la densidad óptica deseable en un tiempo adecuado. No obstante, este problema se soluciona con la lisozima T7, cuyo gen se encuentra en el plásmido pLysS (Studier, 1991).

2.4.4. Efecto de la solubilidad según el tiempo y temperatura de expresión

El tiempo, así como la temperatura de expresión, aumentan la cantidad de proteína soluble expresada dependiendo del sistema y el tipo de proteína expresada (Vera *et al*, 2007). Se sabe que temperaturas bajas de expresión aumentan la solubilidad de las proteínas al alterar el recambio metabólico de la bacteria y de esta manera garantizar la disponibilidad de chaperonas para el correcto plegamiento de la proteína de interés. Se ha demostrado que el correcto plegamiento de las proteínas depende de la energía de activación y a temperaturas bajas se ve favorecida (Naganathan *et al*, 2007).

2.5. Modelamiento estructural de proteínas

El uso de la bioinformática para predecir y modelar proteínas ha aumentado en los últimos años. El uso de programas como Swiss PDB Viewer, Visual Molecular Dynamics (VMD) y I-TASSER se utiliza principalmente para modelar la estructura globular de las proteínas y hallar sus propiedades fisicoquímicas como la hidrofobicidad (porcentaje de regiones hidrofóbicas en la superficie de una proteína), así como para calcular el porcentaje de similitud e identidad entre dos estructuras (Guex & Peitsch, 1997; Kopp & Schwede, 2004; Waterhouse *et al*, 2018). El porcentaje de identidad entre dos estructuras proteicas se define como el porcentaje de conservación a nivel de la secuencia de aminoácidos entre dos proteínas, mientras que el porcentaje de similitud compara las propiedades fisicoquímicas en cada aminoácido de dos secuencias primarias de dos proteínas (Guex & Peitsch, 1997; Kopp & Schwede, 2004; Waterhouse *et al*, 2018). En adición, los valores de GMQE (*Global Model Quality Estimation*), QMEAN (*Quality score*) y RMSD (*Root-Mean-Square Deviation of atomic positions*) son muy importantes para predecir la calidad de un modelo. De la misma manera, el servidor web TM-align usa un algoritmo para comparar la estructura terciaria de dos proteínas al superponerlas, TM-align brinda el TM-score que indica el nivel de concordancia entre dos estructuras, TM-score tiene valores entre 0 y 1 (Zhang & Skolnick, 2005).

2.5.1. Parámetros de evaluación del modelo

2.5.1.1. GMQE

Se define como la estimación de la calidad entre el modelo o plantilla (cristal de una proteína) y la proteína a modelar experimentalmente, sus valores se encuentran entre 0 y 1 (Benkert *et al*, 2011; Waterhouse *et al*, 2018).

2.5.1.2. QMEAN

Evalúa tanto la calidad global (toda la estructura) como la local (por residuo), sus valores alrededor de cero indican una alta concordancia entre el modelo y la proteína

experimental, valores menores a -4.0 nos indican una baja calidad del modelamiento (Benkert *et al*, 2011; Waterhouse *et al*, 2018).

2.5.1.3. RMSD

Nos indica una medida cuantitativa de similitud entre los átomos de dos o más estructuras proteicas; por ello, mientras menor sea el valor de RMSD mejor es el modelo estimado con respecto a la plantilla (Sargsyan *et al*, 2017).

2.5.2. Parámetros de solubilidad

Para evaluar los parámetros de solubilidad de una proteína modelada *in silico*, se utiliza el porcentaje de residuos hidrofóbicos e hidrofílicos, SASA (*Solvent-accessible surface area*), y el software protein-Sol. SASA se define como el área de superficie accesible al solvente y su valor se determina en el programa VMD (Ali *et al*, 2014); mientras que protein-Sol predice la solubilidad de una proteína y nos indica el valor de QuerySol, el cual si es mayor a 0.45 la proteína tendrá una solubilidad más alta que la proteína de *E. coli* soluble promedio (Hebditch *et al*, 2017).

3. PLANTEAMIENTO DEL PROBLEMA Y OBJETIVOS

3.1. Planteamiento del problema

La proteína ClpC1 es una proteasa responsable de la degradación de péptidos mal plegados. ClpC1 pertenece a la familia de las chaperonas de shock térmico *HSP 100* y está presente en múltiples organismos; y en las bacterias juega un rol fundamental en su virulencia y supervivencia (Elsholz *et al*, 2017; Luong *et al*, 2011; Rouquette *et al*, 1998; Turgay *et al*, 1997).

ClpC1 es esencial para *M. tuberculosis* y por ello, es blanco de fármacos antituberculosos y posiblemente interviene en el mecanismo de acción de la PZA, antibiótico que forma parte del tratamiento de primera línea, efectivo contra las micobacterias en estado latente (Njire *et al*, 2016). Se ha mostrado que ClpC1 incrementa la degradación de la aspartato descarboxilasa (PanD) cuando POA, la molécula activa de

PZA, se une a PanD; sin embargo, no queda claro el modo de acción de ClpC1 (Gopal *et al*, 2019, 2020b). Las evidencias muestran que POA se estaría uniendo a PanD pero se desconoce la relación de POA con ClpC1 así como los residuos de PanD y ClpC1 que estarían participando en este mecanismo. El conocimiento de los detalles de este mecanismo permitiría diseñar nuevos fármacos antituberculosos y mejorar la detección de cepas resistentes a PZA.

En el estudio de la participación de ClpC1 en el mecanismo molecular de POA-PanD se requiere la proteína recombinante en forma activa debido a que se estudiarán las interacciones con otras moléculas; y eso se garantiza obteniendo la proteína en forma soluble. Las proteínas recombinantes pueden producirse en un porcentaje de solubilidad; esto va a depender de varios factores como la composición de sus aminoácidos, el sistema de expresión, el tiempo y temperatura de inducción, la concentración del inductor, entre otros (Wilkinson & Harrison, 1991; Idicula-Thomas & Balaji, 2005). Si una proteína es expresada principalmente en forma insoluble, debe ser producida en forma denaturada y solubilizada posteriormente a través de un complejo proceso con baja productividad. La determinación de la solubilidad de una proteína recombinante es importante debido a que va a permitir conocer la cantidad de proteína recombinante activa y por lo tanto diseñar de una manera eficiente todos los experimentos necesarios.

Previamente, ClpC1 de *M. tuberculosis* fue producido en el sistema de *E. coli* con el plásmido pET28a(+) pero no hay evidencia acerca del rendimiento (concentración) y porcentaje de solubilidad que se obtiene en estas condiciones (Kandror *et al*, 2014). De la misma manera, se trató de replicar esta expresión utilizando la cepa BL21(DE3)pLysS, en donde se obtuvo ClpC1 de manera insoluble.

La obtención de ClpC1 de *M. tuberculosis* de forma soluble en *E. coli* ayudaría a reducir los costos, tiempo y a acelerar futuros estudios sobre su función; ya sea dentro del mecanismo de POA u otras drogas.

En el presente trabajo se expresó ClpC1 de *M. tuberculosis* en diferentes sistemas de expresión de *E. coli*, tiempo y temperatura de inducción y concentración del inductor (IPTG) para determinar las mejores condiciones en donde se obtiene mayor concentración y ClpC1 soluble.

3.2. Objetivos generales

- Comparar el rendimiento (concentración) y porcentaje de solubilidad estimados de la proteína ClpC1 de *M. tuberculosis* expresada de manera recombinante en diferentes sistemas de expresión en *E.coli*.
- Evaluar los parámetros de solubilidad del modelo bioinformático de ClpC1 y ClpC1- MBP.

3.3. Objetivos específicos

- Clonar el gen *Rv3596c* en los plásmidos pET28a (+) y pMAL-cHT en *E. coli*.
- Expresar ClpC1 de *M. tuberculosis* en pET28a(+)-ClpC1 y pMAL-cHT-ClpC1 en diferentes cepas de *E.coli*, concentración de IPTG, tiempo y temperatura de expresión.
- Comparar la concentración y porcentaje de solubilidad de ClpC1 recombinante en las todas condiciones de expresión.
- Verificar la solubilidad de la proteína ClpC1 al remover MBP de ClpC1-MBP.
- Modelar *in silico* ClpC1 y ClpC1-MBP por homología usando como base el monómero de ClpC1 de *B. subtilis* para hallar los parámetros de solubilidad de ambas proteínas.

4. METODOLOGÍA

4.1. Diseño de primers *Rv3596c*

Se diseñaron los primers forward y reverse para el clonar el gen *Rv3596c* en los vectores pET28a(+) y pMAL-cHT utilizando el protocolo *In-Fusion* (Takara Bio- USA) (Laboratories Inc, 2011). Cada primer contiene secuencias de corte de enzimas específicas, *NcoI* (forward) y *HindIII* (reverse) para el vector pET28a(+) y *EcoRI* (forward) y *HindIII* (reverse) para el vector pMAL-cHT. Se evaluó longitud, contenido de GC y la temperatura de melting en la herramienta OligoAnalyzer 3.1 (Owczarzy *et al*, 2008).

4.2. Amplificación del gen *Rv3596c*

Se realizó un PCR convencional para la amplificación del gen *Rv3596c* en un termociclador T100 Thermal Cycler (Bio-Rad). La reacción de amplificación se realizó en un volumen final de 20 µl con 0.4U Phusion DNA Polymerase (Thermo Fisher Scientific), 1.5 mM de MgCl₂, 200 µM de cada dNTPs, 0.5 µM de cada primer, y 50 ng del ADN de *M. tuberculosis*. La amplificación se logró mediante un ciclo de denaturación inicial a 98°C por 2 minutos, 30 ciclos de denaturación a 98°C por 45 segundos, hibridación a 72°C por 30 segundos y extensión a 72°C por 1 minuto, y finalmente un último ciclo de extensión final a 72°C por 10 minutos. La amplificación se corroboró a través de una electroforesis de gel de agarosa al 1 % y TAE 1X a un voltaje constante de 95 V por 1 hora.

4.3. Clonación con *In-fusion* del gen *Rv3596c*

El vector de clonación pET28a(+) se digirió con las enzimas de restricción *NcoI* y *HindIII* mientras que el vector pMAL-CHT con *EcoRI* y *HindIII*. Se corroboró la digestión mediante electroforesis en gel de agarosa al 1%. Los fragmentos de DNA correspondientes a los plásmidos digeridos se purificaron por el kit Zymoclean Gel DNA Recovery Kit (Zymo Research). Finalmente, estos fueron ligados y clonados siguiendo las indicaciones de fabricante del protocolo de In-Fusion 5X (NEB) (Laboratories Inc, 2011). Las células de *E. coli* Novablue fueron tratadas con CaCl₂ para luego ser transformadas con los plásmidos mediante shock térmico. Luego, para la verificación de las colonias transformadas, se seleccionaron aleatoriamente 7 – de cada plásmido clonado- para extraer y digerir los plásmidos con cada enzima mencionada. Posteriormente, se verificó la digestión a través de una electroforesis en gel de agarosa al 1%.

4.4. Secuenciación del gen *Rv3596c*

La secuenciación de los plásmidos recombinantes se realizó por el método de Sanger (Macrogen USA). La tabla 2 muestra los primers utilizados para cada plásmido clonado.

Los datos de la secuenciación fueron analizados en el programa *Molecular Evolutionary Genetics Analysis* (MEGA) (<https://www.megasoftware.net/docs>).

Tabla 2. Lista de primers utilizados para la secuenciación de *Rv3596c* en pMAL-cHT y pET28a(+).

Plásmido	Primers de secuenciación
pMAL-cHT	<i>Rv3596c_Sec_Fw_5'</i> -GCTCTCCGTCTACGTCCTTG-3'
	<i>Rv3596c_Sec_Fw2_5'</i> -GAGTTCGATGAGAAGATCGCCG-3'
	pMAL Fw 5' TGTCGATGAAGCCCTGAAAGACGC 3'
	pMAL Rv 5' AAC TGT TGG GAA GGG CGA TCG GTG 3'
pET28a(+)	<i>Rv3596c_Sec_Fw_5'</i> -GCTCTCCGTCTACGTCCTTG-3'
	<i>Rv3596c_Sec_Fw2_5'</i> -GAGTTCGATGAGAAGATCGCCG-3'
	pET28a(+) Fw 5' AGGAGATATA CCATG GGC TTC GAA CGA TTT ACC GAC 3'
	pET28a(+) Rv 5' GTGCGGCCGCAAGCTTCCGCGCGGCCGG 3'

4.5. Expresión de ClpC1 recombinante

Se transformaron las cepas de *E. coli* BL21(DE3)pLysS, Rosetta(DE3)pLysS y Lemo(DE3)pLysS con los plásmidos recombinantes obtenidos en la sección 5.3. Para la expresión de ClpC1 en pET28a (+) y pMAL- cHT, se cultivó una colonia de las células transformadas en 5 ml de caldo LB con kanamicina (40 µg/ ml) y cloranfenicol (30 µg/ml) para pET28a(+)-ClpC1, y una colonia (de cada cepa) en 5 ml de caldo LB con ampicilina (100 ng/ ml) y cloranfenicol (30 µg/ml) para pMAL-cHT-ClpC1 a 37 °C 225 rpm durante 18 horas. Luego se transfirió 2.5 ml del cultivo a un matraz con 50 ml de caldo LB con los respectivos antibióticos a 37°C en agitación hasta llegar a un OD₆₀₀ de 0.6-0.8. Se tomó 1 ml de muestra de caldo, para utilizarlo como control (no inducido), y se indujo la

expresión con IPTG a una concentración final de 0.25, 0.5, 1 y 2 mM. Se evaluó la expresión incubando los cultivos por 4, 4 y 10 horas, a 30°C, 37°C y 16°C respectivamente, para cada concentración de IPTG y sistema de expresión. Transcurrido el tiempo de inducción se tomó 500 µL de cada cultivo para utilizarlo como control inducido. En cada condición ClpC1 fue expresada por triplicado. La Tabla 3 muestra las diferentes condiciones probadas.

Tabla 3. Tabla resumen de las diferentes condiciones probadas en cada expresión.

Tipo de cepa de <i>E. coli</i>	Tipo de plásmido	Temperatura (C°)/ Tiempo (h)	Concentración de IPTG (mM)			
Bl21(DE3)pLysS	pET28a(+)-ClpC1	16/10	0.25	0.5	1	2
		30/4	0.25	0.5	1	2
		37/4	0.25	0.5	1	2
	pMAL-cHT-ClpC1	16/10	0.25	0.5	1	2
		30/4	0.25	0.5	1	2
		37/4	0.25	0.5	1	2
Lemo(DE3)pLysS	pET28a(+)-ClpC1	16/10	0.25	0.5	1	2
		30/4	0.25	0.5	1	2
		37/4	0.25	0.5	1	2
	pMAL-cHT-ClpC1	16/10	0.25	0.5	1	2
		30/4	0.25	0.5	1	2
		37/4	0.25	0.5	1	2
Rosetta(DE3)pLysS	pET28a(+)-ClpC1	16/10	0.25	0.5	1	2
		30/4	0.25	0.5	1	2
		37/4	0.25	0.5	1	2
	pMAL-cHT-ClpC1	16/10	0.25	0.5	1	2
		30/4	0.25	0.5	1	2
		37/4	0.25	0.5	1	2

4.6. Estimación de la concentración relativa de ClpC1 recombinante en los sistemas de expresión.

La estimación de la concentración relativa de ClpC1 se realizó comparando la concentración de las proteínas en la fracción no inducida y extracto crudo. La concentración de proteínas fue determinada después de normalizar las muestras por la cantidad de sedimento bacteriano. La normalización se realizó centrifugando las fracciones a 13 000 rpm por 1 minuto; el sobrenadante fue eliminado y el sedimento fue resuspendido en un volumen de binding buffer (20 mM Tris-HCL, 200mM NaCl y 1mM DTT pH 7.4) en proporción al tamaño observado (Figura 4).

Se tomó una alícuota de 20 μ l de la fracción no inducida, la cual fue hervida a 95°C por 5 min. La fracción inducida fue sonicada 3 veces (1 segundo ON/ 1 segundo OFF por 1 minuto en intervalos de 5 min de descanso a temperatura ambiente) para obtener el extracto crudo, y se guardó una alícuota de 20 μ l, la cual fue hervida a 95°C por 5 min (Figura 4).

La concentración de proteína de la fracción no inducida y extracto crudo fue determinada mediante la técnica de Bradford. Esta técnica incluye la adición del colorante Coomassie Brilliant Blue G- 250 a las muestras y la lectura a 595 nm en el lector de ELISA (Kruger, 1994). La curva estándar se construyó usando diluciones de BSA (5 μ g/ ml, 10 μ g/ ml, 20 μ g/ml, 40 μ g/ml, 80 μ g/ ml, 100 μ g/ ml y 200 μ g/ ml). Del mismo modo, se cuantificaron las alícuotas de las fracciones no inducidas y extracto crudo mediante diluciones de 1/50, 1/100, 1/300/ y 1/500.

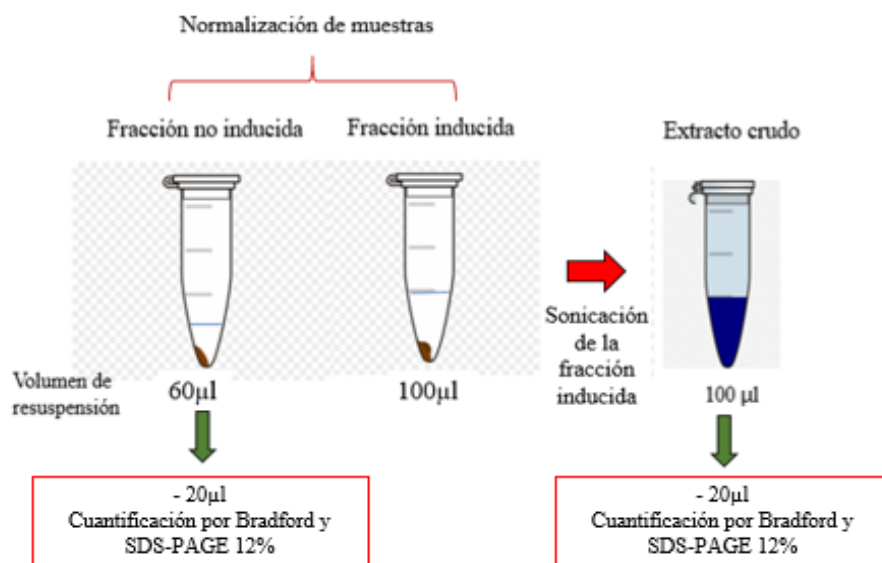


Figura 4. Normalización y determinación de la concentración de las proteínas en las fracciones no inducidas e inducidas.

Finalmente, la concentración relativa de la proteína ClpC1 se estimó de la diferencia de la concentración total de proteínas entre el extracto crudo menos la concentración total de proteína de la fracción no inducida empleando la siguiente fórmula:

$$\text{Concentración relativa de ClpC1 (mg/ml)} = F.\text{ext (mg/ml)} - F.\text{NI (mg/ml)}$$

$$F.\text{ext} = \text{Fracción del extracto crudo (mg/ml)}$$

$$F.\text{NI} = \text{Fracción no inducida (mg/ml)}$$

Cada medición se realizó por triplicado, se calculó el promedio y se comparó el resultado de cada sistema de expresión.

4.7. Estimación del porcentaje relativo de solubilidad de ClpC1 recombinante en los sistemas de expresión.

La solubilidad de ClpC1 fue estimada de manera aproximada considerando que la mayor parte de la proteína soluble corresponde a ClpC1. La solubilidad fue calculada como el porcentaje de proteína en la fracción soluble con respecto a la concentración de la proteína en el extracto crudo.

Una alícuota de 20 µl de extracto crudo fue reservada; el resto del volumen fue centrifugado dos veces a 12 500 rpm por 20 min a 4°C. El sobrenadante conteniendo la fracción soluble fue recuperado (Figura 5). El precipitado- fracción insoluble -se

resuspendió en buffer A (20 mM de Tris-HCl, 200mM NaCl y 8M urea pH 7.4). La concentración de proteínas del extracto crudo y de la fracción soluble fueron determinadas mediante la técnica de Bradford descrito anteriormente por triplicado.

El porcentaje solubilidad fue verificado en algunas muestras mediante la observación de la banda de ClpC1 en gel SDS-PAGE 12%.

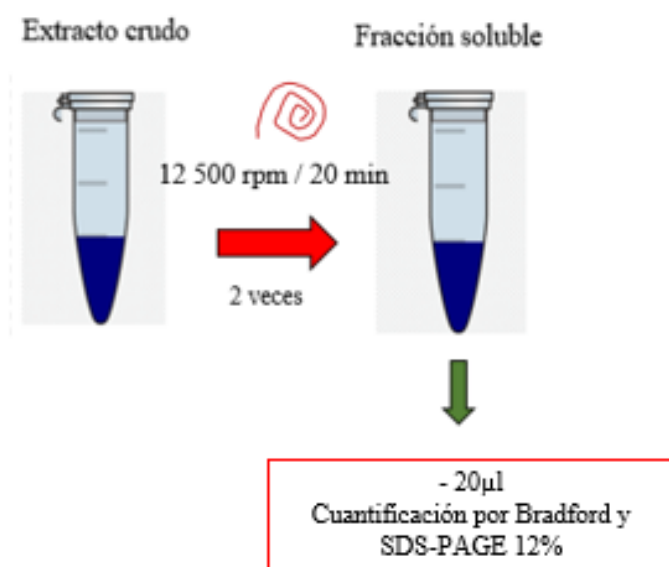


Figura 5. Procedimiento de la fracción del extracto crudo y fracción soluble.

El porcentaje relativo de solubilidad de ClpC1 fue obtenido usando la siguiente fórmula:

$$\text{Porcentaje relativo de ClpC1 soluble} = F. \text{ Sol (mg/ml)} / F. \text{ ext (mg/ml)} * 100\%$$

$$F. \text{ Sol} = \text{Fracción Soluble (mg/ml)}$$

$$F. \text{ ext} = \text{Fracción del extracto crudo (mg/ml)}$$

4.8. Análisis estadístico

Los datos estadísticos descriptivos como la media e intervalos de confianza se obtuvieron en la herramienta Statistica (StatSoft). Se calculó un intervalo de confianza al 95% (95% IC) para cada la concentración y porcentaje de solubilidad de ClpC1 en cada condición de expresión evaluada. Los gráficos de barra se realizaron en el programa Sigmaplot (Systat Software Inc.).

4.9. Purificación de ClpC1-MBP y remoción de MBP de ClpC1.

Se seleccionó la mejor condición de expresión en donde se produce mayor concentración y porcentaje de proteína soluble, para realizar una prueba piloto de purificación y corte TEV de ClpC1-MBP. Para ello, se expresó nuevamente ClpC1 por el protocolo antes descrito. La fracción soluble se filtró con un filtro de 0.45 μ M y purificó por MBPTrap HP (5ml) siguiendo el protocolo del fabricante (GE Healthcare). La fracción eluída se concentró por ultra-Amicon 100 KDa (Millipore). El corte TEV (NEB) se realizó con el protocolo del fabricante a temperaturas de 4 y 30°C por 72 y 24 horas respectivamente.

4.10. Modelamiento estructural de ClpC1 con y sin MBP.

Se modeló la estructura de la proteína ClpC1 de *M. tuberculosis* mediante el método de homología a partir de su secuencia completa (GenBank: ALB20844.1). Para ello, se usó el servidor web SwissModel con la configuración por defecto (Kopp & Schwede, 2004). Se usó como plantilla el monómero del cristal de la proteína ClpC1 de *B.subtilis*. Dicho monómero fue previamente extraído del complejo MecA121-ClpC1-485 (PDB: 3PXG; resolución de 3.654Å) (Wang *et al*, 2011) usando la herramienta VMD (<https://www.ks.uiuc.edu/Research/vmd/>).

Para evaluar la integridad del modelo por homología, se realizó un alineamiento estructural global entre este y la región N-terminal del cristal de la proteína ClpC1 de *M. tuberculosis* (PDB: 3WDB) (Vasudevan *et al*, 2013). Se usó el servidor web TM-align (<https://zhanglab.ccmb.med.umich.edu/TM-align/>) con los parámetros por defecto (Y.

Zhang & Skolnick, 2005). De la misma manera, se modeló ClpC1-MBP mediante el programa I-TASSER (<https://zhanglab.dcm.b.med.umich.edu/I-TASSER/>).

4.11. Estimación de los parámetros de solubilidad de ClpC1 con y sin MBP.

Por último, para determinar los parámetros de solubilidad de ClpC1 y ClpC1-MBP, se halló el porcentaje de residuos hidrofóbicos e hidrofílicos; así como el área total y SASA; y el QuerySol de cada modelo mediante los programas I-TASSER, VMD y protein-Sol (<https://protein-sol.manchester.ac.uk/>), respectivamente.

5. RESULTADOS

5.1. Diseño de primers para el gen *Rv3596c*

Se diseñaron dos sets de primer uno para cada vector de expresión, cada uno con diferentes características, como se muestra en la tabla 4.

Tabla 4. Características de los primers diseñados para la clonación de *Rv3596c* en los plásmidos pET28a(+) y pMAL-cHT. Sitios de restricción: *NcoI* (CCA TGG), *HindIII* (AAGCTT) y *EcoRI* (GAATTC).

Plásmido	Primer	Longitud (bp)	Tm (C°)	GC (%)
pET28a(+)	5'-AGG AGA TAT <u>ACC ATG</u> GGC TTC GAA CGA TTT ACC GAC-3'	34	75.4	73.50
	5'-GTG CGG CCG C <u>AA GCT</u> TCC GCG CGG CCG G-3'	28	77.3	82.10
pMAL-Cht	5'- TCA CCA TCA C <u>GA ATT</u> CAT GTT CGA ACG ATT TAC CGA CC 3'	38	64.4	44.70
	5'- GGC CAG TGC C <u>AA GCT</u> TCT ACC GCG CGG CCG GCT C 3'	34	75.4	73.50

5.2. Amplificación del gen *Rv3596c*

La amplificación del gen *Rv3596c* mostro una banda única de aproximadamente 2574 bp en la electroforesis empleando agarosa (figuras 6 y 7).

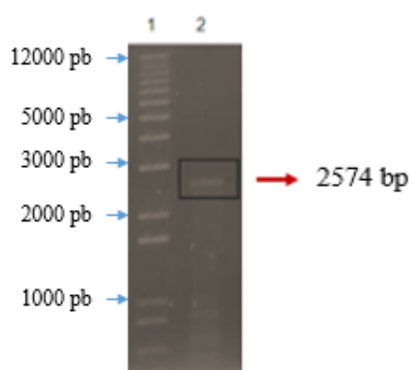


Figura 6. Amplificación del gen *Rv3596c* con los primers diseñados para el plásmido pET28a(+). Carril 1: Marcador 1Kb plus, carril 2: Gen *Rv3596c* (2547 bp).

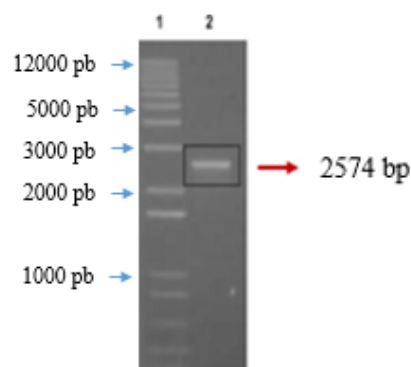


Figura 7. Amplificación del gen *Rv3596c* con los primers diseñados para el plásmido pMAL-cHT. Carril 1: Marcador 1Kb plus, carril 2: Gen *Rv3596c* (2547 bp).

5.3. Clonación con *In-fusion* del gen *Rv3596c*

Con respecto a las cepas *E. coli* Novablue transformadas con pET28a(+)-ClpC1, las clonas 3 y 6 presentan el plásmido recombinante. La figura 8 muestra el gel de electroforesis de las clonas mencionadas, en donde se observa el plásmido pET28a(+) linealizado (5369 bp) y el inserto *Rv3596c* (2547 bp) en los carriles 7 y 10, respectivamente para las clonas 3 y 6. En las cepas *E. coli* Novablue transformadas con pMAL-cHT-ClpC1, las clonas 5 y 6 presentan el plásmido recombinante. En la figura 9, los carriles 7 y 10, respectivamente para las clonas 5 y 6, muestran el plásmido pMAL-cHT linealizado (6679 bp) y el inserto *Rv3596c* en dos tamaños, debido a que el inserto tiene un sitio de reconocimiento para la enzima *EcoRI*. El primer tamaño es de aproximadamente 1700 bp y el segundo de 847 bp.

Finalmente, el gen *Rv3596c* se clonó de manera correcta, sin mutaciones y en la dirección adecuada en ambos plásmidos (pET28a(+)) y pMAL-cHT) (Anexo 2).

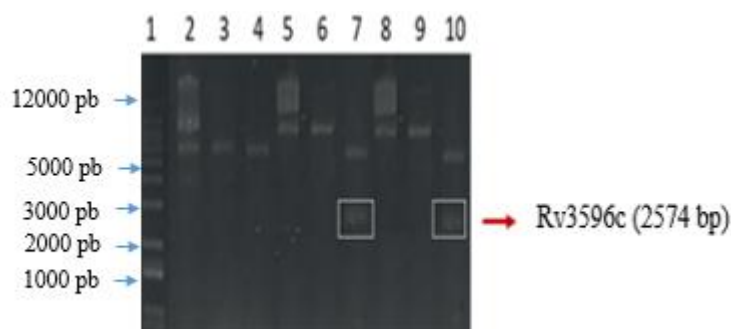


Figura 8. Digestión simple (SD) con *HindIII* y digestión doble (DD) con *HindIII* y *NcoI* del plásmido pET28a(+)-ClpC1. 1: Marcador 1Kb Plus, 2: pET28a(+) sin digerir, 3: pET28a(+) SD, 4: pET28a(+) DD, 5: pET28a(+), clona 3, 6: pET28a(+), clona 3 SD, 7: pET28a(+), clona 3 DD, 8: pET28a(+), clona 6, 9: pET28a(+), clona 6 SD, 10: pET28a(+), clona 6 DD. SD: Simple digestión; DD: Doble digestión. Bp: pares de bases

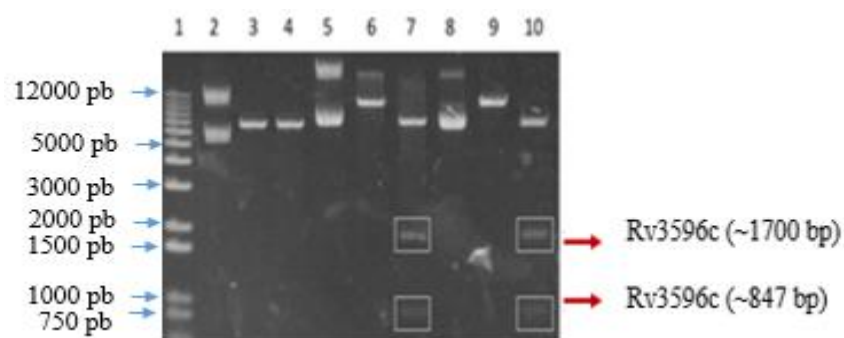


Figura 9. Digestión simple (SD) con *Hind*III y digestión doble (DD) con *Hind*III y *Eco*RI del plásmido pMAL-cHT-ClpC1. 1: Marcador 1Kb Plus, 2: pMAL-cHT sin digerir, 3: pMAL-cHT SD, 4: pMAL-cHT DD, 5: pMAL-cHT clona 5, 6: pMAL-cHT, clona 5 SD, 7: pMAL-cHT, clona 5 DD, 8: pMAL-cHT, clona 6, 9: pMAL-cHT, clona 6 SD, 10: pMAL-cHT, clona 6 DD. SD: Simple digestión; DD: Doble digestión.

5.4. Expresión de ClpC1 recombinante

La proteína se expresó eficientemente y de manera distinta en las diferentes condiciones probadas (Tabla 3). ClpC1 expresada en las tres cepas de *E. coli* transformadas con el plásmido pET28a(+)-ClpC1 en las diferentes temperaturas, tiempo y concentración de IPTG; se expresó en mayor cantidad de manera insoluble (Figura 10). Por otro lado, la expresión de ClpC1 en *E. coli* transformada con el plásmido pMAL-cHT-ClpC1 fue mayor en la fracción soluble. ClpC1 pesa 92 kDa y al expresarla utilizando el plásmido pMAL-cHT, se adiciona en el extremo C-terminal la proteína de unión a maltosa (MBP) la cual pesa alrededor de 42.5 kDa, por lo cual el peso total de ClpC1 recombinante es de 134.5 kDa (Figura 11).

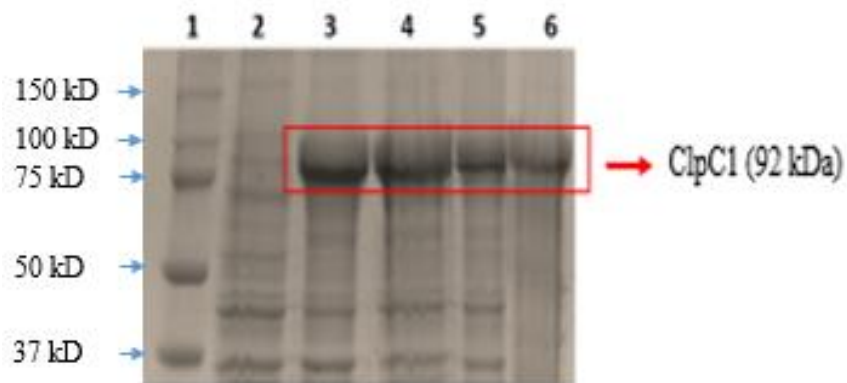


Figura 10. Expresión de ClpC1 recombinante en la cepa BL21(DE3)pLysS con el plásmido pET28a(+)-ClpC1. 1: Marcador de peso molecular de proteínas, 2: Fracción no inducida, 3: Fracción inducida; 4: Extracto crudo; 5: Fracción soluble, 6: Fracción insoluble.

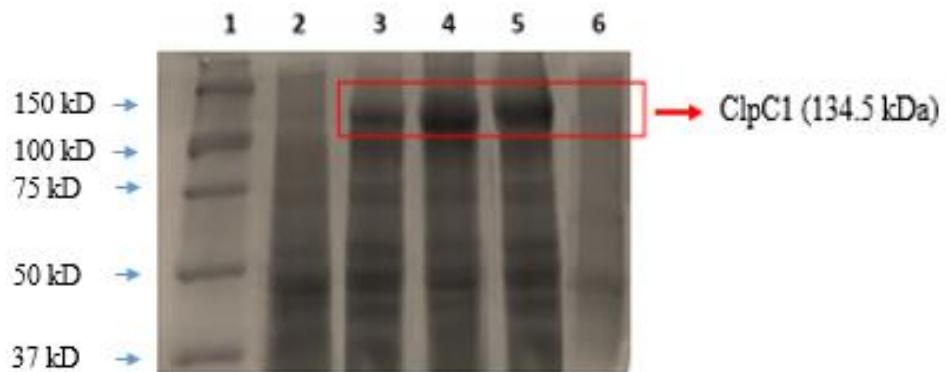


Figura 11. Expresión de ClpC1 recombinante en la cepa BL21(DE3)pLysS clonada con el plásmido pMAL- cHT-ClpC1. 1: Marcador de peso molecular de proteínas, 2: Fracción no inducida, 3: Fracción inducida; 4: Extracto crudo; 5: Fracción soluble, 6: Fracción insoluble.

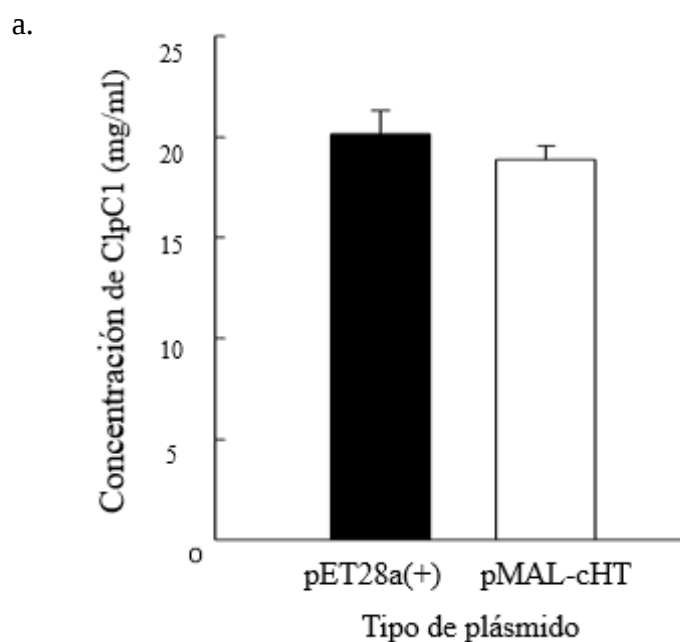
5.5. Comparación de la concentración y porcentaje de solubilidad de ClpC1 recombinante en los sistemas de expresión.

La concentración promedio total de ClpC1 fue mayor en las cepas transformadas con el plásmido pET28a(+)-ClpC1; sin embargo, se obtuvo mayor porcentaje de ClpC1 soluble en las cepas con el plásmido pMAL-cHT-ClpC1. Entre las cepas utilizadas, la cepa BL21(DE3)pLysS con el plásmido pMAL-cHT-ClpC1 evidenció una mayor concentración, así como un mayor porcentaje de ClpC1 soluble. Además, la temperatura y tiempo óptimo

en la cual se obtiene mayor concentración y solubilidad de ClpC1 fue de 16 grados/10 horas a una concentración final de 0.5 mM de IPTG.

En la figura 12a, se observa que la concentración de ClpC1 fue de 20,15 mg/ml (IC 95% 17,89 - 22,42) en las cepas con pET28a(+)-ClpC1 y de 18,86 mg/ml (IC 95% 17,47 - 20,24) en cepas con pMAL-cHT-ClpC1.

Por otro lado, el porcentaje de solubilidad obtenido en las cepas con pMAL-cHT-ClpC1 fue del 56 % (IC 95% 53,42 - 58,71) a comparación de las cepas con pET28a(+)-ClpC1 que fue del 27.61% (IC 95% 25,91 - 29,31) (figura 12b). Por ello, la concentración de ClpC1 fue mayor en cepas transformadas el plásmido pET28a(+)-ClpC1 mientras que el porcentaje de ClpC1 soluble fue mayor en cepas transformadas el plásmido pMAL-cHT-ClpC1.



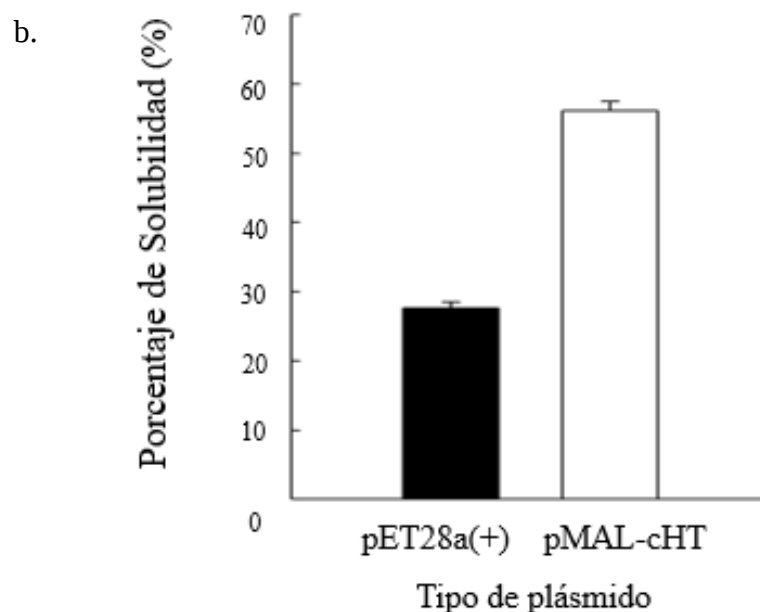


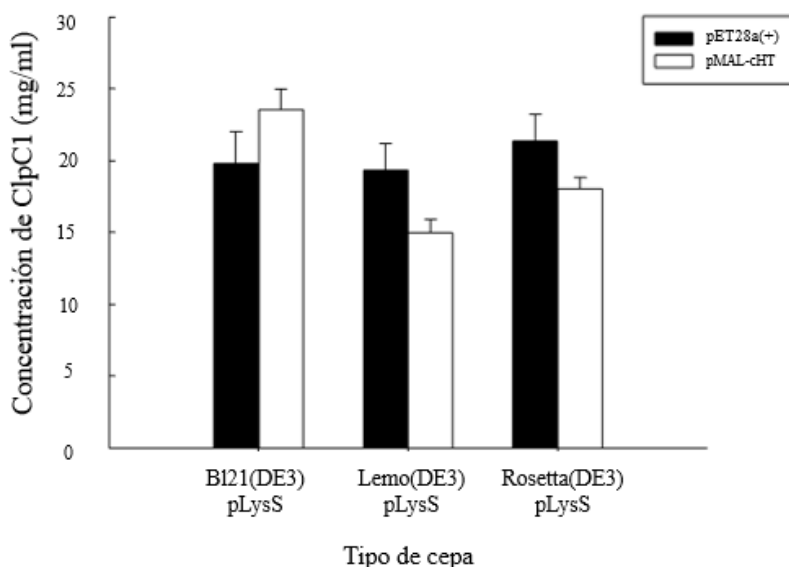
Figura 12. Estimación de la concentración total promedio y porcentaje de solubilidad de ClpC1 entre los plásmidos pET28a(+) y pMAL- cHT. (a) Promedio de la concentración total de ClpC1 obtenida con el plásmido pET28a(+) fue de 20,15 mg/ml (IC 95% 17,89 - 22,42) mientras que con pMAL-cHT fue de 18,86 mg/ml (IC 95% 17,47 - 20,24). (b) El porcentaje de solubilidad de ClpC1 recombinante obtenido con pET28a(+) fue del 27,61% (IC 95% 25,91 - 29,31) mientras que con pMAL-cHT fue de 56% (IC 95% 53,42 - 58,71).

Entre las cepas con pET28a(+)-ClpC1, la mayor concentración de ClpC1 fue en la cepa Rosetta(DE3)pLysS con 21,37 mg/ml (IC 95% 17,62 - 25,12), seguido por la cepa BL21(DE3)pLysS con 19,79 mg/ml (IC 95% 15,26 - 24,32) y, por último, Lemo(DE3)pLysS con 19,31 mg/ml (IC 95% 15,49 - 23,13). Por otro lado, la mayor concentración de ClpC1 en la cepa BL21(DE3)pLysS con pMAL-cHT-ClpC1 fue de 23,55 mg/ml (IC 95% 20,69 - 26,40), seguido por la cepa Rosetta(DE3)pLysS 18,03 mg/ml (IC 95% 16,46 - 19,62) y, finalmente, Lemo(DE3)pLysS 14,99 mg/ml (IC 95% 13,16 - 16,83) (Figura 13a).

En relación al porcentaje de solubilidad, se alcanzó el mayor porcentaje en las cepas con pMAL-cHT-ClpC1 a comparación de las cepas con pET28a(+)-ClpC1. En las cepas BL21(DE3) pLysS, Lemo(DE3) pLysS y Rosetta(DE3) pLysS con pMAL-cHT-ClpC1 se obtuvo 71,84 % (IC 95% 67,95 - 75,74), 51 % (IC 95% 49,67 - 52,36) y 45,3% (IC 95%

43,25 - 47,43), respectivamente. Asimismo, en las cepas BL21(DE3) pLysS, Lemo(DE3)pLysS y Rosetta (DE3)pLysS con pET28a(+)-ClpC1 se obtuvo 27.9 % (IC 95% 25,54 - 30,32), 28.4% (IC 95% 25,22 - 31,59) y 26.5% (IC 95% 23,08 - 29,91), respectivamente (Figura 13b).

a.



b.

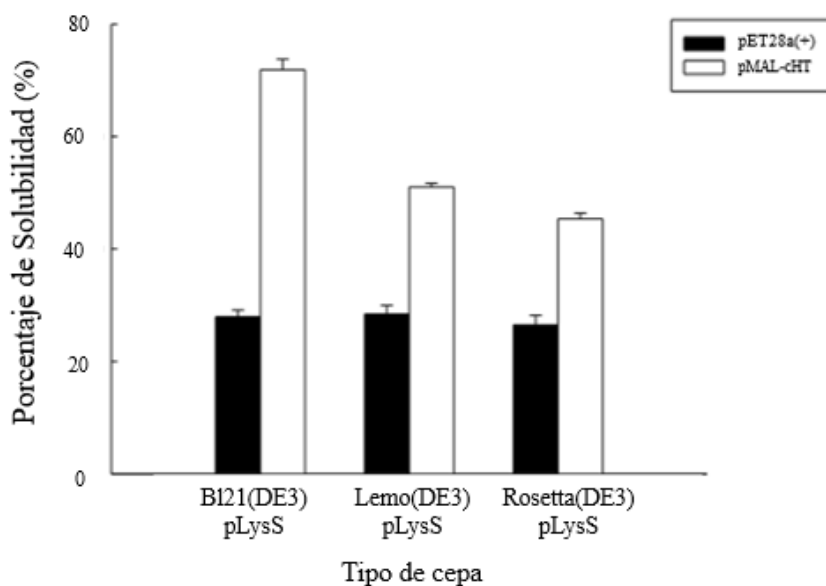


Figura 13. Estimación de la concentración total promedio y porcentaje de solubilidad de ClpC1 recombinante en las cepas BL21(DE3)pLysS, Lemo(DE3)pLysS y Rosetta(DE3)pLysS con los plásmidos pET28a (+) y pMAL- cHT. (a) Promedio de la

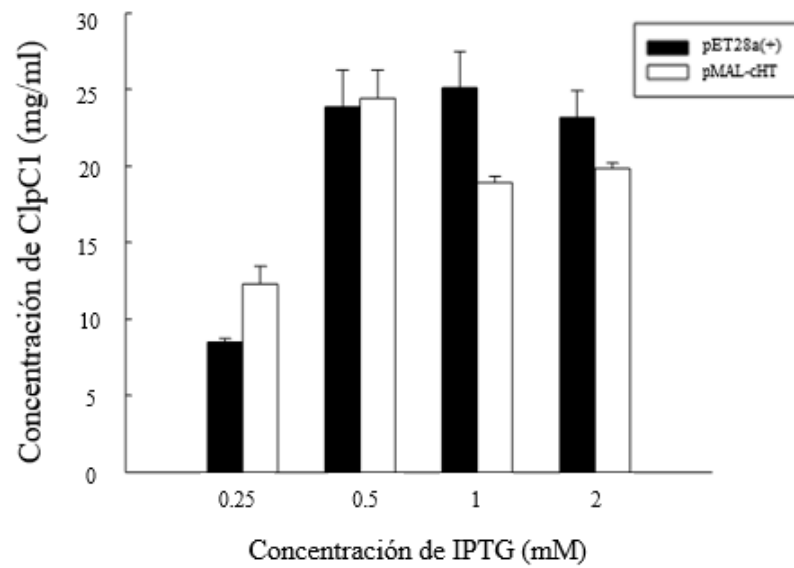
concentración de ClpC1 expresada en las cepas con pET28a(+) y pMAL-cHT. Se obtuvo mayor concentración de ClpC1 en la cepa BL21(DE3)pLysS 23,55 mg/ml (IC 95% 20,69 - 26,40) con pMAL-cHT.(b) Porcentaje de solubilidad de ClpC1 expresada en las cepas con pET28a (+) y pMAL-cHT. El mayor porcentaje que se obtuvo fue en las cepas BL21(DE3)pLysS con pMAL-cHT con 71,84 % (IC 95% 67,95 - 75,74).

Referente a la estimación de la concentración y porcentaje de solubilidad de ClpC1 expresada en las cepas transformadas con pET28a(+)-ClpC1 y pMAL-cHT-ClpC1a diferentes concentraciones de IPTG, se evidenció que la mayor concentración de ClpC1 fue de 25,09 mg/ml (IC 95% 20,22 -29,97) a una concentración final de 1mM de IPTG en cepas con pET28a(+)-ClpC1. Mientras que en cepas con pMAL-cHT-ClpC1se obtuvo una concentración de 24,39 mg/ml (IC 95% 20,54 - 28,25) a una concentración final de IPTG de 0.5 mM.

Se obtuvieron en cepas transformadas con pET28a(+)-ClpC1 8,49 mg/ml (IC 95% 7,95 - 9.02), 23,86 mg/ml (IC 95% 18,92 -28,80) y 23,18 mg/ml (IC 95% 19,61 - 26,75), a las concentraciones finales de 0.25, 0.5 y 2 mM de IPTG respectivamente. De la misma manera, en cepas transformadas con pMAL-cHT-ClpC1se produjeron 12,31 mg/ml (IC 95% 9.95 - 14,68), 18,91 mg/ml (IC 95% 18,04 - 19,77) y 19,82 mg/ml (IC 95% 19,02 - 20,62) de concentración de ClpC1 a las concentraciones de 0.25, 1 y 2 mM de IPTG (Figura 14a).

En base al porcentaje de solubilidad, en las cepas con pET28a(+)-ClpC1 se muestra 34, 25% (IC 95% 31,28 - 37,23), 30,12 % (IC 95% 27,45 - 32,79), 25,47 % (IC 95% 22,26 - 28,69) y 20,60% (IC 95% 17,8 - 23,40) a las concentraciones finales de 0.25, 0.5, 1 y 2 mM de IPTG respectivamente. Mientras que las cepas con pMAL-cHT-ClpC1 muestran 60, 82 % (IC 95% 54,39 - 67,26), 55,32 % (IC 95% 49,50 - 61,14), 55,04 % (IC 95% 49,45 - 60,63) y 53,07 % (IC 95% 49,65 - 56,49) a las concentraciones finales de 0.25, 0.5, 1 y 2 mM de IPTG respectivamente (Figura 14b).

a.



b.

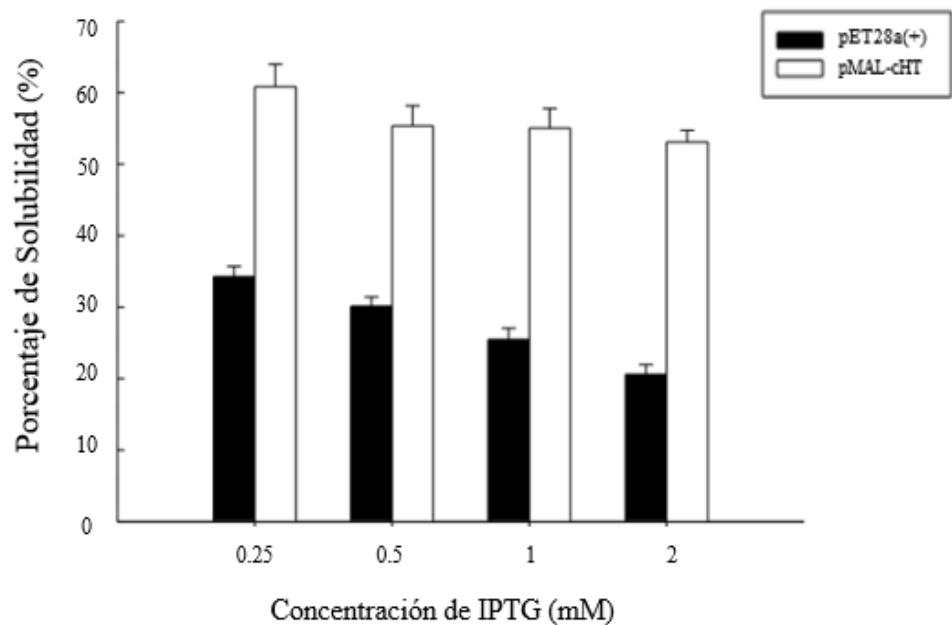


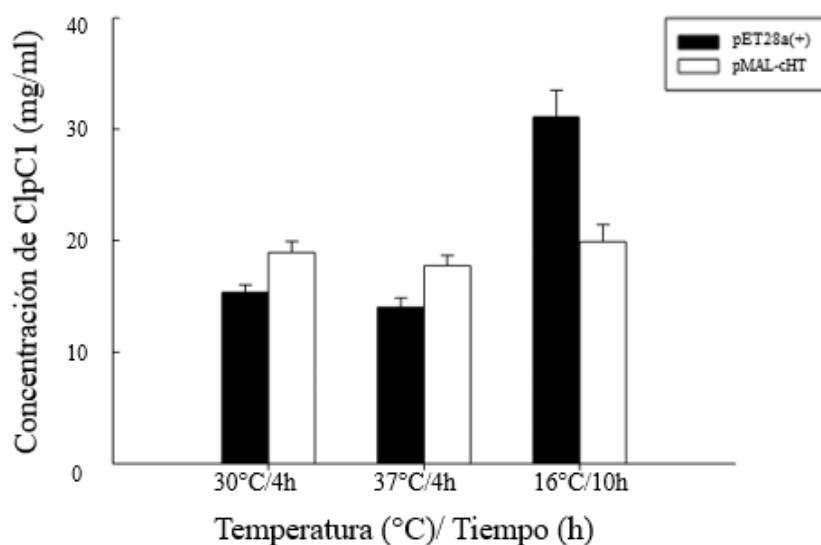
Figura 14. Estimación de la concentración total promedio y porcentaje de solubilidad de ClpC1 referente a las diferentes concentraciones de IPTG en cepas clonadas con pET28a(+) y pMAL-cHT. (a) Promedio de la concentración de ClpC1 expresada en las cepas con pET28a(+) y pMAL-cHT. Se obtuvo la mayor concentración de ClpC1 en las cepas con pET28a(+) 25,09 mg/ml (IC 95% 20,22 -29,97) a una concentración final de 1 mM de IPTG. (b) Porcentaje de solubilidad de ClpC1 expresada en las cepas con pET28a

(+) y pMAL-cHT. Las cepas con pMAL-cHT mostraron un 60,82 % (IC 95% 54,39 - 67,26) a una concentración final de 0.25 mM de IPTG.

En cuanto a la concentración y porcentaje estimados con respecto a la temperatura y al tiempo de expresión, se evidenció que la mayor concentración de ClpC1 fue 31,08 mg/ml (IC 95% 26,26 - 35,92) a 16°C por 10 horas en cepas con pET28a(+)-ClpC1, mientras que en cepas transformadas con pMAL-cHT-ClpC1 se obtuvo como mayor concentración de ClpC1 19,93 mg/ml (IC 95% 16,79 - 23,05) en las mismas condiciones descritas. Las concentraciones producidas en las cepas con pET28a(+)-ClpC1 fueron 15,34 mg/ml (IC 95% 13,86 - 16,83) y 14,03 mg/ml (IC 95% 12,38 - 15,70) en las condiciones de 30°C/ 4 horas y 37°C/ 4 horas respectivamente. Igualmente, las concentraciones de ClpC1 en cepas con pMAL-cHT-ClpC1, fueron de 18,94 mg/ml (IC 95% 16,83 - 21,05) y 17,72 mg/ml (IC 95% 15,74 - 19,69) en las condiciones de 30°C/4 horas y 37°C/4 horas respectivamente (Figura 15a).

Acerca del porcentaje de solubilidad de ClpC1, se obtuvo como mayor porcentaje 60,08% (IC 95% 54,17 - 65,98) en cepas transformadas con pMAL-cHT-ClpC1a condiciones de temperatura y tiempo de 16°C/10 horas de expresión. En las cepas con pET28a(+)-ClpC1, se obtuvieron 31,39 % (IC 95% 28,28 - 34,49), 28, 12% (IC 95% 25,15 - 31,10) y 23,32% (IC 95% 20,99 - 25,65) a condiciones de 37°C /4 horas, 30°C/4 horas y 16°C/10 horas respectivamente. De igual modo, en cepas con pMAL-cHT-ClpC1, los porcentajes fueron de 53,17 % (IC 95% 49,0 - 57,33) y 54,95 % (IC 95% 51,50 - 58,39) a condiciones de 30°C/4 horas y 37°C/4 horas respectivamente (Figura 15 b).

a.



b.

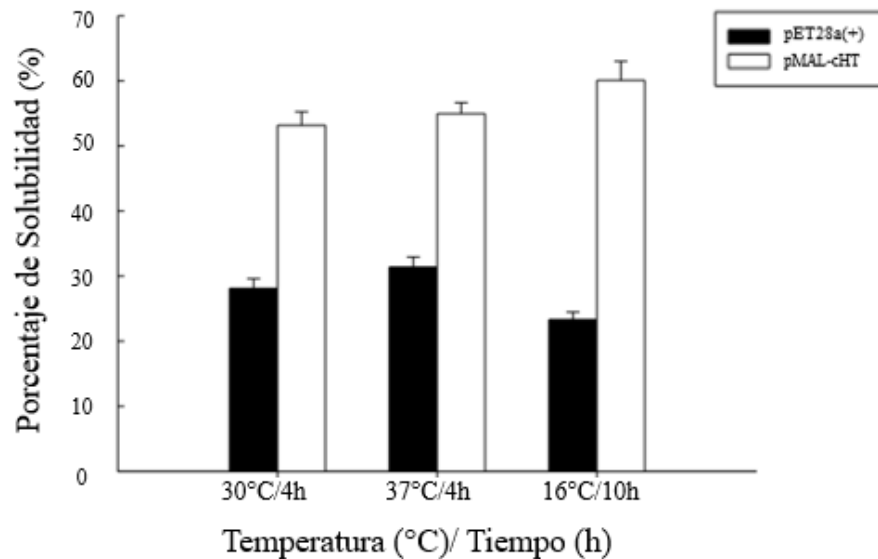


Figura 15. Estimación de la concentración total promedio y porcentaje de solubilidad de ClpC1 concernientes a las diferentes temperaturas y tiempos de expresión en cepas clonadas con pET28a (+) y pMAL-cHT. (a) Promedio de la concentración de ClpC1 expresada en las cepas con pET28a (+) y pMAL-cHT. Se obtuvo la mayor concentración de ClpC1 en las cepas con pET28a(+) 31,08 mg/ml (IC 95% 26,26 - 35,92) a condiciones de 16°C/10 horas. (b) Porcentaje de solubilidad de ClpC1 expresada en las cepas con pET28a (+) y pMAL-cHT. El mayor porcentaje se obtuvo en cepas con pMAL-cHT 60,08% a condiciones de 16°C/10 horas.

Al analizar el efecto de las interacciones de cada condición descrita, ClpC1 en pET28a(+) se expresó como mayor concentración 45,65 mg/ml (IC 95% 44,67 - 46,64) en la cepa BL21(DE3)pLysS con 0.5 mM de IPTG a 16 °C por 10 horas y con mayor porcentaje de solubilidad 44.09% (IC 95% 43,84 - 44,35) en la cepa BL21(DE3)pLysS con 0.25 mM de IPTG a 37°C por 4 horas. Del mismo modo, ClpC1 en pMAL-cHT expresó como mayor concentración 45,55 mg/ml (IC 95% 44,69 - 46,41) en la cepa BL21(DE3)pLysS con 0.5 mM de IPTG a 16°C por 10 horas y con mayor porcentaje de solubilidad 93.22% (IC 95% 92.33 - 94.11) en la cepa BL21(DE3)pLysS con 0.25 mM de IPTG a 16°C por 10 horas (Anexo 1).

Por último, las mejores condiciones para expresar ClpC1 de manera recombinante en *E. coli* es en la cepa BL21(DE3)pLysS transformadas con pMAL-cHT-ClpC1, a una concentración final de 0.5 mM de IPTG por 10 horas a 16°C. Con estas características, se obtuvo como concentración total promedio 45,55 mg/ml (IC 95% 44,69 - 46,41) y 86.93% (IC 95% 80,41 - 93,45) de porcentaje de solubilidad (Anexo 1).

5.6. Purificación de ClpC1-MBP y remoción de MBP de ClpC1.

ClpC1-MBP fue purificada, usando las condiciones seleccionadas (BL21(DE3)pLysS, pMAL-cHT-ClpC1, 0.5 mM de IPTG, 16°C/10h), por cromatografía de afinidad utilizando la columna MBPTrap-HP; sin embargo, se evidenció la presencia de otras proteínas en la fracción eluída (Figura 16). Asimismo, MBP fue removida de ClpC1 por la proteasa TEV en los diferentes temperaturas y tiempos probados (4°C/72 h y 30°C/24 h) (Figura 16). ClpC1 fue obtenida en forma soluble al remover MBP, sin embargo, al almacenarla por aproximadamente 2 meses, se observó una precipitación.

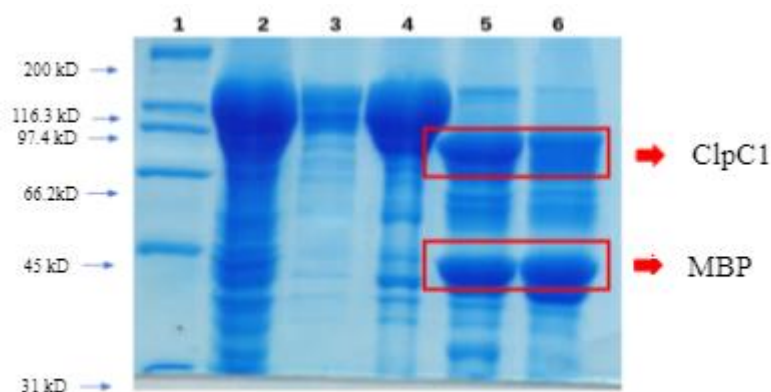


Figura 16. Purificación y corte de ClpC1-MBP. 1: Marcador de peso molecular de proteínas, 2: Fracción soluble, 3: Fracción Eluída, 4: Muestra concentrada; 5: Corte TEV 4°C/72 horas, 6: Corte TEV 30°C/24 horas.

5.7. Modelamiento estructural de ClpC1 con y sin MBP

Al comparar las estructuras de ClpC1 con/ sin MBP, la estructura modelada de ClpC1 sin MBP mostró como porcentaje de identidad a nivel de la secuencia de aminoácidos entre el modelo de ClpC1 y la plantilla un 63.35%. Asimismo, el modelo generado por homología reportó un GMQE de 0.44 y un score de calidad (QMEAN) de -5.32. El posterior alineamiento estructural reportó un porcentaje de identidad de 33.4% y similitud de 42.1%. De la misma manera, se obtuvo un RMSD de 1.36 Å y un T-Score de 0.90589 (Figura 17). Por otro lado, el modelo de ClpC1 con MBP presentó un RMSD de 11 Å y un T-Score de 0.57 (Figura 18).

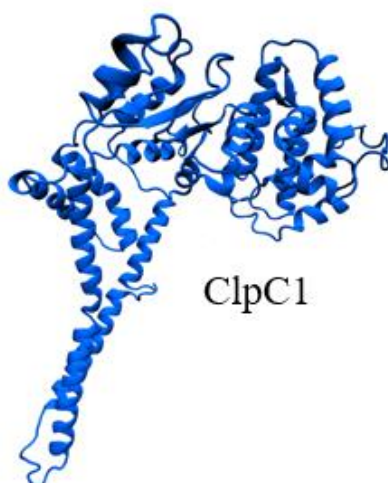


Figura 17. Estructura tridimensional de ClpC1 sin MBP.

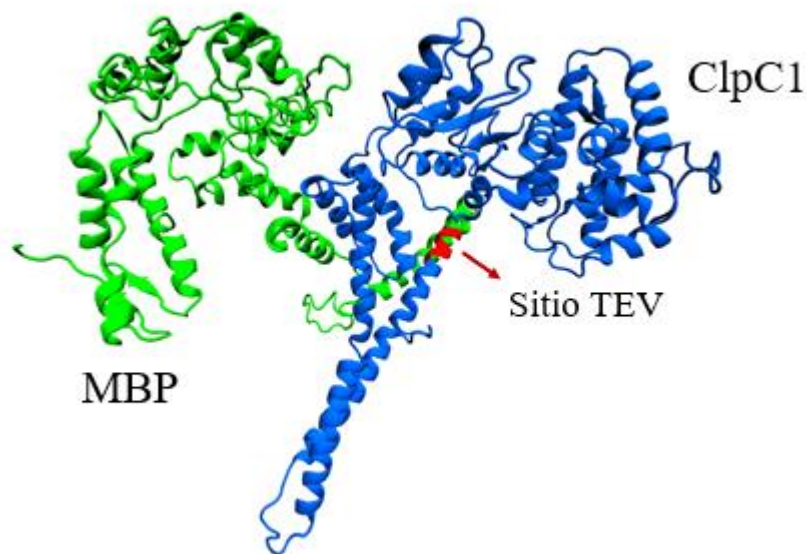


Figura 18. Estructura tridimensional de ClpC1 con MBP. ClpC1 se muestra de color azul, sitio TEV de color rojo y MBP de color rojo.

5.8. Estimación de los parámetros de solubilidad de ClpC1 con y sin MBP

ClpC1 presenta 848 residuos de los cuales 175 son hidrofóbicos, lo que corresponde un 20.6% de residuos hidrofóbicos y 79.4% de residuos hidrofílicos (figura 19); mientras que ClpC1 con MBP - Rv3596c clonado en pMAL-cHT- presenta 1241 residuos, de los cuales 344 residuos son hidrofóbicos, lo que representa el 27.8% de residuos hidrofóbicos y 72.2 % de residuos hidrofílicos (figura 20). Por otro lado, el área superficial correspondiente a ClpC1 presentó un valor de 5150.7 Å² del cual el SASA corresponde el 23.845.1 Å² (46.29%); mientras que el área superficial correspondiente a ClpC1 con MBP presentó un valor de 94143 Å² del cual el SASA corresponde el 42525 Å² (45.17%). Finalmente, los valores de QuerySol para cada modelo fue mayor a 0.45, ClpC1 presentó un valor de 0.56 y ClpC1 con MBP 0.59 (Tabla 5).

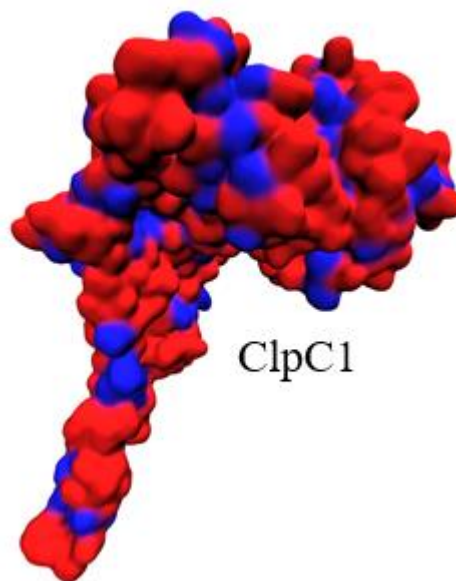


Figura 19. Porcentaje de residuos hidrofóbicos e hidrofílicos de ClpC1 sin MBP. Los residuos hidrofóbicos se encuentran en color azul (20.6%) que los hidrofílicos de color rojo (79.4%).

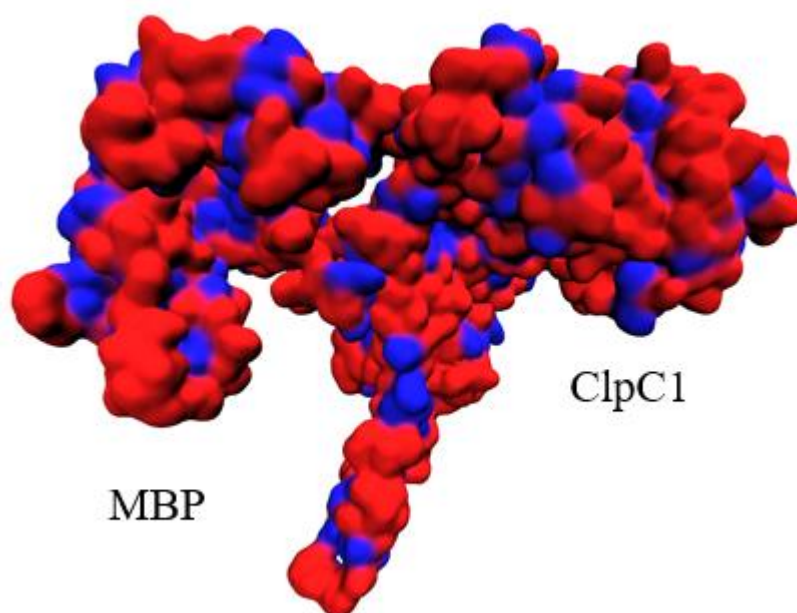


Figura 20. Porcentaje de residuos hidrofóbicos e hidrofílicos de ClpC1 con MBP. Los residuos hidrofóbicos se encuentran en color azul (27.8%) que los hidrofílicos de color rojo (72.2%).

Tabla 5. Resumen de los parámetros de solubilidad de ClpC1 sin y con MBP.

Proteínas	Número de Residuos	Naturaleza de los residuos		SASA (%)	Query Sol
		Hidrofílico (%)	Hidrofóbico (%)		
ClpC1	848	79.4	20.6	46.29	0.56
ClpC1+MBP	1241	72.2	27.8	45.17	0.59

6. DISCUSIÓN

En el presente trabajo se clonó y expresó de manera exitosa la proteína ClpC1 de *M. tuberculosis* en las diferentes condiciones y vectores de expresión en *E. coli*. De la misma manera, se logró comparar la concentración y solubilidad de la proteína recombinante ClpC1 de *M. tuberculosis* en todos los sistemas de expresión. Además, se modeló de manera correcta ClpC1 con y sin MBP y se hallaron los parámetros de solubilidad.

Los resultados obtenidos muestran que se logró clonar de manera eficiente el gen *Rv3596c* en los plásmidos pET28a(+) y pMAL-cHT siguiendo el protocolo *In-fusion* siguiendo las recomendaciones del fabricante (Laboratories Inc, 2011). La secuenciación evidenció la correcta clonación, dirección y ninguna mutación en el gen (Anexo 2). Solo el gen *Rv3596c* fusionado con el gen de *maltE* (Maltose Binding Protein) (clonado en pMAL-cHT) se expresó en mayor porcentaje de forma soluble en las diferentes condiciones evaluadas (sistema de expresión, concentración de IPTG y temperatura), en otras palabras, ClpC1 + MBP se encontró en mayor cantidad en la fracción soluble (Figura 11). Se demostró que la mejor condición para lograr una mayor concentración estimada así como porcentaje de solubilidad fue al expresar ClpC1 en el plásmido pMAL-cHT en la cepa BL21(DE3)pLysS a una concentración final de 0.5 mM de IPTG por 10 horas a una temperatura de 16°C (Anexo 1).

Una posible explicación por la cual se obtuvo mayor concentración relativa de ClpC1 producida en pET28a(+) se debe principalmente al promotor T7 el cual es más eficaz (unión a la RNA pol) que el promotor tac presente en pMAL-cHT y en consecuencia se produce mayor cantidad de mRNA (de Boer *et al*, 1983; Rosano &

Ceccarelli, 2014; Samuelson, 2011). El plásmido pET28a(+) contiene el origen de replicación del plásmido pBR322, el cual se considera que es de bajo número de copias debido a que solo se encuentran 15-20 copias por bacteria, al igual que el origen ColE1 de pMAL-cHT, por lo cual no hay diferencias entre el número de copias de plásmidos entre ellos (Lee *et al*, 2006; Schmidt & Inselburg, 1982).

Por el contrario, se demostró que el mayor porcentaje de solubilidad relativo de ClpC1 se dio con pMAL-cHT. Como ya se mencionó, el plásmido pMAL- cHT añade a la proteína de interés la proteína MBP, lo cual aumenta la solubilidad de la proteína. El mecanismo por el cual MBP aumenta la solubilidad de una proteína aún no se conoce en su totalidad, pero se postula que MBP mejora la estabilidad de las proteínas fusión debido a que actúa como una “chaperona” al promover el plegamiento de la proteína fusión (Kapust & Waugh, 1999). Adicionalmente, se demostró la proteína fusión (proteína de interés + MBP) es más afín a las chaperonas GroELS (Bernier *et al*, 2018; Kapust & Waugh, 1999).

Para tener acceso suficiente a estas chaperonas es muy importante que *E.coli* se encuentre en un estado metabólico óptimo, este se alcanza con la inducción de la expresión cuando se tiene un OD₆₀₀ entre 0.4-0.6. De la misma manera, se sabe que MBP es muy estable y “enmascara” la insolubilidad de la proteína; por ello, la solubilidad neta de la proteína fusión se mantiene. Asimismo, no se sabe si los agregados proteicos de las proteínas fusión se forman antes o después del plegamiento de MBP, este punto es crítico para entender si MBP presenta otro mecanismo adicional en la solubilidad de las proteínas fusión.

Por otro lado, la cepa BL21(DE3)pLysS es utilizada para aumentar la solubilidad de proteínas tóxicas o insolubles (Angius *et al*, 2018), en nuestros resultados esta cepa produjo mayor porcentaje de solubilidad de ClpC1 aunque esta no sea considerada tóxica. La concentración de IPTG óptima para la expresión soluble de ClpC1 fue 0.5 mM, esto se explica debido a la capacidad de las chaperonas de la bacteria. Las concentraciones altas de IPTG (> 1mM) pueden limitar la disponibilidad de las chaperonas y en consecuencia la formación de agregados proteicos, en otras palabras, proteínas insolubles. Por último, temperaturas bajas de expresión como 16 ° C mejoran la solubilidad de las proteínas al alterar el recambio metabólico de la bacteria y de esta manera garantizar la disponibilidad de chaperonas para el correcto plegamiento de la proteína de interés (Vera *et al*, 2007).

Además, se sabe que la energía de activación durante el plegamiento de las proteínas depende de la temperatura, a temperaturas bajas se ve favorecida (Naganathan *et al*, 2007).

Por otra parte, el modelo de ClpC1 presentó valores de RMSD de 1.63 Å y T-Score de 0.9, los cuales se encuentran dentro del rango establecido por la literatura; al contrario del modelo ClpC1-MBP (RMSD de 11^ª y T-Score de 0.57) (Zhang, 2005; Sargsyan, 2017). Con respecto a los parámetros de solubilidad, los resultados se asemejan ligeramente a los obtenidos experimentalmente. La solubilidad de una proteína se ve influenciada por la naturaleza de sus aminoácidos, proteínas con un mayor porcentaje de residuos no polares, hidrofóbicos en su superficie, generalmente tienden a precipitar cuando se producen de manera recombinante (Idicula-Thomas & Balaji, 2005). Asimismo, proteínas con alto porcentaje de residuos hidrofílicos en su superficie tienden a ser más estables y solubles (Strub *et al*, 2004). Cuando una proteína se desestabiliza, tiende a formar cuerpos de inclusión los cuales se forman por interacciones hidrofóbicas inespecíficas entre los polipéptidos en formación, estas interacciones se ven favorecidas por el alto porcentaje de residuos hidrofóbicos que puede poseer la estructura primaria de una proteína (Ventura & Villaverde, 2006).

El porcentaje de residuos hidrofílicos fue mayor en ClpC1 sin MBP (Tabla 4), por ende ClpC1-MBP presentar mayor cantidad de residuos hidrofóbicos, sin embargo ClpC1-MBP *in vitro* es más soluble que ClpC1 sin MBP. Aunque ClpC1-MBP presente más residuos hidrofóbicos, este resultado no nos indica si la mayoría de estos se encuentran en mayor proporción en la superficie de la proteína.

Asimismo, el valor de SASA fue mayor en ClpC1 sin MBP, este resultado es controversial, debido a que nos muestra que hay mayor superficie accesible al solvente en ClpC1 sin embargo en esta superficie pueden estar tanto los residuos hidrofóbicos como hidrofílicos. En relación al valor de QuerySol, este fue ligeramente mayor en ClpC1 -MBP (Tabla 4), se puede concluir que al adicionar MBP aumenta la solubilidad global de la proteína. Esto concuerda con los resultados experimentales obtenidos.

Con respecto a la purificación de ClpC1-MBP, se evidencia claramente la presencia de proteínas contaminantes, debido a ello es necesario una segunda purificación para obtener ClpC1 más pura. De la misma manera, al remover MBP de ClpC1, se verificó la solubilidad de ClpC1. Sin embargo, esta precipitó al almacenarse por un tiempo de dos meses. Por ello es necesario encontrar las mejores condiciones de almacenamiento de ClpC1. Aunque ClpC1 permanezca soluble después del corte, esta podría ser

biológicamente inactiva. Debido a ello, para corroborar su correcto plegamiento y función son necesarios futuros experimentos de su actividad.

Existen otros factores que afectan la solubilidad de ClpC1, como el pH, la fuerza iónica y el tipo de solvente que no han sido estudiados en este trabajo. Se necesitan futuros estudios para determinar el rol de MBP en la solubilidad de ClpC1.

7. CONCLUSIONES

- El gen Rv3596c fue clonado correctamente en los plásmidos pET28a (+) y pMAL-cHT en *E. coli*.
- ClpC1 en pET28a(+) fue expresada en mayor concentración en la cepa BL21(DE3)pLysS con 0.5 mM de IPTG a 16°C por 10 horas y con mayor solubilidad en la cepa BL21(DE3)pLysS con 0.25 mM de IPTG a 37°C por 4 horas.
- ClpC1 en pMAL-cHT fue expresada en mayor concentración en la cepa BL21(DE3)pLysS con 0.5 mM de IPTG a 16°C por 10 horas y con mayor solubilidad en la cepa BL21(DE3)pLysS con 0.25 mM de IPTG a 16°C por 10 horas.
- Las mejores condiciones para obtener mayor concentración de ClpC1 soluble fue empleando la cepa BL21(DE3)pLysS transformada con el plásmido pMAL-cHT-ClpC1 a una concentración final de 0.5 mM de IPTG, 16°C y 10 horas de expresión.
- Se verificó la presencia de ClpC1 soluble después de remover MBP de ClpC1-MBP.
- Los valores de los parámetros de calidad del modelo de ClpC1 de *M. tuberculosis* se encontraron entre los rangos establecidos por la literatura, mientras que en el de ClpC1 -MBP no.
- El porcentaje de residuos hidrofílicos fue mayor en la estructura de ClpC1 (79.4%) a comparación de ClpC1 -MBP (72.2%).
- La solubilidad global de ClpC1 -MBP fue mayor a la de ClpC1.

8. RECOMENDACIONES

En este trabajo, se logró encontrar las mejores condiciones para aumentar la expresión de ClpC1 de manera soluble, esto ayudará a obtener mayor cantidad y calidad de proteína después de su purificación. Es necesario pasos adicionales antes de evaluar funcionalmente a ClpC1, la proteína debe purificada por afinidad, usando una resina de amilosa, y luego, ser cortada para separar MBP de la estructura nativa de ClpC1, esto es fácil y sencillo de realizar gracias a la proteasa TEV, la cual actúa a diferentes temperaturas adecuándose de esta manera a la susceptibilidad de la proteína de interés como se corroboró en la prueba piloto realizada. Una vez removida MBP de ClpC1, será necesario realizar una purificación por 6X-His Tag y de esta manera obtener ClpC1 de una manera más pura y realizar futuros experimentos de su actividad. Además, la obtención de ClpC1 nativa contribuirá a futuros estudios que deseen investigar su rol dentro del mecanismo de pirazinamida u otros fármacos antituberculosos.

Adicionalmente, es necesario medir el crecimiento bacteriano mediante densidad óptica (OD_{600}) para normalizar las fracciones de una manera más precisa. Por último, la limitación más importante de este trabajo fue que no se logró purificar la proteína ClpC1 de las fracciones solubles e insolubles en todas las condiciones descritas, esto es importante determinar la concentración final de ClpC1 de una forma más precisa. De la misma manera, los valores de los parámetros de calidad del modelo bioinformático de ClpC1-MBP (RMSD Y T-Score) se encontraron fuera del rango establecido por la literatura. El valor de RMSD podría disminuir al comparar el modelo de ClpC1-MBP con una estructura cristalizada de ClpC1 en unión con MBP.

9. REFERENCIAS BIBLIOGRÁFICAS

- Ali, M. S., & Baek, K. H. (2020). Co-suppression of NbClpC1 and NbClpC2, encoding clp protease chaperons, elicits significant changes in the metabolic profile of *Nicotiana benthamiana*. *Plants*, 9(2), 259. <https://doi.org/10.3390/plants9020259>
- Ali, S., Hassan, M., Islam, A., & Ahmad, F. (2014). A Review of Methods Available to Estimate Solvent-Accessible Surface Areas of Soluble Proteins in the Folded and Unfolded States. *Current Protein & Peptide Science*, 15(5), 456–476. <https://doi.org/10.2174/1389203715666140327114232>
- Angius, F., Iliaia, O., Amrani, A., Suisse, A., Rosset, L., Legrand, A., Abou-Hamdan, A., Uzan, M., Zito, F., & Miroux, B. (2018). A novel regulation mechanism of the T7 RNA polymerase based expression system improves overproduction and folding of membrane proteins. *Scientific Reports*, 8(1), 1–11. <https://doi.org/10.1038/s41598-018-26668-y>
- Benkert, P., Biasini, M., & Schwede, T. (2011). Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics*, 27(3), 343–350. <https://doi.org/10.1093/bioinformatics/btq662>
- Bernier, S. C., Cantin, L., & Salesse, C. (2018). Systematic analysis of the expression, solubility and purification of a passenger protein in fusion with different tags. *Protein Expression and Purification*, 152, 92–106. <https://doi.org/10.1016/j.pep.2018.07.007>
- Choules, M. P., Wolf, N. M., Lee, H., Anderson, J. R., Grzelak, E. M., Wang, Y., Ma, R., Gao, W., McAlpine, J. B., Jin, Y.-Y., Cheng, J., Lee, H., Suh, J.-W., Duc, N. M., Paik, S., Choe, J. H., Jo, E.-K., Chang, C. L., Lee, J. S., ... Cho, S. (2019). Rifomycin Targets ClpC1 Proteolysis in *Mycobacterium tuberculosis* and *M. abscessus*. *Antimicrobial Agents and Chemotherapy*, 63(3), e02204-18. <https://doi.org/10.1128/AAC.02204-18>
- Clarke, A. K., & Eriksson, M. J. (1996). The cyanobacterium *Synechococcus* sp. PCC 7942 possesses a close homologue to the chloroplast ClpC protein of higher plants. *Plant Molecular Biology*, 31(4), 721–730. <https://doi.org/10.1007/BF00019460>
- de Boer, H. A., Comstock, L. J., & Vasser, M. (1983). The tac promoter: a functional hybrid derived from the trp and lac promoters. *Proceedings of the National Academy of Sciences of the United States of America*, 80(1), 21–25.

<https://doi.org/10.1073/pnas.80.1.21>

- Deller, M. C., Kong, L., & Rupp, B. (2016). Protein stability: a crystallographer's perspective. *Acta crystallographica. Section F, Structural biology communications*, 72(Pt 2), 72–95. <https://doi.org/10.1107/S2053230X15024619>
- Donovan, R., Robinson, C., & Glick, B. (1996). Optimizing inducer and culture conditions for expression of foreign proteins under the control of the lac promoter. *Journal of Industrial Microbiology*, 16(3), [145–154. https://doi.org/10.1007/BF01569997](https://doi.org/10.1007/BF01569997)
- Drew, D., Fröderberg, L., Baars, L., & De Gier, J. W. L. (2003). Assembly and overexpression of membrane proteins in Escherichia coli. In *Biochimica et Biophysica Acta - Biomembranes* (Vol. 1610, Issue 1, pp. 3–10). Elsevier. [https://doi.org/10.1016/S0005-2736\(02\)00707-1](https://doi.org/10.1016/S0005-2736(02)00707-1)
- Elsholz, A. K. W., Birk, M. S., Charpentier, E., & Turgay, K. (2017). Functional diversity of AAA+ protease complexes in Bacillus subtilis. In *Frontiers in Molecular Biosciences* (Vol. 4, Issue JUL, p. 44). Frontiers Media S.A. <https://doi.org/10.3389/fmolb.2017.00044>
- Feldman, D. E., & Frydman, J. (2000). Protein folding in vivo: The importance of molecular chaperones. *Current Opinion in Structural Biology*, 10(1), 26–33. [https://doi.org/10.1016/S0959-440X\(99\)00044-5](https://doi.org/10.1016/S0959-440X(99)00044-5)
- Gao, W., Kim, J. Y., Anderson, J. R., Akopian, T., Hong, S., Jin, Y. Y., Kandror, O., Kim, J. W., Lee, I. A., Lee, S. Y., McAlpine, J. B., Mulugeta, S., Sunoqrot, S., Wang, Y., Yang, S. H., Yoon, T. M., Goldberg, A. L., Pauli, G. F., Suh, J. W., ... Cho, S. (2015). The cyclic peptide ecumicin targeting CLpC1 is active against Mycobacterium tuberculosis in vivo. *Antimicrobial Agents and Chemotherapy*, 59(2), 880–889. <https://doi.org/10.1128/AAC.04054-14>
- Gavriš, E., Sit, C. S., Cao, S., Kandror, O., Spoering, A., Peoples, A., Ling, L., Fetterman, A., Hughes, D., Bissell, A., Torrey, H., Akopian, T., Mueller, A., Epstein, S., Goldberg, A., Clardy, J., & Lewis, K. (2014). Lassomycin, a ribosomally synthesized cyclic peptide, kills mycobacterium tuberculosis by targeting the ATP-dependent protease ClpC1P1P2. *Chemistry & Biology*, 21(4), 509–518. <https://doi.org/10.1016/j.chembiol.2014.01.014>
- Gopal, P., Grüber, G., Dartois, V., & Dick, T. (2019). Pharmacological and Molecular Mechanisms Behind the Sterilizing Activity of Pyrazinamide. In *Trends in*

- Pharmacological Sciences* (Vol. 40, Issue 12, pp. 930–940). Elsevier Ltd.
<https://doi.org/10.1016/j.tips.2019.10.005>
- Gopal, P., Nartey, W., Ragunathan, P., Sarathy, J., Kaya, F., Yee, M., Setzer, C., Manimekalai, M. S. S., Dartois, V., Grüber, G., & Dick, T. (2017). Pyrazinoic Acid Inhibits Mycobacterial Coenzyme A Biosynthesis by Binding to Aspartate Decarboxylase PanD. *ACS Infectious Diseases*, 3(11), 807–819.
<https://doi.org/10.1021/acsinfecdis.7b00079>
- Gopal, P., Sarathy, J. P., Yee, M., Ragunathan, P., Shin, J., Bhushan, S., Zhu, J., Akopian, T., Kandror, O., Lim, T. K., Gengenbacher, M., Lin, Q., Rubin, E. J., Grüber, G., & Dick, T. (2020a). Pyrazinamide triggers degradation of its target aspartate decarboxylase. *Nature Communications*, 11(1), 1–10.
<https://doi.org/10.1038/s41467-020-15516-1>
- Gopal, P., Sarathy, J. P., Yee, M., Ragunathan, P., Shin, J., Bhushan, S., Zhu, J., Akopian, T., Kandror, O., Lim, T. K., Gengenbacher, M., Lin, Q., Rubin, E. J., Grüber, G., & Dick, T. (2020b). Pyrazinamide triggers degradation of its target aspartate decarboxylase. *Nature Communications*, 11(1), 1–10.
<https://doi.org/10.1038/s41467-020-15516-1>
- Gopal, P., Tasneen, R., Yee, M., Lanoix, J.-P., Sarathy, J., Rasic, G., Li, L., Dartois, V., Nuermberger, E., & Dick, T. (2017). In Vivo-Selected Pyrazinoic Acid-Resistant Mycobacterium tuberculosis Strains Harbor Missense Mutations in the Aspartate Decarboxylase PanD and the Unfoldase ClpC1. *ACS Infectious Diseases*, 3(7), 492–501. <https://doi.org/10.1021/acsinfecdis.7b00017>
- Gottesman, S., Squires, C., Pichersky, E., Carrington, M., Hobbs, M., Mattick, J. S., Dalrymple, B., Kuramitsu, H., Shiroza, T., Foster, T., Clark, W. P., Ross, B., Squires, C. L., & Maurizi, M. R. (1990). Conservation of the regulatory subunit for the Clp ATP-dependent protease in prokaryotes and eukaryotes. *Proceedings of the National Academy of Sciences of the United States of America*, 87(9), 3513–3517.
<https://doi.org/10.1073/pnas.87.9.3513>
- Gottesman, S., Wickner, S., & Maurizi, M. R. (1997). Protein quality control: Triage by chaperones and proteases. *Genes and Development*, 11(7), 815–823.
<https://doi.org/10.1101/gad.11.7.815>
- Guex, N., & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis*, 18(15), 2714–

2723. <https://doi.org/10.1002/elps.1150181505>

- Halperin, T., Ostersetzer, O., & Adam, Z. (2001). ATP-dependent association between subunits of Clp protease in pea chloroplasts. *Planta*, 213(4), 614–619. <https://doi.org/10.1007/s004250100527>
- Hebditch, M., Carballo-Amador, M. A., Charonis, S., Curtis, R., & Warwicker, J. (2017). Protein-Sol: A web tool for predicting protein solubility from sequence. *Bioinformatics*, 33(19), 3098–3100. <https://doi.org/10.1093/bioinformatics/btx345>
- Horwich, A. L., Weber-Ban, E. U., & Finley, D. (1999). Chaperone rings in protein folding and degradation. *Proceedings of the National Academy of Sciences of the United States of America*, 96(20), 11033–11040. <https://doi.org/10.1073/pnas.96.20.11033>
- Idicula-Thomas, S., & Balaji, P. V. (2005). Understanding the relationship between the primary structure of proteins and its propensity to be soluble on overexpression in *Escherichia coli*. *Protein Science*, 14(3), 582–592. <https://doi.org/10.1110/ps.041009005>
- Kandror, O., Yang, S.-H., Goldberg, A. L., Anderson, J. R., Lee, I.-A., Suh, J.-W., Jin, Y.-Y., Sunoqrot, S., Lee, S.-Y., Mulugeta, S., Cho, S., Wang, Y., Hong, S., Franzblau, S. G., Yoon, T.-M., Pauli, G. F., Kim, J.-Y., Gao, W., Akopian, T., ... Kim, J.-W. (2014). The Cyclic Peptide Ecumicin Targeting ClpC1 Is Active against *Mycobacterium tuberculosis* In Vivo. *Antimicrobial Agents and Chemotherapy*, 59(2), 880–889. <https://doi.org/10.1128/aac.04054-14>
- Kang, S. G., Ortega, J., Singh, S. K., Wang, N., Huang, N. N., Steven, A. C., & Maurizi, M. R. (2002). Functional proteolytic complexes of the human mitochondrial ATP-dependent protease, hClpXP. *Journal of Biological Chemistry*, 277(23), 21095–21102. <https://doi.org/10.1074/jbc.M201642200>
- Kapust, R. B., & Waugh, D. S. (1999). *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Science*, 8(8), 1668–1674. <https://doi.org/10.1110/ps.8.8.1668>
- Kar, N. P., Sikriwal, D., Rath, P., Choudhary, R. K., & Batra, J. K. (2008). *Mycobacterium tuberculosis* ClpC1: Characterization and role of the N-terminal domain in its function. *FEBS Journal*, 275(24), 6149–6158. <https://doi.org/10.1111/j.1742-4658.2008.06738.x>
- Kirstein, J., Schlothauer, T., Dougan, D. A., Lilie, H., Tischendorf, G., Mogk, A.,

- Bukau, B., & Turgay, K. (2006). Adaptor protein controlled oligomerization activates the AAA + protein ClpC. *EMBO Journal*, 25(7), 1481–1491. <https://doi.org/10.1038/sj.emboj.7601042>
- Komar, A. A., Lesnik, T., & Reiss, C. (1999). Synonymous codon substitutions affect ribosome traffic and protein folding during in vitro translation. *FEBS Letters*, 462(3), 387–391. [https://doi.org/10.1016/S0014-5793\(99\)01566-5](https://doi.org/10.1016/S0014-5793(99)01566-5)
- Kopp, J., & Schwede, T. (2004). The SWISS-MODEL repository of annotated three-dimensional protein structure homology models. *Nucleic Acids Research*, 32(DATABASE ISS.), D230–D234. <https://doi.org/10.1093/nar/gkh008>
- Kruger, N. J. (1994). The Bradford method for protein quantitation. In *Methods in molecular biology (Clifton, N.J.)* (Vol. 32, pp. 9–15). Methods Mol Biol. <https://doi.org/10.1385/0-89603-268-x:9>
- Kuipers, G., Karyolaimos, A., Zhang, Z., Ismail, N., Trinco, G., Vikström, D., Slotboom, D. J., & de Gier, J. W. (2017). The tunable pReX expression vector enables optimizing the T7-based production of membrane and secretory proteins in *E. coli*. *Microbial Cell Factories*, 16(1), 226. <https://doi.org/10.1186/s12934-017-0840-4>
- Laboratories Inc, C. (2011). *In-Fusion® HD Cloning Kit User Manual*. www.clontech.com
- Lee, C., Kim, J., Shin, S. G., & Hwang, S. (2006). Absolute and relative QPCR quantification of plasmid copy number in *Escherichia coli*. *Journal of Biotechnology*, 123(3), 273–280. <https://doi.org/10.1016/j.jbiotec.2005.11.014>
- Lee, H., & Suh, J. W. (2016). Anti-tuberculosis lead molecules from natural products targeting *Mycobacterium tuberculosis* ClpC1. *Journal of Industrial Microbiology and Biotechnology*, 43(2–3), 205–212. <https://doi.org/10.1007/s10295-015-1709-3>
- Leodolter, J., Warweg, J., & Weber-Ban, E. (2015). The *Mycobacterium tuberculosis* ClpP1P2 protease interacts asymmetrically with its ATPase partners ClpX and ClpC1. *PLoS ONE*, 10(5), 1–21. <https://doi.org/10.1371/journal.pone.0125345>
- Lilie, H., Schwarz, E., & Rudolph, R. (1998). Advances in refolding of proteins produced in *E. coli*. *Current Opinion in Biotechnology*, 9(5), 497–501. [https://doi.org/10.1016/S0958-1669\(98\)80035-9](https://doi.org/10.1016/S0958-1669(98)80035-9)
- Lindquist, S., & Craig, E. A. (1988). The heat -shock proteins. *Annual Review of Genetics*, 22, 631–677.

- Lunge, A., Gupta, R., Choudhary, E., & Agarwal, N. (2020). The unfoldase ClpC1 of mycobacterium tuberculosis regulates the expression of a distinct subset of proteins having intrinsically disordered termini. *Journal of Biological Chemistry*, 295(28), 9455–9473. <https://doi.org/10.1074/jbc.ra120.013456>
- Luong, T. T., Sau, K., Roux, C., Sau, S., Dunman, P. M., & Lee, C. Y. (2011). *Staphylococcus aureus* ClpC divergently regulates capsule via *sae* and *codY* in strain Newman but activates capsule via *codY* in strain UAMS-1 and in strain Newman with repaired *saeS*. *Journal of Bacteriology*, 193(3), 686–694. <https://doi.org/10.1128/JB.00987-10>
- Makrides, S. C. (1996). Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiological Reviews*, 60(3), 512–538. <https://doi.org/10.1128/mr.60.3.512-538.1996>
- Marsee, J. D., Ridings, A., Yu, T., & Miller, J. M. (2018). Mycobacterium tuberculosis ClpC1 N-terminal domain is dispensable for adaptor protein-dependent allosteric regulation. *International Journal of Molecular Sciences*, 19(11), 3651. <https://doi.org/10.3390/ijms19113651>
- Maurizi, M. R., Clark, W. P., Kim, S. H., & Gottesman, S. (1990). Clp P represents a unique family of serine proteases. *Journal of Biological Chemistry*, 265(21), 12546–12552. [https://doi.org/10.1016/s0021-9258\(19\)38379-6](https://doi.org/10.1016/s0021-9258(19)38379-6)
- McGillivray, A., Golden, N. A., & Kaushal, D. (2015). The Mycobacterium tuberculosis Clp gene regulator is required for in vitro reactivation from hypoxia-induced dormancy. *The Journal of Biological Chemistry*, 290(4), 2351–2367. <https://doi.org/10.1074/jbc.M114.615534>
- Mogk, A., Mayer, M. P., & Deuerling, E. (2010). Mechanisms of Protein Folding: Molecular Chaperones and Their Application in Biotechnology. *ChemInform*, 33(46), no-no. <https://doi.org/10.1002/chin.200246272>
- Msadek, T., Kunst, F., & Rapoport, G. (1994). MecB of *Bacillus subtilis*, a member of the ClpC ATPase family, is a pleiotropic regulator controlling competence gene expression and growth at high temperature. *Proceedings of the National Academy of Sciences of the United States of America*, 91(13), 5788–5792. <https://doi.org/10.1073/pnas.91.13.5788>
- Naganathan, A. N., Doshi, U., & Muñoz, V. (2007). Protein folding kinetics: Barrier effects in chemical and thermal denaturation experiments. *Journal of the American*

- Chemical Society*, 129(17), 5673–5682. <https://doi.org/10.1021/ja0689740>
- Nikaido, H. (1994). Maltose transport system of *Escherichia coli*: An ABC-type transporter. *FEBS Letters*, 346(1), 55–58. [https://doi.org/10.1016/0014-5793\(94\)00315-7](https://doi.org/10.1016/0014-5793(94)00315-7)
- Njire, M., Tan, Y., Mugweru, J., Wang, C., Guo, J., Yew, W., Tan, S., & Zhang, T. (2016). Pyrazinamide resistance in *Mycobacterium tuberculosis*: Review and update. *Advances in Medical Sciences*, 61(1), 63–71. <https://doi.org/10.1016/j.advms.2015.09.007>
- Owczarzy, R., Tataurov, A. V., Wu, Y., Manthey, J. A., McQuisten, K. A., Almagbrazi, H. G., Pedersen, K. F., Lin, Y., Garretson, J., McEntaggart, N. O., Sailor, C. A., Dawson, R. B., & Peek, A. S. (2008). IDT SciTools: a suite for analysis and design of nucleic acid oligomers. *Nucleic Acids Research*, 36(Web Server issue). <https://doi.org/10.1093/nar/gkn198>
- Pan, S. H., & Malcolm, B. A. (2000). Reduced background expression and improved plasmid stability with pET vectors in BL21 (DE3). *BioTechniques*, 29(6), 1234–1238. <https://doi.org/10.2144/00296st03>
- Raran-Kurussi, S., Cherry, S., Zhang, D., & Waugh, D. S. (2017). Removal of affinity tags with TEV protease. In *Methods in Molecular Biology* (Vol. 1586, pp. 221–230). Humana Press Inc. https://doi.org/10.1007/978-1-4939-6887-9_14
- Raran-Kurussi, S., & Waugh, D. S. (2012). The Ability to Enhance the Solubility of Its Fusion Partners Is an Intrinsic Property of Maltose-Binding Protein but Their Folding Is Either Spontaneous or Chaperone-Mediated. *PLoS ONE*, 7(11), e49589. <https://doi.org/10.1371/journal.pone.0049589>
- Richter, K., Haslbeck, M., & Buchner, J. (2010, October 22). The Heat Shock Response: Life on the Verge of Death. *Molecular Cell*, 40(2), 253–266. <https://doi.org/10.1016/j.molcel.2010.10.006>
- Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: Advances and challenges. *Frontiers in Microbiology*, 5(APR), 1–17. <https://doi.org/10.3389/fmicb.2014.00172>
- Rouquette, C., De Chastellier, C., Nair, S., & Berche, P. (1998). The ClpC ATPase of *Listeria monocytogenes* is a general stress protein required for virulence and promoting early bacterial escape from the phagosome of macrophages. *Molecular Microbiology*, 27(6), 1235–1245. <https://doi.org/10.1046/j.1365->

[2958.1998.00775.x](#)

- Samuelson, J. (2011). Production of Membrane Proteins: Strategies for Expression and Isolation: Expression systems. *Production of Membrane Proteins: Strategies for Expression and Isolation*, 11–35. http://www.wiley-vch.de/books/sample/3527327290_c01.pdf
- Sargsyan, K., Grauffel, C., & Lim, C. (2017). How Molecular Size Impacts RMSD Applications in Molecular Dynamics Simulations. *Journal of Chemical Theory and Computation*, 13(4), 1518–1524. <https://doi.org/10.1021/acs.jctc.7b00028>
- Sassetti, C. M., Boyd, D. H., & Rubin, E. J. (2003). Genes required for mycobacterial growth defined by high density mutagenesis. *Molecular Microbiology*, 48(1), 77–84. <https://doi.org/10.1046/j.1365-2958.2003.03425.x>
- Schein, C. H. (1990). Solubility as a function of protein structure and solvent components. *Bio/Technology*, 8(4), 308–317. <https://doi.org/10.1038/nbt0490-308>
- Schlegel, S., Löfblom, J., Lee, C., Hjelm, A., Klepsch, M., Strous, M., Drew, D., Slotboom, D. J., & De Gier, J. W. (2012). Optimizing membrane protein overexpression in the Escherichia coli strain Lemo21(DE3). *Journal of Molecular Biology*, 423(4), 648–659. <https://doi.org/10.1016/j.jmb.2012.07.019>
- Schmidt, L., & Inselburg, J. (1982). ColE1 copy number mutants. *Journal of Bacteriology*, 151(2), 845–854. <https://doi.org/10.1128/jb.151.2.845-854.1982>
- Schmitz, K. R., & Sauer, R. T. (2014). Substrate delivery by the AAA+ ClpX and ClpC1 unfoldases activates the mycobacterial ClpP1P2 peptidase. *Molecular Microbiology*, 93(4), 617–628. <https://doi.org/10.1111/mmi.12694>
- Shilling, P. J., Mirzadeh, K., Cumming, A. J., Widesheim, M., Köck, Z., & Daley, D. O. (2020). Improved designs for pET expression plasmids increase protein production yield in Escherichia coli. *Communications Biology*, 3(1), 1–8. <https://doi.org/10.1038/s42003-020-0939-8>
- Singh, A., Upadhyay, V., Upadhyay, A. K., Singh, S. M., & Panda, A. K. (2015). Protein recovery from inclusion bodies of Escherichia coli using mild solubilization process. *Microbial Cell Factories* 2015 14:1, 14(1), 1–10. <https://doi.org/10.1186/S12934-015-0222-8>
- Sjögren, L. L. E., MacDonald, T. M., Sutinen, S., & Clarke, A. K. (2004). Inactivation of the clpC1 gene encoding a chloroplast Hsp100 molecular chaperone causes growth retardation, leaf chlorosis, lower photosynthetic activity, and a specific

- reduction in photosystem content. *Plant Physiology*, 136(4), 4114–4126.
<https://doi.org/10.1104/pp.104.053835>
- Squires, C. L., Pedersen, S., Ross, B. M., & Squires, C. (1991). ClpB is the Escherichia coli heat shock protein F84.1. *Journal of Bacteriology*, 173(14), 4254–4262.
<https://doi.org/10.1128/jb.173.14.4254-4262.1991>
- Squires, C., & Squires, C. L. (1992). The Clp proteins: Proteolysis regulators or molecular chaperones? In *Journal of Bacteriology* (Vol. 174, Issue 4, pp. 1081–1085). <https://doi.org/10.1128/jb.174.4.1081-1085.1992>
- Stevens, R. C. (2000). Design of high-throughput methods of protein production for structural biology. In *Structure* (Vol. 8, Issue 9, pp. 177–185). Current Biology Ltd.
[https://doi.org/10.1016/S0969-2126\(00\)00193-3](https://doi.org/10.1016/S0969-2126(00)00193-3)
- Strub, C., Alies, C., Lougarre, A., Ladurantie, C., Czaplicki, J., & Fournier, D. (2004). Mutation of exposed hydrophobic amino acids to arginine to increase protein stability. *BMC biochemistry*, 5, 9. <https://doi.org/10.1186/1471-2091-5-9>
- Studier, F. W. (1991). Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *Journal of Molecular Biology*, 219(1), 37–44.
[https://doi.org/10.1016/0022-2836\(91\)90855-Z](https://doi.org/10.1016/0022-2836(91)90855-Z)
- Sun, P., Tropea, J. E., & Waugh, D. S. (2011). Enhancing the solubility of recombinant proteins in escherichia coli by using hexahistidine-tagged maltose-binding protein as a fusion partner. *Methods in Molecular Biology*, 705, 259–274.
https://doi.org/10.1007/978-1-61737-967-3_16
- Sun, Q., Li, X., Perez, L. M., Shi, W., Zhang, Y., & Sacchettini, J. C. (2020). The molecular basis of pyrazinamide activity on Mycobacterium tuberculosis PanD. *Nature Communications*, 11(1), 1–7. <https://doi.org/10.1038/s41467-019-14238-3>
- Turgay, K., Hamoen, L. W., Venema, G., & Dubnau, D. (1997). Biochemical characterization of a molecular switch involving the heat shock protein ClpC, which controls the activity of ComK, the competence transcription factor of Bacillus subtilis. *Genes and Development*, 11(1), 119–128.
<https://doi.org/10.1101/gad.11.1.119>
- Vasudevan, D., Rao, S. P. S., & Noble, C. G. (2013). Structural Basis of Mycobacterial Inhibition by Cyclomarin A. *Journal of Biological Chemistry*, 288(43), 30883–30891. <https://doi.org/10.1074/jbc.M113.493767>
- Ventura, S., & Villaverde, A. (2006). Protein quality in bacterial inclusion

- bodies. *Trends in biotechnology*, 24(4), 179–185.
<https://doi.org/10.1016/j.tibtech.2006.02.007>
- Vera, A., González-Montalbán, N., Arís, A., & Villaverde, A. (2007). The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. *Biotechnology and Bioengineering*, 96(6), 1101–1106.
<https://doi.org/10.1002/bit.21218>
- Walls, D., & Loughran, S. T. (2011). Tagging recombinant proteins to enhance solubility and aid purification. In *Methods in molecular biology (Clifton, N.J.)* (Vol. 681, pp. 151–175). Methods Mol Biol. https://doi.org/10.1007/978-1-60761-913-0_9
- Wang, F., Mei, Z., Qi, Y., Yan, C., Hu, Q., Wang, J., & Shi, Y. (2011). Structure and mechanism of the hexameric MecA-ClpC molecular machine. *Nature*, 471(7338), 331–337. <https://doi.org/10.1038/nature09780>
- 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(W1), W296–W303.
<https://doi.org/10.1093/nar/gky427>
- Weinhäupl, K., Brennich, M., Kzmaier, U., Lelievre, J., Ballell, L., Goldberg, A., Schanda, P., & Fraga, H. (2018). The antibiotic cyclomarin blocks arginine-phosphate-induced millisecond dynamics in the N-terminal domain of ClpC1 from *Mycobacterium tuberculosis*. *Journal of Biological Chemistry*, 293(22), 8379–8393. <https://doi.org/10.1074/jbc.RA118.002251>
- Wilkinson, D. L., & Harrison, R. G. (1991). Predicting the solubility of recombinant proteins in *Escherichia coli*. *Bio/technology* (Nature Publishing Company), 9(5), 443–448. <https://doi.org/10.1038/nbt0591-443>
- William Studier, F., Rosenberg, A. H., Dunn, J. J., & Dubendorff, J. W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods in Enzymology*, 185(C), 60–89. [https://doi.org/10.1016/0076-6879\(90\)85008-C](https://doi.org/10.1016/0076-6879(90)85008-C)
- Wolf, N. M., Lee, H., Choules, M. P., Pauli, G. F., Phansalkar, R., Anderson, J. R., Gao, W., Ren, J., Santarsiero, B. D., Lee, H., Cheng, J., Jin, Y. Y., Ho, N. A., Duc, N. M., Suh, J. W., Abad-Zapatero, C., & Cho, S. (2019). High-Resolution Structure of ClpC1-Rufomycin and Ligand Binding Studies Provide a Framework to Design

- and Optimize Anti-Tuberculosis Leads. *ACS Infectious Diseases*, 5(6), 829–840.
<https://doi.org/10.1021/acsinfecdis.8b00276>
- Yee, M., Gopal, P., & Dick, T. (2016). Missense mutations in the unfoldase ClpC1 of the caseinolytic protease complex are associated with pyrazinamide resistance in *Mycobacterium tuberculosis*. *Antimicrobial Agents and Chemotherapy*, AAC.02342-16. <https://doi.org/10.1128/AAC.02342-16>
- Zhang, Shoudong, Zhang, H., Xia, Y., & Xiong, L. (2018). The caseinolytic protease complex component CLPC1 in *Arabidopsis* maintains proteome and RNA homeostasis in chloroplasts. *BMC Plant Biology*, 18(1), 1–16.
<https://doi.org/10.1186/s12870-018-1396-0>
- Zhang, Shuo, Chen, J., Shi, W., Cui, P., Zhang, J., Cho, S., Zhang, W., & Zhang, Y. (2017). Mutation in *clpC1* encoding an ATP-dependent ATPase involved in protein degradation is associated with pyrazinamide resistance in *Mycobacterium tuberculosis*. *Emerging Microbes & Infections*, 6(2), e8.
<https://doi.org/10.1038/emi.2017.1>
- Zhang, Y., & Skolnick, J. (2005). TM-align: A protein structure alignment algorithm based on the TM-score. *Nucleic Acids Research*, 33(7), 2302–2309.
<https://doi.org/10.1093/nar/gki524>

10. ANEXOS

ANEXO 1

Resultado del efecto de las interacciones de tipo de plásmidos, cepa, concentración de IPTG y temperatura en las concentraciones y porcentaje de solubilidad de ClpC1.

Interacción				Concentración de proteína (mg/ml)	Solubilidad (%)
Plásmido	Tipo de cepa	[I] IPTG (mM)	Temperatura (C°) y tiempo (h)		
pET28a (+)-ClpC1	BL21(DE3) pLysS	0.25	16 y 10	7	24.74
	BL21(DE3) pLysS	0.25	30 y 4	9.27	31.08
	BL21(DE3) pLysS	0.25	37 y 4	8.92	44.09
	BL21(DE3) pLysS	0.5	16 y 10	45.65	22.25
	BL21(DE3) pLysS	0.5	30 y 4	14.63	35.44
	BL21(DE3) pLysS	0.5	37 y 4	12.81	29.38
	BL21(DE3) pLysS	1	16 y 10	44.87	19.21
	BL21(DE3) pLysS	1	30 y 4	16.64	22.49

BL21(DE3) pLysS	1	37 y 4	12.08	34.46
BL21(DE3) pLysS	2	16 y 10	33.98	27.26
BL21(DE3) pLysS	2	30 y 4	19.03	21.66
BL21(DE3) pLysS	2	37 y 4	12.61	23.13
Lemo(DE3)pLysS	0.25	16 y 10	7.54	22.28
Lemo(DE3)pLysS	0.25	30 y 4	8.37	31.43
Lemo(DE3)pLysS	0.25	37 y 4	11.58	42.4
Lemo(DE3)pLysS	0.5	16 y 10	38.74	30.7
Lemo(DE3)pLysS	0.5	30 y 4	16.072	42.19
Lemo(DE3)pLysS	0.5	37 y 4	13.93	34.64
Lemo(DE3)pLysS	1	16 y 10	40.66	21.27
Lemo(DE3)pLysS	1	30 y 4	16.46	20.6
Lemo(DE3)pLysS	1	37 y 4	12.07	31.57

Lemo(DE3)pLysS	2	16 y 10	33.77	8.34
Lemo(DE3)pLysS	2	30 y 4	18.1	23.87
Lemo(DE3)pLysS	2	37 y 4	14.48	31.67
Rosetta(DE3)pLysS	0.25	16 y 10	7.85	38.28
Rosetta(DE3)pLysS	0.25	30 y 4	7.7182	32.42
Rosetta(DE3)pLysS	0.25	37 y 4	8.1472	41.6
Rosetta(DE3)pLysS	0.5	16 y 10	38.14	22.52
Rosetta(DE3)pLysS	0.5	30 y 4	18.03	21.47
Rosetta(DE3)pLysS	0.5	37 y 4	16.77	32.54
Rosetta(DE3)pLysS	1	16 y 10	37.14	20.91
Rosetta(DE3)pLysS	1	30 y 4	18.39	41.81
Rosetta(DE3)pLysS	1	37 y 4	27.5646	16.94
Rosetta(DE3)pLysS	2	16 y 10	37.74	22.15

	Rosetta(DE3)pLysS	2	30 y 4	21.44	13.05
	Rosetta(DE3)pLysS	2	37 y 4	17.49	14.3
	BL21(DE3) pLysS	0.25	16 y 10	22.07	93.23
	BL21(DE3) pLysS	0.25	30 y 4	19.5	82.66
	BL21(DE3) pLysS	0.25	37 y 4	18.57	65.03
	BL21(DE3) pLysS	0.5	16 y 10	45.55	86.93
	BL21(DE3) pLysS	0.5	30 y 4	33.55	59.24
	BL21(DE3) pLysS	0.5	37 y 4	29.83	68.89
pMAL-cHT-ClpC1	BL21(DE3) pLysS	1	16 y 10	23.54	84.6
	BL21(DE3) pLysS	1	30 y 4	18.77	68.41
	BL21(DE3) pLysS	1	37 y 4	15.15	63.84
	BL21(DE3) pLysS	2	16 y 10	18.63	60.55
	BL21(DE3) pLysS	2	30 y 4	19.41	61.47

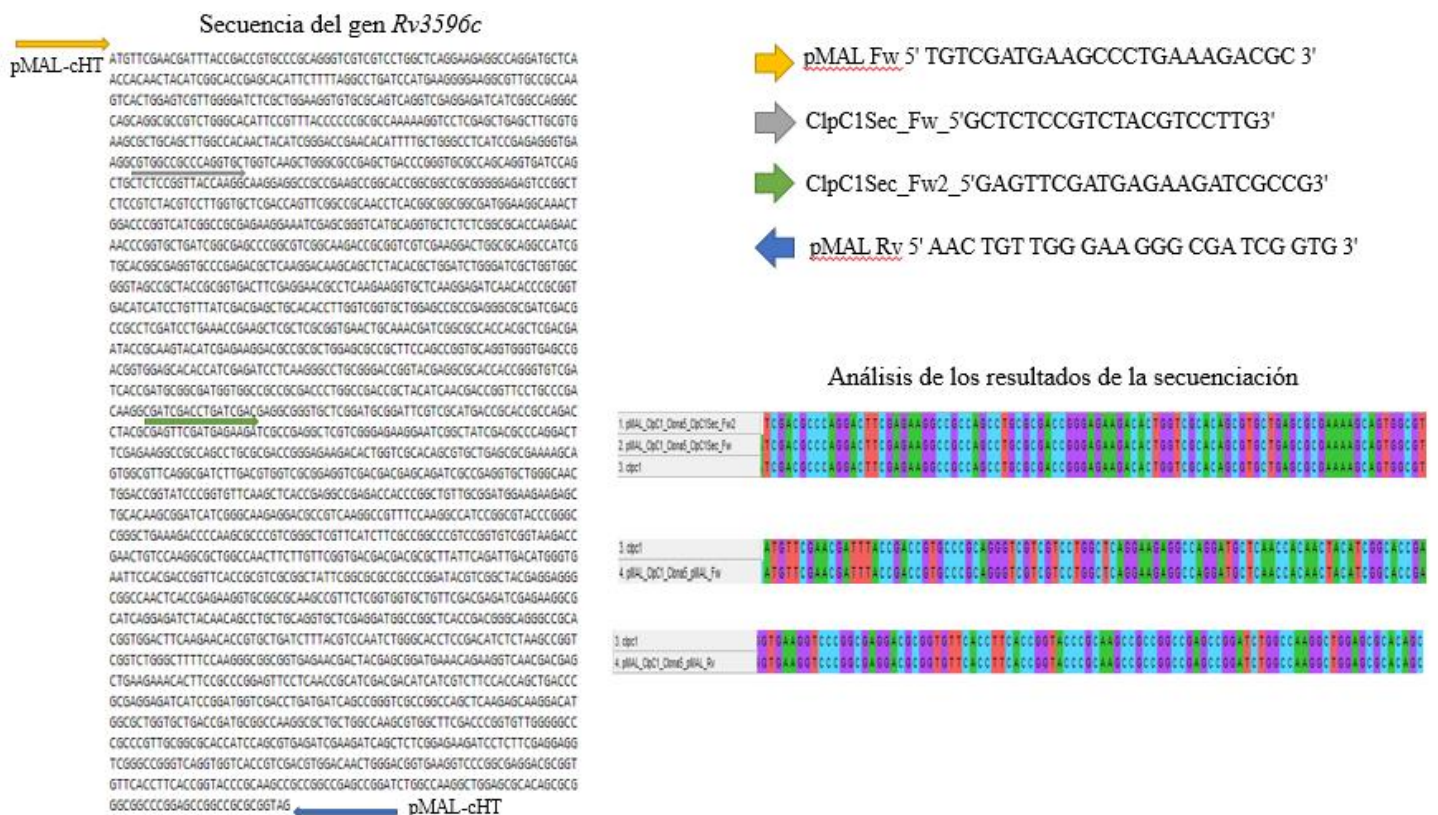
BL21(DE3) pLysS	2	37 y 4	17.98	67.27
Lemo(DE3)pLysS	0.25	16 y 10	6.06	50.81
Lemo(DE3)pLysS	0.25	30 y 4	6.49	48.42
Lemo(DE3)pLysS	0.25	37 y 4	6.12	52.48
Lemo(DE3)pLysS	0.5	16 y 10	17.54	44.85
Lemo(DE3)pLysS	0.5	30 y 4	17.44	49.74
Lemo(DE3)pLysS	0.5	37 y 4	13.05	59.9
Lemo(DE3)pLysS	1	16 y 10	18.07	46.57
Lemo(DE3)pLysS	1	30 y 4	20.14	49.15
Lemo(DE3)pLysS	1	37 y 4	19.62	54.66
Lemo(DE3)pLysS	2	16 y 10	19.27	52.72
Lemo(DE3)pLysS	2	30 y 4	17.99	49.25
Lemo(DE3)pLysS	2	37 y 4	18.16	53.68

Rosetta(DE3)pLysS	0.25	16 y 10	9.09	60.64
Rosetta(DE3)pLysS	0.25	30 y 4	11.05	43.02
Rosetta(DE3)pLysS	0.25	37 y 4	11.83	51.18
Rosetta(DE3)pLysS	0.5	16 y 10	19.86	46.85
Rosetta(DE3)pLysS	0.5	30 y 4	21.77	38.6
Rosetta(DE3)pLysS	0.5	37 y 4	20.96	42.87
Rosetta(DE3)pLysS	1	16 y 10	18.55	43.68
Rosetta(DE3)pLysS	1	30 y 4	18.93	43.85
Rosetta(DE3)pLysS	1	37 y 4	17.43	40.64
Rosetta(DE3)pLysS	2	16 y 10	20.89	49.54
Rosetta(DE3)pLysS	2	30 y 4	22.17	44.23
Rosetta(DE3)pLysS	2	37 y 4	23.91	38.99

ANEXO 2

Con la finalidad de secuenciar todo el gen *Rv3596c* en los plásmidos clonados, se diseñaron un par de primers para pMAL-cHT y pET28a(+) y dos primers forward para *Rv3596c*. En ambos resultados, *Rv3596c* se clonó sin mutaciones, en la dirección correcta y con un 100% de identidad con respecto a la secuencia nativa de *Rv3596c*.

La siguiente figura muestra el análisis de los resultados de la secuenciación en pMAL-cHT-ClpC1.



De la misma manera, el análisis de los resultados de la secuenciación en pET28a(+)- ClpC1 se muestran en la siguiente figura.

