

**UNIVERSIDAD PERUANA CAYETANO HEREDIA**

**FACULTAD DE CIENCIAS Y FILOSOFÍA “ALBERTO CAZORLA  
TALLERÍ”**



**Expresión, purificación y actividad enzimática de la proteína recombinante Catepsina L5 de *Fasciola hepatica* (rCatL5Fh) obtenida en un sistema de expresión procariota (*E. coli* BL21).**

Yerson Jahel Buiza Urdanivia

Tesis para optar por el título profesional de LICENCIADO en  
BIOLOGÍA

Lima, Perú

2021

**MIEMBROS DEL JURADO QUE APROBARON LA TESIS:**

**LUIS DE STEFANO BELTRAN:                   PRESIDENTE**

**ANA COLAROSI SALINAS:                   VOCAL**

**ROCÍO INGA PEÑA:                   SECRETARIO**

**ASESOR:** Dr. José Ronald Espinoza Babilón

**CO-ASESORA:** MSc. Jenniffer Angélica Espinoza Ramírez

**DEDICATORIA**

A mis padres Gregorio y Gloria

Por el inmenso apoyo y amor desde siempre.

A mis hermanos,

Por todos los momentos buenos y malos siempre junto a mí.

## **AGRADECIMIENTOS**

A mis asesores Dr. José Espinoza Babilón y Maestra Jenniffer Espinoza Ramírez por la enseñanza brindada y las recomendaciones durante el desarrollo del trabajo.

A la Dra. Patricia Herrera por el apoyo y la confianza brindada para el desarrollo del proyecto.

A todos los miembros del laboratorio Unidad de Biotecnología Molecular, que con sus recomendaciones y ayudas se hizo posible la culminación de este trabajo, y además de brindarme su inapreciable amistad: Teresa Barreto, Luis Laura, Mabel Pereda, Henry Becerra.

El presente trabajo de licenciatura se realizó gracias a que fui recipiente de una beca del programa beca 18 del Ministerio de Educación que financió mis estudios de pregrado en la carrera de biología.

El proyecto “Nuevo agente biológico recombinante con actividad fibrinolítica para su uso en la disolución de trombos”, Contrato N° 142-2017-FONDECYT, financió mis prácticas pre-profesionales y los recursos para el desarrollo del presente trabajo.



## RESUMEN

*Fasciola hepatica* produce, durante la infección, una serie de cisteínas tipo catepsina L la cual permiten al parásito la invasión, alimentación y evasión del sistema inmune. El tremátodo expresa 7 isoformas de Catepsina tipo L, siendo catepsina L5 (CatL5Fh) secretada en mayor proporción en el estadio adulto; por lo que se considera que esta proteinasa está involucrada en la supervivencia y alimentación ya que puede degradar componentes proteicos del hospedero definitivo. En la Unidad de Biotecnología Molecular (UBM) se ha determinado que CatL5Fh puede degradar ST3 (codificado en UBM) por lo cual resulta de interés evaluar esta cisteína proteinasa como un potencial agente de aplicación biomédico. Debido a que la purificación de CatL5Fh nativa a partir de parásitos adultos requiere de inversión de costo y tiempo resultando en bajos rendimientos de la proteinasa pura; el objetivo del presente estudio es expresar rCatL5Fh en el sistema de expresión del vector pET-26b(+) en las cepas BL21 y Rosetta de *Escherichia coli*, purificar y caracterizar su actividad con el sustrato ST3. rCatL5Fh fue clonada y expresada por fermentación con diferentes concentraciones de IPTG y en distintos tiempos para controlar la formación de cuerpos de inclusión. El más alto rendimiento de producción de rCatL5Fh purificada que se obtuvo con BL21 fue de 4.5 mg por 100 ml de cultivo con 1 mM de IPTG, 37 °C por 16 horas y con Rosetta fue de 6 mg por 100 ml con 0.25 mM de IPTG, 37 °C por 8 horas. Sin embargo, rCatL5Fh purificada por distintos procedimientos no tuvo actividad proteolítica por lo que se sugiere que esta proteinasa requiere transitar por el sistema de secreción para lograr las modificaciones post-traduccionales y el plegamiento adecuado de su forma activa, procesos que no ocurren en los sistemas de expresión procariotas.

**Palabras claves:** *Escherichia coli*, recombinante, proteasa, *Fasciola hepatica*, catepsina L5.

## ABSTRACT

*Fasciola hepatica* produces, during infection, a series of cathepsin L-type cysteines which allow the parasite to invade, feed, and escape the immune system. The trematode expresses 7 isoforms of Cathepsin type L, being cathepsin L5 (CatL5Fh) secreted in a higher proportion in the adult stage; therefore, it is considered that this proteinase is involved in survival and feeding since it can degrade protein components of the definitive host. At the Molecular Biotechnology Unit (UBM) it has been determined that CatL5Fh can degrade ST3 (encoded in UBM), therefore it is of interest to evaluate this cysteine proteinase as a potential agent for biomedical application. Because the purification of native CatL5Fh from adult parasites requires investment of cost and time resulting in low yields of pure proteinase; The objective of the present study is to express rCatL5Fh in the expression system of the vector pET-26b (+) in the BL21 and Rosetta strains of *Escherichia coli*, to purify and characterize its activity with the ST3 substrate. rCatL5Fh was cloned and expressed by fermentation with different concentrations of IPTG and at different times to control the formation of inclusion bodies. The highest production yield of purified rCatL5Fh that was obtained with BL21 was 4.5 mg per 100 ml of culture with 1 mM IPTG, 37 ° C for 16 hours and with Rosetta it was 6 mg per 100 ml with 0.25 mM IPTG, 37 ° C for 8 hours. However, rCatL5Fh purified by different procedures did not have proteolytic activity, which is why it is suggested that this proteinase requires transit through the secretion system to achieve post-translational modifications and proper folding of its active form, processes that do not occur in systems of prokaryotic expression.

**Key words: *Escherichia coli*, recombinant, protease, *Fasciola hepatica*, cathepsin L5.**



## TABLA DE CONTENIDO

LISTA DE FIGURAS	6
LISTA DE TABLAS	8
<b>CAPÍTULO I</b>	
INTRODUCCIÓN	9
I.1 <i>Fasciola hepatica</i>	9
I.1.1 Morfología y clasificación taxonómica	9
I.1.2 Ciclo de vida	10
I.1.3 Organización genómica	12
I.1.4 Proteómica	12
I.1.4.1 Cisteína proteinasas	16
I.1.4.1.1 Proceso de activación de cisteína proteasas	17
I.1.4.1.1.1 Auto activación	17
I.1.4.1.1.2 Trans-activación	18
I.1.4.1.2 Mecanismo catalítico	19
I.1.4.2 Catepsinas	20
I.1.4.2.1 Catepsina L	21
I.1.4.2.2 Interacción de catepsina L con su sustrato	23
I.1.4.2.3 Catepsina L5 (CatL5Fh)	25
I.1.4.2.4 Importancia del estudio	26
I.2 Proteína recombinante expresada en <i>E. coli</i>	27
I.3 Purificación de proteína recombinante	29
<b>CAPÍTULO II</b>	
OBJETIVO	31
II.1 Objetivo general	31
II.2 Objetivos específicos	31

## **CAPÍTULO III**

MATERIALES Y MÉTODOS	32
III.1 Clonamiento de la secuencia CatL5Fh hacia el vector pET-26b(+) y transformación del plásmido a <i>E. coli</i> BL21.	32
III.2 Purificación de plásmido pET-26b(+) con CatL5Fh y corroboración de la integridad de la secuencia.	33
III.3 Expresión de la proteína recombinante Catepsina L5 (rCatL5Fh).	34
III.3.1 Inducción	35
III.4 Lisis	35
III.5 Purificación de rCatL5Fh	35
III.5.1 Purificación a condición denaturante (Pellet sonicado)	35
III.5.2 Purificación sin agentes caotrópicos (Sobrenadante sonicado)	36
III.5.3 Purificación a condiciones híbridas	37
III.5.4 Purificación y re-ensamblaje de rCatL5Fh mediante cromatografía de filtración utilizando gel de Sephadex G-25	38
III.6 Cuantificación	38
III.7 Análisis de SDS-PAGE	39
III.8 Activación de la proteína recombinante Catepsina L5	39
III.9 Análisis y comparación de actividad enzimática de la proteína recombinante Catepsina L5 con la nativa.	39
III.10 Modelamiento estructural de rCatL5Fh	40

## **CAPÍTULO IV**

RESULTADOS	41
IV.1 Expresión de la proteína recombinante	43
IV.2 Purificación de rCatL5Fh	48
IV.3 Cuantificación y actividad enzimática de la proteína rCatL5Fh purificada	52
IV.4 Alineamiento estructural entre rCatL5Fh y AF271385.1	55

## **CAPÍTULO V**

DISCUSIÓN	57
-----------	----

## **CAPÍTULO VI**

CONCLUSIÓN

65

## **CAPÍTULO VII**

REFERENCIAS BIBLIOGRÁFICAS

66

## LISTA DE FIGURAS

Figura 1: Diagrama de <i>Fasciola hepatica</i> adulta y huevo	9
Figura 2: Ciclo de vida de <i>F. hepatica</i>	11
Figura 3: Clasificación funcional de proteinasas derivados de <i>F. hepatica</i> juvenil y adulto	14
Figura 4: Clasificación funcional de proteinasas solubles, juvenil y adultos	15
Figura 5: Representación esquemática del pro-dominio y el dominio maduro de las cisteínas proteinasas	17
Figura 6: Mecanismo de autoactivación de catepsina B en presencia de GAG	18
Figura 7: Trans-activación mediado por otra molécula	19
Figura 8: Representación estructural de la cisteína proteinasas	20
Figura 9: Mecanismo catalítico de hidrólisis peptídico catalizado por cisteína proteinasa	20
Figura 10: Perfil de expresión de las catepsinas B e isoformas de L en cada etapa durante la infección de la <i>F. hepatica</i> a su hospedero definitivo	21
Figura 11: Familias de catepsinas L de <i>F. hepatica</i>	22
Figura 12: Localización del sub-sitio en el sitio activo de la catepsina L	23
Figura 13: Nomenclatura de la interacción del sustrato con el sitio activo	24
Figura 14: Estructura 3D de la Catepsina L5 desde <i>Fasciola hepática</i> , resaltando los residuos que poseen altas fluctuaciones	26
Figura 15: Sistema de expresión de proteína pET	28
Figura 16: Representación esquemática de inmovilización del ion Niquel ( $Ni^{+2}$ ) en una matriz NTA y su interacción con el Tag de Poli-Histidina	30
Figura 17: Representación estructural de Imidazol y el residuo de Histidina	30
Figura 18: pET-26b(+) con inserto CatL5Fh.	33
Figura 19: Alineamiento de secuencias de rCatL5Fh con la nativa	42
Figura 20: Curva de crecimiento de <i>E. coli</i> BL21 y Rosetta	43
Figura 21: Expresión de rCatL5Fh en muestras insolubles del cultivo a 37 °C de BL21.	44
Figura 22: Expresión de rCatL5Fh en las muestras insolubles y solubles del cultivo a 37 °C de BL21.	45

Figura 23: Expresión de rCatL5Fh en las muestras insolubles del cultivo a 30 °C de BL21	46
Figura 24: Expresión de rCatL5Fh en las muestras insolubles como solubles del cultivo a 30 °C de BL21	47
Figura 25: Expresión de rCatL5Fh en las muestras insolubles del cultivo a 37 °C de Rosetta	47
Figura 26: Análisis de la expresión de rCatL5Fh mediante Western Blot utilizando anti penta His para marcaje de recombinante	48
Figura 27: Purificación a condiciones desnaturalizantes del cultivo BL21 inducida por 16 horas a 37 °C con 1 mM de IPT	49
Figura 28: Purificación de rCatL5Fh del cultivo de BL21 inducida a 1 y 0.1 mM de IPTG	50
Figura 29: Purificación de rCatL5Fh a condiciones híbridas del cultivo BL21 inducida por 8 horas a 37 °C con 0.25 mM	50
Figura 30: Purificación de rCatL5Fh con Sephadex G-25 del cultivo de BL21 inducida por 8 horas a 37 °C con 0.25 mM	51
Figura 31: Purificación de rCatL5Fh a partir de la cepa Rosetta inducida a 0.25 mM de IPTG de incubada a 37 °C	52
Figura 32: Ensayo de activación de rCatL5Fh purificadas a condiciones desnaturalizantes y no desnaturalizantes	53
Figura 33. Alineamiento estructural entre la proteína recombinante y nativa con visión A) frontal y B) posterior.	55
Figura 34: Aminoácidos distintos entre la recombinante y la nativa ubicados alrededor del sitio de corte entre pre y pro-dominio.	56

## LISTA DE TABLAS

Tabla 1: Ubicación de los aminoácidos en el sub-sitio S2 de las catepsinas L	25
Tabla 2: Preparación de concentraciones finales del estándar BSA	38
Tabla 3: Resumen de la mezcla 1X Dye Reagent con el estándar y muestras	39
Tabla 4: Resumen de la inducción, purificación, concentración y actividad enzimática de la proteína recombinante Catepsina L5 y nativa	54

# CAPÍTULO I

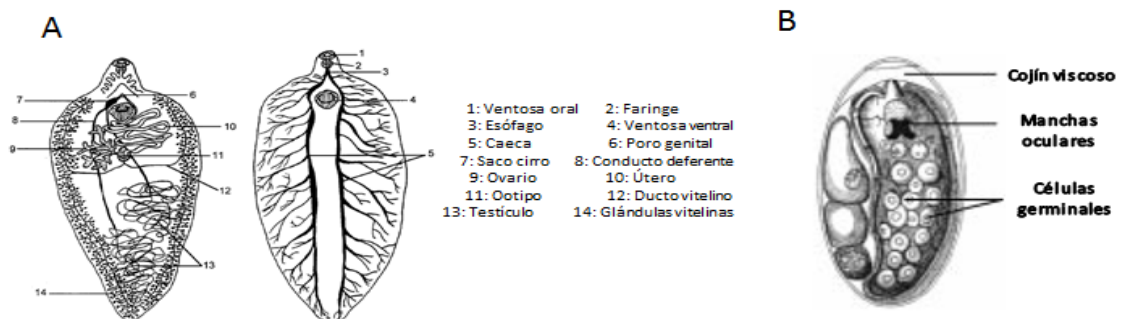
## INTRODUCCIÓN

### I.1 *Fasciola hepatica*

#### I.1.1 Morfología y clasificación taxonómica

*Fasciola hepatica*, conocida como duela del hígado, es un endoparásito hermafrodita perteneciente al filo de platelmintos, clase tremátoda, subclase digenea, orden plagiorchiida y familia fasciolidae (1,2). Este parásito tiene una simetría bilateral con forma de hoja aplanada dorsoventralmente, presenta aparato digestivo incompleto, aparato excretor con protonefridios y ventosas tanto en la boca como en la zona ventral lo cual le permite adherirse a los tejidos del conducto biliar (figura 1A); además, al ser hermafrodita, presenta autofecundación y fertilización cruzada incrementando la diversidad y el flujo genético (1-3).

En el estadio adulto, la *Fasciola hepatica* mide alrededor de 17 a 29 mm de largo y 9 a 12 mm de ancho; de acuerdo a ello (16), es considerado uno de los tremátodos más grandes, por detrás de la *F. gigantica*, a nivel mundial (2,4,5,7,8). Los huevos de la *F. hepatica* miden en promedio 140  $\mu\text{m}$  de largo y 75  $\mu\text{m}$  de ancho (figura 1B) (2,4,6-8,16).



**Figura 1.** A) Diagrama de *F. hepatica* con sus estructuras principales y B) del huevo completamente desarrollado con un miracidio listo para la eclosión. (adaptado de Stuart J. Andrews 1999) (2).

### **I.1.2 Ciclo de vida**

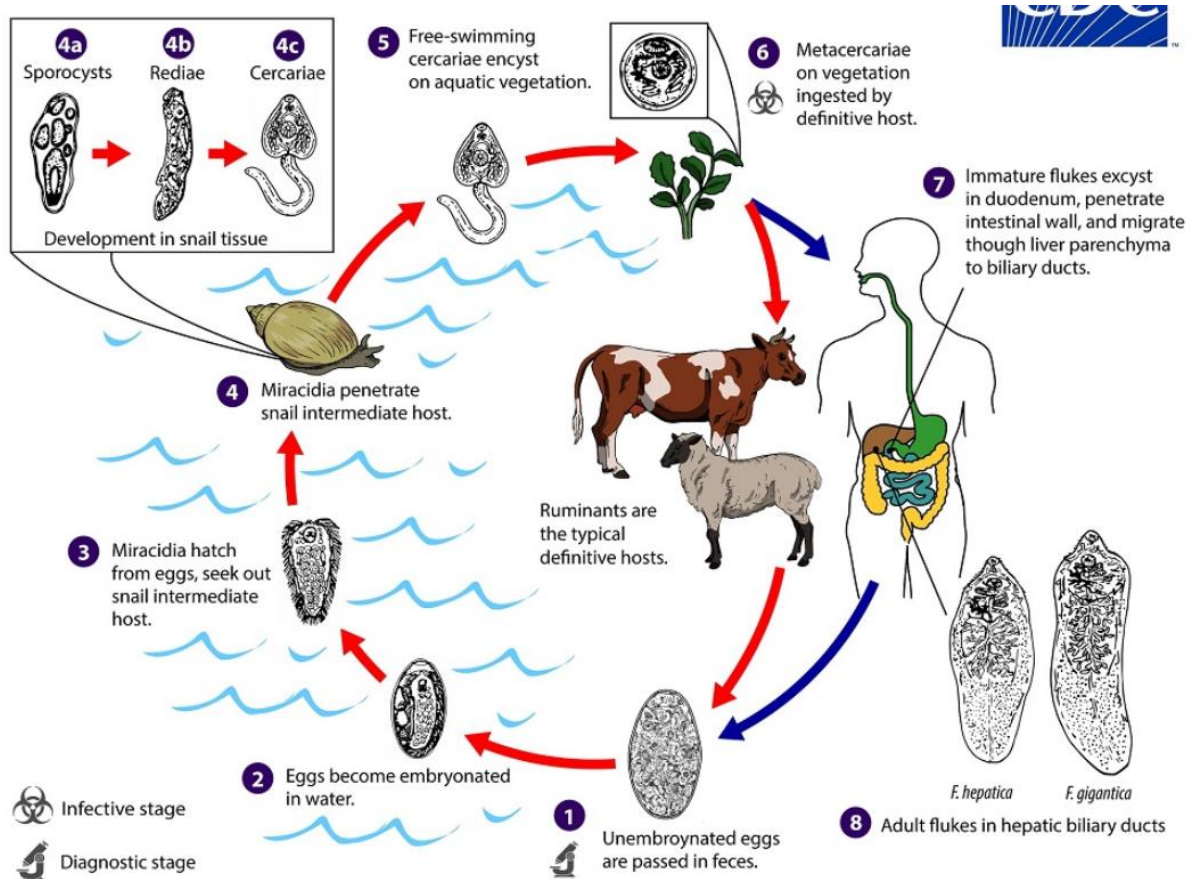
La fascioliasis es una enfermedad parasitaria causada por la *F. hepatica*. La infección es un problema de salud pública global debido a que el parásito infecta humanos y también constituye un serio problema para la ganadería por la alta prevalencia de infección en vacunos, ovejas, camélidos, conejos, entre otros causando disminución en la producción de carne, leche, lana y ocasionando pérdidas significativas en la economía ganadera de los países. (2,4,9-14).

El ciclo de vida del tremátodo (figura 2) se inicia con la excreción de huevos, en las heces, por los animales infectados (2,9,15-18). Los huevos libres en el agua, en condiciones óptimas de temperatura (10-30 °C), luz, humedad, pH y tensión de O<sub>2</sub> favorece la emergencia de los miracidios ciliados que nadan en busca de su hospedero intermediario (caracol del género *Lymnaea* - principal, *Galba*, *Fossaria* y *Pseudosuccinea*) (8,15-22). Existen más de 20 especies del género *Lymnaea* que actúan como transmisores intermediarios del tremátodo. Durante la infección del caracol, el miracidium pierde su capa ciliada y desarrolla un esporoquiste que consiste de múltiples células germinales que dan lugar a una redia el cual será desarrollada y liberada por la explosión de la pared del esporoquiste en las glándulas digestivas (hígado) del caracol (15-22). Al igual que el esporoquiste, la redia consiste de múltiples células germinales que se multiplicarán y darán lugar a la etapa final larval denominada cercaria; en este estadio, el parásito presenta una cola larga que le permite moverse en el medio exterior (15-21).

La cercaria sale del caracol y nada libremente en el agua para adherirse a objetos u hojas acuáticas (15-18). Después de la adherencia, pierde su cola y se enquistas como metacercaria (15,16). Esta última es la forma infectante del parásito y el hospedero definitivo adquiere la infección por la ingesta de la metacercaria adheridas en las hojas o flotando en el agua. Luego de la ingestión, la capa del quiste es destruido en el intestino delgado del hospedero (por proteasas del estómago, sales biliares, tensión de CO<sub>2</sub>, altas temperaturas) y el tremátodo juvenil es liberado para penetrar la pared intestinal, entrar a la cavidad abdominal, atravesar el espacio peritoneal y buscar el hígado (a través de respuesta a estímulos) (2,15,16). La *F. hepatica* juvenil migra por el parénquima hepático del hígado para finalmente alojarse en el ducto biliar, donde alcanza la madurez



sexual e inicia la producción de huevos. El huevo alcanza el duodeno por los canales biliares y son excretados con las heces y así reiniciar otro ciclo de vida (figura 2). Se ha observado de forma poco frecuente infecciones ectópicas de *F. hepatica*, donde el parásito se localiza en otros órganos del hospedero como el corazón, vasos sanguíneos, pulmón, cavidad pleural, cerebro, páncreas, etc (2,15,18). En humanos, la maduración desde metacercaria a un tremátodo adulto toma entre 3 a 4 meses luego de la ingesta de metacercarias (23).



**Figura 2: Ciclo de vida de *Fasciola hepatica* (23).** Reproducida desde <http://www.dpd.cdc.gov/dpdx>

### **I.1.3 Organización genómica**

La *F. hepatica* es un tremátodo di-genético que presenta un alto potencial zoonótico ya que tiene alta habilidad para infectar un amplio rango de mamíferos como los vacunos, camélidos, ovinos, caprinos y humanos; además, este parásito tiene la capacidad de adaptarse de forma rápida a otros hospederos como canguros y roedores (24-25). Su alta infectividad está correlacionada con la expresión de biomoléculas del parásito durante distintos estadios de desarrollo que atraviesa (16,24-28).

*F. hepatica* tiene un genoma de 1.3 Gbp con 10 pares de cromosomas, 32% de secuencias de ADN repetitivas, y además retrotransposones (16,24,25). Por otro lado, el parásito presenta un elevado grado de polimorfismo genético lo que explica el alto potencial evolutivo y la alta adaptabilidad en un amplio rango de hospederos intermediarios y definitivos (16,25,28,29). La expresión de los genes varía según la etapa de infección en el que se encuentre el tremátodo en el hospedero definitivo o intermediario; esta característica puede aprovecharse para buscar drogas que interfieren con el ciclo de vida del parásito (16,30).

La *F. hepatica* tiene la capacidad de auto-fertilización, una forma de endogamia que influencia en la frecuencia de alelos en una población; esta característica tiene una ventaja en expandir los alelos de resistencia en caso los rasgos sean recesivos, pero ésta puede tener desventajas debido a que la heterocigosidad y diversidad genética dentro del parásito individual disminuiría. Contrariamente, el amplio rango de hospederos del tremátodo permite al parásito tener mayor flujo genético y mantener la diversidad genética (25,31).

### **I.1.4 Proteómica**

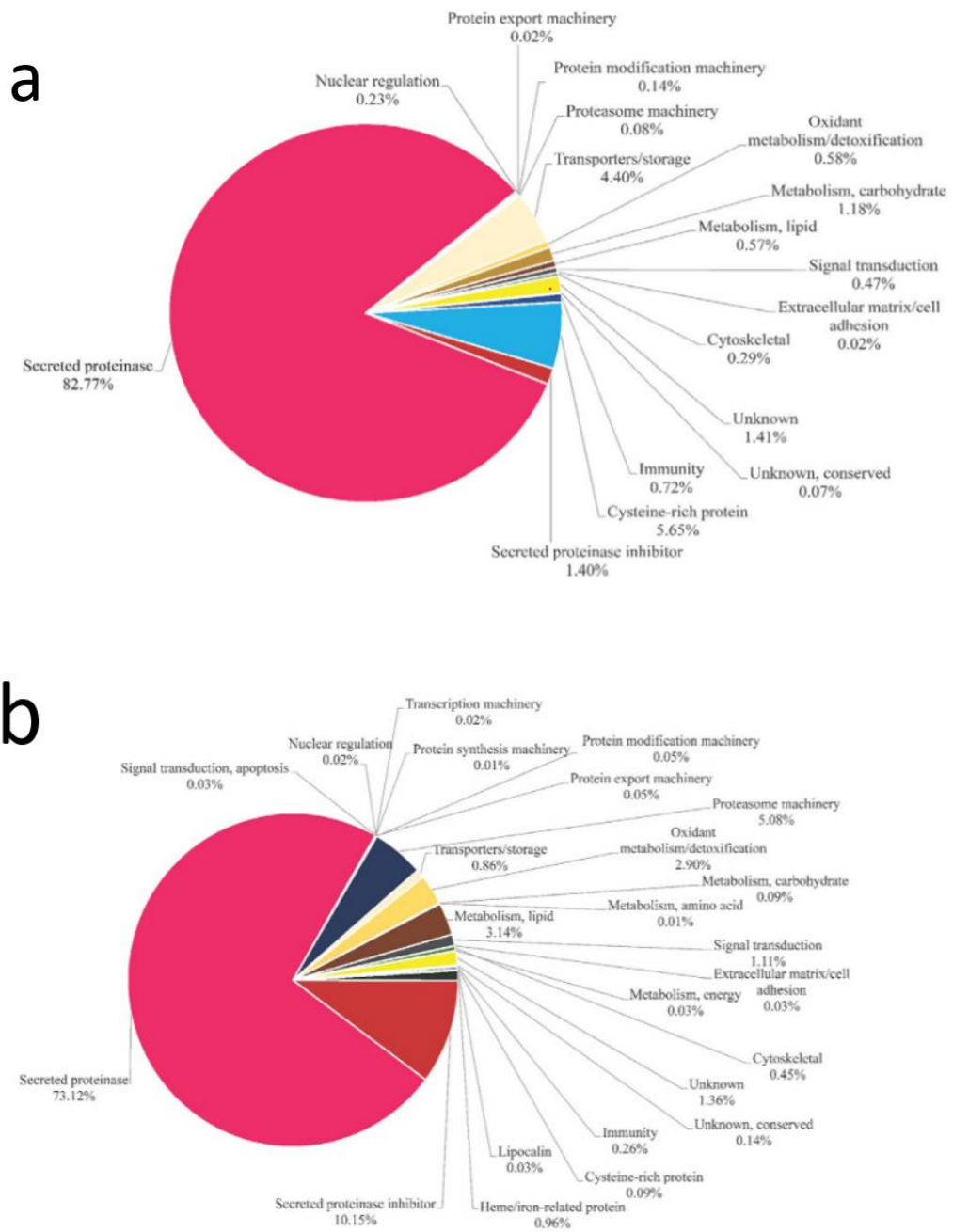
Se reporta que *F. hepatica* contiene 22,676 genes codificantes de proteínas (16) que se expresan específicamente durante la etapa de infección dentro de su hospedero ya sea intermediario o definitivo. Se han identificado enzimas (en mayor porcentaje) y otras proteínas para la evasión de la respuesta inmune del hospedador (16,27,30,32).

Se ha realizado el análisis proteómico de los estadios de *F. hepatica* durante la infección a su hospedero definitivo, se comparó el conjunto de las proteínas expresadas en tremátodos adultos, juvenil y proteínas solubles somáticas (27,32). En el análisis de expresión relativa, se encontró que la proteína más abundante en juvenil y adultos son las proteinasas, representando el 83 y 73% del total de proteínas, respectivamente (figura 3 a,b). En el caso de proteínas solubles somáticas como catepsinas L y B representaron un 51% del contenido total de proteínas (25-27,32,33), seguido de legumainas (20%), metaloproteinasas (17%), serina proteinasas (8%), y calpaína (2.8%) (figura 4a).

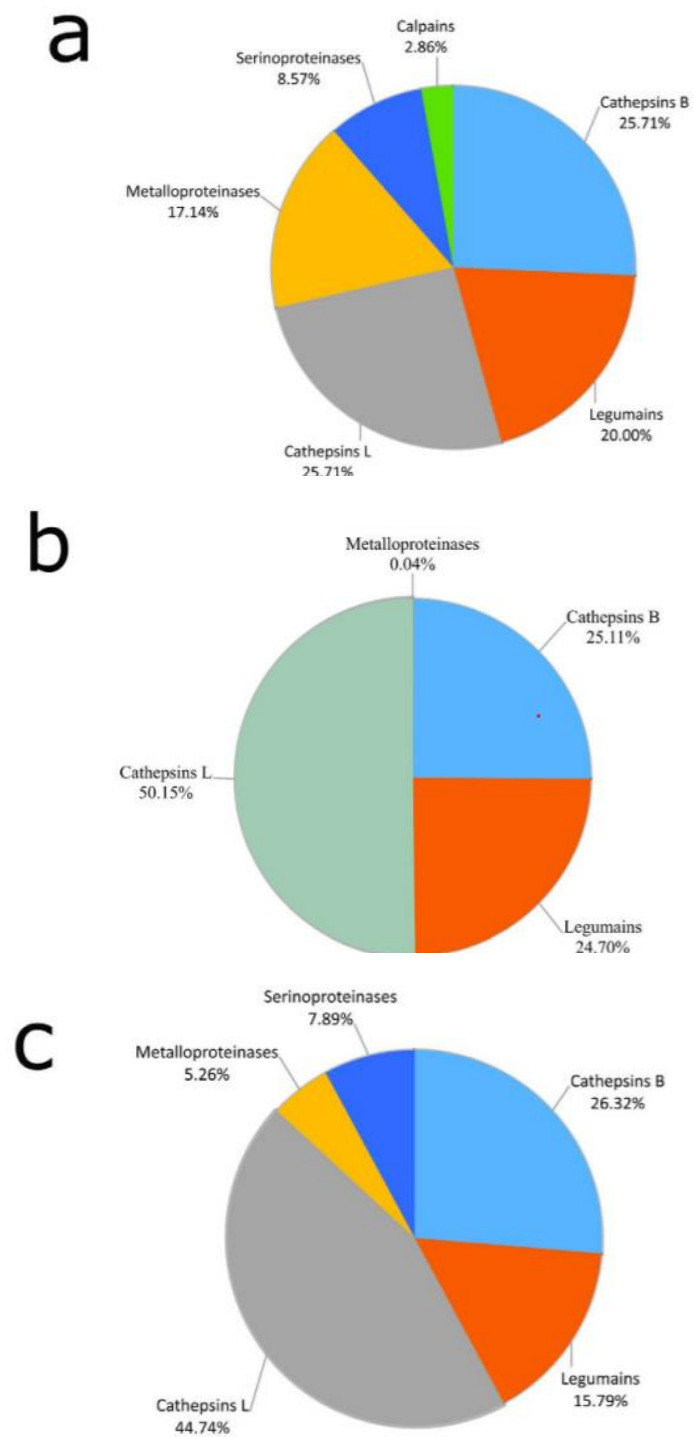
En la expresión proteica desde el estadio juvenil, la familia de catepsina L fue el grupo más representado dentro de las cisteínas proteinasas, comprendiendo un 50 % del total, seguido por catepsina B (25%) y legumainas (24%) (figura 4b) (32).

Para la expresión proteica del tremátodo adulto, la clase más abundante de proteinasas fue incluso relacionado a la familia de cisteína proteinasa, con un 44% de catepsinas L, seguido por catepsina B (26%) y legumainas (15%) (figura 4c).

Debido a que las catepsinas son las proteinasas secretadas en mayor proporción durante la infección, se sugiere que estas presentan roles importantes para la invasión y supervivencia del parásito dentro de su hospedero (25-27,32,33).



**Figura 3:** Clasificación funcional de proteinasas derivados de *F. hepatica* identificados (a) productos de juveniles (b) y de adultos (32).



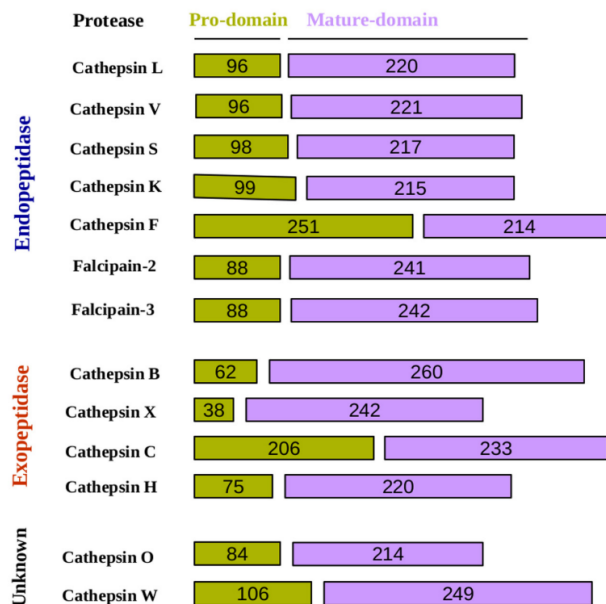
**Figura 4:** Clasificación funcional de proteinasas; (a) proteínas solubles somático, (b) proteínas de juveniles y (c) de adultos (32).

#### **I.1.4.1 Cisteína proteinasas**

Las cisteínas proteinasas son enzimas que se encuentran distribuidas desde procariotas hasta eucariotas superiores. Estas proteínas son endopeptidasas pertenecientes al clan CA y se caracterizan por tener en el sitio activo los residuos de Cis, His y Asn (figura 8) (34-39). En *F. hepatica*, las cisteína proteinasas están involucradas en la hidrólisis de hemoglobina, invasión de glóbulos rojos, salida del parásito, procesamiento de proteínas de superficie, evasión del sistema inmune, etc. Estas moléculas son blancos terapéuticos para el desarrollo de nuevos agentes para el tratamiento de la fascioliasis (16,39-46).

Las familias de la cisteína proteinasas se clasifican por la homología de secuencias y la especificidad por el sustrato. Las familias más abundantes de cisteína proteasas son la papaína y las catepsinas (34-37,43,47). El grupo principal de cisteína proteinasas en los mamíferos y en los humanos son las catepsinas, que presentan 11 isoformas (catepsinas B, C, F, H, K, L, O, S, V, X y W) (35,48) y que cumplen diferentes funciones celulares. En *F. hepatica*, la catepsina L (CatLFh) y B son las cisteína proteinasas expresadas en mayor proporción durante la infección y juegan un rol clave en la supervivencia del tremátodo en el hospedero definitivo (figura 4) (32,37,43,48). En efecto, *F. hepatica* expresa 7 isoformas de catepsinas L: CatL1Fh, CatL2Fh, CatL3Fh, CatL4Fh, CatL5Fh, CatL6Fh y CatL7Fh (37,43,49).

Las cisteínas proteinasas son expresadas como precursores inactivos (zimógeno) para prevenir la autodegradación y la degradación de proteínas de forma no controlada, característica que comparten otras enzimas proteolíticas. Los zimógenos de las catepsinas de *F. hepatica* contienen un pro-dominio que bloquea la entrada del sustrato al sitio activo, direcciona la proteína para su ensamblaje y media la localización intracelular (16,32,37,39,49-51). La interacción entre el pro-dominio y el dominio maduro se da a través de los residuos de F<sup>56</sup>p, F<sup>63</sup>p, F<sup>71</sup>p, Y<sup>146</sup>p y Y<sup>151</sup>p donde p representa los residuos del pro-dominio y los aminoácidos mencionados al dominio maduro (37,52). El pro-dominio contribuye, incluso, a la clasificación de las isoformas de la catepsina en tres grupos: Exopeptidasas, endopeptidasas y un grupo no conocido (figura 5).



**Figura 5:** Representación esquemática del pro-dominio y el dominio maduro de las cisteínas proteinasas (37).

#### I.1.4.1.1 Proceso de activación de las cisteínas proteasas

La activación del zimógeno involucra la escisión del pro-dominio. El proceso puede darse por diferentes vías tales como autoactivación, transactivación o ambos a la vez (37).

##### I.1.4.1.1.1 Auto activación

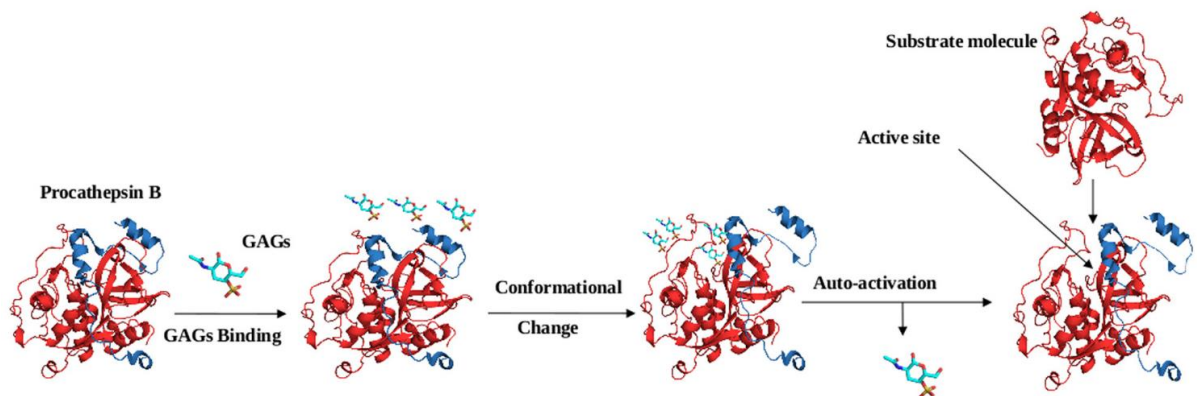
En el caso de catepsina L, el procesamiento inicia con la ruptura de puentes salinos conservados, donde la presencia de cisteína en la posición 25 es esencial (16). Los puentes salinos se rompen debido a la protonación del grupo carboxilo; este paso desencadena la ruptura del núcleo hidrofóbico dirigiendo a la disociación del pro-dominio del sitio activo e iniciándose entonces el proceso de autoactivación (37,54).

La auto-activación de catepsinas es una combinación de procesos uni y bimoleculares (54). El proceso unimolecular involucra la disociación del pro-dominio del sitio activo; el proceso bimolecular es el paso siguiente en el cual la misma catepsina activada procesa y activa las otras catepsinas a través de uno o varios pasos de catálisis (autoactivación) (37,54). Este proceso depende del pH, generalmente ocurre en medio ácido (37,53).

Los glicosaminoglicanos (GAG) tales como GAG sulfatados, heparina, sulfato de heparano, sulfatos de condroitina A, B, C, E y otros polisacáridos cargados negativamente modulan la activación de cisteína proteasas, principalmente de catepsinas L y B, incluso a pH neutro (37,38,55). Para la activación de catepsina B, GAG se une a residuos cargados positivamente en el bucle de eclosión entre el dominio maduro y el pro-dominio; la interacción resulta en un cambio conformacional al pro-dominio, el cual expone el sitio activo para la catálisis de otras pro-catepsinas B (figura 6) (37,55).

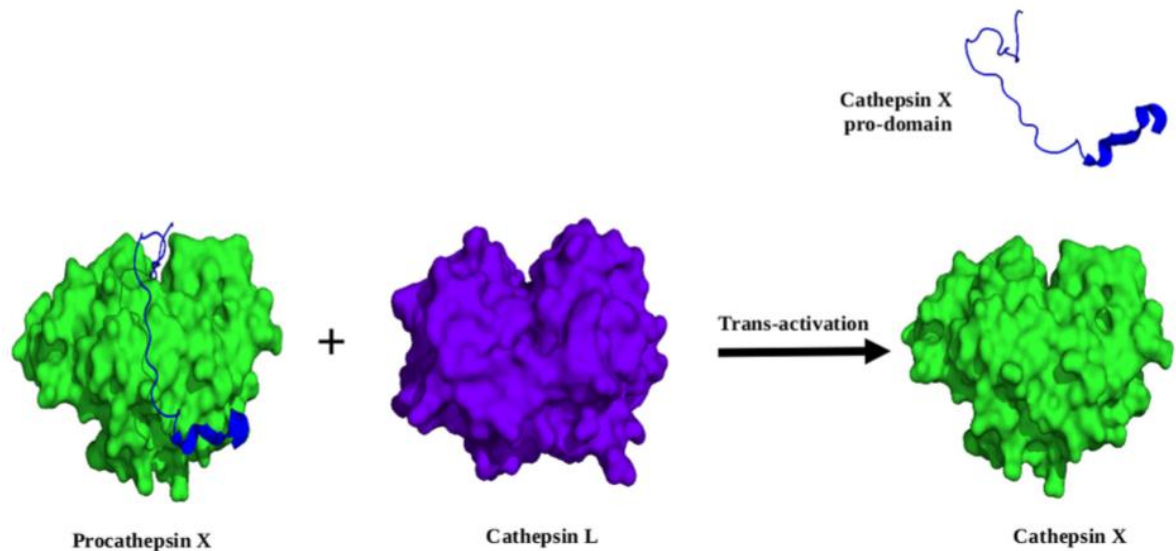
#### I.1.4.1.1.2 Trans-activación

Trans-activación involucra la escisión por otras moléculas de la misma enzima u otras proteasas que escinden específicamente los residuos de unión entre el pro-dominio y el sitio activo. Por ejemplo, las catepsinas C y X, requieren de endopeptidasas como catepsinas L o S para su activación; catepsina S requiere de la degradación de su pro-dominio por la catepsina L para su activación (37,56) (figura 7).



**Figura 6:** Mecanismo de autoactivación de catepsina B en presencia de GAG. Unión de GAG induce un cambio conformacional de la pro-catepsina desenmascarando el sitio activo y permitiendo la entrada de sustratos (37).

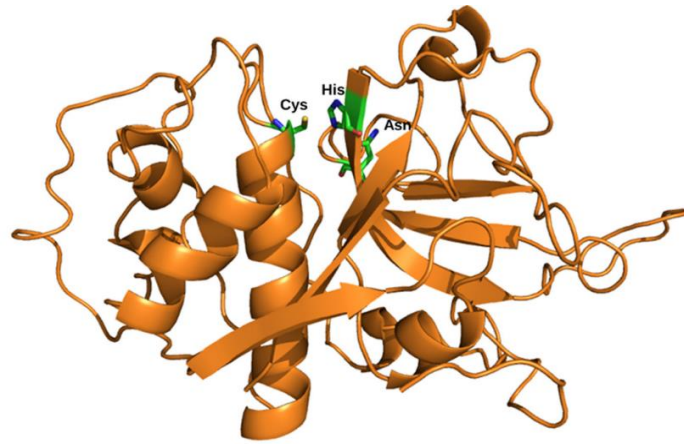




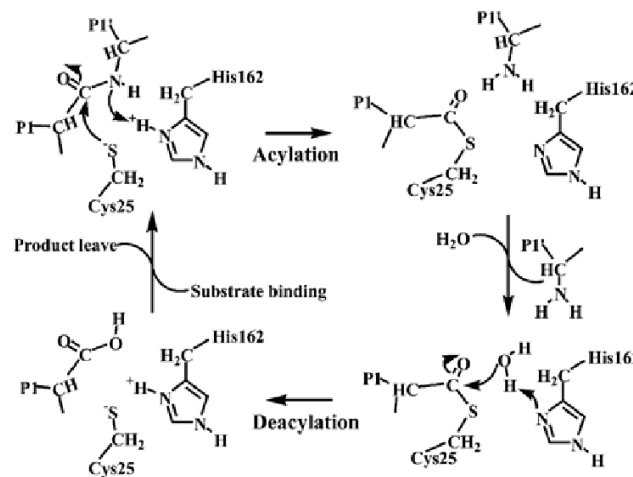
**Figura 7:** Transactivación mediado por otra molécula. Pro-catepsina X consiste de un pro-dominio (azul) unido covalentemente al sitio activo; Catepsina L trans-activa pro-catepsina X por el clivaje del pro-dominio (37).

#### I.1.4.1.2 Mecanismo de catálisis

Las cisteínas proteinasas contienen en su sitio activo los residuos de Cys25 - His162 - Asn182 (figura 8). La histidina162 actúa como donador de protón y potencia la nucleofilicidad del residuo cisteína25 (16,37,50,57,58). El residuo tiol de cisteína25 nucleofílico ataca al carbono del enlace peptídico reactivo, produciendo el primer intermediario tioéster tetraédrico con el lanzamiento de un fragmento amina o amino terminal del sustrato (figura 9) (16,37,57,58). El intermediario es estabilizado mediante enlaces de hidrógeno entre el oxianión del sustrato y un residuo de glutamina altamente conservado. Seguidamente, el enlace tioéster es hidrolizado para producir un ácido carboxílico desde el fragmento de sustrato restante (figura 9) (16,37,57,58). En el caso de la asparagina, este es crucial en el proceso catalítico ya que ayuda a orientar al residuo de Cis25 y además convierte a esta misma en un mejor aceptor de electrones a través de los efectos electrostáticos (16,37,57,58).



**Figura 8:** Representación estructural de la cisteína proteinasas. Dominio maduro (naranja) presentando sus residuos del sitio activo: Cys25, His162 y Asn182 (verde) importantes para la catálisis de su sustrato (37).



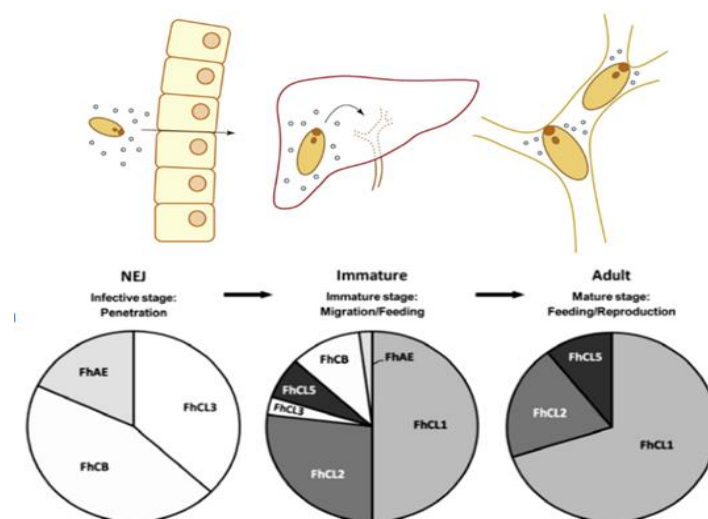
**Figura 9:** Mecanismo catalítico de hidrólisis peptídico catalizado por cisteína proteinasa (16).

#### I.1.4.2 Catepsinas

*F. hepatica* secreta catepsinas tipo L y B, las cuales son importantes para el éxito en la infección y supervivencia del parásito. CatL3Fh es considerada como la más antigua por análisis de su secuencia (37,59). CatL3Fh, por duplicación, generó CatL2Fh; de forma similar, la duplicación subsecuente de CatL2Fh dio origen a CatL1Fh, y luego esta última por el mismo mecanismo generó CatL5Fh (37,59).

Las isoformas de catepsinas B y L son secretadas diferencialmente durante el ciclo de vida del parásito donde estimulan la respuesta humoral de anticuerpos en el hospedero mamífero infectado, siendo considerado como antígenos inmunodominantes en la fascioliasis humana. La catepsina B y L3 son expresadas en el estadio de metacercaria, la catepsina B, L1, L2, L3 son expresadas en el estadio juvenil y catepsinas L1, L2 y L5 son secretadas en mayor proporción en la etapa adulta (figura 10) (15,25,26,27,59-62).

Las catepsinas L pertenecen a un grupo monofilético dentro del linaje enzimático que da origen a las catepsinas L, S y K de mamíferos; por ende, se considera que las proteasas de la *F. hepatica* son idénticas, en un 40-55%, a las catepsinas Ls de los vertebrados (63-66).

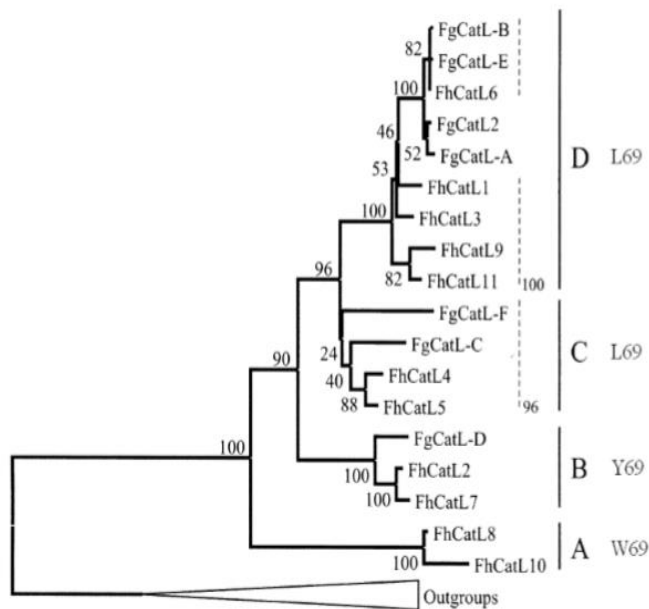


**Figura 10:** Perfil de expresión de las catepsinas B e isoformas de L en cada etapa durante la infección de la *F. hepatica* a su hospedero definitivo. NEJ (JND); FhCL (CatLFh); FhCB (CatBFh) (Modificado de 27,61,64)

#### I.1.4.2.1 Catepsina tipo L

La caracterización de varias isoformas de catepsinas indica que difieren en la especificidad por el sustrato de acuerdo al residuo de aminoácido presente en su centro activo (49); a partir de ello, las familias de las catepsinas se dividen en 4 clados (figura 11) en la cual cada uno presenta una afinidad específica para un sustrato determinado (67,68).

Las catepsinas Ls son blancos para el desarrollo de drogas anti *Fasciola* porque son esenciales para lograr la infección a su hospedero y para el ciclo de vida del parásito; de acuerdo a su importancia, estas proteínas están involucrados en: potenciar la producción de huevos en la *F. hepatica* adulta (cuando estas proteasas se encuentran en los órganos reproductivos del tremátodo) (63,67,69,70); digerir un determinado rango de componentes del hospedero para la alimentación y migración (67,71); tienen la capacidad de degradar inmunoglobulinas del hospedero para contrarrestar citotoxicidad celular dependiente de los anticuerpos (ADCC) (72); supresión de la proliferación de linfocitos T; reducción de la expresión de CD4 y células T humanas y, modula la inmunidad a través de la supresión de la respuesta inmune tipo Th1 (65,73,74). Estas funciones evidencian que las catepsinas Ls juegan roles importantes en la evasión del sistema inmune y en la alimentación del tremátodo dentro del hospedero; además, aumenta la virulencia del parásito mejorando la entrada, migración, y permanencia en el hospedero (45,71,73,69 y 74).

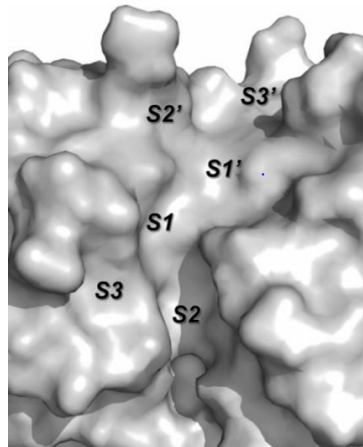


**Figura 11.** Familias de catepsinas L de *F. hepatica*. Árbol filogenético de catepsinas Ls. Las proteasas son divididas hacia 4 distintos clados, donde el residuo del aminoácido en la posición 69 es conservado en cada clado (68).

Con respecto a las funciones específicas de las catepsinas L, se caracterizó que tanto CatL1Fh como CatL2Fh pueden degradar matriz extracelular y moléculas de membrana basal incluyendo fibronectina, laminina, hemoglobina y colágeno (65,67,73,75). Así mismo, la catepsina L2 tiene la capacidad de degradar fibrinógeno, para producir un coágulo de fibrina, previniendo el excesivo sangrado en los puntos de alimentación en el ducto biliar; alternativamente, también puede producir aglutinación previniendo el acceso de células inmunológicas a la superficie del parásito o al sitio de infección (16,67,69,73,75-77).

#### **I.1.4.2.2 Interacción de catepsinas L con su sustrato**

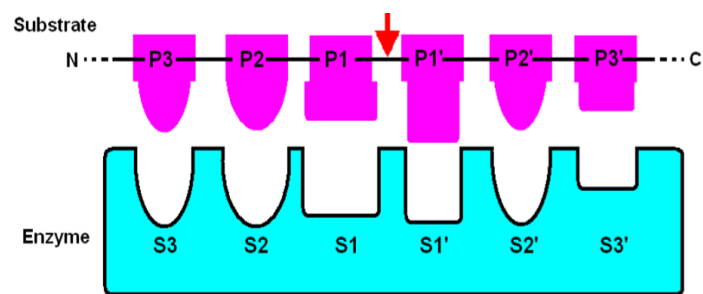
Las cisteínas proteinasas presentan sub-sitios en su sitio activo que están formados por residuos que le brindan un arreglo estructural característico y especificidad para un sustrato determinado (16,67,71,78 y 79). En la figura 12, se presenta la localización de estos sub-sitios en la ranura del sitio activo de la catepsina L, donde S1', S2' y S3' están direccionados al C-terminal, mientras que S1, S2 y S3 están cercanos al N-terminal (16,51,71,78,79).



**Figura 12:** Localización del sub-sitio en el sitio activo de la catepsina L (79).

Los residuos del sustrato que interactúan con cada sub-sitio son conocidos como P1', P2', P3', P1, P2 y P3 (116,51,71,78,79). El mecanismo de interacción se muestra en la figura 13, donde la flecha roja indica el sitio de corte del sustrato; de acuerdo a ello, el sub-sitio inmediatamente N-terminal al sitio de corte es S1, mientras que el sub-sitio inmediatamente C-terminal al sitio de corte es S1' (16,51,68,71,78-80).

Las catepsinas Ls tienen residuos pequeños, extensos y ramificados en la posición P1', residuos básicos e hidrofóbicos en la posición P3. La interacción eficaz enzima-sustrato depende de una buena interacción entre el sub-sitio S2 y P2 (interacción S2-P2) (51,67,71,78-80). En efecto, las variantes de aminoácidos presentes en S2 van a determinar la especificidad de cada isoforma de catepsina L sobre su sustrato (45), teniendo en cuenta que S2 comúnmente muestra una preferencia por aminoácidos hidrófobos alifáticos o aromáticos comúnmente leucina o fenilalanina (16,51,68,71,78-80).



**Figura 13:** Nomenclatura de la interacción del sustrato con el sitio activo de la enzima (67,80).

De acuerdo a lo mencionado anteriormente, se caracterizó siete posiciones de los residuos de aminoácidos dentro de la región S2 importantes para determinar la especificidad (49). Estos aminoácidos son ubicados en la posición 67, 68, 133, 157, 158, 160 y 205 (tabla 1). En consecuencia, ellos determinan el tipo de sustrato a reconocer por las distintas isoformas de catepsina L (16,51,68,71,78-80). A partir de ello, se encontró que CatL1Fh con Leu67, Val157 y Leu205 tiene una preferencia fuerte por aceptar Leucina hidrofóbica en la posición P2, mientras que CatL2Fh con Tyr67, Leu157 y Leu205 también puede aceptar Leucina, pero con una alta preferencia por Prolina (57,63,64,71-73). En conclusión, los aminoácidos ubicados en la posición 67, 157 y 205 son mucho más importantes para la determinación de especificidad de sustrato de las distintas isoformas de catepsinas L.

Tabla 1: Ubicación de los aminoácidos en el sub-sitio S2 de las catepsinas L.

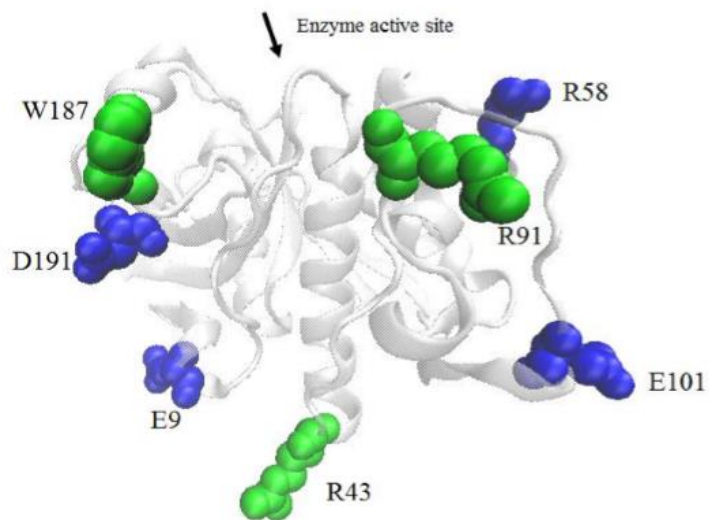
	67	68	133	Residues			
				157	158	160	205
Human cathepsin L	Leu	Met	Ala	Met	Asp	Gly	Ala
Human cathepsin K	Tyr	Met	Ala	Leu	Asn	Ala	Leu
Adult: FhCL 1A	Leu	Met	Ala	Val	Asn	Ala	Leu
FhCL 1B	Leu	Met	Ala	Leu	Asn	Ala	Leu
FhCL 2	Tyr	Met	Ala	Leu	Thr	Ala	Leu
FhCL 5	Leu	Met	Ala	Leu	Asn	Gly	Leu
Juvenile: FhCL 3	Trp	Met	Ala	Val	Thr	Ala	Val
FhCL 4	Phe	Met	Ala	Leu	Asn	Ala	Phe

Comparación de residuos del sitio activo S2 que contribuye a unión diferenciado de sustratos en catepsina L de *F. hepática* y catepsinas L y K de humanos. Los residuos del S2 coloreados en rojo son importantes porque confieren especificidad para su sustrato (71).

#### I.1.4.2.1.1 Catepsina L5 (CatL5Fh)

La catepsina L5 (CatL5Fh; figura 15) es una proteinasa secretada por *F. hepatica* durante la infección del hospedero definitivo, los niveles de esta enzima en el estadio juvenil son menores que la etapa adulta (figura 10) (16,26,27,32,49,61,64,67,68,71,81). El gen que codifica la proteína consta con un fragmento de lectura abierta de 981 nucleótidos que codifica al zimógeno de peso molecular 37.1 kDa, la enzima madura de 24.4 kDa (49). CatL5Fh tiene una gran similitud de secuencia con CatL4Fh, con 97% de identidad a nivel de aminoácidos; también presenta una similitud debajo del 80% con CatL2Fh y CatL7Fh (16,49,67,71).

CatL5Fh consta de Leu67, Leu157 y Leu205 en el sub-sitio S2 lo cual le permite tener la capacidad alta de digerir hemoglobina, inmunoglobulina, laminina, fibronectina y colágeno (16, 64, 67, 71). La proteinasa se caracteriza por presentar preferencia única para el ácido aspártico y leucina en la posición P2 del sustrato (16,64,67,71,82). Catepsina L5 comparte un rango de sustratos con otras isoformas de catepsinas L, especialmente con L1 y L2; sin embargo, el perfil de preferencia de sustrato sugiere que la habilidad de acomodar en P2 a Leucina y Ácido aspártico es único de la CatL5Fh (16, 64, 67, 71, 82-85).



**Figura 14:** Estructura 3D de la Catepsina L5 desde la *Fasciola hepatica* resaltando los residuos que poseen altas fluctuaciones. Los residuos coloreados en verde presentan una mayor flexibilidad, incluso mayor que la catepsina L1 (16).

#### I.1.4.2.1.2 Importancia del estudio

CatL5Fh es una proteinasa secretada en la etapa adulta de la *F. hepatica* que cumple un rol en la supervivencia del tremátodo en el hospedero ya que participa en la evasión del sistema inmune y en la alimentación del helminto. La Catepsina L5 se caracteriza por degradar el sustrato ST3 (sustrato codificado en el laboratorio de Unidad de Biotecnología Molecular – Universidad Peruana Cayetano Heredia) la cual le convierte en un potente agente biomédico.

Para continuar con el desarrollo de CatL5Fh como una innovación biomédica, se requiere contar con altas cantidades de CatL5Fh nativa; sin embargo, esta proteinasa es recuperada de forma limitada a partir de los parásitos adultos con alta inversión de tiempo y dinero; por ende, es necesario utilizar la tecnología de ADN recombinante para lograr un mayor rendimiento de la enzima activa y continuar el desarrollo de su potencial aplicación biomédica. A partir de ello, el objetivo del presente trabajo es expresar, purificar y determinar la actividad enzimática de la proteína recombinante catepsina L5 (rCatL5Fh) en un sistema de expresión procariota. Se elige sistema de expresión procariota debido a que anteriormente un grupo de investigadores lograron expresar una cisteína proteinasa en un sistema de *E. coli* (112) por lo cual se plantea la hipótesis de



que algunas cisteínas proteasas (como CatL5Fh) probablemente no necesitan de modificaciones post-traduccionales para presentar actividad enzimática.

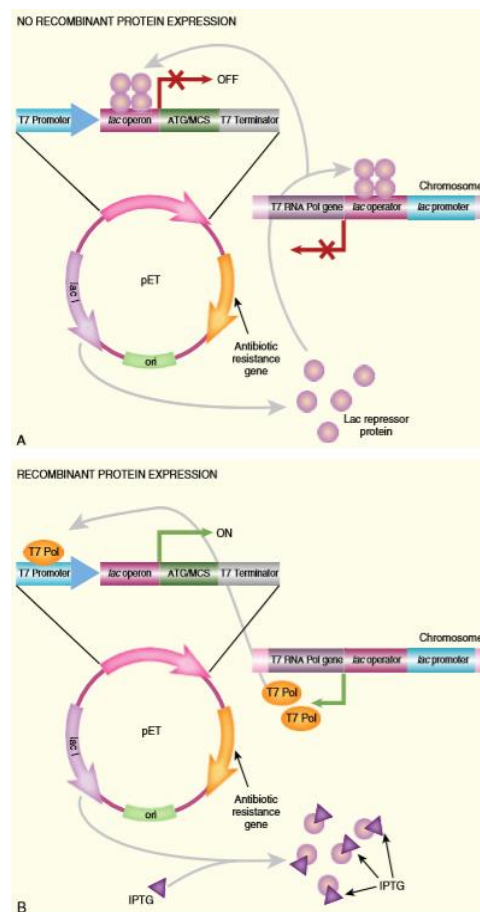
## **I.2 Proteína recombinante expresada en *E. coli***

*E. coli* es una bacteria ampliamente utilizada para la producción de proteínas recombinantes. Las ventajas de este sistema de expresión son: i) el amplio rango de vectores con que se cuenta, ii) relativamente fácil control de la expresión del gen, iii) tasas altas de rendimiento de la producción de proteína recombinante, iv) la producción puede ser diseñada para la secreción hacia el periplasma de la célula (86,87). En este sistema se puede clonar tanto genes procariotas como eucariotas (86-89).

El objetivo de optar por la tecnología de ADN recombinante es para producir grandes cantidades de la proteína de interés; sin embargo, la sobreproducción de la proteína recombinante puede ser dañina para la célula hospedera que forma cuerpos de inclusión como mecanismo de defensa por la toxicidad generada por la sobreexpresión de la proteína (86). Los cuerpos de inclusión son agregados de proteínas mal plegadas y no funcionales. En ese sentido, para disminuir la formación de cuerpos de inclusión, se usan vectores de expresión en los que se pueda controlar cuándo y cuánta proteína va a producir el hospedero; a partir de ello, se conoce que los vectores pET y pBAD tienen mecanismos para cambiar de un estado encendido a apagado de la producción de la proteína, incluso, pBAD puede modular la cantidad de proteína producida (91-92).

Los vectores de expresión (plásmidos) tienen tres características principales: secuencia ori (origen de auto replicación de ADN); gen de resistencia a antibióticos (el cual permite la selección del hospedero transformado) y una región polylinker (sitio múltiple de clonación) en donde se inserta la secuencia de interés (86,90). Adicionalmente, los plásmidos tienen distintas secuencias que le permiten controlar la expresión del gen de interés; como, por ejemplo, pET tiene un promotor híbrido T7/lac (upstream al gen de interés) y un terminador T7 (downstream al gen de interés) (86). La transcripción desde el promotor híbrido requiere del ARN polimerasa T7 codificada en el genoma de *E. coli*; tanto el ARN polimerasa como el promotor híbrido tienen una región denominada operador lac al cual normalmente (altas concentraciones de glucosa) se une el tetrámero de la proteína represora del operón lac para inhibir la transcripción (figura 15A) (86).

La activación de la transcripción depende de la presencia del inductor galactosa. La galactosa sirve tanto como inductor y como recurso de carbono; esta última característica hace que disminuya la capacidad de la galactosa para inducir por lo que es muy inestable de su concentración; a partir de ello, para una inducción eficaz se usa el análogo de la galactosa, Isopropil- $\beta$ -D-1-tiogalactopiranosido (IPTG). Esta molécula es estable ya que no es usado como recurso de carbono y solamente sirve para inducir la expresión (91). IPTG al igual que la galactosa, se une a la proteína represora lac inhibiendo la formación del tetrámero y en consecuencia inhibiendo la unión al operador lac (figura 15B). De acuerdo a ello, la síntesis de la proteína recombinante es controlada por la concentración de IPTG y por el tiempo de incubación post-inducción (86,91).



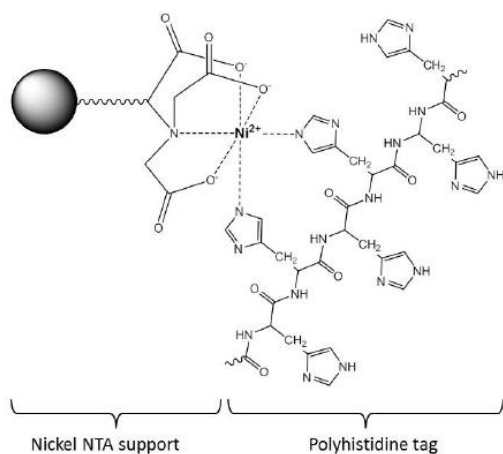
**Figura 15:** Sistema de expresión de proteína pET. (A) La proteína recombinante es clonada en la inserto polylinker (MCS) pero no se expresa en las células no inducidas. (B) El IPTG es agregado, este se une al represor *lac* y abandona la unión al operador y desde ambos promotores permitiendo la expresión de la ARN polimerasa T7 y la expresión de la proteína recombinante simultáneamente (86).

Para optimizar la detección y purificación de la proteína recombinante, a la secuencia clonada se le adiciona en el extremo 5' o 3' una secuencia que codifica polipéptidos de fusión (tag). Los tags comúnmente usados son glutatión S transferasa (GST) y poli-histidina (6xHis); estos péptidos de fusión permiten obtener la proteína recombinante con alta pureza ya que se pueden unir y disociar con facilidad de sus ligandos inmovilizados en una matriz en fase estacionaria (matriz de glutatión y partículas de Níquel, respectivamente). Finalmente, los tags pueden detectarse usando inmunoensayo; en caso de 6xHis tag, no es necesario removerlo de la proteína recombinante ya que es poco inmunogénico y no altera significativamente la estructura de la proteína porque es muy pequeño (86-92).

### **I.3 Purificación de proteína recombinante**

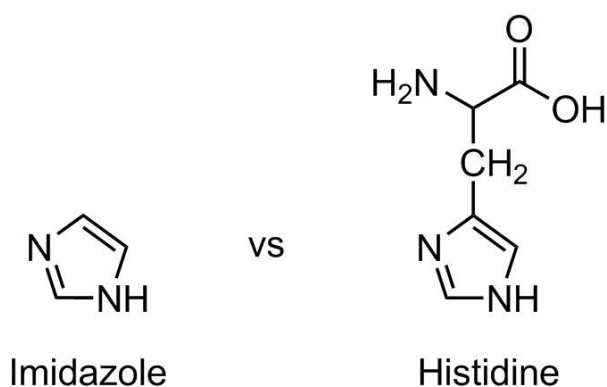
La proteína recombinante puede contar de un tag (marcaje) en los extremos N o C terminal. En muchas ocasiones se usa la Poli-Histidina como tag donde el número de residuos puede variar desde 4 a 10. His-tag permite la purificación de la recombinante mediante cromatografía de afinidad con Ni u otro metal inmovilizado en agarosa (IMAC) (94-96).

IMAC es una técnica que inmoviliza iones metálicos de transición en una matriz de resina usando un agente quelante tales como ácido iminodiacético o ácido nitrilotriacético (NTA). Los iones más comunes para la purificación de proteína con Poli-His son  $\text{Ni}^{+2}$ ,  $\text{Co}^{+2}$ ,  $\text{Cu}^{+2}$  y  $\text{Zn}^{+2}$ ; el primero es el más utilizado porque tiene mayores rendimientos de purificación (94). La figura 16 muestra esquemáticamente el ion Níquel inmovilizado y la interacción de la misma con el nitrógeno del anillo del residuo de histidina de la proteína recombinante (93-96). La interacción es favorecida a pH 8 porque el nitrógeno y el ion Ni se encuentran desprotonados (94). La presente gráfica ilustra la interacción de la proteína marcada; entonces, si a la columna se le agrega el lisado crudo de un cultivo de células, las probabilidades de interacción con el ion son más altas en la proteína recombinante porque la unión es más fuerte ya que tiene varios residuos de histidinas (6x o 10x), mientras que las otras proteínas pueden ser eluidas durante el lavado incluso aquellos que tienen residuos de histidina en su estructura (93-96).



**Figura 16.** Representación esquemática de la inmovilización del ion Níquel ( $\text{Ni}^{+2}$ ) en una matriz NTA; además, se representa la interacción del ion con dos de los residuos del Tag de Poli-Histidina (94).

Se agrega una solución de imidazol para eluir la proteína recombinante unida a la resina (Figura 17) (pKa 6) (94). El imidazol forma parte de la histidina y compite en la interacción con el Ni cargado de la matriz. La cola de poli-His de la enzima recombinante tiene una alta afinidad por el ion Ni por lo que se requiere una alta concentración de imidazol para eluir la proteína inmovilizada en la columna. Si la alta concentración de imidazol no es suficiente para eluir la proteína, bajar el pH favorece el proceso ya que el ion metálico puede protonarse y desestabilizar la interacción electrostática con el grupo imidazol (93-96).



**Figura 17.** Representación estructural de Imidazol y la Histidina.

## **CAPÍTULO II**

### **OBJETIVO**

#### **II.1 Objetivo general**

- Expresar, purificar y determinar actividad enzimática de la proteína recombinante catepsina L5 de la *Fasciola hepatica* (CatL5Fh) obtenida en un sistema de expresión procariota con el vector pET-26b(+) en las cepas BL21 y Rosetta de *Escherichia coli*.

#### **II.2 Objetivos específicos**

- Determinar la secuencia de CatL5Fh clonado en el plásmido pET-26b(+).
- Expresar la proteína recombinante CatL5Fh en *E. coli* BL21 con distintas condiciones de temperatura y concentración de inductor.
- Purificar la proteína recombinante CatL5Fh de *E. coli* mediante la utilización de distintos métodos que favorecen el plegamiento de la enzima.
- Analizar la actividad enzimática de la proteína recombinante CatL5Fh y compararla con la nativa.

## CAPÍTULO III

### MATERIALES Y MÉTODOS

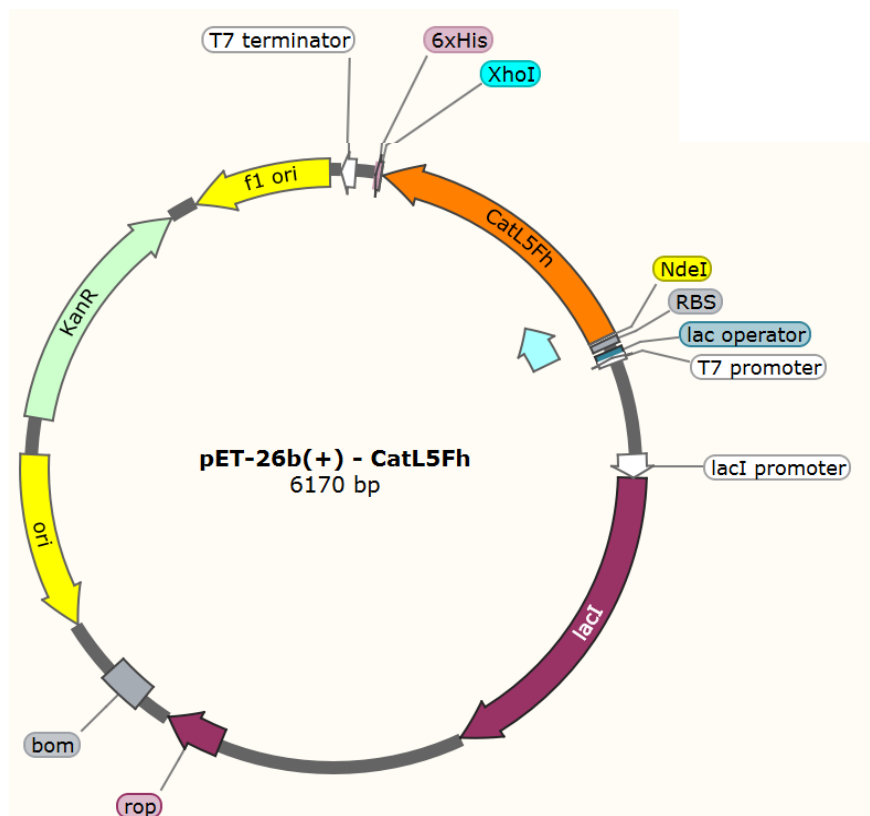
#### III.1 Clonamiento de la secuencia CatL5Fh hacia el vector pET-26b(+) y transformación del plásmido a *E. coli* BL21.

Es importante mencionar que en la Unidad de Biotecnología Molecular (UBM), previo a este trabajo, se realizó el clonamiento del inserto rCatL5Fh (981 bp) hacia el vector pET-26b(+) (figura 18); a la vez, también se transformó el vector en *E. coli* BL21. Brevemente, la preparación de BL21 quimio-competente fue iniciado por inoculación de una colonia de las células en medio SOB (super optimal broth). El cultivo fue incubado toda la noche a 37 °C a 250 rpm. Cumplido el tiempo, 1 mL fue inoculado en un matraz con 100 mL de medio SOB fresco y fue incubado bajo las mismas condiciones de cultivo hasta obtener un OD<sub>600</sub> de 0.6. Luego, el matraz fue incubado en hielo por 10 minutos y las células fueron alicuotados en tubos 50 mL frío para centrifugar a 2500 x g por 10 min a 4 °C. El sobrenadante fue descartado y el pellet se resuspendió en 40 mL de Buffer de transformación química bacteriana frío. La suspensión fue centrifugada y el pellet nuevamente fue resuspendido en 5 mL de Buffer de transformación química bacteriana frío e incubado por 5 min. La suspensión fue mezclada suavemente con 187 µL de DMSO (dimetilsulfóxido) altamente puro e incubado por 5 min en hielo. Finalmente, 100 µL de la suspensión celular fue alicuotado en tubos estériles fríos y almacenados a -80 °C.

Para la ligación, previamente tanto el vector como la secuencia de pro-catepsina L5 de *F. hepatica* (de forma separada) fueron digeridos con 8 U de NdeI y 2 U de XhoI por 1 h a 37 °C. El plásmido doblemente digerido fue tratado con 1 U de fosfatasa alcalina por incubación de 1 h a 37 °C para remover grupo fosfato 5' y minimizar autoligación. Finalmente, la ligación se llevó a cabo incubando por 16 h y 16 °C 50 ng de pET-26b(+), 50 ng de CatL5Fh y 2000 U de ligasa T4. La figura 18 ilustra la ligación correspondiente.

Finalmente, se realizó la transformación de BL21. En primer lugar, la cepa fue descongelado lentamente en hielo por 10 min. La suspensión celular fue mezclada

suavemente por agitación y se agregó 4  $\mu\text{L}$  de ligación. La mezcla se colocó en hielo por 30 min. Luego, las células fueron sometidos a choque térmico por 30 segundos a 42  $^{\circ}\text{C}$  e inmediatamente colocadas en hielo por 2 min. Posteriormente, se agregó 250  $\mu\text{L}$  de medio SOB y se incubó a 37  $^{\circ}\text{C}$  por 1h a 220 rpm. Las células fueron plaqueadas en agar LB con Kanamicina (LB/Kan) e incubadas a 37  $^{\circ}\text{C}$  por toda la noche. Al siguiente día, una colonia de células fue inoculada en medio LB/Kan e incubadas a 37  $^{\circ}\text{C}$  por toda la noche a 250 rpm. El inóculo luego fue alicuotado y criopreservado a -80  $^{\circ}\text{C}$  con glicerol al 15%.



**Figura 18:** pET-26b(+)-CatL5Fh. 6.1 Kpb, Kan<sup>R</sup>. XhoI y NdeI fueron usados para la clonación.

### III.2 Purificación del plásmido pET-26b(+)-CatL5Fh y corroboración de la integridad de la secuencia.

Inicialmente, las clonas transformadas fueron plaqueadas por agotamiento en placas LB/Kan (30  $\mu\text{g}/\text{mL}$ ) e incubadas toda la noche a 37  $^{\circ}\text{C}$ . Cumplido el tiempo, se realizó el pre-inóculo de una colonia en 10 mL de medio LB/Kan (30  $\mu\text{g}/\text{mL}$ ) incubando toda

la noche a 37°C y a 250 rpm. Luego, se purificó el plásmido utilizando el kit de DNA Wisard Plus SV Minipreps. Brevemente, se centrifugó el cultivo en tubos de 1.5 mL por 5 minutos a 13000 rpm y se descartó el sobrenadante. Se re-suspendió el pellet con 250 µL de buffer de re-suspensión, se le añadió 250 µL de buffer de lisis y se mezcló mediante inversión. Seguidamente, se agregó 10 µL de solución de proteasa alcalina y se incubó por 5 minutos a temperatura ambiente. Cumplido el tiempo, se paró la reacción agregando 350 µL de solución de neutralización. Posteriormente, se centrifugó la muestra a 13000 rpm por 10 minutos a temperatura ambiente y el sobrenadante fue transferido a la columna previamente colocada en el tubo colector. Se centrifugó la muestra por 1 minutos a 1300 rpm. El lavado se realizó dos veces utilizando 750 y 250 µL de solución de lavado y centrifugando a 1300 rpm por 1 minuto. Finalmente, se realizó la elución del plásmido con 100 µL de agua de PCR y mediante centrifugación a 1300 rpm por 3 minutos. Se cuantificó el ADN extraído por Nanodrop.

Una vez purificada el plásmido, con la colaboración de la empresa MACROGEN se secuenció el inserto utilizando los primers Forward (5'-CGGTCCCATATGTCAAATGACGATTTGTG-3') y Reverse (5'-CGTTCTCGAGCGGAAATTGTGCCACCATC-3'). Se corrigió la integridad de la secuencia comparando el producto Forward y Reverse usando el programa DNA baser assembler; luego, se realizó la traducción de las secuencias nucleotídicas usando ExpAsy translate tool con diferentes Open Reading Frame. Finalmente, se alineó las secuencias de aminoácidos del wild type (GenBank AF271385) con la recombinante utilizando el programa ClustalW-Tools > Pairwise Sequence Alignment (alineamiento múltiple) y Protein BLAST.

### **III.3 Expresión de la proteína recombinante catepsina L5 (rCatL5Fh).**

Para la expresión de la proteína recombinante, inicialmente la cepa BL21 recombinante fue de-crio preservada mediante sembrío en placas LB/Kan (30 µg/mL) con incubación a 37 °C toda la noche. Luego, una colonia fue cultivada en 3 mL de LB/Kan y se incubó 16 horas a 37 °C con agitación constante de 250 rpm. Cumplido el tiempo, del pre-inóculo se extrajo 200 µL y se sembró en 100 mL de medio LB/Kan con incubación a 37 °C, 250 rpm hasta obtener un OD<sub>600nm</sub> entre 0,4 - 0,6 para la inducción correspondiente con IPTG.



### **III.3.1 Inducción**

La expresión de la proteína recombinante fue inducida a diferentes concentraciones de IPTG (0.1, 0.25 y 1 mM) y a distintas temperaturas (30 y 37 °C), con el objetivo de determinar la condición que forme cuerpos de inclusión en menor proporción. Las inducciones se incubaron a distintos tiempos (2, 4, 6, 8 y 16 horas) con agitación constante de 250 rpm. Cumplido cada tiempo de inducción, se cosecharon las células de *E. coli* mediante centrifugación a 13000 rpm por 3 minutos. Se conservaron las muestras a -20 °C hasta la lisis celular.

### **III.4 Lisis**

Las células cosechadas fueron re-suspendidas en Buffer de lisis (50mM NaH<sub>2</sub>PO<sub>4</sub> y 5mM de Tris-HCl pH=8.0); con posterior adición de lisozima (1mg/mL) e incubación de 1 hora en frío. Cumplido el tiempo, las células fueron colocadas a -20°C para garantizar la lisis.

Para romper las células, las muestras fueron sonicadas tres veces con 30 pulsos de un segundo cada uno (un pulso de un segundo y un segundo de descanso) con un minuto de descanso para cada 30 pulsos, hasta que la muestra se torne fluida y clara. Finalmente, se centrifugó el sonicado a 13000 rpm por 30 min a 4 °C y se guardó tanto el pellet como el sobrenadante a -20 °C. Se analizó la expresión mediante SDS-PAGE al 12%.

### **III.5 Purificación de rCatL5Fh**

#### **III.5.1 Purificación a condiciones desnaturizantes (pellet sonicado).**

El pellet del sonicado inicialmente fue re-suspendido en Buffer B (100mM NaH<sub>2</sub>PO<sub>4</sub>, 10 mM de Tris-HCl y 8 M de Urea, pH=8.0) (extracto A). La purificación de la proteína recombinante fue realizada por cromatografía de afinidad utilizando matriz de níquel la cual presenta afinidad por la histidina. Para ello, 2.0 mL de matrix +Ni-NTA (al 50%) fue equilibrada dos veces con 2 volúmenes de Buffer B. Luego, la matriz fue mezclada con 4 mL del extracto A e incubada a 4°C por 16 horas en agitación constante. Cumplido el tiempo, se cargó la muestra en una columna cromatográfica y lavada con Buffer C (100mM NaH<sub>2</sub>PO<sub>4</sub>, 10 mM de Tris-HCl y 8 M de Urea, pH=6.3), Buffer D (100mM NaH<sub>2</sub>PO<sub>4</sub>, 10 mM de Tris-HCl y 8 M de Urea, pH=5.9) y Buffer E (100mM NaH<sub>2</sub>PO<sub>4</sub>, 10 mM de Tris-HCl y 8 M de Urea, pH=4.5); en todos los lavados se colectaron

fracciones de 1 mL para su análisis de SDS-PAGE. Por último, para disociar la proteína recombinante de la matriz, se agregó Buffer E + imidazol a una concentración final de 200 mM. Se colectó las eluciones en fracciones de 1 mL y se mantuvo en hielo. La purificación fue analizada mediante SDS-PAGE al 12%. Finalmente, se hizo un pool de todas las fracciones con la proteína purificada y se realizó el re-plegamiento de la proteína mediante diálisis en dilución gradual de urea (6 a 1 M) en 20 mM de Tris-HCl y 5 mM de buffer EDTA pH 8. Para finalizar el re plegamiento, la urea fue eliminada mediante diálisis usando PBS pH 7.4.

### **III.5.2 Purificación sin agentes caotrópicos (sobrenadante sonificado).**

Inicialmente se preparó la matriz níquel (ProBond) pipeteando 2 mL de resina hacia una columna de purificación; luego se permitió que la resina asiente completamente por gravedad mediante incubación por 10 minutos. Se aspiró el sobrenadante y se agregó 6 mL de agua estéril, se mezcló por inversión y se permitió que la resina asiente por gravedad mediante incubación de 10 minutos. Se quitó el sobrenadante y se adicionó 6 mL de Buffer de unión (50 mM  $\text{NaH}_2\text{PO}_4$ , 0.5 M NaCl, 10 mM imidazol, pH 8); se re-suspendió la resina suavemente por inversión y se permitió que la resina asiente por incubación de 10 min. Se eliminó el sobrenadante y se repitió una vez más el procedimiento con Buffer de unión.

De forma seguida, se agregó 8 mL de sobrenadante sonificado a la columna de níquel preparado y se permitió la interacción de la muestra con la matriz incubando en agitación constante por 1 hora. Cumplido el tiempo, se permitió asentar a la resina por gravedad y cuidadosamente se aspiró el sobrenadante. Luego, la columna fue lavada 5 veces con 8 mL de Buffer de lavado (50 mM  $\text{NaH}_2\text{PO}_4$ , 0.5 M NaCl, 20 mM imidazol, pH 8). Terminado este proceso, se colocó la columna verticalmente y se eluyó la proteína con 12 mL de Buffer de elución (50 mM  $\text{NaH}_2\text{PO}_4$ , 0.5 M NaCl, 250 mM imidazol, pH 8). Finalmente, se colectó las fracciones de eluciones (1 mL aproximadamente) para el análisis de purificación mediante SDS-PAGE al 12%.

### **III.5.3 Purificación a condiciones híbridas**

Para este sistema de purificación, se realizó la lisis celular a condiciones desnaturizantes con buffer conteniendo guanidina pH 7.8, en incubación de 10 minutos a temperatura ambiente. Para asegurar la lisis, se realizó la sonicación en hielo con tres pulsos de 5 segundos a alta intensidad. Finalmente, las células se centrifugaron a 3000 x g por 15 minutos y se pasó a purificar la recombinante a partir del sobrenadante.

En primer lugar, se preparó la columna cargando 2 mL de resina hacia la columna de purificación. Se permitió que la resina asiente por gravedad durante 10 minutos. Se aspiró el sobrenadante y se agregó 6 mL de agua destilada y se mezcló la resina por inversión; luego, se permitió que la resina asiente para seguidamente aspirar el sobrenadante. Se equilibró la resina con 6 mL de Buffer de unión desnaturizante (20 mM NaH<sub>2</sub>PO<sub>4</sub>, 0.5 M NaCl, 8 M urea pH 7.8) y se mezcló por inversión; posteriormente se aspiró el sobrenadante y se repitió una vez más el proceso.

Una vez preparado la resina, la purificación se inició agregando 8 mL del lisado celular a la columna. Se permitió la unión de la muestra y la matriz de níquel a temperatura ambiente por 30 minutos y agitación constante. Cumplido el tiempo, se permitió que la resina asiente por gravedad y luego el sobrenadante fue aspirado. Posteriormente, se lavó la columna con 4 mL de Buffer de unión desnaturizante (20 mM NaH<sub>2</sub>PO<sub>4</sub>, 0.5 M NaCl, 8 M urea, pH 7.8) con incubación de 2 minutos. Se aspiró el sobrenadante y se repitió el proceso. Lavada la columna, se dejó que la resina asiente para aspirar el sobrenadante y luego agregar 8 mL de buffer de lavado nativo (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 0.5 M NaCl, 20 mM imidazol, pH 8); se mezcló por inversión y la resina fue permitida que asiente para eliminar el sobrenadante. Posteriormente se colocó la columna de forma vertical y se agregó 12 mL de buffer de elución nativo (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 0.5 M NaCl, 250 mM imidazol, pH 8); se colectó fracciones de 1 mL para analizar la purificación mediante SDS-PAGE al 12 %.

### III.5.4 Purificación y re-ensamblaje de rCatL5Fh mediante cromatografía de filtración utilizando gel de Sephadex G-25.

Se agregó 3 mL de rCatL5Fh purificada a condiciones desnaturizantes a una columna con 8.3 mL de sephadex G-25 pre equilibrado con Buffer de re-ensamblaje (5 mM CaCl<sub>2</sub>, 0.02% SDS, 100 mM Tris HCl, pH 8). En este tipo de purificación, se elimina la urea y se permite el re-ensamblaje de la enzima a su conformación nativa de acuerdo a los componentes del Buffer. Analizar mediante SDS-PAGE al 12% la purificación.

### III.6 Cuantificación

La cuantificación de la proteína fue realizada mediante el método de Bradford. Se emplearon 7 estándares de BSA, un blanco y las muestras. Se preparó inicialmente los estándares según la tabla 2.

El reactivo de Bradford (Biorad ®) fue diluido 1/4 con agua Milli Q. Luego, se mezcló tanto los estándares como las muestras problemas con el reactivo según la tabla 3. Se realizó la mezcla en microplacas de ELISA, y se generó la ecuación lineal en base a la relación concentración de proteína y absorbancia. Finalmente, se determinó la concentración de las muestras a partir de la absorbancia y utilizando la ecuación lineal.

Tabla 2. Preparación de concentraciones finales del estándar BSA.

Tubo	Volumen de estándar (µL)	Fuente de estándar	Volumen diluyente (µL)	Concentración final de proteína (µg/mL)
1	20	stock 2 mg/mL	0	2000
2	30	stock 2 mg/mL	10	1500
3	20	stock 2 mg/mL	20	1000
4	20	tubo 2	20	750
5	20	tubo 3	20	500
6	20	tubo 4	20	250
7	20	tubo 5	20	125

Tabla 3: Resumen de la mezcla 1X Dye Reagent con el estándar y muestra.

	<b>Volumen</b>	<b>Volumen de 1X Dye Reagent</b>	<b>Volumen total de reacción</b>
<b>Estándar del tubo 1 al 7</b>	3 uL	147 uL	150 uL
<b>Muestra</b>	3 uL	147 uL	150 uL

### **III.7 Análisis de SDS-PAGE**

La calidad de pureza de la proteína recombinante catepsina L5 fue analizada mediante SDS-PAGE al 12% a condiciones desnaturizantes.

### **III.8 Activación de la proteína recombinante Catepsina L5**

5 µg del zimógeno rCatL5Fh fue activado utilizando el buffer conteniendo 50 mM de EDTA, 100 mM de DTT, Buffer citrato 0.1 M a pH 5.5, para un volumen final de 50 µL. También se probó la activación con 2 ug de CatL5Fh nativa y se analizó el proceso mediante SDS-PAGE al 15%.

### **III.9 Análisis y comparación de actividad enzimática de la proteína recombinante catepsina L5 con la nativa.**

Se determinó la actividad enzimática de la rCatL5Fh a pH 5.5 con azocoll como sustrato.

En un tubo eppendorf de 1.5 mL se pesó 5 mg de azocoll. Luego se agregó 700 uL de Buffer citrato 0.2 M pH 5.5. Posteriormente, se agregó 50 uL de DTT 100 mM y finalmente se adicionó 50 uL de la muestra pre activada. Se mezcló suavemente y el tubo fue incubado 37 °C por 16 horas hasta la liberación de un color rojizo. Cumplido el tiempo, se centrifugó a 14000 rpm por 3 minutos (o 6500 rpm por 5 minutos) y se extrajo el sobrenadante para la medición de la absorbancia a 520 nm.

### **III.10 Modelamiento estructural de rCatL5Fh.**

Se determinó la estructura tridimensional de rCatL5Fh mediante modelamiento por homología. Para ello se utilizó el servidor online SWISS MODEL y se tuvo como estructura molde a la proteína cristalizada catepsina L1 de *F. hepatica* (2O6X). La resolución estructural de esta proteína es de 1.4 Å y el porcentaje de identidad con rCatL5Fh es de 92.48%; por ende, se consideran homólogos y la estructura predicha estaría bien resuelta. Para tener la estructura óptima, se realizó la minimización de energía usando el programa VMD (Visual Molecular Dynamics). Dicho programa trabaja en conjunto con MAMD (Nanoscale Molecular Dynamics) en la cual usa como campos de fuerza el compatible a CHARM++; además, presenta un modelo de agua tipo TIP4. La minimización de energía fue realizada calculando el centro de masa de rCatL5Fh, teniendo las fuerzas iónicas de NaCl 0.154 M, a temperatura de 37°C y por 500 nanosegundos.

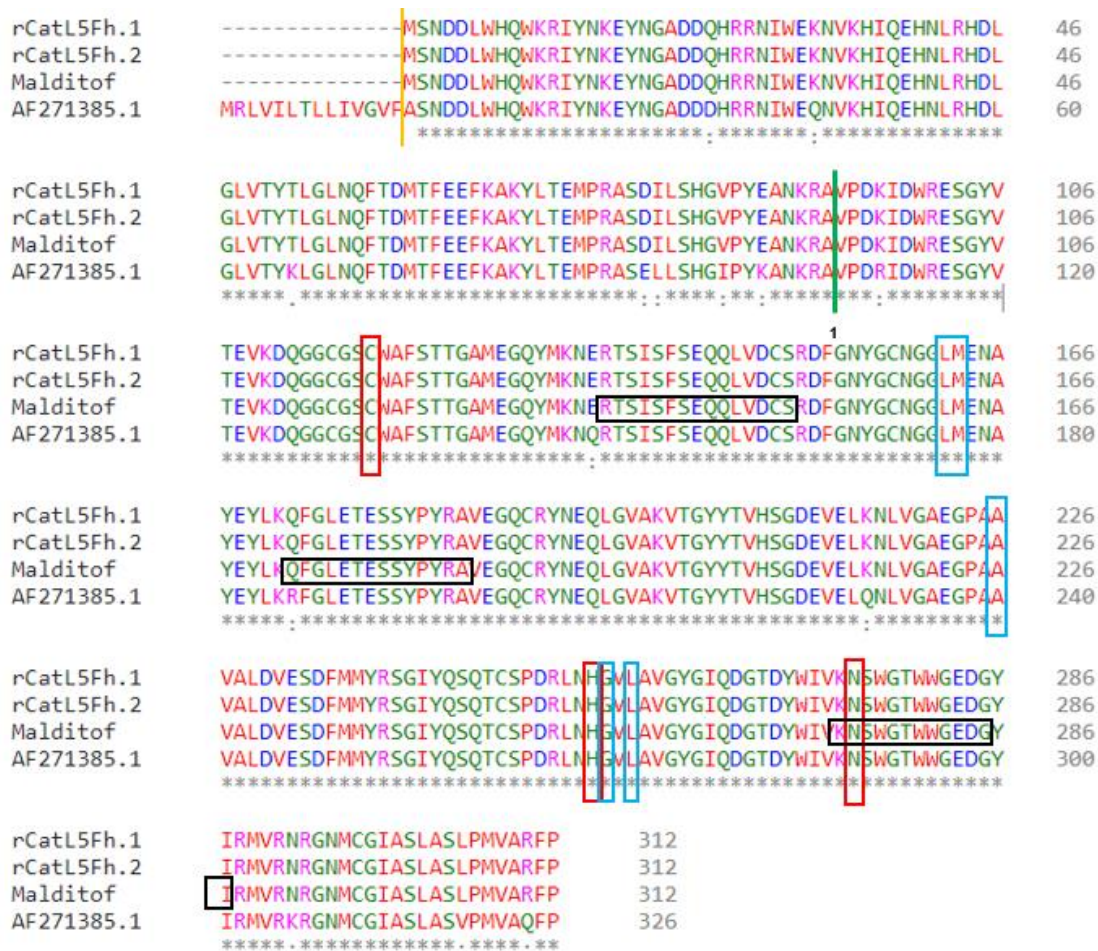
La estructura de rCatL5Fh con la mínima energía potencial fue visualizada con el programa VMD y además fue alineada con la estructura construida a partir de la secuencia reportada por los australianos. El alineamiento estructural se realizó en VMD y dicho proceso fue conllevado referenciando el carbono alfa como punto de comparación. VMD usa el programa STAMP (Structural alignment of multiple proteins) para el alineamiento estructural. Finalmente, usando el panel de visualización de VMD, se muestran las estructuras alineadas. Si en caso existiera una variación en los aminoácidos entre ambas secuencias, se selecciona el residuo específico y se visualiza el efecto sobre la estructura según la posición en el espacio tridimensional.

## CAPÍTULO IV

### RESULTADOS

En primer lugar, la secuencia de nucleótidos de la proteína Catepsina L5 de la *F. hepatica* fue clonada en el vector de expresión pET-26b(+) (Figura 18); y transformado en *E. coli* BL21. Para continuar con la expresión de la proteína clonada, se analizó la secuencia del inserto; para ello, se purificó el plásmido y a partir de esta, se secuenció el inserto en la empresa MACROGEN. La concentración de plásmido purificada de las colonias 1 y 2 fue de 112.9 y 139.7 ng/uL, respectivamente. La secuencia de nucleótido del inserto fue traducida con diferentes Open Reading Frame y se alineó los residuos de la recombinante con la secuencia correspondiente en el GenBank (AF271385). El análisis de alineamiento mediante Protein BLAST mostró un porcentaje de identidad de 95.5% y una cobertura de 99%, lo que indica posibles mutaciones. El alineamiento de las secuencias con ClustalW mostró 14 mutaciones en las secuencias respecto a AF271385, los residuos que variaron y sus posiciones fueron las siguientes: R4K, Q43E, R79Q, Q123K, K199N, V212L, Q217R y otras 7 mutaciones ubicadas en la región del pro-dominio (Figura 19).

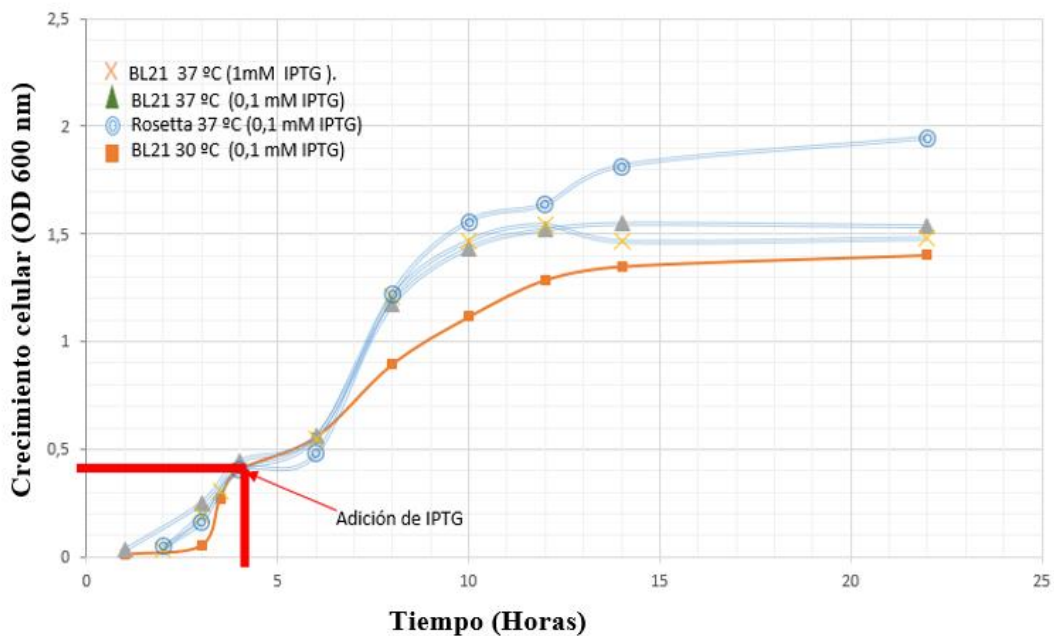
Debido a que el alineamiento de secuencias muestra que las mutaciones ocurridas no presentan efecto significativo sobre la estructura de la proteína, se continuó con la expresión de la recombinante mediante la inducción con distintas concentraciones de IPTG (0.1, 0.25 y 1 mM) y a distintas temperaturas (30 y 37 °C). Previo a la expresión, el cultivo de *E. coli* fue monitoreado a 600 nm para determinar la curva de crecimiento en las distintas temperaturas indicadas y a la vez especificar el tiempo de inducción durante la fase exponencial; siendo como referencia que el OD para la inducción es alrededor de 0.4-0.6. En la figura 20 se observa que entre 4 a 5 horas el cultivo de *E. coli* alcanza el OD indicado. Antes de la inducción, todos los cultivos fueron incubados a 37 °C con agitación orbital a 250 rpm; luego, post-inducción, cada uno es incubado a la temperatura correspondiente según la condición de expresión de rCat15Fh asignada.



**Figura 19.** Alineamiento de secuencia de rCatL5Fh con la nativa. Recuadro rojo, aminoácidos importantes del sitio activo; recuadro celeste, residuos importantes para el arreglo del sub-sitio; recuadro negro, péptidos determinados mediante espectrofotometría de masas (MaldiTof, comunicación personal J. Espinoza); línea verde, primer residuo del dominio maduro y línea naranja, aminoácidos entre pre y pro dominio. Segmente anterior al residuo 1 es el pro-dominio.

La inducción con distintas concentraciones de IPTG a 37 °C no muestra variación en la tasa de crecimiento, lo que indica que las distintas concentraciones del inductor no tienen efecto sobre el crecimiento celular; sin embargo, el cultivo a 30 °C si disminuye la tasa de crecimiento bacteriano. Por otro lado, se evidencia que la cepa de Rosetta presenta una mayor tasa de crecimiento a 37°C.

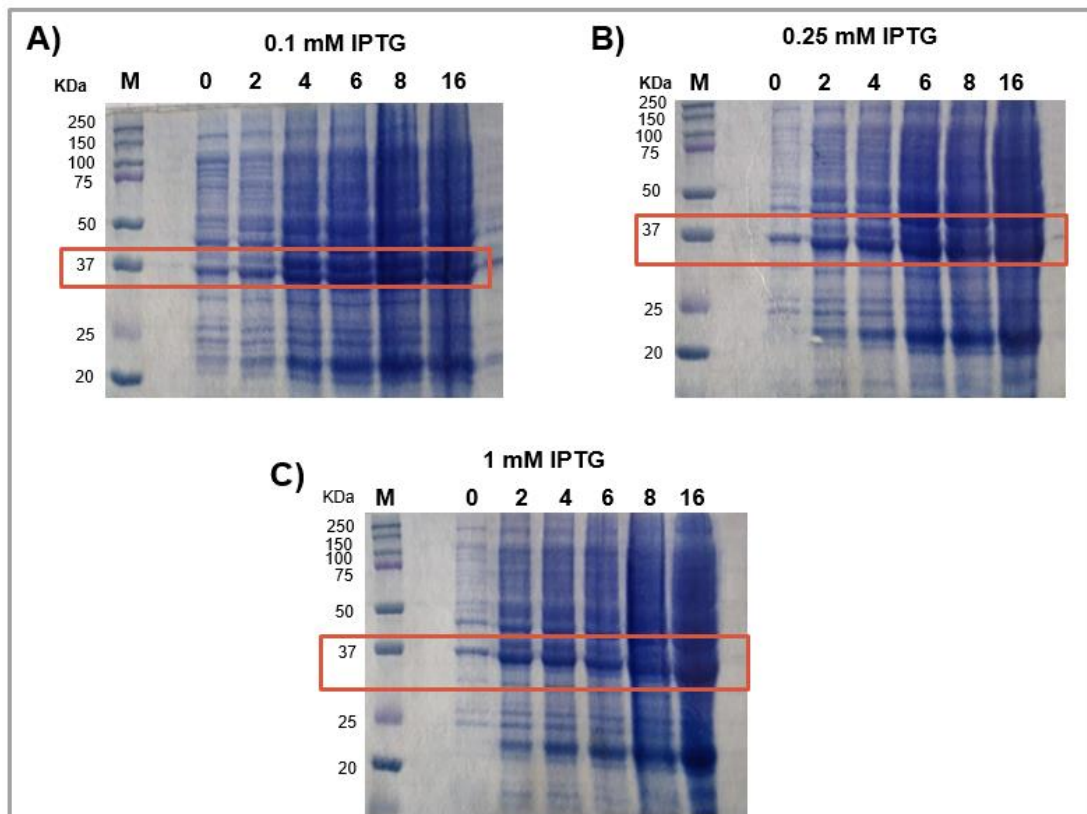




**Figura 20:** Curva de crecimiento de *E. coli* BL21 y Rosetta. Crecimiento celular a 37°C, 250 rpm hasta la adición de IPTG. Línea roja indica el tiempo de inducción (adición de IPTG). Post infección, las cepas son incubadas a las temperaturas distintas como indica la leyenda.

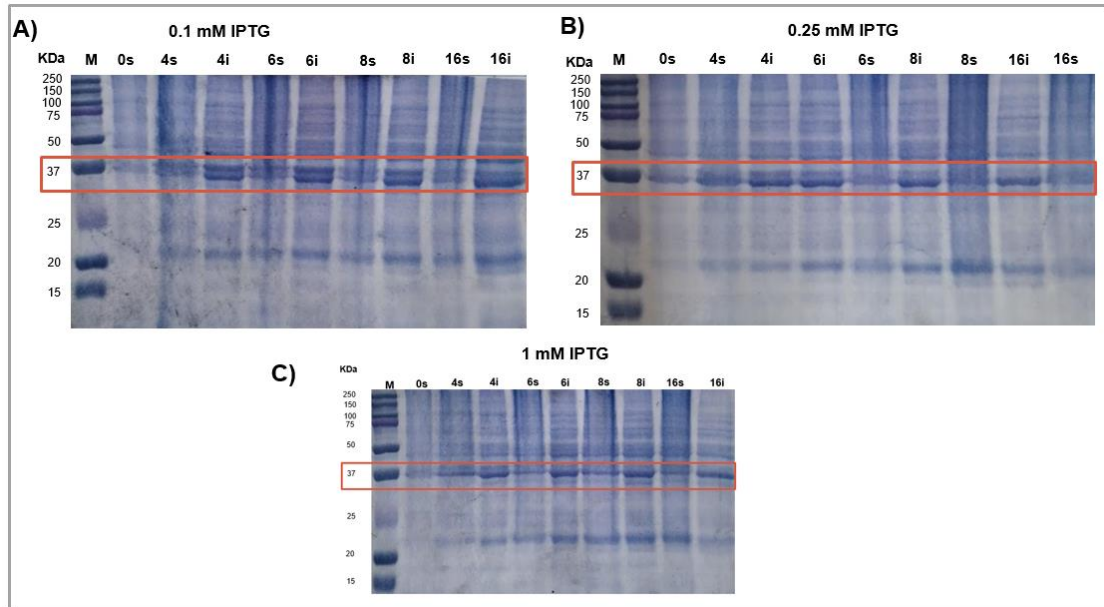
#### IV.1 Expresión de la proteína recombinante

La expresión de rCatL5Fh se realizó en dos temperaturas (30 °C y 37 °C), con diferentes concentraciones de IPTG (0.1, 0.25 y 1 mM) y en distintos tiempos de incubación post-inducción. En primer lugar, se analizó la expresión a 37 °C, la temperatura en la cual BL21 presenta una tasa más alta de crecimiento. Para ello, aproximadamente a las 4 horas y media de crecimiento (pre-inducción) se le agregó el inductor y la expresión de la proteína fue evaluada a distintos tiempos de incubación (2, 4, 6, 8 y 16 horas post infección) mediante SDS-PAGE al 12% tanto de la fracción soluble como insoluble (figura 21 y 22). La figura 21 muestra la expresión de rCatL5Fh en la parte insoluble; esta es mostrada en las bandas que incrementan en intensidad a nivel de 37 kDa en los distintos tiempos post inducción, además las bandas son más intensas cuando las concentraciones de IPTG son altas.



**Figura 21.** Expresión de rCatL5Fh en las muestras insolubles del cultivo a 37°C de BL21. El carril M corresponde al marcador Precision plus protein Dual Color estándar; los carriles 0, 2, 4, 6, 8 y 16 referencian las horas de incubación post-inducción con IPTG. Las muestras fueron inducidas con **A)** 0.1, **B)** 0.25 y **C)** 1 mM de IPTG. SDS PAGE al 12 %.

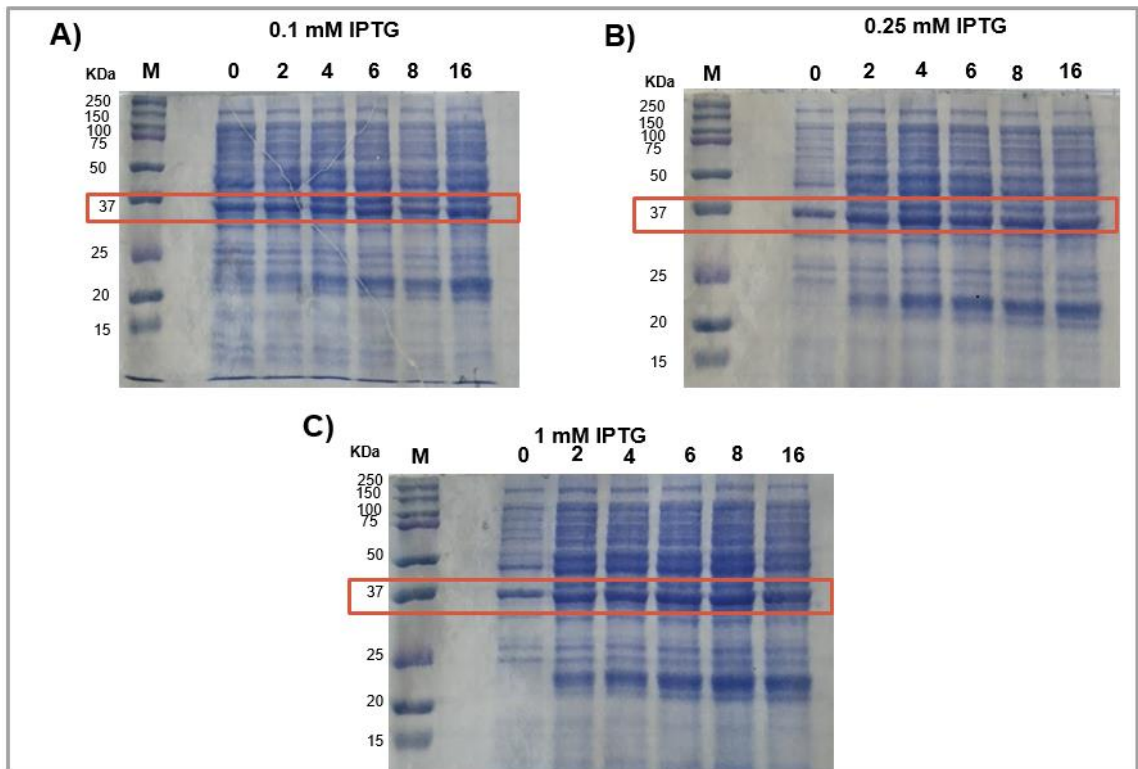
Por otro lado, la figura 22 muestra la expresión de rCatL5Fh tanto en la fracción soluble como insoluble del cultivo BL21 a 37 °C. En ella se muestra que no existe mucha variación de expresión de la recombinante entre los distintos tiempos de incubación post inducción, pero sí se evidencia que mientras mayor sea la concentración de IPTG, mayor es la recuperación de la proteína en la fracción soluble.



**Figura 22:** Expresión de rCatL5Fh en las muestras insolubles (i) como solubles (s) del cultivo a 37 °C de BL21. Las muestras fueron normalizadas por absorbancia a 600 nm. El carril M corresponde al marcador Precision plus protein Dual Color estándar; los carriles 0, 4, 6, 8 y 16 corresponden a horas de incubación post-inducción con IPTG. Las muestras fueron inducidas con **A)** 0.1, **B)** 0.25 y **C)** 1 mM de IPTG. SDS-PAGE al 12%.

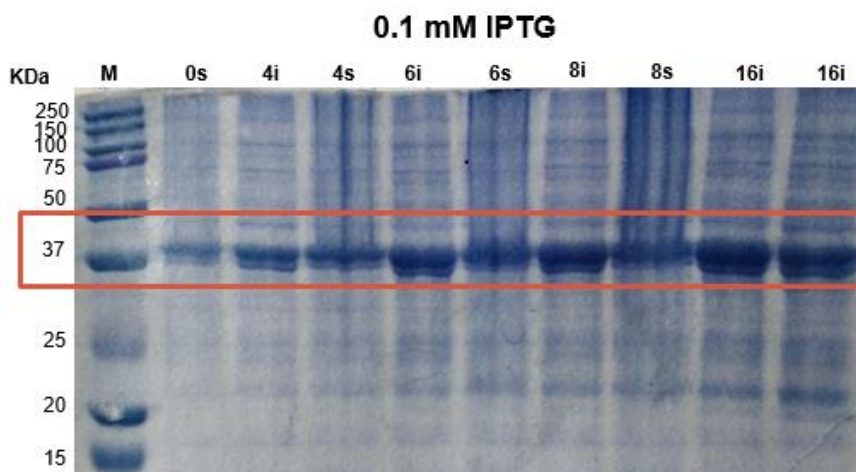
Debido a que se estima que la generación de cuerpos de inclusión inhabilita la expresión de la recombinante correctamente plegada, se realizó la expresión de la recombinante CatL5Fh en la cepa BL21, pero a una temperatura de inducción de 30 °C. De acuerdo a ello, la figura 23 muestra la expresión en la fracción insoluble, mientras que la figura 24 muestra la expresión tanto en la fracción insoluble como soluble a distintos tiempos de inducción y a diferentes concentraciones de IPTG.

En esta sección, se parte de la hipótesis que a menor concentración de IPTG y a menor temperatura de inducción, se produzca una mayor concentración de la proteína recombinante correctamente plegada en la fracción soluble.



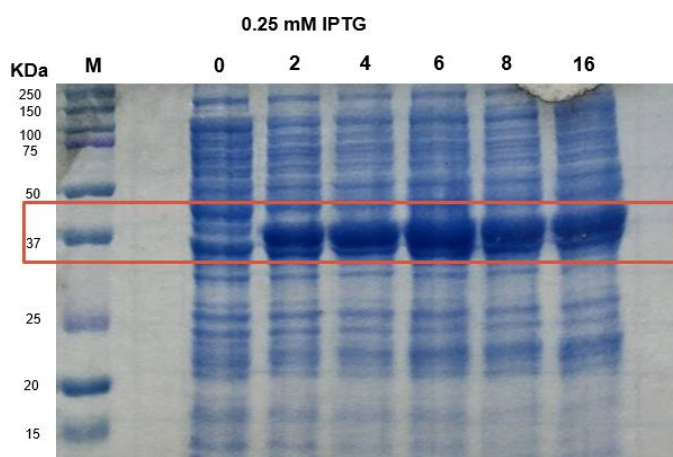
**Figura 23.** Expresión de rCatL5Fh en las muestras insolubles del cultivo a 30 °C de BL21. Las muestras fueron normalizadas por absorbancia a 600 nm. El carril M corresponde al marcador Precision plus protein Dual Color estándar; los carriles 0, 2, 4, 6, 8 y 16 corresponden a horas de incubación post-inducción con IPTG. Las muestras fueron inducidas con **A)** 0.1, **B)** 0.25 y **C)** 1 mM de IPTG. SDS-PAGE al 12%

No se observó una diferencia significativa en la expresión de la rCatL5Fh inducida a distintas concentraciones de IPTG y a distintos tiempos de inducción a 30 °C (figura 23). Por ello, se evaluó la expresión de la recombinante en la fracción soluble solo de la inducción a 0.1 mM de IPTG (figura 24) donde se estimaba mayor proporción de expresión. Los resultados indican una buena expresión en la fracción soluble entre las 6 y 8 horas de incubación post-inducción.



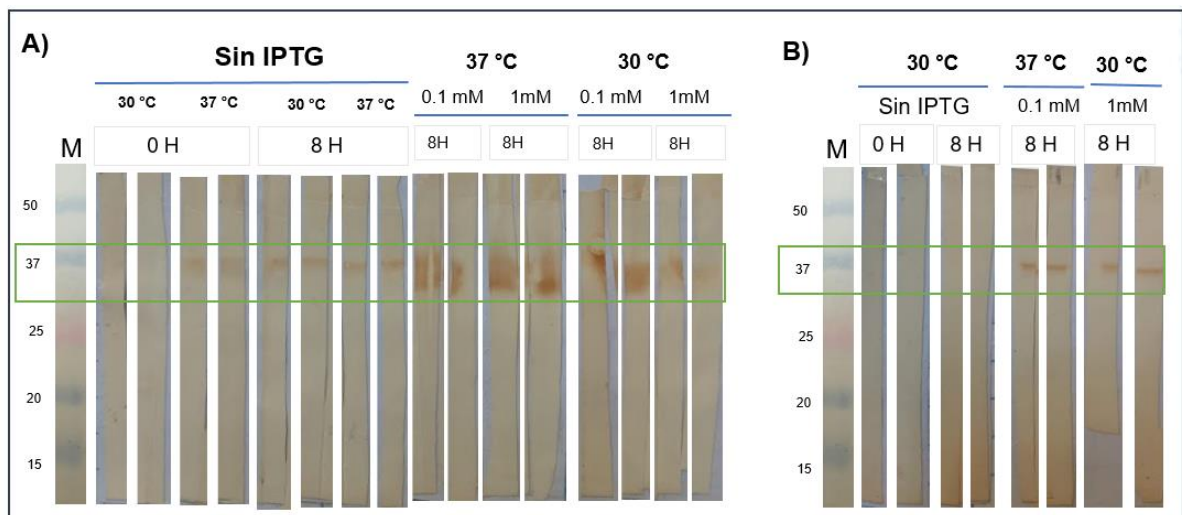
**Figura 24.** Expresión de rCatL5Fh en las muestras insolubles (i) como solubles (s) del cultivo a 30 °C de BL21. Las muestras fueron normalizadas por absorbancia a 600 nm. El carril M corresponde al marcador Precision plus protein Dual Color estándar; los carriles 0, 4, 6, 8 y 16 corresponden a horas de incubación post-inducción con IPTG a 0.1 mM. SDS-PAGE al 12%.

La expresión de rCatL5Fh fue evaluada usando la cepa Rosetta de *E. coli* que puede simular algunas modificaciones post traduccionales simples que posiblemente ocurren en la proteína nativa. De acuerdo a ello, la figura 25 muestra la buena expresión de la recombinante en las distintas horas de incubación post inducción a 37 °C.



**Figura 25.** Expresión de rCatL5Fh en las muestras insolubles del cultivo a 37°C en Rosetta. Las muestras fueron normalizadas por absorbancia a 600 nm. El carril M corresponde al marcador Precision plus protein Dual Color estándar; los carriles 0, 2, 4, 6, 8 y 16 corresponden a horas de incubación post-inducción con IPTG a 0.25 mM. SDS-PAGE al 12%.

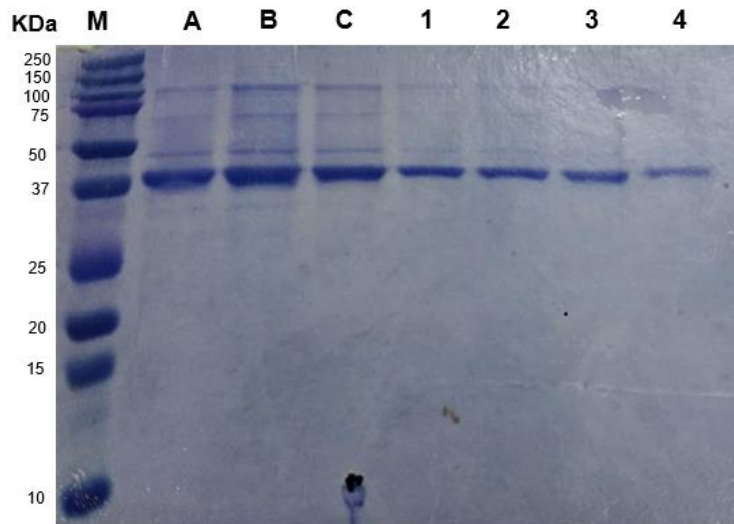
Finalmente, para determinar si la proteína recombinante se encuentra en las fracción soluble e insoluble, se realizó Western Blot utilizando anti penta-His para identificar la rCatL5Fh. La figura 26 muestra que la proteína recombinante se encuentran tanto en la fracción soluble como insoluble y a distintas concentraciones de IPTG como a distintas temperaturas de inducción (30 y 37 °C). Este último resultado evidencia la expresión de la rCatL5Fh a las distintas condiciones de inducción.



**Figura 26.** Análisis de la expresión de rCatL5Fh mediante Western Blot utilizando anti penta His para marcaje de la recombinante. El carril M corresponde al marcador Precision plus protein Dual Color estándar; los carriles 0H y 8H corresponden a 0 y 8 horas de incubación post-inducción evaluados en la fracción **A)** insoluble y **B)** soluble.

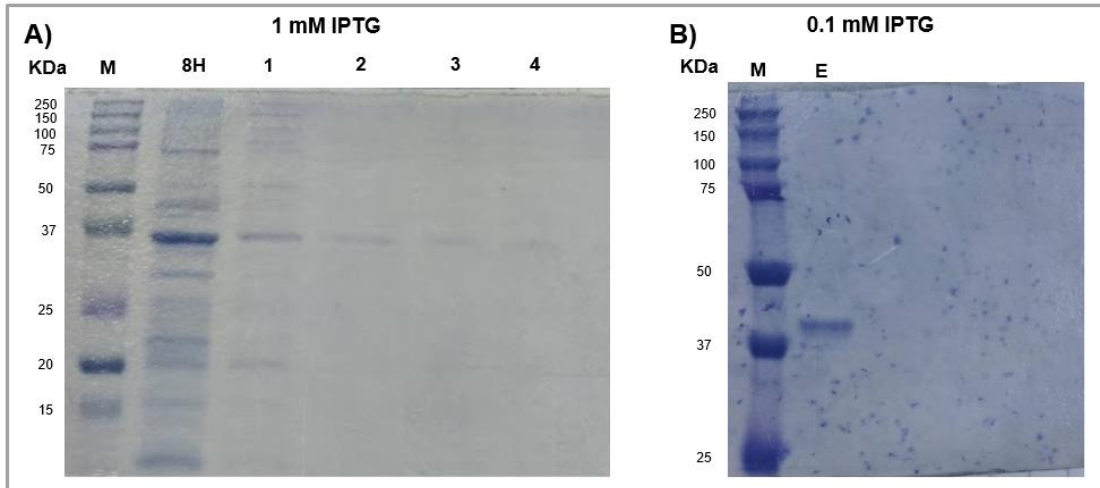
## IV.2 Purificación de rCatL5Fh

La purificación de la recombinante se realizó en distintos métodos, esto debido a obtener la proteína de forma activa. A partir de ello, en primer lugar, se purificó rCatL5Fh a condiciones desnaturizantes a partir de la fracción insoluble (pellet) (figura 27); la purificación en este caso se realizó a partir de la muestra incubada a 16 horas post-inducción a 37 °C y a 1 mM de IPTG. La figura 27 muestra la purificación de la proteína recombinante, las fracciones de los cuatro últimos carriles fueron colectadas en un solo pool para seguir con el proceso de re-plegamiento de la proteína siguiendo el protocolo descrito anteriormente.

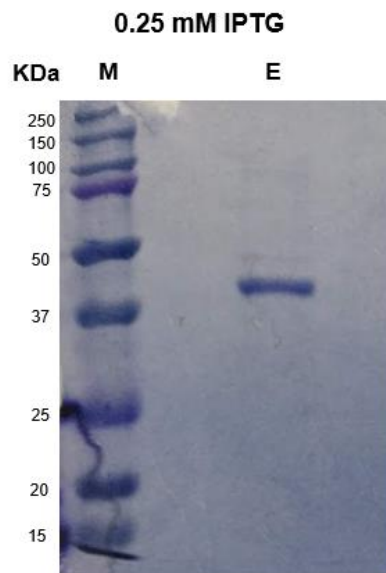


**Figura 27.** rCatL5Fh purificada a condición desnaturalizante de cultivo BL21 inducida por 16 horas a 37 °C con 1 mM de IPTG. Carril M corresponde al marcador Precision plus protein Dual Color estándar; carril A, B y C representan las fracciones de lavado de columna; carril 1, 2, 3 y 4 corresponden a la proteína purificada una vez agregada el buffer de elución. SDS-PAGE al 12 %.

En segundo lugar, se purificó rCatL5Fh a condiciones nativas a partir de la fracción soluble de la muestra inducida con IPTG 1 y 0.1 mM a 37 y 30 °C, respectivamente. Se realizó dicha purificación ya que se conjetura que la proteína correctamente plegada se encuentra en la fracción soluble por la menor temperatura de inducción y menor concentración de IPTG, se tiene más concentración de la recombinante en la fracción soluble. La figura 28 muestra la purificación de la rCatL5Fh a partir de la fracción soluble tanto de la inducción a 30 como a 37 °C. En tercer lugar, la recombinante fue purificada a condiciones híbridas a partir de la fracción insoluble de la inducción a 37 °C con 0.25 mM de IPTG y a 8 horas de incubación. La purificación a estas condiciones involucra la utilización de buffer desnaturalizante y posteriormente buffer nativo (descrito en la sección de métodos) con el objetivo de favorecer el plegamiento de la proteína durante la purificación (Figura 29).



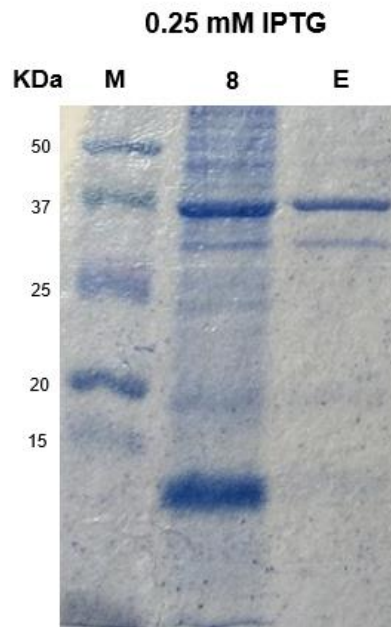
**Figura 28.** Purificación de rCatL5Fh de cultivo BL21 inducida a 1 y 0.1 mM de IPTG. Carril M corresponde al marcador Precision plus protein Dual Color estándar; carril 8H, expresión de la recombinante inducida a 8 horas con 1 mM de IPTG; carril 1, lavado de la columna de purificación; carril 2, 3 y 4 corresponde a la proteína purificada; carril E, corresponde a rCatL5Fh purificada. Las muestras fueron incubadas pos inducción a **A)** 37°C y **B)** 30 °C, respectivamente. SDS-PAGE al 10%



**Figura 29.** Purificación de rCatL5Fh a condiciones híbridas de cultivo BL21 inducida por 8 horas a 37 °C con 0.25 mM de IPTG. Carril M corresponde al marcador Precision plus protein Dual Color estándar; carril E la proteína purificada una vez agregada el buffer de elución. SDS-PAGE al 12 %.

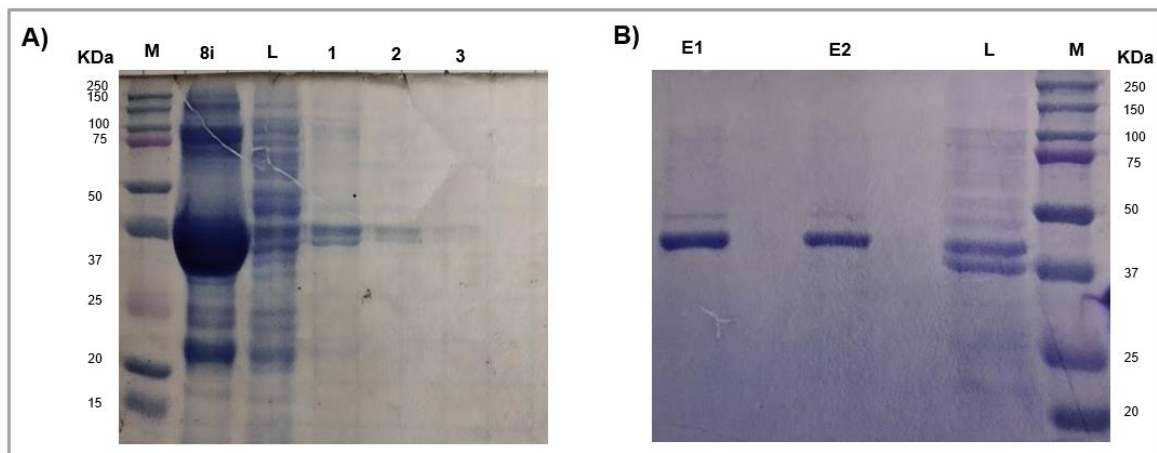


En cuarto lugar, la proteína recombinante fue purificada utilizando Sephadex G-25 como matriz. Para este método, se partió del supuesto que durante la purificación con los buffers que se utilizan, la proteína realiza el plegamiento teniendo en cuenta que favorecer dicho proceso conllevaría a tener la enzima pura y de forma activa. En la figura 30 muestra la proteína recombinante purificada.



**Figura 30.** Purificación de rCatL5Fh con Sephadex G-25 de cultivo BL21 inducida por 8 horas a 37 °C con 0.25 mM de IPTG. Carril M corresponde al marcador Precision plus protein Dual Color estándar; carril 8, inducción de la expresión a 8 horas de incubación a 37 °C; carril E, proteína purificada una vez agregada el buffer de elución. SDS.PAGE al 12%.

Finalmente, se purificó rCatL5Fh a partir de la fracción soluble e insoluble de la cepa Rosetta inducida por 8 horas con 0.25 mM de IPTG y a 37 °C de incubación. La purificación a partir de la fracción insoluble se realizó a condiciones híbridas y denaturalizantes; para este último involucró el proceso de re-plegamiento de la proteína mediante diálisis en dilución gradual de urea (6 a 1 M) en 20 mM de Tris-HCl y 5 mM de buffer EDTA pH 8. Se finalizó el re-plegamiento eliminando la urea mediante diálisis con PBS pH 7.4 (Figura 31 B). La purificación a partir de la fracción soluble fue realizada de acuerdo a las condiciones nativas (Figura 31 A).



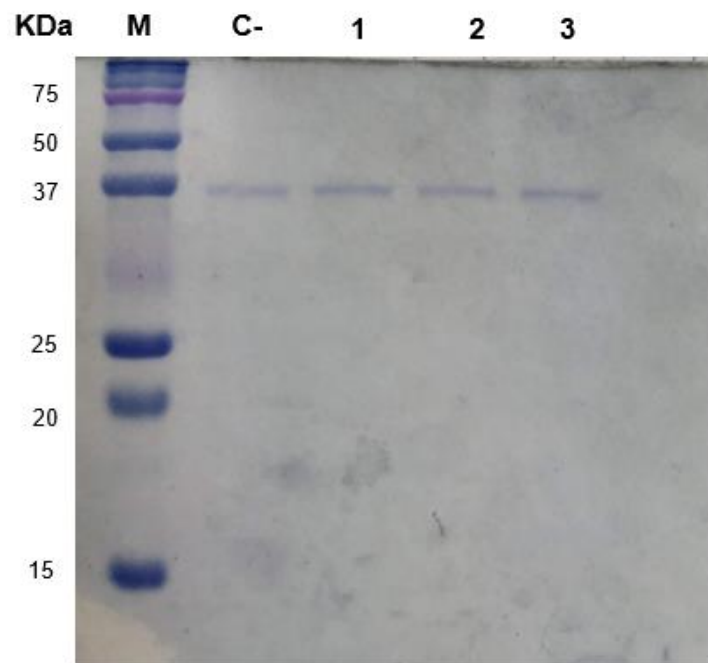
**Figura 31.** Purificación de rCatL5Fh a partir del cultivo de Rosetta inducida a 0.25 mM de IPTG e incubada a 37 °C. Carril M corresponde al marcador Precision plus protein Dual Color estándar. **A)** carril 8i, fracción insoluble de la expresión inducida a 8 horas; Carril L, fracción de lavado durante la purificación; carril 1, 2 y 3 presentan la proteína purificada a condiciones nativas. **B)** Carril E1, proteína purificada a condiciones desnaturizantes; carril E2, proteína purificada a condiciones híbridas. SDS-PAGE al 10%

#### **IV.3 Cuantificación y actividad enzimática de la proteína rCatL5Fh purificada.**

Como se describió anteriormente, tanto la expresión como la purificación de la recombinante fue realizada utilizando varias condiciones con el objetivo de obtener la enzima pura de forma activa. De acuerdo a ello, en la tabla siguiente se muestra las actividades enzimáticas de todas las enzimas purificadas e incluso del control positivo que corresponde a la CatL5Fh nativa aislada desde la *F. hepatica*. La actividad fue evaluada con Azocoll, un sustrato general para cisteínas proteasas. De acuerdo a los resultados (Tabla 4), no se evidencia actividad enzimática en ninguna de las rCatL5Fhs purificadas ya que las absorbancias a 520 son muy bajas en comparación a los controles positivos.

La evaluación de actividad enzimática corresponde también determinar si la proteína se está activando (corte del pro-dominio), para ello se activó rCatL5Fh en un medio reductor con 100 mM de IPTG y otro con el medio reductor más la enzima nativa activada. El resultado esperado de activación del zimógeno involucra la presencia del

pro-rCatL5Fh (37 kDa), rCatL5Fh (24 kDa) y el dominio pro (10 kDa aproximadamente) en SDS-PAGE al 15%. Sin embargo, la figura 32 muestra que no hay ningún proceso de activación por lo que el pro dominio no se está cortando y puede ser la principal causa de la falta de actividad enzimática.



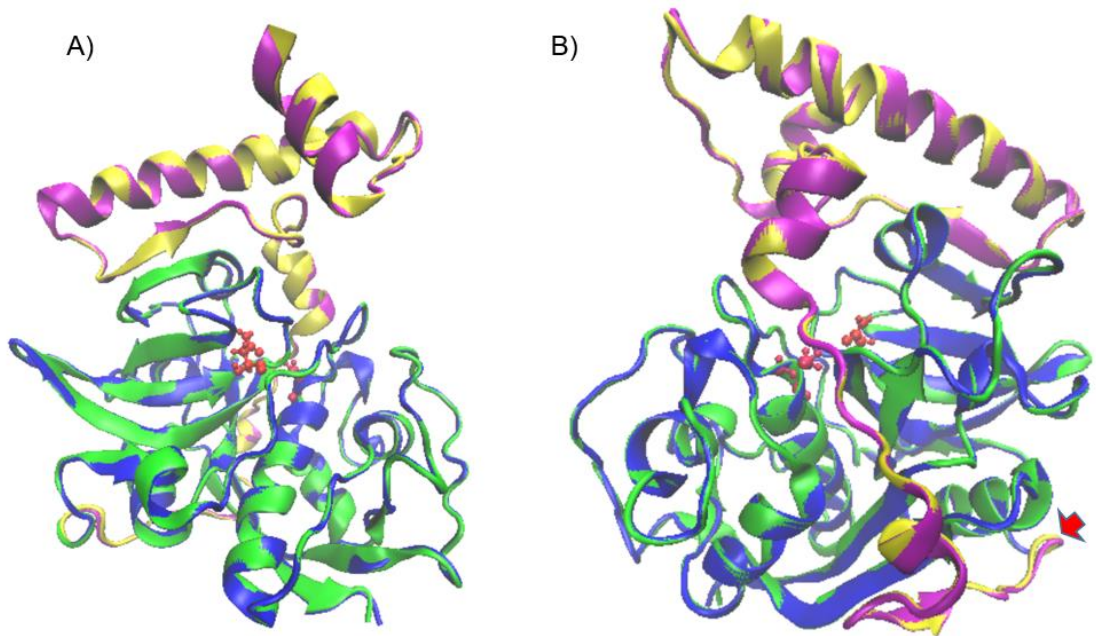
**Figura 32.** Ensayo de activación de rCatL5Fh purificadas a condiciones desnaturizantes y no desnaturizantes. Carril M corresponde al marcador Precision plus protein Dual Color estándar; carril C-, control negativo (sin DTT ni catepina activada); carril 1, rCatL5Fh + DTT 100mM; carril 2, rCatL5Fh purificada a condiciones no desnaturizantes + DTT 100mM + CatL5Fh nativa activada y carril 3 rCatL5Fh purificada a condiciones desnaturizantes + DTT 100mM + CatL5Fh nativa activada. SDS-PAGE al 15%.

**Tabla 4:** Producción de rCatL5Fh con diferentes condiciones de fermentación y procedimiento de purificación.

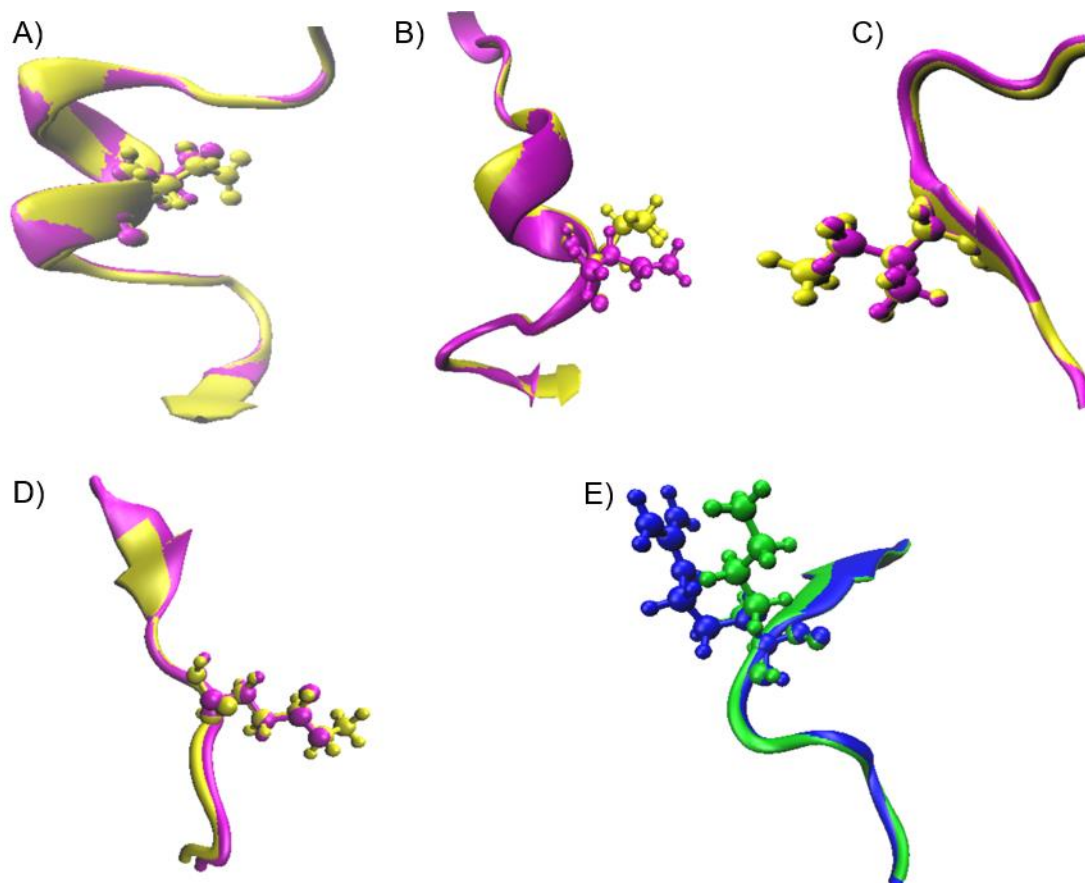
Cepa <i>E. coli</i>	Tamaño de cultivo	Temperatura fermentación	Tiempo fermentación	Concentración de IPTG	Fuente de purificación	Condición de purificación	Rendimiento (mg)	Actividad enzimática OD/min/mg
BL21	100 mL	37 °C	16 Horas	1 mM	Pellet	Denaturante	4.5	0
				0.25 mM	Pellet	Híbrido	3.1	0
				0.25 mM	Pellet	Con Sphadex	3	0
		1 mM	Sobrenadante	Nativo (Sin agentes caotrópicos)	0.6	0		
		0.1 mM	Sobrenadante	Nativo (Sin agentes caotrópicos)	0.7	0		
		0.25 mM	Pellet	Denaturante	6	0		
Rosetta		37 °C		0.25 mM	Pellet	Híbrido	3.4	0
				0.25 mM	Sobrenadante	Nativo (Sin agentes caotrópicos)	1	0
Control positivo (CatL5Fh native)								
							0.9 por lote	0.177
							0.97 por lote	0.169

#### IV.4 Alineamiento estructural entre rCatL5Fh y AF271385.1

Para determinar cuál es el efecto sobre la estructura de los aminoácidos variantes que están alrededor del sitio de corte del pro y dominio maduro, se realizó modelamiento estructural de rCatL5Fh y de la secuencia reportada por Smooker et al (AF271385.1) (49). Además de ello, se realizó el alineamiento estructural entre ambos y se evaluó las diferencias estructurales que podrían ser mediadas por las variaciones de aminoácidos entre las secuencias. La figura 33 muestra el alineamiento en la cual no se observa una diferencia clara entre ambas, por lo que a grandes rasgos los 14 aminoácidos variantes no presentan un efecto significativo sobre la estructura. Se realizó también la evaluación de los 5 aminoácidos variantes que están alrededor del sitio de corte del pro-dominio. La figura 34 muestra cómo los grupos R de los aminoácidos distintos se ubican en el espacio tridimensional de forma similar lo que en estructura no tendría mucho efecto. Además, el hecho de estar cerca del sitio de corte, estos aminoácidos podrían no presentar efecto negativo sobre el reconocimiento del sustrato ya que, entre ambas secuencias, los residuos presentan propiedades fisicoquímicas muy similares.



**Figura 33.** Alineamiento estructural entre la proteína recombinante y nativa con visión A) frontal y B) posterior. Dominio maduro (Verde) y pro-dominio (magenta) de rCatL5Fh; dominio maduro (azul) y pro-dominio (amarillo) de AF271385.1 y flecha roja, sitio de corte entre pro y dominio maduro.



**Figura 34:** Aminoácidos distintos entre la recombinante y la nativa ubicados alrededor del sitio de corte entre pre y pro-dominio (detallado en la figura 9). Dominio maduro (Verde) y pro-dominio (magenta) de rCatL5Fh; dominio maduro (azul) y pro-dominio (amarillo) de AF271385.1. La gráfica muestra alineamiento entre A) Ácido aspártico (magenta) y ácido glutámico (amarillo); B) Isoleucina (magenta) y leucina (amarillo); C) Valina (magenta) e isoleucina (amarillo); D) Ácido glutámico (magenta) y lisina (amarillo); E) Lisina (verde) y arginina (azul).

## CAPÍTULO V

### DISCUSIÓN

El objetivo del presente estudio fue producir rCatL5Fh en un sistema procariota y caracterizar la actividad enzimática de la proteína recombinante. Para ello, inicialmente se clonó la secuencia codificante de CatL5Fh en el vector de expresión pET26-b(+). Este vector de expresión presenta en la región upstream del gen clonado, un promotor T7, una secuencia de operador *lac* adyacente (que controla la expresión de la recombinante por IPTG) (97), secuencia Shine-Dalgarno para favorecer la interacción del ARNm con la subunidad 30S del ribosoma y la secuencia que codifica para un tag de purificación poli-histidina (6-His) (98). Las características de pET26-b(+) permiten expresar la proteína recombinante a escala en forma controlada mediante el inductor IPTG.

Debido a que la clonación y la transformación del plásmido fue realizado previo a este estudio, se verificó la integridad de la secuencia para continuar con el sistema de expresión; para ello, se aisló el plásmido recombinante desde BL21 y se secuenció el inserto utilizando primers forward y reverse específicos para CatL5Fh. Finalmente, se alineó la secuencia primaria de aminoácidos de la recombinante con la nativa reportada por el grupo de los australianos (AF271385.1). El alineamiento mostró algunas variaciones en la secuencia con respecto a la reportada, que, por un lado, pueden ser propias de la secuencia de la cepa peruana o por algunas mutaciones ocurridas durante la transformación o producción en la BL21 recombinante (Figura 19). Al realizar el análisis de las variaciones encontradas en el alineamiento; se plantea la primera hipótesis referenciando que podrían ser parte de los productos de mutación ocurrida en el hospedero de *E. coli*. Estudios anteriores reportan que la sustitución de residuos con aminoácidos con propiedades fisicoquímicas similares es más probable que mutaciones dando lugar aminoácidos muy diferentes. Esto significaría que estas sustituciones no presentarían mayores efectos en la dinámica tanto local (movimiento de residuos de aminoácidos) como global (movimiento de dominios) de la proteína la cual están regidas por las interacciones tipo puentes de hidrógeno, las fuerzas de Van der Waals e hidrofóbica entre cada átomo de cada aminoácido (102,103). De forma general, cada residuo, según sus características favorece el arreglo estructural de la proteína

independientemente gracias a la densidad electrónica de sus grupos funcionales o por las interacciones que presente frente a otros residuos (102,103). De acuerdo a lo mencionado y debido a la similitud fisicoquímica de los residuos sustituidos, se espera que las mutaciones no tengan efecto negativo en la actividad enzimática de la proteína CatL5Fh, en caso estas mutaciones hayan ocurrido en el hospedero recombinante.

La segunda hipótesis planteada respecto a las variaciones encontradas rige con enfoque a las secuencias que podrían ser propias de las cepas tanto de Australia como de Perú. De ser diferentes estas secuencias, los aminoácidos que difieren no aportarían como efecto negativo sobre la estructura de las proteínas, sino más bien podrían contribuir a una función específica de cada variante la cual podría estar relacionada a un rasgo característico de cada cepa como parte de la adaptación de la *F. hepatica* a los diferentes ambientes.

Para determinar si estas variaciones pueden tener un efecto significativo sobre la estructura o función de la enzima; en primer lugar se evaluó la correcta posición de los tres aminoácidos (Cys, His y Asn) (37) importantes para la actividad catalítica de la enzima; en segundo lugar, debido a que la enzima CatL5Fh necesita el arreglo estructural de sus sub-sitios bien definidos para poder interactuar de forma correcta con sus sustratos específicos, se evaluó también si estos aminoácidos presentan alguna variación con respecto a la reportada (67,68). En efecto, no se identificó ningún cambio en los 6 aminoácidos esenciales en las siguientes posiciones de los sub-sitios: 67, 68, 133, 156, 157 y 205 (recuadro celeste de la figura 19) ni en los tres aminoácidos del sitio activo. Cabe recalcar que CatL5Fh es una enzima que se sintetiza *in vivo* como zimógeno, lo que implica la presencia de un pro-dominio y dominio maduro; a partir de ello, se identificó 7 variaciones en ambos dominios. El análisis de la región madura presenta que los aminoácidos que difieren son fisicoquímicamente similares (R5K, Q43E, R79Q, Q123K, K199N, V212L, Q217R) lo que sugiere que dichos cambios no tendrían un efecto significativo sobre la estructura y función de la enzima (99-101).

Por otro lado, de los 7 aminoácidos que variaron en el pro dominio, 6 residuos tienen la propiedad fisicoquímica similar; pero la variación de lisina por treonina tienen propiedades muy distintas. Debido a que la mayor parte de los aminoácidos son



idénticos a la nativa, se espera que el pro-dominio expresado presenta la función característica de inhibición sobre la actividad del dominio maduro. Finalmente, debido a que la secuencia de la recombinante fue considerada homóloga a la reportada (95.5% de identidad y 99% de cobertura), se continuó con el sistema de expresión de la recombinante.

Ya que rCatL5Fh fue expresada a distintas concentraciones de IPTG, inicialmente se evaluó la citotoxicidad de IPTG sobre el crecimiento bacteriano post-inducción. Los resultados mostrados en la figura 20 indican que no existe variación en la tasa de crecimiento de *E. coli* post-inducción frente a las distintas concentraciones de IPTG. Ello no concuerda con reportes que indican que IPTG presenta un efecto negativo sobre la tasa de crecimiento bacteriano, haciendo referencia que la molécula causa un efecto tóxico y/o impone una carga metabólica a la célula debido a la expresión del gen recombinante o toxicidad proteica (104). De la misma forma se evaluó la tasa de crecimiento a las temperaturas de 30 y 37°C incubadas luego de la inducción; en la figura 20 se evidencia que existe una mayor tasa de crecimiento a 37°C, esto debido a que es la temperatura óptima de crecimiento de *E. coli*. Sin embargo, a 30°C la bacteria presenta menor tasa debido a que las velocidades de los procesos metabólicos disminuyen y la bacteria sufre un proceso de adaptación lo cual indirectamente implica la disminución en expresión de la proteína recombinante (105).

En el presente estudio, se evaluó la expresión de rCatL5Fh con distintas concentraciones de IPTG y a distintas temperaturas y tiempos de fermentación. Considerando que la expresión de la proteína recombinante en forma activa y soluble se obtendría a menor concentración de IPTG y a menor temperatura de fermentación (30°C); y teniendo en cuenta que condiciones contrarias (37°C de fermentación) puede conllevar al sistema generar niveles altos de expresión de la recombinante la cual pueden ser tóxicos para la bacteria y generar cuerpos de inclusión (agregación de proteínas mal plegadas) como parte de su sistema de defensa del propio hospedero (106,107). Esta última es una desventaja para la producción de la recombinante en su forma activa como es el caso de rCatL5Fh.

Se encontró una mayor expresión de la proteína recombinante rCatL5Fh de 37 kDa a mayor concentración de IPTG y a 37°C de fermentación (Figura 21); además, se observa un incremento en la expresión de la proteína recombinante cuando el tiempo de incubación post-inducción es mayor, teniendo en cuenta que a las 8 horas ya se tiene una buena expresión. Por otro lado, se pudo expresar la recombinante en la fracción soluble (figura 22) donde se evidencia que, a mayor concentración de inductor, mayor fue la síntesis en dicha fracción. De la misma forma se realizó la expresión a 30°C, en el cual se pudo expresar la recombinante tanto en la fracción insoluble (figura 23) como soluble (figura 24); para este último, se espera que la concentración proteica en la fracción soluble sea mayor. Finalmente, para confirmar la expresión de la recombinante en ambas fracciones (soluble e insoluble), se realizó Western Blot con el anticuerpo conjugado anti poli-histidina que se une a la cola de histidina de la recombinante con alta especificidad. En efecto, la figura 26 muestra la eficaz expresión de rCatL5Fh a las temperaturas de 30 y 37 °C, con distintas concentraciones de IPTG y a 8 horas de incubación post-inducción.

Se realizó la expresión de rCatL5Fh en distintas condiciones de fermentación para obtener la enzima en su forma activa; adicionalmente, se evaluaron cuatro métodos para la purificación de la recombinante. El primer método de purificación fue a partir de la fracción insoluble (figura 27) de la recombinante expresada a 37 °C y 16 horas de incubación; la proteína purificada luego se enfrentó a un proceso de re-plegamiento mediante diálisis y utilizando componentes que permiten realizar el intercambio del agente caotrópico (urea). Este proceso de re- plegamiento proteico fue realizado en una cisteína proteasa expresada en *E. coli*, obteniendo la forma activa (108). El segundo método de purificación fue a partir de la fracción soluble de la inducción a 30 y 37 °C (figura 28). Se realizó la purificación a partir de la fuente soluble debido a que los estudios anteriores mencionan que la probabilidad de encontrar la proteína recombinante en su forma correctamente plegada es muy alta en dicha fracción (109-111). El tercer método de purificación se realizó a condiciones híbridas (Figura 32), quiere decir que en el proceso de fijación a la matriz de Níquel se utilizó Buffer con componentes desnaturizantes (guanidina), luego, durante la elución se utilizó buffer que no tienen agentes desnaturizantes (urea, guanidina) para favorecer el re-plegamiento de la proteína durante la elución. Este proceso de obtención de la enzima

en su forma correctamente plegada fue descrito por la empresa NOVEX en su sistema de Probond Purifications System como una variante de la metodología de purificación.

El cuarto método de purificación de rCatL5Fh fue a partir de la fracción insoluble utilizando Sephadex G-25 como matriz (Figura 30) (112) que permite remover la urea y favorecer el re-plegamiento de la proteína recombinante durante la elución.

Se obtuvo diferentes cantidades de la proteína recombinante empleando distintos métodos de purificación; siendo la purificación en condiciones desnaturalizantes la que permitió recuperar la mayor cantidad de rCatL5Fh. Los distintos métodos empleados tuvieron como objetivo evaluar las condiciones para producir la proteína rCatL5Fh en forma activa; sin embargo, no se observó actividad enzimática con azocoll en las proteínas purificadas.

Por otro lado, rCatL5Fh también fue expresado en la cepa Rosseta de *E. coli*. que puede realizar simples modificaciones como son las O-glicosilaciones; las cuales pueden ser suficientes para que rCatL5Fh presente actividad. La expresión y purificación fue exitosa tanto de la fracción soluble como insoluble; sin embargo, no se observó actividad enzimática con Azocoll. Dicho resultado sugiere que las modificaciones realizadas para CatL5Fh nativa son complejas donde involucra la intervención de retículo endoplasmático y aparato de Golgi.

La ausencia de actividad enzimática podría deberse a distintos factores que involucran la secuencia, sistema de activación o modificaciones post-traduccionales. Las catepsinas en general necesitan realizar el corte de su pro-dominio para que puedan presentar actividad; además, se tiene en cuenta que la activación (que involucra desestabilizar la interacción entre el pro-dominio y el dominio maduro) se realiza disminuyendo el pH (ácido), utilizando agentes reductores o incluyendo en la reacción catepsinas activadas de la misma familia. Para evaluar la activación de rCatL5Fh se realizó tres condiciones: el control, que corresponde a la recombinante sin agente reductor ni catepsina nativa activada; ensayo 1, rCatL5Fh + DTT y ensayo 2, rCatL5Fh + CatL5Fh nativa activada + DTT. El resultado esperado del presente ensayo es visualizar dos o tres bandas en SDS-PAGE al 15% que corresponden al pro-rCatL5Fh (37 Kda), rCatL5Fh madura (24

KDa) y el dominio pro (alrededor de 10 KDa). La figura 32 muestra que no ocurrió activación por lo que se plantea dos aproximaciones: la activación de la recombinante no ocurrió debido a que los aminoácidos variantes que están alrededor del sitio de corte entre el pro y dominio maduro alteran este proceso (residuo 1 figura 19) o que requiere de las modificaciones post-traduccionales para la activación y presentar actividad enzimática.

Para responder el primer caso y determinar el efecto estructural de cada variante, se construyó la estructura tridimensional de rCatL5Fh y de la secuencia AF271385.1 mediante modelamiento por homología utilizando SWISS MODEL. Para determinar el efecto de cada variante sobre la estructura, se realizó alineamiento estructural entre rCatL5Fh y AF271385.1. En la figura 33 (anexo) se muestra la alineación estructural; y en la figura 34 (anexo) se muestra el alineamiento de las 5 variantes de aminoácidos que están alrededor del sitio de corte entre el pro y dominio maduro (figura 19 - residuo 1). El alineamiento estructural muestra que los aminoácidos distintos no causan una variación drástica sobre la estructura secundaria de la proteína, lo que conlleva a reforzar que la no activación no será efecto de estas variantes sino más bien de otra variable como pueden ser las modificaciones pos-traduccionales. Por otro lado, al encontrar que los aminoácidos distintos no son críticos para la estructura, se sugiere que dichos residuos serían parte de cada cepa tanto peruana como australiana y que estaría relacionado a la adaptación de cada una de ella con el tipo de ambiente en el que se encuentran. Para corroborar la secuencia peruana, el grupo de investigación anteriormente realizó una espectrometría de masas (Maldi-tof) donde se determinó tres principales péptidos correspondiente a CatL5Fh (figura 19); sin embargo, estos péptidos están ubicados fuera de las regiones que presentan variaciones por lo que no podemos concluir con certeza que las secuencias de la cepa peruana y australiana son distintas.

Para tener más claro la diferencia de secuencias entre ambas cepas se sugiere realizar un micro-secuenciamiento. Esta última consiste en aislar ARN desde la *Fasciola hepatica* usando un método de isocianato/fenol/cloroformo, posteriormente separar los fragmentos de ARN mediante electroforesis a condiciones denaturantes, y luego ligar los fragmentos de tamaño apropiado usando adaptadores tanto del 5' como de 3'. Los adaptadores permiten en el siguiente paso generar el cADN mediante transcripción reversa para posteriormente realizar una PCR. Finalmente se realiza el secuenciamiento guiado por un chip en donde se detecta los nucleótidos que van amplificando durante la PCR (116). Lo que se espera con este tipo de secuenciamiento es corroborar con mayor detalle la secuencia de la cepa peruana y comprobar que los aminoácidos distintos son propios de las variantes y por ende los aminoácidos alrededor de la secuencia pro y dominio maduro no tendrían efecto negativo sobre la activación de la enzima.

Finalmente, los análisis realizados en el presente estudio, sugiere que rCatL5Fh requiere de la expresión en un sistema eucariota para simular las modificaciones post-traduccionales probables la cual permitan obtener la enzima con una buena estabilidad y que sea funcionalmente activa (113-115). Esta cisteína proteinasa requiere de la formación de enlaces di-sulfuros para tener un correcto plegamiento (91). Se realizó varios métodos de re-plegamiento para favorecer este proceso, pero no se pudo obtener la enzima de forma activa. De acuerdo a ello, se recomienda expresar esta enzima en un sistema de expresión eucariota justamente para simular las modificaciones post-traduccionales que ocurre en la síntesis y secreción de CatL5Fh en el parásito. Es probable que expresar en un sistema eucariota nos permita obtener la enzima de forma activa ya que algunos grupos de investigadores lo realizaron (16) y además porque la caracterización de la secuencia del pro-dominio hace referencia que el péptido sintetizado es conllevado por el camino de secreción como es retículo endoplasmático y aparato de golgi donde se producen las modificaciones pos traduccionales especiales para proteínas secretadas.

Sin embargo, si en posteriores trabajos se pretende seguir con la expresión en sistemas procariotas, es necesario evaluar la actividad enzimática de CatL5Fh nativa después de haber pasado por todos los sistemas de purificación empleadas en la purificación de la recombinante. Si en caso se logra actividad, se corrobora que los buffers o los sistemas

de purificación permite recuperar la proteína en su forma estructuralmente activa y de tal forma corrobora que la no actividad de la recombinante sería parte de las modificaciones pos-traduccionales. En caso no existiera actividad de la nativa, se podría modificar los buffers con el objetivo de recuperar la estructura conformacional activa de la nativa y a la vez de la recombinante. El trabajo para este último sería arriesgado ya que los buffers usados en los sistemas de purificación no presentan ninguna propiedad de interrumpir con la estructura de la proteína.

## CAPÍTULO VI

### CONCLUSIÓN

La proteína rCatL5Fh se expresó en el sistema de expresión pET-26b(+) en las cepas BL21 y Rosetta de *E. coli* obteniéndose los más altos rendimientos de producción con BL21 de 4.523 mg por 100 ml de cultivo con 1 mM IPTG, 37 °C por 16 h y con Rosetta de 5.995 mg por 100 ml con 0.25 mM IPTG, 37°C por 8 h.

rCatL5Fh no presentó actividad enzimática evaluada con el sustrato Azocoll; por lo que no se pudo caracterizar la actividad enzimática sobre el sustrato ST3.

Se sugiere que el sistema de expresión bacteriano no permite obtener la cisteína proteasa de forma activa ya que esta no presenta modificaciones pos-traducción presentes probablemente en CatL5Fh nativa. Por lo que se recomienda usar un sistema de expresión eucariota para poder expresar la proteinasa funcionalmente activa.

## CAPÍTULO VII

### REFERENCIAS BIBLIOGRÁFICAS

1. Galaktionov K-V., Dobrovolskij AA. The Biology and Evolution of Trematodes. Fried Bernard GT, editor. The Biology and Evolution of Trematodes. Russia; 2003.
2. Lotfy W, Brant S, Dejong RJ, Le TH, Demiaszkiewics A, et al. Evolutionary Origins, Diversification, and Biogeography of Liver Flukes (Digenea, Fasciolidae). *Med Sch.* 2008;248–55.
3. Beesley NJ, Williams DJL, Paterson S, Hodgkinson J. *Fasciola hepatica* demonstrates high levels of genetic diversity, a lack of population structure and high gene flow: possible implications for drug resistance. *Int J Parasitos.* 2017;47(1):11–20. Doi.org/10.1016/j.ijpara.2016.09.007
4. Valero MA, Panova M, Andre AV. Phenotypic analysis of adults and eggs of *Fasciola hepatica* by computer image analysis system. *Helminthology.* 2005;217–25.
5. Valero MA, Perez-Crespo I, Khoubbane M, Artigas P, Panova M, Ortiz P, et al. *Fasciola hepatica* phenotypic characterization in Andean human endemic areas: Valley versus altiplanic patterns analysed in liver flukes from sheep from Cajamarca and Mantaro, Peru. *Infect Genet Evol.* 2012;12(2):403–10 Doi.org/10.1016/j.meegid.2012.01.009
6. Valero MA, Panova M, Comes AM, Fons R, Mas-Coma S. Patterns in size and shedding of *Fasciola hepatica* eggs by naturally and experimentally infected murid rodents. *J Parasitol.* 2002;88(2):308–13.
7. Hussein ANA, Hassan IM, Khalifa RMA. Development and hatching mechanism of *Fasciola* eggs, light and scanning electron microscopic studies. *Saudi J Biol Sci.* 2010;17(3):247–51. Doi.org/10.1016/j.sjbs.2010.04.010
8. Periago M V., Valero MA, El Sayed M, Ashrafi K, El Wakeel A, Mohamed MY, et al. First phenotypic description of *Fasciola hepatica/Fasciola gigantica* intermediate forms from the human endemic area of the Nile Delta, Egypt. *Infect Genet Evol.* 2008;8(1):51–8.
9. Beesley NJ, Camide C, Charlier J, Flynn RJ, Hodgkinson JE, Martínez moreo A, et al. *Fasciola* and fasciolosis in ruminants in Europe : Identifying research needs. 2018;65(April 2017):199–216.
10. Rondelaud D, Vignoles P, Abrous M, Dreyfuss G. The definitive and intermediate hosts of *Fasciola hepatica* in the natural watercress beds in central



- France. Parasitol Res. 2001;87(6):475–8.
11. Rondelaud D, Dreyfuss G. *Fasciola hepatica* : The influence of the definitive host on the characteristics of infection in the snail *Lymnaea truncatula* . Parasite. 1995;2(3):275–80.
  12. Arbabi M, Nezami E, Hooshyar H, Delavari M. Epidemiology and economic loss of fasciolosis and dicrocoeliosis in Arak, Iran. Vet World. 2018;11(12):1648–55.
  13. Espinoza J, Terashima A, Herrera Patricia ML. Fasciolosis humana y animal en el Perú: impacto en la economís de las zonas endémicas. Med Exp Salud pública. 2010;4:604–12.
  14. Bennett R, IJpelaar J. Updated estimates of the costs associated with thirty four endemic livestock diseases in Great Britain: A note. J Agric Econ. 2005;56(1):135–44.
  15. Moazeni M, Ahmadi A. Controversial aspects of the life cycle of *Fasciola hepatica*. Exp Parasitol. 2016;169:81–9. Doi.org/10.1016/j.exppara.2016.07.010
  16. Yap HY. Function and molecular interactions of the *Fasciola hepatica* cathepsin L5 protease. Sch Sci. 2018;(February).
  17. Hurtrez-Boussés, Meunier C; Durand P, Renaud F. Dynamics of host–parasite interactions: the example of population biology of the liver fluke (*Fasciola hepatica*). Microbes Infect. 2001;3:841–9.
  18. Despommier DD, Karapelou JW. Parasite Life Cycles. Parasite Life Cycles. 1988. 69–90 p.
  19. Lindquist HDA, Cross JH. Helminths. Infect Dis (Auckl). 2017;1763–79.
  20. Moazeni M, Ansari-Lari M, Masoodfar M, Hosseinzadeh S, Mootabi Alavi AM. Lethal effect of high temperatures on the eggs of *Fasciola hepatica*. Iran J Vet Res. 2010;11(2):168–73.
  21. Yakhchali M, Bahramnejad K. Inhibition effect of pH on the hatchability of *fasciola* miracidia under laboratory conditions. Iran J Parasitol. 2016;11(1):30–4.
  22. Cañete R, Yong M, Sánchez J, Wong L, Gutiérrez A. Population dynamics of intermediate snail hosts of *Fasciola hepatica* and some environmental factors in San Juan y Martinez Municipality, Cuba. Mem Inst Oswaldo Cruz. 2004;99(3):257–62.
  23. Boix N, Teixidó E, Vila-Cejudo M, Ortiz P, Ibáñez E, Gomez-Catalan J, et al. Risk assessment for human embryonic development of triclabendazole residues in milk and cheese in the diet of a rural population in Cajamarca ( Peru ): A preliminary approach. Recent Adv Pharm Sci VI. 2016;661(2):37–47.

24. Cwiklinski K, Dalton JP. Advances in *Fasciola hepatica* research using ‘omics’ technologies. *Int J Parasitol.* 2018;48(5):321–31.
25. Cwiklinski K, Dalton JP, Dufresne PJ, La Course J, Williams DJL, Hodgkinson J, et al. The *Fasciola hepatica* genome: Gene duplication and polymorphism reveals adaptation to the host environment and the capacity for rapid evolution. *Genome Biol.* 2015;16(1):1–13.
26. Cwiklinski K, Jewhurst H, McVeigh P, Barbour T, Maule AG, Tort J, et al. Infection by the helminth parasite *Fasciola hepatica* requires rapid regulation of metabolic, virulence, and invasive factors to adjust to its mammalian host. *Mol Cell Proteomics* 2018.
27. Robinson MW, Menon R, Donnelly SM, Dalton JP, Ranganathan S. An Integrated Transcriptomics and Proteomics Analysis of the Secretome of the Helminth Pathogen *Fasciola hepatica*. *Mol Cell Proteomics.* 2009;8(8):1891–907.
28. Hernández-González A, Valero ML, Pino MS del, Oleaga A, Siles-Lucas M. Proteomic analysis of in vitro newly excysted juveniles from *Fasciola hepatica*. *Mol Biochem Parasitol.* 2010;172(2):121–8.
29. Mas-Coma S, Bargues MD, Valero MA. Fascioliasis and other plant-borne trematode zoonoses. *Int J Parasitol.* 2005;35(11–12):1255–78.
30. Liu Q, Huang SY, Yue DM, Wang JL, Wang Y, Li X, et al. Proteomic analysis of *Fasciola hepatica* excretory and secretory products (FhESPs) involved in interacting with host PBMCs and cytokines by shotgun LC-MS/MS. *Parasitol Res.* 2017;116(2):627–35.
31. Vázquez AA, Lounnas M, Sánchez J, Alba A, Milesi A, Hurtrez-Boussès S. Genetic and infective diversity of the liver fluke *Fasciola hepatica* (Trematoda: Digenea) from Cuba. *J Helminthol.* 2016;90(6):719–25.
32. Di Maggio LS, Tirloni L, Pinto AFM, Diedrich JK, Yates JR, Benavides U, et al. Across intra-mammalian stages of the liver fluke *Fasciola hepatica*: A proteomic study. *Sci Rep.* 2016;6(September):1–14.
33. Morphew RM, Wright HA, LaCourse EJ, Woods DJ, Brophy PM. Comparative proteomics of excretory-secretory proteins released by the liver fluke *Fasciola hepatica* in sheep host bile and during in vitro culture ex host. *Mol Cell Proteomics.* 2007;6(6):963–72.
34. Rawlings ND, Barrett AJ. Families of cysteine peptidases. *Methods Enzymol.* 1994;244(C):461–86.
35. Cianni L, Feldmann CW, Gilberg E, Gütschow M, Juliano L, Leitão A, et al. Can Cysteine Protease Cross-Class Inhibitors Achieve Selectivity? *J Med Chem.*

- 2019;62(23):10497–525.
36. Barrett AJ, Rawlings ND. Evolutionary lines of cysteine peptidases. *Biol Chem.* 2001;382(5):727–33.
  37. Verma S, Dixit R, Pandey KC. Cysteine proteases: Modes of activation and future prospects as pharmacological targets. *Front Pharmacol.* 2016;7(APR):1–12.
  38. Novinec M, Lenarčič B, Turk B. Cysteine cathepsin activity regulation by glycosaminoglycans. *Biomed Res Int.* 2014;2014.
  39. Klaudia B. WS. Proteases: structure and function. *Proteases: Structure and Function.* 2013. 1–564 p.
  40. McKerrow JH. The diverse roles of cysteine proteases in parasites and their suitability as drug targets. *PLoS Negl Trop Dis.* 2018;12(8):1–3.
  41. Grzonka Z, Jankowska E, Kasprzykowski F, Kasprzykowska R, Łankiewicz L, Wiczek W, et al. Structural studies of cysteine proteases and their inhibitors. *Acta Biochim Pol.* 2001;48(1):1–20.
  42. McGonigle L, Mousley A, Marks NJ, Brennan GP, Dalton JP, Spithill TW, et al. The silencing of cysteine proteases in *Fasciola hepatica* newly excysted juveniles using RNA interference reduces gut penetration. *Int J Parasitol.* 2008;38(2):149–55.
  43. Sajid M, McKerrow JH. Cysteine proteases of parasitic organisms. *Mol Biochem Parasitol.* 2002;120(1):1–21.
  44. Selzer PM, Pingel S, Hsieh I, Ugele B, Chan VJ, Engel JC, et al. Cysteine protease inhibitors as chemotherapy: Lessons from a parasite target. *Proc Natl Acad Sci U S A.* 1999;96(20):11015–22.
  45. McVeigh P, Maule AG, Dalton JP, Robinson MW. *Fasciola hepatica* virulence-associated cysteine peptidases: A systems biology perspective. *Microbes Infect.* 2012;14(4):301–10. [Doi.org/10.1016/j.micinf.2011.11.012](https://doi.org/10.1016/j.micinf.2011.11.012)
  46. Dalton JP, Robinson MW, Mulcahy G, O'Neill SM, Donnelly S. Immunomodulatory molecules of *Fasciola hepatica*: Candidates for both vaccine and immunotherapeutic development. *Vet Parasitol.* 2013;195(3–4):272–85.
  47. Rawlings ND, Barrett AJ. Introduction: The Clans and Families of Cysteine Peptidases. Vol. 2, *Handbook of Proteolytic Enzymes.* 2013. 1743–1773 p.
  48. Rossi A, Deveraux Q, Turk B, Sali A. Comprehensive search for cysteine cathepsins in the human genome. *Biol Chem.* 2004;385(5):363–72.
  49. Smooker PM, Whisstock JC, Irving JA, Siyaguna S, Spithill TW, Pike RN. A single amino acid substitution affects substrate specificity in cysteine proteinases from *Fasciola hepatica*. *Protein Sci.* 2000;9(12):2567–72.

50. Coulombe R, Grochulski P, Sivaraman J, Ménard R, Mort JS, Cygler M. Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. *EMBO*. 1996;15(20):5492–503.
51. Turk V, Stoka V, Vasiljeva O, Renko M, Sun T, Turk B, et al. Cysteine cathepsins: From structure, function and regulation to new frontiers. *Biochim Biophys Acta - Proteins Proteomics*. 2012;1824(1):68–88.
52. Cygler M, Sivaraman J, Grochulski P, Coulombe R, Storer AC, Mort JS. Structure of rat procathepsin B: Model for inhibition of cysteine protease activity by the proregion. *Structure*. 1996;4(4):405–16.
53. Vernet T, Berti PJ, De Montigny C, Musil R, Tessier DC, Menard R, et al. Processing of the papain precursor. The ionization state of a conserved amino acid motif within the pro region participates in the regulation of intramolecular processing. Vol. 270, *Journal of Biological Chemistry*. 1995. p. 10838–46.
54. Rozman J, Stojan J, Kuhelj R, Turk V, Turk B. Autocatalytic processing of recombinant human procathepsin B is a bimolecular process. *FEBS Lett*. 1999;459(3):358–62.
55. Caglič D, Pungerčar JR, Pejler G, Turk V, Turk B. Glycosaminoglycans facilitate procathepsin B activation through disruption of propeptide-mature enzyme interactions. *J Biol Chem*. 2007;282(45):33076–85.
56. Dahl SW, Halkier T, Lauritzen C, Dolenc I, Pedersen J, Turk V, et al. Human recombinant pro-dipeptidyl peptidase I (cathepsin C) can be activated by cathepsins L and S but not by autocatalytic processing. *Biochemistry*. 2001;40(6):1671–8.
57. Rzychon M, Chmiel D, Stec-Niemczyk J. Modes of inhibition of cysteine proteases Circled white star. *Acta Biochim Pol*. 2004;51(4):861–73.
58. Ma, Shuhua; Devi-Kesavan, Lakshmi; Gao J. Molecular Dynamics Simulations of the Catalytic Pathway of a Cysteine Protease: A Combined QM/MM Study of Human Cathepsin K. *Chem Rev*. 2008;23(1):1–7.
59. Robinson MW, Tort JF, Lowther J, Donnelly SM, Wong E, Xu W, et al. Proteomics and phylogenetic analysis of the cathepsin L protease family of the helminth pathogen *Fasciola hepatica*: Expansion of a repertoire of virulence-associated factors. *Mol Cell Proteomics*. 2008;7(6):1111–23.
60. Lowther J, Robinson MW, Donnelly SM, Xu W, Stack CM, Matthews JM, et al. The importance of pH in regulating the function of the *Fasciola hepatica* cathepsin L1 cysteine protease. *PLoS Negl Trop Dis*. 2009;3(1):1–11.
61. Robinson MW, Dalton JP, Donnelly S. Helminth pathogen cathepsin proteases: it's a family affair. *Trends Biochem Sci*. 2008;33(12):601–8.

62. Robinson MW, Corvo I, Jones PM, George AM, Padula MP, To J, et al. Collagenolytic activities of the major secreted cathepsin L peptidases involved in the virulence of the helminth pathogen, *Fasciola hepatica*. PLoS Negl Trop Dis. 2011;5(4).
63. Maule AG, Marks NJ. Parasitic flatworms: Molecular biology, biochemistry, immunology and physiology. Parasitic Flatworms: Molecular Biology, Biochemistry, Immunology and Physiology. 2006. 1–448 p.
64. Stack C, Dalton JP, Robinson MW. The phylogeny, structure and function of trematode cysteine proteases, with particular emphasis on the *Fasciola hepatica* cathepsin L family. Adv Exp Med Biol. 2011;712:116–35.
65. Brady CP, Dowd AJ, Tort J, Roche L, Condon B, O'Neill SM, et al. The cathepsin L-like proteinases of liver fluke and blood fluke parasites of the trematode genera *Fasciola* and *Schistosoma*. Biochem Soc Trans. 1999;27(4):740–5.
66. Tort J, Brindley PJ, Knox D, Wolfe KH, Dalton JP. Proteinases and associated genes of parasitic helminths. Adv Parasitol. 1999;43:161–266: <http://www.ncbi.nlm.nih.gov/pubmed/10214692>
67. Norbury LJ. Structure , Function and Evolutionary Studies of *Fasciola* Cathepsin L-like Proteases. Eng Technol. 2008;(August).
68. Irving JA, Spithill TW, Pike RN, Whisstock JC, Smooker PM. The evolution of enzyme specificity in *Fasciola* spp. J Mol Evol. 2003;57(1):1–15.
69. Dalton JP, Neill SO, Stack C, Collins P, Walshe A, Sekiya M, et al. *Fasciola hepatica* cathepsin L-like proteases: Biology, function, and potential in the development of first generation liver fluke vaccines. Int J Parasitol. 2003;33(11):1173–81.
70. McManus DP, Dalton JP. Vaccines against the zoonotic trematodes *Schistosoma japonicum*, *Fasciola hepatica* and *Fasciola gigantica*. Parasitology. 2006;133(SUPPL.2).
71. Dalton MWRJP. Cysteine proteases of pathogenic organisms. Vol. 6, Angewandte Chemie International Edition. 1967. 116–131 p.
72. Carmona C, Dowd AJ, Smith AM, Dalton JP. Cathepsin L proteinase secreted by *Fasciola hepatica* in vitro prevents antibody-mediated eosinophil attachment to newly excysted juveniles. Mol Biochem Parasitol. 1993;62(1):9–17.
73. Dalton JP, Mulcahy G. Parasite vaccines - A reality? Vet Parasitol. 2001;98(1–3):149–67.
74. O'Neill SM, Mills KHG, Dalton JP. *Fasciola hepatica* cathepsin L cysteine proteinase suppresses Bordetella pertussis-specific interferon- $\gamma$  production in vivo. Parasite Immunol. 2001;23(10):541–7.

75. Berasaín P, Goñi F, McGonigle S, Dowd A, Dalton JP, Frangione B, et al. Proteinases secreted by *Fasciola hepatica* degrade extracellular matrix and basement membrane components. *J Parasitol.* 2015;83(1):1–5.
76. Dowd AJ, McGonigle S, Dalton JP. *Fasciola hepatica* Cathepsin L Proteinase Cleaves Fibrinogen and Produces a Novel Type of Fibrin Clot. *Eur J Biochem.* 1995;232(1):241–6.
77. Dowd AJ, Tort J, Roche L, Ryan T, Dalton JP. Isolation of a cDNA encoding *Fasciola hepatica* cathepsin L2 and functional expression in *Saccharomyces cerevisiae*. *Mol Biochem Parasitol.* 1997;88:163–74.
78. Berger A, Schechter I. Mapping the active site of papain with the aid of peptide substrates and inhibitors. *Philos Trans R Soc Lond B Biol Sci.* 1970;257(813):249–64.
79. Chowdhury SF, Joseph L, Kumar S, Tulsidas SR, Bhat S, Ziomek E, et al. Exploring Inhibitor Binding at the S' Subsites of Cathepsin L. *Am Chem Soc.* 2008;1361–8.
80. Harper E, Berger A. On the size of the active site in proteases: Pronase. *Biochem Biophys Res Commun.*
81. Norbury LJ, Beckham S, Pike RN, Grams R, Spithill TW, Fecondo J V, et al. Adult and juvenile *Fasciola* cathepsin L proteases: Different enzymes for different roles. *Biochimie.* 2011;93(3):604–11.
82. Norbury LJ, Hung A, Beckham S, Pike RN, Spithill TW, Craik CS, et al. Analysis of *Fasciola* cathepsin L5 by S2subsite substitutions and determination of the P1-P4 specificity reveals an unusual preference. *Biochimie.* 2012;94(5):1119–27: <http://dx.doi.org/10.1016/j.biochi.2012.01.011>
83. Sivaraman J, Nägler DK, Zhang R, Ménard R, Cygler M. Crystal structure of human procathepsin X: A cysteine protease with the proregion covalently linked to the active site cysteine. *J Mol Biol.* 2000;295(4):939–51.
84. Martínez-Sernández V, Perteguer MJ, Hernández-González A, Mezo M, González-Warleta M, Orbegozo-Medina RA, et al. Comparison of recombinant cathepsins L1, L2, and L5 as ELISA targets for serodiagnosis of bovine and ovine fascioliasis. *Parasitol Res.* 2018;117(5):1521–34.
85. Corvo I, Ferraro F, Merlino A, Zuberbühler K, O'Donoghue AJ, Pastro L, et al. Substrate specificity of cysteine proteases beyond the S2 pocket: Mutagenesis and molecular dynamics investigation of *Fasciola hepatica* Cathepsins L. *Front Mol Biosci.* 2018;5(APR):1–11.
86. Clark DP. NJP. *Biotechnology: Applying the Genetic Revolution. Secunda. Vol. 11, Biotechnology: Second Edition.* 2008. 1–820 p.

87. AA E. The recombinant protein Handbook: Protein Amplification and simple purification. In: amersham pharmacia biotech. 2019. p. 1–108.
88. Tripathi NK, Shrivastava A. Recent Developments in Bioprocessing of Recombinant Proteins: Expression Hosts and Process Development. *Front Bioeng Biotechnol.* 2019;7.
89. Ravi I, Baunthiyal M, Saxena J. Advances in biotechnology. *Adv Biotechnol.* 2013;9788132215:1–259.
90. Rapley R. Recombinant DNA and genetic analysis. Principles and Techniques of Biochemistry and Molecular Biology, Sixth Edition. 2005. 225–291 p.
91. Rosano GL, Ceccarelli EA. Recombinant protein expression in *Escherichia coli*: Advances and challenges. *Front Microbiol.* 2014;5(APR):1–17.
92. Fisher AC, Haitjema CH, Guarino C, Çelik E, Endicott CE, Reading CA, et al. Production of secretory and extracellular N-linked glycoproteins in *Escherichia coli*. *Appl Environ Microbiol.* 2011;77(3):871–81.
93. Riguero V, Clifford R, Dawley M, Dickson M, Gastfriend B, Thompson C, et al. Immobilized metal affinity chromatography optimization for poly-histidine tagged proteins. *J Chromatogr A.* 2020;1629:461505.
94. Joshua A, Bornhorst; Joseph J F. Purification of Proteins Using Polyhistidine Affinity Tags. *Biotechnol to PCR.* 2000;326:245–54.
95. Valenti LE, De Pauli CP, Giacomelli CE. The binding of Ni(II) ions to hexahistidine as a model system of the interaction between nickel and His-tagged proteins. *J Inorg Biochem.* 2006;100(2):192–200.
96. Schoonen L, Van Esterik KS, Zhang C, Ulijn R V., Nolte RJM, Hest JCMV. Alternative application of an affinity purification tag: Hexahistidines in ester hydrolysis. *Sci Rep.* 2017;7(1):2–10.
97. Dubendorf JW, Studier FW. Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor. *J Mol Biol.* 1991;219(1):45–59.
98. Shilling PJ, Mirzadeh K, Cumming AJ, Widesheim M, Köck Z, Daley DO. Improved designs for pET expression plasmids increase protein production yield in *Escherichia coli*. *Commun Biol.* 2020;3(1).
99. Rodrigues CHM, Pires DEV, Ascher DB. DynaMut: Predicting the impact of mutations on protein conformation, flexibility and stability. *Nucleic Acids Res.* 2018;46(W1):W350–5.
100. William RP, An introduction to sequence similarity (“home”) searching. 2013, Institutes natinal of health. doi:10.1002/0471250953.bi0301s42.

101. Weber CC, Whelan S. Physicochemical amino acid properties better describe substitution rates in large populations. *bioRxiv*. 2018;
102. Kucukkal, Tugba; Petukh Marharyta; Li LAE. Structural and Physico-Chemical Effects of Disease and Non- Disease nsSNPs on Proteins. *Physiol Behav*. 2015;176(5):18–24.
103. Creixell P, Schoof EM, Tan CSH, Linding R. Mutational properties of amino acid residues: Implications for evolvability of phosphorylatable residues. *Philos Trans R Soc B Biol Sci*. 2012;367(1602):2584–93.
104. Larentis AL, Nicolau JFMQ, Esteves GDS, Vareschini DT, De Almeida FVR, Dos Reis MG, et al. Evaluation of pre-induction temperature, cell growth at induction and IPTG concentration on the expression of a leptospiral protein in *E. coli* using shaking flasks and microbioreactor. *BMC Res Notes*. 2014;7(1):1–13.
105. Kumar P, Libchaber A. Pressure and temperature dependence of growth and morphology of *Escherichia coli*: Experiments and stochastic model. *Biophys J* 2013;105(3):783–93.
106. Papaneophytou CP, Kontopidis G. Statistical approaches to maximize recombinant protein expression in *Escherichia coli*: A general review. *Protein Expr Purif*. 2014;94:22–32.
107. Quiroga A. Optimización del cultivo de *Escherichia coli* para la producción de cutinasas recombinantes. *Univ Chile*. 2010;2–24.
108. Ma C, Liang K, Tang L, He S, Liu X, He M, et al. Identification and characteristics of a cathepsin L-like cysteine protease from *Clonorchis sinensis*. *Parasitol Res*. 2019;118(3):829–35.
109. Garcia E. Regulation of recombinant protein solubility and conformational quality in *Escherichia coli*. 2009;99–210.
110. Gutiérrez-González M, Farías C, Tello S, Pérez-Etcheverry D, Romero A, Zúñiga R, et al. Optimization of culture conditions for the expression of three different insoluble proteins in *Escherichia coli*. *Sci Rep*. 2019;9(1):1–11.
111. San-Miguel T, Pérez-Bermúdez P, Gavidia I. Production of soluble eukaryotic recombinant proteins in *E. coli* is favoured in early log-phase cultures induced at low temperature. *Springerplus*. 2013;2(1):1–4.
112. Quintas-Granados LI, Orozco E, Brieba LG, Arroyo R, Ortega-López J. Purification, refolding and autoactivation of the recombinant cysteine proteinase EhCP112 from *Entamoeba histolytica*. *Protein Expr Purif*. 2009;63(1):26–32.
113. Ryšlavá H, Doubnerová V, Kavan D, Vaněk O. Effect of posttranslational modifications on enzyme function and assembly. *J Proteomics*. 2013;92:80–109.



114. McVeigh P, Cwiklinski K, Garcia-Campos A, Mulcahy G, O'Neill SM, Maule AG, et al. In silico analyses of protein glycosylating genes in the helminth *Fasciola hepatica* (liver fluke) predict protein-linked glycan simplicity and reveal temporally-dynamic expression profiles. *Sci Rep.* 2018;8(1):1–15.
115. Katunuma N. Posttranslational Processing and Modification of Cathepsins and Cystatins. *J Signal Transduct.* 2010;2010:1–8.
116. Slatko BE, Gardner AF, Ausubel FM. Overview of next generation sequencing technologies (and bioinformatics) in cancer. *Mol Biol.* 2018;122(1):1–15.