



**CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS
AVANZADOS DEL INSTITUTO
POLITÉCNICO NACIONAL**

**UNIDAD ZACATENCO DEPARTAMENTO DE BIOTECNOLOGÍA Y
BIOINGENIERÍA**

**Estudio de la modificación de los giros L4 y L6 de la
chagasina por epítomos de TSA-1 en la solubilidad de las
quimeras recombinantes y su actividad inhibitoria**

Tesis que presenta

Yasmin Irene Rodríguez Gavaldón

Para obtener el grado de

MAESTRO EN CIENCIAS

EN LA ESPECIALIDAD DE BIOTECNOLOGÍA

Director de la Tesis: **Jaime Ortega López**

México, D.F.

Febrero, 2021

La presente investigación se llevó a cabo en el Laboratorio de Biotecnología de proteínas (No. 23) del Departamento de Biotecnología y Bioingeniería (DBB) del Centro de Investigación y de Estudios Avanzados (Cinvestav) del Instituto Politécnico Nacional (IPN), bajo la dirección del Dr. Jaime Ortega López, Investigador Titular del DBB, y con la asesoría de la Dra. María del Carmen Montes Horcasitas, Investigadora Titular en el DBB y del Dr. Edgar Morales Ríos Investigador Titular adscrito al Departamento de Bioquímica del Cinvestav-IPN. Para la realización de este trabajo se contó con el apoyo del Cinvestav-IPN (FIDSC2018/268), del Consejo Nacional de Ciencia y Tecnología (CONACyT, proyectos 269657 y A1-S-34224) y de la Fundación Carlos Slim (SLIM-W03. PO 5600999400). Agradezco al CONACyT por el apoyo económico otorgado para la realización de mis estudios de maestría (Becario CONACyT No. 784949) durante el periodo de septiembre 2018 a agosto 2020.

AGRADECIMIENTOS

Agradezco a mi director de tesis, **Dr. Jaime Ortega López**, por su apoyo, paciencia y guía profesional durante el desarrollo de este proyecto, así como por brindarme todas las facilidades para su ejecución.

También, agradezco a mi comité tutorial conformado por la **Dra. María del Carmen Montes Horcasitas**, y el **Dr. Edgar Morales Ríos** por sus invaluable aportes al proyecto.

Adicionalmente, agradezco a la **Dra. Rossana Arroyo Verástegui** por ayudar con el equipo que fue importante y necesario para culminar los experimentos de tesis. A mis compañeros de laboratorio el **Dr. Gerardo Reséndiz Cardiel** y a la **M. en C. Claudia Ivonne Flores Pucheta**, por compartirme su valioso conocimiento y orientarme y guiarme en la ejecución de este proyecto.

A mis compañeros y amigos **Yeily Adriana Rangel Basto** y **Juan Carlos Sánchez Bermúdez**, por siempre apoyarme y estar conmigo en todo este proceso.

Y finalmente a mis padres, por brindarme su apoyo incondicional y motivarme siempre a seguir adelante, sin ellos nada de esto sería posible.

ÍNDICE

INDICE DE TABLAS	5
INDICE DE FIGURAS	6
GLOSARIO	7
RESUMEN	8
ABSTRACT	9
1. INTRODUCCIÓN	10
1.1 ENFERMEDAD DE CHAGAS	10
1.2 <i>TRYPANOSOMA CRUZI</i>	12
1.2.1 MORFOLOGÍA	12
1.2.2 CICLO DE VIDA	14
1.2.3 MANIFESTACIONES CLÍNICAS	16
1.3 TRATAMIENTO FARMACOLÓGICO	17
1.4 VACUNA TERAPÉUTICA BASADA EN TSA-1	17
1.5 CHAGASINA: INHIBIDOR DE CISTEÍNA PROTEINASAS Y ANDAMIO MOLECULAR DE TSA-1	21
1.5.1 INTERACCIÓN CHAGASINA-CATEPSINA L	23
1.5.2 CHAGASINA COMO ANDAMIO MOLECULAR PARA EXPRESAR TSA-1	27
2. JUSTIFICACIÓN	29
3. HIPÓTESIS	31
4. OBJETIVOS	31
4.1 OBJETIVO GENERAL	31
4.2 OBJETIVOS ESPECÍFICOS	31
5. ESTRATEGIA EXPERIMENTAL	32
6. MATERIALES Y MÉTODOS	33
6.1 PREDICCIÓN DE MODELOS 3D Y SOLUBILIDAD DE LAS QUIMERAS Q1-Q50	33
6.2 SÍNTESIS Y SUBCLONACIÓN DE LA PROTEÍNA QUIMÉRICA MULTIEPÍTOPO Q4	36
6.3 ANÁLISIS EXPRESIÓN Y SOLUBILIDAD DE LA CHAGASINA Y QUIMERAS RECOMBINANTES Q4, Q8, Q20 Y Q45 EN <i>E. COLI</i> SOLUBL21	38
6.4 EXPRESIÓN EN <i>E. COLI</i> SOLUBL21 DE CHAGASINA DE <i>T. CRUZI</i> Y QUIMERAS RECOMBINANTES A UNA ESCALA DE 1 L	39
6.4.1 EXPRESIÓN DE CHAGASINA EN <i>E. COLI</i> SOLUBL21 EN 1 L	39
6.4.2 EXPRESIÓN DE LAS QUIMERAS Q4, Q8, Q20, Q45 EN <i>E. COLI</i> SOLUBL21 EN 1L	40
6.5 PURIFICACIÓN DE CHAGASINA DE <i>T. CRUZI</i> Y QUIMERAS Q4, Q8, Q20 Y Q45 POR MÉTODOS CROMATOGRÁFICOS.	40
6.5.1 PURIFICACIÓN DE CHAGASINA POR CROMATOGRAFÍA DE AFINIDAD A NÍQUEL E INTERCAMBIO IÓNICO.	40
6.5.2 PURIFICACIÓN DE LAS PROTEÍNAS QUIMÉRICAS Q4, Q8, Q20 Y Q45 POR CROMATOGRAFÍA DE AFINIDAD A NÍQUEL.	41
6.6 CONCENTRACIÓN Y CUANTIFICACIÓN DE PROTEÍNAS PURIFICADAS.	42
6.7 ENSAYOS DE ACTIVIDAD INHIBITORIA Y OBTENCIÓN DE LA CONSTANTE DE INHIBICIÓN (K _i) DE LA CHAGASINA Y LAS QUIMERAS RECOMBINANTES	43
7. RESULTADOS	45
7.1 PREDICCIÓN DE MODELOS 3D Y SOLUBILIDAD DE LAS QUIMERAS Q1-Q50	45
7.2 EXPRESIÓN EN <i>E. COLI</i> SOLUBL21 DE CHAGASINA DE <i>T. CRUZI</i> Y QUIMERAS RECOMBINANTES A UNA ESCALA DE 1 L	51
7.2.1 EXPRESIÓN DE CHAGASINA EN <i>E. COLI</i> SOLUBL21	51
7.2.2 EXPRESIÓN DE LAS QUIMERAS Q8, Q20, Q45 Y Q4 EN <i>E. COLI</i> SOLUBL21	53
7.3 PURIFICACIÓN DE CHAGASINA DE <i>T. CRUZI</i> Y QUIMERAS Q4, Q8 Y Q45	55
7.3.1 PURIFICACIÓN DE CHAGASINA POR CROMATOGRAFÍA DE AFINIDAD A NÍQUEL E INTERCAMBIO IÓNICO	55

7.3.2 PURIFICACIÓN DE LA PROTEÍNA QUIMÉRICA Q4, Q8, Q20 Y Q45 POR CROMATOGRAFÍA DE AFINIDAD A NÍQUEL	57
7.4 CONCENTRACIÓN Y CUANTIFICACIÓN DE PROTEÍNAS PURIFICADAS	61
7.5 ENSAYOS DE ACTIVIDAD INHIBITORIA	62
8. DISCUSIÓN	65
8.1 ANÁLISIS DE SOLUBILIDAD	65
8.2 ANÁLISIS DE ACTIVIDAD INHIBITORIA	69
9. CONCLUSIÓN	72
10. PERSPECTIVAS	73
11. BIBLIOGRAFÍA	74
12. APÉNDICE	78
12.1 APÉNDICE 1	78
12.2 APÉNDICE 2	79
12.3 APÉNDICE 3	81
12.4 APÉNDICE 4	83
13. METODOLOGÍAS GENERALES	86
13.1 ALMACENAMIENTO A LARGO PLAZO DE CEPAS BACTERIANAS (GLICEROL)	86
13.2 PREPARACIÓN DE CÉLULAS QUIMIOCOMPETENTES DE <i>E. COLI</i>	87
13.3 TRANSFORMACIÓN DE CÉLULAS QUÍMICAMENTE COMPETENTES DE <i>E. COLI</i>	88
13.4 SELECCIÓN DE CLONAS TRANSFORMANTES DE <i>E. COLI</i> PARA CONFIRMAR LA PRESENCIA DE LA CONSTRUCCIÓN DE INTERÉS.	88
13.5 MEDIOS Y CONDICIONES DE CULTIVO PARA <i>E. COLI</i>	89
13.6 ANÁLISIS DE PROTEÍNAS POR SDS-PAGE	89

INDICE DE TABLAS

TABLA 1. SECUENCIA CONSENSO DE EPÍTOPOS LINEALES DE TSA-1	20
TABLA 2. AMINOÁCIDOS ESENCIALES DE LA CHAGASINA EN LA INTERACCIÓN CON CATEPSINA L.	26
TABLA 3. PARÁMETROS DE CHAGASINA Y QUIMERAS CHAGASINA-EPÍTOPOS DE TSA-1, POTENCIALMENTE SOLUBLES	28
TABLA 4. DISEÑO DE PROTEÍNA QUIMÉRICA Q4.	37
TABLA 5. COMPARACIÓN DE PORCENTAJE DE SOLUBILIDAD DE LOS SERVIDORES CCSOLOMICS, PROTEINSOL, SODOPE, SOLART, CAMSOL Y AGGRESKAN.	48
TABLA 6. CONCENTRACIÓN DE CHAGASINA (CH) Y QUIMERAS Q4, Q8, Q45	61
TABLA 7. COMPARACIÓN DE LA CONSTANTE DE INHIBICIÓN APARENTE (K_{IAPP}) Y LA CONSTANTE DE INHIBICIÓN (K_I) DE LA CHAGASINA Y QUIMERAS RECOMBINANTES SOBRE LA CP CATEPSINA L.	64
TABLA 8. ESQUEMA DE PREPARACIÓN DE LAS REACCIONES DE ACTIVIDAD DE CP CATEPSINA L CON LA CHAGASINA RECOMBINANTE.	79
TABLA 9. ESQUEMA DE PREPARACIÓN DE LAS REACCIONES DE ACTIVIDAD DE CP CATEPSINA L CON LA QUIMERA Q4 RECOMBINANTE.	79
TABLA 10. ESQUEMA DE PREPARACIÓN DE LAS REACCIONES DE ACTIVIDAD DE CP CATEPSINA L CON LA QUIMERA Q8 RECOMBINANTE.	80
TABLA 11. ESQUEMA DE PREPARACIÓN DE LAS REACCIONES DE ACTIVIDAD DE CP CATEPSINA L CON LA QUIMERA Q45 RECOMBINANTE.	80
TABLA 12. PREDICCIÓN BIOINFORMÁTICA DE SOLUBILIDAD DE LAS PROTEÍNAS QUIMÉRICAS MULTIEPÍTOPO. 83	
TABLA 13. SOLUCIÓN DE GLICEROL PARA ALMACENAMIENTO A LARGO PLAZO.	86

INDICE DE FIGURAS

FIGURA 1. CASOS REPORTADOS DE ENFERMEDAD DE CHAGAS EN DIFERENTES PAÍSES NO ENDÉMICOS Y SU TRANSMISIÓN	11
FIGURA 2. ESTADIOS MORFOLÓGICOS DE <i>TRYPANOSOMA CRUZI</i>	13
FIGURA 3. CICLO VITAL DE <i>TRYPANOSOMA CRUZI</i> .	15
FIGURA 4. MODELO TRIDIMENSIONAL DE TSA-1 Y LA UBICACIÓN DE LOS EPÍTOPOS E1-E5.	20
FIGURA 5. ESTRUCTURA DE RESONANCIA MAGNÉTICA NUCLEAR (NMR) DE LA CHAGASINA DE <i>T. CRUZI</i> .	22
FIGURA 6. VISTA DEL MODO DE UNIÓN DE LA CHAGASINA EN COMPLEJO CON CP CATEPSINA L.	24
FIGURA 7. INTERACCIÓN DE GIROS L2, L4 Y L6 DE CHAGASINA CON LA ENZIMA CISTEÍNA PROTEINASA (CATEPSINA L).	26
FIGURA 8. MAPA DEL VECTOR DE EXPRESIÓN PCRI 8A UTILIZADO.	37
FIGURA 9. MODELO DE LA ESTRUCTURA 3D DE LA CHAGASINA Y SECUENCIA DE AMINOÁCIDOS DE LAS QUIMERAS RECOMBINANTES Q4, Q8, Q20 Y Q45.	47
FIGURA 10. REGIONES SOLUBLES E INSOLUBLES EN LAS ESTRUCTURAS 3D DE LA	48
FIGURA 11. ESTRUCTURAS 3D EN PERSPECTIVA IZQUIERDA Y DERECHA DEL SERVIDOR AGGREGAN(A3D) CON REGIONES SOLUBLES E INSOLUBLES DE LAS QUIMERAS RECOMBINANTES	50
FIGURA 12. EXPRESIÓN DE 1L DE CHAGASINA EN <i>E. COLI</i> SOLUBL21. ANÁLISIS POR SDS PAGE AL 16 % DE LA EXPRESIÓN DE CHAGASINA.	52
FIGURA 13. ANÁLISIS POR SDS PAGE 16 % DE LA EXPRESIÓN DE LAS PROTEÍNAS A) Q4, B) Q8, C) Q20 Y D) Q45 EN <i>E. COLI</i> SOLUBL21	54
FIGURA 14. PURIFICACIÓN DE CHAGASINA EN SOLUBL21. ANÁLISIS ELECTROFORÉTICO DE LA PURIFICACIÓN DE CHAGASINA POR AFINIDAD A NÍQUEL E INTERCAMBIO IÓNICO.	56
FIGURA 15. PURIFICACIÓN POR CROMATOGRAFÍA DE AFINIDAD A NÍQUEL DE LA PROTEÍNA Q4.	58
FIGURA 16. PURIFICACIÓN POR CROMATOGRAFÍA DE AFINIDAD A NÍQUEL DE LA PROTEÍNA Q8.	59
FIGURA 17. PURIFICACIÓN POR CROMATOGRAFÍA DE AFINIDAD A NÍQUEL DE LA PROTEÍNA Q45.	60
FIGURA 18. ANÁLISIS SDS PAGE DE LA CHAGASINA, QUIMERA Q4, Q8 Y Q45 CONCENTRADA POR FILTRACIÓN EN MEMBRANA AMICON® ULTRA-4.	61
FIGURA 19. INHIBICIÓN DE LA ACTIVIDAD PROTEOLÍTICA DE CATEPSINA-L DE HÍGADO HUMANO POR LAS PROTEÍNAS RECOMBINANTES CHAGASINA, Q4, Q8 Y Q45.	63
FIGURA 20. MAPA DEL VECTOR DE EXPRESIÓN PET-19B.	78
FIGURA 21. GRÁFICO DE LA RECTA $(V_0/V_1)^{-1}$ VS $[I]$ PARA LA OBTENCIÓN Y CÁLCULO DE $K_I(APP)$ Y K_I DE CHAGASINA.	81
FIGURA 22. GRÁFICO DE LA RECTA $(V_0/V_1)^{-1}$ VS $[I]$ PARA LA OBTENCIÓN Y CÁLCULO DE $K_I(APP)$ Y K_I DE LA QUIMERA Q4.	81
FIGURA 23. GRÁFICO DE LA RECTA $(V_0/V_1)^{-1}$ VS $[I]$ PARA LA OBTENCIÓN Y CÁLCULO DE $K_I(APP)$ Y K_I DE LA QUIMERA Q8.	82
FIGURA 24. GRÁFICO DE LA RECTA $(V_0/V_1)^{-1}$ VS $[I]$ PARA LA OBTENCIÓN Y CÁLCULO DE $K_I(APP)$ Y K_I DE LA QUIMERA Q45.	82

GLOSARIO

aa: Aminoácidos

Q: Quimera

ADN: Ácido desoxirribonucleico

kDa: Kilodalton

μm: Micrómetros

pI: Punto isoeléctrico

DO₆₀₀: Densidad óptica a 600 nm

mM: Milimolar

μM: Micromolar

pM: Picomolar

LB: Luria-Bertani

rpm: Revoluciones por minuto

IPTG: Isopropil-B-D-galactopiranosido

SDS-PAGE: Electroforesis en gel de

poliacrilamida y dodecil sulfato de sodio

PBS: Amortiguador de fosfato salino

PMSF: Fluoruro de fenilmetilsulfonilo

VC: Volumen de columna

TEV: Tobacco Etch Virus

GFP: Proteína verde fluorescente

His₆: Tag de seis histidinas

E-64: trans epoxisuccinilo-L-leucilo-amida (4)

guanidino butano

PDB: Protein Data Bank

Z-Phe-Arg-MCA: Z: carbobenzoxi, AMC: 7-amino-4-metilcumarina

EDTA: Acido etilendiaminotetraacético

RESUMEN

La enfermedad de Chagas es causada por el parásito protozoario *Trypanosoma cruzi*. Esta enfermedad puede ocasionar fallas cardiacas en estadios crónicos y es un problema de salud pública de magnitud global. Debido a la baja eficacia del tratamiento farmacológico en etapas crónicas, actualmente se buscan nuevas alternativas terapéuticas basadas en vacunas de proteínas recombinantes. Ensayos en modelos murinos y caninos han demostrado que el antígeno TSA-1 es buen candidato para el desarrollo de estas vacunas. La producción heteróloga de TSA-1 es eficiente en *E. coli*, pero se expresa en agregados de proteína mal plegada conocidos como cuerpos de inclusión. Como alternativa para evitar la etapa de repliegamiento, se propuso usar un andamio molecular para generar proteínas quiméricas que incluyan los epítomos de TSA-1 y que se expresen mayoritariamente en forma soluble. La chagasina, un inhibidor de cisteína proteinasas (CPs), es un factor de virulencia de *T. cruzi*, se expresa en forma soluble en *E. coli* y es un potencial como andamio molecular. En un estudio previo en el grupo de trabajo se generaron quimeras de la chagasina incorporando epítomos de TSA-1 en los giros L4 y L6. Se sabe que los residuos conservados de L4 y L6 interactúan con las CPs y contribuyen a la inhibición de alta afinidad.

El objetivo de este proyecto fue estudiar el efecto de las modificaciones en las secuencias de los giros L4 y L6 de la chagasina en: a) la solubilidad de las quimeras recombinantes, estimada por métodos bioinformáticos; b) su expresión en la cepa de *E. coli* SoluBL21; c) su purificación por cromatografía y d) la inhibición de la actividad enzimática de la catepsina L por las quimeras recombinantes Q4, Q8, Q20 y Q45.

Se demostró que los servidores SoDoPe, Solart, CamSol y AggresCan son una alternativa recomendable para predecir la solubilidad de proteínas recombinantes a expresarse en *E. coli*. También se observó que la solubilidad de la proteína mejora cuando se expresa en la cepa de *E. coli* SoluBL21 para la expresión de las proteínas a 25 °C y que la modificación en giros L4 y L6 de la chagasina por los epítomos de TSA-1, afecta negativamente la solubilidad y actividad inhibitoria de las quimeras recombinantes.

ABSTRACT

Chagas disease is caused by the protozoan parasite *Trypanosoma cruzi*. This disease can cause heart failure in chronic stages and is a public health problem of global magnitude. Due to the low efficacy of pharmacological treatment in chronic stages, new therapeutic alternatives based on recombinant protein vaccines are currently being sought. Assays in murine and canine models have shown that the TSA-1 antigen is a good candidate for the development of these vaccines. Heterologous production of TSA-1 is efficient in *E. coli* but is expressed in aggregates of misfolded protein known as inclusion bodies. As an alternative to avoid the refolding step, it was proposed to use a molecular scaffold to generate chimeric proteins that include TSA-1 epitopes and that are mostly expressed in soluble form. Chagasin, a cysteine proteinase (CP) inhibitor, is a virulence factor of *T. cruzi*, is expressed in soluble form in *E. coli*, and is a potential molecular scaffold. In previous studies, chagasin chimeras were generated incorporating TSA-1 epitopes in the L4 and L6 turns. Conserved L4 and L6 residues are known to interact with CPs and contribute to high affinity inhibition.

The objective of this project was to study the effect of modifications in chagasin L4 and L6 turn sequences on: a) the solubility of recombinant chimeras, estimated by bioinformatics methods; b) its expression in the *E. coli* strain SoluBL21; c) its purification by chromatography and d) the inhibition of the enzymatic activity of cathepsin L by the recombinant chimeras Q4, Q8, Q20 and Q45.

The SoDoPe, Solart, CamSol and AggresCan servers were shown to be a recommended alternative to predict the solubility of recombinant proteins to be expressed in *E. coli*. It was also observed that the solubility improves, when the strain of *E. coli* SoluBL21 is used for the expression of the proteins at 25 ° C and that the modification in L4 and L6 turns of the chagasin by the TSA-1 epitopes, negatively affects the solubility and inhibitory activity of recombinant chimeras.

1. INTRODUCCIÓN

1.1 Enfermedad de chagas

La enfermedad de Chagas es causada por el parásito protozoario *Trypanosoma cruzi*, la cual se transmite principalmente de forma vectorial por la picadura del insecto de la subfamilia Triatominae del género *Triatoma* conocido coloquialmente como “chinche besucona”. Esta enfermedad es un problema de salud pública y de acuerdo a los datos de la Organización Mundial de la Salud (OMS) afecta alrededor de 7 millones de personas distribuidas por toda América, desde el sur de Estados Unidos hasta Argentina, localizadas mayormente en áreas habitadas por personas con bajos recursos económicos (Murillo, 2018). Inicialmente, la enfermedad de Chagas estaba confinada a las zonas rurales de la Región de las Américas, principalmente América Latina y debido a la mayor movilidad de la población en los últimos decenios, la mayoría de las personas infectadas ha pasado a vivir en entornos urbanos (proceso de urbanización) y la enfermedad se ha ido detectando cada vez más en los Estados Unidos de América, Canadá, en muchos países europeos y en algunos africanos, del Mediterráneo Oriental y del Pacífico Occidental (**Figura 1**).

Actualmente, la enfermedad de Chagas, es clasificada como una de las principales enfermedades tropicales desatendidas (NTDs, por sus siglas en inglés) y se cataloga como una enfermedad crónica silenciosa y potencialmente mortal. En México se estima que 1.1 millones de personas están infectadas y las mayores tasas de incidencia se registraron en Yucatán, Oaxaca e Hidalgo. De acuerdo con la Organización Panamericana de la Salud (OPS), existen un total de 18 áreas endémicas dentro del país ubicadas en el sureste y estas áreas incluyen los estados de Oaxaca, Jalisco, Yucatán, Chiapas, Veracruz, Puebla, Guerrero, Hidalgo y Morelos,

todas ellas con áreas rurales. Así mismo, es la parasitosis de mayor importancia tanto por su morbilidad-mortalidad debido a su impacto socioeconómico con pérdidas de alrededor de 7.2 mil millones de dólares anuales (Ramsey et al., 2003; Dumonteil et al., 2012). Las principales rutas de transmisión en países no endémicos son la transfusión de sangre, el trasplante de órganos, por el consumo de alimentos contaminados con el parásito o por el paso congénito (Dorn et al., 2007).

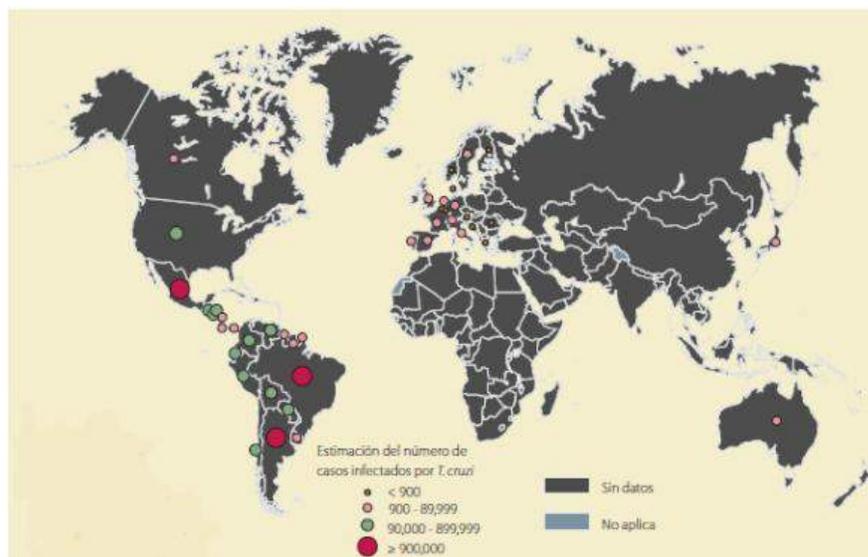


Figura 1. Casos reportados de enfermedad de Chagas en diferentes países no endémicos y su transmisión. En América Latina se considera en riesgo de infección a un mínimo de 110 millones de individuos en 21 países (Argentina, Brasil, Chile, Uruguay, Paraguay, Perú, Ecuador, Bolivia, Venezuela, Colombia, Guyana Francesa, Guyana, Surinam, Costa Rica, El Salvador, Honduras, Nicaragua, Panamá, Belice, Guatemala y México). Los movimientos poblacionales han modificado el perfil epidemiológico de esta enfermedad y la han convertido en un riesgo mundial, y emergente en países no endémicos (Estados Unidos, Canadá, Australia, Japón, Francia, España, Bélgica, Portugal, Suiza, Gran Bretaña, Irlanda del Norte, Italia, Alemania, Austria, Croacia, Dinamarca, Luxemburgo, Noruega, los Países Bajos, Rumania y Suecia) (Organización Mundial de la Salud, 2013).

1.2 *Trypanosoma cruzi*

Trypanosoma cruzi es un parásito flagelado perteneciente a la familia de los tripanosomátidos, incluida en el orden de los cinetoplástidos de la clase Zoomastigina (Deras, 2011).

1.2.1 Morfología

Las diferentes etapas morfológicas de *T. cruzi* se dividen en amastigotes, epimastigotes y tripomastigotes, las cuales son consecuencia de una serie de cambios metabólicos y estructurales en respuesta a los mecanismos adaptativos que el parásito ejerce para infectar a sus diferentes hospederos. Estas etapas se diferencian entre sí principalmente por la posición del cinetoplasto, una estructura exclusiva del orden Kinetoplastida en cuyo interior se localiza el ADN mitocondrial, con relación al núcleo y la presencia o no del aparato flagelar.

Los amastigotes son la forma intracelular del parásito, y se encuentran en los tejidos de los hospederos mamíferos infectados (**Figura 2A**). Poseen un cinetoplasto anterior al núcleo y carecen de flagelo por lo que su movimiento es mayormente rotatorio. Tienen un diámetro aproximado de ~2 a 6.5 μm y se multiplican por fisión binaria longitudinal cada 12 horas, hasta provocar la lisis de las células infectadas. A su vez, los epimastigotes de ~20 a 40 μm se encuentran en el tracto digestivo de los insectos vectores (**Figura 2B**). Su cinetoplasto se localiza en la porción anterior de la célula y son flagelados. Poseen una intensa actividad replicativa por fisión binaria longitudinal. El tercer estadio corresponde a los tripomastigotes, de ~ 17 μm de diámetro (**Figura 2C**). Es la etapa del ciclo vital de *T. cruzi* que carece de

capacidad replicativa, cuentan con flagelo y es la forma extracelular e infecciosa del parásito. Se le encuentra tanto en los insectos vectores como en el hospedero final, en forma de tripomastigotes metacíclicos y tripomastigotes sanguíneos, respectivamente. Al ser flagelados les confiere una gran capacidad de desplazamiento (de Souza et al., 2010).

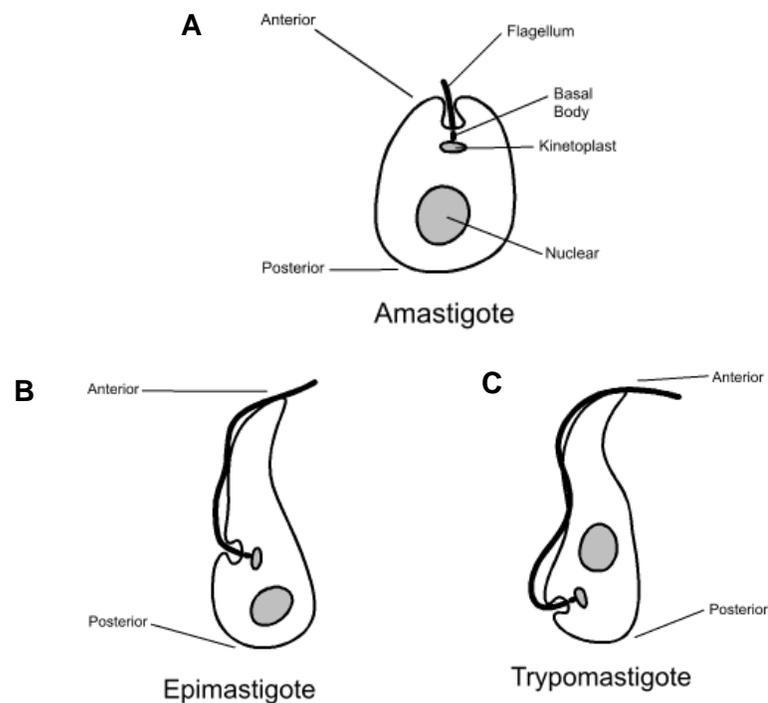


Figura 2. Estadios morfológicos de *Trypanosoma cruzi*. A) Amastigote. B) Epimastigote. C) Trypomastigote. Se diferencian principalmente por la posición del cinetoplasto con relación al núcleo y la presencia o no de la estructura flagelar. (Adaptado de Alba Soto & González Cappa, 2019) .

1.2.2 Ciclo de vida

El parásito es transmitido a los humanos principalmente a través de una gran variedad de especies de artrópodos hematófagos, de hábitos nocturnos e intradomiciliarios que pertenecen a la subfamilia *Triatominae* (triatominos), mejor conocidos como chinches besuconas. Hay más de 100 especies diferentes de triatominos, entre ellas *T. barberi*, *T. dimidiata* y *M. pallidipennis* son considerados como los vectores más comunes en México por su amplia distribución geográfica y capacidad para transmitir al parásito (Ramsey et al., 2015). Los triatominos al momento de alimentarse de sangre expulsan tripomastigotes metacíclicos en sus deyecciones, por lo que los parásitos entran al torrente sanguíneo por acción mecánica por parte del hospedero, que al rascarse arrastra las heces hacia heridas y/o mucosas.

Cuando entran en una célula humana, se convierten en amastigotes. Esta es una etapa reproductiva a través de la mitosis. Después de la reproducción, una gran cantidad de amastigotes se encuentran en la célula infectada, formándose pseudoquistes. El amastigote se convierte de nuevo en tripomastigote sanguíneo y la célula se rompe. El tripomastigote sanguíneo vuelve a infectar otra célula repitiéndose el ciclo de multiplicación. Una vez que ingresa a torrente sanguíneo, y después de presentar cambios morfológicos, *T. cruzi* puede afectar diversos órganos, principalmente corazón, esófago y colon.

Cuando el triatomino pica a un hospedador infectado, algunos tripomastigotes sanguíneos pasan a él a través de la sangre al intestino del insecto, posteriormente se transforman en epimastigotes, los cuales constituyen una segunda etapa reproductiva. Después de la reproducción a través de mitosis, los epimastigotes pasan al recto. Allí se convierten en

tripomastigotes metacíclicos y se evacúan a través de las heces. Las heces pueden infectar a un nuevo hospedador, repitiéndose el ciclo (**Figura 3**) (Moretti et al., 2020).

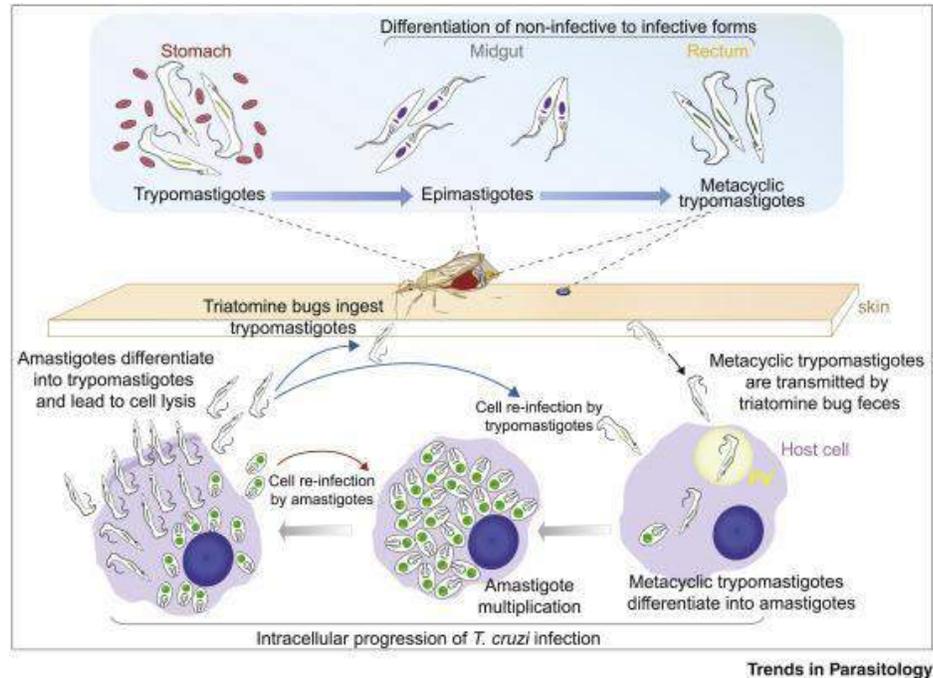


Figura 3. Ciclo vital de *Trypanosoma cruzi*. El triatomino vector infectado libera heces y parásitos de *T. cruzi* cerca del sitio de la picadura. El parásito en forma de tripomastigote metacíclico invade las células, se diferencian en amastigotes y se multiplican por fisión binaria, posteriormente se liberan en forma de tripomastigote sanguíneo para infectar al huésped e ingresan al torrente sanguíneo cuando la enfermedad de Chagas aguda se vuelve crónica. La replicación se reanuda solo cuando los parásitos ingresan a otra célula o son ingeridos por el triatomino al alimentarse de sangre humana que contiene parásitos circulantes. Los tripomastigotes ingeridos se transforman en epimastigotes en el intestino del vector. Los parásitos se multiplican y diferencian en el intestino medio en tripomastigotes metacíclicos infecciosos.

1.2.3 Manifestaciones clínicas

La enfermedad se caracteriza por tres fases: aguda, indeterminada o latente y crónica. El periodo de incubación es de 7 a 14 días (Murgueitio, 2017). Según la OMS, la fase aguda dura aproximadamente dos meses después de contraerse la infección. Durante esa fase aguda circulan por el torrente sanguíneo una gran cantidad de parásitos, pero en la mayoría de los casos no hay síntomas o estos son leves y no específicos. En menos del 50 % de las personas picadas por un triatomino se puede presentar fiebre, dolor de cabeza, agrandamiento de ganglios linfáticos, palidez, dolores musculares, dificultad para respirar, hinchazón y dolor abdominal o torácico.

Durante la fase crónica, los parásitos permanecen ocultos principalmente en el músculo cardíaco y digestivo. Hasta un 30 % de los pacientes sufren trastornos cardíacos y hasta un 10 % presentan alteraciones digestivas (agrandamiento del esófago o del colon), neurológicas o mixtas. Con el paso de los años, la infección puede causar muerte súbita por arritmias o insuficiencia cardíaca progresiva como consecuencia de la destrucción del músculo cardíaco y sus inervaciones.

1.3 Tratamiento farmacológico

Se han ensayado múltiples fármacos para el tratamiento de esta enfermedad, sin embargo, los únicos fármacos que por razones éticas y de eficacia se prescriben son: Nifurtimox [NF] (Lampit®) y Benznidazol [BNZ] (Ragonil®, Rochagan®, Radanil®) (Reyes PA, 2007). Ambos medicamentos son eficaces casi al 100 % para curar la enfermedad si se administran al comienzo de la infección en la etapa aguda. Sin embargo, su eficacia disminuye a medida que transcurre el tiempo de la infección, y las reacciones adversas son más frecuentes en edades avanzadas. Además, presentan otras múltiples limitaciones en el manejo terapéutico como contraindicación durante el embarazo, elevado costo, prolongados regímenes de medicación y múltiples efectos secundarios que resultan en el abandono del tratamiento (Pérez-Molina et al., 2015). Debido a la baja eficacia que ha presentado el tratamiento farmacológico en etapas crónicas y a las reacciones adversas, se han buscado nuevas alternativas terapéuticas basadas en vacunas de proteínas recombinantes.

1.4 Vacuna terapéutica basada en TSA-1

La proteína de superficie de *T. cruzi*, TSA-1 (Trypomostigote Surface Antigen) por sus siglas en inglés, está implicada en el proceso de penetración a las células del hospedero al igual que el antígeno Tc24. Mediante ensayos en modelos murinos y caninos con vacunas de ADN o proteínas recombinantes se ha demostrado que los antígenos TSA-1 y Tc24 son buenos

candidatos para el desarrollo de una vacuna terapéutica contra la enfermedad de Chagas (Dumonteil et al., 2004; Dumonteil et al., 2012; Martinez-Campos et al., 2015) y despiertan la respuesta de linfocitos de personas infectadas (Villanueva-Lizama et al., 2018), mejorando la respuesta inmunitaria del cuerpo.

El antígeno Tc24 se produce en forma soluble, estable y con alto rendimiento en *E. coli* (Biter et al., 2018). Por otro lado, TSA-1, una proteína de 85 kDa perteneciente a la familia de las trans-sialidasas pero sin actividad enzimática (Pereira-Chioccola & Schenkman, 1999), se expresa de manera eficiente en *E. coli* pero en forma de agregados de proteína mal plegada conocidos como cuerpos de inclusión. Estos agregados de proteína mal plegada que se generan *in vivo* durante la producción, limita la expresión de TSA-1 a gran escala debido al paso de replegamiento que se requiere (de la Cruz et al., 2019).

El amino terminal de TSA-1 es el más inmunogénico y capaz de conferir protección contra la infección (Wrightsmann et al., 1994; De la Cruz et al., 2019). En la **Tabla 1** se muestra la secuencia consenso de los 5 principales epítopos antigénicos lineales de TSA-1 en diferentes poblaciones de *T. cruzi* (Knight et al., 2014) y en la **Figura 4** se muestra la estructura 3D de la proteína TSA-1 y la ubicación de sus respectivos epítopos de linfocitos T (E1 a E5).

Tabla 1. Secuencia Consenso de Epítomos Lineales de TSA-1

Epítopo	Secuencia (aa)*	Número de aa	Posición**
E1	KLFPEVIDL	9	TSA-1 ⁽¹⁹²⁻²⁰⁰⁾
E2	IVMGNGTLV	9	TSA-1 ⁽³³⁴⁻³⁴²⁾
E3	RLTEELNTI	9	TSA-1 ⁽⁵²⁶⁻⁵³⁴⁾
E4	FVDYNTIV	9	TSA-1 ⁽⁶¹⁷⁻⁶²⁵⁾
E5	LLYSDDALHL	10	TSA-1 ⁽⁵⁰¹⁻⁵¹⁰⁾

* Los aminoácidos más solubles se encuentran en azul, los neutros en negro y los altamente insolubles en rojo oscuro (Niwa et al., 2009; Chan et al., 2013).

**Los números en paréntesis indican la posición del epítopo en la secuencia de la TSA-1.

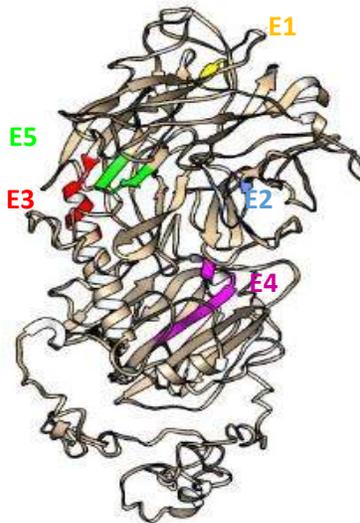


Figura 4. Modelo tridimensional de TSA-1 y la ubicación de los epítomos E1-E5. Los epítomos E1 (amarillo), E2 (azul), E4 (rosa) y E5 (verde) se encuentran localizados en estructuras de hoja β -plegada a excepción del epítopo E3 (rojo) que se encuentra justamente en una α -hélice. El modelo de la estructura 3D de TSA-1 se obtuvo usando el servidor I-Tasser (<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>) (Yang et al., 2015) y visualizado con el software UCSF Chimera 1.13.1 (<https://www.cgl.ucsf.edu/chimera/>) (Pettersen et al., 2004).

1.5 Chagasina: Inhibidor de cisteína proteinasas y andamio molecular de TSA-1

La chagasina de *T. cruzi*, es un inhibidor natural de la superfamilia del Clan CA, de las cisteín-proteinasas (CP) de la Familia C1, y actúa como el principal regulador fisiológico de la cruzipaína, una CP que se expresa durante todo el ciclo vital de *T. cruzi*, y que además actúa como un factor de virulencia cuando el parásito infecta al humano.

La chagasina es una cadena polipeptídica de 110 residuos de aminoácidos (12.440 kDa) sin similitud de secuencia con otros grupos conocidos de inhibidores de las CP. Esta proteína inactiva de manera potente y reversible a enzimas (CP) similares a la papaína, como la cruzipaína del propio parásito o a la catepsina L del huésped humano, mediante la formación de complejos de unión estrecha, en estequiometría 1:1 (Monteiro et al., 2001). Esta proteína junto con inhibidores similares a ella que se encuentran en otros protozoos y bacterias comprenden la familia de las chagasinas (Clan XI - Familia I42) (Rigden et al., 2002; Sanderson et al., 2003).

La chagasina adopta una estructura de β sándwich similar a la inmunoglobulina. Su estructura monomérica está compuesta casi exclusivamente por láminas- β (β 1- β 8) y seis giros (L1-L6) que pueden ser susceptibles de mutaciones específicas que permitan la inserción de los determinantes antigénicos de TSA-1 u otro antígeno. Los giros L1, L3 y L5 son compactos (de 4, 3 y 11 residuos de aminoácidos, respectivamente) y no protruyen hacia el exterior de la proteína; mientras que los giros L2, L4 y L6 se encuentran en el extremo opuesto de la enzima dispuestos de manera secuencial con una alineación casi coplanar. L2 está formado por los

residuos N29-P34, L4 por P59-G68 y L6 está compuesto por los aminoácidos R91-S100, siendo el más grande y formando dos vueltas. Los dos planos de hoja β son casi antiparalelos entre sí. La hélice 3_{10} y cuatro de los seis giros (L1, L2, L3 y L5) se cruzan de una hoja β a la otra. El espacio entre las láminas se llena con una gran cantidad de cadenas laterales hidrófobas (Figura 5).

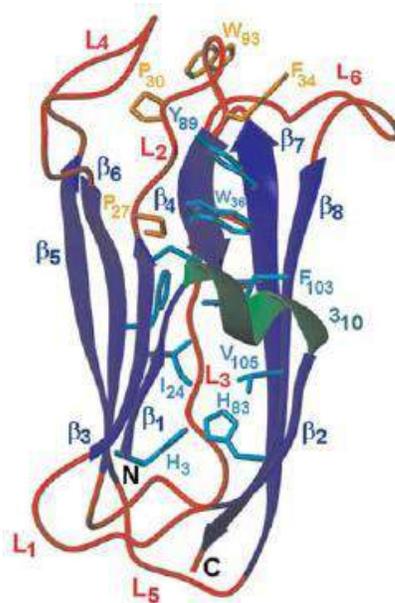


Figura 5. Estructura de resonancia magnética nuclear (NMR) de la chagasina de *T. cruzi*. Láminas β ($\beta_1 - \beta_8$) en azul, 3_{10} -hélice en verde y los giros (L1 – L6) en rojo (Salmon et al., 2006) (PDB: 2NNR).

1.5.1 Interacción chagasina-catepsina L

Las enzimas de la familia C1, como la catepsina L, generalmente funcionan en todas las células como componentes del sistema de degradación lisosomal, participando en la renovación de proteínas. Además, participan en una serie de funciones especializadas, como escisiones proteolíticas que activan pro-hormonas, regulación de presentación de antígeno, etc. (Redzynia et al., 2008).

La estructura del complejo chagasina-catepsina L proporciona detalles de cómo la proteína del parásito inhibe una enzima del huésped de posible importancia en la defensa, ya que la catepsina L es una enzima blanco potencial durante la infección por *T. cruzi* (Ljunggren et al., 2007).

En la **Figura 6A - B** se muestra la chagasina en complejo con la catepsina L. Una cuña inhibitoria compuesta por tres giros L2, L4 y L6, forma varios contactos responsables de la unión de alta afinidad (K_i , 39pM) (Ljunggren et al., 2007) a la enzima. Los tres giros interactúan con la ranura catalítica, con el bucle central L2 insertado directamente en el centro catalítico. Los giros L4 y L6 abarcan la molécula enzimática de ambos lados y exhiben patrones diferentes de reconocimiento proteína-proteína.

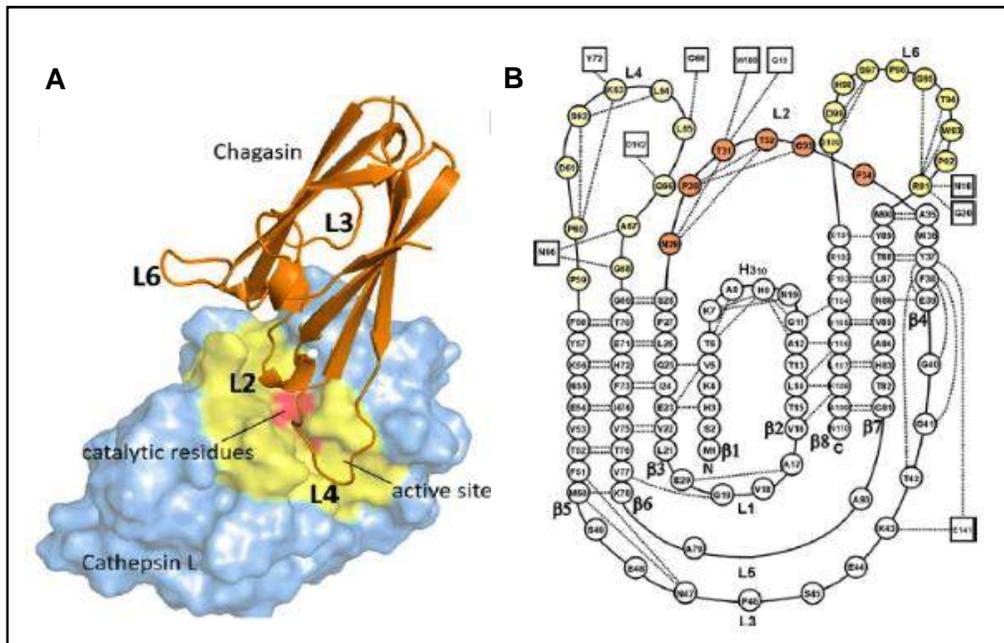


Figura 6. Vista del modo de unión de la chagasina en complejo con CP catepsina L.

A) Representación de los giros de interacción L2, L4 y L6 de chagasina (naranja) en complejo con catepsina L (azul) en su sitio activo (amarillo). **B)** Interacciones del enlace de hidrógeno de la cadena principal con los residuos de catepsina L (cuadrados). Los giros que forman la superficie que interactúa con la catepsina L son de color amarillo claro (L4), naranja (L2) y dorado (L6). Adaptado de Toman et al., 2020 y Ljunggren *et al.*, 2007.

Las interacciones de cada giro con la enzima tienen características diferentes (**Figura 7A-C**). El giro L2 se une al sitio activo de la proteasa, formando enlaces de hidrógeno que involucran los residuos T31 y T32 (**Figura 7A**). Las interacciones en L4 están compuestas por la formación de una hoja β antiparalela de residuos de chagasina L65, G66 y A67 con residuos de la enzima (G67-G68) (**Figura 7B**) (Redzyna *et al.*, 2008). L6 interactúa con la enzima a través de enlaces de hidrógeno (R91), contactos hidrófobos (P92) e interacciones π (W93). Primero, W93 interactúa con un grupo de residuos aromáticos de la enzima (W189 y W193) que sirven para posicionar el elemento N187 de la tríada catalítica de catepsina L (C25-H163-N187). El segundo elemento R91 proporciona comunicación directa entre los giros L6 y L2 formando dos enlaces de hidrógeno con el grupo carbonilo de T32 de L2, que se encuentra junto al residuo de bloqueo del sitio activo T31 (Ljunggren *et al.*, 2007; Redzyna *et al.*, 2008) que además, interactúa con el N18 crucial de la molécula de catepsina L. El tercer elemento de L6, P92, da forma al giro para interacciones óptimas con la enzima al formar contactos hidrófobos con las cadenas laterales de F145 y L144 (**Figura 7C**) y (**Tabla 2**).

Una interacción adicional entre L2 y L6 de chagasina es proporcionada por el grupo carbonilo de la cadena principal de M90 y el átomo de nitrógeno de A35 (Ljunggren *et al.*, 2007).

Se ha reportado que el residuo W93 de chagasina es esencial para la inhibición de catepsina L y que una mutación W93A disminuye 110 veces la afinidad de unión por la enzima. Por otro lado, una doble mutación T31A / T32A en el giro L2 aumenta ligeramente la afinidad de unión por la enzima (4 veces) (dos Reis *et al.*, 2008).

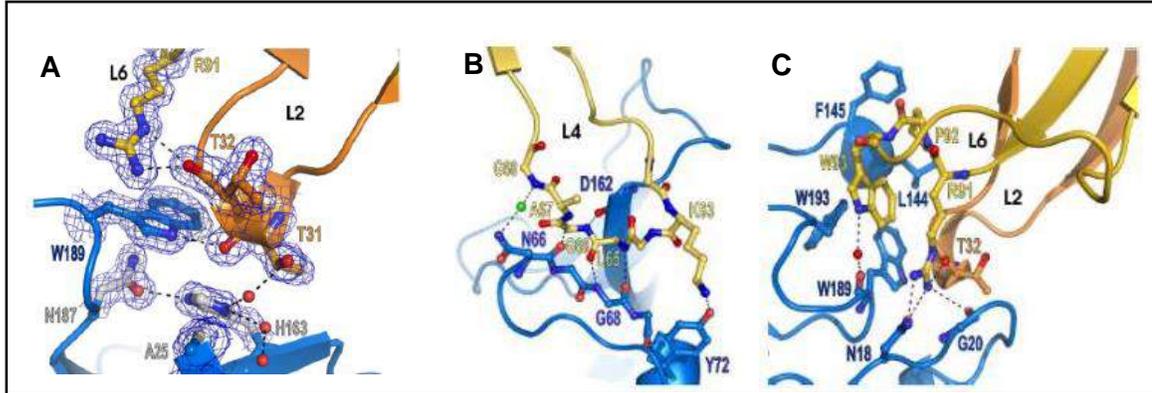


Figura 7. Interacción de giros L2, L4 y L6 de chagasina con la enzima cisteína proteinasa (catepsina L). Detalles de las interacciones enzimáticas de los giros **A) L2**, **B) L4** y **C) L6**. Las cadenas laterales de los residuos cruciales para las interacciones enzimáticas están representadas por palos y bolas. La enzima es de color azul y los giros L4, L2 y L6 que interactúan con la enzima en amarillo claro, naranja y dorado, respectivamente. Adaptado de Ljunggren et al., 2007.

Tabla 2. Aminoácidos esenciales de la chagasina en la interacción con catepsina L.

	Secuencia Aminoácidos					
Chagasina	10	20	30	40	50	60
	MSHKVTKAHN	GATLTVAVGE	LVEIQLPSNP	TTGFAWYFEG	GTKESPNESE	FTVENKYFPP
	70	80	90	100	110	
	<u>DSKL</u> <u>LGAGG</u> T	EHFHVTVKAÄ	GTHAVNLT YM	R <u>PW</u> <u>TG</u> PSHDS	ERFTVYLKAN	

Los residuos subrayados representan los que fueron modificados en el giro L4 y L6. Se señalan de color rojo los esenciales para la interacción con la enzima.

1.5.2 Chagasina como andamio molecular para expresar TSA-1

En un estudio previo en el grupo de trabajo de Nava, 2018 se trabajó en una alternativa para evitar la expresión de TSA-1 en cuerpos de inclusión y, por ende, el paso de replegamiento. Se propuso el uso de la chagasina como potencial andamiaje molecular que se logra expresar en forma soluble en *E. coli* para la inserción de los 5 principales epítopos antigénicos de TSA-1 (Epitope grafting) (Capelli et al., 2017) y producir proteínas quiméricas recombinantes con sus propiedades antigénicas. Las modificaciones se realizaron en los giros L4 y L6 de la chagasina, diseñando 50 quimeras en total. En la **Tabla 3** se indican tres quimeras (Q8, Q20 y Q45) que por el análisis *in silico* presentaron una alta solubilidad para su expresión en *E. coli* (>90%). Solo tres quimeras (Q8, Q20 y Q45) se expresaron en *E. coli* BL21, sin embargo, únicamente la quimera Q8 se expresó en forma soluble y las dos restantes como cuerpos de inclusión (Nava,2018).

Tabla 3. Parámetros de chagasina y quimeras chagasina-epítomos de TSA-1, potencialmente solubles.

Quimera	L4	L6	A.A.	MW (Da)	pI teórico	Coefficiente de extinción molar (M-1 cm-1)	<i>E. coli</i> Solubilidad (%)
Chagasina			110	12039.51	6.58	16960	100
Q8	E5	E2	139	15253.10	6.28	14440	100
Q20	E4	E5	139	15467.26	6.14	15930	99.5
Q45	E1 Y E2	E5	153	16673.77	6.14	14440	62.7

*Adaptada de (Nava, 2018).

El objetivo de este proyecto será estudiar el efecto de las modificaciones en las secuencias de los giros L4 y L6 de la chagasina en la solubilidad de las quimeras recombinantes, estimada por métodos bioinformáticos; su expresión en la cepa de *E. coli* SoluBL21; su purificación por cromatografía y la inhibición de la actividad enzimática de la catepsina L por las quimeras recombinantes.

2. JUSTIFICACIÓN

La enfermedad de Chagas afecta alrededor de 7 millones de personas distribuidas por toda América, desde el sur de Estados Unidos hasta Argentina, localizadas mayormente en áreas habitadas por personas con bajos recursos económicos. Así mismo, es la parasitosis de mayor importancia tanto por su morbilidad-mortalidad debido a su impacto socioeconómico con pérdidas de alrededor de 7.2 mil millones de dólares anuales.

Debido a la baja eficacia que ha presentado el tratamiento farmacológico contra esta enfermedad en etapas crónicas y a las reacciones adversas, se han buscado nuevas alternativas terapéuticas basadas en vacunas de proteínas recombinantes.

Cabe destacar que los altos niveles de expresión y solubilidad de proteínas son dos requisitos principales para la producción exitosa de proteínas recombinantes. Sin embargo, la producción de éstas es un proceso desafiante debido a que el 50 % no se logran expresar y la otra mitad son insolubles (cuerpos de inclusión). Estas fallas obstaculizan la producción e investigación de proteínas, con implicaciones particulares para estudios estructurales, funcionales y farmacéuticos, que requieren soluciones de proteínas solubles y concentradas. Por lo que, la predicción de solubilidad y la ingeniería de proteínas para mejorarla es un área activa de investigación y continuamente se tiene que estar re-evaluando.

Este es el caso del antígeno TSA-1 de *T. cruzi*, el cual ha demostrado ser un buen candidato para el desarrollo de una vacuna contra la enfermedad de Chagas, pero se ha logrado expresar sólo como cuerpos de inclusión.

En un estudio previo en el grupo de trabajo (Nava, 2018) se determinó como una opción potencial una vacuna terapéutica basada en proteínas quiméricas recombinantes que exhibieran en su superficie los mejores determinantes antigénicos inmunoestimulantes de la proteína TSA-1 de *T. cruzi* y, además, que se logaran expresar solubles en *E. coli*. Para ello se utilizó como proteína de andamiaje molecular la chagasina. Posteriormente se insertaron en los giros L4 y L6 de esta proteína combinaciones de 5 diferentes epítomos (E1, E2, E3, E4 y E5) de TSA-1, diseñando 51 proteínas quiméricas recombinantes.

Es por ello, que debido a las modificaciones que se realizaron en los giros L4 y L6 de la chagasina por epítomos de TSA-1 es necesario conocer el efecto que ejerce en su expresión de forma soluble y la manera en que se afecta su capacidad inhibitoria como un excelente inhibidor natural de la cisteína proteinasa catepsina L. Por lo que en este proyecto se propone, hacer un análisis bioinformático detallado con distintos servidores de predicción de solubilidad de las 50 quimeras diseñadas y con base a los datos se seleccionará una de las quimeras con mayor solubilidad para determinar la actividad inhibitoria junto con Q8, Q20 y Q45 previamente expresadas en *E.coli* (Q8 soluble, Q20 y Q45 en cuerpos de inclusión) para conocer la mejor alternativa en el desarrollo de una vacuna contra la enfermedad de Chagas.

3. HIPÓTESIS

Modificaciones en los giros L4 y L6 de la chagasina con epítomos de TSA-1 cambiarán sus propiedades de solubilidad y actividad inhibitoria de la catepsina L humana.

4. OBJETIVOS

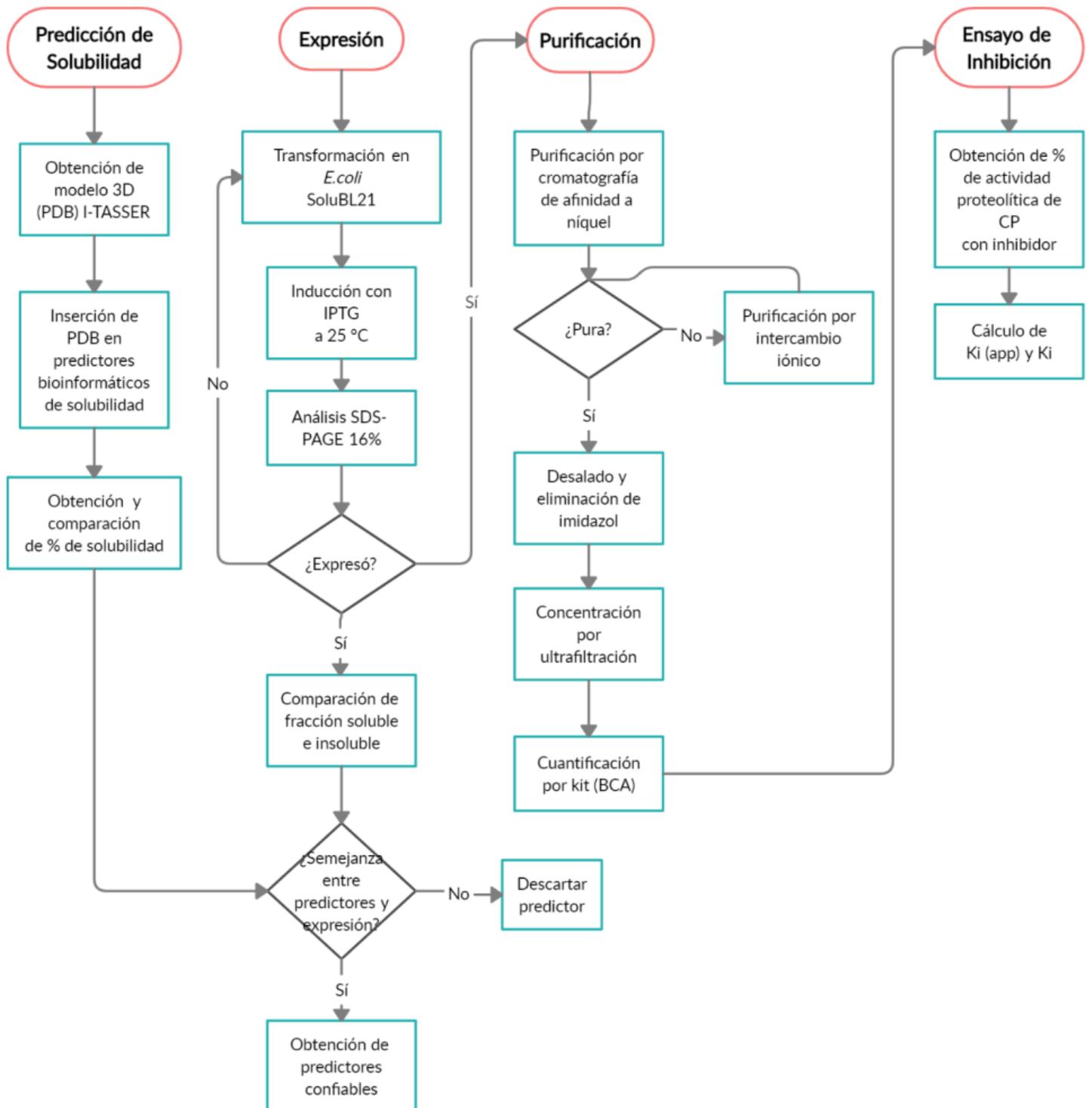
4.1 OBJETIVO GENERAL

Determinar el efecto de las modificaciones en giros L4 y L6 en la chagasina de *T. cruzi* con los epítomos de TSA-1 en la solubilidad y actividad inhibitoria de la catepsina L humana.

4.2 OBJETIVOS ESPECÍFICOS

- Predicción de modelos 3D y solubilidad de las quimeras recombinantes mediante distintos servidores en línea.
- Síntesis química y clonación de la quimera Q4 en vector pCri-8a.
- Expresión y purificación de la chagasina y quimeras recombinantes Q4, Q8, Q20 y Q45 en *E. coli* SoluBL21.
- Determinación de la actividad inhibitoria mediante ensayos de inhibición enzimática de la catepsina L por las quimeras Q4, Q8, Q20 y Q45.

5. ESTRATEGIA EXPERIMENTAL



6. MATERIALES Y MÉTODOS

6.1 Predicción de modelos 3D y solubilidad de las quimeras Q1-Q50

Se realizó la predicción de los modelos tridimensionales de la proteína chagasina nativa y las quimeras Q4, Q8, Q20 y Q45 que se obtuvieron usando el servidor I-TASSER (<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>) (Yang et al., 2015). Los modelos fueron visualizados con el software UCSF Chimera 1.13.1 (<http://www.rbvi.ucsf.edu/chimera>) (Pettersen et al., 2004).

Por otro lado, el análisis predictivo de la solubilidad en *E. coli* de las 51 quimeras recombinantes se determinó mediante distintos servidores como CcSolomics (http://service.tartagliolab.com/update_submission/268986/d32a3118a2), Proteinsol (<https://protein-sol.manchester.ac.uk/>) y SoDoPE (<https://tisigner.com/sodope>) los cuales realizan la determinación de acuerdo a la secuencia de aminoácidos de la proteína y a sus características fisicoquímicas como hidrofobicidad/ hidrofiliidad y hélice alfa para calcular los perfiles de propensión de cada proteína.

Además se utilizó Solart (babylone.ulb.ac.be/SOLART/) y CamSol structural (<http://www-mvsoftware.ch.cam.ac.uk/index.php/camsolstrucorr>) que determinan el puntaje de solubilidad de residuos de acuerdo a su estructura 3D y AggresCan (A3D) 2.0 (<http://biocomp.chem.uw.edu.pl/A3D2/>) que predice la propensión de agregación y solubilidad en estructuras de proteínas.

El servidor ccSolomics emplea un modelo para la predicción de solubilidad de proteínas expresadas en *E. coli* basado en las características fisicoquímicas de los aminoácidos como: espiral/desorden, hidrofobicidad, hidrofiliidad, propensiones de hoja β y hélice α . Estos

parámetros se combinan en una puntuación de propensión de solubilidad dando valores de porcentaje 0 a 100 % (Agostini et al., 2012).

En cuanto a ProteinSol, un servidor de la Universidad de Manchester, es un predictor que procesa la secuencia de aminoácidos y calcula la solubilidad prevista con base a un algoritmo que considera 35 características de secuencia, 20 composiciones de aminoácidos; 7 compuestos: K-R, D-E, K+R, D+E, K+R-D-E, K+R+D+E, F+W+Y; y 8 características adicionales como: longitud, pI, hidropatía (Kyte & Doolittle, 1982), carga absoluta a pH 7, propensión al plegamiento (Uversky et al., 2000), desorden (Linding et al., 2003), entropía de secuencia y propensión a la hoja β (Costantini et al., 2006). Las correlaciones de estos parámetros se registran como solubilidad escalada prevista en un rango de 0 a 1. Se tomó como 100% de solubilidad el valor de 1.

Otro servidor utilizado en este proyecto fue SoDoPE (Soluble Domain for Protein Expression por sus siglas en inglés). Usa un conjunto de 20 valores para los residuos de aminoácidos estándar para puntuar la solubilidad de las proteínas. Estos valores se derivan de los factores B normalizados (Smith et al., 2003) utilizando el algoritmo de Nelder-Mead. La solubilidad de las proteínas se calcula utilizando una métrica llamada SWI (Solubility-Weighted Index por sus siglas en inglés) que van de un valor de 0 a 1. Se tomó el valor de 1 como 100 % de solubilidad. El servidor Solart se basa en potenciales estadísticos, así como en el tamaño de la proteína, área de superficie accesible al solvente y composición de estructura secundaria y terciaria. Proporciona una puntuación de solubilidad escalada con valores cercanos a cero que indican proteínas propensas a agregarse, mientras que valores cercanos a 130 designan proteínas solubles (Hou et al., 2020).

El servidor de Vendruscolo Lab de la Universidad de Cambridge (Método CamSol), utiliza la estructura 3D para calcular el perfil de solubilidad estructuralmente corregido. Predice las regiones de baja solubilidad en la superficie de las proteínas en donde las puntuaciones van de -1.5 a 1.5 tomando en cuenta que valores mayores que 1 (azul) denotan regiones altamente solubles, mientras que las puntuaciones menores que -1 (rojo) son poco solubles (Sormanni *et al.*, 2014).

Por último, el servidor AggresCan3D, integra la información 3D de las estructuras de proteínas y evalúa la contribución de las regiones propensas a la agregación de proteínas (APR por sus siglas en inglés) expuestas a solventes. La propensión a la agregación se calcula para las regiones esféricas centradas en cada residuo de carbono α , utilizando la escala de agregación de aminoácidos intrínsecos del método Aggrescan. Un valor de puntuación mínimo (azul) representa una proteína más soluble en el contexto estructural y un valor de puntuación máximo (rojo) representa una proteína más propensa a la agregación en el contexto estructural (Kuriata *et al.*, 2019) .

6.2 Síntesis y subclonación de la proteína quimérica multiepítopo Q4

Se utilizaron los diseños de las 51 proteínas quiméricas multiepítopo con la sustitución de los aminoácidos de los giros L4 y L6 de la chagasina de *T. cruzi* (ID: Q966X9. Uniprot en www.uniprot.org) por los epítomos E1, E2, E4 y E5 de la proteína TSA-1 (Nava, 2018). Se seleccionó la quimera Q4, la cual contenía los epítomos E5 y E1 para su síntesis y clonación (**Tabla 4**).

El gen codificante se diseñó *in silico*, añadiendo en los extremos 5' el sitio de restricción NcoI y en el extremo 3' el sitio de restricción XhoI. Además, se introdujo el codón de paro TAA para incrementar la eficiencia en la terminación de la traducción en el hospedero. Finalmente, la secuencia del gen fue optimizada para su expresión en *E. coli* y sintetizada por la empresa Synbio Technologies en el vector pUC57, generándose así la construcción pUC57-Q4.

Se transformaron células químicamente competentes de la cepa DH5 α de *E. coli* mediante choque térmico, con la construcción pUC57-Q4 para su propagación y la posterior subclonación del gen en el vector pCri 8a. Este último sería utilizado para la expresión de la proteína de interés, el cual contiene una etiqueta de histidinas (His6-tag) para la purificación por afinidad a níquel, los sitios de restricción NcoI y XhoI, y un codón de paro TAA (**Figura 8**).

Tabla 4. Diseño de proteína quimérica Q4.

	Secuencia de aminoácidos	Secuencia de nucleótidos
Q4		5'ATGGCGAGCCATAAAGTGACCAAAGCGCATAACGGCGGACCTGACCGTGG
E5-E1	MASHKVTKAHNGATLTVAVGELVEIQLPNSNPT	CGGTGGGCGAACTGGTGGAAATTCAGCTGCCGAGCAACCCGACCACCGGCTTT
	TGFAWYFEGG T KESP N ESMFTVENKYFP PLLY	GCCTGGTATTTTGAAGGCGGCACCAAGAAAGCCCGAACGAAAGCATGTTTAC
	SDDALHLGG TEHFHVTVKAAGTHAVNLT YMR	CGTGAAAAACAAATATTTTCCGCCGCTGCTGTATAGCGATGATGCGCTGCATCT
	PKLFPEVIDLS ERFTVYLKAN	GGGCGGCACCGAACATTTTCATGTGACCGTGAAGCGCGGGCACCATGCGG
		TGAACCTGACCTATATGCGCCCGAAACTGTTCCGGAAGTGATTGATCTGAGCG
		AACGCTTTACCGTGTATCTGAAAAGCGAACTAA 3'

Las secuencias de aminoácidos de los giros se encuentran señaladas como: L2 (azul), L3 (morado), L4 (verde) y L6 (naranja). L4 y L6 se encuentran subrayados y los residuos conservados se encuentran sin subrayar.

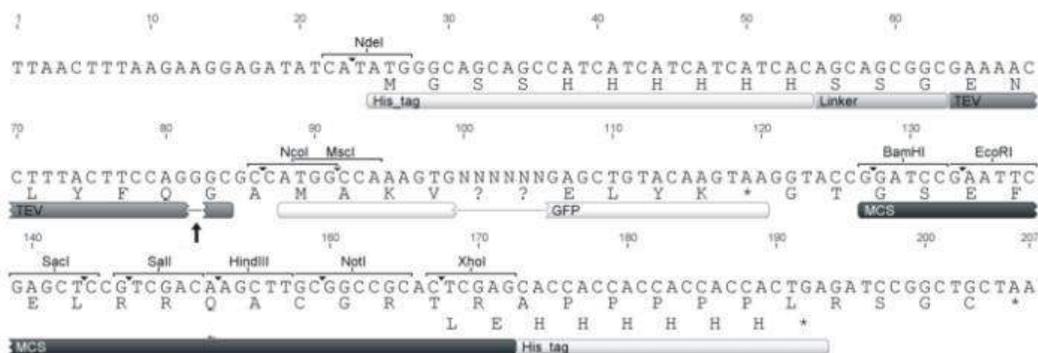


Figura 8. Mapa del vector de expresión pCri 8a utilizado. GFP: proteína verde fluorescente; His6-tag: etiqueta de histidinas; TEV: sitio de escisión de la proteinasa; codón de paro TAA y sitios de restricción NcoI y XhoI. Adaptado de (Goulas et al., 2014).

6.3 Análisis expresión y solubilidad de la chagasina y quimeras recombinantes Q4, Q8, Q20 y Q45 en *E. coli* SoluBL21

El gen de la chagasina que se encuentra clonado en el vector pET19b entre los sitios de restricción NdeI y BamHI (**Apéndice 1**) fue amablemente donado por el Dr. Luis Gabriel Brieba de Castro del Laboratorio Nacional de Genómica para la Biodiversidad (LANGEBIO) en la Unidad Irapuato del Cinvestav-IPN. La chagasina, Q4, Q8, Q20 y Q45 se transformaron en células quimiocompetentes de *E. coli* SoluBL21 por choque térmico para comprobar su expresión, en las construcciones pET19b-chagasina, pCri 8a-Q4, pCri 8a-Q8, pCri 8a-Q20 y pCri 8a-Q45. Después, se inocularon 3 clonas distintas de cada construcción en tubos de ensaye con 5 mL de medio LB más 100 µg/mL de ampicilina para chagasina y 50 µg/mL de kanamicina para las quimeras. Los cultivos se incubaron a 37°C con agitación orbital de 250 rpm hasta que alcanzaron una densidad óptica a 600 nm (DO600) de 0.5-1 y se indujo la expresión de las proteínas con 1 mM de IPTG. El crecimiento se continuó a 25 °C durante 16 h. El análisis de la expresión se realizó por medio de electroforesis en geles de poliacrilamida al 16 %, para lo cual se prepararon muestras de 1 mL del medio de cultivo antes y después de las 16 h de inducción, de tal forma que tuvieron una DO600=1. Cada muestra se centrifugó a 13000 rpm durante 20 min, los sobrenadantes se descartaron y se reservaron los paquetes celulares. Para las muestras denominadas “sin inducir” e “inducida”, los botones celulares se resuspendieron con 75 µL de agua y 25 µL de amortiguador de Laemmli 4x. Por otro lado, el paquete celular restante se lisó con 250 µL de solución BugBuster (Merck) en agitación durante 20 min, después la mezcla se centrifugó y se separó el sobrenadante del sedimento colocándolo en un tubo diferente. Del sobrenadante se tomaron 75 µL que se mezclaron con 25 µL de

amortiguador de Laemmli 4x y se denominó “fracción soluble”. El sedimento se resuspendió con 75 µL de agua y 25 µL de amortiguador de Laemmli 4x y se denominó “fracción insoluble”. Una vez comprobada la expresión, se prepararon glicerolos de cada cultivo para su almacenamiento a -80 °C y posterior expresión en 1 L.

6.4 Expresión en *E. coli* SoluBL21 de chagasina de *T. cruzi* y quimeras recombinantes a una escala de 1 L

6.4.1 Expresión de chagasina en *E. coli* SoluBL21 en 1 L

A partir de construcciones pET19b-chagasina transformadas en células químicamente competentes de *E. coli* SoluBL21, se sembraron en placas de agar LB con 100 µg/mL de ampicilina y se preparó un pre-inóculo de 5 ml, a partir del cual se tomó el volumen necesario para inocular 1 L de medio LB (100 µg/mL de ampicilina) y se incubaron a 37 °C y 250 rpm hasta una DO 600 de 0.5-1, momento en el cual se indujo la expresión con 1 mM de IPTG a 25 °C y 250 rpm por 16 h. La biomasa resultante se recuperó por centrifugación a 8000 rpm durante 30 min. El análisis de la expresión se realizó por geles de electroforesis SDS-PAGE 16 %, en donde se analizaron muestras de cultivo sin inducir, inducida, fracción soluble e insoluble. Posteriormente, la biomasa se llevó a un proceso de lisis para su posterior purificación (apartado 6.5.1).

6.4.2 Expresión de las quimeras Q4, Q8, Q20, Q45 en *E. coli* SoluBL21 en 1L

A partir de clonas de *E. coli* SoluBL21 con las construcciones pCri 8a-Q8, pCri 8a-Q20, pCri 8a-Q45 y pCri-Q4 se prepararon para cada una, un inóculo de 5 mL de medio LB con 50 µg/mL de kanamicina, a partir del cual se tomaron los volúmenes necesarios para inocular 1 L de medio LB (50 µg/mL de kanamicina). El procedimiento se llevó a cabo al igual que en el apartado 6.4.1.

6.5 Purificación de chagasina de *T. cruzi* y quimeras Q4, Q8, Q20 y Q45 por métodos cromatográficos.

6.5.1 Purificación de chagasina por cromatografía de afinidad a níquel e intercambio iónico.

La biomasa obtenida durante la inducción de la expresión de la chagasina se resuspendió a una proporción de 10 mL de amortiguador A (20 mM Tris-HCl H 8.0, 300 mM NaCl, 1 mM PMSF, 5 % glicerol) por cada gramo de biomasa y se lisó con lisozima a una concentración final de 0.1 mg/mL. Posteriormente, la mezcla se incubó a 37°C por 30 min a 250 rpm y se sonicó en baño de hielo durante 10 ciclos de 30 segundos a 30 % de amplitud y 30 segundos de reposo. El lisado se centrifugó a 13000 rpm por 30 min y el sobrenadante se filtró a través de membranas de 0.22 µm. La fracción insoluble se almacenó a -80 °C y la purificación de la chagasina de la fracción soluble se llevó a cabo en el equipo NGC Chromatography System (*Bio-Rad*) en una columna HisTrap FF de 1 mL. La columna se equilibró con 10 VC de amortiguador A y posteriormente la muestra se pasó a un flujo de 1 mL/min y se hicieron lavados con concentraciones crecientes de imidazol: 10 mM, 20 mM, 50 mM y 100 mM de 5 VC cada uno

y la elución de la proteína se llevó a cabo con 5 VC de amortiguador B (20 mM Tris-HCl H 8.0, 300 mM NaCl, 1 mM PMSF, 5 % glicerol y 500 mM imidazol) (Nava, 2018). Las fracciones que mostraron una señal de UV se analizaron por SDS-PAGE al 16 % y las fracciones donde eluyó la chagasina se colectaron y diluyeron en relación 1:10 con amortiguador A2 (20 mM Tris-HCl pH 8.0, 1 mM PMSF y 5 % glicerol) para disminuir la fuerza iónica y someterlas a un segundo paso de purificación por intercambio iónico usando una columna Hitrap QFF de 1 mL (GE-Hetlhcare). La columna se equilibró con 5 VC de regulador A2, la muestra se pasó por la columna a 1 mL/min. La chagasina se eluyó mediante un gradiente lineal de 0-1 M de NaCl con un amortiguador B2 (20 mM Tris-HCl pH 8.0, 1M NaCl, 1 mM de PMSF y 5 % de glicerol) a 1 mL/min. Las fracciones conteniendo la proteína se colectaron y el amortiguador se cambió a PBS 1x con 5 % glicerol. La proteína se almacenó a 4 °C hasta realizar los ensayos de inhibición (Apartado 6.7).

6.5.2 Purificación de las proteínas quiméricas Q4, Q8, Q20 y Q45 por cromatografía de afinidad a níquel.

La biomasa obtenida durante la inducción de las proteínas quiméricas Q4, Q8, Q20 y Q45 se sometió a un proceso de lisis celular para liberar la proteína recombinante al medio. En todos los casos, la biomasa a lizar se resuspendió con 50 mL de amortiguador A (PBS 1x, 5 % de glicerol y 1x SigmaFAST Proteasa Inhibitor Cocktail) por cada gramo de biomasa y se adicionó lisozima a una concentración final de 0.1 mg/mL. Posteriormente, la mezcla se incubó a 37°C por 30 min a 250 rpm y se sonicó en baño de hielo durante 10 ciclos de 30 segundos a 30 % de amplitud y 30 segundos de reposo. El lisado se centrifugó a 13000 rpm durante 30 min

obteniéndose una fase insoluble (que se reservó a -20 °C para su posterior análisis) y una fracción soluble que se filtró a través de una membrana de 0.22 µm de tamaño de poro. Con el filtrado se efectuó un procedimiento de cromatografía de afinidad a níquel en un equipo NGC Chromatography System (*Bio-Rad*) en una columna pre-empacada Profinity IMAC-Ni *Bio-Rad* de 5 mL. La metodología se llevó a cabo según (Nava, 2018), en donde la columna se equilibró con 10 VC del amortiguador A a 1 mL/min, la muestra se aplicó a 0.5 mg/mL y se realizaron lavados a diferentes concentraciones de imidazol de 12.5, 25, 50, 75 y 100 mM para eliminar contaminantes. La elución se realizó con 5 VC de amortiguador B (PBS 1x, 5% de glicerol, 250 mM imidazol y 1x SigmaFAST Proteasa Inhibitor Cocktail) a 1 mL/min. Las muestras que demostraron una señal de UV se analizaron por SDS-PAGE al 16 %. Se colectaron las fracciones de mayor elución de proteína y se almacenaron a 4 °C. El amortiguador se cambió por PBS 1x con 5 % glicerol para su posterior análisis de actividad.

6.6 Concentración y cuantificación de proteínas purificadas.

Previo a cualquier ensayo funcional, la chagasina, Q4, Q8 y Q45 a excepción de Q20, que se lograron purificar eficientemente, se cambiaron a un amortiguador de PBS 1x con 5 % glicerol, y a su vez, se concentró mediante ultrafiltración con Amicon® Ultra-4 Centrifugal Filter Devices de 3 kDa (Sigma-Aldrich) siguiendo el procedimiento recomendado por el proveedor. Finalmente, se determinó la concentración de proteína utilizando un kit de ensayo de proteínas de ácido bicinonínico (BCA) (Pierce) en un espectrofotómetro Multiskan EX (Labsystems) a una absorbancia de 570 nm siguiendo las recomendaciones del fabricante utilizando albúmina de suero bovino (BSA) como estándar y se corrió un gel SDS-PAGE al 16 %

para verificar la integridad y pureza de las proteínas.

6.7 Ensayos de actividad inhibitoria y obtención de la constante de inhibición (K_i) de la chagasina y las quimeras recombinantes

La actividad inhibitoria de cada proteína purificada fue probada usando la enzima catepsina L de humano (h-Cat L) (Sigma-Aldrich), una cisteína proteinasa inhibida por la chagasina, y un sustrato específico para catepsina L, Z-Phe-Arg-MCA (Z-Phe-Arg-7-amido-4-methylcoumarin) que al ser hidrolizado por la enzima libera el 7-amido-4-metilcumarina (AMC) altamente fluorescente. Antes del ensayo enzimático, la enzima fue activada siguiendo las especificaciones del fabricante.

La chagasina fue utilizada en concentraciones finales de 0, 0.005, 0.01, 0.02, 0.1 y 0.2 μM y las quimeras recombinantes Q4, Q8 y Q45 fueron utilizadas a concentraciones de 0.0, 0.5, 1, 2, 4 y 6 μM . Como control positivo de inhibición se utilizó E-64 (E-64: trans epoxisuccinilo-L-leucilo-amida (4) guanidino butano) (Sigma-Aldrich) a una concentración final de 1 μM .

La reacción se llevó a cabo en placas negras de 96 pozos (Nunc) a 37 °C en un amortiguador C (100 mM acetato de sodio pH 5.5, 1 mM EDTA, 50 mM NaCl, 2 mM β -mercaptoetanol) mediante la adición de h-Cat L (0.5 ng/ μL) y el sustrato Z-Phe-Arg-MCA (Sigma-Aldrich) a una concentración final de 10 μM (**Apéndice 2**).

La fluorescencia emitida debido a la liberación de aminometilcumarina (AMC) por la hidrólisis del sustrato se midió a una longitud de onda de excitación de 350 nm y una longitud de onda de emisión de 460 nm, utilizando un lector de microplacas de fluorescencia Gemini EM Fluorescence Microplate Reader. Las Unidades Relativas de Fluorescencia (RFU) se graficaron contra el tiempo para posteriormente determinar el porcentaje de actividad relativa.

La constante de inhibición aparente ($K_{i(app)}$) de cada proteína se determinó midiendo las velocidades de hidrólisis del sustrato en ausencia del inhibidor (V_o) y la velocidad de hidrólisis en presencia de concentraciones variables del inhibidor [V_i] a una concentración de sustrato y enzima fija. Posteriormente, se realizó una gráfica de $(V_o/V_i)-1$ vs $[I]_o$ con una pendiente de $1/K_{i(app)}$. Los valores de $K_{i(app)}$ se calcularon a partir de la ecuación (1) y los valores de K_i a partir de la ecuación (2) (Beynon R & Bound JS, 1982), utilizando una K_m de 2.8 μ M reportada (Brömme et al., 1989) (**Apéndice 3**) como la especificidad de la catepsina L para la hidrólisis de sustratos de péptido metilcumarilamida a pH 5.5.

$$\frac{V_o}{V_i} = 1 + \frac{[I]_o}{K_{i(app)}} \quad (1)$$

$$K_i = K_{i(app)} / \left(1 + \frac{[S]_o}{K_m}\right) \quad (2)$$

7. RESULTADOS

7.1 Predicción de modelos 3D y solubilidad de las quimeras Q1-Q50

Se determinaron los modelos 3D de la chagasina y las proteínas quiméricas multiepítopo diseñadas previamente. Se predijo la solubilidad de las 50 quimeras con distintos servidores (**Apéndice 4**) y se seleccionó a la quimera Q4 por la elevada probabilidad de expresión en forma soluble en *E. coli*. Para los fines de la presente investigación sólo se trabajó con las 3 quimeras Q8, Q20 y Q45 que se generaron en el trabajo previo de Nava, 2018 para su comparación, expresión, purificación y determinación de la actividad (**Figura 9**). Se determinaron los modelos 3D con el servidor I-TASSER de la chagasina y las proteínas quiméricas multiepítopo Q4, Q8, Q20 y Q45 diseñadas previamente.

Se seleccionó el modelo que tuviera un puntaje de confianza de 100 y un porcentaje de identidad mayor (>70 %). El puntaje de confianza es un parámetro que indica la probabilidad (de 9 a 100) que la secuencia y el modelo sean homólogos mientras que el otro parámetro indica el porcentaje de identidad entre la secuencia y el modelo.

En cuanto a las predicciones de solubilidad con base en la secuencia de aminoácidos de chagasina y las proteínas quiméricas multiepítopo Q4, Q8, Q20 y Q45 se realizaron en los servidores ccSolomics, ProteinSol y SoDoPE. Por otro lado, la predicción en Solart, CamSol (structural) y AggresCan (A3D) 2.0 se llevaron a cabo con el modelo estructural (pdb) obtenido anteriormente en I-TASSER. Estos resultados se obtuvieron en porcentajes para fines de comparación.

Como se muestra en la **Tabla 5**, los resultados de CcSolomics sugieren que Q4 y Q8 de las 5 proteínas quiméricas evaluadas tienen la mayor probabilidad de expresión soluble (72 % y 79 %), por el contrario, la que tiene menos probabilidad de ser soluble en la plataforma bacteriana es la chagasina (47 %).

Por otro lado, y en contraste con ccSolomics, el servidor ProteinSol predice que la chagasina y la quimera Q20 son las más solubles con porcentajes de solubilidad de 60 % y 57 % respectivamente, y como la menos propensa a ser soluble la Q45 con un 55 %.

Finalmente, los resultados de los servidores SoDoPE, Solart, CamSol y AggresCan coinciden mayoritariamente en que la chagasina y la quimera Q4 son los más propensos a expresarse de forma soluble en *E. coli* mostrando los valores más altos de porcentaje, seguido de Q8 con un porcentaje menor, después Q45 y finalmente Q20 como el menos propenso a expresarse de forma soluble.

En la **Figura 10** se muestran los modelos tridimensionales de cada quimera, creados por el servidor CamSol (con los modelos estructurales (pdb) obtenidos previamente en I-TASSER) con las regiones de los residuos más solubles en color azul y los residuos insolubles en color rojo. Estos modelos no presentaron una gran diferencia de estructura tridimensional con excepción de Q20 y Q45, sobre todo en las regiones en donde se encuentra el giro L4. La quimera Q20 presenta residuos mayoritariamente insolubles (rojo) en el giro L4 con la inserción del epítipo E4. Por otro lado, la quimera Q45 que presenta dos epítopos (E1 y E2) en el giro L4, cambia su estructura tridimensional pero no logra mostrar residuos insolubles (rojo) como Q20.

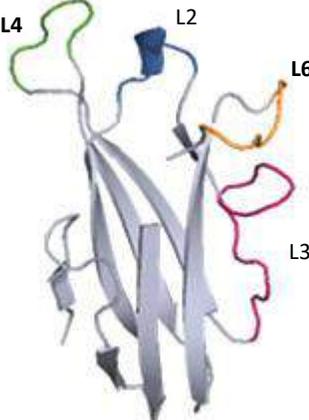
 <p>Chagasina MSHKVTKAHNGATLTVAVGEL VEIQLPSNPTTGFAWYFEGGK ESPNESMFTVENKYFPPDSKLL GAGGTEHFHVTVKAAGTHAVN LTYMRPWTGPSHDSERFTVYLK AN</p>	Q4 E5-E1 MGSSHHHHHSSGENLYFQGA MASHKVTKAHNGATLTVAVGEL VEIQLPSNPTTGFAWYFEGGKESPNESMFTVENKYFPP LLYSDDA LHL <u>GG</u> TEHFHVTVKAAGTHAVNLTYM <u>RP</u> KLFPEVIDL <u>S</u> ERFTVYLK ANLE.
	Q8 E5-E2 MGSSHHHHHSSGENLYFQGA MASHKVTKAHNGATLTVAVGEL VEIQLPSNPTTGFAWYFEGGKESPNESMFTVENKYFPP LLYSDDA LHL <u>GG</u> TEHFHVTVKAAGTHAVNLTYM <u>RP</u> IVMGNGLV <u>S</u> ERFTVYL KANLE.
	Q20 E4-E5 MGSSHHHHHSSGENLYFQGA MASHKVTKAHNGATLTVAVGEL VEIQLPSNPTTGFAWYFEGGKESPNESMFTVENKYFPP FVDYNF TIV <u>GG</u> TEHFHVTVKAAGTHAVNLTYM <u>RP</u> LLYSDDALHL <u>S</u> ERFTVY LKANLE.
	Q45 E1 Y E2- E5 MGSSHHHHHSSGENLYFQGA MASHKVTKAHNGATLTVAVGEL VEIQLPSNPTTGFAWYFEGGKESPNESMFTVENKYFPP KLFPEVI DLGPGPGIVMGNGLV <u>GG</u> TEHFHVTVKAAGTHAVNLTYM <u>RP</u> LLY SDDALHL <u>S</u> ERFTVYLKANLE

Figura 9. Modelo de la estructura 3D de la chagasina y secuencia de aminoácidos de las quimeras recombinantes Q4, Q8, Q20 y Q45. Se muestran con los epítomos de TSA-1 en los giros L4 (aminoácidos en verde) y L6 (aminoácidos en naranja). Los aminoácidos subrayados son los que se mantuvieron conservados de los giros de L4 y L6 de la chagasina (PDB: 2NNR)

Tabla 5. Comparación de porcentaje de solubilidad de los servidores ccSOLomics, ProteinSol, SoDoPE, Solart, CamSol y AggresCan.

	ccSOLomics %	ProteinSol %	SoDoPE %	Solart %	CamSol %	AggresCan %
Chagasina	47.0	60	80	83.8	68.1	100.0
Q4 (E5-E1)	72.0	56	72	83.8	64.8	60.0
Q8 (E5-E2)	79.0	56	69	81.7	60.4	52.5
Q20 (E4-E5)	53.0	57	66	79.6	42.7	39.7
Q45 (E1 y E2-E5)	67.0	55	67	80.4	60.0	41.9

Los números en negritas representan los valores con mayor porcentaje de solubilidad.

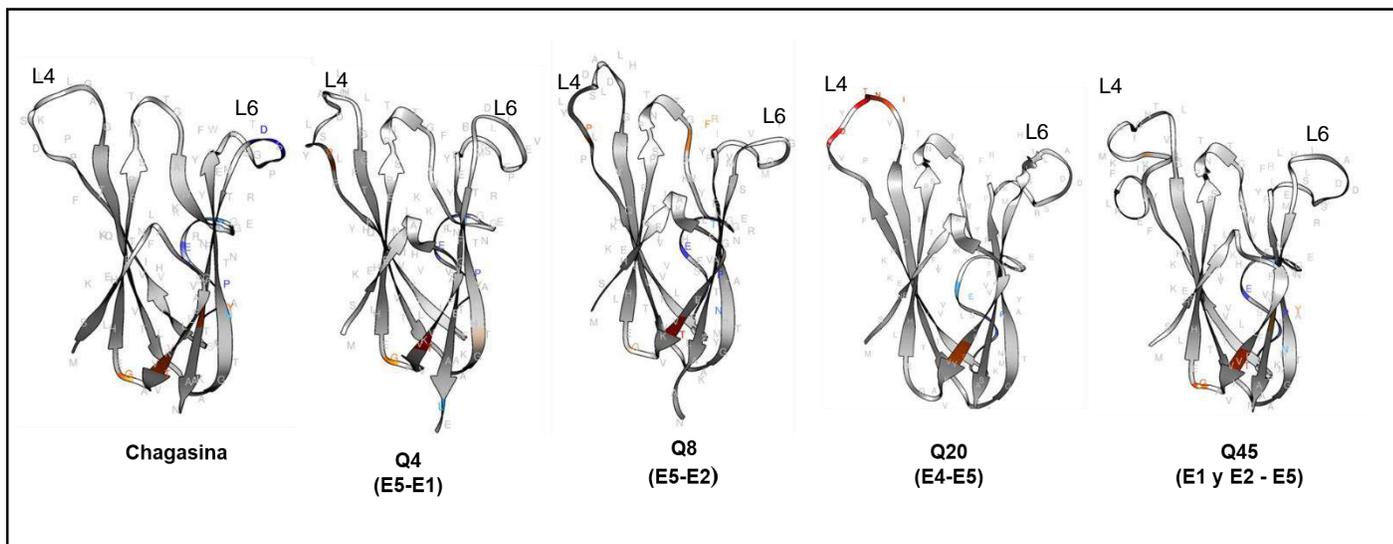


Figura 10. Regiones solubles e insolubles en las estructuras 3D de la chagasina y de las quimeras Q4, Q8, Q20 Y Q45 obtenidas usando el servidor CamSol (Structural). Se usaron los archivos en formato pdb de la chagasina, Q4, Q8, Q20 y Q45 para predecir la solubilidad usando al servidor CamSol (Structural). Las zonas de color rojo representan los residuos que tienden a ser insolubles y de color azul los que tienden a ser solubles.

En la **Figura 11**, se muestran los modelos tridimensionales con la perspectiva I (izquierda) y D derecha de cada quimera, mostrado por el servidor AggresCan (A3D) 2.0 con las regiones de los residuos más solubles en color azul y los residuos insolubles en color rojo. Las regiones más solubles e insolubles varían entre cada quimera, destacando la chagasina como la proteína con mayor coloración azul en los giros L4 y L6, seguido de Q4 y Q8, enfatizando ligeramente Q4 con mayores zonas solubles en L4. La quimera Q45 muestra zonas mayoritariamente insolubles en ambos giros con coloraciones rojas tenues en L4. Finalmente, la quimera Q20 muestra muy pocas zonas solubles (azules) y las zonas insolubles (rojas) se enfatizan en los giros L4.

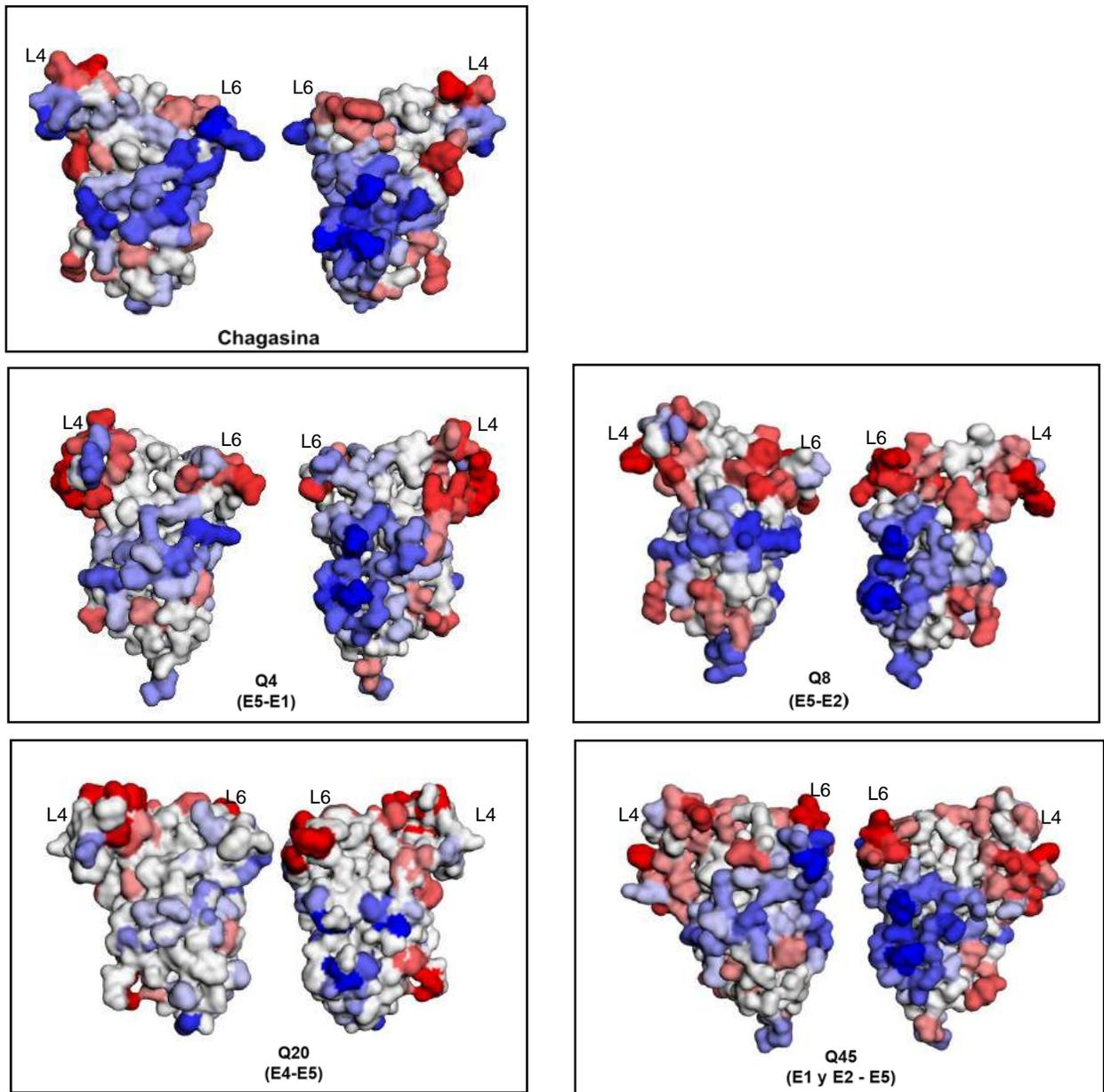


Figura 11. Estructuras 3D en perspectiva izquierda y derecha del servidor AggresCan(A3D) con regiones solubles e insolubles de las quimeras recombinantes. Se usaron los archivos en formato pdb de la chagasina, Q4, Q8, Q20 y Q45 para predecir la solubilidad usando el servidor AggresCan3D. En la escala de colores, el azul oscuro indica que la estructura estática es la más soluble y el rojo oscuro indica que la estructura estática es la más propensa a la agregación en donde las zonas de color rojo representan los residuos que tienden a ser insolubles y de color azul los que tienden a ser solubles.

7.2 Expresión en *E. coli* SoluBL21 de chagasina de *T. cruzi* y quimeras recombinantes a una escala de 1 L

7.2.1 Expresión de chagasina en *E. coli* SoluBL21

En la **Figura 12** se muestra el análisis de expresión y solubilidad de la chagasina producida en la cepa *E. coli* SoluBL21. En la fracción sin inducir (SI) no se observó expresión basal. Por el contrario, si se logró una sobreexpresión después de la inducción con IPTG (I). Para saber si la proteína se expresó correctamente plegada o formando cuerpos de inclusión, se analizaron las fracciones soluble (FS) e insoluble (FI) del extracto proteico, determinándose que la chagasina se acumula en ambas fracciones, pero mayoritariamente en la fracción soluble.

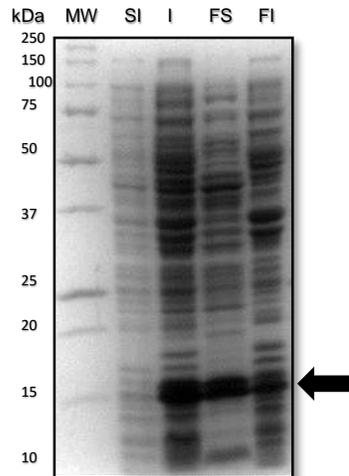


Figura 12. Expresión de 1L de chagasina en *E. coli* SoluBL21. Análisis por SDS PAGE al 16 % de la expresión de chagasina. El cultivo se incubó a 37°C y 250 rpm, y a una DO600 de 0.5-1 se adicionó 1 mM IPTG. Se mantuvo el cultivo durante 16 h a 25°C. La flecha señala las bandas de las muestras de extracto total antes de inducir (SI: Sin inducir), después de la adicción de IPTG (I: inducido), fracción insoluble (FI) y fracción soluble (FS).

7.2.2 Expresión de las quimeras Q8, Q20, Q45 y Q4 en *E. coli* SoluBL21

Con el análisis electroforético se determinó que Q4, Q8, Q20 y Q45 se expresan en forma soluble en la cepa SoluBL21, pero en diferentes proporciones.

En la **Figura 13A** se muestra el análisis de expresión y solubilidad de la quimera Q4. En la fracción sin inducir (carril SI) no se observó expresión basal y la banda de la fracción inducida (carril I) muestra una sobreexpresión después de la adición del IPTG. Las fracciones soluble (carril FS) e insoluble (carril FI) del extracto proteico mostraron que Q4 se produce totalmente en forma soluble. En el análisis de Q8 se visualiza que también se acumula mayoritariamente en forma soluble que de forma insoluble, pero su expresión se muestra ligeramente menor que Q4 (**Figura 13B**). Por otra parte, en el análisis de Q20 (**Figura 13C**) presenta una banda de menor intensidad en la fracción soluble muy similar a la fracción insoluble, lo que sugiere que Q20 tiene una menor sobreexpresión y además se genera mitad en forma soluble y mitad en insoluble. Q45 también mostró bandas similares entre la fracción soluble y la fracción insoluble (**Figura 13D**), pero con una mayor sobreexpresión. Los resultados también indican que la mitad de la expresión es en forma soluble y la otra mitad en cuerpos de inclusión. Todas las quimeras se observaron a un tamaño de ~15 kDa.

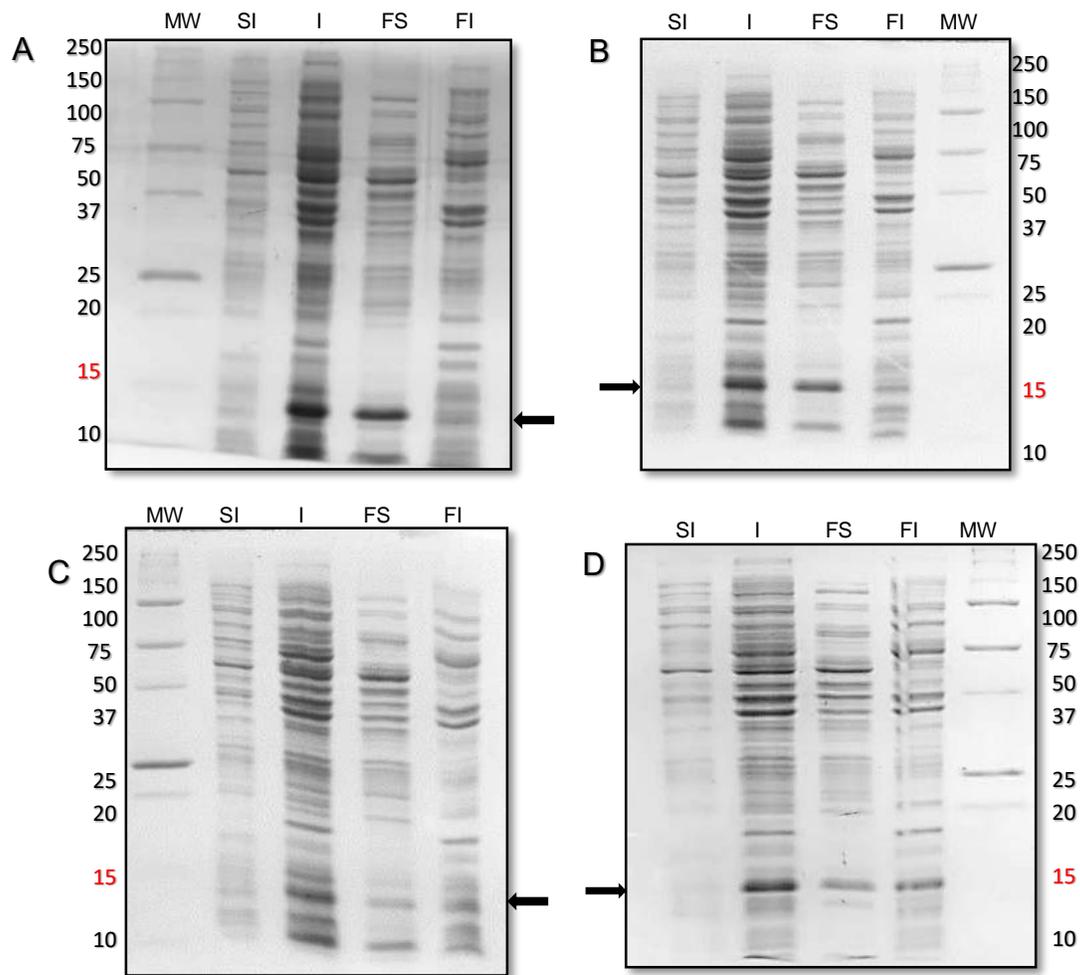


Figura 13. Análisis por SDS PAGE 16 % de la expresión de las proteínas A) Q4, B) Q8, C) Q20 y D) Q45 en *E. coli* SoluBL21. Las clonas de *E. coli* SoluBL21 transformadas con pCri 8a-Q4, pCri 8a-Q8, pCri 8a-Q20 y pCri 8a-Q45 se inocularon en 1 L de LB (50 µg/mL de kanamicina). Los cultivos se incubaron a 37 °C y 250 rpm, y a una DO600 de 0.5-1 se adicionó 1 mM IPTG. Posteriormente se cultivó a 25 °C por 16 h. La flecha señala las bandas de las muestras de extracto total antes de inducir (SI: Sin inducir), después de la adición de IPTG (I: inducido), fracción insoluble (FI) y fracción soluble (FS) de cada quimera.

7.3 Purificación de chagasina de *T. cruzi* y quimeras Q4, Q8 y Q45

7.3.1 Purificación de chagasina por cromatografía de afinidad a níquel e intercambio iónico

En la **Figura 14A** se muestra el análisis en SDS-PAGE 16 % de la muestra no unida (NU) y las muestras eluidas (F3-F7) con 50 mM de imidazol en la purificación por afinidad a níquel. Para una mejor purificación estas fracciones se colectaron y se llevaron a una purificación por intercambio iónico (**Figura 14B**), donde las muestras eluidas con 100-250 mM de NaCl (F35-F45) se colectaron para determinar su concentración y se analizaron por SDS 16 %.

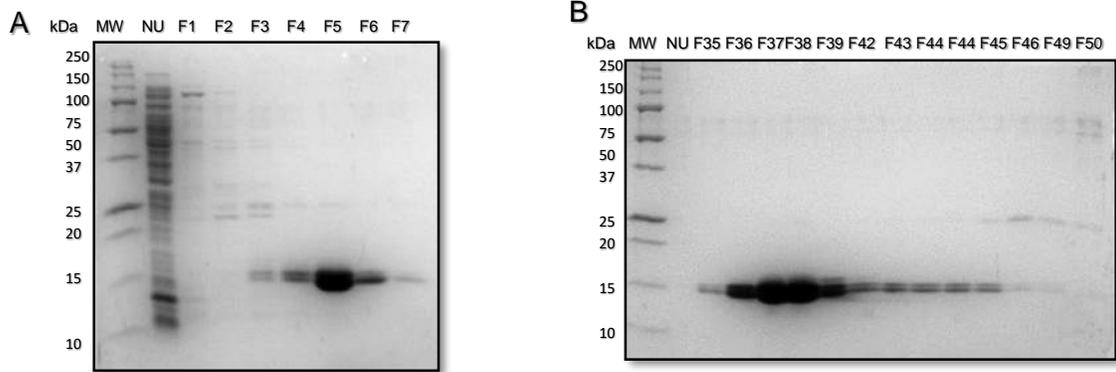


Figura 14. Purificación de chagasina en SoluBL21. Análisis electroforético de la purificación de chagasina por afinidad a níquel e intercambio iónico. A) Análisis por SDS PAGE al 16% de fracciones representativas de la purificación por cromatografía de afinidad a níquel. MW: Marcador de peso molecular. NU: Fracción no unida. F1-F7: Fracciones eluidas analizadas. **B)** Análisis SDS PAGE de fracciones representativas de la purificación por cromatografía de intercambio iónico. MW: Marcador de peso molecular. NU: Fracción no unida. F35-F50: Fracciones eluidas analizadas.

7.3.2 Purificación de la proteína quimérica Q4, Q8, Q20 y Q45 por cromatografía de afinidad a níquel

En la **Figura 15A** se muestra el perfil cromatográfico de la purificación de Q4 en donde se muestra la absorbancia a 280 nm (mAu; línea azul), la fuerza iónica (mS/cm²; línea roja) y concentración de imidazol (%B; línea negra) en función del volumen (mL). Se analizaron en un gel SDS-PAGE, las fracciones que corresponden a los picos con mayor señal de absorbancia que se muestra en la (**Figura 15B**) El primer pico corresponde a la elución de 12.5 mM de imidazol, el segundo a 25 mM, el tercero a 50 mM y el cuarto a 75 mM con las fracciones (F15-F16), (F19-F20), (F21-F31), (F33-F34), respectivamente.

El perfil cromatográfico de la purificación de Q8 se muestra en la **Figura 16A** y el análisis SDS-PAGE de las fracciones de los primeros 3 picos con mayor señal de absorbancia se muestran en la **Figura 16B**, las cuales corresponden a 12.5, 25 y 50 mM de imidazol a las fracciones (F35-F36), (F50-F51) y (F73-F81), respectivamente.

En la **Figura 17A** se muestra el perfil cromatográfico de la purificación de Q45. Los primeros 3 picos de mayor señal de absorbancia se analizaron en un gel SDS-PAGE 16 (**Figura 17B**), los cuales corresponden a la elución con 12.5, 25 y 50 mM de imidazol y corresponde a las fracciones corresponden a (F16), (F20) y (F23-F31), respectivamente.

Los resultados de todos los gels SDS-PAGE de electroforesis mostraron que a 12.5 y 25 mM de imidazol se eluye solamente contaminantes de expresión y que a partir de 50 Mm comienza a eluir la proteína de interés.

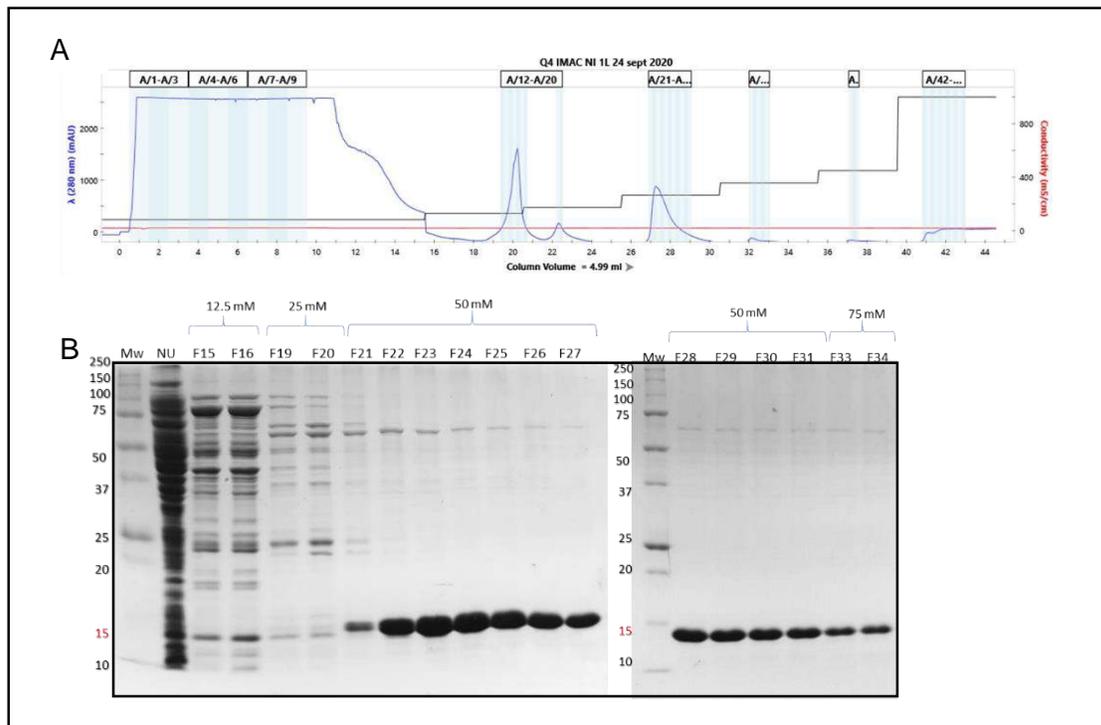


Figura 15. Purificación por cromatografía de afinidad a níquel de la proteína Q4.

Los extractos proteicos de biomasa de Q4 obtenidos después de 16 h de inducción se aplicaron a una columna Bio Scale Profinity IMAC de 5 ml (VC) y después de una etapa de lavado con el amortiguador A (PBS 1x, 5% de glicerol, 1 X SigmaFAST Proteasa Inhibitor Cocktail) la proteína unida a la columna se eluyó con diferentes concentraciones de imidazol. **A)** Perfil cromatográfico de Q4. **B)** Análisis SDS PAGE de los picos de elución de Q4 a diferentes concentraciones de imidazol (12.5, 25, 50 y 75 mM). MW: Marcador de peso molecular. NU: Fracción no unida. F15-F16: Fracciones eluidas a 12.5 mM. F19-F20: Fracciones eluidas a 25 mM. F21-F31: Fracciones eluidas a 50 mM. F33-F34: Fracciones eluidas a 75 mM.

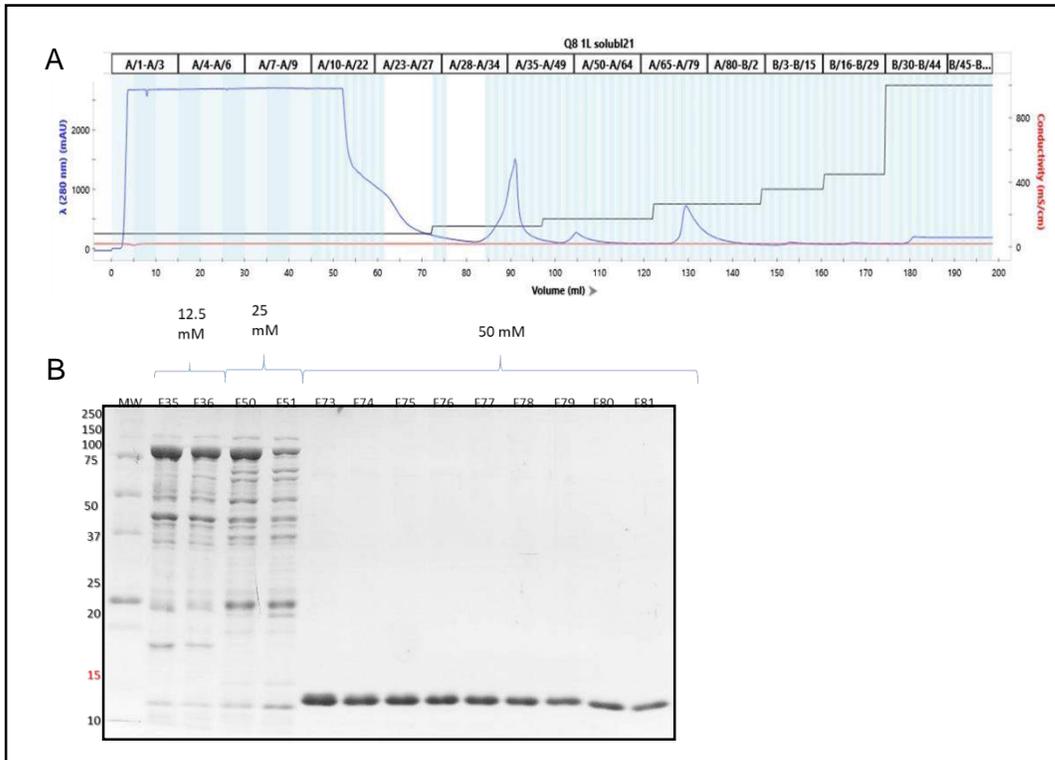


Figura 16. Purificación por cromatografía de afinidad a níquel de la proteína Q8.

Los extractos proteicos de biomasa de Q8 obtenidos después de 16 h de inducción se aplicaron a una columna Bio Scale Profinity IMAC de 5 ml (VC) y después de una etapa de lavado con el amortiguador A (PBS 1x, 5% de glicerol, 1 X SigmaFAST Proteasa Inhibitor Cocktail) la proteína unida a la columna se eluyó con diferentes concentraciones de imidazol. **A)** Perfil cromatográfico de Q4. **B)** Análisis SDS PAGE de los picos de elución de Q8 a diferentes concentraciones de imidazol (12.5, 25 y 50). MW: Marcador de peso molecular. F35-F36: Fracciones eluidas a 12.5 mM. F50-F51: Fracciones eluidas a 25 mM. F73-F81: Fracciones eluidas a 50 mM.

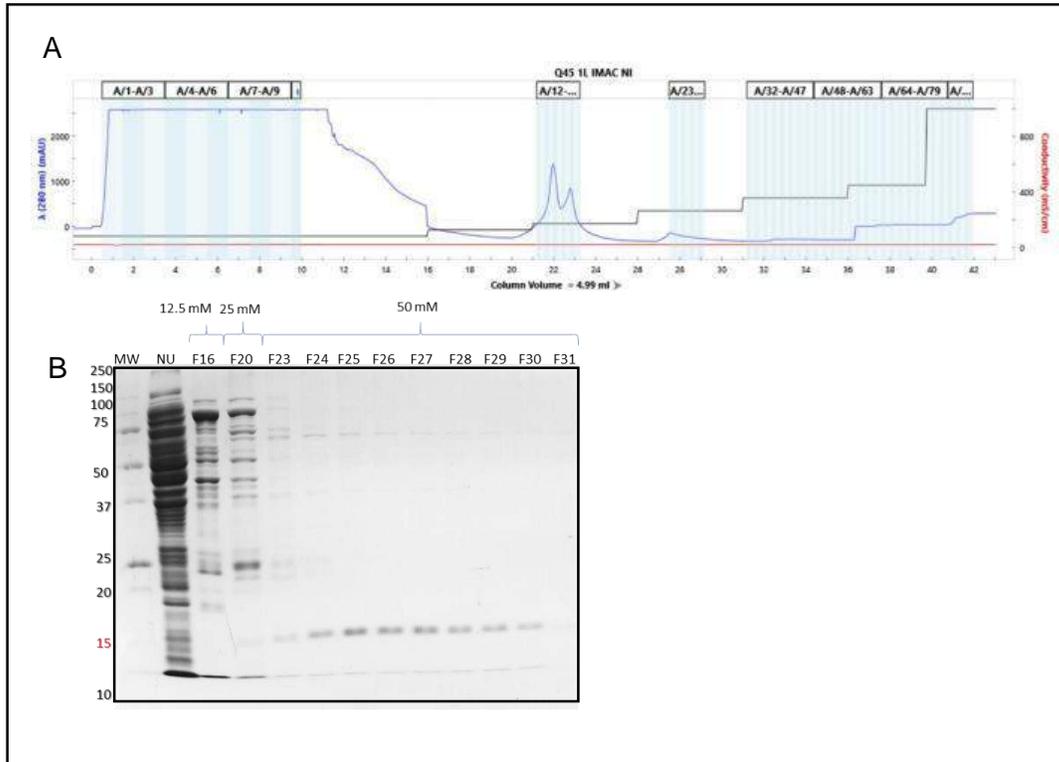


Figura 17. Purificación por cromatografía de afinidad a níquel de la proteína Q45.

Los extractos proteicos de biomasa de Q45 obtenidos después de 16 h de inducción se aplicaron a una columna Bio Scale Profinity IMAC de 5 ml (VC) y después de una etapa de lavado con el amortiguador A (PBS 1x, 5% de glicerol, 1 X SigmaFAST Proteasa Inhibitor Cocktail) la proteína unida a la columna se eluyó con diferentes concentraciones de imidazol.

A) Perfil cromatográfico de Q45. **B)** Análisis SDS PAGE de los picos de elución de Q45 a diferentes concentraciones de imidazol (12.5, 25, 50 mM). MW: Marcador de peso molecular. NU: Fracción no unida. F16: Fracción eluida a 12.5 mM. F20: Fracción eluida a 25 mM. F23-F31: Fracciones eluidas a 50 mM.

7.4 Concentración y cuantificación de proteínas purificadas

Las proteínas previamente purificadas se concentraron por ultrafiltración en membrana y se determinó su concentración por ensayo de BSA para su posterior análisis de actividad inhibitoria. En la **Figura 18** se muestra el gel SDS-PAGE de la chagasina y quimeras Q4, Q8 y Q45. En la **Tabla 6** se muestra la concentración determinada de cada proteína.

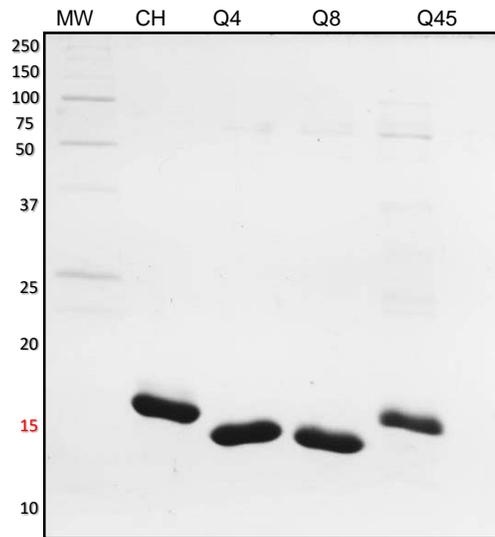


Figura 18. Análisis SDS PAGE de la chagasina, quimera Q4, Q8 y Q45 concentrada por filtración en membrana Amicon® Ultra-4.

Tabla 6. Concentración de chagasina (CH) y quimeras Q4, Q8, Q45.

Proteína	Concentración (μM)
CH	73.07
Q4	88.25
Q8	84.00
Q45	47.99

La cuantificación se realizó mediante kit de ensayo de proteínas de ácido bicinconínico (BCA)

7.5 Ensayos de actividad inhibitoria

La función inhibitoria de la chagasina y las quimeras recombinantes purificadas se probaron con h-Cat L (0.5 ng/ μ L) comercial, utilizando el sustrato fluorogénico Z-Phe-Arg-MCA (10 μ M) específico para esta CP. Las quimeras mostraron un efecto inhibitor dependiente de la concentración sobre la actividad proteolítica. La chagasina como un inhibidor de cisteína proteinasas (0-0.2 μ M) redujo la actividad de h-Cat L hasta en un 80 % (**Figura 19 A**). Las quimeras recombinantes se probaron a concentraciones crecientes de 0-6 μ M. La quimera Q4, la cual tiene insertado el epítipo E5 en el giro L4 y el epítipo E1 en el giro L6, redujo la actividad de h-Cat L hasta en un 80% (**Figura 19 B**). La quimera Q8, con el epítipo E5 insertado en el giro L4 y el epítipo E2 en el giro L6, redujo la actividad de h-Cat L hasta un 90 % (**Figura 19 C**). Por otro lado, la quimera Q45, con la inserción de dos epítipos E1 y E2 en el giro L4 y el epítipo E5 en el giro L6, redujo la actividad de h-Cat L hasta un 95 % (**Figura 19 D**). E-64 (1 μ M) usado como control positivo en todas las reacciones, inhibió la actividad proteolítica al 100 %.

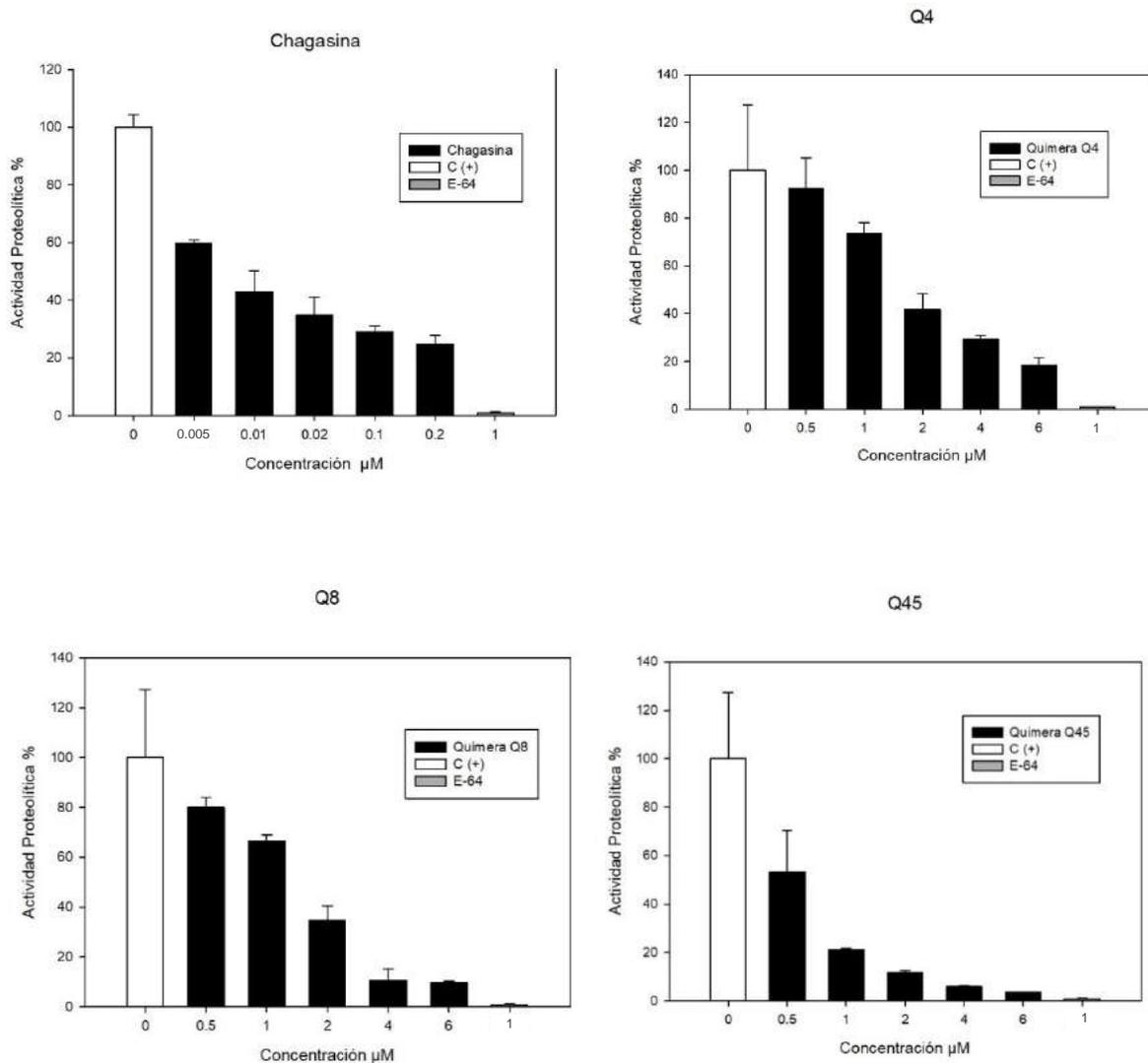


Figura 19. Inhibición de la actividad proteolítica de catepsina-L de hígado humano por las proteínas recombinantes chagasina, Q4, Q8 y Q45. Porcentaje de actividad proteolítica de **A)** Chagasina a diferentes concentraciones (0, 0.005, 0.01, 0.02, 0.1 y 0.2 μM). **B)** Q4, **C)** Q8, y **D)** Q45 a diferentes concentraciones (0, 0.5, 1, 2, 4 y 6 μM) en presencia del sustrato fluorogénico específico (Z-Phe-Arg-MCA) a una concentración final de 10 μM y de la enzima catepsina-L (0.5 $\text{ng}/\mu\text{L}$). Se utilizó E-64 a 1 μM como inhibidor de proteasas de tipo catepsina-L (gris). La primera barra (blanca) representa la actividad proteolítica total sin inhibidores.

En la **Tabla 7** se muestran la constante de inhibición aparente (K_{iapp}) y la constante de inhibición (K_i) obtenidas de cada proteína a partir de la gráfica $(V_o/V_i)-1$ vs $[I]_o$.

Tabla 7. Comparación de la constante de inhibición aparente (K_{iapp}) y la constante de inhibición (K_i) de la chagasina y quimeras recombinantes sobre la CP catepsina L.

Inhibidor	$K_{i(app)} \mu M$	$K_i \mu M$
Chagasina	0.011	0.0024
Q4	1.31	0.29
Q8	0.42	0.09
Q45	0.232	0.05

8. DISCUSIÓN

8.1 Análisis de solubilidad

La solubilidad y la agregación son propiedades cruciales de las proteínas, que pueden asegurar o prevenir su correcto funcionamiento (Trevino et al., 2008). Obtener una mejor comprensión de estos temas es cada vez más importante, ya que la solubilidad de las proteínas es necesaria para mejorar una amplia gama de procesos biotecnológicos y biofarmacéuticos, especialmente cuando se demandan alta concentración de estas proteínas. Es por ello, que en este trabajo primero se realizó un enfoque bioinformático sobre estas características de solubilidad aprovechando la disponibilidad de varios servidores en línea.

Se ha reportado que la composición de aminoácidos influye en la solubilidad de las proteínas. Un ejemplo es el aspartato (D), glutamato (E), serina (S) y lisina (K) que contribuyen más favorablemente a la solubilidad que otros aminoácidos hidrofílicos (Niwa et al., 2009; Chan et al., 2013); los valores de las relaciones Lys / Arg y Glu / Asp se correlacionan con la solubilidad (Warwicker et al., 2014; Chan *et al.*, 2013); mientras que los residuos aromáticos como fenilalanina (F), triptófano (W) y tirosina (Y) están asociados con baja solubilidad (Niwa *et al.*, 2009; Hebditch et al., 2017).

Además, se ha demostrado que las interacciones proteína-proteína y proteína-disolvente juegan un papel clave en las propiedades de solubilidad. En particular, los residuos expuestos a solventes tienen algunas características que están bien correlacionadas con esta propiedad: las proteínas insolubles tienden a tener parches de superficie más grandes que llevan una carga neta positiva que las proteínas solubles (Chan *et al.*, 2013), que se caracterizan en cambio

por una superficie más cargada negativamente (Kramer et al., 2012).

En un trabajo reciente, se mostró que entre todas las interacciones residuo-residuo, los puentes salinos que contienen Lys (K) y las interacciones alifáticas contribuyen más fuertemente que otros a promover la solubilidad, mientras que las interacciones que involucran electrones π deslocalizados favorecen agregación (por ejemplo, interacciones aromáticas, His- π , catión- π , amino- π y anión- π) (Hou *et al.*, 2018).

De acuerdo con los resultados de predicción con diferentes servidores bioinformáticos mostrados en la **Tabla 5**, los predictores con mayor similitud y correlación con la experimentación fueron SODOPE, SOLART, CAMSOL y AGGREGSCAN3D, por lo que se decidió basar la discusión en estos mismos, descartando a CCSOLOMICS Y PROTEIN SOL.

La quimera que presentó la mayor solubilidad (en promedio 69 %) fue Q4. Esto fue corroborado por medio del análisis de la expresión mediante SDS-PAGE (**Figura 13A**) donde se observó que la banda en la fracción soluble fue de mayor intensidad que en la fracción insoluble. Haciendo un análisis de la secuencia de los epítomos insertados en L4 (E5) y L6 (E1), que contienen 10 y 9 residuos de aminoácidos, respectivamente, ambos contienen 3 residuos altamente solubles S, D, D en E5 y K, E, D en E1. (Niwa *et al.*, 2009; Chan *et al.*, 2013).

La segunda quimera que presentó mayor solubilidad según los predictores fue Q8 (en promedio 65 %). Esto fue corroborado por medio del análisis de expresión SDS-PAGE (**Figura 13B**), donde se observó que la banda en la fracción soluble fue de mayor intensidad que en la fracción insoluble, pero de menor cantidad que en Q4. Tanto Q8 y Q4 contienen el epítomo E5 en L4 y difieren en epítomo en L6, E1 en Q4 con tres residuos de aminoácidos (S, D, D) altamente solubles y E2 en Q8 con los residuos G, N, G, T neutros. Por lo tanto, con base en esta

observación era de esperarse que la predicción de la solubilidad de Q8 fuera menor. Interesantemente, los resultados de la expresión son consistentes con esta predicción. Q8 se expresó en forma soluble, pero con menor intensidad que Q4 (Niwa *et al.*, 2009; Chan *et al.*, 2013).

Desde la estimación de los valores de solubilidad de la quimera Q45, que contiene dos epítomos en la posición L4 mostró un comportamiento interesante. A pesar de contar con dos epítomos en L4 (E1-E2) con 18 residuos de aminoácidos y un epítomo en L6 (E5) con 10 residuos de aminoácidos presentó una solubilidad en promedio 62%, menor que Q4 y Q8, pero mayor que Q20. Esta predicción resultó consistente con el análisis de su expresión mediante SDS-PAGE (**Figura 13D**), donde se observó la misma intensidad entre las bandas correspondientes a las fracciones soluble e insoluble, sugiriendo que aproximadamente el 50% de la proteína recombinante se expresó en su forma soluble. Si se analiza la secuencia de los 23 aminoácidos insertados en L4 contiene tres residuos altamente solubles K, E, D y siete neutros G, G, G, G, N, G, T y los 10 residuos insertados en L6 tiene tres residuos S, D, D altamente solubles (Niwa *et al.*, 2009; Chan *et al.* 2013). Además, se conoce que Q45 que contiene tres epítomos, se tomó en cuenta la existencia de una proteína homóloga a la chagasina en *T. brucei* con un giro L4 de 21 aminoácidos de extensión y L6 de 10 aminoácidos (Rigden *et al.*, 2002; dos Reis *et al.*, 2008). En este caso, dos epítomos se ubicaron en el lazo L4 (interespaciados entre sí por el linker GPGPG) y el tercero se ubicó en el lazo L6 (Nava, 2018). *T. brucei* contiene los aminoácidos NNRAPHDNHKNHRRLLVGAGG en L4 y, por otro lado la modificación de chagasina de *T. cruzi*, Q45, los aminoácidos PPKLFPEVIDLGPGPGIVMGNGTLVGG. Por lo que es importante señalar, que difieren en la cantidad de aminoácidos solubles e insolubles, teniendo *T. brucei*

dos altamente solubles D, K y, Q45 tres (K, E, D), además de un aminoácido aromático altamente insoluble F.

Finalmente, la quimera más insoluble fue Q20 (en promedio 56 %). Esto fue corroborado en el análisis mediante SDS-PAGE (**Figura 13C**) de su expresión en *E. coli*, donde se observó que la banda en la fracción soluble muestra una intensidad menor en comparación con las demás quimeras. Resulta interesante observar que la sustitución de la secuencia de E4 en L4, con 9 residuos de aminoácido que contiene 3 residuos aromáticos (F, Y, F) y 3 residuos hidrofóbicos (V, I, V) con un alto índice de insolubilidad, afectaron drásticamente la expresión de esta quimera, aun cuando en L6 se insertó un epítipo con un alto índice de solubilidad (E5). Desde el extracto total se observó una expresión menor a las demás quimeras y por ende en la fracción soluble se observa una banda muy débil. Por ello, la purificación por cromatografía no se pudo llevar a cabo, ya que la cantidad de proteína expresada de forma soluble fue muy poca.

Por otro lado, la expresión en *E. coli* BL21 a 37°C de Q8, Q20 y Q45 reportada por Nava, 2018, se llevó a cabo en forma soluble para Q8, pero Q20 y Q45 se llevaron mayoritariamente en forma insoluble. Los resultados obtenidos en este trabajo con la expresión en *E.coli* SoluBL21 mostraron ser significativamente favorables para la obtención de proteínas recombinantes solubles.

Estos diferentes hallazgos demuestran la importante potencialidad de la información secuencial de aminoácidos como estructural en la comprensión de los mecanismos de solubilidad.

8.2 Análisis de actividad Inhibitoria

Por otro lado, se conoce que la chagasina se encuentra presente en la superficie del parásito *T. cruzi* o fuera de la célula (Monteiro *et al.*, 2001), donde puede interactuar con las cisteína proteinasas humanas, como la catepsina L. Esta enzima es blanco potencial de la chagasina durante la infección de este parásito. Dado que la catepsina L es probable que sea importante en la respuesta del huésped al parásito invasor a través de sus propiedades reguladoras de la presentación de antígeno (Bevec *et al.*, 1996), es posible que la inhibición de la catepsina L mediada por la chagasina mejore las posibilidades de supervivencia del parásito o que el parásito necesite chagasina para regular la acción de su propia catepsina L como la cisteína proteinasa, cruzipaína.

Se ha cristalizado y resuelto la estructura de la chagasina en complejo con la cisteína proteinasa del huésped, catepsina L con una resolución de 1.75 Å. Una cuña inhibidora compuesta por tres *giros* (L2, L4 y L6) de la chagasina forma varios contactos responsables de la unión de alta afinidad (K_i , 0.0039 μM) a la enzima (Ljunggren *et al.*, 2007). De igual forma, en los resultados experimentales obtenidos en este trabajo se obtuvo un valor muy similar al reportado por los autores (K_i , 0.0024 μM).

Además, se conoce que los residuos conservados dentro de L6 parecen realizar interacciones principalmente hidrófobas con las peptidasas y, en última instancia, podrían contribuir a una inhibición de alta afinidad (Miao *et al.*, 2014).

De acuerdo con los resultados de ensayos de inhibición de la chagasina con la CP catepsina L y las quimeras (**Tabla 7**), todas las quimeras mostraron tener un efecto negativo en la actividad

inhibitoria, debido a que se modificaron los residuos más importantes para el anclaje con la enzima, como L65, G66 y A67 en el giro L4 y W93 en el giro L6 (**Tabla 2**) (Redzynia et al., 2009). Como se ha demostrado, el residuo triptófano (W93) es importante, ya que un estudio demostró que al mutarlo por una alanina (A93) se disminuye 100 veces la afinidad por catepsina L (Dos Reis *et al.*, 2008).

La chagasina mostró tener la mayor capacidad de inhibición de todas las proteínas analizadas, con una K_i de 0.0024 μM . Este resultado era de esperarse debido a que es el inhibidor natural sin ninguna modificación en sus secuencias de aminoácidos. Posteriormente, la quimera que presentó mayor inhibición de la actividad de CP fue Q45 (E1-E2 y E5) con una K_i de 0.05 μM . Esto pudiera deberse a que es la quimera con mayor número de aminoácidos hidrofóbicos, en particular, en el giro L6 contiene un aminoácido aromático (Y) que podría interactuar con W189 y W193 de la catepsina L (Ljunggren *et al.*, 2007; Miao *et al.*, 2014). Con respecto a L4, contiene 2 epítomos y como se observa en la **Figura 10**, se puede especular que cuenta con un mayor número de aminoácidos que pueden ser susceptibles a la interacción con la catepsina L. En particular contiene 3 leucinas (L) que aumentan la probabilidad de interacción con G67 y G68 de la catepsina L (Ljunggren *et al.*, 2007).

Por otro lado, la segunda quimera que presentó inhibición a catepsina L fue Q8 (E5 y E1), con una constante de inhibición de 0.09 μM . Contiene un menor número de aminoácidos hidrofóbicos en L4 y L6 que la quimera Q45, lo que podría afectar con la interacción de los aminoácidos hidrofóbicos de la enzima, además de que no contiene aminoácidos aromáticos que pudieran ejercer un efecto estérico.

Por último, la quimera con menor inhibición a catepsina L fue Q4 (E5 y E1), con una K_i de 0.29 μM . Aunque contiene un aminoácido aromático (Y) en el giro L6, la mayoría de sus aminoácidos son hidrofílicos por lo que de alguna forma se sugiere que afecta la interacción hidrofóbica entre inhibidor-CP.

En conjunto, los resultados obtenidos en este trabajo demuestran la utilidad de usar los recursos bioinformáticos disponibles más recientes, para predecir la solubilidad de proteínas que se desean expresar en *E. coli* y en específico de modificaciones a proteínas de andamiaje, como la chagasina, para la expresión de epítomos de antígenos de difícil expresión.

Por otra parte, en el desarrollo de fármacos para las infecciones parasitarias, una preocupación general es evitar la inhibición de enzimas similares del huésped. Los seres humanos expresan varias catepsinas, entre las que la catepsina L exhibe una alta identidad de secuencia con la cruzipaína (43% de identidad en el dominio catalítico) (Toman et al., 2020). Todas las quimeras mostraron un cambio importante en la K_i (0.05, 0.09, 0.29 μM) con respecto a la chagasina (0.0024 μM), por lo tanto, de acuerdo a este criterio, sería más conveniente la producción de una quimera recombinante que no inhiba o inhiba menos la catepsina L de humano. De a las quimeras estudiadas en este trabajo, la Q4 sería la más indicada como antígeno y además de fue la que se expresó con el mejor grado de solubilidad. Por otra parte, desde el punto de vista del número de epítomos de TSA-1, Q45 sería la más indicada, pero fue la que más inhibió a la catepsina L. Es posible que todas pudieran ser aptas para su uso como antígenos, lo cual se tendría que descartar en ensayos en animales de laboratorio o bien modificar también el giro L2 de la chagasina y encontrar una quimera que no inhiba la catepsina L.

9. CONCLUSIÓN

- Los valores de solubilidad de Chagasina y las quimeras Q4, Q8, Q20, y Q45 estimadas con los servidores SoDoPe, Solart, CamSol y AggresCan correlacionaron con la expresión de las proteínas recombinantes en *E. coli*, lo que sugiere que son una alternativa útil y recomendable de incluir en las estrategias experimentales para predecir la solubilidad de proteínas recombinantes a expresarse en *E. coli*.
- La expresión en *E. coli* SoluBL21 a 25°C mejoró la solubilidad.
- La modificación de los giros L4 y L6 de la chagasina por epítomos de TSA-1, afectó negativamente la solubilidad (CH>Q4>Q8,>Q45) y actividad inhibitoria de las quimeras recombinantes (CH>Q45>Q8>Q4).

10.PERSPECTIVAS

- Probar diferentes alternativas de expresión que mejoren la solubilidad de la proteína quimérica Q20, como el cambio de temperatura y tiempo de expresión. Además, realizar ensayos de inhibición para comprobar la hipótesis de que, al tener un mayor número de aminoácidos hidrofóbicos en los giros, puede inhibir más la actividad de la catepsina L.
- Probar in-silico todas las mutaciones en especial en L2 el epitopo E3, el cual es el más hidrofílico de todos los epítomos.
- Expresar las quimeras Q16, Q17, Q18, para evaluar completamente el efecto de la posición del epitopo E5 en la solubilidad de las quimeras y comparar con los predictores seleccionados.
- Evaluar el potencial para el control inmunoterapéutico de las quimeras recombinantes en modelos murinos y comparar su eficiencia con TSA-1

11. BIBLIOGRAFÍA

- Agostini, F., Vendruscolo, M., & Tartaglia, G. G. (2012). Sequence-based prediction of protein solubility. *Journal of Molecular Biology*, 421(2–3), 237–241. <https://doi.org/10.1016/j.jmb.2011.12.005>
- Alba Soto, C. D., & González Cappa, S. M. (2019). Trypanosoma cruzi Journey from the Insect Vector to the Host Cell. In J. M. Altcheh & H. Freilij (Eds.), *Chagas Disease: A Clinical Approach* (pp. 25–59). Springer International Publishing. https://doi.org/10.1007/978-3-030-00054-7_2
- Bevec, T., Stoka, V., Pungercic, G., Dolenc, I., & Turk, V. (1996). Major histocompatibility complex class II-associated p41 invariant chain fragment is a strong inhibitor of lysosomal cathepsin L. *The Journal of Experimental Medicine*, 183(4), 1331–1338. <https://doi.org/10.1084/jem.183.4.1331>
- Beynon R & Bound JS. (1982). A simple method for displaying the hydrophobic character of a protein. *Journal of Molecular Biology*, 157(1), 105–132. [https://doi.org/10.1016/0022-2836\(82\)90515-0](https://doi.org/10.1016/0022-2836(82)90515-0)
- Biter, A. B., Weltje, S., Hudspeth, E. M., Seid, C. A., McAtee, C. P., Chen, W.-H., Pollet, J. B., Strych, U., Hotez, P. J., & Bottazzi, M. E. (2018). Characterization and Stability of Trypanosoma cruzi 24-C4 (Tc24-C4), a Candidate Antigen for a Therapeutic Vaccine Against Chagas Disease. *Journal of Pharmaceutical Sciences*, 107(5), 1468–1473. <https://doi.org/10.1016/j.xphs.2017.12.014>
- Brömme, D., Steinert, A., Friebe, S., Fittkau, S., Wiederanders, B., & Kirschke, H. (1989). The specificity of bovine spleen cathepsin S. A comparison with rat liver cathepsins L and B. *The Biochemical Journal*, 264(2), 475–481. <https://doi.org/10.1042/bj2640475>
- Capelli, R., Marchetti, F., Tiana, G., & Colombo, G. (2017). SAGE: A Fast Computational Tool for Linear Epitope Grafting onto a Foreign Protein Scaffold. *Journal of Chemical Information and Modeling*, 57(1), 6–10. <https://doi.org/10.1021/acs.jcim.6b00584>
- Chan, P., Curtis, R. A., & Warwicker, J. (2013). Soluble expression of proteins correlates with a lack of positively-charged surface. *Scientific Reports*, 3(1), 3333. <https://doi.org/10.1038/srep03333>
- Costantini, S., Colonna, G., & Facchiano, A. M. (2006). Amino acid propensities for secondary structures are influenced by the protein structural class. *Biochemical and Biophysical Research Communications*, 342(2), 441–451. <https://doi.org/10.1016/j.bbrc.2006.01.159>
- de la Cruz, J. J., Villanueva-Lizama, L., Dzul-Huchim, V., Ramírez-Sierra, M.-J., Martínez-Vega, P., Rosado-Vallado, M., Ortega-Lopez, J., Flores-Pucheta, C. I., Gillespie, P., Zhan, B., Bottazzi, M. E., Hotez, P. J., & Dumonteil, E. (2019). Production of recombinant TSA-1 and evaluation of its potential for the immuno-therapeutic control of Trypanosoma cruzi infection in mice. *Human Vaccines & Immunotherapeutics*, 15(1), 210–219. <https://doi.org/10.1080/21645515.2018.1520581>
- de Souza, W., de Carvalho, T. M. U., & Barrias, E. S. (2010). Review on Trypanosoma cruzi: Host Cell Interaction. *International Journal of Cell Biology*, 2010. <https://doi.org/10.1155/2010/295394>
- Deras, L. (2011). Enfermedad de Chagas, en: Ruesga ZEA, Saturno CG. *Cardiología. El Manual Moderno, México.*, 1357–1369.
- Dorn, P. L., Perniciaro, L., Yabsley, M. J., Roellig, D. M., Balsamo, G., Diaz, J., & Wesson, D. (2007). Autochthonous transmission of Trypanosoma cruzi, Louisiana. *Emerging Infectious Diseases*, 13(4), 605–607. <https://doi.org/10.3201/eid1304.061002>
- dos Reis, F. C. G., Smith, B. O., Santos, C. C., Costa, T. F. R., Scharfstein, J., Coombs, G. H., Mottram, J. C., & Lima, A. P. C. A. (2008). The role of conserved residues of chagasin in the inhibition of cysteine peptidases. *FEBS Letters*, 582(4), 485–490. <https://doi.org/10.1016/j.febslet.2008.01.008>
- Dumonteil, E., Bottazzi, M. E., Zhan, B., Heffernan, M. J., Jones, K., Valenzuela, J. G., Kamhawi, S., Ortega, J., de Leon Rosales, S. P., Lee, B. Y., Bacon, K. M., Fleischer, B., Slingsby, B. T., Cravioto, M. B., Tapia-Conyer, R., & Hotez, P. J. (2012). Accelerating the development of a therapeutic vaccine for human Chagas disease: rationale and prospects. *Expert Review of Vaccines*, 11(9), 1043–1055. <https://doi.org/10.1586/erv.12.85>
- Dumonteil, E., Escobedo-Ortegon, J., Reyes-Rodriguez, N., Arjona-Torres, A., & Ramirez-Sierra, M. J. (2004). Immunotherapy of Trypanosoma cruzi infection with DNA vaccines in mice. *Infection and Immunity*, 72(1),

- 46–53. <https://doi.org/10.1128/iai.72.1.46-53.2004>
- Goulas, T., Cuppari, A., Garcia-Castellanos, R., Snipas, S., Glockshuber, R., Arolas, J. L., & Gomis-Rüth, F. X. (2014). The pCri System: a vector collection for recombinant protein expression and purification. *PLoS One*, *9*(11), e112643. <https://doi.org/10.1371/journal.pone.0112643>
- Hebditch, M., Carballo-Amador, M. A., Charonis, S., Curtis, R., & Warwicker, J. (2017). Protein-Sol: a web tool for predicting protein solubility from sequence. *Bioinformatics (Oxford, England)*, *33*(19), 3098–3100. <https://doi.org/10.1093/bioinformatics/btx345>
- Hou, Q., Kwasigroch, J. M., Rooman, M., & Pucci, F. (2020). SOLart: a structure-based method to predict protein solubility and aggregation. *Bioinformatics*, *36*(5), 1445–1452. <https://doi.org/10.1093/bioinformatics/btz773>
- Knight, J. M., Zingales, B., Bottazzi, M. E., Hotez, P., & Zhan, B. (2014). Limited antigenic variation in the *Trypanosoma cruzi* candidate vaccine antigen TSA-1. *Parasite Immunology*, *36*(12), 708–712. <https://doi.org/10.1111/pim.12130>
- Kramer, R. M., Shende, V. R., Motl, N., Pace, C. N., & Scholtz, J. M. (2012). Toward a molecular understanding of protein solubility: increased negative surface charge correlates with increased solubility. *Biophysical Journal*, *102*(8), 1907–1915. <https://doi.org/10.1016/j.bpj.2012.01.060>
- Kuriata, A., Iglesias, V., Pujols, J., Kurcinski, M., Kmiecik, S., & Ventura, S. (2019). *Aggrescan3D (A3D) 2.0: prediction and engineering of protein solubility*. *47*(May), 300–307. <https://doi.org/10.1093/nar/gkz321>
- Kyte, J., & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, *157*(1), 105–132. [https://doi.org/10.1016/0022-2836\(82\)90515-0](https://doi.org/10.1016/0022-2836(82)90515-0)
- Linding, R., Russell, R. B., Neduva, V., & Gibson, T. J. (2003). GlobPlot: Exploring protein sequences for globularity and disorder. *Nucleic Acids Research*, *31*(13), 3701–3708. <https://doi.org/10.1093/nar/gkg519>
- Ljunggren, A., Redzynia, I., Alvarez-Fernandez, M., Abrahamson, M., Mort, J. S., Krupa, J. C., Jaskolski, M., & Bujacz, G. (2007). Crystal Structure of the Parasite Protease Inhibitor Chagasin in Complex with a Host Target Cysteine Protease. *Journal of Molecular Biology*, *371*(1), 137–153. <https://doi.org/10.1016/j.jmb.2007.05.005>
- Martinez-Campos, V., Martinez-Vega, P., Ramirez-Sierra, M. J., Rosado-Vallado, M., Seid, C. A., Hudspeth, E. M., Wei, J., Liu, Z., Kwityn, C., Hammond, M., Ortega-López, J., Zhan, B., Hotez, P. J., Bottazzi, M. E., & Dumonteil, E. (2015). Expression, purification, immunogenicity, and protective efficacy of a recombinant Tc24 antigen as a vaccine against *Trypanosoma cruzi* infection in mice. *Vaccine*, *33*(36), 4505–4512. <https://doi.org/10.1016/j.vaccine.2015.07.017>
- Miao, Q., Santamaria, C., Bailey, D., Genest, J., Ward, B. J., & Ndao, M. (2014). Apolipoprotein A-I truncations in Chagas disease are caused by cruzipain, the major cysteine protease of *Trypanosoma cruzi*. *The American Journal of Pathology*, *184*(4), 976–984. <https://doi.org/10.1016/j.ajpath.2013.12.018>
- Monteiro, A. C. S., Abrahamson, M., Lima, A. P. C. A., Vannier-Santos, M. A., & Scharfstein, J. (2001). Identification, characterization and localization of chagasin, a tight-binding cysteine protease inhibitor in *Trypanosoma cruzi*. *Journal of Cell Science*, *114*(21), 3933–3942.
- Moretti, N. S., Mortara, R. A., & Schenkman, S. (2020). *Trypanosoma cruzi*. *Trends in Parasitology*, *36*(4), 404–405. <https://doi.org/10.1016/j.pt.2019.10.002>
- Murgueitio, C. (2017). Parasitosis extraintestinales. Enfermedad de Chagas. En: Pinilla RAE, Murgueitio CR, Calderón VCM, Mesa UA, Villalobos GW, Espinosa TF y col. Compendio de Terapéutica. Evidencia actual. *Asociación Colombiana de Medicina Interna-Edit. Médica Celsus*, 769–770.
- Murillo, G. (2018). Enfermedad de Chagas (tripanosomiasis americana). *Med Int Méx*, *34*(6), 959–970. <https://doi.org/10.24245/mim.v34i6.2217>
- Niwa, T., Ying, B.-W., Saito, K., Jin, W., Takada, S., Ueda, T., & Taguchi, H. (2009). Bimodal protein solubility distribution revealed by an aggregation analysis of the entire ensemble of *Escherichia coli* proteins. *Proceedings of the National Academy of Sciences*, *106*(11), 4201 LP – 4206. <https://doi.org/10.1073/pnas.0811922106>
- Organización Mundial de la Salud. (2013). *Sustaining the drive to overcome the global impact of neglected tropical diseases: second WHO report on neglected tropical diseases*.
- Pereira-Chioccola, V. L., & Schenkman, S. (1999). Biological role of *Trypanosoma cruzi* trans-sialidase. *Biochemical Society Transactions*, *27*(4), 516–518. <https://doi.org/10.1042/bst0270516>
- Pérez-Molina, J. A., Perez, A. M., Norman, F. F., Monge-Maillou, B., & López-Vélez, R. (2015). Old and new

- challenges in Chagas disease. *The Lancet. Infectious Diseases*, 15(11), 1347–1356.
[https://doi.org/10.1016/S1473-3099\(15\)00243-1](https://doi.org/10.1016/S1473-3099(15)00243-1)
- Petterson, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13), 1605–1612. <https://doi.org/10.1002/jcc.20084>
- Ramsey, J. M., Peterson, A. T., Carmona-Castro, O., Moo-Llanes, D. A., Nakazawa, Y., Butrick, M., Tun-Ku, E., la Cruz-FÃ©\copyrightlix, K. de, & Ibarra-CerdeÃ \pma, C. N. (2015). Atlas of Mexican Triatominae (Reduviidae: Hemiptera) and vector transmission of Chagas disease. *MemÃ \textthreesuperiorrias Do Instituto Oswaldo Cruz*, 110, 339–352. http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0074-02762015000300339&nrm=iso
- Ramsey J.M., Ordoñez R, Tello Lopez A, Pohls J.L, Sanchez V., P. A. T. (2003). Actualidades sobre la epidemiología de la enfermedad de chagas en México. In *Iniciativa para la Vigilancia y el Control de la Enfermedad De Chagas*. (Primera ed, pp. 85–105).
- Redzynia, I., Ljunggren, A., Abrahamson, M., Mort, J. S., Krupa, J. C., Jaskolski, M., & Bujacz, G. (2008). Displacement of the occluding loop by the parasite protein, chagasin, results in efficient inhibition of human cathepsin B. *Journal of Biological Chemistry*, 283(33), 22815–22825.
<https://doi.org/10.1074/jbc.M802064200>
- Redzynia, I., Ljunggren, A., Bujacz, A., Abrahamson, M., Jaskolski, M., & Bujacz, G. (2009). Crystal structure of the parasite inhibitor chagasin in complex with papain allows identification of structural requirements for broad reactivity and specificity determinants for target proteases. *FEBS Journal*, 276(3), 793–806.
<https://doi.org/10.1111/j.1742-4658.2008.06824.x>
- Reyes PA. (2007). Cardiopatía chagásica crónica, en: Vargas BJ. In *Tratado de Cardiología. Sociedad Mexicana de Cardiología-Intersistemas*.
- Rigden, D. J., Mosolov, V. V., & Galperin, M. Y. (2002). Sequence conservation in the chagasin family suggests a common trend in cysteine proteinase binding by unrelated protein inhibitors. *Protein Science*, 11(8), 1971–1977. <https://doi.org/10.1110/ps.0207202>
- Salmon, D., Do Aido-Machado, R., Diehl, A., Leidert, M., Schmetzer, O., Lima, A. P. C. D. A., Scharfstein, J., Oschkinat, H., & Pires, J. R. (2006). Solution structure and backbone dynamics of the Trypanosoma cruzi cysteine protease inhibitor chagasin. *Journal of Molecular Biology*, 357(5), 1511–1521.
<https://doi.org/10.1016/j.jmb.2006.01.064>
- Sanderson, S. J., Westrop, G. D., Scharfstein, J., Mottram, J. C., & Coombs, G. H. (2003). Functional conservation of a natural cysteine peptidase inhibitor in protozoan and bacterial pathogens. *FEBS Letters*, 542(1–3), 12–16. [https://doi.org/10.1016/s0014-5793\(03\)00327-2](https://doi.org/10.1016/s0014-5793(03)00327-2)
- Smith, D. K., Radivojac, P., Obradovic, Z., Dunker, A. K., & Zhu, G. (2003). Improved amino acid flexibility parameters. *Protein Science : A Publication of the Protein Society*, 12(5), 1060–1072.
<https://doi.org/10.1110/ps.0236203>
- Sormanni, P., Aprile, F. A., & Vendruscolo, M. (2014). The CamSol Method of Rational Design of Protein Mutants with Enhanced The CamSol Method of Rational Design of Protein Mutants with Enhanced Solubility. *Journal of Molecular Biology*, October. <https://doi.org/10.1016/j.jmb.2014.09.026>
- Trevino, S. R., Scholtz, J. M., & Pace, C. N. (2008). Measuring and increasing protein solubility. *Journal of Pharmaceutical Sciences*, 97(10), 4155–4166. <https://doi.org/10.1002/jps.21327>
- Uversky, V. N., Gillespie, J. R., & Fink, A. L. (2000). Why are “natively unfolded” proteins unstructured under physiologic conditions? *Proteins*, 41(3), 415–427. [https://doi.org/10.1002/1097-0134\(20001115\)41:3<415::aid-prot130>3.0.co;2-7](https://doi.org/10.1002/1097-0134(20001115)41:3<415::aid-prot130>3.0.co;2-7)
- Villanueva-Lizama, L. E., Cruz-Chan, J. V., Aguilar-Cetina, A. D. C., Herrera-Sanchez, L. F., Rodriguez-Perez, J. M., Rosado-Vallado, M. E., Ramirez-Sierra, M. J., Ortega-Lopez, J., Jones, K., Hotez, P., Bottazzi, M. E., & Dumonteil, E. (2018). Trypanosoma cruzi vaccine candidate antigens Tc24 and TSA-1 recall memory immune response associated with HLA-A and -B supertypes in Chagasic chronic patients from Mexico. *PLoS Neglected Tropical Diseases*, 12(1), e0006240. <https://doi.org/10.1371/journal.pntd.0006240>
- Warwicker, J., Charonis, S., & Curtis, R. A. (2014). Lysine and Arginine Content of Proteins: Computational Analysis Suggests a New Tool for Solubility Design. *Molecular Pharmaceutics*, 11(1), 294–303.
<https://doi.org/10.1021/mp4004749>
- Wrightsmann, R. A., Dawson, B. D., Fouts, D. L., & Manning, J. E. (1994). Identification of immunodominant

epitopes in *Trypanosoma cruzi* trypomastigote surface antigen-1 protein that mask protective epitopes. *Journal of Immunology (Baltimore, Md. : 1950)*, *153*(7), 3148–3154.

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. *Nature Methods*, *12*(1), 7–8. <https://doi.org/10.1038/nmeth.3213>

12.APÉNDICE

12.1 Apéndice 1

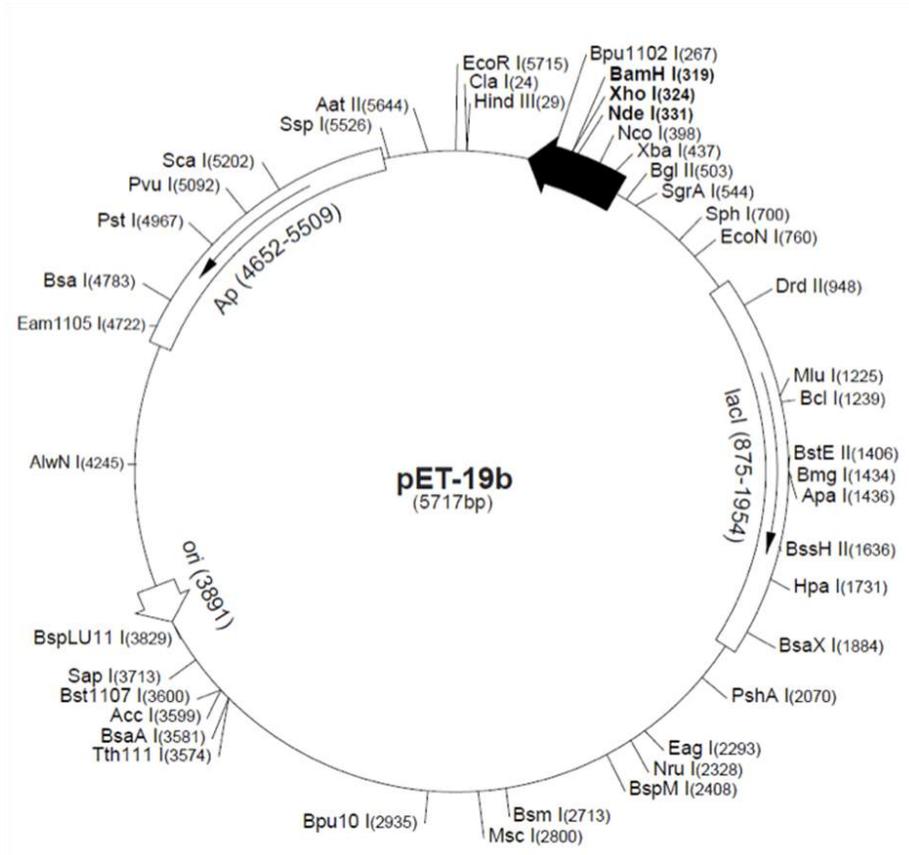


Figura 20. Mapa del vector de expresión pET-19b.

12.2 Apéndice 2

Tabla 8. Esquema de preparación de las reacciones de actividad de CP catepsina L con la chagasina recombinante.

Chagasina (μM)	*Dilución 1:100			*Dilución 1:50		
	0	0.005	0.01	0.02	0.1	0.2
	μL	μL	μL	μL	μL	μL
Buffer 1X	50	50	50	50	50	50
β -met (2mM)	4	4	4	4	4	4
h-Cat L ** (0.5 ng/ μL)	2	2	2	2	2	2
Substrato ** (10 μM)	4	4	4	4	4	4
Chagasina	0	1	3	3	14	27
H2O	140	139	137	137	126	113

*La chagasina se partió de un stock de 73.07 μM con una dilución de 1:100 para 0.005 y 0.01 μM y 1:50 para 0.02, 0.1 y 0.2 μM

** h-Cat L se partió de un stock de 50 ng/ μL y el substrato de un stock de 500 μM .

Tabla 9. Esquema de preparación de las reacciones de actividad de CP catepsina L con la quimera Q4 recombinante.

Q4 (μM)	*Dilución 1:5					
	0	0.5	1	2	4	6
	μL	μL	μL	μL	μL	μL
Buffer 1X	50	50	50	50	50	50
β -met (2mM)	4	4	4	4	4	4
h-Cat L ** (0.5 ng/ μL)	2	2	2	2	2	2
Substrato ** (10 μM)	4	4	4	4	4	4
Q4	0	6	2	5	9	14
H2O	140	134	138	135	131	126

*La quimera Q4 se partió de un stock de 88.25 μM con una dilución de 1:5 para 0.5 μM .

** h-Cat L se partió de un stock de 50 ng/ μL y el substrato de un stock de 500 μM .

Tabla 10. Esquema de preparación de las reacciones de actividad de CP catepsina L con la quimera Q8 recombinante.

*Dilución 1:5						
Q8 (μM)	0	0.5	1	2	4	6
	μL	μL	μL	μL	μL	μL
Buffer 1X	50	50	50	50	50	50
β-met (2mM)	4	4	4	4	4	4
h-Cat L ** (0.5 ng/μL)	2	2	2	2	2	2
Substrato ** (10μM)	4	4	4	4	4	4
Q8	0	6	2	5	10	14
H2O	140	134	138	135	130	126

*La quimera Q8 se partió de un stock de 84 μM con una dilución de 1:5 para 0.5 μM .

** h-Cat L se partió de un stock de 50 ng/ μL y el substrato de un stock de 500 μM .

Tabla 11. Esquema de preparación de las reacciones de actividad de CP catepsina L con la quimera Q45 recombinante.

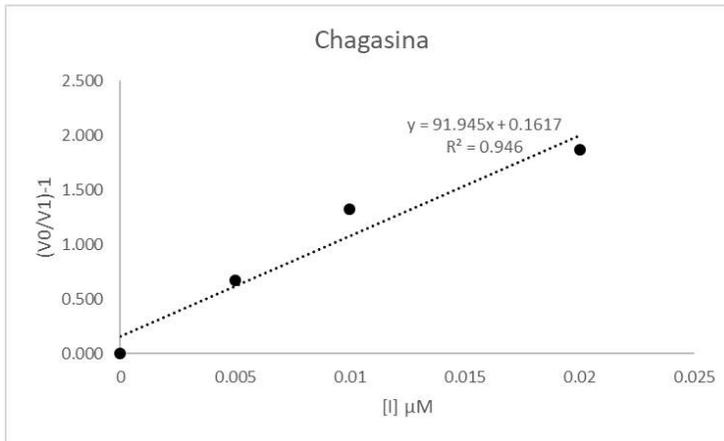
Dilución 1:5						
Q45 (μM)	0	0.5	1	2	4	6
	μL	μL	μL	μL	μL	μL
Buffer 1X	50	50	50	50	50	50
β-met (2mM)	4	4	4	4	4	4
h-Cat L ** (0.5 ng/μL)	2	2	2	2	2	2
Substrato ** (10μM)	4	4	4	4	4	4
Q45	0	10	4	8	17	25
H2O	140	130	136	132	123	115

*La quimera Q45 se partió de un stock de 47.99 μM con una dilución de 1:5 para 0.5 μM .

** h-Cat L se partió de un stock de 50 ng/ μL y el substrato de un stock de 500 μM .

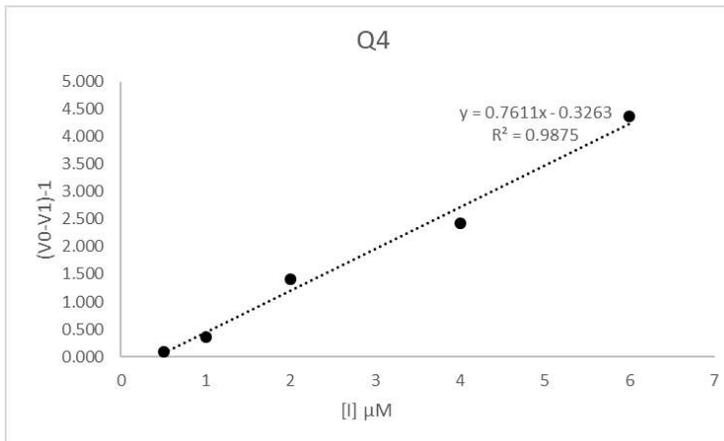
.

12.3 Apéndice 3



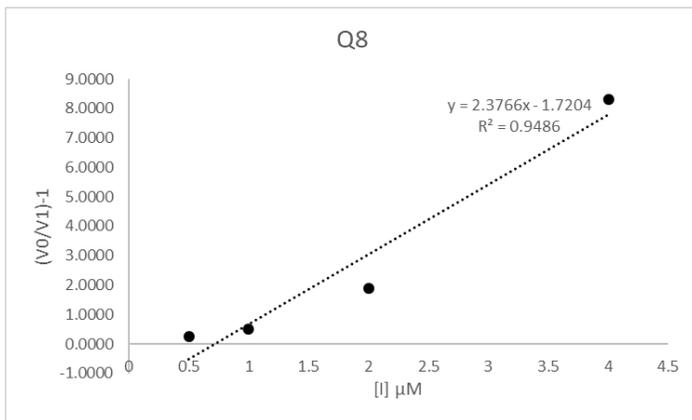
$1/Ki(\text{app}) =$	91.945
$Ki(\text{app}) =$	0.011 μM
$Km =$	2.8 μM
$[S] =$	10 μM
$Ki = \frac{Ki(\text{app})}{(1 + [S]/Km)}$	
$Ki =$	0.0024 μM

Figura 21. Gráfico de la recta $(V_0/V_1)-1$ vs $[I]$ para la obtención y cálculo de $Ki(\text{app})$ y Ki de chagasina.



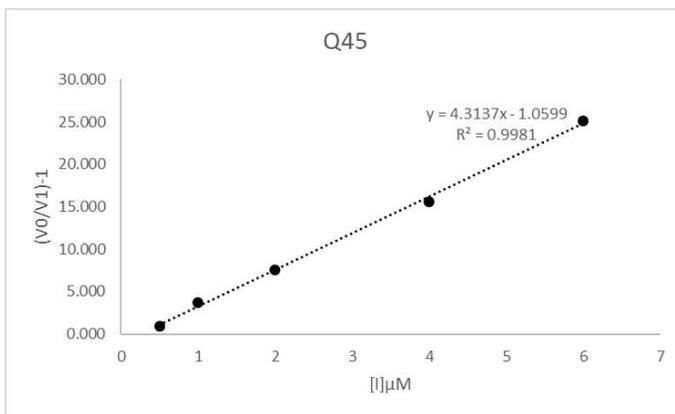
$1/Ki(\text{app}) =$	0.7611
$Ki(\text{app}) =$	1.31 μM
$Km =$	2.8 μM
$[S] =$	10 μM
$Ki = \frac{Ki(\text{app})}{(1 + [S]/Km)}$	
$Ki =$	0.287 μM

Figura 22. Gráfico de la recta $(V_0/V_1)-1$ vs $[I]$ para la obtención y cálculo de $Ki(\text{app})$ y Ki de la quimera Q4.



$1/Ki(app) =$	2.3766
$Ki(app) =$	0.42 μM
$Km =$	2.8 μM
$[S] =$	10 μM
$Ki = \frac{Ki(app)}{(1 + [S]/Km)}$	
$Ki =$	0.09 μM

Figura 23. Gráfico de la recta $(V_0/V_1)-1$ vs $[I]$ para la obtención y cálculo de $Ki(app)$ y Ki de la quimera Q8.



$1/Ki(app) =$	4.3137
$Ki(app) =$	0.232 μM
$Km =$	2.8 μM
$[S] =$	10 μM
$Ki = \frac{Ki(app)}{(1 + [S]/Km)}$	
$Ki =$	0.05 μM

Figura 24. Gráfico de la recta $(V_0/V_1)-1$ vs $[I]$ para la obtención y cálculo de $Ki(app)$ y Ki de la quimera Q45.

12.4 Apéndice 4

Tabla12. Predicción bioinformática de solubilidad de las proteínas quiméricas multiepitopo.

	ccSOL omics	Protein Sol		SoDoPE		Solart	CamSol		AggresCan3D	
		Score		Score			Score		Score	
	%	0 a 1	%	0 a 1	%	%	-1.5 a 1.5	%	%	
Chagasina	47	0.604	60	0.8017	80	83.8	0.5433	68	-0.6265	100
Q1(E2-E1)	67	0.553	55	0.7182	72	81.7	0.3517	62	-0.3819	61
Q2(E3-E1)	85	0.576	58	0.7857	79	81.2	0.5788	69	-0.5879	94
Q3(E4-E1)	54	0.557	56	0.6892	69	82.0	-0.1780	44	-0.2169	35
Q4(E5-E1)	72	0.556	56	0.7229	72	83.8	0.4434	65	-0.376	60
Q5(E1-E2)	71	0.553	55	0.7182	72	81.9	0.3877	63	-0.369	59
Q6(E3-E2)	39	0.553	55	0.7539	75	79.9	0.4462	65	-0.5994	96
Q7(E4-E2)	69	0.553	55	0.6496	65	80.4	-0.3107	40	-0.3412	54
Q8(E5-E2)	79	0.557	56	0.686	69	81.7	0.3114	60	-0.3286	52
Q9(E1-E3)	53	0.576	58	0.7857	79	79.3	0.5692	69	-0.4277	68
Q10(E2-E3)	85	0.553	55	0.7539	75	79.4	0.400	63	-0.3762	60
Q11(E4-E3)	69	0.566	57	0.7273	73	79.9	-0.1291	46	-0.4409	70
Q12(E5-E3)	84	0.608	61	0.758	76	77	0.4921	66	-0.4574	73
Q13(E1-E4)	52	0.557	56	0.6892	69	76.8	-0.1002	47	-0.299	48
Q14(E2-E4)	62	0.553	55	0.6496	65	78.5	-0.2687	41	-0.3513	56
Q15(E3-E4)	66	0.566	57	0.7273	73	78.7	-0.0416	49	-0.5641	90
Q16(E5-E4)	80	0.571	57	0.6556	66	79.3	-0.1743	44	-0.5176	83
Q17(E1-E5)	66	0.556	56	0.7298	73	82	0.4767	66	-0.4551	73
Q18(E2-E5)	87	0.557	56	0.686	69	80.2	0.308	60	-0.3343	53

Q19(E3-E5)	33	0.608	61	0.758	76	82.5	0.5349	68	-0.683	109
Q20(E4-E5)	53	0.571	57	0.6556	66	79.6	-0.2185	43	-0.2488	40
Q21(E2 Y E3 - E1)	34	0.553	55	0.7366	74		0.4184	64		
Q22(E4 Y E5 - E1)	54	0.582	58	0.6428	64		-0.3640	38		
Q23(E2 Y E5 - E1)	36	0.546	55	0.6716	67		0.1774	56		
Q24(E3 Y E4 - E1)	87	0.583	58	0.7109	71		-0.1650	44		
Q25(E2 Y E4 - E1)	68	0.553	55	0.6371	64		-0.5057	33		
Q26 (E3 Y E5 - E1)	55	0.557	56	0.7407	74		0.4981	67		
Q27(E1 Y E3 - E2)	81	0.553	55	0.7366	74		0.4099	64		
Q28(E4 Y E5- E2)	71	0.569	57	0.6042	60		-0.4909	34		
Q29(E1 Y E5- E2)	69	0.546	55	0.6716	67		0.3375	61		
Q30(E3 Y E4-E2)	67	0.553	55	0.7108	71		-0.2925	40		
Q31(E1 Y E4- E2)	75	0.553	55	0.6371	64		-0.2947	40		
Q32(E3 Y E5- E2)	76	0.564	56	0.7078	71		0.3711	62		
Q33(E1 Y E2- E3)	82	0.553	55	0.7366	74		0.3884	63		
Q34(E4 Y E5- E3)	60	0.597	60	0.6807	68		-0.3172	39		
Q35(E1 Y E5- E3)	71	0.557	56	0.7407	74		0.5113	67		
Q36(E2 Y E4- E3)	52	0.553	55	0.6756	68		-0.4587	35		
Q37(E1 Y E4- E3)	71	0.583	58	0.7109	71		-0.1202	46		
Q38(E2 Y E5- E3)	69	0.564	56	0.7078	71		0.2242	57		
Q39(E1 Y E2- E4)	83	0.553	55	0.6371	64		-0.2551	41		
Q40(E3 Y E5- E4)	82	0.597	60	0.6807	68		-0.0960	47		
Q41(E1 Y E5- E4)	74	0.582	58	0.6428	64		-0.1296	46		
Q42(E2 Y E3- E4)	63	0.553	55	0.6756	68		-0.1781	44		
Q43(E1 Y E3- E4)	82	0.583	58	0.7109	71		-0.0591	48		
Q44(E2 Y E5- E4)	72	0.569	57	0.6042	60		-0.4166	36		

Q45(E1 y E2-E5)	67	0.546	55	0.6716	67	80.4	0.3000	60	-0.2623	42
Q46(E3 Y E4- E5)	66	0.597	60	0.6807	68		-0.2043	43		
Q47(E1 Y E4- E5)	44	0.582	58	0.6428	64		-0.2065	43		
Q48(E2 Y E3- E5)	71	0.564	56	0.7078	71		0.3767	63		
Q49(E1 Y E3- E5)	74	0.557	56	0.7407	74		0.4952	67		
Q50(E2 Y E4- E5)	52	0.569	57	0.6042	60		-0.5436	32		

13.METODOLOGÍAS GENERALES

13.1 Almacenamiento a largo plazo de cepas bacterianas (gliceroles)

Inocular una colonia (de un crecimiento fresco) en tubo con 5 mL de medio LB pH 7.0 (complementado con 5 μ L de Kan50mg/mL). Posteriormente, incubar por 16 h a 37°C, 220 rpm e inocular a una dilución 1:100 (500 μ L) en un nuevo matraz con 50 mL de LB + 50 μ L de Kan 50mg/mL hasta que el cultivo alcance la fase logarítmica o fase estacionaria.

Colocar en cada vial 1 mL de la solución estéril de glicerol (**Tabla 14**) y adicionar 1 mL del cultivo bacteriano. Es importante mezclar perfectamente la suspensión para evitar la formación de cristales de hielo que pueden disminuir la viabilidad de las células. Congelar rápidamente en etanol-hielo seco y almacenar a -80 °C

Tabla 13. Solución de glicerol para almacenamiento a largo plazo.

Solución de glicerol	(100 mL)
Glicerol	65%
MgSO ₄	0.1 M
Tris Base pH 8	0.025 M
ddH ₂ O	22.5 mL

13.2 Preparación de células quimiocompetentes de *E. coli*

A partir de un glicerol de referencia se realizó un cultivo en una caja con medio LB más agar bacteriológico sin antibiótico de selección, durante 16 h a 37°C. Posteriormente, se tomó una colonia aislada con crecimiento evidente y se cultivó en un tubo con 5 mL de medio LB sin antibiótico de selección, durante 16 h a 37°C y 220 rpm. Después se tomó 1 mL del cultivo para inocular un matraz con 50 mL de medio LB y se incubó a 37°C y 220 rpm hasta alcanzar una DO_{600} de 0.6. El cultivo se centrifugó a 13000 rpm por 10 min a 4°C, se descartó el sobrenadante, la pastilla se recuperó y se resuspendió con 5 mL de $CaCl_2$ (0.1 M) estéril. Después se procedió a incubar en hielo durante 2 h y se centrifugó a 10,000 rpm por 10 min, la pastilla se resuspendió con 5 mL de $CaCl_2$ (0.1 M) estéril y se agregaron 88 μ L de DMSO por cada mL de células en $CaCl_2$. Se hicieron alícuotas de 100 μ L de células competentes y se almacenaron a -80°C hasta su uso (Sambrook et al, 1989).

13.3 Transformación de células químicamente competentes de *E. coli*

La transformación de células quimiocompetentes de *E. coli* (DH5 α) y SoluBL21, se realizó por el método de choque térmico. Una alícuota de 100 μ L de células competentes se mezcló en condiciones de esterilidad con un volumen de plásmido, construcción o ligación según fue el caso, y se incubaron a 4°C durante 30 min en baño de hielo. Posteriormente, se realizó el choque térmico en un baño maría a 42°C durante 45 segundos e inmediatamente después se incubaron a 4°C durante 3 min en baño de hielo. Luego, se adicionaron 450 mL de medio LB sin antibiótico de selección, el cultivo se incubó a 37°C y 220 rpm durante una hora y se sembró en placas con medio LB más el antibiótico de selección a 37°C por 16 h.

13.4 Selección de clonas transformantes de *E. coli* para confirmar la presencia de la construcción de interés.

Las colonias resultantes de la transformación con cada una de las construcciones que presentaron resistencia al antibiótico de selección fueron crecidas en 5 mL de medio LB con la adición de antibiótico (ampicilina 100 μ g/mL o kanamicina 50 μ g/mL) y se incubaron por 16 h a 37°C y 200 rpm en una incubadora con agitación orbital, posteriormente se les extrajo el DNAp utilizando del kit “QIAprep Spin Miniprep” (QIAGEN). Con el DNAp obtenido se hicieron digestiones con las enzimas de restricción XhoI y NcoI para confirmar la presencia del inserto correspondiente al gen Q4.

13.5 Medios y condiciones de cultivo para *E. coli*

Para la obtención de proteína recombinante de *E. coli* se utilizó la cepa SoluBL21. El crecimiento en esta etapa se realizó en el medio LB (1% Triptona, 0.5% Extracto de levadura y 1% NaCl), adicionando 100 µg/mL de ampicilina ó 50 µg/mL de kanamicina como antibiótico de selección. Las placas de selección poseían el mismo medio adicionado, con el antibiótico de selección más 1.5% de agar bacteriológico.

13.6 Análisis de proteínas por SDS-PAGE

Las muestras se analizaron en gel de poliacrilamida al 16% de 0.75 mm de espesor con una cámara de electroforesis vertical *Mini-Protean® II Cell (Bio-Rad)*, utilizando como amortiguador de corrida SDS-Tris-Glicina (25 mM Tris base, 192 mM glicina, 0.1% SDS). Se utilizaron los marcadores de peso molecular *250 (Bio-Rad)*, se inyectaron 10 µL por cada muestra y se corrieron a 85 V por 20 min seguido de 90 V por 150 min. Las bandas se visualizaron por tinción con azul de Coomassie. La fotodocumentación de las imágenes se llevó a cabo utilizando un fotodocumentador *Gel Doc XR + Gel Documentation System (Bio-Rad)*.

Para el análisis de proteínas de las muestras mediante SDS-PAGE, se adicionó 100 µL del amortiguador de muestra 2x SDS-PAGE a 100 µL de muestra y se hirvieron por 5 min. Las muestras procesadas de esta manera se analizaron en un gel de poliacrilamida al 15% de 0.75 mm de espesor en una cámara de electroforesis vertical (*Bio-Rad Mini-Protean® II Cell*). Se utilizaron los marcadores de peso molecular “Precision Plus Protein Standards” (*Bio-Rad*), y las bandas se visualizaron por tinción con el colorante azul de Coomassie.

