



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

**Unidad Zacatenco**

**Departamento de Biotecnología y Bioingeniería**

**Caracterización enzimática de las Peptidil-Prolil Isomerasas  
recombinantes de *Trypanosoma cruzi* (TcCyP19, TcCyP22 y  
TcCySEC) a partir de su expresión en *Escherichia coli***

Tesis que presenta

**Daniel Jiménez Guerra**

Para obtener el grado de

**MAESTRO EN CIENCIAS**

**EN LA ESPECIALIDAD DE BIOTECNOLOGÍA**

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

México, Ciudad de México

Noviembre, 2022

La presente investigación se llevó a cabo en el Laboratorio de Biotecnología de Proteínas (Laboratorio N° 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. Dra. Rossana Arroyo Verástegui Investigadora Titular adscrito al Departamento de Infectómica y Patogénesis Molecular del Cinvestav-IPN, de la Dra. Aurora Antonio Pérez profesora base del Departamento de Bioingeniería del Instituto Tecnológico y de Estudios Superiores de Monterrey – Campus Estado de México (ITESM-CEM), así como del Dr. Rodolfo Marsch Moreno, Investigador Titular en el DBB. 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). Adicionalmente, agradezco al CONACyT por el apoyo económico otorgado para la realización de mis estudios de maestría durante el periodo de septiembre 2020 a agosto 2022 (CONACyT N° CVU 289362).

## AGRADECIMIENTOS

Agradezco a mi director de tesis, el **Dr. Jaime Ortega López** por el apoyo, estoicismo y conocimientos proporcionados durante el desarrollo de este proyecto lleno de retos que no hubiera podido superar sin usted, y por supuesto le agradezco por brindarme todos los recursos para cumplir con este objetivo profesional y personal.

También, agradezco a los miembros de mi comité tutorial conformado por la **Dra. Rossana Arroyo Verástegui**, la **Dra. Aurora Antonio Pérez** y el **Dr. Rodolfo Marsch Moreno** por sus innumerables e invaluable aportes al proyecto, todos y cada uno de ellos me han transportado a este momento, a este logro.

Adicionalmente, agradezco a la **Dra. Rosa Elena Cárdenas Guerra** por su apoyo en las etapas de análisis bioinformático y en la etapa de subclonación, a la **M. en C. Claudia Ivonne Flores Pucheta**, al **IBT. Octavio Flores** por compartirme su valioso conocimiento y experiencia que me permitieron conseguir cada uno de los objetivos de este proyecto. Agradeciendo también a cada uno de ustedes el volver esta etapa en una gran aventura.

A los compañeros y amigos que tuve el privilegio de conocer en el Cinvestav, algunos incluso fuera del DBB, a todos ellos muchas gracias por esos momentos que serán irrepetibles, pero ahora son parte de mi vida. Especialmente gracias a ti **Aimme** mi pareja, mi cómplice y la mejor amiga que jamás he tenido, esto no sería posible sin ti.

Finalmente, gracias a mis padres y hermanos por estar para mí en todo momento, sin ustedes esto nunca hubiera sucedido. Solo encuentro una manera agradecerles, y es seguir creciendo emocional, personal, y profesionalmente con los valores y principios que me han otorgado a lo largo de mi vida. Esto es para ustedes, **nadxiee' laatu stale binni lidxe'**.

## ÍNDICE DE CONTENIDO

ÍNDICE DE TABLAS .....	7
ÍNDICE DE FIGURAS.....	8
GLOSARIO .....	10
RESUMEN .....	12
ABSTRACT.....	13
1. INTRODUCCIÓN .....	14
1.1. Proteínas Recombinantes y su Conformación Nativa.....	14
1.2. Sistemas de Replegamiento de Proteínas.....	15
1.3. Peptidil-Prolil Isomerasas .....	17
1.4. PPIasas de T. cruzi.....	20
2. JUSTIFICACIÓN.....	22
3. HIPÓTESIS.....	22
4. OBJETIVOS.....	23
4.1. Objetivo General.....	23
4.2. Objetivos Específicos.....	23
5. ESTRATEGIA EXPERIMENTAL .....	24
6. MATERIALES Y MÉTODOS.....	26
6.1. Análisis Bioinformático .....	26
6.2. Síntesis de Genes .....	28
6.3. Subclonación de los ORF .....	29
6.4. Análisis de la expresión y solubilidad de las proteínas recombinantes ....	31
6.4.1. Tratamiento de Muestras para SDS-PAGE .....	32

6.4.2.	Análisis por SDS - PAGE.....	32
6.4.3.	Inmunodetección de Proteínas por Ensayos de Western Blot.....	33
6.5.	Producción de PPIasas Recombinantes.....	34
6.6.	Purificación por Cromatografía de Afinidad a Níquel .....	35
6.6.1.	Purificación a partir de Fracción Soluble .....	35
6.6.2.	Purificación a partir de Fracción Insoluble.....	36
6.6.3.	Análisis de las Muestras Colectadas.....	37
6.6.4.	Cuantificación de las Proteínas Purificadas .....	37
6.7.	Ensayo de Actividad de PPIasa por Replegamiento de RNasa T1.....	38
7.	RESULTADOS .....	40
7.1.	Análisis Bioinformático .....	40
7.2.	Síntesis de Genes .....	55
7.3.	Subclonación de ORF.....	55
7.4.	Análisis de la expresión y solubilidad de las proteínas recombinantes ....	60
7.4.1.	Inmunodetección de Proteínas por Ensayo de Western Blot .....	66
7.5.	Producción de Proteínas Recombinantes .....	68
7.6.	Purificación por Cromatografía de Afinidad a Níquel .....	68
7.6.1.	Purificación de TcCyP19 a partir de Fracción Soluble.....	68
7.6.2.	Purificación de TcCyP22 a partir de Fracción Insoluble. ....	72
7.6.3.	Purificación de TcCySEC a partir de Fracción Soluble. ....	75
7.7.	Ensayo de Actividad de PPIasa por Replegamiento de RNasa T .....	78
8.	DISCUSIÓN.....	81
9.	CONCLUSIÓN .....	89

<b>10.</b>	<b>PERSPECTIVAS.....</b>	<b>90</b>
<b>11.</b>	<b>REFERENCIAS.....</b>	<b>91</b>
<b>12.</b>	<b>APÉNDICES.....</b>	<b>97</b>
<b>12.1.</b>	<b>Apéndice 1 .....</b>	<b>97</b>
<b>12.2.</b>	<b>Apéndice 2 .....</b>	<b>98</b>
<b>12.3.</b>	<b>Apéndice 3 .....</b>	<b>99</b>
<b>12.4.</b>	<b>Apéndice 4 .....</b>	<b>99</b>
<b>12.5.</b>	<b>Apéndice 5 .....</b>	<b>99</b>
<b>12.6.</b>	<b>Apéndice 6.....</b>	<b>100</b>
<b>12.7.</b>	<b>Apéndice 7 .....</b>	<b>102</b>
<b>12.8.</b>	<b>Apéndice 8 .....</b>	<b>103</b>
<b>12.9.</b>	<b>Apéndice 9.....</b>	<b>105</b>
<b>13.</b>	<b>METODOLOGÍAS GENERALES .....</b>	<b>106</b>
<b>13.1.</b>	<b>Preparación de células quimiocompetentes de E. coli. ....</b>	<b>106</b>
<b>13.2.</b>	<b>Transformación de células químicamente competentes de E. coli.....</b>	<b>106</b>
<b>13.3.</b>	<b>Selección de clonas transformantes de E. coli para confirmar la presencia de la construcción de interés. ....</b>	<b>107</b>
<b>13.4.</b>	<b>Almacenamiento a largo plazo de cepas bacterianas (gliceroles).....</b>	<b>107</b>
<b>13.5.</b>	<b>Medios de cultivo para E. coli. ....</b>	<b>108</b>

## ÍNDICE DE TABLAS

<b>Tabla 1. Números de acceso de TcCyP19, TcCyP22 y TcCySEC en las diferentes bases de datos.....</b>	<b>40</b>
<b>Tabla 2. Parámetros del análisis de BLAST de TcCyP19, TcCyP22 y TcCySEC Vs PPIasas ya cristalizadas.....</b>	<b>41</b>
<b>Tabla 3. Parámetros Físicoquímicos de TcCyP19, TcCyP22 y TcCySEC determinados en ProtParam.....</b>	<b>46</b>
<b>Tabla 4. Parámetros de validación del modelo 3D de la TcCyP19 obtenido del servidor Swiss-Model y sus modelos 3D optimizados.....</b>	<b>48</b>
<b>Tabla 5. Parámetros de validación del modelo 3D de la TcCyP19 obtenido del servidor trRosetta y sus modelos 3D optimizados. ....</b>	<b>49</b>
<b>Tabla 6. Parámetros de validación del modelo 3D de la TcCyP19 obtenido del servidor I-TASSER y sus modelos 3D optimizados.....</b>	<b>49</b>
<b>Tabla 7. Parámetros de validación del modelo 3D de la TcCyP22 obtenido del servidor Swiss-Model y sus modelos 3D optimizados.....</b>	<b>50</b>
<b>Tabla 8. Parámetros de validación del modelo 3D de la TcCyP22 obtenido del servidor trRosetta y sus modelos 3D optimizados. ....</b>	<b>51</b>
<b>Tabla 9. Parámetros de validación del modelo 3D de la TcCyP22 obtenido del servidor I-TASSER y sus modelos 3D optimizados.....</b>	<b>51</b>
<b>Tabla 10. Solubilidad y Agregación de TcCyP19, TcCyP22 y TcCySEC determinada a partir de su secuencia de aminoácidos.....</b>	<b>54</b>
<b>Tabla 11. Solubilidad y Agregación de TcCyP19, TcCyP22 y TcCySEC determinada a partir de su modelo 3D.....</b>	<b>55</b>
<b>Tabla 12. Concentración de ADNp de las construcciones pUC57-ppiasa .....</b>	<b>56</b>
<b>Tabla 13. Concentración de ADNp, inserto o vector linealizado de clonas seleccionadas.....</b>	<b>57</b>
<b>Tabla 14. Ensayos de Actividad de PPIasa .....</b>	<b>78</b>
<b>Tabla 15. Determinación de la actividad de PPIasa (Eficiencia Catalítica).....</b>	<b>79</b>
<b>Tabla 16. Condiciones de Reacción de Doble Digestión.....</b>	<b>99</b>
<b>Tabla 17. Condiciones de Reacción de Ligación y Controles .....</b>	<b>99</b>

## ÍNDICE DE FIGURAS

<b>Figura 1. Isomerización cis/trans de un enlace peptídico N-terminal a prolina. ....</b>	<b>18</b>
<b>Figura 2. Galería estructural de las cuatro clases principales de PPIasas. ....</b>	<b>19</b>
<b>Figura 3. Secuencia de aminoácidos de las PPIasas de <i>T. cruzi</i> seleccionadas. ....</b>	<b>21</b>
<b>Figura 4. Diagrama General del Proyecto. ....</b>	<b>25</b>
<b>Figura 5. Alineamiento de Secuencias de TCCyP19 y TCCyP22 Vs PPIasas cristalizadas de diversas especies .....</b>	<b>43</b>
<b>Figura 6. Alineamiento de Secuencias de TCCyP19 y TCCyP22 Vs PPIasas cristalizadas de parásitos. ....</b>	<b>44</b>
<b>Figura 7. Alineamiento de Secuencias entre TCCyP19, TcCyP22 y TcCySEC.....</b>	<b>45</b>
<b>Figura 8. Alineamiento de secuencias entre TCCyP19 y TcCyP22.....</b>	<b>45</b>
<b>Figura 9. Modelo 3D de la proteína cristalizada TcMIP. ....</b>	<b>47</b>
<b>Figura 10. Modelo 3D óptimos de la TcCyP19 y TcCyP22 .....</b>	<b>53</b>
<b>Figura 11. Análisis del ADN de las clonas positivas a las construcciones pCri8a-tccyp19, pCri8a-tccyp22 y pCri8a-tccysec. ....</b>	<b>59</b>
<b>Figura 12. Curva de crecimiento de células BL21 (DE3) que expresan rTcCyP19, rTcCyP22 y rTcCySEC.....</b>	<b>61</b>
<b>Figura 13. Análisis por SDS-PAGE de la expresión de la proteína rTcCyP19.....</b>	<b>63</b>
<b>Figura 14. Análisis por SDS-PAGE de la expresión de la proteína rTcCyP22.....</b>	<b>64</b>
<b>Figura 15. Análisis por SDS-PAGE de la expresión de la proteína rTcCySEC. ....</b>	<b>65</b>
<b>Figura 16. Inmunodetección de las proteínas rTcCyP19, rTcCyP22 y rTcCySEC. ....</b>	<b>67</b>



<b>Figura 17. Perfil cromatográfico de la purificación de la proteína rTcCyP19. ....</b>	<b>70</b>
<b>Figura 18. Análisis por SDS-PAGE de la purificación de rTcCyP19 por afinidad a níquel.....</b>	<b>71</b>
<b>Figura 19. Perfil cromatográfico de la purificación de la proteína rTcCyP22. ....</b>	<b>73</b>
<b>Figura 20. Análisis por SDS-PAGE de la purificación de rTcCyP22 por afinidad a níquel.....</b>	<b>74</b>
<b>Figura 21. Perfil cromatográfico de la purificación de rTcCySEC.....</b>	<b>76</b>
<b>Figura 22. Análisis por SDS-PAGE de la purificación de rTcCySEC por afinidad a níquel.....</b>	<b>77</b>
<b>Figura 23. Cinética de Replegamiento de la RNasa T1 utilizando rTcCyP19 o rTcCySEC.....</b>	<b>80</b>
<b>Figura 24. Mapa del Vector de Clonación pU57.....</b>	<b>97</b>
<b>Figura 25. Secuencias de genes y aminoácidos. ....</b>	<b>98</b>
<b>Figura 26. Mapa del Vector de Expresión pCri8a. ....</b>	<b>99</b>
<b>Figura 27. Análisis de la Doble Digestión de pUC57-tccyp19 coli, pUC57-tccyp22 coli y pUC57-tccysec coli.....</b>	<b>104</b>

## GLOSARIO

<b>Abs:</b> Absorbancia	<b>His<sub>6</sub>:</b> Etiqueta de seis histidinas
<b>ADGroEL:</b> Dominio Apical de la chaperonina GroEL	<b>IPTG:</b> Isopropil-B-D-galactopiranosido
<b>ADN:</b> Ácido desoxirribonucleico	<b>kDa:</b> kilo Dalton
<b>ADNp:</b> ADN plasmídico	<b>LB:</b> Luria-Bertani
<b>AU:</b> Unidades de Absorbancia	<b>m:</b> pendiente de la recta
<b>βME:</b> β-Mercaptoetanol	<b>mAU:</b> mili unidades de Absorbancia
<b>CBD:</b> Dominio de Unión a Celulosa	<b>min:</b> minuto (s)
<b>CI:</b> Cuerpos de Inclusión	<b>MIP:</b> Potenciador de la Infectividad a Macrófagos, por sus siglas en inglés.
<b>CsA:</b> Ciclosporina A	<b>μm:</b> Micrómetros
<b>CyPD:</b> Ciclofilina D mitocondrial en humano	<b>μM:</b> Micromolar
<b>DLS:</b> Dispersión Dinámica de Luz	<b>mM:</b> Milimolar
<b>ADNsa:</b> Desoxirribonucleasa	<b>MSA:</b> Alineamiento de Secuencias Múltiple
<b>g:</b> gramo (s)	<b>NCBI:</b> National Center for Biothechnology Information
<b>GFP:</b> Proteína verde fluorescente	<b>nm:</b> nanómetro
<b>GRAVY:</b> Gran promedio de hidropatía	<b>nmol:</b> nano mol
<b>EMBL-EBI:</b> Laboratorio Europeo de Biología Molecular del Instituto Europeo de Bioinformática	<b>OD<sub>600</sub>:</b> Densidad óptica a 600 nm
<b>FKBP:</b> Proteínas de unión a FK506	<b>ORF:</b> Marco Abierto de Lectura
<b>h:</b> hora (s)	<b>PBS:</b> Amortiguador de fosfato salino
	<b>PDB:</b> Protein Data Bank

**pI:** Punto isoeléctrico

**PMSF:** Fluoruro de fenilmetilsulfonilo

**PIasa:** Peptidil-Protil Isomerasas

**PTPA:** Activadores de la fosfatasa PP2A

**RCSB PDB:** Protein Data Bank de RCSB

**RNasa T:** Ribonucleasa T

**rpm:** Revoluciones por minuto

**rTcCyP19:** Proteína recombinante de la TcCyP19 que incluye las etiquetas de fusión proporcionadas por el vector pCri8a

**rTcCyP22:** Proteína recombinante de la TcCyP22 que incluye las etiquetas de fusión proporcionadas por el vector pCri8a

**rTcCySEC:** Proteína recombinante de la TcCySEC que incluye las etiquetas de fusión proporcionadas por el vector pCri8a

**s:** segundo

**sd:** Desviación Estándar por sus siglas en inglés

**SDS-PAGE:** Electroforesis en gel de poliacrilamida y dodecil sulfato de sodio

**SoDoPE:** Soluble Domain for Protein Expression, por sus siglas en inglés

**SWI:** Solubility-Weighted Index, por sus siglas en inglés

**t<sub>1/2</sub>:** Tiempo de vida media

**TEV:** Tobacco Etch Virus

**TriTrypDB:** Base de Datos para Tripanosomátidos

**UFC:** Unidades Formadoras de Colonias

**UniProt:** Base de Datos de Proteínas. UniProt Knowledgebase.

**VC:** Volumen de columna

**EDTA:** Acido etilendiaminotetraacético

## RESUMEN

La producción de proteínas recombinantes es un “tema de gran de interés” en el ámbito farmacéutico, por lo cual la optimización en la producción de estos productos biotecnológicos es importante. El sistema de expresión heteróloga más utilizado es *Escherichia coli* debido a sus diferentes ventajas, sin embargo, existe la desventaja de la expresión de la proteína recombinante en una conformación diferente a la nativa. La TSA-1 es considerada un buen antígeno vacunal contra la enfermedad de Chagas causada por el parásito *Trypanosoma cruzi*; sin embargo, al ser expresada en *E. coli* mal ésta se pliega inadecuadamente conduciéndose a la formación de cuerpos de inclusión. Estudios previos sugieren que el mal plegamiento de la TSA-1 recombinante se debe a la *cis/trans* isomerización de los enlaces que preceden a la prolina. Por lo cual el objetivo de este proyecto fue expresar en *E. coli* algunas peptidil-prolil isomerasas (PPIasas) de *T. cruzi*, purificarlas y determinar su eficiencia catalítica para su uso potencial en el replegamiento de la TSA-1, o cualquier otra proteína de difícil expresión en su conformación nativa que posea varias prolinas. En el presente estudio de las proteínas TcCyP19, TcCyP22 y TcCySEC de *T. cruzi* se realizó un análisis *in silico* para determinar sus parámetros fisicoquímicos, predecir su solubilidad y dominio de PPIasa. Después, se subclonaron los ORF de estas PPIasas en el vector de expresión pCri8a, estos constructos se transformaron en células de *E. coli* BL21 (DE3). La expresión de estas PPIasas se indujo con 1 mM de isopropil- $\beta$ -D-tiogalactopiranosido (IPTG) en 1 L de cultivo a 37°C por 5 h. La TcCyP19 y TcCySEC se expresaron en forma soluble y su purificación se llevó a cabo por cromatografía de afinidad a níquel. Finalmente, los ensayos de actividad de PPIasa de rTcCyP19 y rTcCySEC por replegamiento de la RNasa T1, mostraron que a pH 8.0 y 10 °C la eficiencia catalítica de la rTcCyP19 fue 165,008.4  $\Delta F/s$  nmol, que es mayor al promedio obtenido para la rTcCySEC de 13,051.9  $\Delta F/s$  nmol.

## ABSTRACT

The production of recombinant proteins is a major topic of interest in the pharmaceutical field, so the optimization in the production of these biotechnological products is important. The most widely used heterologous expression system is *Escherichia coli*. Due to its various advantages, however, there is the disadvantage of recombinant protein expression in a different conformation than the native one. TSA-1 is considered a good vaccine antigen against Chagas disease caused by the *Trypanosoma cruzi* parasite; however, however, when expressed in *E. coli* it misfolds leading to the formation of inclusion bodies. Previous studies suggest that misfolding of the recombinant TSA-1 is due to *cis/trans* isomerization of the bonds preceding proline. Therefore, the aim of this project is to express in *E. coli* peptidyl-prolyl isomerases (PPIases) from *T. cruzi*, purify them and determine their catalytic efficiency for potential use in the refolding of TSA-1, or any other protein with several prolines and difficult to express in its native conformation. In the present study of *T. cruzi* TcCyP19, TcCyP22 and TcCySEC proteins, *in silico* analysis was performed to determine their physicochemical parameters, predict their solubility and PPIase domain. The ORFs of these PPIases were then subcloned into the pCri8a expression vector, these constructs were transformed into *E. coli* BL21 (DE3) cells. Expression of these PPIases was induced with 1 mM isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) in 1 L culture at 37°C for 5 h. rTcCyP19 and rTcCySEC were expressed in soluble fraction and purification was carried out by nickel affinity chromatography. Finally, PPIase activity assays of rTcCyP19 and rTcCySEC by RNase T1 refolding showed that at pH 8.0 and 10 °C the catalytic efficiency of rTcCyP19 was 165,008.4  $\Delta F/s$  nmol, that is higher than the average obtained for rTcCySEC of 13,051.9  $\Delta F/s$  nmol.

## 1. INTRODUCCIÓN

### 1.1. Proteínas Recombinantes y su Conformación Nativa

La producción de proteínas recombinantes crece a una tasa anual superior al 7.96% y se estima que para el año 2030 alcanzará un valor mayor a los 566.82 miles de millones de dólares solamente en el sector farmacéutico de este tipo de productos biotecnológicos (Biospace, <https://www.biospace.com/article/protein-therapeutics-market-size-to-surpass-usd-566-82-billion-by-2030/?keywords=novo+nordisk>). En 2017, más de 600 productos biotecnológicos basados en proteínas recombinantes se encontraban aprobados por la Administración de Alimentos y Medicamentos (FDA) de EE.UU. y muchos más estaban pendientes de su aprobación (Usmani et al., 2017). La producción de proteínas recombinantes se ha expandido desde el sistema de expresión heteróloga más conocido y utilizado que es *Escherichia coli*, hasta otras plataformas como levaduras, células de insecto, células vegetales y células de mamífero. Cuando no se requiere de modificaciones postraduccionales *E. coli* continúa siendo la primera opción para la expresión de proteínas recombinantes debido a la disponibilidad de cepas y/o vectores, rápido y fácil control de crecimiento, bajo costo, entre otras ventajas (Sanchez-Garcia et al., 2016). Sin embargo, la mayor limitante de este sistema es cuando la proteína no alcanza su conformación nativa y se acumula en agregados de proteína mal plegada conocidos como Cuerpos de Inclusión (CI) (Antonio-Pérez et al., 2012b). Con el propósito de resolver este problema se han modificado cepas y condiciones de cultivo que facilitan el correcto plegamiento “*in vivo*” y se han desarrollado protocolos para el replegamiento “*in vitro*” de muchas proteínas de gran interés, pero de difícil expresión en su conformación nativa (Antonio-Pérez et al., 2014; Antonio-Pérez et al., 2012a; Antonio-Pérez et al., 2012b; Chow et al., 2006a; Chow et al., 2006b; Shuid et al., 2017). Basado en lo anterior, es necesario contar con protocolos de replegamiento a partir de los CI o protocolos de clonación,

expresión y purificación de dichas proteínas recombinantes con el propósito de producirlas en su conformación nativa, permitiendo así su posterior caracterización y aplicación en los campos de interés para dicho producto de origen biotecnológico.

### **1.2. Sistemas de Replegamiento de Proteínas.**

Estudios previos han demostrado que el dominio apical de la chaperonina GroEL (ADGroEL) en fusión traduccional con el dominio de unión a celulosa (CBD) de *Cellulomonas fimi*, es una alternativa eficiente para la preparación de biocatalizadores de plegamiento (Ramon-Luing et al., 2006). Esta metodología se extendió a las disulfuro oxidoreductasas DsbA y DsbC donde se demostró que, junto con el dominio ADGroEL, inmovilizadas en celulosa asisten el replegamiento cromatográfico oxidativo de la lisozima (Antonio-Perez & Ortega-Lopez, 2010). Adicionalmente, se ha encontrado que la cinética de replegamiento de la lisozima se pueden describir por un modelo de replegamiento sencillo de tres estados, confirmando que la fusión del dominio de unión a celulosa con el dominio ADGroEL y las disulfuro oxido-reductasas DsbA y DsbC es un alternativa eficiente para su purificación e inmovilización en celulosa (Antonio-Pérez et al., 2012b). Se ha observado que la ADGroEL, DsbA y DsbC inmovilizadas en celulosa soportan altas concertaciones de desnaturalizantes y agentes reductores, lo que permite su uso en una columna de replegamiento cromatográfico al menos hasta seis ciclos de replegamiento, sin modificar su capacidad catalítica (Antonio-Pérez et al., 2012a).

Este sistema de replegamiento cromatográfico se probó con éxito para el replegamiento de una hialuronidasa de tarántula, así como la cisteína proteinasa EhCP112, expresadas en *E. coli* como CI. Sin embargo, no se obtuvieron los resultados esperados en el replegamiento de las proteínas recombinantes TvLEGU-1 de *Trichomonas vaginalis* así como la TSA-1 de *Trypanosoma cruzi* (Antonio-Perez et al., 2010), consideradas factor de virulencia y antígenos potenciales para nuevos métodos de diagnóstico y/o el desarrollo

de vacunas para la enfermedad de Chagas en el caso de TSA-1 (Dumonteil et al., 2020; Quijano-Hernandez et al., 2013; Zamudio-Prieto et al., 2014; Zamudio-Prieto et al., 2013) y tricomoniasis para TvLEGU-1 (Ramon-Luing et al., 2010; Rendon-Gandarilla et al., 2013; Resendiz-Cardiel et al., 2017).

La TSA-1 es una proteína de la familia de las transialidasas localizada en la superficie de *T. cruzi*, la cual participa en la interacción huésped-parasito y es un antígeno potencial para el desarrollo de la primera vacuna terapéutica contra la enfermedad de Chagas (Dumonteil et al., 2012; Dumonteil et al., 2022; Dumonteil et al., 2020; Dzul-Huchim et al., 2022; Villanueva-Lizama et al., 2018). La TSA-1 se expresa con altos rendimientos en *E. coli*, pero como CI. Protocolos de replegamiento para obtener esta proteína a partir de CI han permitido demostrar en pruebas pre-clínicas que TSA-1 es efectivamente un buen antígeno vacunal contra la enfermedad de Chagas (Villanueva-Lizama et al., 2018) al igual que otro candidato como el antígeno vacunal, Tc24-C4 (Biter et al., 2018). Sin embargo, las pruebas de estabilidad y el análisis mediante dispersión dinámica de luz (DLS) indican que TSA-1 se encuentra en agregados multiméricos cuasi-estables, por lo que es imperativo producirlas en su conformación nativa con el objetivo de avanzar en su caracterización funcional y en su aplicación como antígenos potenciales para nuevos métodos de diagnóstico y/o el desarrollo de vacunas.

Estudios previos de la TvLEGU-1, otra proteína de gran interés para el diagnóstico de tricomoniasis, en los cuales se determinó la actividad enzimática de la legumina en clonas de *Pichia pastoris* que expresaron esta cisteína proteinasa con diferente número de copias del gen TvLEGU-1. En este estudio se encontró que la actividad enzimática fue mayor en clonas con un bajo número de copias y a una temperatura de inducción sub-óptima (16°C). En clonas con alto número de copias la actividad enzimática fue menor y la proteína recombinante se acumuló posiblemente en el retículo endoplásmico provocando lisis celular (Resendiz-Cardiel et al., 2017). Resultados similares se han obtenido en la

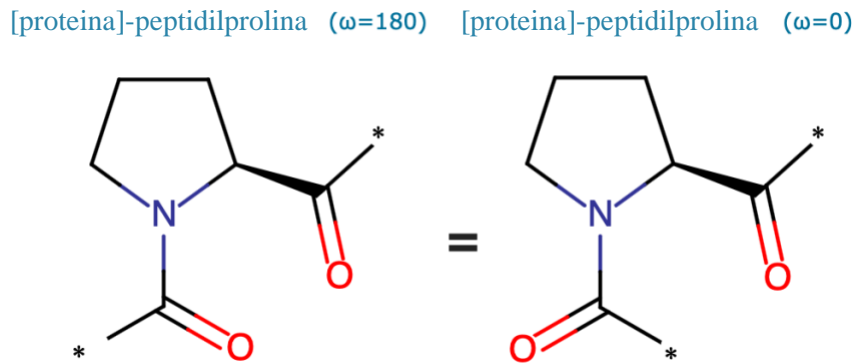


expresión de TSA-1 de *T. cruzi*. Lo anterior sugiere que en las clonas con bajo número de copias e inducción a baja temperatura la síntesis y el plegamiento de la TvLEGU-1 tiene una tasa equivalente que permite secretarlas en su conformación nativa, sin la acumulación de proteína mal plegada.

En clonas con un alto número de copias en *P. pastoris*, la tasa de síntesis de la proteína recombinante rebasa a la tasa del plegamiento, por lo cual la proteína mal plegada se acumula en el retículo endoplásmico hasta provocar lisis celular. TSA-1 y TvLEGU-1 contienen en su secuencia varias cisteínas y un gran número de prolina que posiblemente afecten la cinética de plegamiento (“in vivo”) durante su expresión heteróloga o en su replegamiento libre o asistido. El hecho de que aun en clonas con un alto número de copias se detecte actividad enzimática de esta legumafina sugiere que la cis/trans isomerización de los enlaces que preceden a la prolina, catalizada por las Peptidil-Prolil Isomerasas (PPIasa) de este microorganismo es el paso limitante para el correcto plegamiento de la TvLEGU-1, interviniendo incluso más que la formación de puentes disulfuro (Resendiz-Cardiel et al., 2017).

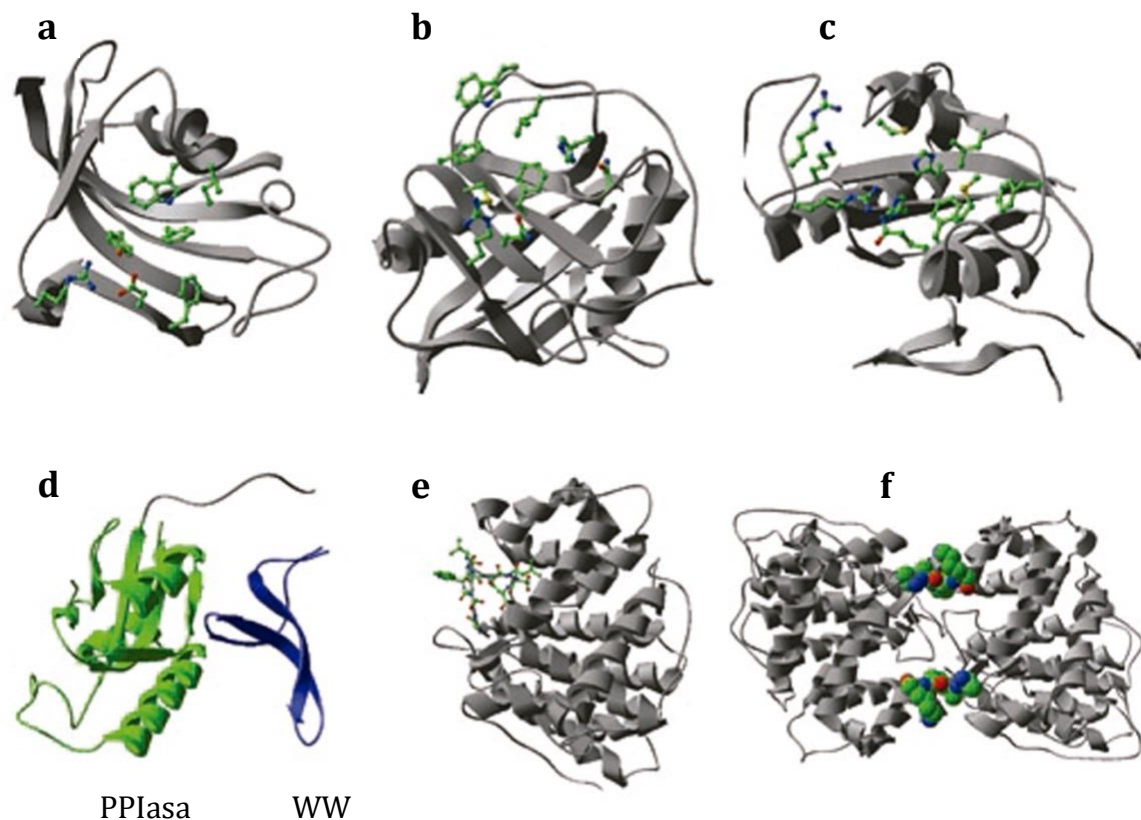
### **1.3. Peptidil-Prolil Isomerasas**

Como se ha mencionado anteriormente, los cambios conformacionales en las unidades que constituyen a una cadena peptídica son esenciales para el correcto proceso de plegamiento de dicha proteína y así finalmente alcance su estructura tridimensional nativa. Se ha establecido que este proceso de plegamiento, además de espontáneo o autónomo, requiere la intervención de enzimas como auxiliares de plegamiento existiendo diversos tipos de estas proteínas que catalizan esta interconversión conformacional. La interconversión de los isómeros cis y trans de los enlaces N-terminal que preceden a un residuo de prolina es catalizado por las peptidil-prolil cis/trans isomerasas (Fischer, 1994; Fischer et al., 1989; Selles et al., 2021) tal cual se representa en la Figura 1.



**Figura 1. Isomerización cis/trans de un enlace peptídico N-terminal a prolina.** Tomado de la información de cualquier PPIasa en UniProt.

Actualmente se sabe, que las PPIasas son proteínas ubicuas y altamente conservadas en su secuencia. Adicionalmente, esta superfamilia de proteínas está constituida por cuatro familias de proteínas no homólogas identificadas como: i) ciclofilinas, ii) proteínas de unión a FK506 (FKBP), iii) activadores de la fosfatasa PP2A (PTPA) y iv) parvulinas (Fanghanel & Fischer, 2004; Jordens et al., 2006; Lu et al., 2007). La diferenciación entre estas cuatro familias podría abordarse considerando la diferencia en su estructura terciaria, su secuencias o características en la misma, organismos en los que se presentan, inhibidores o proteína a la cual se unen, actividad enzimática adicional a PPIasa, entre otros aspectos. Sin embargo, el común denominador es justamente que todas poseen una secuencia de aminoácidos altamente conservada que les otorga la capacidad catalítica para realizar la interconversión entre los *cis/trans* de los enlaces peptidil-prolil (Figura 2).



**Figura 2. Galería estructural de las cuatro clases principales de PPIasas.** (a–f). Se muestran plegamientos globales representativos de las PPIasas FKBP12 (a), CypA (b), Pin1 (c, d) y PTPA (e, f). Aunque FKBP12, CypA y Pin1 no están relacionadas estructuralmente, comparten sitios activos similares (se muestran los residuos clave en color verde), lo que sugiere la posibilidad de que estas enzimas aparentemente no relacionadas puedan compartir características comunes en su acción catalítica. Una característica única de Pin1 es la separación de sus funciones de unión y catálisis en distintos dominios, WW y PPIasa, respectivamente (d). A diferencia de todas las demás PPIasas conocidas, el dominio catalítico de PTPA es un pliegue de hélice alfa (e) y su sitio activo se encuentra en la interfaz de un dímero inducido por sustrato (f). Tomado de (Lu et al., 2007).

#### 1.4. PPIasas de *T. cruzi*

Basado en lo anterior, *T. cruzi* tiene una familia de proteínas ubicuas y altamente conservadas llamadas ciclofilinas perteneciente a la superfamilia de las PPIasas, la cual cuenta con un familia de genes que codifican para 15 parálogos de estas ciclofilinas, de las cuales existen 8 ortólogos que se encuentran en otros microorganismos. El peso molecular de estas PPIasas va de los 19 a los 110 kDa y su localización celular es amplia, encontrándose en el citoplasma, membrana, retículo endoplásmico, en la vía secretora, núcleo, entre otros compartimentos o sitios celulares (Dimou et al., 2017; Moro et al., 1995; Rego et al., 2015). Estas PPIasas se encuentran ampliamente distribuidas en la naturaleza encontrándose en virus, bacterias y eucariotas superiores, teniendo además actividad como factores de virulencia en muchos microorganismos infecciosos (Rego et al., 2015).

Basado en lo descrito en la literatura se han seleccionado tres PPIasas descritas para *T. cruzi* (Perrone et al., 2018; Potenza et al., 2006): la ciclofilina TcCyP19 (Bell et al., 2006; Rego et al., 2015), la ciclofilina TcCyP22 (Perrone et al., 2018) y la FKBP TcCySEC (Moro et al., 1995; Pereira et al., 2002); a continuación se especifica su *RefSeq* o Genbank del NCBI, el número de aminoácidos, su peso molecular, la cepa de la cual proviene esta proteína, y la ubicación de esta PPIasa en el parásito. **TcCyP19** (NCBI Reference Sequence XP\_821578) de 177 aminoácidos, cepa CL Brener, 18,781 kDa y ubicación en el citoplasma de *T. cruzi*. **TcCyP22** (Genbank AAQ55215) de 196 aminoácidos, cepa CL Brener, 21,374 kDa y ubicación en el citoplasma de *T. cruzi*. **TcCySEC** (Genbank PWV17357) de 196 aminoácidos, cepa TCC, 22,136 kDa y ubicación en el citoplasma y membrana de *T. cruzi*, también conocida como TcMIP (Macrophage infectivity potentiator), ya que participa como factor de virulencia en el proceso de infectividad del parásito. La Figura 3 muestra la secuencia de aminoácidos de interés señalando el dominio PPIasa (sombreado verde).

TcCyP19	10	20	30	40	50	60					
	MSYKPHHATV	PTNPKV	FFDV SIGGQSAGRV	VFELFADAVP	KTAENFRALC	TGEKNFGYAG					
	70	80	90	100	110	120					
	SGFHRIIPQF	MCQGGDFTNH	NGTGGRSIYG	EKFADESFAG	KAGKHFGLGT	LSMANAGPNT					
130	140	150	160	170							
						NGSQFFICTA	PTQWLDGKHV	VFGQVLEGIE	VVKAMEAVGS	QTGKTSKPVK	IEASGQL
TcCyP22	10	20	30	40	50	60					
	MFSRTWFWAQ	RKLPFYPI	KNPLV	FFEIS IGAQPAGRVE	MELFKDAVPK	TAENFRALCT					
	70	80	90	100	110	120					
	GEKGVGRSGK	ALCYKGSKFH	RVIPQFMCQG	GDFTNNGTG	GESIYGMKFP	DESFAGRAGK					
	130	140	150	160	170	180					
	HFGPGTLSMA	NAGPNTNGSQ	FFICTAATEW	LDGKHVVFQ	ITKGYEVIK	VEANGSRSGA					
190											
						TRQPILITDC	GEVKNN				
TcCySEC	10	20	30	40	50	60					
	MHRENYFPKM	AFCLLGVLFL	SCITSVQTVS	GDAASHEERM	NNYRKRVGRL	FMEQKAAQPD					
	70	80	90	100	110	120					
	AVKLPSGLVF	QRIARGSGKR	APAI	DDKCEV	HYTGRLRDGT	VFDSSRERK	PTTFRPNEVI				
	130	140	150	160	170	180					
	KGWTEALQLM	REGDRWRLFI	PYDLAYGVTG	GGGMIPPYSP	LEFDVELISI	KDGGKGR	TAE				
190											
							EVDEILRKA	EDREDM			

**Figura 3. Secuencia de aminoácidos de las PPIasas de *T. cruzi* seleccionadas.** Dominio de PPIasa remarcado en color verde. En el caso de la TcCySEC, el dominio MIP se identifica del aminoácido 30 al 196 (identificado por las letras rojas), mientras el péptido señal del aminoácido 1 al 29 remarcado en color morado.

## **2. JUSTIFICACIÓN**

*E. coli* es la plataforma de expresión de proteínas recombinantes de mayor uso en los laboratorios de investigación debido a sus diferentes ventajas. Sin embargo, cuando la sobreexpresión de la proteína recombinante rebasa la maquinaria celular de plegamiento se ocasiona la acumulación de proteína mal plegada, resultando en la formación de CI. Por lo cual, si se caracteriza que las proteínas TcCyP19, TcCyP22 y TcCySEC expresadas de manera recombinante tienen actividad de peptidil-prolil isomerasa, se podría probar que la cis/trans isomerización de prolinas es el paso limitante en el plegamiento de TSA-1 u otras proteínas de difícil expresión en su conformación nativa, y así confirmar que las PPIasas del propio microorganismo (*T. cruzi*) catalizan de manera eficiente la cis/trans isomerización de los enlaces que preceden prolinas. Resultando en un avance importante en la expresión y/o replegamiento de proteínas de difícil expresión como TSA-1, así como otras proteínas con un gran número de prolinas en su secuencia.

## **3. HIPÓTESIS**

Las Peptidil-Prolil-Isomerasas (PPIasa) TcCyP19, TcCyP22 y TcCySEC de *Trypanosoma cruzi*, se expresan de manera recombinante y en su forma activa, en la plataforma de *Escherichia coli*.

## 4. OBJETIVOS.

### 4.1. Objetivo General

- Expresar las PPIasas rTcCyP19, rTcCyP22 y rTcCySEC de *T. cruzi* en *E. coli* y determinar su actividad de peptidil-prolil-isomerasa.

### 4.2. Objetivos Específicos

- Realizar el análisis bioinformático de las proteínas TcCyP19, TcCyP22 y TcCySEC con el objetivo de obtener datos *in silico* que permitan validar, interpretar y optimizar los resultados de la etapa experimental.
- Sub-clonar los ORF que codifican para TcCyP19, TcCyP22 y TcCySEC en el vector de expresión pCri8a, a partir de las secuencias previamente optimizadas para el uso de codones en *E. coli* y clonadas en el vector pUC57.
- Expresar las PPIasas rTcCyP19, rTcCyP22 y rTcCySEC utilizando el vector de expresión pCri8a en *E. coli*.
- Purificar las PPIasas rTcCyP19, rTcCyP22 y rTcCySEC por Cromatografía de Afinidad a Níquel.
- Determinar la actividad enzimática de las PPIasas rTcCyP19, rTcCyP22 y rTcCySEC con el fin de determinar su potencial uso como auxiliares en el plegamiento de las proteínas recombinantes de difícil expresión como TSA-1.

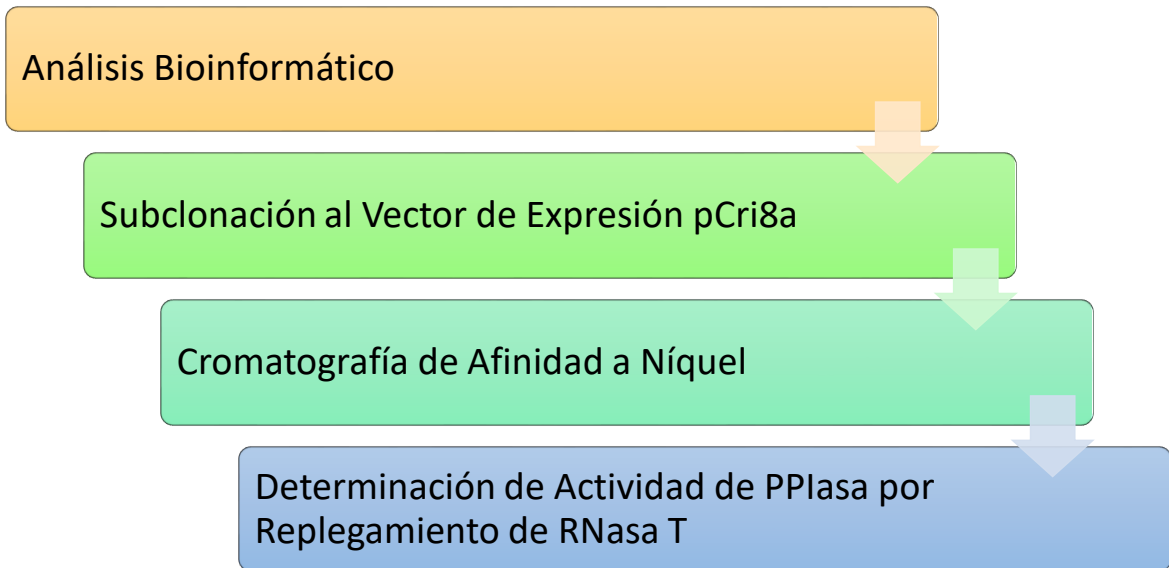
## 5. ESTRATEGIA EXPERIMENTAL

El presente proyecto de investigación se ha dividido en las siguientes etapas (Figura 4):

- I.** Análisis bioinformático de TcCyP19, TcCyP22 y TcCySEC
  - i. Búsqueda de secuencias y números de acceso de las proteínas de interés en las Bases de Datos: NCBI, UniProt, TriTrypDB y RCSB PDB
  - ii. Análisis de BLAST en el servidor del NCBI y TriTrypDB
  - iii. Alineamiento de Secuencias en *Clustal Omega*
    - a) TcCyP19 , TcCyP22 y TcCySEC Vs PPIasas de otros organismos
    - b) TcCyP19, TcCyP22 y TcCySEC Vs PPIasas de otros parásitos
    - c) TcCyP19 Vs TcCyP22 Vs TcCySEC
    - d) TcCyP19 Vs TcCyP22
  - iv. Determinación de Parámetros Físicoquímicos de las proteínas TcCyP19, TcCyP22 y TcCySEC
  - v. Modelado 3D por homología de las proteínas TcCyP19, TcCyP22 y TcCySEC en Swiss-Model, trRosetta, I-Tasser, incluyendo el refinamiento y validación de los modelos obtenidos.
  - vi. Determinación de solubilidad y agregación de las proteínas de interés a partir de su secuencia en los servidores Protein-Sol, SoDoPe, CamSol y Aggrescan, así como de los modelos 3D en los servidores CamSol y Aggrescan 3D.
- II.** Sub-clonación de los Marcos Abiertos de Lectura (ORF) que codifican para las proteínas TcCyP19, TcCyP22 y TcCySEC a partir de la secuencia previamente optimizadas para el uso de codones en *E. coli* y clonadas en el vector pUC57.
- III.** Expresión de las PPIasas recombinantes rTcCyP19, rTcCyP22 y rTcCySEC utilizando el vector de expresión pCri8a.



- IV. Purificación por Cromatografía de Afinidad a Níquel rTcCyP19, rTcCyP22 y rTcCySEC.
- V. Determinación de la actividad de PPIasa de rTcCyP19, rTcCyP22 y rTcCySEC por medio del replegamiento de la RNasa T.



**Figura 4. Diagrama General del Proyecto.**

## 6. MATERIALES Y MÉTODOS

### 6.1. Análisis Bioinformático

La búsqueda de las secuencias de aminoácidos y los números de acceso de las proteínas TcCyP19, TcCyP22 y TcCySEC se realizó en las bases del NCBI (<https://www.ncbi.nlm.nih.gov>), UniProt (<https://www.uniprot.org>), TriTrypDB (<https://tritrypdb.org/tritrypdb/app>) y RCSB PDB (<https://www.rcsb.org>), esto con el fin de obtener los datos necesarios para los análisis bioinformáticos posteriores.

Con el fin de comparar las secuencias de estas proteínas entre ellas mismas así como versus otros organismos se realizó un análisis de BLAST de secuencias de aminoácidos de TcCyP19, TcCyP22 y TcCySEC en el servidor BLAST NCBI (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) y en la base de datos para *T. cruzi* y otros organismos tripanosomátidos, TriTrypDB (<https://tritrypdb.org/tritrypdb/app>).

A continuación, se realizó el alineamiento de las secuencias múltiples (MSA) de aminoácidos de TcCyP19, TcCyP22 y TcCySEC comparadas con otras secuencias de PPIasas utilizando el programa *Clustal Omega* en el servidor EMBL-EBI del Instituto Europeo de Bioinformática (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) esto con el objetivo de evaluar las regiones homólogas y/o conservadas, las sustituciones conservativas, semi-conservativas y no conservativas, así como las deleciones y/o inserciones entre las diferentes secuencias. Los alineamientos se realizaron en los siguientes escenarios: TcCyP19 y TcCyP22 y TcCySEC Vs PPIasas de otros organismos; TcCyP19, TcCyP22 y TcCySEC Vs PPIasas de otros parásitos; TcCyP19 Vs TcCyP22 Vs TcCySEC; y TcCyP19 Vs TcCyP22.

Posteriormente, se utilizaron las secuencias de aminoácidos de las proteínas de interés para determinar los parámetros fisicoquímicos como pI, Coeficiente de extinción a 280 nm medidos en agua, absorbancia 0.1 % considerando que todas las Cisteínas se encuentran

en su estado reducido, Tiempo de vida media ( $t_{1/2}$ ) en *E. coli in vivo*, índice de inestabilidad, índice alifático, gran promedio de hidropatía (GRAVY) entre otros parámetros en el servidor de ProtParam (<https://web.expasy.org/protparam/>)(Gasteiger E., 2005).

Adicionalmente, se realizó la predicción de los modelos tridimensionales por homología, a partir de las secuencias de TcCyP19 y TcCyP22 utilizando tres diferentes servidores: 1) Swiss-Model (<https://swissmodel.expasy.org>), 2) trRosetta (<https://yanglab.nankai.edu.cn/trRosetta/>) y 3) I-TASSER (<https://zhanggroup.org/I-TASSER/>)(Yang et al., 2015). Los modelos obtenidos de estos tres servidores fueron refinados y optimizados en el servidor 3DRefine (<http://sysbio.rnet.missouri.edu/3Drefine/>). Con el propósito de elegir el mejor modelo refinado y optimizado se utilizó la plataforma SAVES v6.0 (<https://saves.mbi.ucla.edu>). Todos los modelos fueron visualizados y personalizados (formato) con el software Pymol 2.5.0 (<https://pymol.org/2/>).

Debido a las diferencias en el fundamento, magnitud, escala y ponderación de cada análisis de validación utilizado (ERRAT, VERIFY 3D, PROVE y PROCHECK) se homologaron los parámetros obtenidos de los cuatro análisis de validación definiendo el modelo 3D con el mejor valor para cada uno de los análisis de validación y se le otorgó el valor de 100. Posteriormente se calculó la desviación para el resto de los modelos 3D respecto al valor de 100 para el mismo análisis. Concluida la homologación para todos los parámetros, se promediaron los valores obtenidos para los cuatro análisis de validación correspondientes al mismo modelo 3D. Finalmente se seleccionó el modelo con el valor promedio más cercano a 100, es decir, el modelo 3D óptimo.

El modelo 3D de la TcCySEC ya se encuentra reportado (Pereira et al., 2002) con el número de acceso **1JVV** en la base de datos RCSB PDB. La secuencia de dicha proteína corresponde a la proteína cristalizada TcMIP (ver sección 6.1 y Figura 9).

La solubilidad en *E. coli* de TcCyP19, TcCyP22 y TcCySEC se determinó con los servidores CamSol Intrinsic (<https://www-cohsoftware.ch.cam.ac.uk/index.php>), Protein-Sol (<https://protein-sol.manchester.ac.uk>) y SoDoPe (<https://tisigner.com/sodope>) los cuales realizan la determinación de acuerdo con la secuencia de aminoácidos de la proteína y a sus características fisicoquímicas como hidrofobicidad/hidrofilicidad y alfa hélice para calcular el perfil de propensión de solubilidad de cada proteína; adicionalmente se determinó la tendencia de agregación utilizando el servidor Aggrescan (<http://bioinf.uab.es/aggrescan/>) que predice la propensión de agregación de secuencias proteicas.

Por otra parte, se realizó el análisis predictivo de solubilidad a partir de los modelos 3D en los servidores CamSol Structurally Corrected (<https://www-cohsoftware.ch.cam.ac.uk/index.php>) que calcula el perfil de solubilidad basado en la estructura, considerando la proximidad entre los aminoácidos y su exposición a disolventes; así como el análisis predictivo de la propensión de agregación en estructuras de proteínas en Aggrescan 3D (<http://biocomp.chem.uw.edu.pl/A3D2/>).

## 6.2. Síntesis de Genes

Se solicitó a la compañía Synbio Technologies LLC la síntesis y clonación por EcoRV de los genes que codifican para TcCyP19, TcCyP22 y TcCySEC utilizando el vector pUC57 (**Apéndice 1**) que es un vector de clonación de plásmidos de uso común en *E. coli*. La longitud del vector es de 2710 pb y se aísla de la cepa DH5 $\alpha$  de *E. coli* mediante procedimientos estándar. (GenScript, 2021; SnapGene, 2021). En el caso del gen que codifica para TcCySEC no se le incluyó el péptido señal presente en TcMIP por lo que la secuencia sintetizada fue de 515 pb a diferencia del gen que original que posee 596 pb.

El **Apéndice 2** describe las secuencias de los genes que codifican para cada PPIasa recombinante contenidos dentro del vector pUC57, entre los sitios de restricción NcoI

(5'→3', ccatgg) y XhoI (3'→5', ctcgag) añadidos en los extremos de la secuencia. Todas las secuencias de los genes han sido optimizadas para el uso de codones en *E. coli* especificando en cada secuencia de ADN la región optimizada. Adicionalmente se muestra la secuencia de aminoácidos de cada proteína recombinante señalando el dominio de PPIasa.

### 6.3. Subclonación de los ORF

Los ORF que codifican para TcCyP19, TcCyP22 y TcCySEC se subclonaron en el vector de expresión **pCri8a** basado en la literatura revisada (Goulas et al., 2014), así como la experiencia y disponibilidad de vectores de expresión en el grupo de trabajo. El **Apéndice 3** muestra las principales características del vector de expresión pCri8a, incluyendo las etiquetas de seis Histidinas (His<sub>6</sub>), linker y sitio de corte para la proteínasa TEV, así como el peso molecular que cada secuencia le otorga a la proteína recombinante.

Células químicamente competentes de la cepa DH5α de *E. coli* se transformaron con la construcción pUC57-*ppiasa* (-*tccyp19*, -*tccyp22* o -*tccysec*) mediante choque térmico para su propagación. Adicionalmente se transformaron células DH5α para el vector pCri8a. Las células transformantes con la construcción pUC57-*ppiasa* se sembraron en cajas con agar Luria-Bertani (LB) + ampicilina [100µg/mL] mientras que las correspondientes al vector se sembraron en cajas con agar LB + kanamicina [50µg/mL] incubando todas durante 16 h a 37°C. Posteriormente se tomó una sola colonia de cada plásmido y se inoculó en los correspondientes tubos de ensayo con 5 mL de medio LB + kanamicina [50µg/mL], se realizó la purificación del DNA plasmídico (ADNp) utilizando el Kit de Purificación de ADNp (midiprep) (MACHEY-NAGEL, 2021). El ADNp purificado se cuantificó por medición de absorbancia a 260 y 280 nm en el equipo NanoDrop 2000 (*Thermo-Scientific*) evaluando adicionalmente la calidad del ADNp por electroforesis en gel de agarosa al 1%.

La Doble Digestión de 3µg de ADNp (Inserto y Vector) se realizó en los sitios de restricción NcoI y XhoI generando extremos cohesivos, procedimiento realizado a 37°C, 250 rpm durante 16 h; los componentes, volúmenes y condiciones de la reacción de doble digestión se detallan en el **Apéndice 4**. Los productos resultantes de la doble digestión se analizaron por electroforesis en gel de agarosa al 2%. Las bandas correspondientes a los insertos (*tccyp19*, *tccyp22* y *tccysec*) y al vector pCri8a linealizado se recuperaron de acuerdo con el protocolo de extracción de ADNp en gel con centrifuga del Kit Spin QIAquick® (QIAGEN, 2020b). El ADN extraído se cuantificó por medición de absorbancia a 260 y 280 nm en el equipo NanoDrop 2000, evaluando también su calidad por electroforesis en gel de agarosa al 1%.

Utilizando los valores de concentración obtenidos previamente y considerando el tamaño del inserto de interés y del vector, se realizó el cálculo de femto moles (fmol) que contiene 1 µL de muestra. Posteriormente, se procedió a la ligación de los ORF objetivo dentro del vector de expresión pCri8a, realizando todas las ligaciones en una relación molar 3:1 (inserto/vector), utilizando la ADN T4 ligasa (NEB) y de acuerdo con lo descrito en el **Apéndice 5**. Al concluir la incubación se inactivó a la T4 Ligasa calentando a 65°C durante 10 min e inmediatamente se transformaron por choque térmico en células DH5α de *E. coli* quimiocompetentes. La transformación se realizó para los productos de ligación, así como a los controles de religación y ADNp, utilizando el protocolo estándar para obtención de transformantes. La siembra se realizó en placas con agar LB + kanamicina [50µg/mL] incubando durante 16 h a 37°C. A las colonias crecidas se les realizó la extracción de ADNp (miniprep) basado en lo descrito en el protocolo de QIAprep® Miniprep (QIAGEN, 2020a). La calidad del ADNp se evaluó por electroforesis en gel de agarosa al 1% y la cuantificación por medición de absorbancia a 260 y 280 nm en el equipo NanoDrop 2000. La doble digestión del ADNp se realizó en un volumen total de reacción de 15 µL utilizando las

enzimas NcoI y XhoI a 37°C, 250 rpm durante 1 h, analizando el patrón de digestión por electroforesis en gel de agarosa al 2% para discriminar las clonas recombinantes. Las clonas que resultaron positivas con las construcciones pCri8a-*ppiasa* (-*tccyp19*, -*tccyp22* o -*tccysec*) se prepararon para su almacenamiento a -80 °C hasta su uso.

#### **6.4. Análisis de la expresión y solubilidad de las proteínas recombinantes**

Células de *E. coli* BL21 (DE3) se transformaron mediante procedimientos estándar utilizando ADNp con la construcción de interés (pCri8a-*ppiasa*). De las clonas obtenidas se prepararon cultivos con medio LB + kanamicina (50 µg/mL) los cuales se conservaron en alícuotas de 250 µL + 250 µL glicerol al 80%, almacenando las muestras con la pCri8a-*ppiasa* (-*tccyp19*, -*tccyp22* o -*tccysec*) a -80°C.

Se realizó una Prueba de Inducción a partir de un cultivo bacteriano con 5 mL de medio LB + kanamicina (50 µg/mL) partiendo de uno de los gliceroles previamente elaborados y conservados a -80 °C. inoculando a una razón de 1:100 (Glicerol-medio LB), se incubó a 200 rpm y 37 °C, durante el tiempo necesario hasta alcanzar una Densidad Óptica a 600 nm (OD<sub>600</sub>) entre 0.6 y 1.0 unidades de absorbancia (AU). Al alcanzar el valor de OD<sub>600</sub> en este rango se tomó el volumen de muestra equivalente a 1.0 AU en tubos de 1.5 mL, que se centrifugaron a 13,000 rpm, a 4 °C durante 5 min. Posteriormente se agregó al cultivo IPTG a una concentración final de 1 mM incubando bajo las mismas condiciones de crecimiento por 16 h. Al concluir el periodo de inducción, se midió la OD<sub>600</sub> de cada cultivo bacteriano con las construcciones de interés, tomando las muestras para su posterior tratamiento y análisis por SDS-PAGE.

Una vez confirmado que las células BL21 (DE3) expresaban las proteínas rTcCyP19, rTcCyP22 y rTcCySEC se realizó un escalamiento del proceso para estandarizar las condiciones y tiempo de inducción, así como confirmar en que fracción se encontraba la proteína de interés. Se utilizó un pre-inóculo en tubo con 7 mL de medio LB + kanamicina

(50 µg/mL) al cual se le adicionó 10 µL del glicerol de células *E. coli* BL21 (DE3) con la construcción incubando a 200 rpm y 37 °C durante 16 h, se midió la OD<sub>600</sub> del pre-inóculo y se transfirieron a 5 mL a un matraz de 250 mL con 100 mL de medio LB + kanamicina (50 µg/mL). Posteriormente, se midió la OD<sub>600</sub> hasta obtener un valor entre 0.4 y 0.7 AU, una vez alcanzado se tomó una muestra del medio equivalente a 1 AU y se adicionó al cultivo IPTG a una concentración final de 1 mM. Incubando en las mismas condiciones de crecimiento por 5 h., se midió la OD<sub>600</sub> cada hora posterior a la inducción con IPTG, tomando las muestras para su tratamiento, análisis por SDS-PAGE e inmunodetección.

#### **6.4.1. Tratamiento de Muestras para SDS-PAGE**

Las muestras correspondientes a la Fracción No Inducida (FNI) se centrifugaron a 13,000 rpm a 4 °C durante 5 min, se adicionó 45 µL de Bug Buster (Merck) a la pastilla obtenida, resuspendiendo y agitando a temperatura ambiente por 20 min. Al concluir el periodo de inducción se tomaron dos muestras cuyo volumen sea equivalente a 1.0 AU, correspondientes a la Proteína Total (PT) y a la Fracción Insoluble (FI), las cuales se centrifugaron a 13,000 rpm, a 4 °C por 5 min. A continuación, se agregó a cada muestra 45 µL de Bug Buster (*Merk*), resuspendiendo y agitando a temperatura ambiente por 20 min. Los tubos destinados para el análisis de la Fracción Soluble (FS) y FI se centrifugaron a 13,000 rpm a 4 °C por 30 min y al finalizar se transfirió el sobrenadante (FS) a otro tubo de 1.5 mL. Al pellet obtenido (FI) se agregaron 45 µL de Buffer de Solubilización (8M Urea) resuspendiendo con pipeta y vortex. Se agregaron 15 µL de Buffer de LaemmLi 4x con 10 % de β-mercaptoetanol (βME) para trabajar todas las muestras en condiciones reducidas.

#### **6.4.2. Análisis por SDS - PAGE**

Las muestras se analizaron en gel de poliacrilamida al 15% de 0.75 mm de espesor con una cámara de electroforesis vertical *Mini-Protean® II Cel (Bio-Rad)*, utilizando como amortiguador de corrida SDS-Tris-Glicina (25 mM Tris base, 192 mM glicina, 0.1% SDS).



Se utilizó como marcador de Peso Molecular Precision Plus Protein™ (*Bio-Rad*), se inyectaron 8 µL por cada muestra previamente tratada de acuerdo con la sección 6.4.1 y se corrieron a 85 V por 20 min seguido de 90 V por 150 min. Las bandas se visualizaron por tinción con azul brillante de Coomassie. Posteriormente, se cambió la solución por solución desteñidora, realizando los cambios de solución necesarios hasta que se apreciaron claramente las bandas de cada carril. La documentación de las imágenes se llevó a cabo utilizando un fotodocumentador *Gel Doc XR + Gel Documentation System* (*Bio-Rad*), utilizando la lámpara de luz visible.

#### **6.4.3. Inmunodetección de Proteínas por Ensayos de Western Blot**

La inmunodetección se llevó a cabo mediante el ensayo Western Blot utilizando anticuerpos  $\alpha$ -6xHis. Las muestras resueltas por SDS-PAGE en un gel al 15 % se transfirieron a una membrana de nitrocelulosa mediante electrotransferencia durante 20 min a 40 V, seguido de 60 min a 100 V. Al término, la membrana se tiñó con una solución de Rojo de Ponceau S para verificar la completa transferencia de las proteínas. Posteriormente, se realizaron lavados con PBS-T (PBS 1x y 0.1% de Tween 20) para la eliminación del colorante y se enjuagó con PBS 1x (138 mM NaCl, 3 mM KCl, 8.1 mM Na<sub>2</sub>HPO<sub>4</sub> y 1.5 mM KH<sub>2</sub>PO<sub>4</sub>). A continuación, se incubó la membrana durante 1 h con solución de bloqueo (PBS-T y leche descremada en polvo al 10% p/v) a 37 °C y 100 rpm; al concluir se realizaron lavados con PBS-T y se incubó durante 18 h con el anticuerpo primario “Mouse  $\alpha$ -His<sub>6</sub>” (Invitrogen) a una dilución de 1:3,000 en una solución de incubación (PBS-T y 5% p/v de leche descremada en polvo), a 4 °C y 100 rpm, realizando lavados con PBS-T. Al término, se incubó la membrana por 1 h a temperatura ambiente utilizando el anticuerpo secundario “Goat  $\alpha$ -Mouse” (Invitrogen) acoplado a la peroxidasa de rábano picante (HRP), a una dilución 1:3,000 en solución de incubación (PBS-T y 2.5%

p/v de leche descremada en polvo), a 37 °C y 100 rpm, realizando lavados con PBS-T. El revelado de la señal inmunorreactiva se realizó por colorimetría, incubando la membrana con una mezcla de dos soluciones previamente preparadas (ambas protegidas de la luz): A) 12.5 mL de TBS 1x (1 mM Tris-HCl pH 7.5 y 500 mM de NaCl) + 7.5 µL de H<sub>2</sub>O<sub>2</sub> y B) 2.5 mL de metanol (frío a -20 °C) + 7.5 mg de HRP. La reacción se detuvo desechando la mezcla de revelado y sustituyéndola por agua destilada.

### **6.5. Producción de PPIasas Recombinantes**

Una vez estandarizadas las condiciones de crecimiento y tiempo de inducción se procedió al escalamiento del proceso de producción de las proteínas recombinantes de interés partiendo de un pre-inóculo en un matraz de 250 mL con 100 mL de medio LB + kanamicina (50 µg/mL) al cual se le adicionaron 200 µL de glicerol de células *E. coli* BL21 (DE3) con la construcción de interés, incubando a 200 rpm y 37 °C durante 22 h. Se monitoreó la OD<sub>600</sub> del pre-inóculo y posteriormente se transfirió el cultivo correspondiente al pre-inóculo a un matraz Tunair™ de 2.5 litros que contenían 1 L de medio *Terrific Broth* (TB) + kanamicina (50 µg/mL). Se monitoreo la OD<sub>600</sub> hasta obtener un valor entre 0.4 y 0.7 AU. En este punto se tomó una muestra cuyo volumen sea equivalente a 1.0 AU, correspondiente a la FNI, para su posterior tratamiento (ver sección 6.4.1) y análisis por SDS-PAGE (ver sección 6.4.2). Posteriormente, se adicionó al cultivo IPTG a una concentración final de 1 mM, incubando a 200 rpm y 37 °C durante 5 h y se midió la OD<sub>600</sub> del cultivo bacteriano. La cosecha se realizó centrifugando el volumen total del cultivo bacteriano a 9,000 rpm a 10 °C durante 20 min. Al pellet resultante se le realizó un lavado con PBS 1x a razón de 5 mL por gramo de biomasa, centrifugando posteriormente a 4200 rpm a 10 °C durante 20 min.

Al pellet celular se le adicionó Buffer de Lisis (Tris-HCl pH 8.0 50 mM, NaCl 20 mM, βME 1 mM, PMSF 1 mM, Glicerol 10%) a razón de 10 mL/g de biomasa, se incubó durante 30 min

a 37°C y 200 rpm y al finalizar se adicionó lisozima a una concentración final de 0.5 mg/mL de lisado, incubando bajo las mismas condiciones. Finalmente, se adicionó ADNasa a una concentración final de 1 µg/mL de lisado, incubando bajo las mismas condiciones. El lisado resultante se colocó en un sonicador trabajando al 40% de amplitud en tiempos de encendido/apagado de 30 s “ON” y 30 s “OFF” durante 6 min. En este punto se tomó una muestra de 45 µL de lisado, correspondiente a la FS y FI, para su posterior tratamiento (ver sección 6.4.1) y análisis por SDS-PAGE (ver sección 6.4.2). Posteriormente las fracciones soluble e insoluble del lisado se separaron por centrifugación a 13,000 rpm durante 30 min a 10 °C. El sobrenadante, correspondiente a la FS, se filtró por una membrana de 0.45 µm y posteriormente por una membrana de 0.22 µm para proceder con la etapa de purificación (ver sección 6.6.1). Al pellet obtenido, correspondiente a la FI, se le realizaron lavados y solubilización para su posterior purificación (ver sección 6.6.2), lo anterior considerando la fracción en que se encuentre la proteína recombinante de interés y basado en los resultados obtenidos en el análisis de la expresión y solubilidad de las proteínas recombinantes.

## **6.6. Purificación por Cromatografía de Afinidad a Níquel**

Debido a que el vector de expresión pCri8a añade a las proteínas recombinantes etiquetas His<sub>6</sub> en los extremos NH<sub>2</sub>- y COOH-, se utilizó como método de purificación la Cromatografía por Afinidad a Níquel tanto si se partió de la FS o FI en que se encuentren las proteínas rTcCyP19, rTcCyP22 y rTcCySEC. En ambas circunstancias, la purificación se realizó en un cromatógrafo NGC Chromatography System (*Bio-Rad*).

### **6.6.1. Purificación a partir de Fracción Soluble**

Se utilizó una columna pre-empacada HisTrap™ High Performance (GE) de 1 mL y bajo un esquema de purificación considerando el volumen de columna (VC) requerido para las etapas de equilibrio, aplicación de la muestra, lavados y elución. Siento un esquema de

purificación individualizado para cada proteína recombinante; utilizando en todas las situaciones dos Buffers (A y B) los cuales estaban constituidos por Tris-HCl 50 mM pH 8.0, NaCl 200 mM,  $\beta$ ME 1 mM y PMSF 1 mM; variando únicamente su concentración de Imidazol, siendo 5 mM en el Buffer A y 500 mM en el Buffer B. Ambos Buffer se filtraron por 0.22  $\mu$ m y se almacenaron a 4 °C hasta su uso.

### **6.6.2. Purificación a partir de Fracción Insoluble**

Los lavados de la FI se realizaron a una razón de 10 mL del Buffer aplicable por gramo de biomasa; realizando entre cada lavado la resuspensión completa del pellet con la ayuda de una pipeta, seguido de agitación durante 20 min a temperatura ambiente y finalmente centrifugación a 13,000 rpm a 10 °C durante 30 min. Primero: tres lavados utilizando un Buffer que contiene Urea 2M, NaCl 100 mM, Tritón X-100 0.1% y Tris-HCl pH 8.0 20 mM. A continuación: tres lavados utilizando un Buffer que contiene Cloruro de Guanidinio 2M, DTT 1 mM y Tris-HCl pH 8.0 20 mM. Finalmente, tres lavados utilizando un Buffer que contiene Tritón X-114 1%, DTT 1 mM y Tris-HCl pH 8.0 20 mM.

Al concluir los lavados de la FI se procedió a la resuspensión de la muestra utilizando un Buffer de Solubilización que contiene Urea 8M, Tris-HCl pH 8.0 50 mM, NaCl 200 mM,  $\beta$ ME 1 mM e Imidazol 5 mM; utilizando 5 mL de dicho Buffer de Solubilización por gramo de biomasa. Posteriormente, se centrifugó la muestra a 13,000 rpm durante 30 min a 10 °C y se filtró el sobrenadante por una membrana de 0.45  $\mu$ m y finalmente por una de 0.22  $\mu$ m. Posteriormente, se procedió a la purificación utilizando una columna con 1 mL de resina *GE Fast Flow 6 Sepharose Ni* y bajo un esquema de purificación considerando el VC requerido para las etapas de equilibrio, aplicación de la muestra, lavados y elución. Siento un esquema de purificación individualizado para cada proteína recombinante; utilizando en todas las situaciones dos Buffers (A y B) los cuales estaban constituidos por Urea 8M, Tris-HCl 50 mM pH 8.0, NaCl 200 mM y  $\beta$ ME 1 mM; variando únicamente su concentración

de Imidazol, siendo 5 mM en el Buffer A y 500 mM en el Buffer B. Ambos Buffer se filtraron por 0.22  $\mu\text{m}$  y se almacenaron a 4 °C hasta su uso.

### **6.6.3. Análisis de las Muestras Colectadas**

Basado en los cromatogramas obtenidos de cada purificación, se trataron y analizaron por SDS-PAGE las muestras colectadas que presentaron un pico o cresta en los valores de Absorbancia a 280 nm (mAU). El tratamiento de las muestras y análisis por SDS-PAGE se realizó conforme a lo indicado en las secciones 5.4.1 y 5.4.2, respectivamente.

### **6.6.4. Cuantificación de las Proteínas Purificadas**

Las fracciones colectadas en la purificación, que contenían a las proteínas recombinantes de interés (rTcCyP19, rTcCyP22 y rTcCySEC), se encontraban en el Buffer de Elución (Buffer B) por lo cual se eliminó el imidazol que interfiere en el proceso de cuantificación al transferirlas a un Buffer de Almacenamiento (Tris-HCl 20 mM pH 8.0, NaCl 100 mM, Azida de Sodio 0.02%, Glicerol 10%) utilizando una columna PD-10 de acuerdo con el protocolo por gravedad indicado en el Manual de PD-10 Desalting Columns (GE-Healthcare, 2007). Las muestras colectadas de la PD-10 se cuantificaron por densitometría en gel de poliacrilamida al 15% por métodos estandarizados en el equipo de trabajo. En resumen, se cargó una proteína estándar en cinco concentraciones ascendentes, más las muestras de la proteína a cuantificar en diferentes diluciones, posteriormente se resolvió por SDS-PAGE un gel al 15%. Una vez concluido, se realizó el teñido, desteñido y obtención de una imagen del gel con un fotodocumentador *Gel Doc XR + Gel Documentation System* (Bio-Rad) utilizando la lámpara de luz visible. La fotografía obtenida se analizó utilizando el software ImageLab (Bio-Rad) analizando intensidad de las bandas de esta proteína estándar y generando una curva estándar con un valor de  $R^2 > 0.99$  a partir de los valores de concentración ya conocidos, considerando el volumen e intensidad de la banda el software otorga un valor de concentración a cada muestra de la proteína de interés

resuelta en el gel. Finalmente, para la determinación de concentración final se debe considerar el factor de dilución y debe ser similar entre todas las muestras independientemente de dicho factor. Los valores determinados se expresaron en unidades micromolar ( $\mu\text{M}$ ). A partir de la concentración de cada muestra se determinó el volumen de muestra requerido para el ensayo de actividad de PPIasa. Adicionalmente se realizó un análisis por Dispersión Dinámica de Luz (DLS) para conocer el tamaño de las proteínas recombinantes obtenidas utilizando el equipo Zetasizer Nano (*Malvern*).

### **6.7. Ensayo de Actividad de PPIasa por Replegamiento de RNasa T1**

Las fracciones colectadas en la purificación, que contenían a las proteínas recombinantes de interés (rTcCyP19, rTcCyP22 y rTcCySEC), se encontraban en el Buffer de Elución (Buffer B) las cuales se transfirieron a un Buffer de Replegamiento (Tris-HCl 20 mM pH 8.0, NaCl 100 mM, Azida de Sodio 0.02%, Glicerol 10%) utilizando una columna PD-10 de acuerdo con el protocolo por gravedad indicado en el Manual de PD-10 Desalting Columns (GE-Healthcare, 2007). Las muestras colectadas de la PD-10 se cuantificaron por densitometría utilizando gel de poliacrilamida al 15% utilizando procedimientos estándar. Los valores determinados se expresaron en unidades micromolar ( $\mu\text{M}$ ), y a partir de la concentración de cada muestra se determinó el volumen de muestra requerido para el ensayo de actividad de PPIasa.

La RNasa T1 se seleccionó para el ensayo de actividad de PPIasa, ya que es una proteína de 104 residuos que contiene cuatro enlaces peptidil-prolil, dos de estos preceden a la Pro-39 y Pro-55, que se encuentran en su conformación cis en la proteína nativa. Esta proteína posee una lenta cinética de plegado, la cual está determinada por la isomerización de estos dos residuos. Adicionalmente, en la proximidad de los mencionados enlaces cis peptidil-prolil, la RNasa T1 contiene el Trp-59, que puede utilizarse para monitorear el ensayo por fluorescencia. Se utilizó RNasa T (Invitrogen) stock a una concentración de  $50\mu\text{M}$

desnaturalizándola en una concentración final de 8 M de Urea durante 16 h a temperatura ambiente.

El ensayo de actividad se realizó utilizando en una celda de 1 mL donde se añadió 50 - 75  $\mu\text{M}$  de la proteína recombinante a evaluar (rTcCyP19, rTcCyP22 o rTcCySEC), contenida en Buffer de Replegamiento (BR: Tris-HCl 20 mM pH 8.0) y finalmente RNasa T desnaturalizada a una concentración final de 2  $\mu\text{M}$ . Este ensayo incluyó tres condiciones: Control negativo (-) que contenía a la proteína recombinante + BR; Blanco sin PPIasa constituido por RNasa T + BR; y Condición problema compuesto por la proteína recombinante + BR + RNasa T. En cualquiera de las condiciones, la mezcla de componentes se realizó utilizando una pipeta y realizando la medición inmediatamente.

El replegamiento de la RNasa T se monitoreó mediante la fluorescencia de triptófano (longitud de onda de excitación de 268 nm, slit de 2 nm y longitud de onda de emisión de 325 nm, slit de 5 nm) a 10 °C utilizando un espectrofluorómetro FluoroMax-3 (JOBIN YVON, HORIBA). La eficiencia catalítica de la proteína de interés para el replegamiento de RNasa T1 se calculó a partir de la parte lineal del gráfico obtenido, resultando de la diferencia entre el valor de las pendientes (m, expresada en IFR/s) de la condición problema menos la condición blanco sin PPIasa y dividido entre el número de nano moles (nmol) contenido en 1 mL de reacción de la condición evaluada, expresando el resultado como  $\Delta\text{IFR s}^{-1} \text{ nmol}^{-1}$ . El protocolo anterior se fundamentó en lo descrito para protocolos para evaluar la actividad de PPIasa a través del replegamiento de la RNasa T1 (Bose et al., 1994; Zemanova et al., 2020).

## 7. RESULTADOS

### 7.1. Análisis Bioinformático

Con el propósito de iniciar el análisis bioinformático previamente se obtuvo información importante respecto a las secuencias de TcCyP19, TcCyP22 y TcCySEC en diferentes bases de datos. Como resultado de esta búsqueda se encontró información relevante respecto a las ciclofilinas que se analizaron experimentalmente en el presente proyecto. Estos resultados incluyeron la existencia de otras ciclofilinas de *T. cruzi*, incluyendo las PPIasas de interés en este proyecto. Adicionalmente se confirmó que la TcCySEC ya ha sido cristalizada (PDB ID 1JVW), por lo cual fungió como ciclofilina de referencia para su comparación respecto a la TcCyP19 y TcCyP22 en las diferentes etapas del presente proyecto. En la Tabla 1 se indican los números de acceso obtenidos de la búsqueda de cada secuencia en las diferentes bases de datos:

**Tabla 1. Números de acceso de TcCyP19, TcCyP22 y TcCySEC en las diferentes bases de datos.**

PPIasa	NCBI	UniProt	TriTrypDB Gene	PDB
<b>TcCyP19</b>	XP_821578.1	Q4E4L9	TcCLB.506925.300	No Cristalizada
<b>TcCyP22</b>	AAQ55215.1	Q6V7K7	TcCLB.510259.50	No Cristalizada
<b>TcCySEC/TcMIP</b>	PWV17357.1	Q09734	C3747_2g288	1JVW

Se realizó un análisis de BLAST de las secuencias de TcCyP19, TcCyP22 y TcCySEC, en los servidores NCBI y TriTrypDB con el propósito de comparar las proteínas de interés con otras proteínas ya cristalizadas y posteriormente identificar aquellas con la mayor identidad en la secuencia, así como otros parámetros. Los resultados obtenidos arrojaron resultados interesantes respecto a la homología de las proteínas de interés con PPIasas ya cristalizadas. Para la TcCyP19 existen cuatro ciclofilinas con los mejores parámetros



obtenidos del análisis de BLAST. En el caso de la TcCyP22 se obtiene como resultado que la PPIasa con los mejores parámetros es 2HQJ de *L. major*, siendo este resultado consistente en ambos servidores. En el caso de TcCySEC se identificó previamente que ya ha sido cristalizada y la secuencia corresponde a la de TcMIP sin el péptido señal, por lo cual los resultados obtenidos en el análisis de BLAST son consistentes. En la Tabla 2 se indican las ciclofilinas que poseen los mejores parámetros para cada PPIasa objetivo.

**Tabla 2. Parámetros del análisis de BLAST de TcCyP19, TcCyP22 y TcCySEC Vs PPIasas ya cristalizadas.**

PPIasa	PDB ID	Descripción	Query Cover (%)	E-value	Indentidad (%)
<b>TcCyP19</b>	6FK1_A	PPIasa de <i>M. musculus</i>	92	4e-84	73.05
	1AK4_A	PPIasa de <i>H. sapiens</i>	92	4e-83	72.29
	4N1M_A	PPIasa de <i>H. sapiens</i>	92	5e-83	72.29
	2X25_B <sup>a</sup>	PPIasa de <i>H. sapiens</i>	97	-	72.00
<b>TcCyP22</b>	2HQJ_A	PPIasa de <i>L. major</i>	90	2e-89	71.75
<b>TcCySEC / TcMIP</b>	1JVW_A	PPIasa de <i>T. cruzi</i>	85 <sup>a</sup> -97	2e-122	100

<sup>a</sup> Parámetros obtenidos únicamente en la base de datos de TriTrypDB

Con el propósito de identificar si existen sitios o regiones conservadas entre las proteínas de interés y PPIasas cristalizadas de otros organismos, se realizó el MSA de las PPIasas de interés comparando sus secuencias en cuatro condiciones diferentes: *i*) Vs PPIasas cristalizadas de otras especies (Figura 5), *ii*) Vs PPIasas cristalizadas de parásitos (Figura 6), *iii*) TcCyP19 Vs TcCyP22 Vs TcCySEC (Figura 7) y *iv*) TcCyP19 Vs TcCyP22 (Figura 8). En un primer intento el MSA incluyó en su análisis a la TcCySEC (Figura 7), observando que aunque existen sitios conservados, así como sustituciones conservadas y semi-conservadas, la cantidad de sitios conservados es menor a si se realiza el alineamiento utilizando únicamente las secuencias de la TcCyP19 y TcCyP22 (Figura 8). La TcCySEC ya ha sido caracterizada y cristalizada demostrando que esta proteína tiene actividad como

PPIasa y MIP, debido a esta dualidad en su actividad el incluirla en un primer análisis del MSA generó variaciones en los resultados, por lo cual no se incluyó la secuencia de la TcCySEC en el MSA de las condiciones *i* y *ii*.

Los MSA de las condiciones *i* y *ii* muestran que las secuencias de las PPIasas de interés presentan un gran número de sitios conservados al compararlas con PPIasas de otras especies (Figura 5) y un número aún mayor de sitios conservados en el alineamiento con PPIasas de parásitos (Figura 6). El MSA de la condición *iii* (Figura 7) demuestra que las tres PPIasas objetivo poseen algunos sitios conservados entre las PPIasas objetivo, aunque al realizar el alineamiento en la condición *iv* (Figura 8) entre la TcCyP19 y TcCyP22 se observa que el número de sitios conservados aumenta, representando la mayor parte de ambas secuencias.



```

6L2B_1|PPIase|Leishmania      -----HHHHHHEPEVTAKVYFDVMIDSEPLGRITIGLFGKDAPLTTENFRQLC 48
1X07_1|PPIase|Trypanosoma    -----MPVVTDKVYFDITIGDEPVGRVVIIGLFGNDVPKTVENFKQLA 42
2PLU_1|PPIase|Cryptosporidium  ----MHHHHHSSGRENLYFQGNPVVYFDISIGQTPAGRITMELFADKVPITAENFRALC 56
307T_1|PPIase|Moniliophthora  -----SHMANVFFNISINDKPEGRIVFKLYDEAVPKTAKNFRCLA 40
2CMT_1|PPIase|SCHISTOSOMA     -----MRTKKQKRNLPRVFFDIRIGNGDAGRIVMELRSDIVPRTAENFRALC 47
5YB9_1|PPIase|Trichomonas    -----GSMLKRPKTFFDISIRGDKVGVKIVFELFNDIVPKTAENFRALC 43
TcCyP19                      -----MASYKPHHATVPTNPKVFFDVSIGGQSAGRVVVFELFADAVPKTAENFRALC 51
TcCyP22                      MAFSRTWFWAQKLPFYPIPNKPLVFFEISIGAQPAGRVEMELFKDAVPKTAENFRALC 60
4JCP_1|PPIase|Brugia         -----GPGSMSRPKVFFDITIGGSNAGRIVMELFADIVPKTAENFRCLC 44
2HQJ_1|PPIase|Leishmania     -----GPGSMTNPKVFFDISIDNKAAGRIVMELYADTVPKTAENFRALC 44
                                .:*:: *      *:: : * . .* *::*: *.

6L2B_1|PPIase|Leishmania      TGEHFG-----YKDSIFHRVIQNFMIQGGDFTNFDGTGGKSIYGEKFADENLNV--- 98
1X07_1|PPIase|Trypanosoma    SGENFG-----YKGSIFHRVIRNFMIQGGDFTNFDGTGGKSIYGRFDDENLKI--- 92
2PLU_1|PPIase|Cryptosporidium TGEKMGQSGKPLCYTGSFFHRIIPQFMIQGGDFTRGDGTGGESIYGSKFRDENFVY--- 113
307T_1|PPIase|Moniliophthora TGQHGFG-----YKDSIFHRVIPQFMLQGGDFTRHNGTGGKSIYGEKFADENFQV--- 90
2CMT_1|PPIase|SCHISTOSOMA     TGERGFG-----YHNCCFHRVIPQFMCQGGDFVKGDGTGGKSIYGRKFDENFQL--- 97
5YB9_1|PPIase|Trichomonas    TGEKGIKSGMPLSYKGTMFHRIIPQFMIQGGDFTRFNGTGGESIYGMKFDENFKV--- 100
TcCyP19                      TGEKNFG-----YAGSGFHRRIIPQFMCQGGDFTNHNGTGGRSIYGEKFADESFAKAG 104
TcCyP22                      TGEKGVRSKALCYKGSKFHRVIPQFMCQGGDFTNNGTGGESIYGMKFPDESFAGRAG 120
4JCP_1|PPIase|Brugia         TGERGMGRSGKLLHYKGSKFHRVIPNFMQLQGGDFTRNGTGGESIYGEKFPDENFQE--- 101
2HQJ_1|PPIase|Leishmania     TGEKGIKSGKPLHYKSSVHRVIPNFMQGGDFTRNGTGGESIYGTTFRDESFSKAG 104
                                :*!.. *      * . ***:* :** *****. :****.**** * **.:

6L2B_1|PPIase|Leishmania      KHFGVGLTSMANAGPNTNGSQFFICTAPTQWLDGKHVVFGQVLEGEVVKAMEAVGS-QT 163
1X07_1|PPIase|Trypanosoma    KHFGVGLTSMANAGPNTNGSQFFICTAPTQWLDGKHVVFGQVLEGEVVKAMEAVGS-QT 163
2PLU_1|PPIase|Cryptosporidium THDAPFLLSMANAGPNTNGSQFFITVPCPWLDPGKHVVFGVLEGEVVKKSIKCGS-QN 172
307T_1|PPIase|Moniliophthora KHTKPGLLSMANAGANTNGSQFFITVPTSWLDGKHVVFGVEIEGLDIVRKEVKGGS-AS 149
2CMT_1|PPIase|SCHISTOSOMA     RHEGGVLSMANSGPNTNGSQFFICTTKCDWLDGKHVVFGRVVDGQNVVKKMESVGS-KS 156
5YB9_1|PPIase|Trichomonas    KHDKPGLLSMANAGPNTNGSQFFITVETPWLDPGHHCVFGQVIEGMDIVKQIESCGT-ES 159
TcCyP19                      KHFGVGLTSMANAGPNTNGSQFFICTAPTQWLDGKHVVFGQVLEGEVVKAMEAVGS-QT 163
TcCyP22                      KHFGVGLTSMANAGPNTNGSQFFICTAATEWLDGKHVVFGQVLEGEVVKAMEAVGS-QT 163
4JCP_1|PPIase|Brugia         KHTGPGVLSMANAGPNTNGSQFFICTAKTEWLDGKHVVFGRVVDGQNVVKKMESVGS-QS 160
2HQJ_1|PPIase|Leishmania     RHTGLGCLSMANAGPNTNGSQFFICTAATPWLDPGKHVVFGRVVDGQNVVKKVRLGS-SS 163
                                *      :****:* :*****: *. ***** * **.: .* : : :

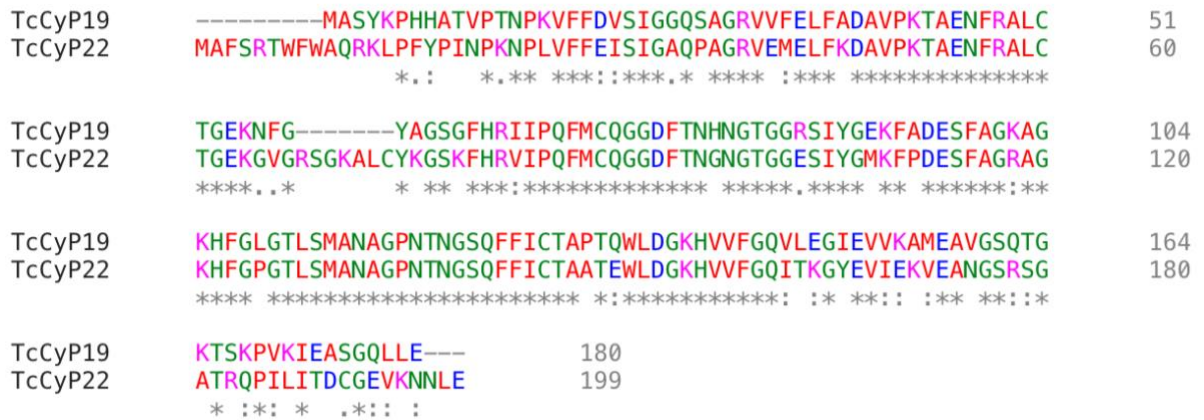
6L2B_1|PPIase|Leishmania      DRPVKPKIVASGEL----- 172
1X07_1|PPIase|Trypanosoma    DKPKKAVKINDCGVL----- 166
2PLU_1|PPIase|Cryptosporidium GKPTKSVKITASGV----- 186
307T_1|PPIase|Moniliophthora GKTNATIKITDCGTV----- 164
2CMT_1|PPIase|SCHISTOSOMA     GKVKEPVVIISRCGELI----- 172
5YB9_1|PPIase|Trichomonas    GRPRAMCMVTDCEGMK----- 175
TcCyP19                      GKTSKPKVIEASGQLLE---- 180
TcCyP22                      GATRQPIILITDCGEVKNLE 199
4JCP_1|PPIase|Brugia         GRTSADIVISDCGQL----- 175
2HQJ_1|PPIase|Leishmania     GKTRSRIVVSDCGEVAADKS 183
                                .      : .*

```

**Figura 6. Alineamiento de Secuencias de TCCyP19 y TCCyP22 Vs PPIasas cristalizadas de parásitos. Sitios o residuos conservados (\*), sustituciones conservadas (:), sustituciones semi-conservadas (.).**



**Figura 7. Alineamiento de Secuencias entre TCCyP19, TcCyP22 y TcCySEC.** Sitios o residuos conservados (\*), sustituciones conservadas (:), sustituciones semi-conservadas (.).



**Figura 8. Alineamiento de secuencias entre TCCyP19 y TcCyP22.** Sitios o residuos conservados (\*), sustituciones conservadas (:), sustituciones semi-conservadas (.).

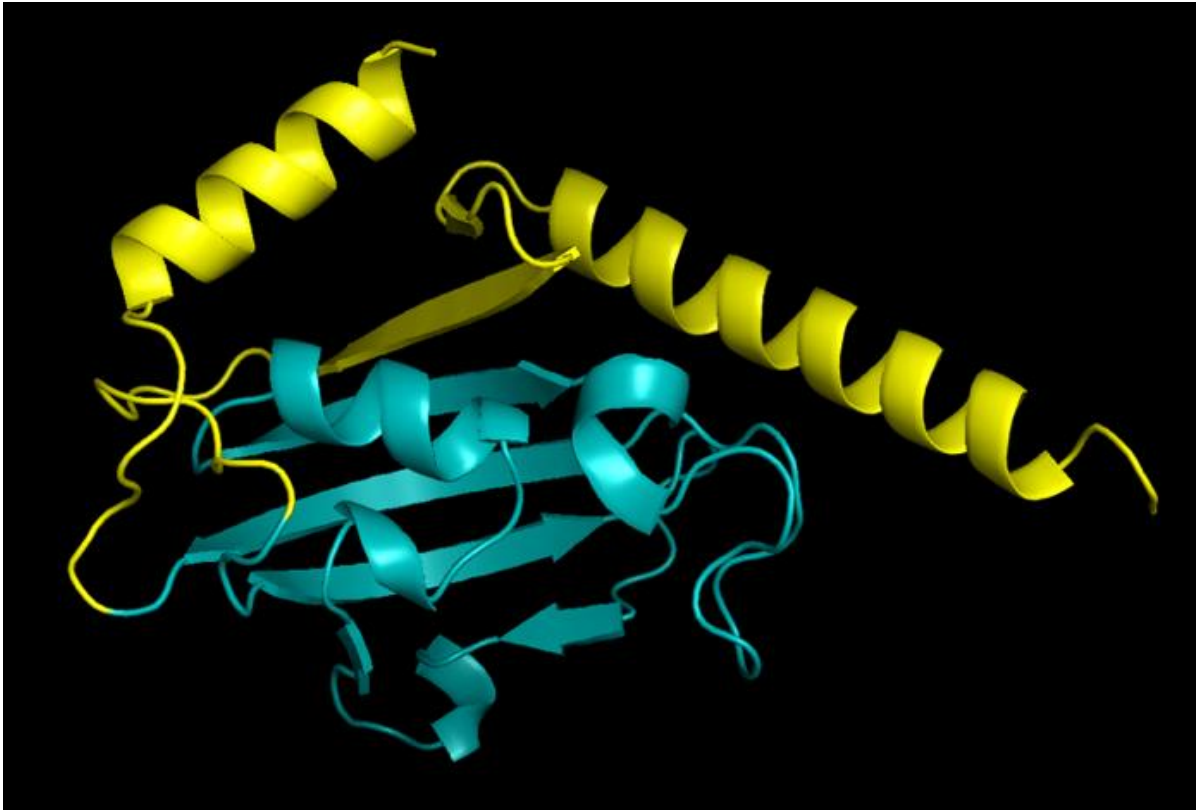
Los parámetros fisicoquímicos obtenidos en ProtParam se interpretaron de acuerdo con lo descrito en “Protein Identification and Analysis Tools on the ExPASy Server” del manual “The Proteomics Protocols Handbook” a partir de las secuencias de cada PPIasa, los datos fisicoquímicos relevantes se describen en la Tabla 3. Por ejemplo, el valor de GRAVY resulta negativo para TcCyP19, TcCyP22 y TcCySEC lo que representa un carácter hidrofílico de las tres PPIasas. Otro dato es el tiempo de vida media que para las tres PPIasas es mayor a 10 h en *E. coli*, la plataforma de expresión que se utilizó en este proyecto.

**Tabla 3. Parámetros Fisicoquímicos de TcCyP19, TcCyP22 y TcCySEC determinados en ProtParam.**

PPIasa	pI	Coefficiente de Extinción <sup>a</sup>	Abs 0.1 % (Cisteínas reducidas)	t½ (h) <sup>b</sup>	Índice de Inestabilidad <sup>c</sup>	Índice Alifático	GRAVY <sup>d</sup>
<b>TcCyP19</b>	7.76	9970	0.522	> 10	19.58	62.89	-0.184
<b>TcCyP22</b>	8.98	22460	1.036	> 10	20.74	59.85	-0.333
<b>TcCySEC / TcMIP</b>	5.83	18450	0.956	> 10	42.91	69.59	-0.709

<sup>a</sup> M-1 cm<sup>-1</sup>, a 280 nm medidos en agua; <sup>b</sup> en *Escherichia coli*, *in vivo*; <sup>c</sup> Sí <40.00 se considera una proteína estable, sí > 40.00 se considera una proteína inestable. <sup>d</sup> Gran promedio de hidropatía (kcal mol<sup>-1</sup>)

Debido a que la TcCySEC ya ha sido caracterizada y cristalizada no se realizó el modelado por homología de esta PPIasa, fundamentado en esto se utilizará el modelo 3D con el PDB ID 1JWV (Figura 9) para los posteriores análisis bioinformáticos.



**Figura 9. Modelo 3D de la proteína cristalizada TcMIP.** Correspondiente con el PDB ID 1JVW obtenido de la base de datos RCSB PDB, también denominada TcCySEC. La región de la ciclofilina en color verde identifica el dominio de PPIasa. En los extremos NH<sub>2</sub>- y COOH- se identifican de color amarillo dos alfa hélices que contienen parte del dominio MIP.

## Modelado, Optimización y Validación

De acuerdo con lo descrito en Materiales y Métodos en la sección 6.1 Análisis Bioinformático, se realizó el modelado por homología, optimización y validación a partir de las secuencias de TcCyP19 y TcCyP22, utilizando proteínas homólogas de estructura conocida. El **Apéndice 6** describe brevemente el fundamento teórico, variables consideradas y la interpretación de cada análisis de validación utilizado.

Las Tablas 4, 5 y 6 indican los parámetros de validación de los modelos obtenidos para la **TcCyP19** en los servidores de Swiss-Model, trRosetta y I-TASSER, respectivamente. En cada tabla se remarca en negritas la fila del modelo con los parámetros de validación óptimos. Se observó que los parámetros obtenidos fueron diferentes en magnitud e interpretación entre los modelos sin optimizar, optimizados y entre servidores, por lo cual la homologación de estos parámetros permitió definir los modelos óptimos, observando que para trRosetta el modelo óptimo es aquel sin optimizar (ver Tabla 5).

**Tabla 4. Parámetros de validación del modelo 3D de la TcCyP19 obtenido del servidor Swiss-Model y sus modelos 3D optimizados.**

Modelo	ERRAT (%)	VERIFY 3D (%) Average Score (3D-1D score $\geq$ 0.2)	PROVE (%) (volume < 3 sd)	PROCHECK (%) <sup>a</sup> Core / Allow / Gener / Disall
Inicial	96.8153	100	3.9	84.3 / 14.9 / 0.7 / 0.0
M1	99.3671	100	3.1	88.1 / 11.9 / 0.0 / 0.0
M2	96.8354	100	2.9	87.3 / 12.7 / 0.0 / 0.0
M3	97.4684	100	2.8	86.6 / 13.4 / 0.0 / 0.0
<b>M4</b>	<b>98.7421</b>	<b>100</b>	<b>2.8</b>	<b>86.6 / 13.4 / 0.0 / 0.0</b>
M5	98.7421	100	2.9	86.6 / 13.4 / 0.0 / 0.0

Valores obtenidos de los análisis de validación ERRAT, VERIFY 3D, PROVE y PROCHECK (ver *Apéndice 6*) para el modelo inicial sin optimizar, así como los cinco modelos optimizados en el servidor 3DRefine (M1-M5). En **negritas** se remarca el modelo con los valores óptimos globales. <sup>a</sup> PROCHECK proporciona los valores correspondientes al porcentaje de residuos que de acuerdo con los Gráficos Ramachandran se encuentran en las regiones más favorecidas (Core), las adicionalmente permitidas (Allow), las generosamente permitidas (Gener) y las no permitidas (Disall).



**Tabla 5. Parámetros de validación del modelo 3D de la TcCyP19 obtenido del servidor trRosetta y sus modelos 3D optimizados.**

<b>Modelo</b>	<b>ERRAT (%)</b>	<b>VERIFY 3D (%)</b> Average Score (3D-1D score $\geq$ 0.2)	<b>PROVE (%)</b> (volume < 3 sd)	<b>PROCHECK (%)</b> <sup>a</sup> Core / Allow / Gener / Disall
<b>Inicial</b>	<b>90.303</b>	<b>100</b>	<b>3.7</b>	<b>82.8 / 15.9 / 0.7 / 0.7</b>
M1	91.2281	100	5.0	84.1 / 13.8 / 1.4 / 0.7
M2	89.4737	100	4.5	84.8 / 13.1 / 1.4 / 0.7
M3	87.7193	100	4.8	84.1 / 13.8 / 1.4 / 0.7
M4	88.3041	100	4.7	84.1 / 13.1 / 2.1 / 0.7
M5	87.7193	100	4.7	84.1 / 13.1 / 2.1 / 0.7

Valores obtenidos de los análisis de validación ERRAT, VERIFY 3D, PROVE y PROCHECK (ver *Apéndice 6*) para el modelo inicial sin optimizar, así como los cinco modelos optimizados en el servidor 3DRefine (M1-M5). En **negritas** se remarca el modelo con los valores óptimos globales. <sup>a</sup> PROCHECK proporciona los valores correspondientes al porcentaje de residuos que de acuerdo con los Gráficos Ramachandran se encuentran en las regiones más favorecidas (Core), las adicionalmente permitidas (Allow), las generosamente permitidas (Gener) y las no permitidas (Disall).

**Tabla 6. Parámetros de validación del modelo 3D de la TcCyP19 obtenido del servidor I-TASSER y sus modelos 3D optimizados.**

<b>Modelo</b>	<b>ERRAT (%)</b>	<b>VERIFY 3D (%)</b> Average Score (3D-1D score $\geq$ 0.2)	<b>PROVE (%)</b> (volume < 3 sd)	<b>PROCHECK (%)</b> <sup>a</sup> Core / Allow / Gener / Disall
Inicial	83.1395	100	4.0	71.7 / 25.5 / 0.7 / 2.1
M1	92.4419	100	4.3	79.3 / 15.9 / 3.4 / 1.4
M2	94.7674	100	4.3	80.7 / 14.5 / 4.1 / 0.7
M3	95.9302	100	4.9	82.1 / 13.8 / 3.4 / 0.7
<b>M4</b>	<b>97.0930</b>	<b>100</b>	<b>4.4</b>	<b>83.4 / 12.4 / 3.4 / 0.7</b>
M5	96.5116	100	4.4	83.4 / 11.7 / 4.1 / 0.7

Valores obtenidos de los análisis de validación ERRAT, VERIFY 3D, PROVE y PROCHECK (ver *Apéndice 6*) para el modelo inicial sin optimizar, así como los cinco modelos optimizados en el servidor 3DRefine (M1-M5). En **negritas** se remarca el modelo con los valores óptimos globales. <sup>a</sup> PROCHECK proporciona los valores correspondientes al porcentaje de residuos que de acuerdo con los Gráficos Ramachandran se encuentran en las regiones más favorecidas (Core), las adicionalmente permitidas (Allow), las generosamente permitidas (Gener) y las no permitidas (Disall).

Respecto a la **TcCyP22**, las Tablas 7, 8 y 9 indican los parámetros de validación de esta proteína para los servidores de Swiss-Model, trRosetta y I-TASSER, respectivamente. En cada tabla se remarca en negritas la fila correspondiente al modelo con los parámetros de validación óptimos. Del mismo modo que con la TcCyP19, el **Apéndice 6** describe brevemente el fundamento teórico, variables consideradas y la interpretación de cada análisis de validación utilizado para el análisis bioinformático de esta proteína.

Se observó que los parámetros obtenidos fueron diferentes en magnitud e interpretación entre los modelos sin optimizar, optimizados y entre servidores, por lo cual la homologación de estos parámetros permitió definir los modelos óptimos, observando que para la TcCyP22 los modelos óptimos de los tres servidores se obtuvieron posterior a la optimización.

**Tabla 7. Parámetros de validación del modelo 3D de la TcCyP22 obtenido del servidor Swiss-Model y sus modelos 3D optimizados.**

<b>Modelo</b>	<b>ERRAT (%)</b>	<b>VERIFY 3D (%)</b> Average Score (3D-1D score $\geq$ 0.2)	<b>PROVE (%)</b> (volume < 3 sd)	<b>PROCHECK (%)</b> <sup>a</sup> Core / Allow / Gener / Disall
Inicial	92.5926	99.43	3.8	87.9 / 12.1 / 0.0 / 0.0
M1	98.2143	99.43	3.7	88.6 / 11.4 / 0.0 / 0.0
M2	97.6190	99.43	3.1	87.9 / 12.1 / 0.0 / 0.0
M3	97.6190	99.43	2.8	89.3 / 10.7 / 0.0 / 0.0
<b>M4</b>	<b>97.0238</b>	<b>99.43</b>	<b>2.6</b>	<b>90.0 / 10.0 / 0.0 / 0.0</b>
M5	97.6190	99.43	2.9	90.7 / 9.3 / 0.0 / 0.0

Valores obtenidos de los análisis de validación ERRAT, VERIFY 3D, PROVE y PROCHECK (ver *Apéndice 6*) para el modelo inicial sin optimizar, así como los cinco modelos optimizados en el servidor 3DRefine (M1-M5). En **negritas** se remarca el modelo con los valores óptimos globales. <sup>a</sup> PROCHECK proporciona los valores correspondientes al porcentaje de residuos que de acuerdo con los Gráficos Ramachandran se encuentran en las regiones más favorecidas (Core), las adicionalmente permitidas (Allow), las generosamente permitidas (Gener) y las no permitidas (Disall).

**Tabla 8. Parámetros de validación del modelo 3D de la TcCyP22 obtenido del servidor trRosetta y sus modelos 3D optimizados.**

<b>Modelo</b>	<b>ERRAT (%)</b>	<b>VERIFY 3D (%)</b> Average Score (3D-1D score $\geq$ 0.2)	<b>PROVE (%)</b> (volume < 3 sd)	<b>PROCHECK (%)</b> <sup>a</sup> Core / Allow / Gener / Disall
Inicial	93.1579	90.45	3.4	86.9 / 13.1 / 0.0 / 0.0
M1	90.0524	91.46	3.5	86.2 / 13.1 / 0.6 / 0.0
M2	87.9581	91.46	3.8	86.2 / 13.1 / 0.6 / 0.0
M3	87.4346	91.46	3.8	86.2 / 13.8 / 0.0 / 0.0
M4	86.9110	91.46	2.9	85.0 / 15.0 / 0.0 / 0.0
<b>M5</b>	<b>86.9110</b>	<b>92.46</b>	<b>2.8</b>	<b>85.0 / 15.0 / 0.0 / 0.0</b>

Valores obtenidos de los análisis de validación ERRAT, VERIFY 3D, PROVE y PROCHECK (ver *Apéndice 6*) para el modelo inicial sin optimizar, así como los cinco modelos optimizados en el servidor 3DRefine (M1-M5). En **negritas** se remarca el modelo con los valores óptimos globales. <sup>a</sup> PROCHECK proporciona los valores correspondientes al porcentaje de residuos que de acuerdo con los Gráficos Ramachandran se encuentran en las regiones más favorecidas (Core), las adicionalmente permitidas (Allow), las generosamente permitidas (Gener) y las no permitidas (Disall).

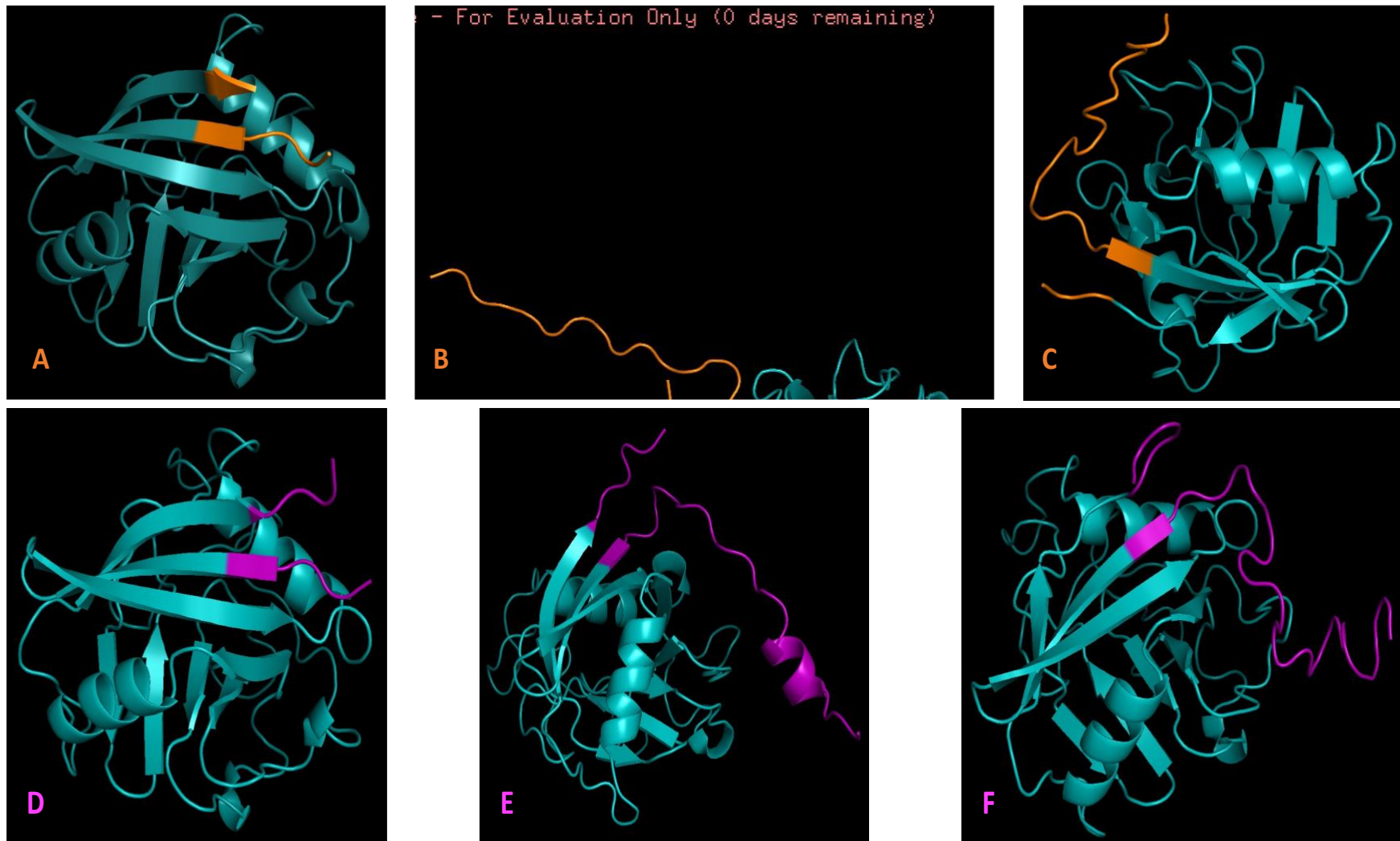
**Tabla 9. Parámetros de validación del modelo 3D de la TcCyP22 obtenido del servidor I-TASSER y sus modelos 3D optimizados.**

<b>Modelo</b>	<b>ERRAT (%)</b>	<b>VERIFY 3D (%)</b> Average Score (3D-1D score $\geq$ 0.2)	<b>PROVE (%)</b> (volume < 3 sd)	<b>PROCHECK (%)</b> <sup>a</sup> Core / Allow / Gener / Disall
Inicial	84.2932	90.95	3.4	76.9 / 18.8 / 3.8 / 0.6
M1	91.0995	90.95	5.2	81.9 / 15.0 / 1.9 / 1.2
<b>M2</b>	<b>90.5759</b>	<b>93.47</b>	<b>4.9</b>	<b>82.5 / 14.4 / 1.2 / 1.9</b>
M3	91.0995	94.47	5.1	81.9 / 15.0 / 1.2 / 1.9
M4	89.0052	94.47	5.5	82.5 / 13.8 / 1.9 / 1.9
M5	88.4817	93.47	5.6	83.1 / 13.1 / 1.9 / 1.9

Valores obtenidos de los análisis de validación ERRAT, VERIFY 3D, PROVE y PROCHECK (ver *Apéndice 6*) para el modelo inicial sin optimizar, así como los cinco modelos optimizados en el servidor 3DRefine (M1-M5). En **negritas** se remarca el modelo con los valores óptimos globales. <sup>a</sup> PROCHECK proporciona los valores correspondientes al porcentaje de residuos que de acuerdo con los Gráficos Ramachandran se encuentran en las regiones más favorecidas (Core), las adicionalmente permitidas (Allow), las generosamente permitidas (Gener) y las no permitidas (Disall).

## **Modelos Óptimos**

Las estructuras óptimas de TcCyP19 son **M4**, **Inicial** y **M4** (resaltadas en negritas en las Tablas 4, 5 y 6, respectivamente), mientras las correspondientes de TcCyP22 son **M4**, **M5** y **M2** (resaltadas en negritas en las Tablas 7, 8 y 9, respectivamente) obtenidas de los servidores de modelado por homología Swiss-Model, trRosetta y I-TASSER, se muestran en la Figura 10 indicando en color verde el dominio de PPIasa. Debido a la variación entre los resultados obtenidos la etapa de modelado, optimización y validación se utilizaron todos los modelos 3D óptimos para la predicción bioinformática de solubilidad y agregación corregida estructuralmente. Lo anterior con el propósito de validar la consistencia o variación en los resultados obtenidos *in silico* versus los datos que se obtuvieron posteriormente en la etapa experimental.



**Figura 10. Modelo 3D óptimos de la TcCyP19 y TcCyP22.** TcCyP19 se identifica para los modelos **A, B, C**; TcCyP22 se identifica para los modelos **D, E, F**. Estructuras obtenidas en los servidores Swiss-Model (**A, D**), trRosetta (**B, E**) e I-TASSER (**C, F**). El dominio de PPIasa se indica en color verde para todos los modelos. Los extremos NH<sub>2</sub>- y COOH- que no forman parte del dominio de PPIasa e identificados de un color distinto al verde en color naranja para TcCyP19 y en color magenta TcCyP22, y no corresponden a algún péptido señal como en el caso de la TcCySEC/TcMIP.

## Determinación *in silico* de Solubilidad y Agregación

Las secuencias de las tres proteínas de interés, así como su correspondiente dominio PPIasa y los modelos 3D óptimos se usaron para la predicción *in silico* de solubilidad y agregación con el propósito de sugerir si la expresión de TcCyP19, TcCyP22 y TcCySEC es soluble o no en *E. coli*. El **Apéndice 7** describe brevemente el fundamento teórico, variables consideradas y la escala de ponderación de cada servidor utilizado en esta etapa del análisis bioinformático. Debido a la variación en los métodos de predicción y los valores obtenidos en los diferentes servidores, se analizaron todos los datos con el propósito de validar que los resultados obtenidos son consistentes y de esta manera confirmar que las proteínas podrían ser solubles. Los valores obtenidos respecto a la secuencia completa de cada proteína y su dominio de PPIasa (Tabla 10), así como los valores obtenidos de los modelos 3D óptimos (Tabla 11) sugieren que las tres proteínas se expresan en forma soluble en *E. coli*, lo anterior fue consistente en todos los servidores de predicción utilizados. Estos resultados fueron comparados con los obtenidos en la etapa experimental al expresar cada una de las proteínas recombinantes.

**Tabla 10. Solubilidad y Agregación de TcCyP19, TcCyP22 y TcCySEC determinada a partir de su secuencia de aminoácidos.**

Proteína	Protein-Sol		SoDoPe		CamSol		Aggrescan	
	Secuencia	PPIasa	Secuencia	PPIasa	Secuencia	PPIasa	Secuencia	PPIasa
<b>TcCyP19</b>	0.471	0.549	0.6476	0.6026	0.563133	0.474322	-0.038	-0.041
<b>TcCyP22</b>	0.473	0.475	0.5739	0.6269	0.676944	0.863757	-0.054	-0.086
<b>TcCySEC</b>	0.585	0.563	0.8115	0.5589	2.514334	1.753254	-0.207	-0.076

Predicción de solubilidad estimada a partir de la secuencia completa de la proteína y utilizando únicamente la secuencia que corresponde al dominio de PPIasa. Utilizando los servidores de Protein-Sol (Soluble >0.45), SoDoPe (1 [soluble] - 0 [insoluble]), CamSol (Soluble > 0.0). Predicción de agregación estimada con el servidor Aggrescan (Agregación > 0).

**Tabla 11. Solubilidad y Agregación de TcCyP19, TcCyP22 y TcCySEC determinada a partir de su modelo 3D**

Modelo 3D	Solubilidad <sup>a</sup>			Agregación <sup>b</sup>		
	TcCyP19	TcCyP22	TcCySEC <sup>c</sup>	TcCyP19	TcCyP22	TcCySEC <sup>c</sup>
<i>Swiss-Model</i>	0.547	0.846	2.363	-0.841	-0.865	-1.270
<i>trRosetta</i>	0.563	0.677	2.363	-0.831	-0.777	-1.270
<i>I-TASSER</i>	0.563	0.677	2.363	-0.789	-0.719	-1.270

<sup>a</sup> Predicción de solubilidad estimada con CamSol (Soluble > 0.0). <sup>b</sup> Predicción de agregación estimada con Aggrescan3D (Agregación > 0). En la TcCySEC se utilizó el modelo cristalizado PDB ID 1JVW.

## 7.2. Síntesis de Genes

Las secuencias de los genes que codifican para cada una de las proteínas de interés se diseñaron para su síntesis por el Dr. Gerardo Reséndiz Cardiel, como se indica en la sección 6.2 de Materiales y Métodos añadiendo en los extremos 5' y 3' de las secuencias nucleotídicas los sitios de restricción para NcoI (CCATG) y XhoI (CTCAGG). Por lo anterior, las proteínas de interés resultantes contienen insertados aminoácidos en el extremo NH<sub>2</sub>-, una Alanina (A) en la posición 2 de las proteínas TcCyP19 y TcCyP22, mientras que en la posición 28 de la TcCySEC se insertó una Metionina (M) considerando que la posición 28 es la posición 1 de la rTcCySEC ya que no contiene al péptido señal. Adicionalmente, en el extremo COOH- se insertó una Leucina (L) y un Ácido glutámico (E). Las secuencias finales que Synbio Technologies LLC sintetizó, así como la secuencia de aminoácidos correspondiente se muestran en el **Apéndice 2**.

## 7.3. Subclonación de ORF

A partir de las construcciones pUC57-*ppiasa* con codones optimizados para su expresión en *E. coli* proporcionadas por el Dr. Gerardo Reséndiz Cardiel (ver **Apéndice 8**), se realizó la extracción y purificación del ADNp con el propósito de obtener muestras de ADNp de mayor volumen y concentración a las proporcionadas por Synbio Technologies LLC.

A partir de estos resultados se realizó la cuantificación de ADNp por absorbancia a 280 nm, correspondientes a las construcciones pUC57-*ppiasa*, provenientes de las Clonas 1 y 2 (Tabla 12) proporcionadas por el Dr. Gerardo Reséndiz Cardiel, así como de las muestras proporcionadas por Synbio Technologies LLC las cuales fueron menores a 20 ng/ $\mu$ l.

**Tabla 12. Concentración de ADNp de las construcciones pUC57-*ppiasa***

	Clona 1 (ng/ $\mu$ l)	Clona 2 (ng/ $\mu$ l)
pUC57- <i>tccyp19</i>	310.3	276.35
pUC57- <i>tccyp22</i>	34.5	57.95
pUC57- <i>tccysec</i>	113.55	69.0

Se evaluó la calidad del ADNp por electroforesis en gel de agarosa al 1% utilizando 1  $\mu$ L de cada muestra correspondiente a la construcción pUC57-*ppiasa coli*, provenientes de la Clona 1, Clona 2, así como de las muestras de Synbio Technologies LLC (figura no incluida). Los resultados muestran una adecuada calidad de las muestras provenientes de las Clonas 1 y 2, sin embargo, las bandas correspondientes a las muestras de Synbio Technologies LLC no se observan claramente ya que de acuerdo con los valores obtenidos en NanoDrop® su concentración es muy baja, por lo anterior se excluyeron estas muestras en la siguiente etapa. El análisis de la Doble Digestión del ADNp de las construcciones pUC57-*ppiasa coli* con las enzimas NcoI y XhoI en un gel de agarosa al 2% muestra la liberación de los fragmentos de 542 pb, 599 pb y 515 pb correspondientes a los insertos de los genes *tccyp19*, *tccy22* y *tccysec*, respectivamente. Basado en estos resultados se seleccionaron las muestras de correspondientes a la Clona 2 para las construcciones pUC57-*tccyp19* y pUC57-*tccyp22*, mientras que la para construcción pUC57-*tccysec* se seleccionó la Clona 1. A partir de estas tres muestras seleccionadas, así como del vector de expresión pCri8a se realizó la transformación en células DH5 $\alpha$  de *E. coli* y se realizó el



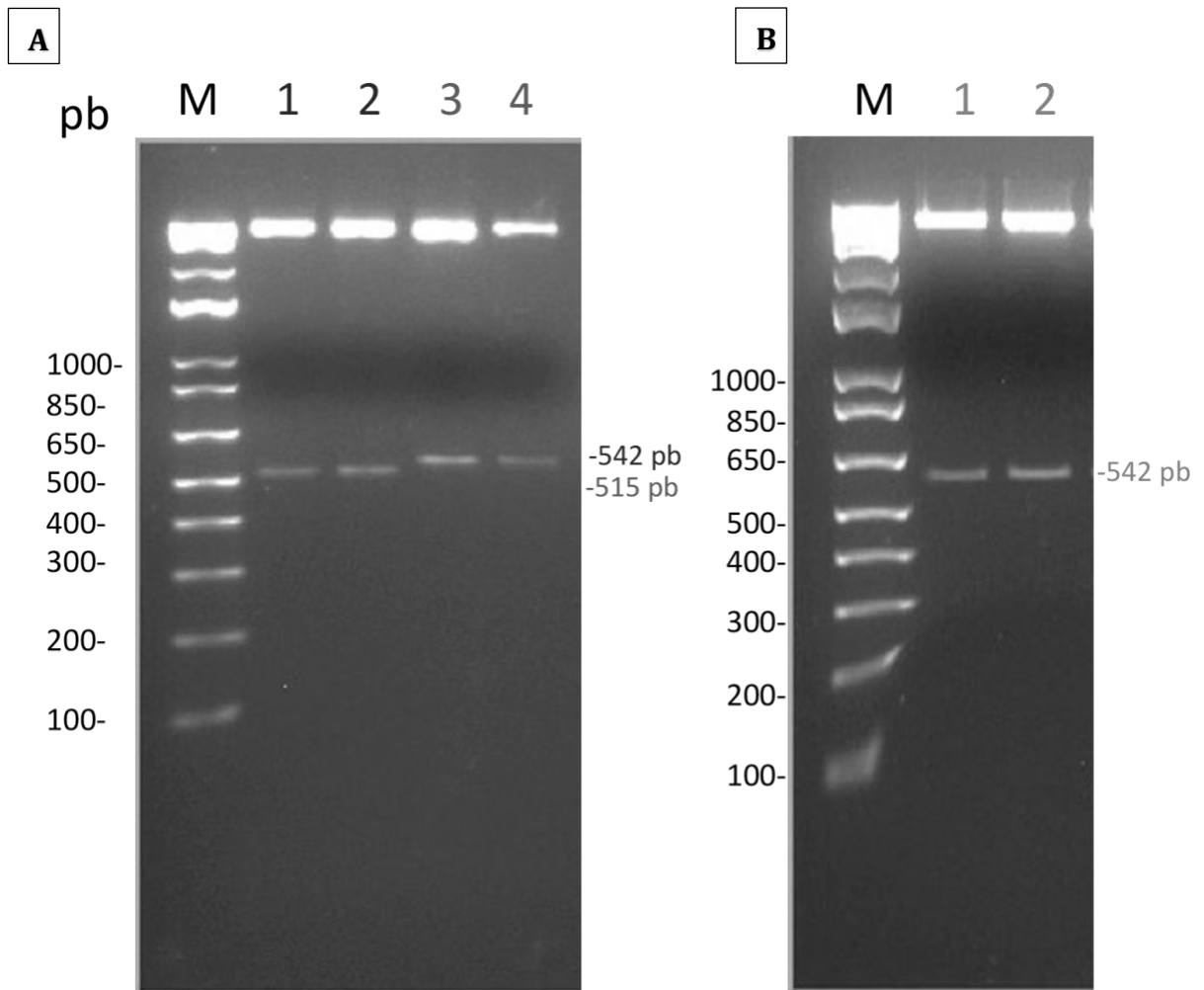
midiprep, eluyendo el ADNp purificado en 550  $\mu$ L de H<sub>2</sub>O grado inyectable. Se analizó la calidad del ADNp por electroforesis en gel de agarosa al 1% donde se observan todas las muestras en el tamaño y calidad esperados. Adicionalmente, a partir del ADNp obtenido de cada midiprep se cuantificó su concentración por absorbancia a 280 nm obteniendo los valores descritos en la Tabla 13, concentraciones de ADNp adecuadas por lo cual se procedió con la siguiente etapa del proceso de subclonación.

De acuerdo con los volúmenes referidos en el **Apéndice 4** se llevó a cabo la reacción de doble digestión utilizando las enzimas de restricción NcoI y XhoI. Posteriormente se analizó la separación de las muestras en geles de agarosa al 2% y se realizó la extracción de ADN en gel con centrifuga, de acuerdo con lo descrito en Materiales y Métodos en la sección 6.3; se analizó la calidad del ADN por electroforesis del inserto y vector eluidas de la extracción en gel confirmando que todas las muestras se encontraban en el tamaño y con la calidad esperados. Adicionalmente, se cuantificó la concentración de ADN extraído por medición de absorbancia a 280 nm obteniendo los valores descritos en la Tabla 13, concentraciones de ADN adecuadas (Inserto / Vector linealizado) por lo cual se procedió con la etapa de ligación.

**Tabla 13. Concentración de ADNp, inserto o vector linealizado de clonas seleccionadas.**

Construcción	ADNp (ng/ $\mu$ l)	Inserto / Vector linealizado (ng/ $\mu$ l)
pUC57-tccyp19	505.95	8.85
pUC57-tccyp22	922.5	15.95
pUC57-tccysec	650.1	11.40
pCri8a	102.2	8.85

De acuerdo con lo descrito en el **Apéndice 6** se realizaron las diferentes reacciones de Ligación con una relación molar 3:1 (inserto/vector) de acuerdo con lo descrito en Materiales y Métodos en la sección 6.3. Posteriormente, utilizando cada producto de ligación (Problema, Control de Religación y Control de ADNp) se procedió a la Transformación de células DH5 $\alpha$  de *E. coli* mediante procedimientos estándar 3. El control positivo de transformación utilizó 1  $\mu$ l de ADNp purificado correspondiente a la construcción Chagasina-pCri8a. Se obtuvo crecimiento de aproximadamente 20 Unidades Formadoras de Colonias (UFC) con la construcción pCri8a-tccyp19, 10 UFC con pCri8a-tccyp22 y 4 UFC con pCri8a-tccysec al concluir 17 h de incubación. Finalmente, se analizaron las clonas de células DH5 $\alpha$  que teóricamente contienen la construcción pCri8a-*ppiasa*, el número de clonas analizadas dependió del total de UFC que se obtuvieron en placa. De acuerdo a estos resultados se analizaron 10 clonas para la construcción pCri8a-tccyp19, 6 clonas para pCri8a-tccyp22 y 4 clonas para pCri8a-tccysec. La Figura 11 presenta los geles de agarosa al 2% observando la doble digestión de las construcciones pCri8a-*ppiasa* obtenidas de dos clonas, observando la liberación de los fragmentos de acuerdo con el tamaño esperado de cada gen que codifica para las proteínas de interés, por lo cual se confirmó que las clonas son positivas para la construcción objetivo.

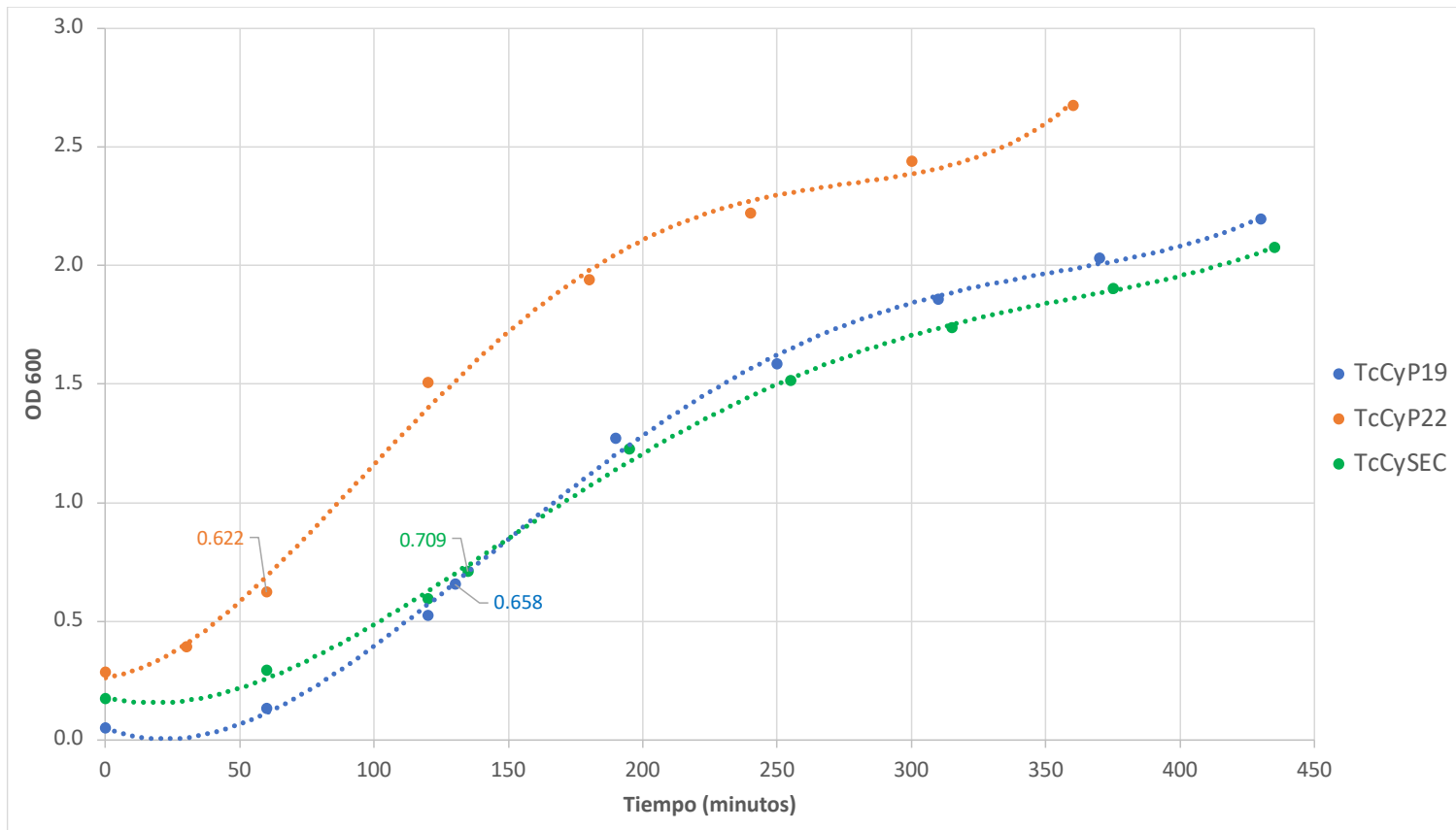


**Figura 11. Análisis del ADN de las clonas positivas a las construcciones pCri8a-tccyp19, pCri8a-tccyp22 y pCri8a-tccysec.** Análisis en geles de agarosa al 2% de la doble digestión utilizando las enzimas de restricción NcoI y XhoI. **Panel A:** pCri8a-tccyp19 (carriles 1-2), pCri8a-tccyp22 (carriles 3-4). **Panel B:** pCri8a-tccysec (carriles 1-2). M: Marcador de peso molecular.

#### **7.4. Análisis de la expresión y solubilidad de las proteínas recombinantes**

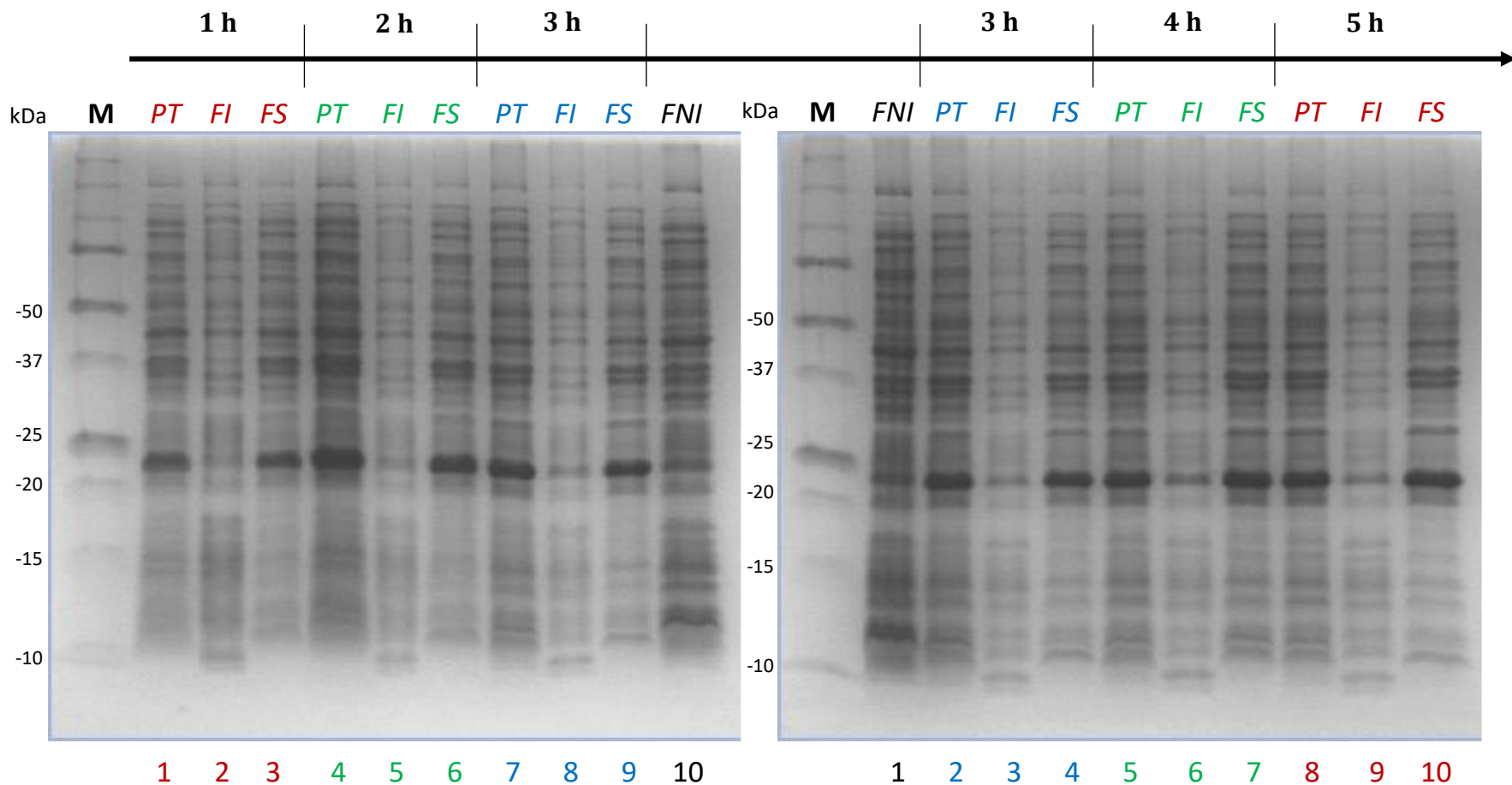
El análisis de las muestras obtenidas en la prueba de inducción sugirió que bajo las condiciones de crecimiento e inducción establecidas en Materiales y Métodos en la sección 6.4, las tres proteínas recombinantes correspondientes a TcCyP19, TcCyP22 y TcCySEC se expresaron en las células *E. coli* BL21(DE3) en el tamaño esperado para cada proteína recombinante de 22 kDa, 25 kDa y 22 kDa, respectivamente; considerando las etiquetas de fusión contenidas en el vector de expresión pCri8a, que suman a cada proteína recombinante un peso molecular aproximado de 3 kDa (véase **Apéndice 3**). A partir de estos resultados se procedió a la caracterización de la expresión y solubilidad de las proteínas rTcCyP19, rTcCyP22 y rTcCySEC en células *E. coli* BL21 (DE3).

Los valores de OD<sub>600</sub> correspondientes a las muestras colectadas desde la adición del inóculo, adición de IPTG 1 mM y durante cada hora hasta las 5 h de inducción, se graficaron y permiten observar en la Figura 12 que la inducción con IPTG 1 mM se realizó aproximadamente a la mitad de la fase exponencial de crecimiento, dentro del rango establecido de 0.6 y 07 AU, lo cual permite que las células continúen creciendo por aproximadamente 2h posteriores a la adición del agente inductor. Después de estas 2 h, se observa el inicio de una fase de crecimiento más lenta ya que las células concentran sus recursos en la producción de la proteína recombinante en cuestión. Este comportamiento se observa en todos los cultivos con las células *E. coli* BL21 (DE3) que contienen la construcción pCri8a-ppiasea y que expresan respectivamente la rTcCyP19, rTcCyP22 y rTcCySEC.

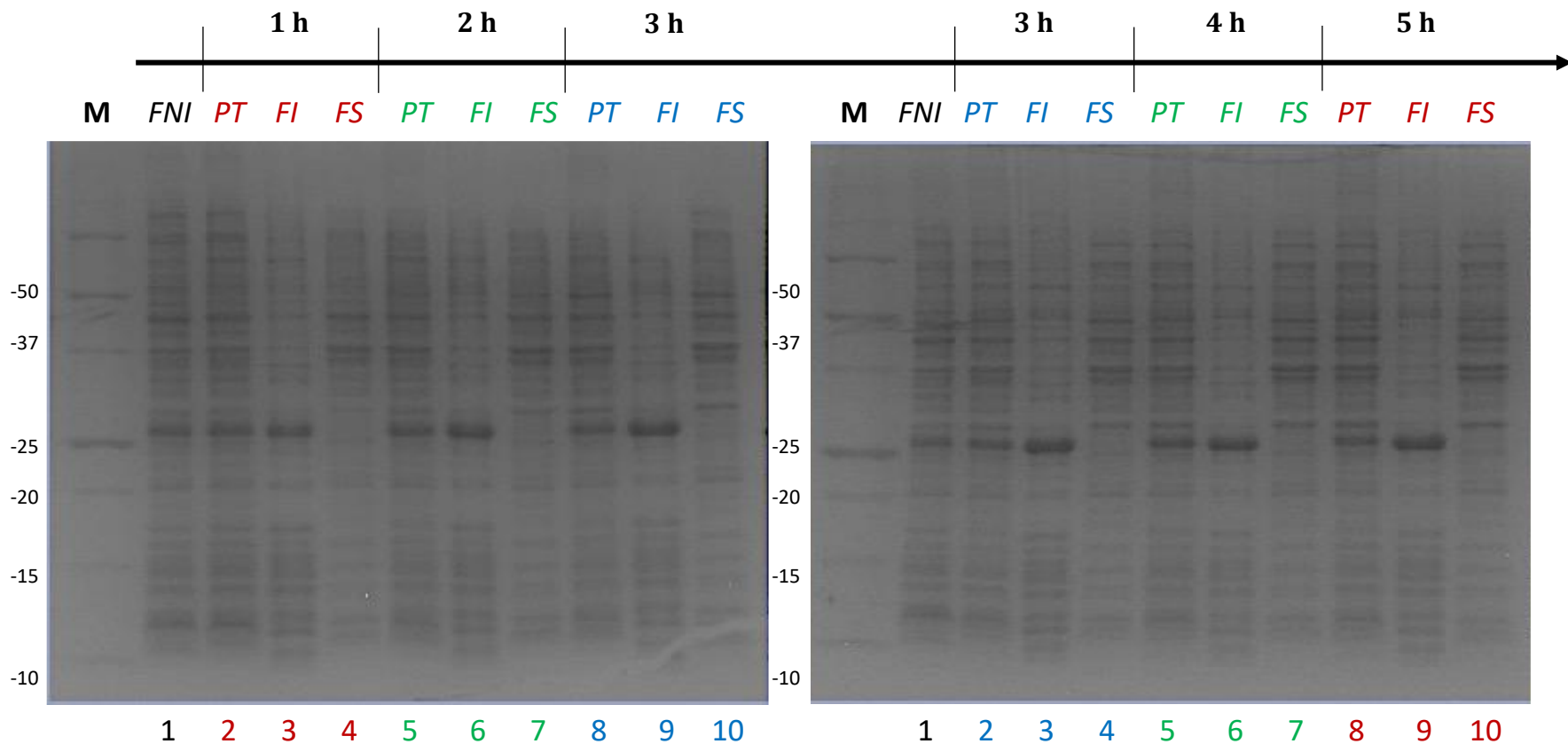


**Figura 12. Curva de crecimiento de células BL21 (DE3) que expresan rTcCyP19, rTcCyP22 y rTcCySEC.** Indicando el valor de la OD<sub>600</sub> al momento de adicionar IPTG 1 mM (agente inductor) al inicio de la fase exponencial de crecimiento de las células BL21 (DE3), observando un crecimiento similar durante aproximadamente 2 h (120 min) posteriores a la adición del agente inductor y cambiando a una fase de crecimiento más lenta después de 2 h (120 min) de inducción.

Basado en los resultados anteriores, se procedió al análisis por SDS-PAGE de las muestras obtenidas para la FNI, así como las muestras de las Fracciones Inducidas a los diferentes tiempos, y bajo las condiciones de crecimiento e inducción descritas en materiales y métodos en la sección 6.4. Este análisis mostró que la expresión de las proteínas recombinantes se presentó desde la primera hora de inducción, incluso cuando aún existe crecimiento de las células *E. coli* BL21 (DE3); permitiendo sugerir que la expresión de las proteínas a las 4 o 5 h de inducción será adecuada para procesos de expresión a mayor escala. Adicionalmente, se muestra que bajo las condiciones de crecimiento e inducción establecidas, las proteínas rTcCyP19 (Figura 13) y rTcCySEC (Figura 15) se encuentran predominantemente en la FS a diferencia de rTcCyP22 (Figura 14) que se encuentra casi en su totalidad en la FI. Estos resultados permiten proceder con el escalamiento de la producción de dichas proteínas recombinantes.

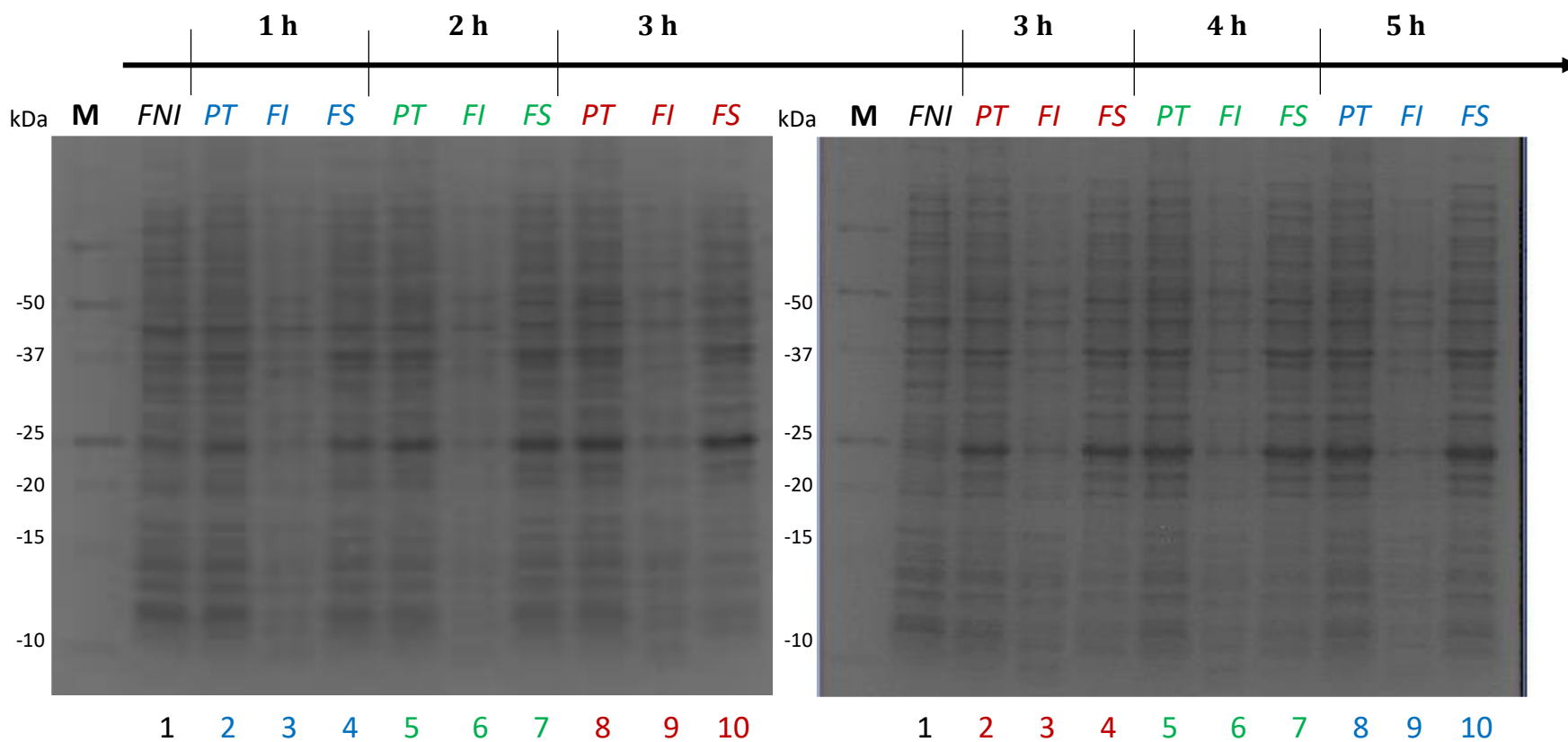


**Figura 13. Análisis por SDS-PAGE de la expresión de la proteína rTcCyP19.** Análisis realizado a partir de las muestras obtenidas del cultivo bacteriano, tomadas cada hora y durante las 5 h posteriores a la inducción con IPTG 1 mM. En la parte superior de la imagen se indica la hora de crecimiento a la cual se tomaron las muestras correspondientes a la PT, FI y FS, las cuales se trataron y posteriormente se resolvieron por electroforesis en un gel de poliacrilamida al 15% de acuerdo a lo indicado en Materiales y Métodos en la secciones 6.4.1 y 6.4.2, respectivamente. FNI: Fracción No Inducida; PT: Proteína Total; FS: Fracción Soluble; FI: Fracción Insoluble; M: Marcador de peso molecular.



**Figura 14. Análisis por SDS-PAGE de la expresión de la proteína rTcCyP22.** Análisis realizado a partir de las muestras obtenidas del cultivo bacteriano, tomadas cada hora y durante las 5 h posteriores a la inducción con IPTG 1 mM. En la parte superior de la imagen se indica la hora de crecimiento a la cual se tomaron las muestras correspondientes a la PT, FI y FS, las cuales se trataron y posteriormente se resolvieron por electroforesis en un gel de poliacrilamida al 15% de acuerdo a lo indicado en Materiales y Métodos en la secciones 6.4.1 y 6.4.2, respectivamente. FNI: Fracción No Inducida; PT: Proteína Total; FS: Fracción Soluble; FI: Fracción Insoluble; M: Marcador de peso molecular.

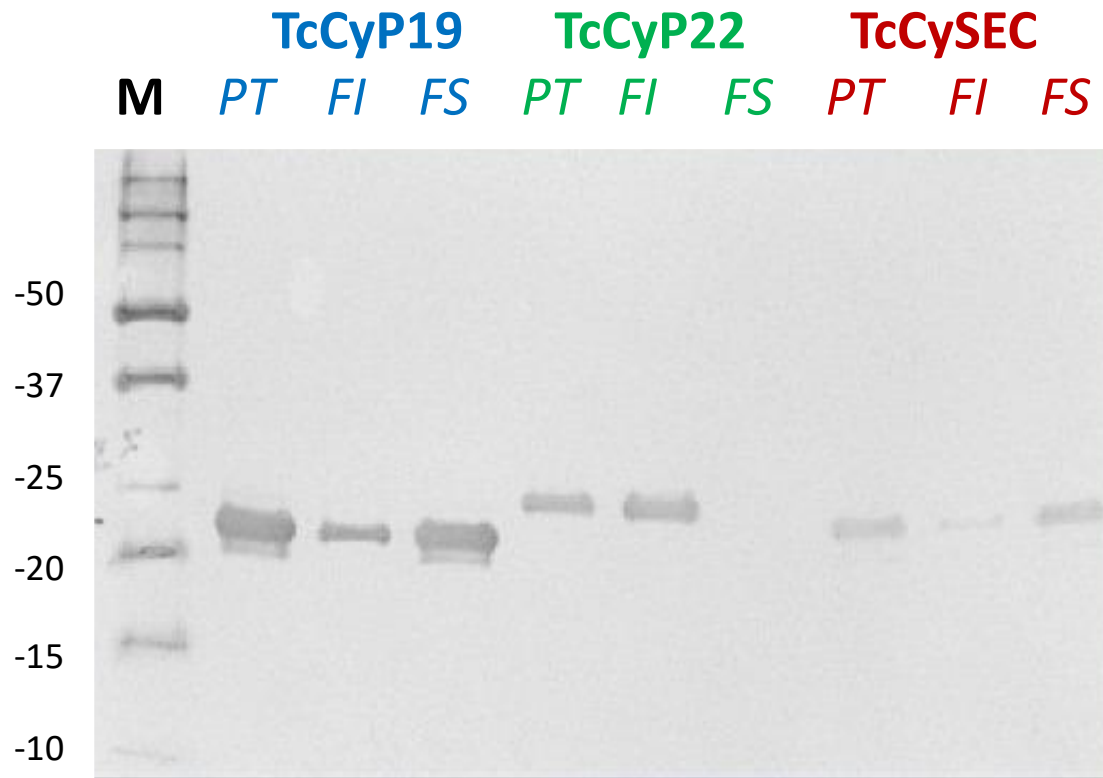




**Figura 15. Análisis por SDS-PAGE de la expresión de la proteína rTcCySEC.** Análisis realizado a partir de las muestras obtenidas del cultivo bacteriano, tomadas cada hora y durante las 5 h posteriores a la inducción con IPTG 1 mM. En la parte superior de la imagen se indica la hora de crecimiento a la cual se tomaron las muestras correspondientes a la PT, FI y FS, las cuales se trataron y posteriormente se resolvieron por electroforesis en un gel de poliacrilamida al 15% de acuerdo a lo indicado en Materiales y Métodos en la secciones 6.4.1 y 6.4.2, respectivamente. FNI: Fracción No Inducida; PT: Proteína Total; FS: Fracción Soluble; FI: Fracción Insoluble; M: Marcador de peso molecular.

#### 7.4.1. Inmunodetección de Proteínas por Ensayo de Western Blot

Considerando que la expresión de las proteínas rTcCyP19, rTcCyP22 y rTcCySEC a las cuatro (4) y cinco (5) h de inducción es similar de acuerdo con el análisis por SDS-PAGE (Figuras 13, 14 y 15) se realizó la inmunodetección en membrana de nitrocelulosa con las muestras obtenidas a las 4 h de inducción y utilizando anticuerpos “Mouse  $\alpha$ -His<sub>6</sub>”, debido a que las proteínas recombinantes poseen dos etiquetas de seis Histidinas en los extremos NH<sub>2</sub>- y COOH- de acuerdo con las etiquetas de fusión proporcionadas por el vector de expresión pCri8a (véase **Apéndice 3**). La Figura 16 muestra la membrana de nitrocelulosa posterior al revelado, indicando las bandas correspondientes a las proteínas rTcCyP19, rTcCyP22 y rTcCySEC que poseen ambas etiquetas His<sub>6</sub>, confirmando además que las tres proteínas se expresan en el tamaño esperado. Adicionalmente, estos resultados permiten confirmar que en el caso de rTcCyP19 y rTcCySEC se encuentran predominantemente en la fracción soluble, existiendo cierta cantidad de la proteína en la fracción insoluble. En el particular caso de la rTcCyP22, los resultados obtenidos permiten confirmar que esta proteína recombinante se encuentra casi en su totalidad en la fracción insoluble, ya que en la membrana no se observó una banda correspondiente a la fracción soluble. Todo lo anterior en consistencia con los resultados obtenidos en el análisis de expresión y solubilidad durante 5 h de inducción.



**Figura 16. Inmunodetección de las proteínas rTcCyP19, rTcCyP22 y rTcCySEC.** Ensayo de Western Blot basado en las etiquetas His<sub>6</sub> que proporciona el vector de expresión pCri8a y realizado en una membrana de nitrocelulosa de acuerdo a lo indicado en Materiales y Métodos en la sección 6.4.3.. PT: Proteína Total; FS: Fracción Soluble; FI: Fracción Insoluble; M: Marcador de peso molecular.

## **7.5. Producción de Proteínas Recombinantes**

La producción de rTcCyP19, rTcCyP22 y rTcCySEc se realizó de acuerdo con lo indicado en Materiales y Métodos en la sección 6.5. En consistencia con los resultados obtenidos previamente en la prueba de inducción, el análisis de la expresión y solubilidad, así como la inmunodetección de dichas proteínas recombinantes, se confirmó que la rTcCyP19 y rTcCySEC se encuentran predominantemente en la fracción soluble, mientras que la rTcCyP22 se encuentra en su totalidad en la fracción insoluble. En las condiciones de producción para las tres proteínas recombinantes de interés se obtuvieron los siguientes pesos en la biomasa, así como sus correspondientes valores de OD<sub>600</sub> al concluir las 5 h de inducción de cada cultivo: 7.6 g y 3.018 AU para rTcCyP19, 11.6 g y 2.757 AU para rTcCyP22, así como 9.8 g y 2.832 AU para rTcCySEc.

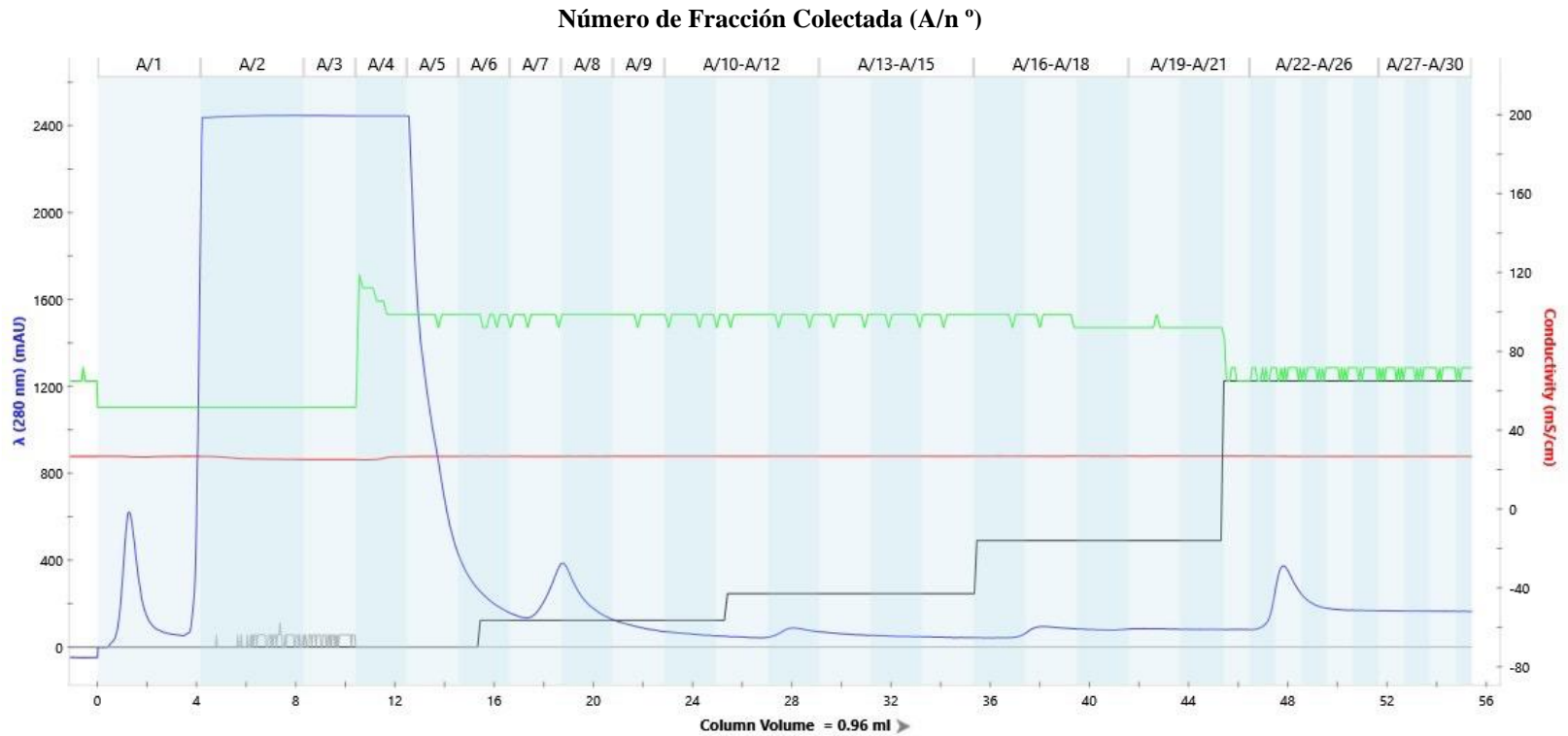
Debido a que la rTcCyP19 y rTcCySEC se encuentran predominantemente en la fracción soluble no requirieron un tratamiento adicional posterior al proceso de lisis. Sin embargo, para el caso de rTcCyP22 se realizaron los lavados a los CI con el propósito de eliminar las impurezas que pudieran contener y posteriormente se resuspendió utilizando Buffer de Solubilización. Una vez obtenidas todas las muestras se procedió con la etapa de purificación considerando la fracción en la cual se encuentra la proteína recombinante de interés.

## **7.6. Purificación por Cromatografía de Afinidad a Níquel**

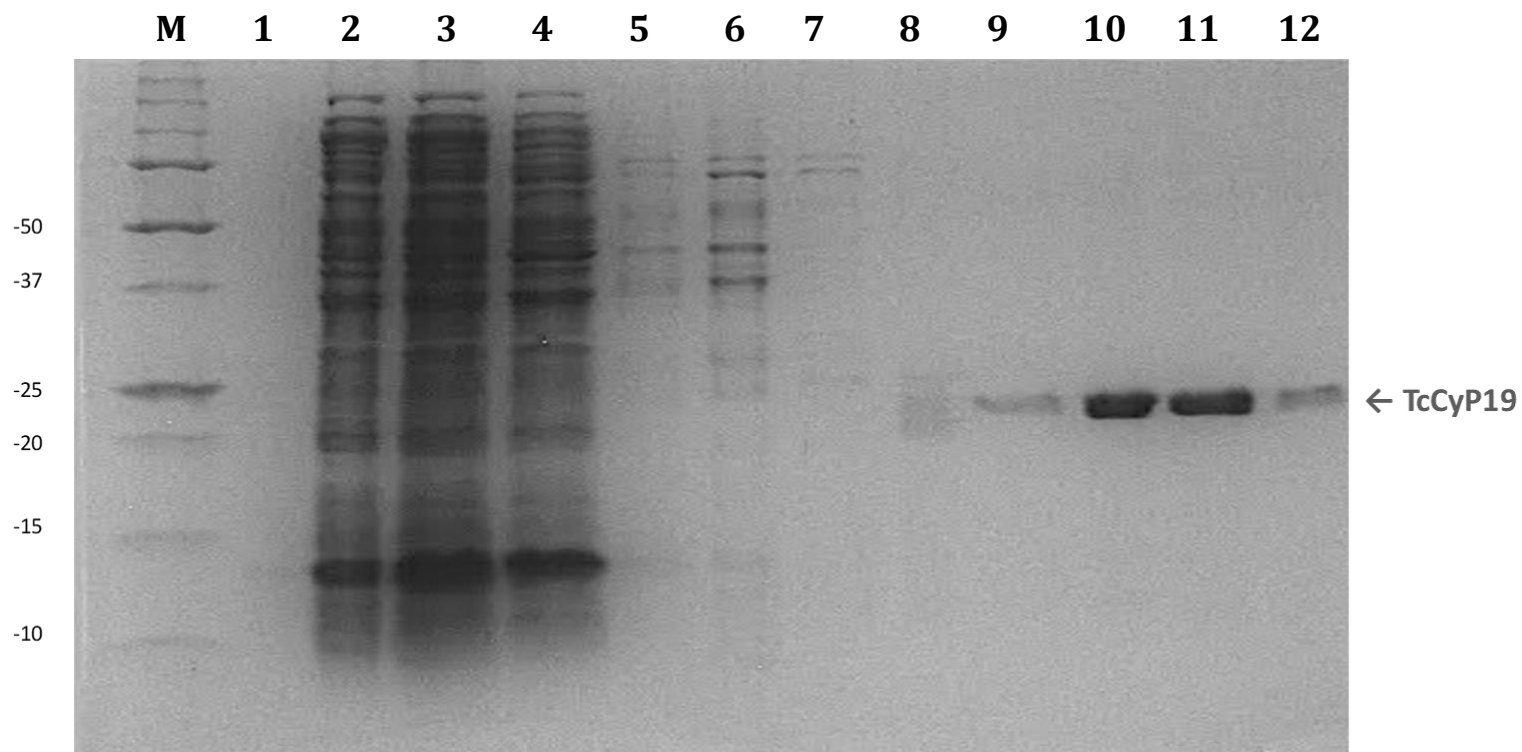
### **7.6.1. Purificación de TcCyP19 a partir de Fracción Soluble.**

La Figura 17 muestra el cromatograma de la purificación por Cromatografía de Afinidad a Níquel obtenido para la proteína rTcCyP19 realizado a partir de la fracción soluble de acuerdo con los resultados obtenidos en la etapa del análisis de la expresión en células *E. coli* BL21 (DE3), así como los resultados obtenidos por inmunodetección. Debido a la

cantidad de proteínas presentes en la muestra se realizaron purificaciones utilizando 10 mL del volumen total de la muestra, las cuales se hicieron pasar por la columna pre-empacada HisTrap™ High Performance (GE) de 1 mL. En resumen, el esquema de purificación específico de esta proteína recombinante incluyó lavados variando el porcentaje del Buffer B, incluyendo 10 VC con 5%, 10 VC con 10%, 10 VC con 20%, y finalmente realizando la elución con 10 VC con 50% de Buffer B. En el cromatograma se observan las crestas en mili unidades de Absorbancia (mAU) correspondientes a la fracción no unida (fracciones 2-5), lavados al 5% (fracciones 7-8), lavados al 10% (fracción 12), lavados al 10% (fracción 17), así como las crestas correspondientes a la elución contenida en las fracciones 22-25. Las muestras colectadas y correspondientes a las crestas obtenidas para los valores de Absorbancia a 280 nm (mAU) se trataron las muestras de acuerdo con lo indicado en la sección 6.4.1, y posteriormente se analizaron por SDS-PAGE de acuerdo con la sección 6.4.2. La Figura 18 muestra que la rTcCyP19 se une por completo a la columna de purificación, no observando la banda de esta proteína en la fracción no unida, posteriormente se observa que los lavados eliminan el resto de las impurezas, observando una ligera liberación de la proteína de interés en el lavado al 20% de Buffer B, y finalmente las muestras correspondientes a la elución muestran a la rTcCyP19 purificada, observando las bandas de mayor intensidad en las fracciones colectadas 23 y 24, que son las correspondientes a las crestas con el mayor valor en mAU para la etapa de elución.



**Figura 17. Perfil cromatográfico de la purificación de la proteína rTcCyP19.** La fracción soluble de 7.6 g de biomasa obtenida después de 5 h de inducción se administró a una columna pre-empacada HisTrap™ High Performance de 1 mL (VC) y realizando posteriormente tres etapas de lavado con el Buffer B (Tris-HCl 50 mM pH 8.0, NaCl 200 mM,  $\beta$ ME 1 mM e Imidazol 500 mM), finalmente la rTcCyP19 unida a la columna se eluyó utilizando una concentración de 250 mM de imidazol. En el cromatograma se muestra la absorbancia a 280 nm (mAU; línea azul), la fuerza iónica (mS/cm; línea roja) y concentración de imidazol (% Buffer B; línea negra) en función del VC (mL). Lo anterior de acuerdo a lo indicado en Materiales y Métodos en la sección 6.6.1. Las fracciones colectadas correspondientes a las crestas obtenidas para los valores de Absorbancia a 280 nm fueron analizadas de acuerdo con lo indicado en la sección 6.6.3. (a) 5% de Buffer B (25 mM de imidazol), (b) 10% de Buffer B (50 mM de imidazol), 20% de Buffer B (100 mM de imidazol). No se observa un cambio relevante en los valores de conductividad a lo largo de la purificación.

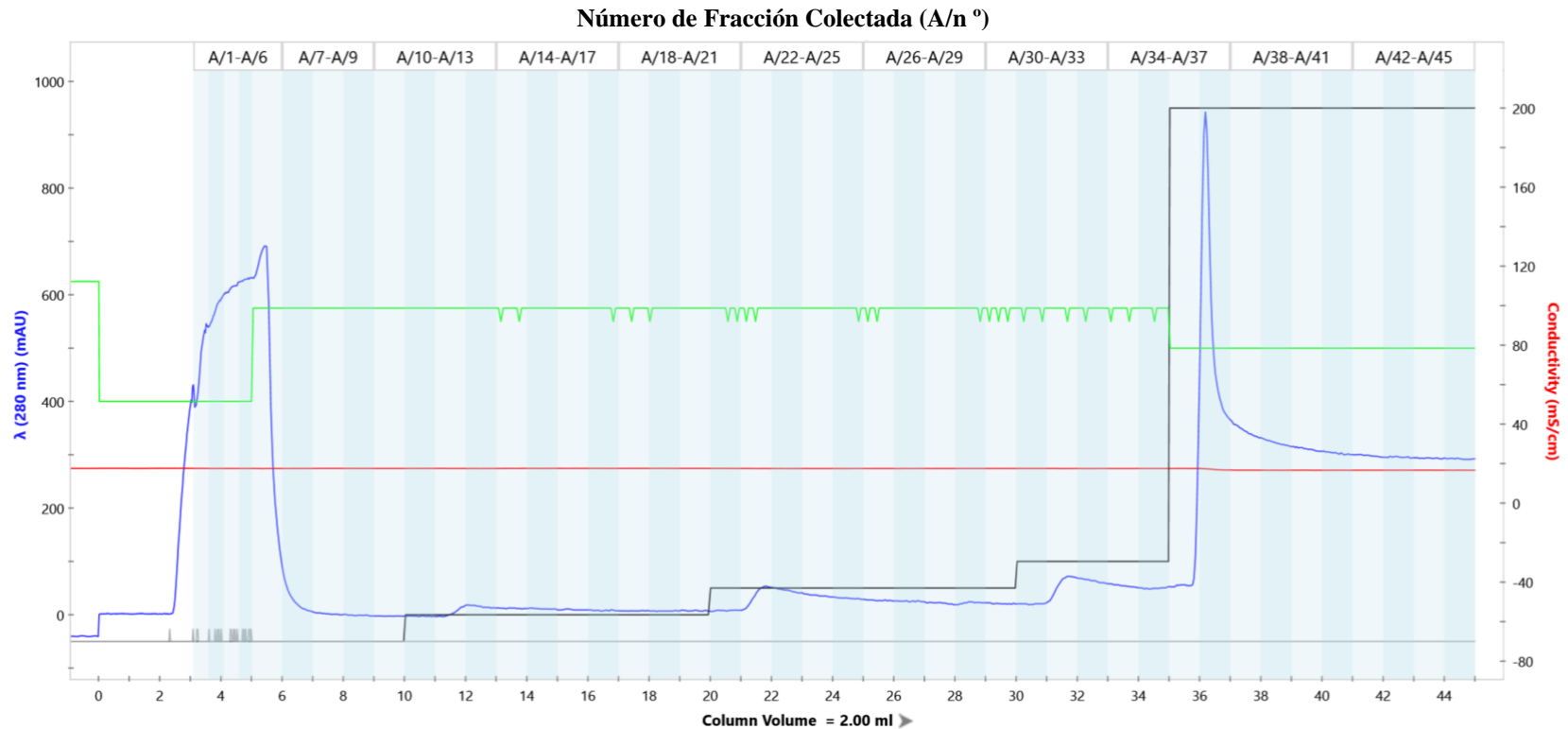


**Figura 18. Análisis por SDS-PAGE de la purificación de rTcCyP19 por afinidad a níquel.** Análisis electroforético de un gel de poliacrilamida al 15% correspondiente a las fracciones representativas de la purificación de la rTcCyP19. **M:** Marcador de peso molecular. **1:** Fracción colectada A/1 (cresta aprox. 600 mAU). **2:** Fracción no unida A/2. **3:** Fracción no unida A/3. **4:** Fracción no unida A/4. **5:** Fracción A/7 (lavado 5% Buffer B). **6:** Fracción A/8 (lavado 5% Buffer B). **7:** Fracción A/12 (lavado 10% Buffer B). **8:** Fracción A/17 (lavado 20% Buffer B). **9:** Fracción A/22 (elución 50% Buffer B). **10:** Fracción A/23 (elución 50% Buffer B). **11:** Fracción A/24 (elución 50% Buffer B). **12:** Fracción A/25 (elución 50% Buffer B).

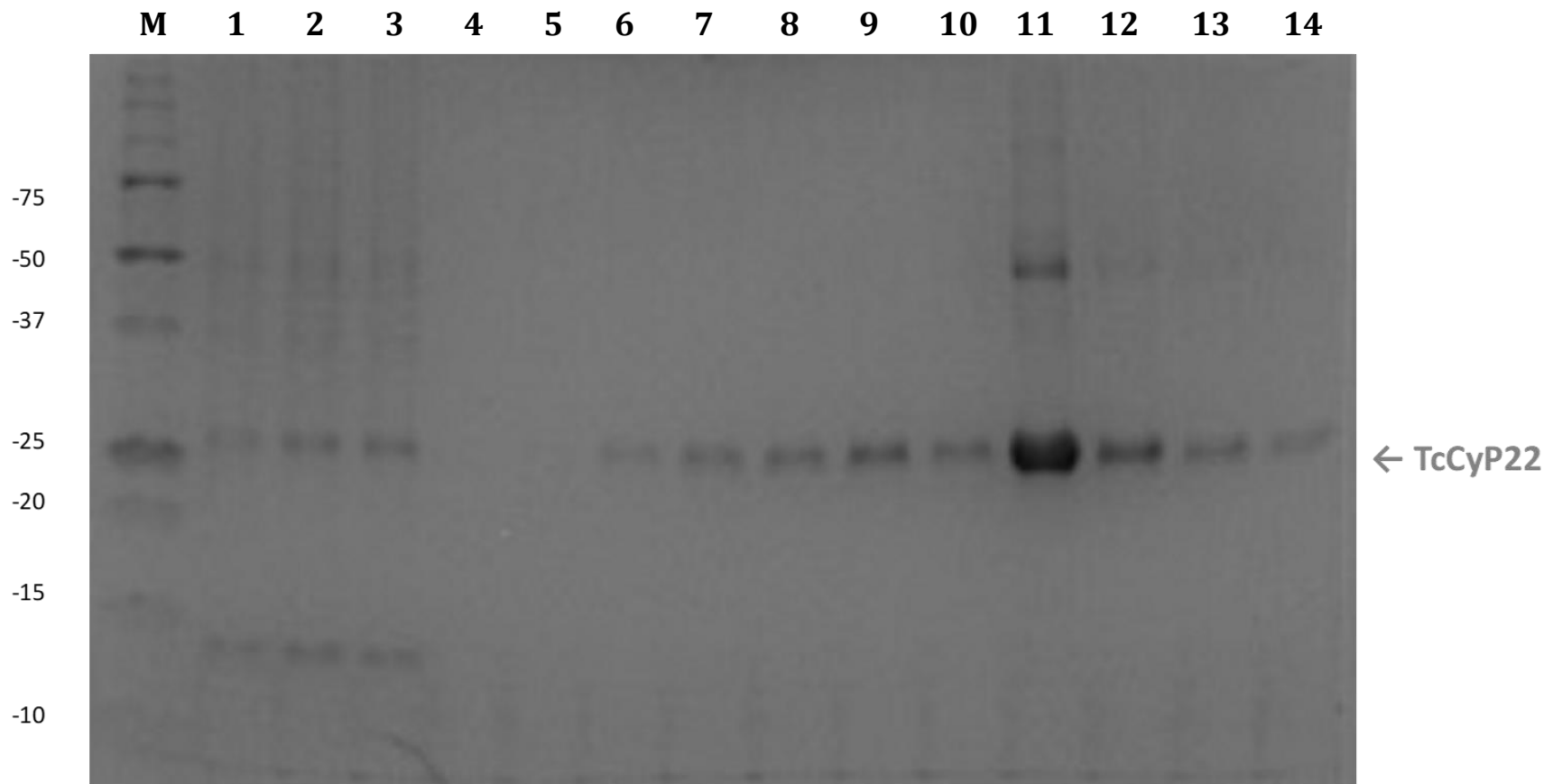
### 7.6.2. Purificación de TcCyP22 a partir de Fracción Insoluble.

La Figura 19 muestra el cromatograma de la purificación por Cromatografía de Afinidad a Níquel obtenido para la proteína rTcCyP22 realizado a partir de la fracción insoluble después de liberarla de los CI de acuerdo con los resultados obtenidos en la etapa del análisis de la expresión en células *E. coli* BL21 (DE3), así como los resultados obtenidos por inmunodetección. Debido a la cantidad de proteínas presentes en la muestra se realizaron purificaciones utilizando 10mL del volumen total de la muestra, las cuales se hicieron pasar por la columna con 1 mL de resina *GE Fast Flow 6 Sepharose Ni*. En resumen, el esquema de purificación específico de esta proteína recombinante incluyó lavados variando el porcentaje del Buffer B, incluyendo 10 VC con 5%, 10 VC con 10%, 5 VC con 15%, y finalmente realizando la elución con 10 VC con 100% de Buffer B. En el cromatograma se observan las crestas en mAU correspondientes a la fracción no unida (fracciones 1-6), lavados al 5% (fracciones 12-13), lavados al 10% (fracciones 22 y 23), lavados al 15% (fracciones 32 y 33), así como las crestas correspondientes a la elución contenida en las fracciones 36-37, sin embargo, se analizó hasta la fracción 40. Las muestras colectadas y correspondientes a las crestas obtenidas para los valores de Absorbancia a 280 nm (mAU) se trataron las muestras de acuerdo con lo indicado en la sección 6.4.1, y posteriormente se analizaron por SDS-PAGE de acuerdo con la sección 6.4.2. La Figura 20 muestra que la rTcCyP22 no se une por completo a la columna de purificación, observando la banda de esta proteína en la fracción no unida, adicionalmente se observa que los lavados, exceptuando el lavado al 5%, incluyen una liberación gradual de la proteína de interés conforme se aumenta el % de Buffer B. Finalmente las muestras correspondientes a la elución muestran a la rTcCyP22 prácticamente purificada más una banda de aproximadamente 50 kDa, observando la banda de mayor intensidad en la fracción colectadas 37, que es la correspondiente a la cresta con el mayor valor en mAU.





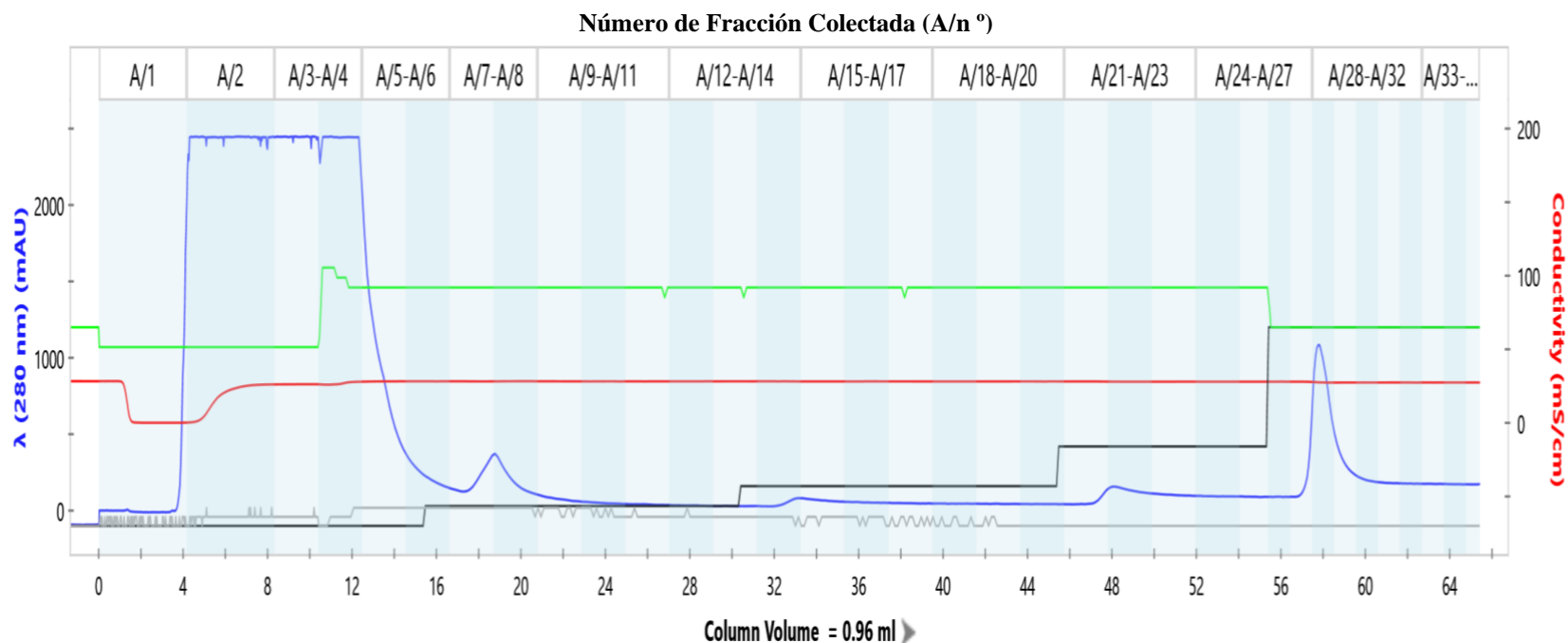
**Figura 19. Perfil cromatográfico de la purificación de la proteína rTcCyP22.** La fracción soluble de 11.6 g de biomasa obtenida después de 5 h de inducción se administró a una columna con 1 mL (VC) de resina *GE Fast Flow 6 Sepharose Ni* y realizando posteriormente tres etapas de lavado con el Buffer B (Tris-HCl 50 mM pH 8.0, NaCl 200 mM,  $\beta$ ME 1 mM e Imidazol 500 mM), finalmente la rTcCyP22 unida a la columna se eluyó utilizando una concentración de 500 mM de imidazol. En el cromatograma se muestra la absorbancia a 280 nm (mAU; línea azul), la fuerza iónica (mS/cm; línea roja) y concentración de imidazol (% Buffer B; línea negra) en función del VC (mL). Lo anterior de acuerdo a lo indicado en Materiales y Métodos en la sección 6.6.2. Las fracciones colectadas correspondientes a las crestas obtenidas para los valores de Absorbancia a 280 nm fueron analizadas de acuerdo con lo indicado en la sección 6.6.3. (a) 5% de Buffer B (25 mM de imidazol), (b) 10% de Buffer B (50 mM de imidazol), 15% de Buffer B (75 mM de imidazol). No se observa un cambio relevante en los valores de conductividad a lo largo de la purificación.



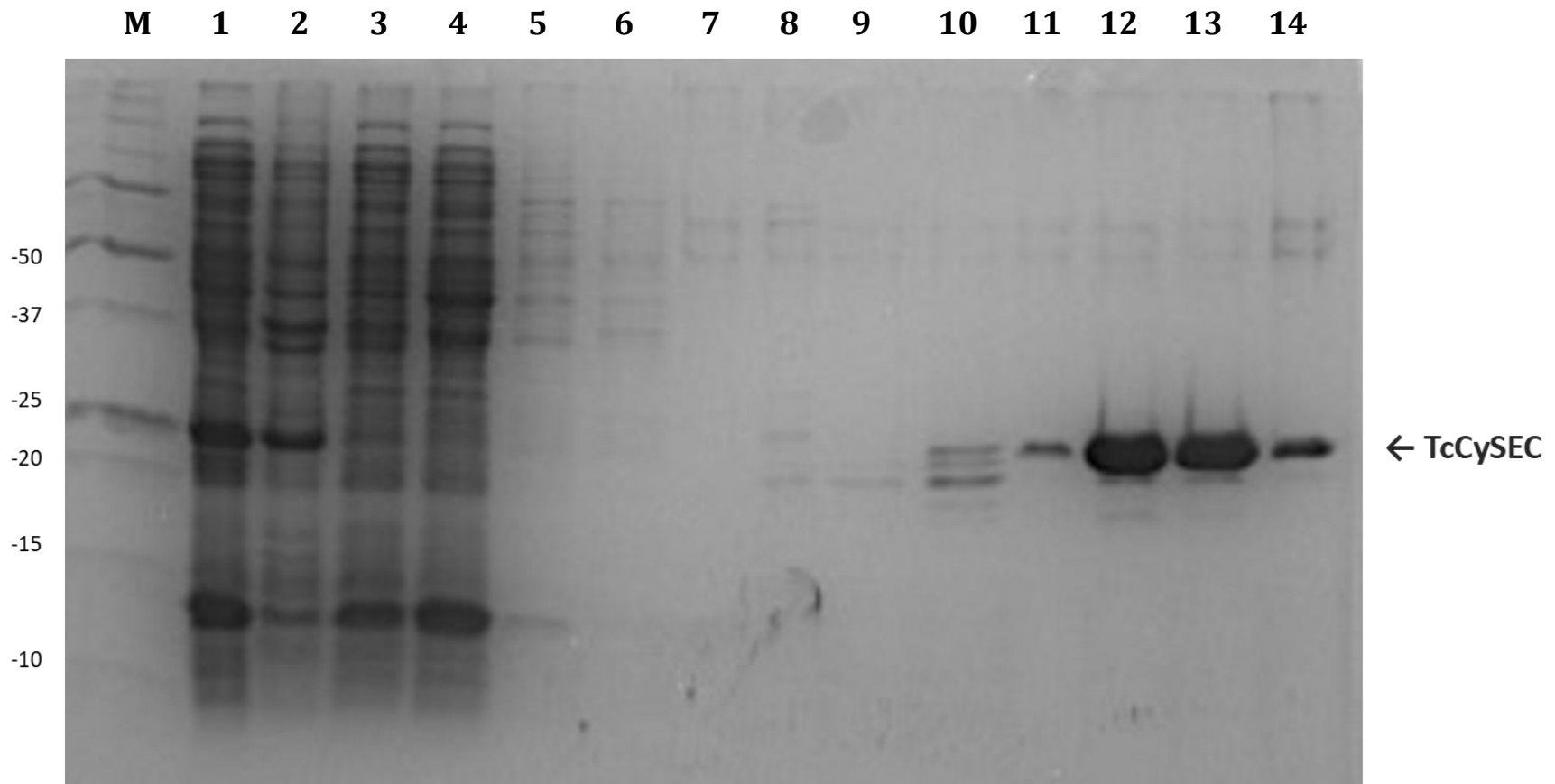
**Figura 20. Análisis por SDS-PAGE de la purificación de rTcCyP22 por afinidad a níquel.** Análisis electroforético de un gel de poliacrilamida al 15% correspondiente a las fracciones representativas de la purificación de la rTcCyP22. **M:** Marcador de peso molecular. **1:** Fracción no unida A/1. **2:** Fracción no unida A/3. **3:** Fracción no unida A/6. **4:** Fracción A/12 (lavado 5% Buffer B). **5:** Fracción A/13 (lavado 5% Buffer B). **6:** Fracción A/22 (lavado 5% Buffer B). **7:** Fracción A/23 (lavado 10% Buffer B). **8:** Fracción A/32 (lavado 15% Buffer B). **9:** Fracción A/33 (lavado 15% Buffer B). **10:** Fracción A/36 (elución 100% Buffer B). **11:** Fracción A/37 (elución 100% Buffer B). **12:** Fracción A/38 (elución 100% Buffer B). **13:** Fracción A/39 (elución 100% Buffer B). **14:** Fracción A/40 (elución 100% Buffer B).

### 7.6.3. Purificación de TcCySEC a partir de Fracción Soluble.

La Figura 21 muestra el cromatograma de la purificación por Cromatografía de Afinidad a Níquel obtenido para la proteína rTcCySEC realizado a partir de la fracción soluble de acuerdo con los resultados obtenidos en la etapa del análisis de la expresión en células *E. coli* BL21 (DE3), así como los resultados obtenidos por inmunodetección. Debido a la cantidad de proteínas presentes en la muestra se realizaron purificaciones utilizando 10 mL del volumen total de la muestra, las cuales se hicieron pasar por la columna pre-empacada HisTrap™ High Performance (GE) de 1 mL. El esquema de purificación específico de esta proteína recombinante incluyó lavados variando el porcentaje del Buffer B, incluyendo 25 VC con 5%, 25 VC con 10%, 10 VC con 20%, y finalmente realizando la elución con 10 VC con 50% de Buffer B. En el cromatograma se observan las crestas en mAU correspondientes a la fracción no unida (fracciones 2-4), lavados al 5% (fracciones 7-8), lavados al 10% (fracciones 14 y 15), lavados al 10% (fracciones 21-22), así como las crestas correspondientes a la elución contenida en las fracciones 27-30. Las muestras colectadas y correspondientes a las crestas obtenidas para los valores de Absorbancia a 280 nm (mAU) se trataron las muestras de acuerdo con lo indicado en la sección 6.4.1, y posteriormente se analizaron por SDS-PAGE de acuerdo con la sección 6.4.2. La Figura 22 muestra que la rTcCySEC se une casi por completo a la columna de purificación, prácticamente no observando la banda de esta proteína en la fracción no unida, posteriormente se observa que los lavados eliminan el resto de las impurezas, observando una ligera liberación de la proteína de interés en la fracción 22 del lavado al 20% de Buffer B, y finalmente las muestras correspondientes a la elución muestran a la rTcCySEC purificada con algunas impurezas de peso molecular similar así como de alto peso, observando las bandas de mayor intensidad en las fracciones colectadas 28 y 29, correspondientes a las crestas con el mayor valor en mAU para la etapa de elución.



**Figura 21. Perfil cromatográfico de la purificación de rTcCySEC.** La fracción soluble de 9.8 g de biomasa obtenida después de 5h de inducción se administró a una columna pre-empacada HisTrap™ High Performance de 1 mL (VC) y realizando posteriormente tres etapas de lavado con el Buffer B (Tris-HCl 50 mM pH 8.0, NaCl 200 mM,  $\beta$ ME 1 mM e Imidazol 500 mM), finalmente la rTcCyP19 unida a la columna se eluyó utilizando una concentración de 250 mM de imidazol. En el cromatograma se muestra la absorbancia a 280 nm (mAU; línea azul), la fuerza iónica (mS/cm; línea roja) y concentración de imidazol (% Buffer B; línea negra) en función del VC (mL). Lo anterior de acuerdo a lo indicado en Materiales y Métodos en la sección 6.6.1. Las fracciones colectadas correspondientes a las crestas obtenidas para los valores de Absorbancia a 280 nm fueron analizadas de acuerdo con lo indicado en la sección 6.6.3. (a) 5% de Buffer B (25 mM de imidazol), (b) 10% de Buffer B (50 mM de imidazol), 20% de Buffer B (100 mM de imidazol). Se observó un ligero cambio en los valores de conductividad al inicio de la purificación, coincidiendo con la etapa de administración de la muestra.



**Figura 22. Análisis por SDS-PAGE de la purificación de rTcCySEC por afinidad a níquel.** Análisis electroforético de un gel de poliacrilamida al 15% correspondiente a las fracciones representativas de la purificación de la rTcCySEC. **M:** Marcador de peso molecular. **1:** Fracción soluble. **2:** Fracción insoluble. **3:** Fracción no unida A/2. **4:** Fracción no unida A/4. **5:** Fracción A/7 (lavado 5% Buffer B). **6:** Fracción A/8 (lavado 5% Buffer B). **7:** Fracción A/14 (lavado 10% Buffer B). **8:** Fracción A/15 (lavado 10% Buffer B). **9:** Fracción A/21 (lavado 20% Buffer B). **10:** Fracción A/22 (lavado 20% Buffer B). **11:** Fracción A/27 (elución 50% Buffer B). **12:** Fracción A/28 (elución 100% Buffer B). **13:** Fracción A/29 (elución 100% Buffer B). **14:** Fracción A/30 (elución 100% Buffer B).

### 7.7. Ensayo de Actividad de PPIasa por Replegamiento de RNasa T

Las muestras obtenidas de la purificación se transfirieron al Buffer de Almacenamiento y posteriormente se realizó la cuantificación por densitometría, de acuerdo con lo descrito en Materiales y Métodos en la sección 6.7. La Tabla 14 describe los volúmenes requeridos para cada uno de los ensayos de actividad realizados en una celda de 1 mL conteniendo la proteína rTcCyP19 a una concentración final 25 nM y la rTcCySEC a 50 nM así como 75 nM, considerando la concentración de las muestras obtenidas por PD-10. Debido a la tendencia de agregación de la rTcCyP22, esta proteína recombinante que se expresa en la fracción insoluble, no se consideró para los ensayos de actividad de PPIasa por lo cual no se describe en la tabla siguiente. El análisis por DLS mostró que tanto rTcCyP19 como rTcCySEC presentaron polidispersión, observando picos homogéneos en el gráfico de distribución por intensidad para la rTcCyP19, pero no en el gráfico de distribución por masa. En el caso de la rTcCySEC no se observaron picos homogéneos en los gráficos de distribución por intensidad o masa (figuras no mostradas).

**Tabla 14. Ensayos de Actividad de PPIasa**

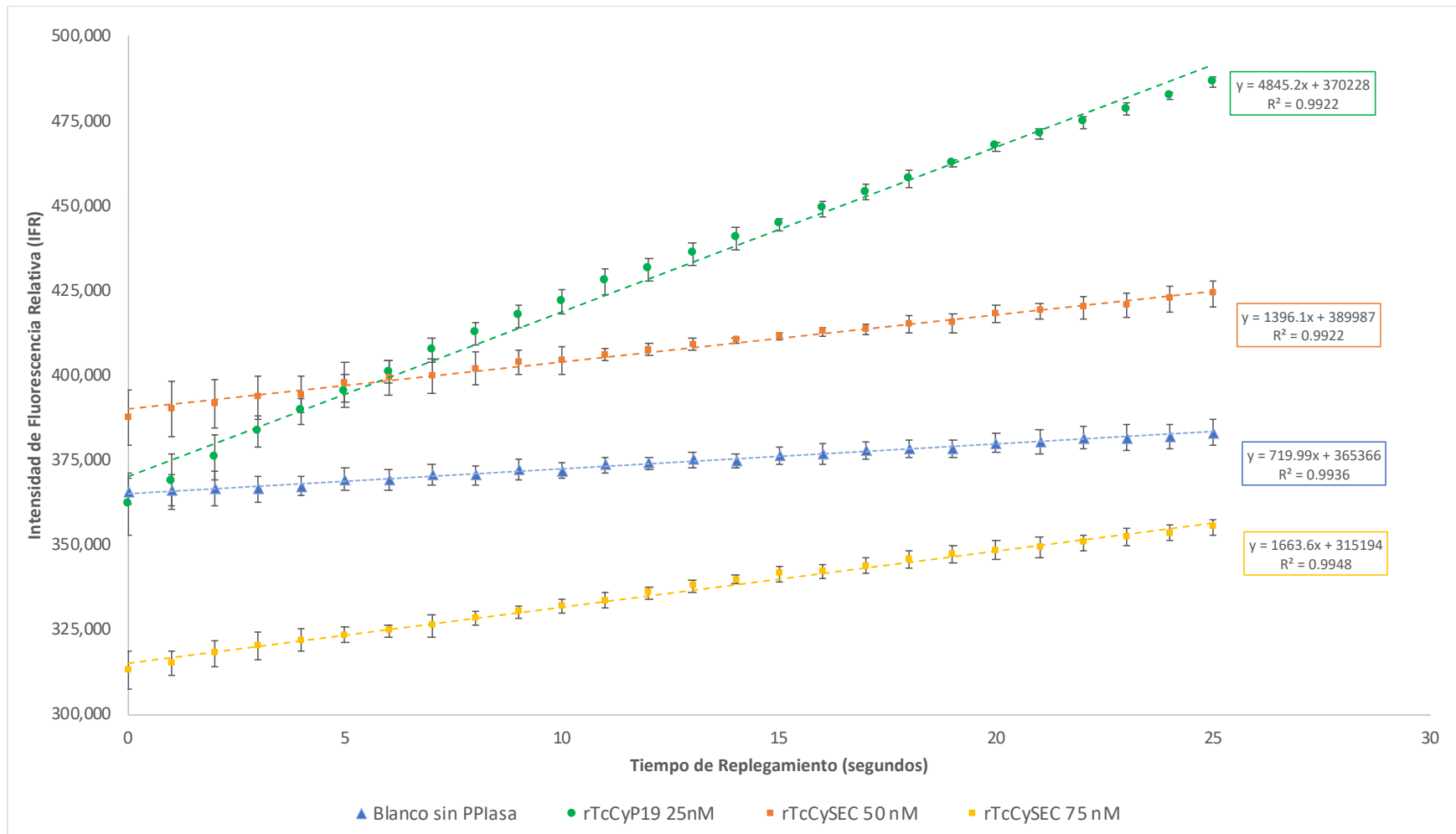
Condición	Tris-HCl 22 mM pH 8.0 (μl) <sup>a</sup>	RNasa T 50 μM desnaturalizada (μl) <sup>b</sup>	H <sub>2</sub> O Milli-Q (μl) <sup>c</sup>	rTcCyP19 1.06 μM (μl) <sup>d</sup>	rTcCySEC 2.06 μM (μl) <sup>e</sup>
Control (-) rTcCyP19	910	---	66.5	23.5 μl	---
Control (-) rTcCySEC	910	---	53.5	---	36.5
Blanco sin PPIasa	910	40	50	---	---
Replegamiento por rTcCyP19	910	40	26.5	23.5	---
Replegamiento por rTcCySEC	910	40	13.5	---	36.5

<sup>a</sup> Tris-HCl 22 mM pH 8.0: Buffer de Replegamiento requerido para una concentración final de 20 mM. <sup>b</sup> RNasa T1 50 μM desnaturalizada a una concentración final de 8M de Urea. <sup>c</sup> H<sub>2</sub>O Milli-Q requerida para aforar la celda de 1 mL. <sup>d</sup> Volumen requerido de la muestra de rTcCyP19 partiendo de una concentración 2.06 μM obtenida por PD-10. <sup>e</sup> Volumen requerido de la muestra de rTcCySEC partiendo de una concentración 1.06 μM obtenida por PD-10.

Los ensayos de actividad se realizaron en la celda de 1 mL en las condiciones descritas en la sección 6.7 y con los volúmenes anteriormente especificados evaluando la variación en la Intensidad de Fluorescencia Relativa (IFR) respecto a los primeros 25 s de la reacción, considerada la parte linear de la curva de cinética de reacción obtenida en el espectrofluorómetro. Posteriormente, se realizó la regresión lineal para obtener la ecuación de la recta. La Figura 23 muestra el gráfico de la cinética enzimática correspondiente al ensayo de actividad por replegamiento de la RNasa T1, así como la ecuación de la recta y su  $R^2$ , el análisis realizado de este gráfico mostró que la pendiente de la recta (m) de mayor magnitud se presentó en el siguiente orden rTcCyP19 25 nM > rTcCySEC 75 nM > rTcCySEC 50 nM > Blanco sin PPIasa. Los valores obtenidos se utilizaron para determinar la eficiencia catalítica de cada una de las proteínas recombinantes considerando los nmol contenidos en 1 mL de reacción para cada una de las concentraciones evaluadas. La Tabla 15 describe los valores utilizados para determinar la actividad de PPIasa expresada en  $\Delta$  IFR/s nmol para cada una de las condiciones evaluadas.

**Tabla 15. Determinación de la actividad de PPIasa (Eficiencia Catalítica)**

Condición	m (IFR/s)	nmol en 1 mL	$\Delta$ IFR/s	Actividad PPIasa ( $\Delta$ IFR/s nmol)
Blanco sin PPIasa	720.0	NA	NA	NA
rTcCyP19 25 nM	4,845.2	0.025	4,125.2	165,008.4
rTcCySEC 50 nM	1,396.1	0.05	676.1	13,522.2
rTcCySEC 75 nM	1,663.6	0.075	943.6	12,581.5



**Figura 23. Cinética de Replegamiento de la RNasa T1 utilizando rTcCyP19 o rTcCySEC.** Valores de IFR obtenidas en el espectrofluorómetro para cada una de las condiciones evaluadas durante los primeros 25 s de reacción. En cada condición evaluada se especifican la ecuación de la recta y  $R^2$ . Blanco sin PPIasa (línea azul). rTcCyP19 25 nM (línea verde). rTcCySEC 50 nM (línea naranja). rTcCySEC 75 nM (línea amarilla).



## 8. DISCUSIÓN

El análisis bioinformático de las proteínas de interés TcCyP19, TcCyP22 y TcCySEC permitió confirmar que, aunque ya estudiadas estas proteínas aún existe información relevante por caracterizar, por ejemplo, los parámetros fisicoquímicos de las proteínas recombinantes, su estructura cristalizada, la solubilidad al expresarse en *E. coli* u otra plataforma de expresión, etc. La búsqueda de los códigos de acceso de las proteínas de interés (ver Tabla 1) permitió consolidar toda la información de las diferentes bases de datos, establecer las condiciones para cada análisis y finalmente interpretar los resultados en cada etapa del análisis bioinformático. En el caso de TcCySEC se confirmó que ya ha sido cristalizada identificándose también como TcMIP (PDB ID 1JVW), considerando que la secuencia de nuestra proteína no posee el péptido señal constituido por 29 aminoácidos. Por lo anterior, el análisis de BLAST (ver Tabla 2) de TcCySEC demostró que, aunque con 100% de identidad en la secuencia, la cobertura (Query Cover) resultó del 85% o 100% dependiendo de sí el análisis se realizó respecto a la proteína cristalizada con o sin péptido señal, que representa el 15% de la secuencia completa de TcMIP. En el caso de TcCyP19 se encontró que al ser ortólogo de la Ciclofilina A humana (Bell et al., 2006; Bua et al., 2001; Rego et al., 2015), posee una alta identidad en la secuencia con PPIasas de otros organismos que por supuesto resultan ser ortólogos de dicha Ciclofilina A. Respecto a la TcCyP22 está es ortóloga de la Ciclofilina D mitocondrial presente en mamíferos (Perrone et al., 2018). Sin embargo, cabe resaltar que los mejores parámetros del análisis de BLAST se obtuvieron para otro parásito, la PPIasa 2HQJ de *L. major*. Considerando lo anterior, el MSA no incluyó a la TcCySEC ya que adicionalmente posee el dominio de MIP dentro del cual se encuentra el dominio de PPIasa que además posee menos aminoácidos en su secuencia respecto a la TcCyP19 y TcCyP22, por lo cual el resultado de los alineamientos en las diferentes condiciones resultaría en un muy bajo número de sitios conservados, en

resumen, sería equivalente a comparar dos enzimas con actividad diferente. Respecto al MSA considerando a la TcCyP19 y TcCyP22, se observa que los sitios conservados, mutaciones conservativas y semi-conservativas son similares en los MSA de las proteínas de interés versus PPIasas de otras especies (Figura 5) y PPIasas de otros parásitos (Figura 6). La Figura 7 del MSA realizado únicamente entre las tres proteínas de interés confirma la desventaja de realizar este análisis al incluir a TcCySEC. Finalmente, el MSA entre la TcCyP19 y TcCyP22 (Figura 8), permitió confirmar que los sitios conservados aumentaron notablemente, permitiendo sugerir que la conservación en la secuencia está relacionada con la especie, manteniéndose las regiones que le otorgan a las proteínas de interés la actividad de PPIasa.

El análisis de los parámetros fisicoquímicos determinados en ProtParam (Gasteiger E., 2005) a partir de la secuencia de las proteínas de TcCyP19, TcCyP22 y TcCySEC permitió sugerir, de acuerdo con la Tabla 3, que las tres proteínas presentan valores similares entre sí. Esto sugiere que las proteínas de interés pueden expresarse en *E. coli* con un  $t_{1/2}$  mayor a 10 h y de manera soluble de acuerdo a los valores de GRAVY. Se obtuvieron valores en su pI similares y mayores a 7.5 para las proteínas TcCyP19 y TcCyP22, considerando que de acuerdo con el análisis de BLAST y MSA son proteínas similares en su secuencia. En contraste, la TcCySEC que es una proteína distinta en su secuencia de acuerdo con los resultados del análisis de BLAST y MSA, su pI resultó de 5.83. Estos valores del pI se consideraron para definir el pH de los amortiguadores utilizados en la etapa de purificación. Conociendo los resultados de búsqueda en las diferentes bases de datos, análisis de BLAST y MSA se realizó el modelado por Homología en los servidores Swiss-Model, trRosetta y I-TASSER, a partir de las secuencias de TcCyP19 y TcCyP22 utilizando modelos ya conocidos de PPIasas ya cristalizadas. A partir de los modelos obtenidos en cada servidor se realizó la optimización y validación de los modelos 3D, definiendo un modelo óptimo para cada servidor (Tablas 4-9), que se utilizaron para la determinación *in*

*in silico* de solubilidad y agregación, predicción corregida estructuralmente. Los resultados basados en su secuencia proteica de acuerdo con lo descrito en la Tabla 10, así como en su estructura 3D de acuerdo en lo descrito en la Tabla 11, demostraron que la TcCySEC es la proteína de mayor solubilidad o menor tendencia de agregación de acuerdo con el análisis *in silico*, apoyado en el fundamento teórico descrito en el **Apéndice 6**. La última etapa del análisis bioinformático permitió sugerir que las proteínas recombinantes de interés se expresarían de manera soluble en *E. coli*, es decir, tienen una baja tendencia de agregación, que al compararlos con los obtenidos en la etapa experimental se observa que solo la TcCyP19 y TcCySEC se expresan de forma soluble, mientras que la TcCyP22 se expresa, en contraste con la predicción *in silico*. En todas las etapas del análisis bioinformático se consideró que la TcCySEC es la proteína más estudiada de las tres proteínas seleccionadas (Moro et al., 1995; Pereira et al., 2002)), por lo cual ésta fungió como proteína de referencia para las diversas etapas del proyecto. La etapa de expresión de TcCyP19 y TcCyP22 permitirá validar esta predicción *in silico*, ya que como se mencionó TcCySEC se expresa en forma soluble. En este proyecto se contó con genes sintéticos para demostrar la expresión de las tres ciclofilinas y así poder confirmar experimentalmente la solubilidad y agregación de estas PPIasas.

La prueba de inducción, el análisis de la expresión y solubilidad de las proteínas recombinantes, así como la inmunodetección de las proteínas recombinantes con las etiquetas de His<sub>6</sub> en los extremos NH<sub>2</sub>- y COOH- que cada proteína posee, permitió definir las condiciones de producción de rTcCyP19, rTcCyP22 y rTcCySEC, confirmando que aunque el análisis bioinformático sugiere que la TcCyP22 se expresa en forma soluble en *E. coli* experimentalmente, la producción de esta proteína recombinante en células *E. coli* BL21 (DE3) confirmó que se encuentra en la fracción insoluble como se muestra en la Figura 14. En cambio, la rTcCyP19 y rTcCySEC en consistencia con los resultados del análisis bioinformático, ambas se expresaron de manera soluble en las células *E. coli* BL21

(DE3) tal cual se muestra en la Figura 13 y la Figura 15, respectivamente. Específicamente para el análisis de la expresión y solubilidad de las tres proteínas recombinantes, se puede observar en la Figura 12 en términos cualitativos que la rTcCyP22 tiene un crecimiento más rápido y distinto a la rTcCyP19 y rTcCySEC que tiene un patrón de crecimiento bastante similar entre ambas. Debido a esta producción acelerada de la proteína rTcCyP22 puede presentarse la situación que la tasa de expresión rebasa la tasa de plegamiento de esta proteína, lo cual da lugar a la formación de CI y por consiguiente que se encuentre en la fracción insoluble, todo lo anterior bajo las condiciones de crecimiento, inducción y expresión descritas en materiales y métodos. Adicionalmente, los pesos de la biomasa (pellet) obtenidos para los cultivos que expresaron rTcCyP19, rTcCyP22 y rTcCySEC fueron de 7.6, 11.6 y 9.8 g, respectivamente, lo que permite sugerir en el caso de rTcCyP22 que debido a que no se obtiene en su conformación nativa, las células de este cultivo ahorran energía permitiendo un crecimiento más rápido en consistencia con lo discutido para la Figura 12.

La producción de rTcCyP19, rTcCyP22 y rTcCySEC en una escala mayor, fue consistente con los resultados obtenidos en la prueba de inducción, el análisis de la expresión y solubilidad, así como la inmunodetección de dichas proteínas recombinantes; considerando que en esta etapa se utilizó medio TB en lugar de medio LB. A partir de las muestras provenientes de la cosecha y lisis celular se procedió a la etapa de purificación observando que la rTcCyP19 es la proteína recombinante que no solo se expresa en condiciones ideales al encontrarse en la fracción soluble, sino también se obtiene con un alto grado de pureza de acuerdo con lo observado en la Figura 17 y Figura 18, obteniendo las muestras colectadas en una concentración adecuada para los ensayos de actividad. Aunque esta proteína ya ha sido expresada y purificada en otros estudios (Rego et al., 2015), este proyecto se diferenció en el protocolo de expresión utilizando medio TB y el protocolo de purificación basada en las etiquetas His<sub>6</sub> que proporciona el vector de

expresión pCri8a. Respecto a la rTcCySEC, donde se partió de la fracción soluble, el protocolo de purificación establecido no permitió obtener la proteína recombinante con el mismo grado de pureza que la rTcCyP19, ya que aunque la cresta correspondiente a la etapa de elución observada en la Figura 21 demuestra la liberación de la proteína de interés, en el análisis electroforético del gel en la Figura 22 se observa que el carril correspondiente a las muestras colectadas en la etapa de elución aún contienen varias impurezas ya sean de peso similar o superior al correspondiente a esta proteína recombinante Sin embargo por la experiencia en el equipo de trabajo esta calidad en la pureza es adecuada para realizar los ensayos de actividad de PPIasa. Un hallazgo relevante en la producción de esta proteína es que al cambiar al Buffer PBS 1X de acuerdo con el protocolo de PD-10, que es el buffer en el que se entregó para la obtención de anticuerpos *Mouse  $\alpha$ -rTcCySEC*, la proteína no muestra tendencia a agregarse lo cual coincide con lo sugerido en el análisis *in silico*. En el caso de la rTcCyP22, donde en este caso se utilizó la fracción insoluble, el protocolo de purificación establecido en condiciones desnaturalizantes no permitió obtener la proteína recombinante con la pureza esperada, ya que aunque la cresta correspondiente a la etapa de elución observada en la Figura 19 demuestra la liberación de la proteína de interés, en el análisis del gel de la Figura 20 se observa que el carril correspondiente a las muestras colectadas en la etapa de elución aún contienen varias impurezas, incluyendo la presencia de una banda con un peso molecular de aproximadamente el doble de la rTcCyP22, lo que pudiera interpretarse como la formación de dímeros, sin embargo, dadas las condiciones desnaturalizantes y reductoras del SDS.PAGE es posible que se trate de otra proteína, pero esto no se confirmó por inmunodetección. Por otra parte, contrario a lo sugerido en el análisis *in silico*, la producción de esta proteína recombinante bajo las condiciones descritas no se encuentra en la fracción soluble. Algunas adecuaciones al protocolo con el propósito de expresarla de manera soluble son inducir a una menor temperatura, utilizar otros medios de cultivo,

cambiar a otras cepas que promuevan su expresión soluble, adicionalmente en las etapas de purificación y replegamiento se podrían utilizar amortiguadores con un pH mayor al pI teórico de 8.98 de la TcCyP22 (ver Tabla 3).

Debido a razones de tiempo se decidió determinar la actividad enzimática de las PPIasas rtCCyP19 y rTcCySEC purificadas a partir de la fracción soluble. Es importante mencionar que en el ensayo de actividad de PPIasa por replegamiento de la RNasa T1 se requirió de una concentración de 25 nM para rTcCyP19 mientras que para rTcCySEC se utilizó el doble (50 nM) y el triple (75 nM) de concentración para obtener pendientes similares. Estos resultados indican que la eficiencia catalítica ( $\Delta$  IFR/s nmol) de la rTcCyP19 es 12 veces superior a la rTcCySEC. Los resultados reportados en la literatura para TcCyP19 y TcMIP (TcCySEC) se obtuvieron mediante el ensayo espectrofotométrico para la determinación de constantes cinéticas de PPIasas (Kofron et al., 1991) al ensayo por replegamiento de la RNasa T1, por lo que los resultados obtenidos en este estudio no se pueden comparar directamente con los reportados en la literatura. Sin embargo, al evaluar la relación de las constantes cinéticas de primer orden reportadas como isomerización catalizada ( $K_{obs} = 0.03949/s$ ) y no catalizada ( $K_{unecat} = 0.0134/s$ ) de TcMIP /TcCySEC) se obtuvo una relación  $K_{obs}/K_{unecat}$  del orden de 2.9 (Moro et al., 1995). Realizando este mismo cálculo para las pendientes (ver Tabla 15) de la condición Blanco sin PPIasa y las condiciones rTcCySEC 50 nM y rTcCySEC 75 nM se obtiene una relación  $m_{SEC}/m_{Blanco}$  del orden de 1.9 y 2.3, respectivamente, lo que sugiere que la actividad obtenida en este estudio es similar a la reportada previamente. En el caso de la TcCyP19, estudios previos (Bua et al., 2001; Bua et al., 2004) mencionan la determinación de la actividad de PPIasa de esta proteína, pero no reportan un valor específico, por lo cual este proyecto es el primero en reportar un valor específico de eficiencia catalítica. Resulta interesante que la relación  $m_{P19}/m_{Blanco}$  de la rTcCyP19 es de 6.7, es decir, 3 veces mayor que la estimada para rTcCySEC (ver Tabla 15). Aunque no fue posible extrapolar los valores de eficiencia catalítica obtenidos en este

proyecto respecto a los resultados reportados en la literatura, si se confirmó que ambas proteínas presentan actividad de PPIasa incluso al expresarlas de manera recombinante utilizando diferentes vectores de expresión y analizar su actividad por métodos, temperaturas, Buffers, pH diferentes. Es importante considerar que los métodos y condiciones utilizados para los ensayos de actividad, así como las proteínas recombinantes obtenidas son distintas por lo cual esta comparación es relativa. Analizando a detalle los resultados obtenidos y las características de cada PPIasa, es posible suponer que, en el caso de la rTcCySEC, perteneciente a la subfamilia de las FKBP, existen diversos factores que pudieran disminuir su eficiencia catalítica, por ejemplo, que su dominio de PPIasa no solo es de menor tamaño en el número de residuos respecto a la rTcCyP19, sino que además este dominio PPIasa se encuentra dentro del dominio MIP (ver Figura 3 y 25). Adicionalmente pudiese presentarse la auto asociación de dos alfa hélices existentes (ver Figura 9) y por supuesto debe considerarse la afinidad de la rTcCySEC al sustrato utilizado, la RNasa T1. Existe información reportada que las FKBP tienden a intervenir predominantemente en otras reacciones catalíticas que en la *cis/trans* isomerización de los enlaces que preceden a la prolina, sobre todo cuando existe más de un dominio presente en la PPIasa como en el caso de la TcCySEC/TcMIP. Además la presencia de otro dominio en la secuencia de las FKBP las hace más propensas a interactuar con otros compuestos presentes en el medio en que se encuentren lo cual introduce una variable más en su actividad como PPIasa, (Schiene & Fischer, 2000; Schiene-Fischer & Yu, 2001). En el caso de rTcCyP19, perteneciente a la subfamilia de las ciclofilinas, está proteína solo cuenta con el dominio PPIasa que representa aproximadamente el 90% de la secuencia, permitiendo suponer que existe una única función para esta enzima al interactuar con el sustrato utilizado y que de acuerdo a los resultados obtenidos permite sugerir una afinidad mayor a la RNasa T1 que la que mostraría la rTcCySEC.

Los resultados obtenidos del análisis por DLS no fueron concluyentes ya que tanto rTcCyP19 y rTcCySEC presentaron polidispersión que probablemente se deben a que ambas muestras no estaban recientemente obtenidas de la purificación y/o que el glicerol presente en el Buffer de Replegamiento generó un ruido que no permitió realizar el análisis adecuadamente.



## 9. CONCLUSIÓN

- El análisis *in silico* de las proteínas TcCyP19, TcCyP22 y TcCySEC sugiere que son PPIasas que se pueden expresar en forma soluble en *E. coli*, lo que sugiere que los servidores utilizados son una estrategia recomendable previo a iniciar la etapa experimental. Este análisis *in silico* de las proteínas TcCyP19, TcCyP22 y TcCySEC indica que su secuencia contiene el dominio de PPIasa.
- La expresión de rTcCyP19 y rTcCySEC fue consistente con los resultados obtenidos *in silico* a excepción de la rTcCyP22 que en las condiciones ensayadas (temperatura, medio de cultivo o volumen de producción) se expresó en forma insoluble.
- La producción de las proteínas rTcCyP19 y rTcCySEC utilizando la plataforma de expresión BL21 (DE3) de *E. coli* permitió la expresión con la concentración y pureza suficientes para confirmar su actividad de PPIasa.
- Los ensayos de actividad de PPIasa indican que la eficiencia catalítica de la TcCyP19 es superior a la de la TcCySEC.

## 10. PERSPECTIVAS

- Modificar las condiciones de expresión que mejoren la solubilidad de la rTcCyP22, como variación de temperatura, tiempo de inducción, medios de cultivo, cepas de expresión e incluso amortiguadores de lisis a diferentes pH.
- Realizar el análisis por DLS con muestras recién purificadas y con una concentración mayor de PPIasa, utilizando un amortiguador que no contenga glicerol.
- Realizar ensayos de actividad con la rTcCyP19 y rTcCySEC a diferentes concentraciones de estas proteínas recombinantes, con el propósito de validar estadísticamente el valor de su eficiencia catalítica.
- Evaluar el potencial de las PPIasas rTcCyP19 y rTcCySEC para el replegamiento asistido de rTSA-1.
- Evaluar la eficiencia de producir la rTSA-1 en su conformación nativa mediante la co-expresión de este antígeno en su conformación nativa junto con la PPIasa con mayor eficiencia catalítica en los ensayos *in vitro*.

## 11. REFERENCIAS

- Antonio-Pérez, A., Aldaz-Martínez, L. M., Meneses-Acosta, A., & Ortega-López, J. (2014). Refolding of lysozyme assisted by molecular chaperones immobilized in cellulose: the operational conditions that affect refolding yields. *Revista mexicana de ingeniería química*, *13*, 83-91.  
[http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S1665-27382014000100007&nrm=iso](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1665-27382014000100007&nrm=iso)
- Antonio-Perez, A., Clement-Carretero, H., Alagon, A., & Ortega-López, J. (2010). Chromatographic refolding of recombinant hyaluronidase assisted by apical domain of Groel, Dsba and DsbC immobilized in cellulose. *Protein Science*, *19*(255), 0961-8368.
- Antonio-Perez, A., & Ortega-Lopez, J. (2010). GroEL Apical Domain, DsbA and DsbC immobilized in cellulose assisted the chromatographic oxidative refolding of lysozyme. *Journal of Biotechnology*, *150*, 85-86.  
<https://doi.org/https://doi.org/10.1016/j.jbiotec.2010.08.220>
- Antonio-Pérez, A., Ramón-Luing, L. A., & Ortega-López, J. (2012a). Chromatographic refolding of rhodanese and lysozyme assisted by the GroEL apical domain, DsbA and DsbC immobilized on cellulose. *J Chromatogr A*, *1248*, 122-129.  
<https://doi.org/10.1016/j.chroma.2012.05.086>
- Antonio-Pérez, A., Rivera-Hernández, T., Aldaz-Martínez, L., & Ortega-López, J. (2012b). Oxidative refolding of lysozyme assisted by DsbA, DsbC and the GroEL apical domain immobilized in cellulose. *Biotechnology and Bioprocess Engineering*, *17*(4), 703-710. <https://doi.org/10.1007/s12257-011-0663-9>
- Bell, A., Monaghan, P., & Page, A. P. (2006). Peptidyl-prolyl cis-trans isomerases (immunophilins) and their roles in parasite biochemistry, host-parasite interaction and antiparasitic drug action. *Int J Parasitol*, *36*(3), 261-276.  
<https://doi.org/10.1016/j.ijpara.2005.11.003>
- 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. *J Pharm Sci*, *107*(5), 1468-1473.  
<https://doi.org/10.1016/j.xphs.2017.12.014>
- Bose, S., Mucke, M., & Freedman, R. B. (1994). The characterization of a cyclophilin-type peptidyl prolyl cis-trans-isomerase from the endoplasmic-reticulum lumen. *Biochem J*, *300* ( Pt 3), 871-875. <https://doi.org/10.1042/bj3000871>
- Bowie, J. U., Luthy, R., & Eisenberg, D. (1991). A method to identify protein sequences that fold into a known three-dimensional structure. *Science*, *253*(5016), 164-170.  
<https://doi.org/10.1126/science.1853201>
- Bua, J., Aslund, L., Pereyra, N., Garcia, G. A., Bontempi, E. J., & Ruiz, A. M. (2001). Characterisation of a cyclophilin isoform in Trypanosoma cruzi. *FEMS Microbiol Lett*, *200*(1), 43-47. <https://doi.org/10.1111/j.1574-6968.2001.tb10690.x>

- Bua, J., Ruiz, A. M., Potenza, M., & Fichera, L. E. (2004). In vitro anti-parasitic activity of Cyclosporin A analogs on *Trypanosoma cruzi*. *Bioorg Med Chem Lett*, *14*(18), 4633-4637. <https://doi.org/10.1016/j.bmcl.2004.07.003>
- Chow, M. K., Amin, A. A., Fulton, K. F., Fernando, T., Kamau, L., Batty, C., Louca, M., Ho, S., Whisstock, J. C., Bottomley, S. P., & Buckle, A. M. (2006a). The REFOLD database: a tool for the optimization of protein expression and refolding. *Nucleic Acids Res*, *34*(Database issue), D207-212. <https://doi.org/10.1093/nar/gkj080>
- Chow, M. K., Amin, A. A., Fulton, K. F., Whisstock, J. C., Buckle, A. M., & Bottomley, S. P. (2006b). REFOLD: an analytical database of protein refolding methods. *Protein Expr Purif*, *46*(1), 166-171. <https://doi.org/10.1016/j.pep.2005.07.022>
- Colovos, C., & Yeates, T. O. (1993). Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci*, *2*(9), 1511-1519. <https://doi.org/10.1002/pro.5560020916>
- Costantini, S., Colonna, G., & Facchiano, A. M. (2006). Amino acid propensities for secondary structures are influenced by the protein structural class. *Biochem Biophys Res Commun*, *342*(2), 441-451. <https://doi.org/10.1016/j.bbrc.2006.01.159>
- de Groot, N. S., Castillo, V., Grana-Montes, R., & Ventura, S. (2012). AGGRESKAN: method, application, and perspectives for drug design. *Methods Mol Biol*, *819*, 199-220. [https://doi.org/10.1007/978-1-61779-465-0\\_14](https://doi.org/10.1007/978-1-61779-465-0_14)
- Dimou, M., Venieraki, A., & Katinakis, P. (2017). Microbial cyclophilins: specialized functions in virulence and beyond. *World J Microbiol Biotechnol*, *33*(9), 164. <https://doi.org/10.1007/s11274-017-2330-6>
- 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 Rev Vaccines*, *11*(9), 1043-1055. <https://doi.org/10.1586/erv.12.85>
- Dumonteil, E., Herrera, C., & Marx, P. A. (2022). Safety and preservation of cardiac function following therapeutic vaccination against *Trypanosoma cruzi* in rhesus macaques. *J Microbiol Immunol Infect*. <https://doi.org/10.1016/j.jmii.2022.09.003>
- Dumonteil, E., Herrera, C., Tu, W., Goff, K., Fahlberg, M., Haupt, E., Kaur, A., Marx, P. A., Ortega-Lopez, J., Hotez, P. J., & Bottazzi, M. E. (2020). Safety and immunogenicity of a recombinant vaccine against *Trypanosoma cruzi* in Rhesus macaques. *Vaccine*, *38*(29), 4584-4591. <https://doi.org/10.1016/j.vaccine.2020.05.010>
- Dzul-Huchim, V. M., Ramirez-Sierra, M. J., Martinez-Vega, P. P., Rosado-Vallado, M. E., Arana-Argaez, V. E., Ortega-Lopez, J., Gusovsky, F., Dumonteil, E., Cruz-Chan, J. V., Hotez, P., Bottazzi, M. E., & Villanueva-Lizama, L. E. (2022). Vaccine-linked chemotherapy with a low dose of benznidazole plus a bivalent recombinant protein vaccine prevents the development of cardiac fibrosis caused by *Trypanosoma cruzi* in chronically-infected BALB/c mice. *PLoS Negl Trop Dis*, *16*(9), e0010258. <https://doi.org/10.1371/journal.pntd.0010258>

- Fanghanel, J., & Fischer, G. (2004). Insights into the catalytic mechanism of peptidyl prolyl cis/trans isomerases. *Front Biosci*, 9, 3453-3478. <https://doi.org/10.2741/1494>
- Fischer, G. (1994). Peptidyl-Prolyl cis/trans Isomerases and Their Effectors. *Angewandte Chemie International Edition in English*, 33(14), 1415-1436. <https://doi.org/https://doi.org/10.1002/anie.199414151>
- Fischer, G., Wittmann-Liebold, B., Lang, K., Kiefhaber, T., & Schmid, F. X. (1989). Cyclophilin and peptidyl-prolyl cis-trans isomerase are probably identical proteins. *Nature*, 337(6206), 476-478. <https://doi.org/10.1038/337476a0>
- Gasteiger E., H. C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. (2005). The Proteomics Protocols Handbook. In J. M. Walker (Ed.), *Protein Identification and Analysis Tools on the ExPASy Server* (1 ed., pp. 571-607). Humana Press. <https://doi.org/https://doi.org/10.1385/1592598900>
- GE-Healthcare. (2007). PD-10 Desalting Columns - Instructions 52-1308-00 BB. In G. H. B.-S. AB (Ed.), (pp. 7): GE Healthcare Bio-Sciences AB.
- GenScript. (2021). *pUC57 plasmid DNA Vector - GenScript*. Retrieved 12/Apr/2021 from [https://www.genscript.com/vector/SD1176-pUC57\\_plasmid\\_DNA.html](https://www.genscript.com/vector/SD1176-pUC57_plasmid_DNA.html)
- Goulas, T., Cuppari, A., Garcia-Castellanos, R., Snipas, S., Glockshuber, R., Arolas, J. L., & Gomis-Ruth, 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>
- Green, M. R., & Sambrook, J. (2021). One-Step Preparation of Competent E. coli: Transformation and Storage of Bacterial Cells in the Same Solution. *Cold Spring Harb Protoc*, 2021(11). <https://doi.org/10.1101/pdb.prot101212>
- Jordens, J., Janssens, V., Longin, S., Stevens, I., Martens, E., Bultynck, G., Engelborghs, Y., Lescrier, E., Waelkens, E., Goris, J., & Van Hoof, C. (2006). The protein phosphatase 2A phosphatase activator is a novel peptidyl-prolyl cis/trans-isomerase. *J Biol Chem*, 281(10), 6349-6357. <https://doi.org/10.1074/jbc.M507760200>
- Kofron, J. L., Kuzmic, P., Kishore, V., Colon-Bonilla, E., & Rich, D. H. (1991). Determination of kinetic constants for peptidyl prolyl cis-trans isomerases by an improved spectrophotometric assay. *Biochemistry*, 30(25), 6127-6134. <https://doi.org/10.1021/bi00239a007>
- Kuriata, A., Iglesias, V., Pujols, J., Kurcinski, M., Kmiecik, S., & Ventura, S. (2019). Aggrescan3D (A3D) 2.0: prediction and engineering of protein solubility. *Nucleic Acids Res*, 47(W1), W300-W307. <https://doi.org/10.1093/nar/gkz321>
- Kyte, J., & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J Mol Biol*, 157(1), 105-132. [https://doi.org/10.1016/0022-2836\(82\)90515-0](https://doi.org/10.1016/0022-2836(82)90515-0)
- Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, 26(2), 283-291. <https://doi.org/https://doi.org/10.1107/S0021889892009944>

- Linding, R., Russell, R. B., Neduva, V., & Gibson, T. J. (2003). GlobPlot: Exploring protein sequences for globularity and disorder. *Nucleic Acids Res*, *31*(13), 3701-3708. <https://doi.org/10.1093/nar/gkg519>
- Lu, K. P., Finn, G., Lee, T. H., & Nicholson, L. K. (2007). Prolyl cis-trans isomerization as a molecular timer. *Nat Chem Biol*, *3*(10), 619-629. <https://doi.org/10.1038/nchembio.2007.35>
- Luthy, R., Bowie, J. U., & Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. *Nature*, *356*(6364), 83-85. <https://doi.org/10.1038/356083a0>
- MACHEREY-NAGEL. (2021). User Manual - Plasmid DNA purification (NucleoBond® Xtra Midi / Maxi). In MACHEREY-NAGEL (Ed.), *Mn-net.com* (Rev. 16 ed., Vol. Rev. 16, pp. 1-47). Web site: MACHEREY-NAGEL.
- Moro, A., Ruiz-Cabello, F., Fernandez-Cano, A., Stock, R. P., & Gonzalez, A. (1995). Secretion by *Trypanosoma cruzi* of a peptidyl-prolyl cis-trans isomerase involved in cell infection. *EMBO J*, *14*(11), 2483-2490. <https://www.ncbi.nlm.nih.gov/pubmed/7540135>
- Morris, A. L., MacArthur, M. W., Hutchinson, E. G., & Thornton, J. M. (1992). Stereochemical quality of protein structure coordinates. *Proteins*, *12*(4), 345-364. <https://doi.org/10.1002/prot.340120407>
- Pereira, P. J., Vega, M. C., Gonzalez-Rey, E., Fernandez-Carazo, R., Macedo-Ribeiro, S., Gomis-Ruth, F. X., Gonzalez, A., & Coll, M. (2002). *Trypanosoma cruzi* macrophage infectivity potentiator has a rotamase core and a highly exposed alpha-helix. *EMBO Rep*, *3*(1), 88-94. <https://doi.org/10.1093/embo-reports/kvf009>
- Perrone, A. E., Milduburger, N., Fuchs, A. G., Bustos, P. L., & Bua, J. (2018). A Functional Analysis of the Cyclophilin Repertoire in the Protozoan Parasite *Trypanosoma Cruzi*. *Biomolecules*, *8*(4). <https://doi.org/10.3390/biom8040132>
- Pontius, J., Richelle, J., & Wodak, S. J. (1996). Deviations from standard atomic volumes as a quality measure for protein crystal structures. *J Mol Biol*, *264*(1), 121-136. <https://doi.org/10.1006/jmbi.1996.0628>
- Potenza, M., Galat, A., Minning, T. A., Ruiz, A. M., Duran, R., Tarleton, R. L., Marin, M., Fichera, L. E., & Bua, J. (2006). Analysis of the *Trypanosoma cruzi* cyclophilin gene family and identification of Cyclosporin A binding proteins. *Parasitology*, *132*(Pt 6), 867-882. <https://doi.org/10.1017/S0031182005009558>
- QIAGEN. (2020a). QIAprep® Miniprep Handbook - Protocol: Plasmid DNA Purification using the QIAprep Spin Miniprep Kit and a Microcentrifuge. In *Qiagen.com*. Qiagen.com: QIAGEN.
- QIAGEN. (2020b). QIAquick® Spin Handbook - Protocol: QIAquick PCR Purification using a Microcentrifuge. In *Qiagen.com* (pp. 1-48). Qiagen.com: QIAGEN.
- Quijano-Hernandez, I. A., Castro-Barcena, A., Vazquez-Chagoyan, J. C., Bolio-Gonzalez, M. E., Ortega-Lopez, J., & Dumonteil, E. (2013). Preventive and therapeutic DNA vaccination partially protect dogs against an infectious challenge with *Trypanosoma cruzi*. *Vaccine*, *31*(18), 2246-2252. <https://doi.org/10.1016/j.vaccine.2013.03.005>

- Ramon-Luing, L. A., Cruz-Migoni, A., Ruiz-Medrano, R., Xoconostle-Cazares, B., & Ortega-Lopez, J. (2006). One-step purification and immobilization in cellulose of the GroEL apical domain fused to a carbohydrate-binding module and its use in protein refolding. *Biotechnol Lett*, 28(5), 301-307. <https://doi.org/10.1007/s10529-005-5714-x>
- Ramon-Luing, L. A., Rendon-Gandarilla, F. J., Cardenas-Guerra, R. E., Rodriguez-Cabrera, N. A., Ortega-Lopez, J., Avila-Gonzalez, L., Angel-Ortiz, C., Herrera-Sanchez, C. N., Mendoza-Garcia, M., & Arroyo, R. (2010). Immunoproteomics of the active degradome to identify biomarkers for *Trichomonas vaginalis*. *Proteomics*, 10(3), 435-444. <https://doi.org/10.1002/pmic.200900479>
- Rego, J. V., Duarte, A. P., Liarte, D. B., de Carvalho Sousa, F., Barreto, H. M., Bua, J., Romanha, A. J., Radis-Baptista, G., & Murta, S. M. (2015). Molecular characterization of Cyclophilin (TcCyP19) in *Trypanosoma cruzi* populations susceptible and resistant to benznidazole. *Exp Parasitol*, 148, 73-80. <https://doi.org/10.1016/j.exppara.2014.11.007>
- Rendon-Gandarilla, F. J., Ramon-Luing Lde, L., Ortega-Lopez, J., Rosa de Andrade, I., Benchimol, M., & Arroyo, R. (2013). The TvLEGU-1, a legumain-like cysteine proteinase, plays a key role in *Trichomonas vaginalis* cytoadherence. *Biomed Res Int*, 2013, 561979. <https://doi.org/10.1155/2013/561979>
- Resendiz-Cardiel, G., Arroyo, R., & Ortega-Lopez, J. (2017). Expression of the enzymatically active legumain-like cysteine proteinase TvLEGU-1 of *Trichomonas vaginalis* in *Pichia pastoris*. *Protein Expr Purif*, 134, 104-113. <https://doi.org/10.1016/j.pep.2017.04.007>
- Sanchez-Garcia, L., Martin, L., Mangues, R., Ferrer-Miralles, N., Vazquez, E., & Villaverde, A. (2016). Recombinant pharmaceuticals from microbial cells: a 2015 update. *Microb Cell Fact*, 15, 33. <https://doi.org/10.1186/s12934-016-0437-3>
- Schiene, C., & Fischer, G. (2000). Enzymes that catalyse the restructuring of proteins. *Curr Opin Struct Biol*, 10(1), 40-45. [https://doi.org/10.1016/s0959-440x\(99\)00046-9](https://doi.org/10.1016/s0959-440x(99)00046-9)
- Schiene-Fischer, C., & Yu, C. (2001). Receptor accessory folding helper enzymes: the functional role of peptidyl prolyl cis/trans isomerases. *FEBS Lett*, 495(1-2), 1-6. [https://doi.org/10.1016/s0014-5793\(01\)02326-2](https://doi.org/10.1016/s0014-5793(01)02326-2)
- Selles, B., Dhalleine, T., Boutilliat, A., Rouhier, N., & Couturier, J. (2021). A Redox-Sensitive Cysteine Is Required for PIN1At Function. *Front Plant Sci*, 12, 735423. <https://doi.org/10.3389/fpls.2021.735423>
- Shuid, A. N., Kempster, R., & McGuffin, L. J. (2017). ReFOLD: a server for the refinement of 3D protein models guided by accurate quality estimates. *Nucleic Acids Res*, 45(W1), W422-w428. <https://doi.org/10.1093/nar/gkx249>
- Smith, D. K., Radivojac, P., Obradovic, Z., Dunker, A. K., & Zhu, G. (2003). Improved amino acid flexibility parameters. *Protein Sci*, 12(5), 1060-1072. <https://doi.org/10.1110/ps.0236203>
- SnapGene. (2021). *pUC57 Sequence and Map - SnapGene*. Retrieved 12/Apr/2021 from [https://www.snapgene.com/resources/plasmid-files/?set=basic\\_cloning\\_vectors&plasmid=pUC57](https://www.snapgene.com/resources/plasmid-files/?set=basic_cloning_vectors&plasmid=pUC57)

- Sormanni, P., Aprile, F. A., & Vendruscolo, M. (2015). The CamSol method of rational design of protein mutants with enhanced solubility. *J Mol Biol*, *427*(2), 478-490.  
<https://doi.org/10.1016/j.jmb.2014.09.026>
- Usmani, S. S., Bedi, G., Samuel, J. S., Singh, S., Kalra, S., Kumar, P., Ahuja, A. A., Sharma, M., Gautam, A., & Raghava, G. P. S. (2017). THPdb: Database of FDA-approved peptide and protein therapeutics. *PLoS One*, *12*(7), e0181748.  
<https://doi.org/10.1371/journal.pone.0181748>
- 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 Negl Trop Dis*, *12*(1), e0006240. <https://doi.org/10.1371/journal.pntd.0006240>
- Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. *Nat Methods*, *12*(1), 7-8.  
<https://doi.org/10.1038/nmeth.3213>
- Zamudio-Prieto, O., Benitez-Cardoza, C., Arroyo, R., & Ortega-Lopez, J. (2014). Conformational changes induced by detergents during the refolding of chemically denatured cysteine protease ppEhCP-B9 from Entamoeba histolytica. *Biochim Biophys Acta*, *1844*(7), 1299-1306. <https://doi.org/10.1016/j.bbapap.2014.04.009>
- Zamudio-Prieto, O., Benitez-Cardoza, C., & Ortega-Lopez, J. (2013). Effect of Artificial Chaperones in Entamoeba Histolytica Cysteine Protease Secondary Structure. *Protein Science*, *22*(146), 0961-8368.
- Zemanova, L., Vaskova, M., Schmidt, M., Roubalova, J., Haleckova, A., Benek, O., & Musilek, K. (2020). RNase T1 Refolding Assay for Determining Mitochondrial Cyclophilin D Activity: A Novel In Vitro Method Applicable in Drug Research and Discovery. *Biochemistry*, *59*(17), 1680-1687.  
<https://doi.org/10.1021/acs.biochem.9b01025>



## 12. APÉNDICES

### 12.1. Apéndice 1

Created with SnapGene®

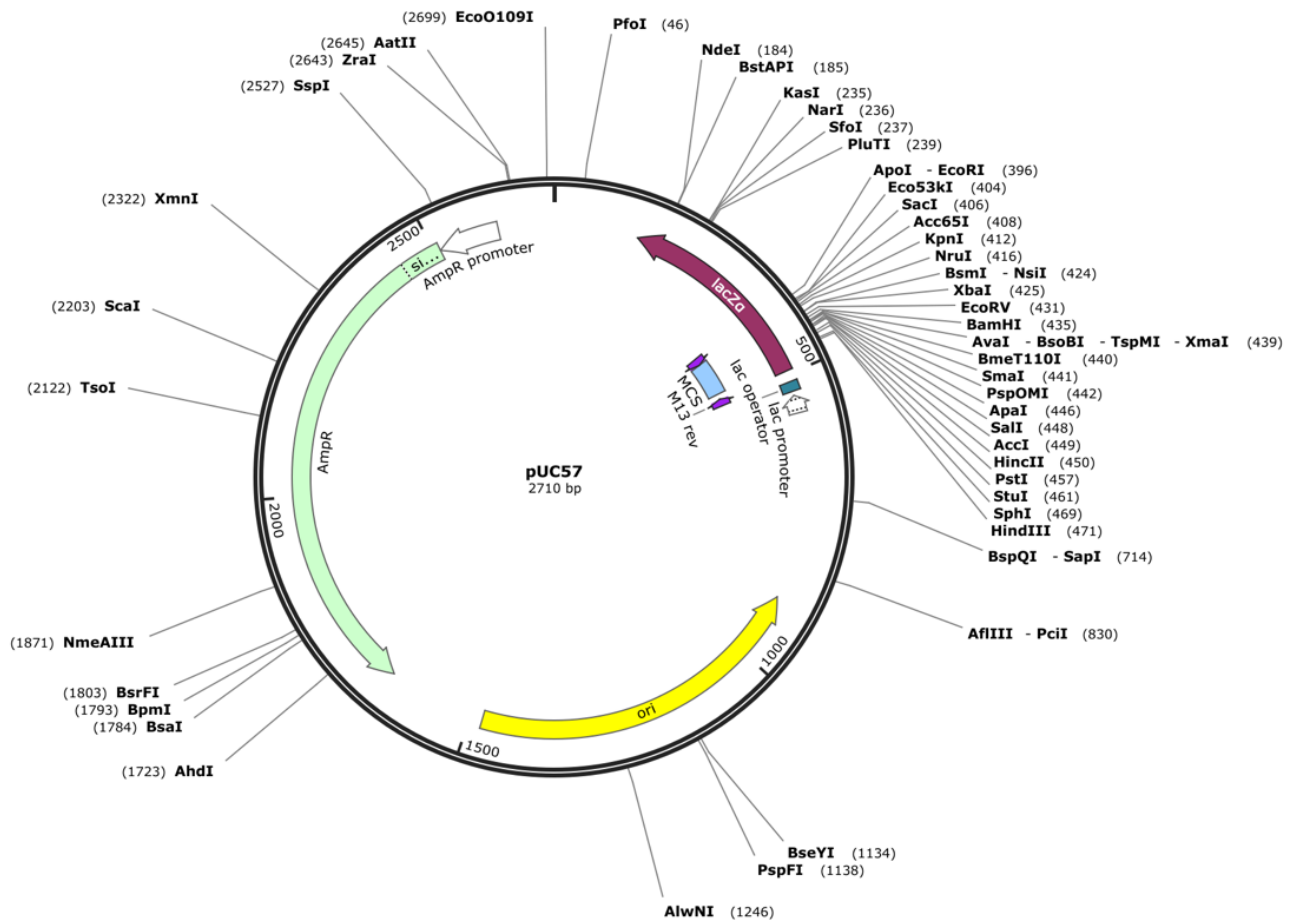


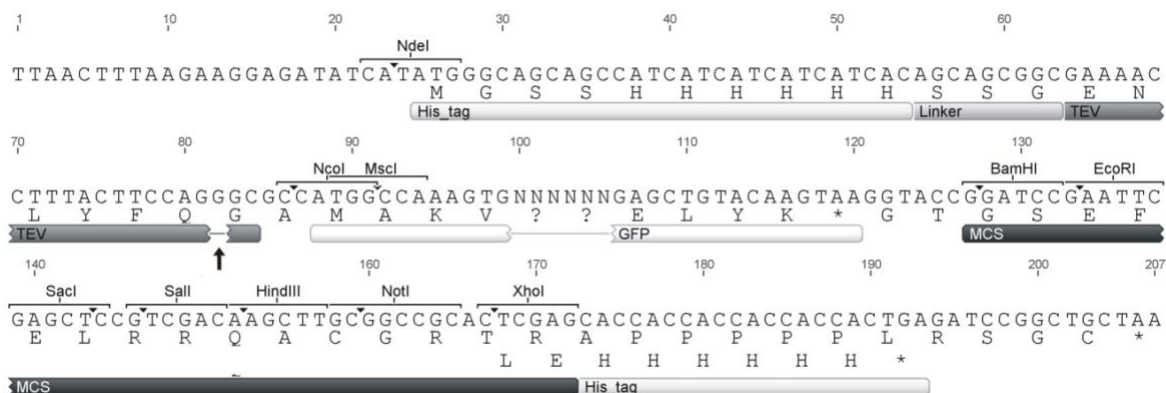
Figura 24. Mapa del Vector de Clonación pU57. (GenScript, 2021; SnapGene, 2021)

## 12.2. Apéndice 2

A	ccatg gcc agt tat aaa ccg cat cac gca acc gtt ccg acc aat ccg aaa gtc ttc ttc gac
	M A S Y K P H H A T V P T N P K V F F D
	gtc agc att ggc ggt caa agc gcg ggc cgc gtt gtt ttt gaa ctg ttt gca gac gcc gtt
	V S I G G Q S A G R V V F E L F A D A V
	ccg aaa acc gca gaa aat ttt cgc gca ctg tgc acc ggc gag aaa aac ttt ggc tac gca
	P K T A E N F R A L C T G E K N F G Y A
	ggt agc ggc ttt cac cgc att att ccg cag ttc atg tgt cag ggc gat ttt acc aac
	G S G F H R I I P Q F M C Q G G D F T N
	cat aac ggc acc ggc ggt cgt agt att tac ggc gaa aaa ttc gcg gac gaa agt ttt gca
	H N G T G G R S I Y G E K F A D E S F A
	ggc aaa gca ggt aaa cac ttt ggt ctg ggc acc ctg agt atg gca aac gca ggt ccg aat
	G K A G K H F G L G T L S M A N A G P N
	acc aac ggt agc cag ttt ttc att tgc acc gct ccg acc cag tgg ctg gac ggt aaa cac
	T N G S Q F F I C T A P T Q W L D G K H
gtt gta ttt ggc cag gtt ctg gaa ggc atc gaa gtt gtc aaa gcg atg gaa gcg gtt ggt	
V V F G Q V L E G I E V V K A M E A V G	
agt caa acc ggc aaa acc agc aaa ccg gtc aaa atc gaa gca tct ggt cag ctg <b>ctc gag</b>	
S Q T G K T S K P V K I E A S G Q L L E	
B	ccatg gcc ttt agt cgt acc tgg ttt tgg gca cag cgt aaa ctg ccg ttc tat ccg atc aac
	M A F S R T W F W A Q R K L P F Y P I N
	ccg aaa aac ccg ctg gtc ttc ttt gaa att agc att ggc gca caa ccg gca ggt cgc gtt
	P K N P L V F F E I S I G A Q P A G R V
	gaa atg gaa ctg ttc aaa gac gcg gtc ccg aaa acc gca gaa aat ttt cgc gca ctg tgt
	E M E L F K D A V P K T A E N F R A L C
	acc ggc gaa aaa ggc gta ggt cgt tct ggt aaa gcg ctg tgc tac aaa ggc agc aaa ttc
	T G E K G V G R S G K A L C Y K G S K F
	cac cgc gtt att ccg caa ttc atg tgt cag ggc gat ttt acc aac ggt aac ggc acc
	H R V I P Q F M C Q G G D F T N G N G T
	ggc ggc gaa tct atc tac ggc atg aaa ttc ccg gac gaa agt ttt gca ggt cgc gca ggt
	G G E S I Y G M K F P D E S F A G R A G
	aaa cat ttt ggt ccg ggc acc ctg agt atg gca aac gca ggt ccg aat acc aac ggt agt
	K H F G P G T L S M A N A G P N T N G S
cag ttt ttc att tgt acc gct gct acc gag tgg ctg gac ggc aaa cac gtt gtt ttt ggc	
Q F F I C T A A T E W L D G K H V V F G	
cag atc acg aaa ggc tac gag gtc atc gag aaa gtc gaa gca aac ggt agt cgt tct ggc	
Q I T K G Y E V I E K V E A N G S R S G	
gca acc cgt caa ccg att ctg att acc gat tgc ggt gaa gtc aaa aac aac <b>ctc gag</b>	
A T R Q P I L I T D C G E V K N N L E	
C	ccatg gtg agt ggt gac gca gca agt cac gaa gaa cgt atg aac aac tac cgt aaa cgc gtt
	M V S G D A A S H E E R M N Y R K R V
	ggt cgc ctg ttt atg gaa cag aaa gct gct cag ccg gat gca gtt aaa ctg cgc tct ggt
	G R L F M E Q K A A Q P D A V K L P S G
	ctg gtt ttt cag cgc att gca cgc ggt tct ggt aaa cgc gca ccg gct att gac gac aaa
	L V F Q R I A R G S G K R A P A I D D K
	tgc gaa gtc cat tat acc ggt cgt ctg cgc gac ggc acc gtt ttt gat agt agt cgc gaa
	C E V H Y T G R L R D G T V F D S S R E
	cgc ggt aaa ccg acc acc ttt cgt ccg aac gaa gtc atc aaa ggt tgg acc gaa gct ctg
	R G K P T T F R P N E V I K G W T E A L
	caa ctg atg cgc gaa ggc gat cgt tgg cgt ctg ttt att ccg tac gat ctg gca tac ggc
	Q L M R E G D R W R L F I P Y D L A Y G
	gtt acc ggc ggt ggc ggt atg att ccg ccg tat agt ccg ctg gaa ttt gac gtt gaa ctg
	V T G G G G M I P P Y S P L E F D V E L
atc agc atc aaa gac ggc ggt aaa ggt cgt acc gca gaa gaa gtc gac gaa att ctg cgc	
I S I K D G G K G R T A E E V D E I L R	
aaa gcg gaa gaa gat cgc gaa gac atg <b>ctc gag</b>	
K A E E D R E D M L E	

**Figura 25. Secuencias de genes y aminoácidos.** A) gen *tcyp19* (542 pb) que codifica para la ciclofilina A de *T. cruzi*, cepa CL Brener (RefSeq XP\_821578). Región optimizada para el uso de codones de *E. coli*: 9 pb-536 pb. B) gen *tcyp22* (599 pb) que codifica para la ciclofilina de 21.3 kDa de *T. cruzi*, cepa CL Brener (Genbank AAQ55215). Región optimizada para el uso de codones de *E. coli*: 9 pb-593 pb. C) gen *tcysec* (515 pb) que codifica para la PPIasa de *T. cruzi* presente en TcMIP, cepa TCC (Genbank PWV17357). Región optimizada para el uso de codones de *E. coli*: 9 pb-509 pb. Sitios de restricción NcoI (5'→3', **ccatgg**) y XhoI (3'→5', **ctcgag**) añadidos en los extremos de la secuencia. Dominio de PPIasa remarcado en **verde**. Dominio de MIP señalado en letras **rojas**.

### 12.3. Apéndice 3



**Figura 26. Mapa del Vector de Expresión pCri8a.** Secuencia de nucleótidos y traducción del sitio de clonación del vector. Parte de la secuencia de GFP que se omite del mapa vectorial está marcada con los símbolos (?) Y (N). Una flecha negra indica el sitio de escisión de la proteinasa TEV. Pesos moleculares: His<sub>6</sub>=830.9 Da, Linker= 286.9, TEV=978 Da, Total de etiquetas de fusión =2,926.7 Da. Los sitios de restricción que se muestran se encuentran una vez (Goulas et al., 2014).

### 12.4. Apéndice 4

**Tabla 16. Condiciones de Reacción de Doble Digestión**

Reactivo	pCri8a (μL)	pUC57-tccyp19 (μL)	pUC57-tccyp22 (μL)	pUC57-tccysec (μL)
ADNp <sup>a</sup>	39.0	5.0	3.5	5.0
NcoI	0.5	0.5	0.5	0.5
XhoI	0.5	0.5	0.5	0.5
rSmartCut Buffer(NEB)	5.0	5.0	5.0	5.0
H <sub>2</sub> O grado inyectable	5.0	39.0	40.5	39.0
<b>Volumen Total</b>	<b>50</b>	<b>50</b>	<b>50</b>	<b>50</b>

<sup>a</sup> Volumen de ADNp obtenido del midiprep y equivalente a 3 μg; <sup>b</sup> volumen de H<sub>2</sub>O necesario para completar el volumen total de la reacción (50 μL). Todas las reacciones se realizaron a 37°C, 250 rpm durante 16 h utilizando el equipo *Thermo Mixer (Thermo Scientific)*.

### 12.5. Apéndice 5

**Tabla 17. Condiciones de Reacción de Ligación y Controles**

Reactivo	Problema tccyp19 (μL)	Problema tccyp22 (μL)	Problema tccysec (μL)	Control Religación (μL)	Control ADNp (μL)
Inserto	6.0	4.0	4.5	-	-
Vector (pCri8a)	3.0	3.0	3.0	3.0	3.0
Buffer Ligasa	2.0	2.0	2.0	2.0	2.0
Ligasa T4	1.0	1.0	1.0	1.0	-
H <sub>2</sub> O grado inyectable	8.0	10.0	10.5	12.8	14.3
<b>Volumen Total</b>	<b>20</b>	<b>20</b>	<b>20</b>	<b>20</b>	<b>20</b>

## 12.6. Apéndice 6.

Fundamento, variables e interpretación de los análisis de validación de los modelos 3D:

**ERRAT** permite diferenciar entre las regiones determinadas correcta e incorrectamente a partir de los modelos 3D de una proteína, basado en la interacción atómica característica. Los átomos se distribuyen de forma no aleatoria entre sí en las proteínas, por lo cual los errores en la construcción del modelo llevan a distribuciones aleatorias de cada átomo, permitiendo distinguir las distribuciones correctas utilizando métodos estadísticos. Lo anterior se logra clasificando los átomos en tres categorías: carbono (C), nitrógeno (N) y oxígeno (O). Existiendo seis combinaciones posibles por pares a través de un enlace no covalente (CC, CN, CO, NN, NO y OO). Este método utiliza una función de error cuadrática para caracterizar el conjunto de interacciones por pares a partir de ventanas deslizantes de nueve residuos en una base de datos de 96 estructuras proteicas conocidas. Finalmente, las regiones de las estructuras de la proteína a analizar que están mal trazadas o mal registradas pueden identificarse mediante el análisis del patrón de interacciones no enlazadas de cada ventana (Colovos & Yeates, 1993).

**VERIFY 3D** se basa en los entornos de la estructura 3D descritos mediante: (i) el área del residuo enterrado en la proteína e inaccesible al disolvente; (ii) la fracción del área de la cadena lateral que está cubierta por átomos polares (O y N); y (iii) la estructura secundaria local. Este método detecta la similitud estructural de actinas y de proteínas con choque térmico de 70 kDa, a pesar de que esta familia de proteínas no comparten similitud con la secuencia a analizar (Bowie et al., 1991), comparando el modelo con su propia secuencia de aminoácidos, con su modelo 3D, calculado a partir de las coordenadas atómicas de los perfiles 3D de estructuras proteicas cristalizadas que coinciden con la secuencia a analizar, otorgando puntuaciones altas (máximo 100). En contraste, los perfiles 3D de modelos de proteínas que se sabe que son erróneos obtienen una puntuación baja. Lo anterior

examinando la puntuación promedio (*Average Score*) de los residuos que constituyen al perfil 3D y evaluando si están dentro la ventana móvil (Luthy et al., 1992).

**PROVE** evalúa los rangos estándar de los volúmenes atómicos y de residuos calculados en 64 estructuras cristalinas de proteínas altamente resueltas y bien refinadas utilizando el procedimiento clásico de Voronoi. Las desviaciones de los volúmenes atómicos respecto a los valores estándar, evaluadas como Z-score del volumen, se utilizan para determinar la calidad de modelos 3D. La puntuación global de la estructura, se calcula con la desviación cuadrática media del Z-score del volumen (*Z-score rms*), que mide la magnitud media de las irregularidades del volumen en la estructura. Interpretando que el *Z-score rms* disminuye a medida que la resolución y el Factor-R mejoran. A partir de la distribución de *Z-score rms* en estructuras con una resolución o Factor-R dados, se determinan los límites normales del *Z-score rms*, observando que estructuras con *Z-score rms* mayor a estos límites se consideran atípicas, mostrando una estereoquímica inusual. *Z-score* absolutos de átomos individuales se utilizan para identificar problemas en regiones específicas dentro de un modelo 3D atípico. Estos *Z-score* se correlacionan con Factores-B atómicos, y los átomos que tienen *Z-score* absolutos  $>3$ , se encuentran en regiones del modelo donde programas como PROCHECK identifican una estereoquímica inusual (Pontius et al., 1996).

**PROCHECK** utiliza las coordenadas de los átomos del esqueleto peptídico para calcular los ángulos de torsión Psi ( $\psi$ ) y Phi ( $\varphi$ ) de la proteína en cuestión. Los datos obtenidos permiten la construcción un Gráfico de Ramachandran, en el que se representa  $\psi$  en el eje de las ordenadas y  $\varphi$  en el eje de las abscisas, visualizando así las regiones energéticamente permitidas para los ángulos diedros del esqueleto de la proteína analizada. Como resultado, es posible determinar cuántos residuos se forman un elemento de estructura secundaria determinado, y cuántos aminoácidos se ven forzados en sus ángulos de torsión basado en la estructura seleccionada en el PDB, asó se permite sugerir que cuantos más aminoácidos presenten ángulos forzados, mayor será la probabilidad de

que la estructura de la proteína analizada no sea la más adecuada, al menos desde análisis *in silico* (Laskowski et al., 1993; Morris et al., 1992).

### **12.7. Apéndice 7**

Fundamento teórico, variables consideradas y escala de ponderación de cada servidor utilizado en la determinación de solubilidad y agregación:

El servidor de la Universidad de Manchester, ProteinSol, es un predictor que analiza la secuencia de aminoácidos y determina la solubilidad basado en un algoritmo que considera 35 características de la 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 de las láminas  $\beta$  (Costantini et al., 2006). Las correlaciones de estos parámetros se ponderan como valores de solubilidad en un rango de 0 a 1.

El servidor Soluble Domain for Protein Expression (SoDoPE) utiliza 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 Solubility-Weighted Index (SWI) otorgando valores en un rango de 0 a 1.

El método CamSol de Vendruscolo Lab de la Universidad de Cambridge, utiliza la secuencia proteica para calcular el perfil de solubilidad, pudiendo realizarlo en un contexto estructuralmente corregido al integrar la información 3D de las estructuras proteicas. Este método predice las regiones de baja solubilidad en la superficie de las proteínas en una escala de -1.5 a 1.5 tomando en cuenta que valores mayores a 1 (azul) corresponden a

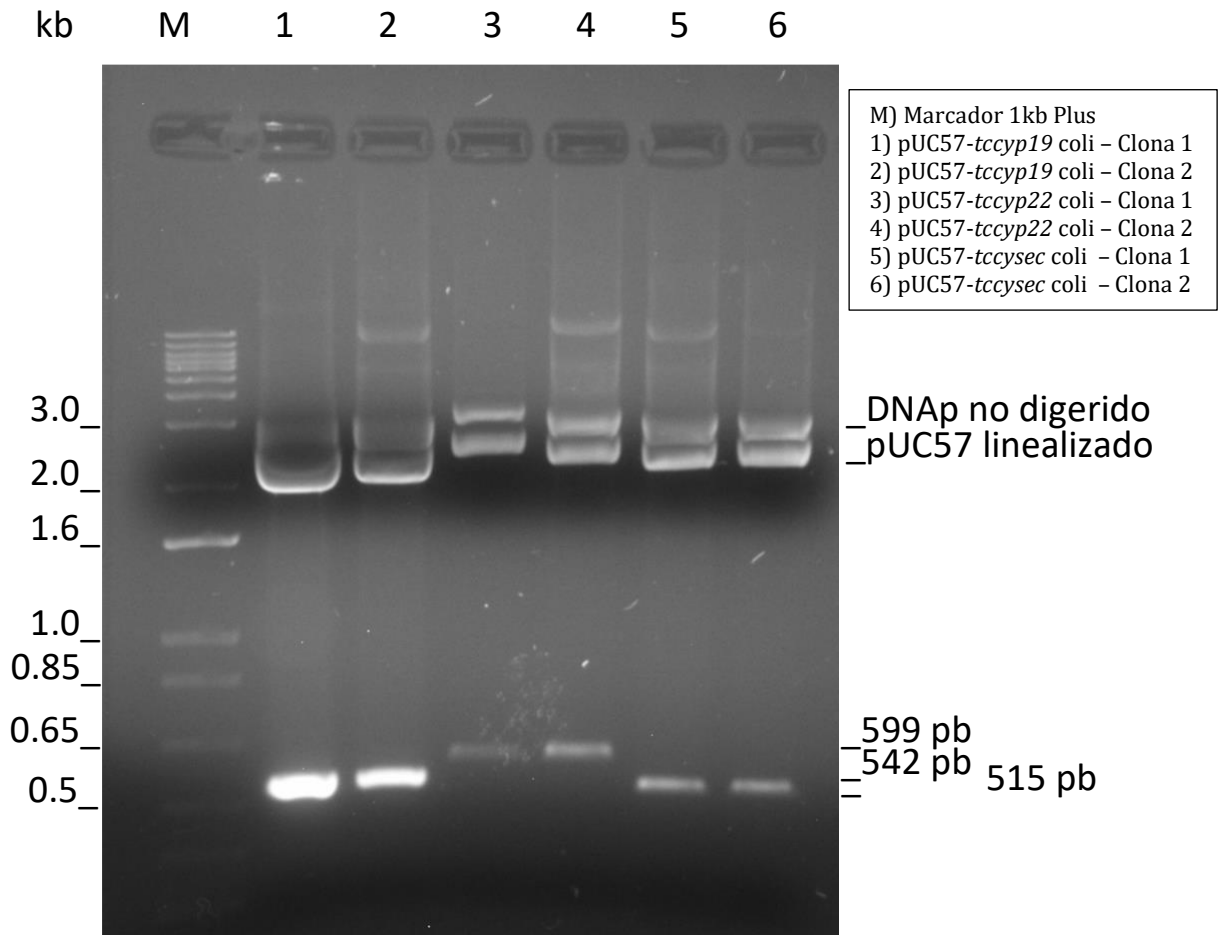
regiones altamente solubles, mientras que los menores que -1 (rojo) son poco solubles (Sormanni et al., 2015).

El servidor AggresCan, se fundamenta en un algoritmo que predice regiones propensas a la agregación en secuencias proteicas, comparando las propiedades de agregación de diferentes proteínas o conjuntos de ellas y analizando el efecto de las mutaciones en las propensiones a la agregación de dichas proteínas (de Groot et al., 2012).

El servidor AggresCan3D, integra la información 3D de las estructuras proteicas y evalúa la contribución de las regiones propensas a la agregación expuestas a solventes. La propensión a la agregación se calcula para las regiones esféricas centradas en cada residuo de carbono  $\alpha$ , utilizando la escala de agregación de aminoácidos intrínsecos del método Aggrescan. Un valor de puntuación mínimo (azul) representa una proteína más soluble y un valor de puntuación máximo (rojo) representa una proteína más propensa a la agregación, ambas consideraciones en el contexto estructural (Kuriata et al., 2019).

### **12.8. Apéndice 8**

La Figura 25 muestra un gel de agarosa al 2% obtenido por el Dr. Gerardo Reséndiz Cardiel, con el análisis de la doble digestión del ADNp correspondiente a las construcciones pUC57-*tccyp19* coli, pUC57-*tccyp22* coli y pUC57-*tccysec* coli utilizando las enzimas NcoI y XhoI, observando la liberación de fragmentos de 542 pb, 599 pb y 515 pb correspondientes a los insertos de los genes *tccyp19*, *tccy22* y *tccysec* coli, respectivamente. Existiendo dos clonas positivas para cada una de las construcciones de interés donadas por el Dr. Gerardo Reséndiz Cardiel.



**Figura 27. Análisis de la Doble Digestión de pUC57-*tccyp19* coli, pUC57-*tccyp22* coli y pUC57-*tccysec* coli.** El análisis electroforético permitió identificar la liberación de los insertos en el tamaño esperado para los genes *tccyp19* (542 pb), *tccyp22* (599 pb) y *tccysec* (515 pb).



## 12.9. Apéndice 9.

### Bases de Datos y Servidores utilizados para el Análisis Bioinformático

BLAST NCBI. (2021). BLAST NCBI. Consultado en: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>

Clustal Omega. (2021) EMBL-EBI. Consultado en: <https://www.ebi.ac.uk/Tools/msa/clustalo/>

ProtParam. (2021). Expasy. Consultado en: <https://web.expasy.org/protparam/>

NCBI. (2021). NCBI. Consultado en: <https://www.ncbi.nlm.nih.gov>

TriTrypDB. (2021). TriTrypDB. Consultado en: <https://tritypdb.org/tritypdb/app>

RCSB PDB. (2021). RCSB. Consultado en: <https://www.rcsb.org>

Swiss-Model. (2021). Expasy. Consultado en: <https://swissmodel.expasy.org>

trRosetta. (2021). Yang Lab. Consultado en: <https://yanglab.nankai.edu.cn/trRosetta/>

I-TASSER. (2021). Zhang Group. Consultado en: <https://zhanggroup.org/I-TASSER/>

3DRefine. (2021). Sysbio. Consultado en: <http://sysbio.rnet.missouri.edu/3Drefine/>

Protein-Sol. (2021). Protein-Sol. Consultado en: <https://protein-sol.manchester.ac.uk>

SoDoPe. (2021). TISIGNER. Consultado en: <https://tisigner.com/sodope>

CamSol. (2021). Chemistry of Health-software. Consultado en: <https://www-cohsoftware.ch.cam.ac.uk/index.php>

Aggrescan. (2021). UAB. Consultado en: <http://bioinf.uab.es/aggrescan/>

Aggrescan3D. (2021). Aggrescan3D. Consultado en: <http://biocomp.chem.uw.edu.pl/A3D2/>

## **13. METODOLOGÍAS GENERALES**

### **13.1. 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 y 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, por 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<sub>600</sub> 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<sub>2</sub> (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<sub>2</sub> (0.1 M) estéril y se agregaron 88 µL de DMSO por cada mL de células en CaCl<sub>2</sub>. Se hicieron alícuotas de 100 µL de células competentes y se almacenaron a -80°C hasta su uso (Green & Sambrook, 2021).

### **13.2. Transformación de células químicamente competentes de *E. coli*.**

La transformación de células quimiocompetentes DH5α y BL21 (DE3) de *E. coli*, se realizó por el método de choque térmico. Una alícuota de 100 µL de células competentes se mezcló en condiciones de esterilidad con un volumen de plásmido, construcción o ligación de acuerdo con la etapa que se estuviera realizando, después se incubó 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 s e inmediatamente después se incubaron a 4°C durante 3 min en baño de hielo. A continuación, se adicionaron 250 mL de medio LB sin antibiótico de selección,

incubando durante 1 h a 37°C y 220 rpm, y finalmente se sembró en placas con medio LB más el antibiótico de selección a 37°C por 16 h.

### **13.3. 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 µg/mL o kanamicina 50 µg/mL) y se incubaron durante 16 h a 37°C y 200 rpm en una incubadora con agitación orbital, posteriormente se extrajo el ADNp utilizando del kit “QIAprep Spin Miniprep” (QIAGEN, 2020a). Utilizando el ADNp obtenido se realizaron dobles digestiones con las enzimas de restricción NcoI y XhoI para confirmar la presencia del inserto a identificar y correspondiente a los genes *tccyp19*, *tccyp22* o *tccysec*.

### **13.4. Almacenamiento a largo plazo de cepas bacterianas (gliceroles)**

Las clonas que resultaron positivas se prepararon en cultivos en medio LB + Kanamicina [50mg/mL] incubando por 16 h a 37°C, 220 rpm hasta que el cultivo alcance la fase logarítmica detectada mediante el monitoreo de la OD<sub>600</sub> hasta que alcance un valor entre 0.4 y 0.7. Posteriormente, se colocó en cada vial 250 µL de solución estéril de glicerol al 80% + 250 µL del cultivo bacteriano. Mezclando 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 hasta su uso.

### **13.5. Medios de cultivo para *E. coli*.**

Las pruebas de inducción y análisis de crecimiento de las proteínas recombinantes se utilizó la cepa BL21 (DE3) de *E. coli*. El crecimiento en esta etapa se realizó en el medio LB (1% Triptona, 0.5% Extracto de levadura y 1% NaCl), adicionando 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. La producción de las proteínas recombinantes utilizó la cepa BL21 (DE3) de *E. coli*. El crecimiento en esta etapa se realizó en el medio TB (1.2% Triptona, 2.4% Extracto de levadura, 0.4% Glicerol, adicionando al final una solución 0.17M KH<sub>2</sub>PO<sub>4</sub>, 0.72M K<sub>2</sub>HP0<sub>4</sub>), adicionando 50 µg/mL de kanamicina como antibiótico de selección. Todos los medios preparados de acuerdo al volumen a utilizar y conforme a lo establecido en la literatura (Green & Sambrook, 2021).