



# **Caracterización de dominios con homología a chaperonas de secreción de bacterias enteropatógenas en tripanosomátidos**

**LUCIA CELAY**

Director: Dr. Carlos A. Buscaglia  
Co-directora: Lic. Camila Centeno Camean

[Abril 2024]

“Ser universitario más que un privilegio es un deber de compromiso con el pueblo para devolver parte de lo que la Nación y la Patria nos dieron”

Cristina Fernandez de Kirchner

## **Agradecimientos**

A la UNSAM, la universidad que me permitió estudiar la carrera que tanto me gusta, y que me enseñó lo hermoso de la educación pública. Gracias a todos sus profesores y gente por el aprendizaje y tiempo compartido.

A Carlos por invitarme a ser parte de su laboratorio. Gracias por darme un lugar para que pueda aprender y crecer como biotecnóloga, y a la vez transmitirme tus conocimientos en el camino.

A Cami por enseñarme absolutamente todo en la mesada y tenerme mucha paciencia. Gracias por hacerme reír todos los días en el labo, escuchar todas mis quejas y entenderme siempre, pero por sobre todo, gracias por ser mi amiga desde el día 1, es hermoso compartir todos nuestros gustos y pensar siempre igual.

A Vir, Guada y Aldi por darme ganas de ir al IIB todos los días, tenerlas a ustedes de compañeras en el labo fue hermoso, los almuerzos con chismes, política y teorías eran mi parte favorita del día.

A la gente del IIB, en especial a Ceci, Gabi Briones, Mariela, Nico, Jacki, Celes, Lucas y Guada. Gracias por ayudarme con todos los experimentos, por preocuparse por cómo estoy y por mostrarme lo valioso que es hacer ciencia en Argentina.

A Juli por acompañarme en todo este último tramo de la carrera con muchos mates y meriendas, compartir nuestras desgracias en el labo siempre fueron un placer. Gracias por todos los chistes malos, los dichos roquenses y las mañas compartidas.

A Cami, Emi y Noe que hicieron que los trabajos en grupo sean muy divertidos. Gracias por compartir infinitas charlas por discord y los quiz de personalidad en el laboratorio, la carrera fue mucho más fácil gracias a ustedes.

A Coti, Juli, Vicki y Cata gracias por formar un grupo de amigas donde podemos hablar mil horas de biotec y también mil horas de nuestros problemas. Las juntadas entre y post cuatrimestre con charlas infinitas, mucho vino y ñoquis me hicieron profundamente feliz.

A la hermosa camada de biotecnología por todo el tiempo compartido en los recreos, en las ochenta horas que estamos en el campus y en las juntadas por todo el conurbano. Me alegra mucho que hayamos armado un grupo tan lindo, donde queremos para todos lo mismo: que se termine esta tortura de carrera.

A Emi por ayudarme a elegir la carrera, por alentarme todos estos años y por festejar cada uno de mis logros como si fueran tuyos. Gracias por entender las ausencias y las cancelaciones, seguir teniéndote como amiga es un privilegio.

A Juli, Tomi y Cami por darme fuerzas para seguir en los momentos más pesados. Su amistad me llena el corazón, juntarnos a comer fideos con manteca y compartir nuestros problemas hizo que todos estos años sean más fáciles.

A Marti y a Pilar por escucharme hablar sobre biotec con tanto amor, todo su interés en el tema y sus preguntas son un mimo al alma. Gracias por abrirme las puertas de su casa siempre, alimentarme y hacerme reír un montón.

A Guido por cuidarme todo este último tramo. Gracias por las risas, los cafecitos y darme el gusto de conocer a Duki y Shaka. Tu compañía me trae paz y me hace más fácil el día a día.

A todos mis amigos gracias por el cariño y la felicidad compartida, la vida sin ustedes no tendría sentido. Les amo ❤️

A Mabel, mi abuela, por todos los mensajes preguntándome cómo va la carrera, gracias por preocuparte porque “siempre estoy estudiando”. Gracias por las milanesas, las albóndigas y las papas fritas que acompañaron los días de estudio, pero más que nada gracias por todo el amor.

A Nati por dejarme vivir en tu casa, hiciste que todos estos años fueran más fáciles. Gracias por hacerme la comida mientras rendía o llegaba de cursar a las once de la noche. La manera en la que me enseñaste sobre la ciencia la voy a llevar siempre conmigo, tu dedicación y esfuerzo para la investigación es completamente admirable y un ejemplo hermoso.

A Clari por ser la mejor hermana. Sos mi persona favorita del mundo, convivir con vos es un placer, gracias por los bailes, las series, los tereres, los cafés, las películas y las cenas. Me encanta que seamos amigas, que pensemos lo mismo y que nos entendamos con solo mirarnos.

A mis papás, Vero y Miguel, por apoyarme todos estos años. A mamá por estar siempre atenta a lo que necesito, gracias por resolver todas mis dudas de la vida adulta, por lavarme la ropa y prepararme milanesas para que me traiga a CABA, por mandarme fotos de loki mientras estoy estresada y por bancarte todos los audios largos donde no módulo. A papá gracias por cumplirme todos los caprichos siempre, por llevarme y traerme de mercedes a capital después de haber trabajado todo el día, gracias por estar siempre dispuesto a escucharme y ayudarme con el estrés facultativo, con una canilla rota o con un depto inundado. Si pude estudiar es todo gracias a ustedes que me bancaron los 6 años permitiéndome el lujo de vivir en capital. No solo eso, sino que también se bancaron mis ausencias, mi falta de tiempo y mi cansancio. Gracias por cuidarme y acompañarme en cada pasito, por alegrarse de todas mis victorias y por preocuparse siempre porque este bien. Este éxito es para ustedes.

**Parte de los resultados mostrados en este trabajo han sido presentados en:**

*Diversidad de dominios con homología a chaperonas de secreción de bacterias enteropatógenas en tripanosomátidos. Celay L., Centeno Cameán C., Durante IM., Briones G., Buscaglia C.A. XXXIII Reunión Anual de la Sociedad Argentina de Protozoología, Noviembre 2022, Buenos Aires, Argentina*

## Resumen

Los tripanosomátidos son un grupo de protozoos parásitos del orden kinetoplastida de gran importancia médica y veterinaria. Entre estas especies, destacan *Leishmania spp.*, *Trypanosoma brucei* y *Trypanosoma cruzi*, los agentes etiológicos de la leishmaniasis, la enfermedad del sueño humana africana y la enfermedad de Chagas, respectivamente. Para estas tres enfermedades tropicales desatendidas, de gran prevalencia y alto impacto en la salud pública global, no existen vacunas o tratamientos apropiados. En este contexto, existe una necesidad urgente de identificar nuevos compuestos y/o blancos de intervención contra estos organismos.

En nuestro laboratorio, se identificó y caracterizó una proteína de *Trypanosoma cruzi*, que se expresa en los estadios replicativos del parásito y está involucrada en la captación de nutrientes. Esta proteína, denominada TCLP1, presenta una estructura compleja, con distintas señales de tráfico/localización sub-celular y tres dominios inferidos de homología: un dominio N-terminal del tipo UBL, un dominio C-terminal PDZ, involucrado en fenómenos de interacción proteína-proteína, y uno al que denominamos 'CEST', ya que muestra homología estructural con chaperonas bacterianas que participan en la secreción de factores de virulencia a través del sistema de secreción tipo III (TIISS).

En rastreos genómicos preliminares logramos identificar una pequeña familia de proteínas presentando el dominio CEST, que hasta la descripción original de TCLP1 se pensaba restringido a bacterias enteropatógenas, dentro del clado de los tripanosomátidos. En el presente trabajo se busca profundizar dichos rastreos y caracterizar algunas de estas proteínas con el fin de comprender aspectos de su evolución, estructura y función.

Para cumplir con este objetivo evaluamos la conservación estructural de los dominios CEST mediante herramientas bioinformáticas, se realizaron alineamientos de secuencias, árboles filogenéticos, predicción de estructuras secundarias y terciarias. A su vez, caracterizamos estructuralmente a los dominios estudiados mediante un modelo experimental basado en el TIISS de *S. Typhimurium*.

Rastreos genómicos exhaustivos revelaron que este dominio se encuentra presente de manera ubicua en el linaje de los tripanosomátidos, y los genes encontrados se delinearon en 3 grupos robustos de homología a los cuales denominamos CEST 1, 2 y 3. Además, los resultados revelaron una significativa similitud estructural global entre los dominios CEST y las chaperonas TIISS bacterianas, sin embargo los dominios CEST conservan sólo alguno de los aminoácidos claves para la actividad de las chaperonas de secreción. Por último, los ensayos funcionales demostraron que los dominios CEST no logran complementar funcionalmente la actividad de dos chaperonas bacterianas evaluadas.

En conjunto, este trabajo de Tesis de Licenciatura proporciona evidencia de la conservación evolutiva y estructural de los dominios CEST en tripanosomátidos. Estos resultados abren nuevas posibilidades para estudiar a nivel molecular el papel de los dominios CEST en la biología de estos parásitos, incluyendo su posible participación en la secreción de proteínas y la virulencia.

# Índice

<b>Agradecimientos</b>	<b>2</b>
<b>Resumen</b>	<b>5</b>
<b>Abreviaturas y definiciones</b>	<b>8</b>
<b>Introducción</b>	<b>10</b>
El orden Kinetoplastida y tripanosomátidos	10
El bolsillo flagelar de los tripanosomátidos	11
TCLP1 (Trypanosomatid CesT-like Protein 1) de <i>T. cruzi</i>	12
Sistema de secreción tipo 3 (TIIIS) de enterobacterias	13
TIIIS de <i>S. Typhimurium</i>	15
Sitio de unión conservado en las chaperonas de secreción	18
Abordaje experimental	19
<b>Objetivos</b>	<b>20</b>
<b>Materiales y métodos</b>	<b>21</b>
Técnicas de análisis <i>in silico</i>	21
Búsqueda y análisis de secuencias	21
Caracterización genómica	21
Caracterización estructural de dominios CEST	21
Soluciones generales	22
Técnicas generales de biología molecular	22
Purificación de ADN plasmídico	22
Purificación de ADN genómico	23
PCR	23
Electroforesis de ADN	23
Digestión de ADN con enzimas de restricción	24
Ligación	24
Preparación de bacterias competentes	24
Transformación por heat shock	25
Técnicas bacteriológicas	25
Cultivo de <i>S. Typhimurium</i>	25
Preparación y electroporación de <i>S. Typhimurium</i>	25
Análisis de secreción de <i>S. Typhimurium in vitro</i>	25
Generación de <i>S. Typhimurium</i> KO utilizando el sistema $\lambda$ Red	26
Técnicas analíticas de proteínas	26
Electroforesis de proteínas en geles de poliacrilamida desnaturalizante (SDS-PAGE)	26
Ensayos de Western blot	26

Ensayos de pull-down	27
<b>Resultados</b>	<b>28</b>
Capítulo 1: Caracterización de una familia de genes con dominios CEST en kinetoplástidos	28
Rastreo de secuencias	28
Análisis filogenéticos	28
Caracterización genómica	30
Capítulo 2: Caracterización <i>in silico</i> de dominios CEST en tripanosomátidos	33
Análisis de la estructura secundaria	33
Selección de secuencias	34
Análisis de la estructura terciaria	34
Capítulo 3: Generación de herramientas para la evaluación funcional de dominios CEST	37
Dominios CEST	37
Efectores bacterianos	40
Líneas de <i>S. Typhimurium</i> genéticamente deficiente para SpaK	42
Capítulo 4: Evaluación funcional de los dominios CEST	47
Sistema SicP/SptP	47
Sistema SpaK/SopE2	48
<b>Discusión</b>	<b>50</b>
<b>Conclusiones</b>	<b>53</b>
<b>Bibliografía</b>	<b>54</b>
<b>Anexos</b>	<b>59</b>



## Abreviaturas y definiciones

- 6xHis: Epitope basado en la secuencia HHHHHH, usado para etiquetar y facilitar la purificación de proteínas.
- aa: Aminoácido. Se utilizará la nomenclatura de 3 letras para designar los aa.
- ADN: Ácido desoxirribonucleico.
- Amp: Ampicilina.
- C-terminal: Extremo de un péptido o proteína en el que se encuentra un grupo carboxilo libre.
- CAT: Cloranfenicol acetil transferasa.
- CEST: Dominio del tipo CesT.
- DMSO: Dimetilsulfóxido.
- D.O.: Densidad óptica.
- DTT: Ditioneitol.
- EDTA: Ácido etilendiaminotetraacético.
- FLAG: Epitope basado en la secuencia DYKDDDK, usado para etiquetar proteínas.
- H<sub>2</sub>O mq: Agua mili-Q.
- IPTG: Isopropil-β-D-1-tiogalactopiranosido.
- Kan: Kanamicina.
- kDa: Kilo Dalton.
- KO (*Knock Out*): Cepa mutada por delección específica de un gen.
- Kpb: Kilo pares de bases.
- kV: Kilovoltio, que equivale a mil voltios.
- LB: Medio de cultivo Luria-Bertani.
- MAD (*Multiwavelength Anomalous Dispersion/Diffraction*): Dispersión/difracción anómala a múltiples longitudes de onda, técnica utilizada en cristalografía de rayos X.
- MCS (*Multiple Cloning Site*): Sitio múltiple de clonado.
- N-terminal: Extremo de un péptido o proteína cuyo grupo amino está libre.
- ON (*Overnight*): Durante la noche.
- ORF (*Open Reading Frame*): Marco abierto de lectura.
- Ori: Origen de replicación.
- PAGE (*Polyacrylamide Gel Electrophoresis*): Electroforesis en gel de poliacrilamida.
- PB: Pares de bases.
- PBS: Buffer fosfato salino.
- PCR (*Polymerase Chain Reaction*): Reacción en cadena de la polimerasa.
- PDZ: PSD95/Dlg1/zo-1, dominio presente en algunas proteínas de eucariotas, presuntamente involucrado en interacciones proteína-proteína.
- RBS (*Ribosome Binding Site*): Sitio de unión al ribosoma.
- RMSD (*root mean square deviation*): Es una medida de similitud entre dos estructuras. Es una diferencia cuadrática media sobre las coordenadas x, y, z de cada átomo.
- Rpm: Revoluciones por minuto.
- SDS (*Sodium Dodecyl Sulphate*): Dodecil sulfato de sodio.
- SDS-PAGE: Electroforesis en gel de poliacrilamida con dodecilsulfato sódico.

- SOB (*Super Optimal Broth*): Medio de cultivo bacteriano súper óptimo.
- TIIIS: Sistema de secreción tipo 3.
- Tag HA: Epitope de la hemaglutinina de influenza humana correspondiente a los aa 98-106 (YPYDVPDYA).
- TE: Buffer Tris-EDTA.
- TEMED: N,N,N,N'-tetrametiletildiamina.
- Tris: 2 amino-2 hidroximetil-1,3 propanodiol.
- UBL (*Ubiquitin-Like Domain*): Dominio parecido a ubiquitina
- WB: Western blot.
- WT (*Wild Type*): Genotipo salvaje.

## Introducción

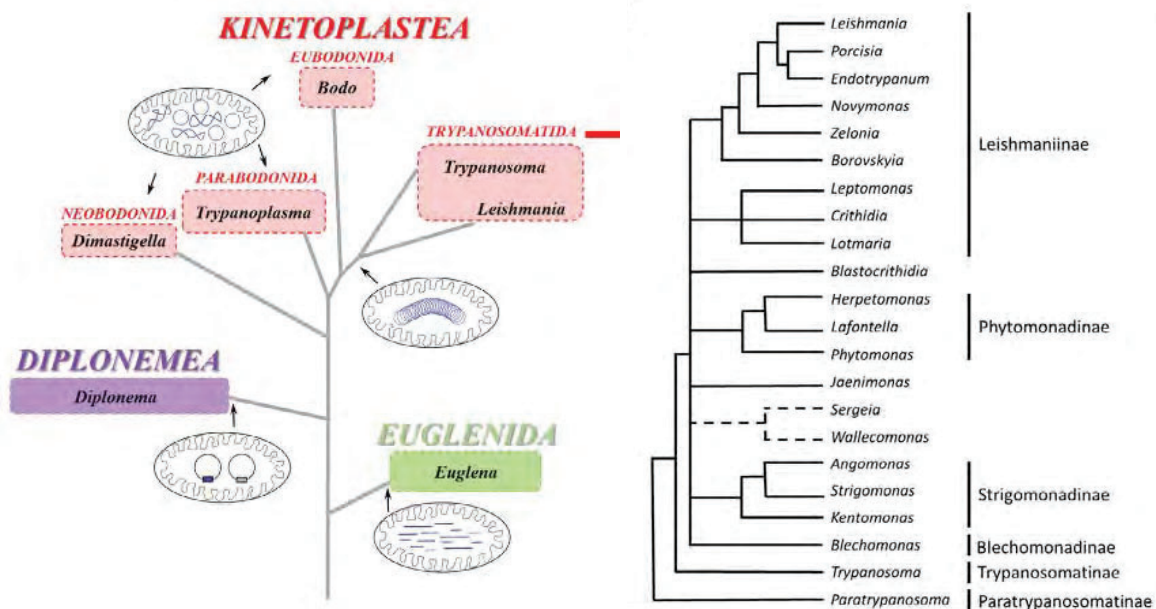
### El orden Kinetoplastida y tripanosomátidos

Los kinetoplástidos son protozoos flagelados que constituyen un grupo extenso y diverso de eucariotas ancestrales, entre los que se cuentan organismos de vida libre y parásitos. Desde el punto de vista de la biología celular, los kinetoplástidos se distinguen por tener glucólisis compartimentalizada en glicosomas, transcripción policistrónica, ausencia casi total de intrones, y un *trans-splicing* de mensajeros maduros generalizado [1], [2]. Otra particularidad es que poseen una organela distintiva, el kinetoplasto, una estructura basófila que forma parte de la mitocondria y donde se concentra el ADN mitocondrial (llamado kDNA).

Los kinetoplástidos se pueden separar en dos grupos principales que incluyen los parásitos obligados (Trypanosomatidae), que se subdividen en los clados 'Trypanosoma' (género *Trypanosoma*) y 'Leishmania' (incluye los géneros *Leishmania*, *Blepharomonas*, *Crithidia*, *Leptomonas*, *Endotrypanum* y *Parcisia*); y un grupo diverso de biflagelados de vida libre, entre los que pueden distinguirse cuatro clados principales (Eubodonida, Parabodonida, Neobodonida y Prokinetoplastida) [3] (Fig 1). Las especies uniflageladas de la familia de los tripanosomátidos son las más estudiadas, debido a su importancia médica y agropecuaria [4]. Los tripanosomátidos son parásitos predominantemente monoxenos (es decir, con un único hospedador) que infectan a invertebrados; sin embargo, son más conocidos por sus miembros dioxenos (es decir, los que tienen un ciclo de vida que alterna entre dos hospedadores), que pueden ser patógenos para una gran variedad de animales, (incluyendo a los seres humanos) y plantas [5]. Entre estas especies, destacan *Leishmania spp.*, *Trypanosoma brucei* y *Trypanosoma cruzi*, los agentes etiológicos de la leishmaniasis, la enfermedad del sueño humana africana y la enfermedad de Chagas, respectivamente [6], tres enfermedades tropicales desatendidas (*Neglected Tropical Diseases*, NTD) de gran prevalencia y alto impacto en la salud pública global. Actualmente no existen vacunas y/o tratamientos apropiados para la infección con tripanosomátidos. Las drogas que se utilizan no son completamente efectivas, presentan efectos adversos y no son sencillas de evaluar en cuanto a su eficacia terapéutica. En este contexto, existe una necesidad urgente de identificar nuevos compuestos y/o blancos de intervención contra estos patógenos [7].

Recientemente, el auge de las nuevas tecnologías de secuenciación, ha permitido obtener un gran repositorio de genomas de tripanosomátidos y otros kinetoplástidos, que han sido depositados en la base de datos pública TritypDB [8]. La disponibilidad de datos genómicos de alta calidad de un conjunto representativo de organismos permite abordar los patrones evolutivos y las capacidades metabólicas de los linajes parásitos, y comprender sobre las características comunes entre los kinetoplástidos parásitos y los de vida libre [9]. En el año 2005 se publicó el genoma de *T. cruzi* [10], junto con el de otros dos tripanosomátidos, también llamados Trityp, de gran importancia para la salud humana: *Trypanosoma brucei* [11] y *Leishmania major* [12]. Esta primera secuenciación del genoma de *T. cruzi* reveló que más del 50% del genoma no es sinténico con el de *T. brucei* y *L. major* y que presenta una gran complejidad. Esta parte del genoma, llamada "disruptiva", está compuesta por secuencias repetitivas, grandes familias multigénicas y transposones [10]. Por otro lado, la secuenciación

de los *Trityps* reveló la presencia de numerosos genes de origen viral y/o bacteriano en sus genomas [13]. En algunos casos, estos genes virales/bacterianos preservan su secuencia y/o estructura original (y posiblemente su función), indicando que son o han sido necesarios para la evolución de estos organismos [14].



**Figura 1. Árbol esquemático de la vida que representa a la clase Kinetoplastea.** Adaptado de [15].

### El bolsillo flagelar de los tripanosomátidos

Una característica distintiva de los tripanosomátidos es la presencia de un único flagelo, crucial para su movilidad, que emerge de una invaginación de la membrana plasmática denominada bolsillo flagelar [16]. Este bolsillo no es meramente una cavidad, sino un sitio activo y estratégico del parásito, implicado en una variedad de procesos como la polaridad celular, la morfogénesis y la replicación. Específicamente, el área del bolsillo flagelar es el único lugar donde se interrumpe la malla del citoesqueleto subcortical responsable de mantener la forma característica de cada estadio de desarrollo [17]. Por consiguiente, el área del bolsillo flagelar es el único sitio del organismo donde tienen lugar los procesos endo- y exocíticos, incluyendo la importación de nutrientes, la exportación de glicoproteínas y glicolípidos de cubierta y el reciclaje de anticuerpos potencialmente líticos unidos a sus moléculas de superficie [17], [18], [19]. Esta singularidad convierte al bolsillo flagelar en una estructura clave para la interacción del parásito con los hospedadores vertebrados e insectos. Hasta ahora, solo se han descrito y caracterizado unas pocas proteínas asociadas al bolsillo flagelar de tripanosomátidos. Muchas de estas proteínas resultaron ser esenciales, ya que su pérdida o deslocalización condujo a defectos morfológicos que perturbaron funciones celulares [20]. Además, la estrecha relación entre la endo-exocitosis y la virulencia, diferenciación y supervivencia en este grupo de parásitos, destaca al bolsillo flagelar y a las moléculas asociadas como blancos terapéuticos de gran valor potencial.

### **TCLP1 (*Trypanosomatid CesT-like Protein 1*) de *T. cruzi***

En nuestro laboratorio, se identificó y caracterizó una proteína de *T. cruzi* que se localiza en el citoplasma subyacente a la zona del bolsillo flagelar de los estadios replicativos del parásito y que estaría involucrada en la captación de macromoléculas necesarias para la nutrición [21]. De hecho, la sobreexpresión homóloga de esta proteína condujo a una reducción de la capacidad endocítica de los epimastigotes de *T. cruzi* y a una mortalidad exacerbada durante la fase estacionaria tardía de los cultivos axénicos [21].

Esta proteína, denominada TCLP1 (*Trypanosomatid CesT-like Protein 1*, TcCLB.510241.10), presenta una estructura compleja, con distintas señales de tráfico/localización sub-celular y tres dominios inferidos de homología: un dominio N-terminal del tipo *Ubiquitin-Like Domain* (UBL) [22], un dominio C-terminal del tipo PSD95/Dlg1/zo-1 (PDZ), involucrado en fenómenos de interacción proteína-proteína [23] y uno al que denominamos 'CEST', ya que muestra homología marginal (18-30% identidad a nivel de aminoácidos, aa) con proteínas del tipo CesT. Las CesT conforman una familia de chaperonas pequeñas que participan en la secreción de factores de virulencia a través del Sistema de Secreción Tipo III (TIIISS) en bacterias enteropatógenas [24]. Brevemente, las chaperonas CesT se unen co-traduccionalmente a uno o varios efectores (dependiendo de si son mono- o multi-específicas), previniendo su degradación en el citoplasma bacteriano y escoltándolos hacia el *pili* de secreción (ver debajo). Aunque no comparten una identidad de secuencia sustancial entre ellas, todas las chaperonas CesT exhiben homología estructural conteniendo, desde el extremo N- al C-terminal el siguiente arreglo: 1  $\alpha$  hélice, 3 hojas  $\beta$ , 1  $\alpha$  hélice, 2 hojas  $\beta$ , 1  $\alpha$  hélice [24]. Todas ellas adoptan una conformación tridimensional similar, usualmente se organizan en dímeros [24] y exhiben un modo común de interacción con sus efectores específicos [25].

El dominio CEST de TCLP1 mostró una identidad de secuencia variable con proteínas del tipo CesT encontradas en diferentes bacterias, como *Escherichia coli* (~12%), *Salmonella enterica* (~17-19%), *Pseudomonas syringae* (~17%), *Yersinia pestis* (~18%). La identidad máxima (~38%), sin embargo, se observó con una proteína CesT de *Chlamydia trachomatis* (Fig 2). *Chlamydia trachomatis* pertenece al grupo Chlamydiae/Verrucomicrobia, compuesto por bacterias patógenas intracelulares obligadas que expresan sistemas de secreción tipo III funcionales durante todo su ciclo infeccioso [26]. Más importante, a pesar de la baja homología de secuencia, modelados de la estructura tridimensional del motivo CEST de TCLP1, mostraron una razonable superposición de las hojas  $\beta$  hidrofóbicas internas y las  $\alpha$  hélices externas con las de la chaperona SicP de *Salmonella enterica* [21].

A partir de rastreos genómicos, logramos identificar una pequeña familia de proteínas presentando el dominio CEST, que hasta la descripción original de TCLP1 se pensaba restringido a bacterias enteropatógenas [21], dentro del clado de los tripanosomátidos. Si bien la posibilidad de evolución (estructural) convergente no puede descartarse, la hipótesis más parsimoniosa para explicar la introgresión del dominio CEST en el linaje de los tripanosomátidos implicaría la adquisición de dicho dominio a partir de una chaperona CesT transferida horizontalmente desde algún miembro de las Chlamydiaceae. La idea de eventos de transferencia horizontal de genes entre las *Chlamydiae* y tripanosomátidos fue postulada previamente en base al análisis genómicos comparativos entre estos grupos; más aún, la

proteína Slc1, sugerida como una de las involucradas en esta transferencia horizontal a tripanosomátidos, es de hecho una chaperona Cest (Tabla 1) [27]. En el mismo sentido, se han observado similitudes entre proteínas de *T. cruzi* y del TIISS de *Salmonella*, en particular genes para la movilización del calcio que median la invasión de *Salmonella* [28], lo que podría también explicarse por eventos de transferencia genética horizontal entre enterobacterias y tripanosomátidos.



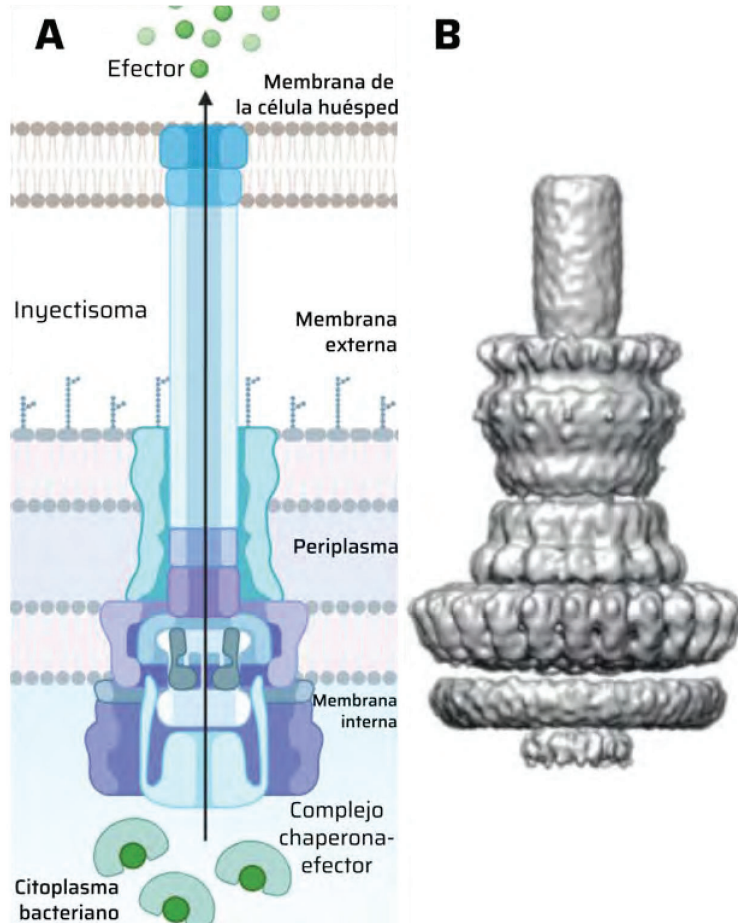
**Figura 2. Alineamiento del motivo CEST de TCLP1 y proteínas Cest bacterianas.** Los residuos conservados están sombreados y la identidad aminoacídica total con respecto a TCLP1 (en %) se indica a la derecha. Adaptado de [21].

### Sistema de secreción tipo 3 (TIISS) de enterobacterias

El TIISS es una nanomáquina empleada para inyectar proteínas 'efectoras' directamente en el citoplasma de células eucariotas huéspedes [29]. Una vez translocados, estos efectores modulan la fisiología de la célula huésped para crear un entorno adecuado para la supervivencia y replicación bacteriana [30]. Los sistemas TIISS están compuestos por alrededor de 25 componentes (principalmente proteínas estructurales, efectores y chaperonas que ayudan en su secreción) que se codifican en isla de patogenicidad 1 (SPI-1), cuya expresión coordinada se induce en respuesta a ciertos estímulos ambientales, como la osmolaridad, el pH y la tensión de oxígeno [31]. Algunos de estos componentes se organizan en una estructura conocida como "inyectisoma" o "complejo de aguja", que consta de múltiples anillos que atraviesan la membrana bacteriana y una extensión en forma de aguja que se extiende varios nanómetros por fuera de la membrana externa de la bacteria (Fig 3). Un canal estrecho, con un ancho de ~28 Å, atraviesa toda esta estructura y proporciona la vía por la cual las proteínas efectoras viajan a través del inyectisoma [29]. Este sistema, evolutivamente relacionado con el flagelo, se conserva en muchas bacterias Gram-negativas que mantienen una relación simbiótica o patógena con un huésped eucariota [31].

Las proteínas destinadas a ser translocadas a través del TIISS (efectores) son dirigidas al 'inyectisoma' mediante chaperonas de secreción que aseguran no solo la especificidad de su translocación, sino que también el orden/jerarquía de translocación durante el proceso de infección [31]. Estos efectores están codificados en las proximidades de sus chaperonas que regulan su expresión a nivel transcripcional [24]. Además, se unen a sus correspondientes chaperonas gracias a la presencia en el N-terminal de un motivo de unión (*chaperona binding site* o CBD) [24]. Además de reconocerlo y escoltarlo hacia el pili, las chaperonas facilitan la

secreción del efector al mantenerlo en una conformación parcialmente desplegada, lo que permite su inserción en el estrecho canal interno del 'inyectisoma' [25]. En ausencia de las chaperonas adecuadas algunas proteínas efectoras pueden secretarse de manera no específica, a través del aparato flagelar [32], o ser rápidamente degradadas en el citoplasma bacteriano [33].



**Figura 3. Representación esquemática del sistema de secreción tipo III.** [A] Diagrama de la inyección de efectores bacterianos en una célula huésped. Las chaperonas CesT (en verde claro) protegen a los efectores en el citoplasma de la bacteria y lo transfieren al cuerpo basal del complejo. [B] Arquitectura general del cuerpo basal, la aguja y la varilla interna del TIISS, obtenida mediante criomicroscopía electrónica (crio-EM) [31].

Existe una gran diversidad de chaperonas TIISS, las que pueden dividirse en al menos dos grupos funcionales con interacciones cualitativamente distintas con sus respectivas proteínas efectoras. El primer grupo comprende a las chaperonas multiespecíficas, que se unen a múltiples efectores, y que, en general, mantienen interacciones más débiles con ellos [34]. En el segundo grupo, se encuentran las chaperonas monoespecíficas, que facilitan el despliegue y translocación de un único efector. En estos casos, el CBD de la proteína efectora envuelve a la pequeña chaperona globular en una conformación desplegada para formar un complejo muy estable [30] (ver debajo). En la Tabla 1 se detallan algunos ejemplos para ambos grupos, presentes en diversas enterobacterias.

Tabla 1: Ejemplos chaperonas TIISS.			
Chaperona	Clasificación	Organismo	Referencia
SicP	Monoespecífica	<i>Salmonella Typhimurium</i>	[33]
Spak	Multiespecífica	<i>Salmonella Typhimurium</i>	[32], [35], [36]
SycE	Monoespecífica	<i>Yersinia spp.</i>	[37]
SycD	Multiespecífica	<i>Yersinia spp.</i>	[38]
SpcU	Monoespecífica	<i>Pseudomonas aeruginosa</i>	[39]
SpcS	Multiespecífica	<i>Pseudomonas aeruginosa</i>	[39]
Slc1	Multiespecífica	<i>Chlamydia trachomatis</i>	[40]
CesD2	Monoespecífica	<i>E.coli</i>	[41]
CesD	Multiespecífica	<i>E.coli</i>	[41]

### TIISS de *S. Typhimurium*

Según los serotipos y los hospedadores, *Salmonella enterica* causa una amplia gama de enfermedades transmitidas mediante alimentos o agua contaminada, y puede causar desde gastroenteritis hasta fiebre tifoidea sistémica [42]. El TIISS codificado en la isla de patogenicidad 1 de *Salmonella* (SPI-1) es necesario para inducir parte de estos síntomas, mediante la invasión de células epiteliales intestinales, la inducción de procesos apoptóticos en macrófagos y la inducción de respuestas proinflamatorias en el hospedador [43].

*S. Typhimurium* se diferencia de otras bacterias enteropatógenas en que sólo algunas de las proteínas efectoras se codifican en las proximidades del SPI-1. El conjunto de estos efectores pueden modular o interferir en una amplia gama de funciones celulares como la dinámica del citoesqueleto de actina y tubulina, la expresión génica, el tráfico vesicular, la muerte celular programada y la progresión del ciclo celular [29]. En este trabajo en particular, nos focalizaremos en 2 efectores del TIISS de *S. Typhimurium*: la fosfatasa SptP, que es unida por la chaperona mono específica SicP [33], y la proteína intercambiadora de guanina SopE2 [32]. SopE2 se une a la chaperona multiespecífica SpaK (también conocida como InvB), involucrada además en la secreción de otros factores de virulencia como SopE [32], SopA [36] y SipA [35].

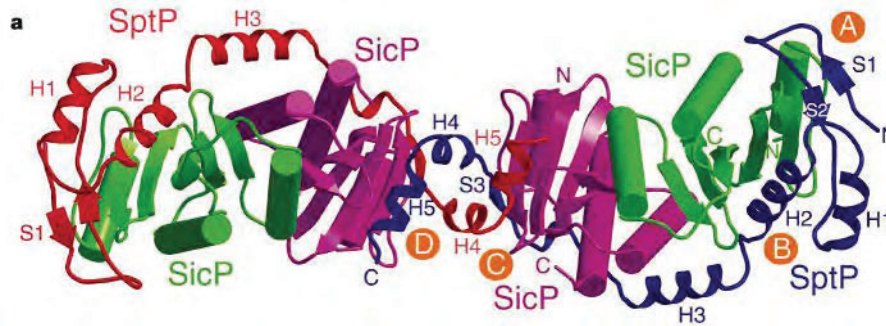
SopE2 activa Rho-GTPasas como Rac1 y Cdc42, promoviendo así la formación de una red de actina altamente ramificada. Además, en conjunto con SopE y SopB, induce la formación de *ruffles* en la membrana de la célula huésped, los cuales son generados por la reorganización del citoesqueleto y son necesarios para la invasión de *S. Typhimurium* [42]. Una vez establecida la infección, es necesario disolver esos *ruffles*, para garantizar la viabilidad celular. Este proceso compensatorio es mayormente promovido por la actividad fosfatasa de SptP [33]. La estricta regulación de la polimerización/despolimerización de los filamentos de actina es vital para la viabilidad celular y la supervivencia intracelular bacteriana.

- Complejo SicP-Sptp

La estructura del complejo SicP-SptP se determinó mediante la técnica de cristalografía de rayos X con dispersión anómala a múltiples longitudes de onda (MAD) [25]. En este co-cristal, la chaperona SicP se dispone en dos conjuntos de homodímeros fuertemente unidos, que a su

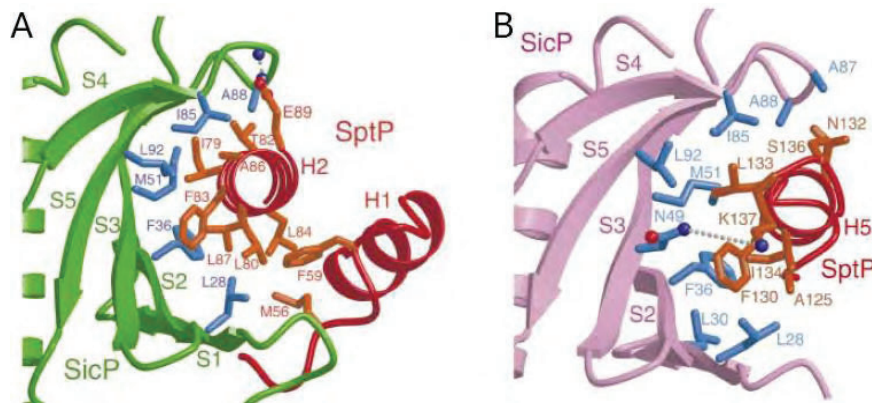


vez se unen a dos moléculas de SptP. Los pares de homodímeros SicP no interactúan entre sí, sino que están unidos por una interfaz molecular formada entre las dos moléculas de SptP, lo que resulta en una estructura tetramérica de chaperonas. Cada molécula de SptP, a su vez, interactúa como un polipéptido desplegado con tres chaperonas SicP (Fig 4).



**Figura 4. Modelado 3D del complejo SicP-SptP de *S. Typhimurium*.** SicP se representa con sus  $\alpha$  hélices como cilindros sólidos y las hojas  $\beta$  como flechas gruesas, mientras que las  $\alpha$  hélices de SptP se muestran como cintas, con las dos cadenas polipeptídicas distintas en rojo y azul. Las moléculas centrales de SicP están en magenta y las exteriores en verde. Las estructuras secundarias de SptP están identificadas como H ( $\alpha$  hélice) y S (hoja  $\beta$ ), y las cuatro regiones de contacto importantes entre SptP y SicP están señaladas (A-D, en naranja) para una molécula de SptP [25].

Múltiples interacciones hidrofóbicas, polares y de puente salino específicas mantienen la unión entre SicP y SptP. Cada molécula de SptP se une a los dímeros de chaperona SicP principalmente a través de cuatro regiones de su dominio CBD (Fig 4, indicadas en naranja). En dos de estas regiones, las interacciones se producen a través de un surco pronunciado formado por la segunda y tercera hoja  $\beta$  de SicP, conocido como surco de unión a hélice (*helix-binding groove*) o HBG, el cual ofrece una amplia superficie de contacto con el efector [25]. Por un lado, el HBG rodea la segunda  $\alpha$  hélice de SptP (Fig 5A), interactuando así con las cadenas laterales de casi todos sus aa. Estos contactos son principalmente hidrofóbicos, centrados en un apilamiento de anillos bencénicos entre los residuos Phe 83 de SptP y Phe 36 de SicP externo. Por otro lado, la  $\alpha$  hélice C-terminal de SptP interactúa con el HBG de SicP central, lo que resulta en un apilamiento entre el residuo Phe 36 de SicP central y Phe 130 de SptP (Fig 5B).



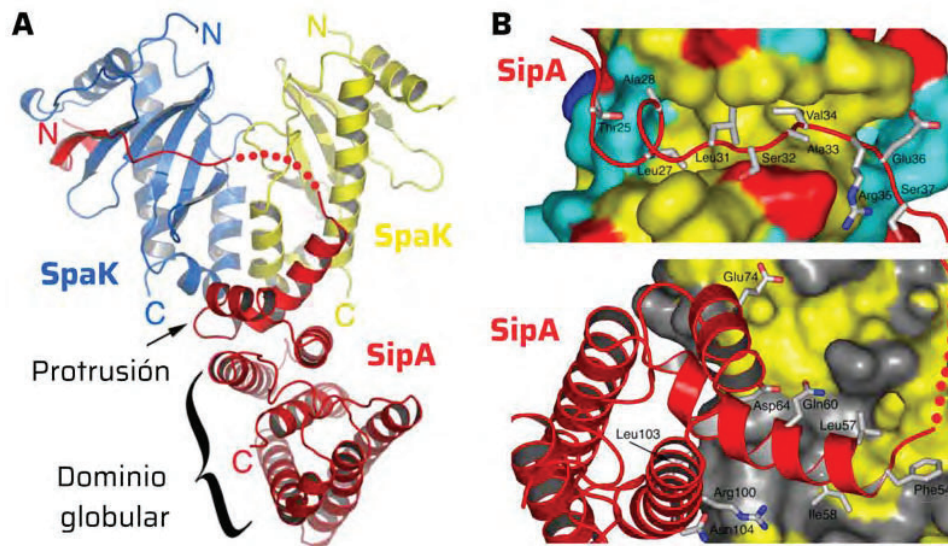
**Figura 5. Interacción del HBG de SicP con el efector SptP de *S. Typhimurium*.** [A] Una  $\alpha$  hélice de SptP (rojo) se une al HBG generado por dos hojas  $\beta$  de SicP externa (verde). [B] Una  $\alpha$  hélice de SptP (rojo) se inserta al HBG de SicP central (magenta). Adaptado de [25].

- Complejo SpaK-SipA

Como se mencionó anteriormente, SpaK es una chaperona multiespecífica que se asocia con varios efectores, incluyendo SopE, SopE2, SipA y SopA. Para comprender su interacción, nos centraremos en el complejo SpaK-SipA, ya que es el único que ha podido ser analizado mediante la técnica de cristalografía de rayos X con MAD [34].

El modelo de interacción entre SipA y SpaK preserva el carácter no globular del efector, y se distingue del resto de los complejos chaperona-efector dado que la interacción ocurre con solo una de las moléculas del homodímero de SpaK (Fig 6A). Numerosos efectores de *S. Typhimurium* interactúan con las chaperonas mediante un polipéptido de estructura desordenada que rodea ambas moléculas del homodímero y establece múltiples contactos con parches hidrofóbicos exteriores de cada chaperona [25]. En el caso de SipA, la región no globular interactúa en cambio con los parches hidrofóbicos conservados de una sola molécula de SpaK, dejando los otros parches de la segunda molécula de SpaK expuestos al solvente [34].

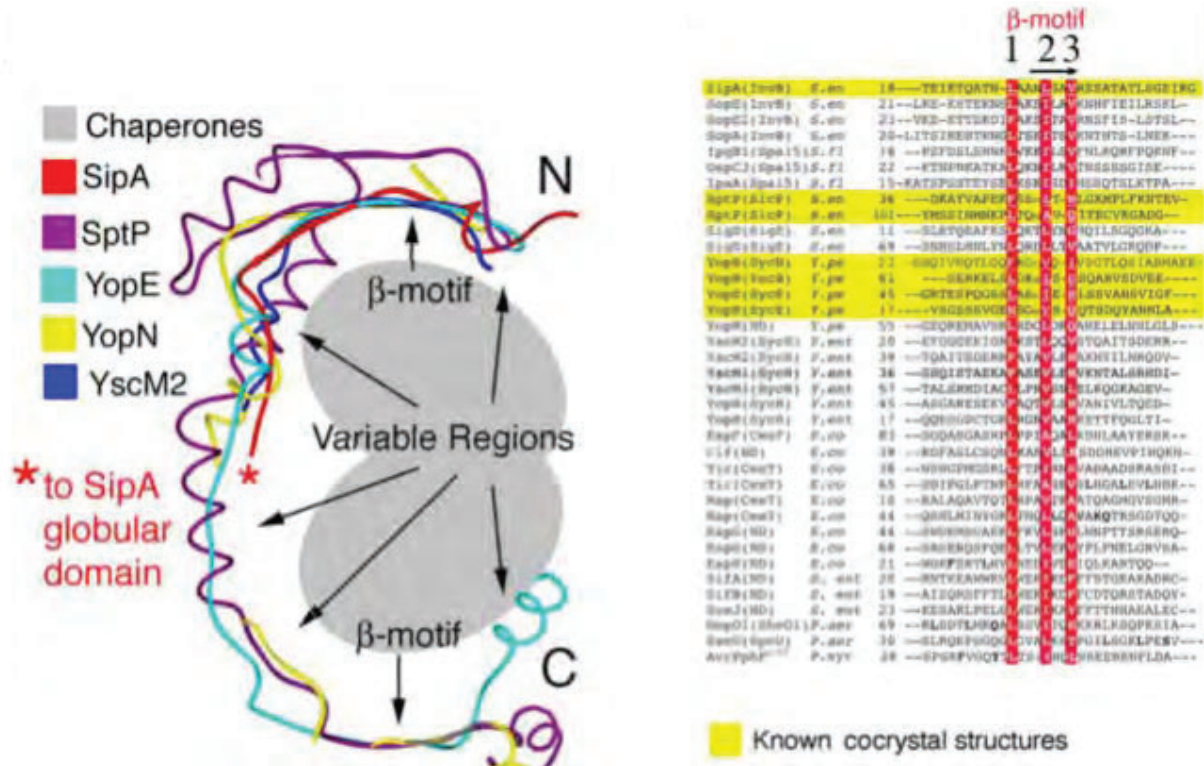
La unión entre SipA y SpaK se genera a partir de múltiples interacciones con los residuos 36-47 de SipA que se extienden sobre el frente de la chaperona (Fig 6B). Por un lado, los residuos Arg35, Glu36 y Ser43 hacen contacto por puentes de hidrógeno con las cadenas laterales de los residuos Asp36, Ser29 y Gln94 de SpaK; por otro lado, se generan contactos de Van der Waals entre las regiones hidrofóbicas de SpaK y el residuo Leu42 de SipA. Además, los residuos 30-36 de SipA se insertan en el bolsillo hidrofóbico de la chaperona, donde los residuos Leu31 y Val34 establecen interacciones con otros residuos no polares de SpaK, mientras que Asn30 y Arg35 de SipA forman puentes de hidrógeno con los residuos Tyr123 y Asp36 de SpaK.



**Figura 6. Estructura cristalina del complejo SipA-SpaK.** [A] La estructura general del complejo SpaK con SipA<sup>22-264</sup> se representa como un diagrama de cinta. Los dos monómeros de SpaK se colorean en azul y amarillo, y SipA se colorea en rojo. Las regiones desordenadas se muestran mediante líneas punteadas. [B] SipA es representado como un diagrama de cinta con cadenas laterales y SpaK como una superficie molecular coloreada de manera que los residuos hidrofóbicos son amarillos, los ácidos son rojos, los básicos son azules y todos los demás residuos polares son cian. Adaptado de [34].

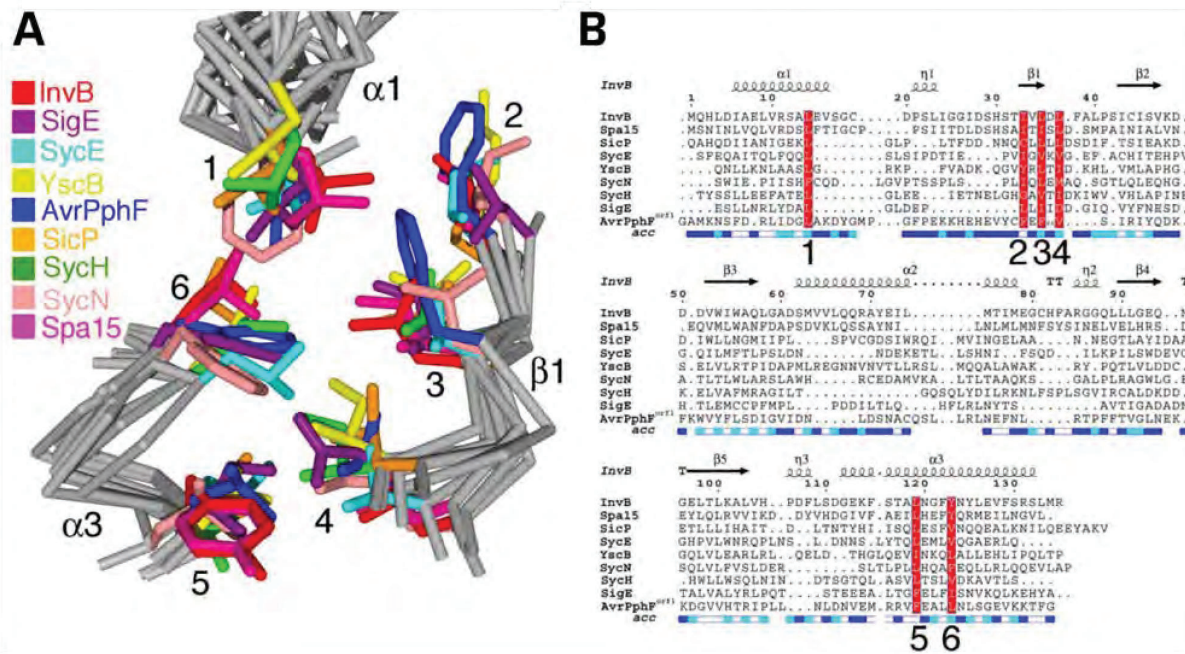
Sitio de unión conservado en las chaperonas de secreción

Tanto los efectores como las chaperonas CesT exhiben un sitio de unión conservado. Seis residuos en las chaperonas CesT mantienen su carácter, ubicación tridimensional y, especialmente, y su afinidad por los residuos del motivo  $\beta$  del CBD (Figura 7) [34]. La primera hoja  $\beta$  desempeña un papel central en esta interacción, ya que establece un apareamiento intermolecular de hojas  $\beta$  antiparalelas. Esta hoja presenta tres residuos hidrofóbicos conservados que interactúan con los residuos del motivo  $\beta$  del CBD (Fig 8B).



**Figura 7. Sitio de unión conservado en los efectores.** Derecha: Superposición de varias estructuras de chaperonas y efectores del TIIISS, se indican el motivo  $\beta$  y las regiones variables. Izquierda: Alineación de secuencias de efectores bacterianos en los que se resalta el motivo  $\beta$ . Adaptado de [34].

El sitio de unión se termina de conformar por una Leu altamente conservada que se encuentra en la "parte superior" del bolsillo hidrofóbico, y por dos residuos hidrofóbicos en el extremo C-terminal de la última  $\alpha$  hélice que delimitan la "parte inferior". Aunque hay algunas variaciones en estos residuos, la naturaleza no polar prevalece, y su disposición tridimensional permanece constante (Fig 8A). En resumen, el motivo  $\beta$  conservado del efector se corresponde con un bolsillo igualmente conservado en las chaperonas.



**Figura 8. Sitio de unión conservado en las chaperonas Cest.** [A] Alineación estructural de varias chaperonas Cest. La cadena principal se muestra en gris, y los residuos de unión al motivo  $\beta$  de los efectores están coloreados según se indica. [B] Alineamiento de secuencias basado en la estructura de nueve chaperonas T355. Se muestran resaltados en rojo los residuos destacados en [A]. Debajo, se muestran los elementos de estructura secundaria de SpaK. Adaptado de [34].

### Abordaje experimental

En el presente trabajo, se buscará profundizar los rastreos genómicos de proteínas CEST en kinetoplástidos y caracterizar algunas de estas proteínas con el fin de comprender aspectos de su evolución, estructura y función. Debido a su posible origen bacteriano, a su similitud estructural con las chaperonas del TIISS (las cuales desempeñan un papel crucial en la transferencia de factores de virulencia al citoplasma de la célula huésped) y a los resultados obtenidos para TCLP1 [21], estas moléculas constituyen atractivos blancos potenciales el desarrollo de nuevas estrategias de intervención en estos patógenos relevantes. Por un lado, para comprender su evolución y para facilitar futuros estudios funcionales, se analizará la distribución de los dominios CEST en el orden kinetoplastida. Por otro lado, para caracterizar la función, desarrollaremos un modelo experimental basado en el TIISS de *S. Typhimurium*, para el que existe una abundancia de conocimiento estructural y funcional. Es importante destacar que *S. Typhimurium* es un microorganismo de patogenicidad limitada (similar a *E. coli*), ampliamente utilizado en laboratorios de investigación biológica y para el que se cuenta con una variedad de herramientas bioquímicas, moleculares y genéticas.

## Objetivos

Considerando el impacto en salud humana y veterinaria de las enfermedades producidas por tripanosomátidos y la necesidad de identificar nuevos blancos de intervención terapéutica, el objetivo principal del presente trabajo es caracterizar estructural y funcionalmente los dominios CEST recientemente identificados en estos organismos. Para desarrollar este objetivo principal, planteamos los siguientes objetivos específicos (OE):

**OE1: Caracterización *in silico* de los dominios CEST de tripanosomátidos.** Se realizará un análisis exhaustivo de la estructura molecular, distribución y evolución de los dominios CEST (y proteínas que los contienen) en los genomas de kinetoplastidos y euglénidos mediante herramientas bioinformáticas.

**OE2: Evaluación funcional de los dominios CEST mediante ensayos de complementación en *S. Typhimurium*.** Se generarán herramientas moleculares y celulares que permitan implementar modelos experimentales en *S. Typhimurium* a partir de las cuales se caracterizará funcionalmente a los dominios CEST.

## Materiales y métodos

### Técnicas de análisis *in silico*

#### Búsqueda y análisis de secuencias

Se utilizaron las secuencias de ADN de distintos dominios CEST como señuelo para la búsqueda por BLASTn en los genomas de tripanosomátidos y bodónidos compilados en la base de datos de acceso público de TritrypDB [8] y en los genomas de Euglenidos compilados en NCBI. Los *hits* fueron filtrados por homología ( $e\ value < 1 \times e^{-13}$ ) y por redundancia (entre secuencias idénticas de una misma especie y luego entre ortólogos entre cepas y/o subespecies de una misma especie) mediante scripts de Python. Los alineamientos de secuencias se generaron y visualizaron mediante ClustalW [44] y MUSCLE-UGENE [45], respectivamente. Se empleó MEGA [46] para obtener árboles filogenéticos, usando el método de máxima verosimilitud, el modelo de evolución molecular de Jones, Taylor y Tornton [47] y 1000 repeticiones Bootstrap. Se utilizó FigTree [48] como visor gráfico de árboles filogenéticos y se infirió la existencia de 3 grupos de homología (llamados CEST 1, 2 y 3) a partir de la topología de los mismos. A su vez, se buscaron dominios auxiliares (UBL, PDZ, etc.) en las proteínas CEST identificadas con InterPro [49]. Se generaron secuencias consenso de aa para cada grupo de homología mediante alineamientos con ClustalW [44] y se realizó la predicción de las estructuras secundarias con JPRED 4 [50], utilizando cada una de las secuencias consenso como *input*.

#### Caracterización genómica

Para cada uno de los 3 tipos de genes CEST encontrados, se evaluó la distribución y el dosaje en genomas de miembros selectos del clado 'Leishmania' (*Leishmania mexicana* y *Crithidia fasciculata*) y del clado 'Trypanosoma' (*T. cruzi* y *T. brucei*) de los tripanosomátidos (Fig 1). Para CEST 2, se analizaron también los genomas de *Paratrypanosoma confusum* y *Parcisia hertigi*. Para el análisis de sintenia, se descargaron de TritrypDB los 5 genes anotados río arriba y 5 río abajo de cada tipo de gen CEST en estos genomas. La posible ortología entre los genes flanqueantes se evaluó considerando su anotación en TritrypDB [8] y también mediante comparaciones de secuencias de ADN, usando BLASTn. Se usó la librería ggplot2 [51] de RStudio para graficar la disposición de los distintos genes CEST en cada genoma.

#### Caracterización estructural de dominios CEST

A partir de los dominios CEST seleccionados se realizaron modelados de la estructura terciaria utilizando la secuencia primaria, mediante AlphaFold2 [52] con los parámetros por *default*. Como métrica de confianza, se usó el test de pLDDT (*predicted local-distance difference test*), que determina las diferencias en las distancias C $\alpha$  (IDDT-C $\alpha$ ) y evalúa la fiabilidad del modelo a nivel local (estimando el error en distancias de un carbono  $\alpha$  (C $\alpha$ ) con C $\alpha$  vecinos en un rango de 15Å). Mediante UCSF Chimera, se graficaron y analizaron los modelos generados y el alineamiento estructural entre los mismos, se consideró el RMSD global como parámetro de comparación. A su vez, se utilizaron estructuras cristalinas de las chaperonas SicP y SpaK de *S. Typhimurium* y se evaluó la homología estructural entre ellas y con los dominios CEST de tripanosomátidos.

**Soluciones generales**

Solución de siembra para muestras de proteínas 5x (CB5x): 50% (v/v) glicerol, 7,7% (p/v) DTT, 10% (p/v) SDS, 0,4 M Tris-HCl pH 6,8, 0,002% (p/v) azul de bromofenol.

PBS: 137 mM NaCl, 2,7 mM KCl, 10mM Na<sub>2</sub>HPO<sub>4</sub>; 2 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.

PI: 25 mM Tris-HCl pH 8, 10mM EDTA (pH 8), 50 mM glucosa.

PII: 0,2 N NaOH, 1% (p/v) SDS.

PIII: 3 M acetato de potasio pH 4,8 o 3 M acetato de sodio pH 5,2.

Solución de Rojo Ponceau: 0,1% (p/v) Rojo Ponceau, 5% (v/v) ácido acético.

Solución de siembra para muestras de ADN 6x (LB6x): 0,25% (p/v) azul bromofenol, 0,25% (p/v) Xylen cyanol, 30 % glicerol.

Solución de transferencia de proteínas: Tris-base glicina pH 7 1X, 20% (v/v) metanol.

Solución Inoue (1lt): 10,88 g MnCl<sub>2</sub>-4H<sub>2</sub>O, 2,2 g CaCl<sub>2</sub>, 18,65 g KCl, 10 ml PIPES (0,5 M pH 6,7), llevar a un litro con agua desionizada. Esterilizar por filtrado.

TBE 5X: 54 g Tris base, 27,5 g ácido bórico, 20 mL 0,5 M EDTA (pH 8.0), llevar a 1 litro con agua desionizada.

PBS-T: PBS, 0,05% (v/v) Tween 20.

Buffer de lisis: NaH<sub>2</sub>PO<sub>4</sub> 50 mM, NaCl 300 mM, Imidazol 10 mM, PMSF 1 mM, Tritón 0,1%

Buffer de unión: NaH<sub>2</sub>PO<sub>4</sub> 50 mM, NaCl 300 mM, Imidazol 10 mM

Buffer de lavado: NaH<sub>2</sub>PO<sub>4</sub> 50 mM, NaCl 300 mM, Imidazol 20 mM/40 mM, PMSF 1 mM

Kanamicina (1000X): 30 mg/ml en H<sub>2</sub>O mq

Ampicilina (1000X): 100 mg/ml en H<sub>2</sub>O mq

Cloranfenicol (1000X): 30 mg/ml en H<sub>2</sub>O mq

Glicerina 10%: 10% (v/v) glicerina, llevando a volumen con agua desionizada

Solución de lisis para ADN: 10 mM Tris-HCl pH 7,5; 100 mM EDTA, SDS 0,1% (p/v)

TE: 10 mM Tris-HCl pH 8, 1 mM EDTA

Tris-glicina 10 x (1 lt): 30,3 g Tris Base, 144 g glicina, llevar a 1 litro con agua desionizada

Tris-glicina-SDS 10X (1 lt): 30,3 gr Tris Base, 144 g glicina, 1% (p/v) SDS, llevar a 1 litro con agua desionizada

SOB (1 lt): 20 gr triptona, 5 gr extracto de levadura, 0,5 gr NaCl, 10 ml KCl 250 mM, llevar a 1 litro con agua desionizada. Autoclavar. Antes de usar agregar 5 ml MgCl<sub>2</sub> 2M.

TSB (1 lt): 17 g peptona de caseína, 3 gr de peptona de harina de soja, 2,5 g de D(+)-glucosa, 5 gr de NaCl, 2,5 K<sub>2</sub>HPO<sub>4</sub>, llevar a 1 litro con agua desionizada. Autoclavar.

L-(+)-Arabinosa 20%: 20 gr L-(+)-Arabinosa, llevar a 100 ml con H<sub>2</sub>O mq. Esterilizar por filtrado.

**Técnicas generales de biología molecular**Purificación de ADN plasmídico

A partir de 3 ml de cultivo de *E. coli* en fase estacionaria, se cosecharon las bacterias por centrifugación a 4000 rpm durante 10 min. Se descartó el sobrenadante y se resuspendió el pellet en 200 µl de solución PI. Se adicionaron 200 µl de solución PII y se homogeneizó cuidadosamente, provocando la ruptura celular y la liberación y desnaturalización de ADN genómico y plasmídico. Luego de 5 min., se agregaron 200 µl de solución PIII frío, se homogeneizó cuidadosamente, se incubó en hielo durante 10 min. y se centrifugó 10 min. a

14000 rpm. Se descartó el precipitado de ADN genómico y restos celulares, recuperando el sobrenadante (fracción en donde se encuentra presente el ADN plasmídico renaturalizado). Se adicionaron 500 µl de cloroformo:alcohol isoamílico (24:1 v/v) emulsionando las fases mediante vortexeo y se centrifugó a 14000 rpm durante 1 min. Se recuperó la fase acuosa y se repitió la extracción orgánica. A continuación, se agregaron 700 µl de isopropanol, se vortexeó y se centrifugó 20 min. a 14000 rpm para sedimentar el ADN plasmídico. El pellet fue lavado con 500 µl de etanol 70% (v/v) y secado en termobloque a 37 °C. Finalmente, el ADN fue resuspendido en 30 µl de H<sub>2</sub>O mq. Alternativamente, se realizaron purificaciones de ADN plasmídico utilizando kits comerciales, como por ejemplo *Zymo PURE Plasmid Miniprep Kit*, siguiendo las instrucciones del fabricante. En algunos casos, para remover contaminantes y concentrar las preparaciones de ADN plasmídico (o bandas recuperadas a partir de geles de agarosa o productos de PCR, ver debajo), se les adicionó 0,1 vol de NaCl 3 M, 3 vol de etanol 100% frío, y se los incubó ON a -20 °C. Posteriormente, se precipitó el ADN por centrifugación a 12000 rpm en frío, se lavó con etanol 70% y finalmente se resuspendió en H<sub>2</sub>O mq.

#### Purificación de ADN genómico

A partir de 5 ml de cultivo de *E. coli* o *S. Typhimurium* en fase estacionaria, se cosecharon las bacterias por centrifugación a 10000 x *g* durante 1 min. Luego, se empleó Quick-DNA™ Miniprep Kit de Zymo para purificar el ADN genómico, siguiendo el protocolo designado para “Suspensiones celulares y muestras digeridas con proteinasa K”.

#### PCR

Las reacciones de PCR fueron llevadas a cabo con Taq ADN polimerasa (ThermoFisher), en las soluciones provistas por los fabricantes con 0,01-1 ng de ADN plasmídico o 10-100 ng de ADN genómico como molde, 2 mM de MgCl<sub>2</sub>, 100 µM de dNTPs y 10 pmol de cada oligonucleótido en un vol final de 50 µl. Las secuencias y características de los distintos oligonucleótidos usados en este trabajo se muestran en la S Tabla 1. Las amplificaciones se realizaron de acuerdo al siguiente esquema:

- 1) Desnaturalización inicial: 10 min. a 95 °C
- 2) 35 ciclos:
  - Desnaturalización: 30 seg. a 95 °C
  - Hibridación de oligonucleótidos: 30 seg. a la temperatura indicada para cada reacción
  - Extensión: 1 min. por cada Kpb a copiar a 72 °C
- 3) Extensión final: 10 min. a 72 °C

Se utilizó el mismo protocolo para realizar *colony* PCR, con la diferencia que el ADN molde provino de una colonia de bacterias.

#### Electroforesis de ADN

Los geles de agarosa se prepararon en TBE 0,5X al 0,8-1,5% (p/v) con 0,5 µg/ml de bromuro de etidio. Las corridas electroforéticas se realizaron con una corriente constante de 100 V por 30-60 min. Antes de la siembra, a las muestras de ADN se les adicionó el volumen necesario de LB6x. Como marcadores de peso molecular se utilizaron uno de 100 pb y otro de 1000 pb (1



kpb) (ambos de PB-L). Luego de la corrida electroforética, los geles fueron visualizados en un transiluminador UV y digitalizados. Para recuperar bandas de ADN a partir de geles de agarosa, se usó el kit comercial QIAEX II Gel Extraction Kit siguiendo indicaciones del fabricante.

#### Digestión de ADN con enzimas de restricción

Se utilizó 1 U de enzima por pmol de ADN a digerir, incubando en las condiciones sugeridas por el fabricante, durante un mínimo de 3 horas, en un volumen final de 50  $\mu$ l. Los volúmenes utilizados para una digestión ejemplo usando 2 enzimas distintas se detallan en la Tabla 2.

<b>Tabla 2: Condiciones generales para reacciones de digestión enzimática de ADN.</b>	
<b>Reactivo</b>	<b>Cantidad/Volumen</b>
Buffer 10X (compatible para enzimas 1 y 2)	5 $\mu$ l
ADN	~ 1 pmol
Enzima 1	1 U
Enzima 2	1 U
H <sub>2</sub> O mq	c.s.p 50 $\mu$ l

#### Ligación

Para las reacciones de ligación se utilizó T4 ADN ligasa (PBL), según especificaciones del fabricante (Tabla 3) y una relación molar inserto:vector de 3:1. Las incubaciones fueron a 16 °C ON, y se transformó con 5  $\mu$ l de reacción en 100  $\mu$ l de células competentes. La ligación de productos de PCR al vector pGEM-T Easy se realizó de acuerdo con lo descrito [53].

<b>Tabla 3: Condiciones generales para reacciones de ligación de ADN.</b>	
<b>Reactivo</b>	<b>Cantidad/Volumen</b>
T4 ADN Ligasa Buffer 10X	1 $\mu$ l
ADN Vector	0,01 pmol
ADN inserto	0,01 pmol
T4 ADN Ligasa	1 - 3 U
H <sub>2</sub> O mq	c.s.p 10 $\mu$ l

#### Preparación de bacterias competentes

Para la preparación de bacterias *E. coli* competentes (cepa DH5 $\alpha$ ) se utilizó el método Inoue [54]. Brevemente, se seleccionó una colonia de una placa crecida ON a 37 °C y se la inoculó en 25 ml de SOB. Se incubó durante 8 horas a 37 °C con agitación (250 rpm) y luego se inocularon 4 ml de este cultivo en un erlenmeyer conteniendo 250 ml de SOB, al que se incubó a 18 °C con agitación ON. Al alcanzar una D.O. a 600 nm de 0,55, se transfirió el cultivo a un baño con hielo por 10 min. Se cosecharon las células por centrifugación a 2500 x *g* durante 10 min. a 4 °C y se resuspendieron suavemente en 80 ml de solución Inoue (esterilizado por filtrado, previamente enfriado). Nuevamente se centrifugaron las bacterias a 2500 x *g* durante 10 min. a 4°C y el

pellet fue suavemente resuspendido en 20 ml de solución Inoue fría. Se agregaron 1,5 ml de DMSO y las bacterias fueron alicuotadas y almacenadas a -70 °C hasta su utilización.

#### Transformación por *heat shock*

Se incubaron 100 µl de bacterias competentes con 2-10 ng del plásmido de interés durante 15 min. en hielo y luego se las sometió a un “shock térmico” a 42 °C durante 90 seg. Las bacterias fueron recuperadas a 37 °C en 1 ml de medio LB durante 1 hora, centrifugadas a 4000 rpm durante 5 min. y cultivadas en placas de Petri con agar 1,5% (p/v) en medio LB suplementado con el/los antibiótico/s correspondiente/s.

### **Técnicas bacteriológicas**

#### Cultivo de *S. Typhimurium*

Para todos los ensayos, las bacterias se cultivaron a 37° C en agitación (200 rpm). Cuando fue necesario, se agregaron antibióticos en las siguientes concentraciones: ampicilina, 100 µg/ml; kanamicina, 50 µg/ml; y cloranfenicol, 30 µg/ml [55].

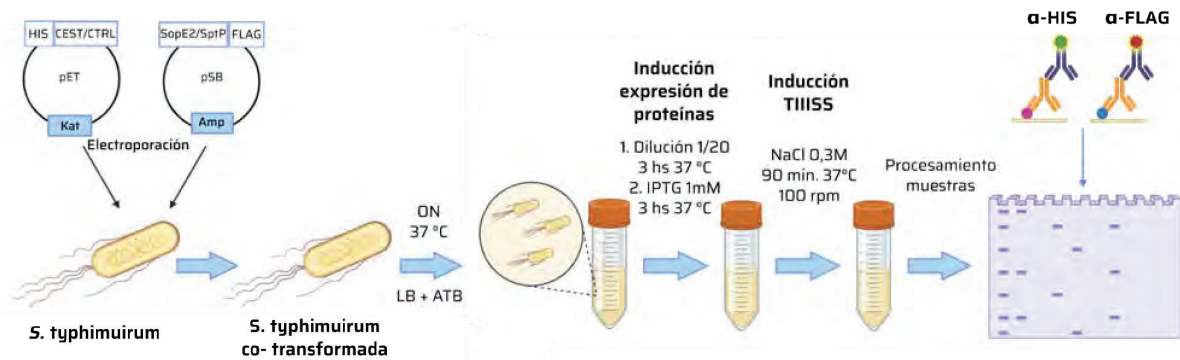
#### Preparación y electroporación de *S. Typhimurium*

Para la preparación de *S. Typhimurium* electrocompetentes se seleccionó una colonia de una placa previamente incubada durante 16-20 horas a 37 °C para inocular 5 ml de medio LB suplementado con los antibióticos apropiados. Se incubó el cultivo ON a 37 °C con agitación (200 rpm), se realizó una dilución 1/100 y se incubó a 37 °C con agitación hasta alcanzar una D.O. a 600 nm de entre 0,4 y 0,6. Se cosecharon las células por centrifugación a 6000 rpm durante 15 min. a 4 °C. Se descartó el sobrenadante, se lavó el pellet con 50 ml de H<sub>2</sub>O mq estéril, y se realizaron dos lavados adicionales con 10 ml de glicerina 10% (v/v) estéril. Finalmente, se lo resuspendió en 300 µl de glicerina 10% (v/v) estéril y se fraccionó en alícuotas de 50 µl. Cada alícuota se electroporó con 200 ng de ADN en cubetas de 0,1 cm con un pulso eléctrico de 2,5 kV. Inmediatamente después, se agregó 1 ml de medio LB y se recuperaron las células por 1 hora, para luego plaquearlas en medio sólido con el/los antibiótico/s apropiado/s.

#### Análisis de secreción de *S. Typhimurium in vitro*

Bacterias co-transformadas (con un ‘plásmido reportero’ y un ‘plásmido chaperona’), seleccionadas con ampicilina y kanamicina, fueron sometidas a un ensayo de inducción del TIISS. Se realizó una dilución 1:20 de cultivos crecidos ON en agitación y luego se indujo la expresión del ‘plásmido chaperona’ mediante la adición de IPTG (1 mM), por 3 horas a 37 °C. A continuación, para promover la expresión y el ensamblaje del TIISS, los cultivos (de 5 ml) se les adicionó NaCl (hasta 0,3 M final), y se crecieron a 37 °C en agitación suave (100 rpm) hasta una densidad óptica de 1 a 600 nm (aproximadamente 90 min.), tal cual lo reportado [33]. Los cultivos se centrifugaron a 14000 rpm durante 20 min. a 4 °C, para separar pellets de sobrenadantes. Los pellets se lavaron con 1 ml de buffer PBS y se lisaron por calentamiento a 100°C durante 10 min. (previa adición de 150 µl de CB 1X) . A los sobrenadantes, conteniendo las proteínas secretadas, se les agregó ácido tricloroacético (hasta 10% v/v), se los incubó ON a 4°C y se los centrifugó a 14000 rpm por 20 min. a 4 °C. Los precipitados de proteínas se lavaron con 1 ml de acetona, se resuspendieron en 25 µl de PBS y se les agregaron 8,3 µl de CB 4X. Las

muestras de pellets y sobrenadantes se resolvieron en geles SDS-PAGE 15% y 10%, respectivamente, y se procesaron por Western blot. Un esquema de los ensayos de complementación funcional se muestra en la Figura 9.



**Figura 9. Esquema de la metodología empleada para los ensayos de secreción en *S. Typhimurium*.**

#### Generación de *S. Typhimurium* KO utilizando el sistema $\lambda$ Red

El sistema  $\lambda$  Red, basado en el bacteriófago  $\lambda$ , permite la recombinación y modificación del ADN en bacterias, y está compuesto por tres proteínas: exo, beta y gam. Gam inhibe las nucleasas RecBCD y previene la degradación de los fragmentos lineales de ADN [56]. Exo es una exonucleasa que genera extremos 3' simple cadena a partir de los fragmentos lineales de ADN, y Bet cataliza la alineación entre el fragmento insertado y la secuencia blanco [57]. *S. Typhimurium* WT electrocompetentes fueron electroporadas con el plásmido pKD46 [20], un vector de expresión regulado por el promotor pAraBad, que es termosensible y contiene el sistema  $\lambda$  Red. Las bacterias fueron seleccionadas por Amp a 30 °C y sometidas a un nuevo protocolo de electrocompetencia. Para la inducción de la expresión del sistema  $\lambda$  Red durante este protocolo, una vez que los cultivos alcanzaron una D.O. a 600 nm de 0,4-0,6 se les agregó L-arabinosa a una concentración final 0,2% p/v y se incubó a 37 °C con agitación por 1 hora [58]. Las bacterias generadas fueron electroporadas con el alelo SpaK mutante y las recombinantes seleccionadas en placas de LB con cloranfenicol. Las colonias CAT<sup>R</sup> fueron plaqueadas nuevamente en LB sin antibióticos e incubadas a 42 °C O.N, lo que permite eliminar el plásmido termosensible pKD46.

#### **Técnicas analíticas de proteínas**

##### Electroforesis de proteínas en geles de poliacrilamida desnaturalizante (SDS-PAGE)

Las muestras se adicionaron con CB (hasta 1X final) y se las calentó por 10 min. a 100 °C. Se sembraron en SDS-PAGE discontinuos (0,1% SDS; 12-15% acrilamida) y se corrieron en solución de Tris-Glicina-SDS a 120 V por 90 min. En todos los casos, se utilizaron marcadores de peso molecular preteñidos de rango 17-200 kDa (PageRuler, ThermoFisher).

##### Ensayos de Western blot

Los geles fueron transferidos a membranas de nitrocelulosa (Pierce) durante 2 horas a 0,2 A en solución de transferencia. La correcta transferencia se verificó mediante la tinción con solución de Rojo Ponceau. Las membranas fueron bloqueadas en PBS-T 0,05% con 1% (m/v) de leche en

polvo descremada durante 1 hora (u ON a 4 °C), e incubadas con el anticuerpo primario diluido 1:5000 en solución de bloqueo por 1 hora. Luego de 3 lavados con PBS-T (10 min. cada uno), las membranas se incubaron con el anticuerpo secundario diluido 1:20000 en solución de bloqueo durante 1 hora, seguido de 3 lavados con PBS-T, y un último lavado en PBS. Todos estos pasos fueron realizados con agitación, a temperatura ambiente. Se escaneó y digitalizó la membrana utilizando el escáner LI-COR Odyssey CLx.

- Anticuerpos utilizados
- $\alpha$ -FLAG: Anticuerpo monoclonal (mAb) de ratón (Invitrogen) reactivo para el epítopo "FLAG".
- $\alpha$ -6xHis: Anticuerpo monoclonal de ratón, cedido gentilmente por la Dra. Daniela Castillo.
- $\alpha$ -ecotin: Anticuerpo policlonal de ratón, cedido gentilmente por el Lic. Lucas Saposnik, del laboratorio de Inmunología, Enfermedades Infecciosas y Desarrollo de Vacunas del IIBio.
- Anticuerpo secundario: IRDye® 800 CW de cabra anti-ratón (LI-COR®)

#### Ensayos de pull-down

Se utilizó una colonia aislada de *S. Typhimurium* co-transformada con un 'plásmido reportero' y un 'plásmido chaperona' para inocular 5 ml de LB (suplementado con ampicilina, kanamicina y cloranfenicol). Al otro día, se realizó una dilución 1:20 del cultivo en 10 ml LB suplementado con antibióticos, se dejó crecer por 3 hs a 37° C, y luego se indujo la expresión del 'plásmido chaperona' mediante la adición de IPTG (1 mM). Luego de 3 hs se centrifugó el cultivo por 10 min. a 4400 rpm, se resuspendió en 2 ml de buffer de lisis y se sonicó por 3 ciclos de 1 min. encendido /1 min. apagado en hielo. Para clarificar la muestra se centrifugó el cultivo durante 30 min. a 12000 rpm a 4 °C. Paralelamente, se equilibraron 25  $\mu$ l de una resina Ni-Sefarosa (Ni-activated Sepharose 4 Fast Flow, Cytiva) con buffer de unión. Una vez centrifugada la muestra, se separaron 60  $\mu$ l (fracción *Input*) y el resto se incubó con la resina equilibrada ON a 4 °C en agitación suave. Al día siguiente se centrifugó la muestra 5 min. a 500 rpm para separar el sobrenadante (fracción *Unbound*) de la resina que se sometió a 5 lavados con 500  $\mu$ l de buffer de lavado durante 5 min. a 4 °C con agitación suave. Una vez terminado el último lavado, se centrifugó la muestra en las mismas condiciones que antes y se resuspendió la resina (fracción *Bound*) en 50  $\mu$ l de CB 2x. Las muestras de *Input* y *Unbound* se prepararon mediante el agregado de 20  $\mu$ l CB 4X para alcanzar una concentración de CB 1X. Todas las muestras fueron crackeadas por 10 min. fueron sembradas en geles SDS-PAGE 15 %, y se procesaron por Western blot.

## Resultados

### Capítulo 1: Caracterización de una familia de genes con dominios CEST en kinetoplástidos

#### Rastreo de secuencias

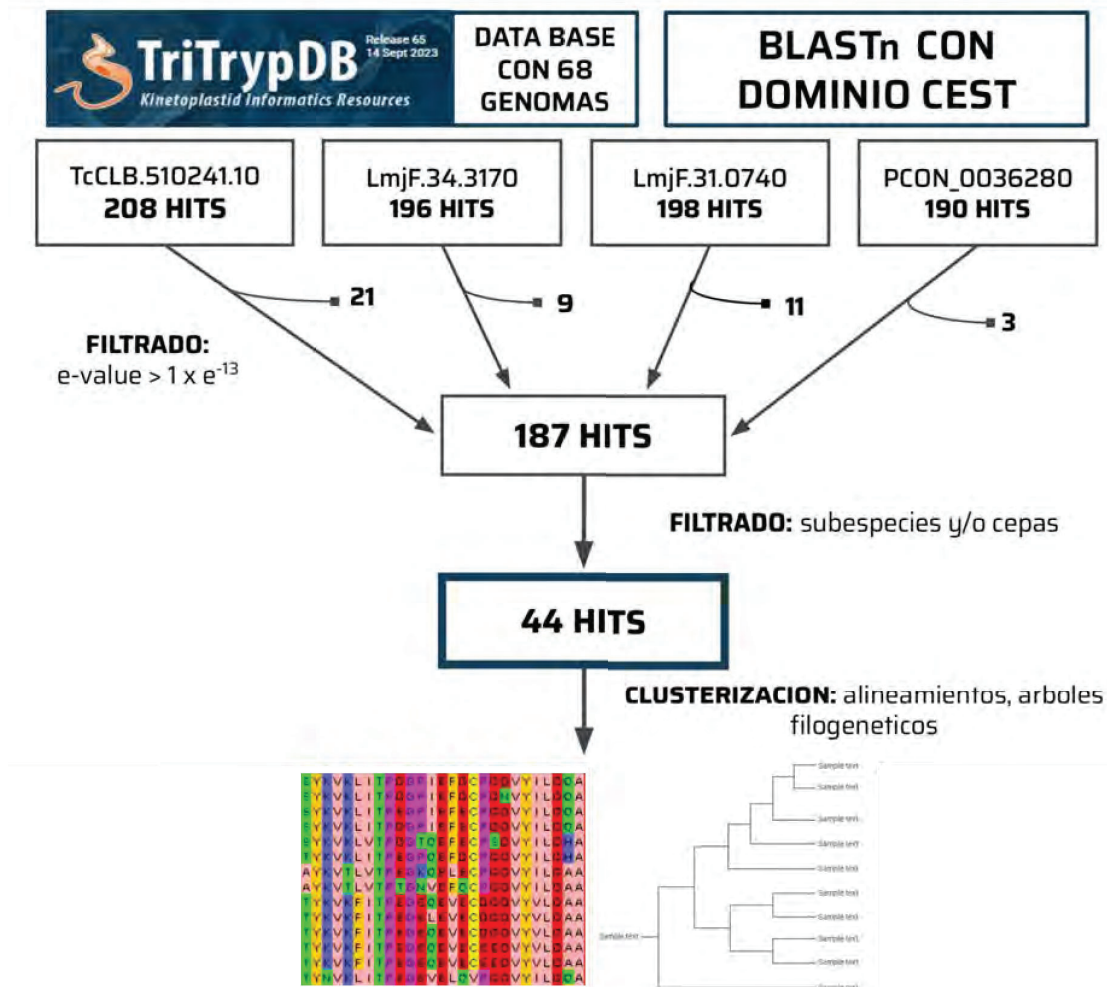
Se usó la secuencia nucleotídica del dominio CEST de TCLP1 (TcCLB.510241.10) como señuelo para la búsqueda de secuencias homólogas en los genomas de kinetoplástidos compilados en la base de datos de acceso público de TritypDB [8]. Esta búsqueda resultó en la identificación de 208 *hits*, distribuidos en el 100% (68/68) de los genomas analizados (Fig 10). Para capturar la mayor diversidad posible, se hizo un alineamiento múltiple con las secuencias identificadas en esta primera búsqueda, se seleccionaron las más divergentes (LmjF.34.3170 y LmjF.31.0740 de *L. major* y PCON\_0036280 de *P. confusum*) y se las usó como señuelo para hacer 3 rastreos independientes sobre las mismas bases de datos y en las mismas condiciones. La integración de estas búsquedas devolvió un total de 187 *hits* redundantes (identificados por más de un señuelo) y 44 *hits* únicos (señuelo-específicos) (Fig 10). Luego de una inspección manual, estos últimos fueron descartados porque presentaron un *e-value* mayor (en el orden de  $1 \times e^{-1}$ ) y una identidad de secuencia menor que los *hits* redundantes (Fig 10). Estos *hits* señuelo-específicos, además, mostraban ser secuencias muy divergentes entre sí. Por último, se realizó un rastreo adicional, similar a los descritos arriba, usando el CEST de TCLP1 como señuelo sobre genomas de Euglénidos depositados en NCBI [59], con resultados negativos (datos no mostrados).

Los 187 *hits* comunes a las 4 búsquedas fueron luego filtrados por redundancia entre cepas y/o subespecies, dependiendo de la especie. Brevemente, en el caso de especies con múltiples cepas (por ejemplo, *T. cruzi*) se seleccionaron las secuencias correspondientes a la cepa de referencia (CL Brener, PMID: 16020725). En el caso de especies con múltiples subespecies (por ejemplo, *Leptomonas* spp y *Leishmania* spp) se seleccionaron las secuencias correspondientes a la subespecie con mejor resolución genómica (*Leptomonas phyrrocoris* [60] y *L. major* [12]). Luego de estos filtrados, se obtuvieron 44 secuencias finales (S Tabla 2) con las que se realizaron los análisis subsiguientes. Para todas ellas se realizó el curado manual del dominio CEST, tomando como punto de partida la ubicación predicha por PFAM y refinando en base a la inspección detallada de los alineamientos (S Tabla 2 y S Fig 1). La Figura 10 esquematiza el protocolo usado para la identificación y curado de dominios CEST en kinetoplástidos.

#### Análisis filogenéticos

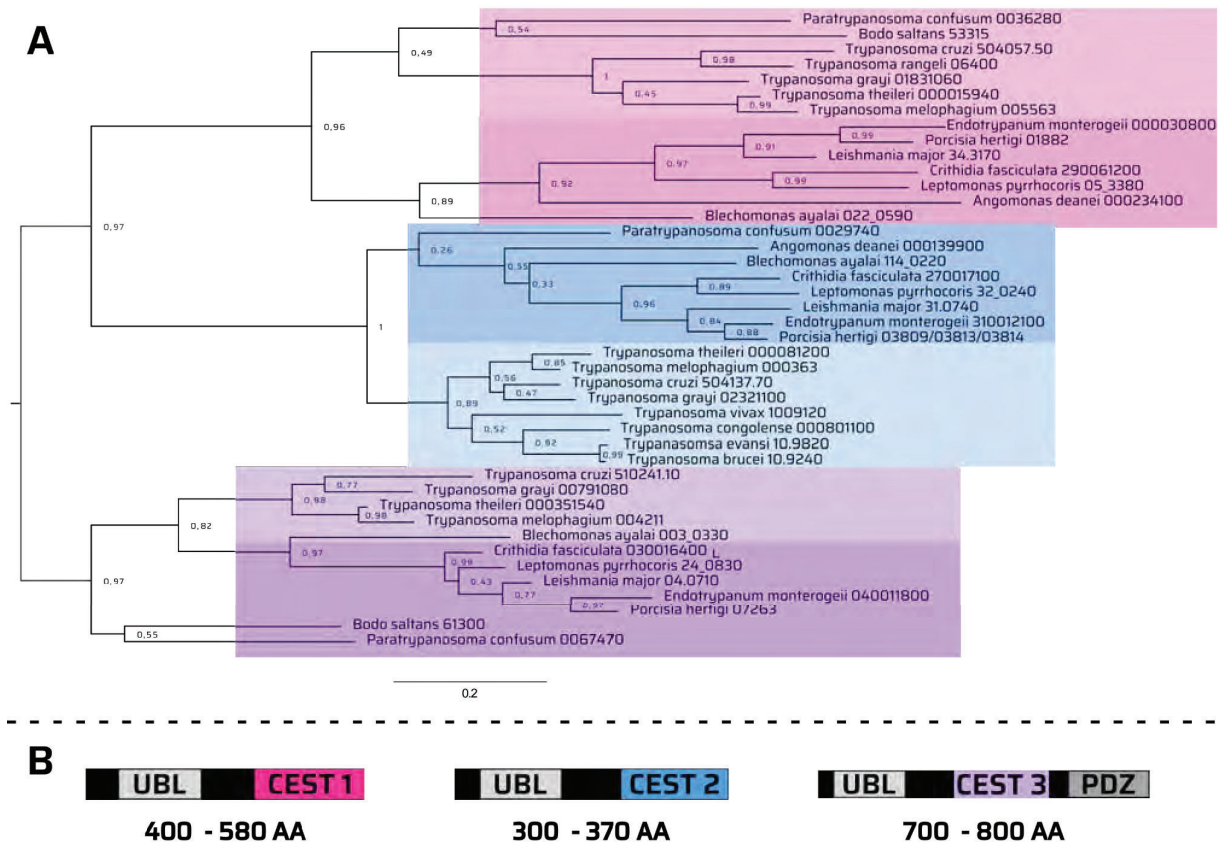
Árboles filogenéticos generados a partir de alineamientos de los dominios CEST seleccionados (S Fig 1) permitieron delinear 3 grupos robustos de homología (llamados CEST 1, 2 y 3), que en principio corresponden a 3 genes distintos (Fig 11). Estos 3 genes, a los que llamamos por simplicidad CEST 1, 2 y 3, tienen distribución diferencial en kinetoplástidos, y para cada uno de ellos pueden distinguirse 2 conjuntos discretos de secuencias en tripanosomátidos: uno que agrupa a las incluidas en las especies del clado 'Trypanosoma' (género *Trypanosoma*) y otro que agrupa a las incluidas en las especies del clado 'Leishmania' (géneros *Leishmania*, *Blechnomonas*, *Crithidia*, *Leptomonas*, *Endotrypanum* y *Porcisia*) (Fig 1). CEST 1 y 3, además, están presentes en *Bodo saltans*, un kinetoplástido fagotrófico de vida libre [61] y en *P. confusum*, un parásito monoxeno recientemente descrito, que constituye la rama más basal de los tripanosomátidos (Fig 11) [13]. CEST 2 es el de distribución más ubicua, con al menos un

representante en todas las especies de tripanosomátidos analizadas, incluyendo *P. confusum*. Interesantemente, no se encontró un representante de CEST 2 en *B. saltans* (Fig 11). Otra particularidad que surge del análisis del árbol filogenético de la Figura 11 es que el grupo salivaria de *Trypanosoma* (tripanosomas transmitidos por la saliva del insecto vector, que contiene especies tales como *T. brucei*, *Trypanosoma evansi* y *Trypanosoma vivax* [63]) presenta solo un gen CEST (CEST 2).



**Figura 10. Esquema del protocolo de obtención de una base de datos curada de las proteínas CEST en kinetoplastidos.** TcCBL: *T. cruzi* CL Brener; LmjF: *L. major* strain Friedlin; PCON: *P. confusum* CUL13.

Si bien existen variaciones lógicas entre alelos de los genes CEST presentes en distintas especies, la estructura general de cada uno de ellos se encuentra muy conservada. Las proteínas CEST 1 y 2 incluyen, además del dominio CEST, un dominio UBL N-terminal, y poseen un tamaño final de 400-580 aa (CEST 1) y de 300-370 aa (CEST 2). Las deducidas a partir de los genes CEST 3, en cambio, incluyen 2 dominios adicionales: un UBL N-terminal y un PDZ en posición C-terminal. Consistente con la ‘anexión’ de este último dominio, las proteínas CEST 3 son las de mayor tamaño promedio (700-800 aa) (Fig 11). Además, los 3 dominios CEST tienen un tamaño similar al de las chaperonas CesT bacterianas (entre 135 y 145 aa) [64], y comparten con ellas un mismo tipo de plegamiento tridimensional (ver Capítulo 2).



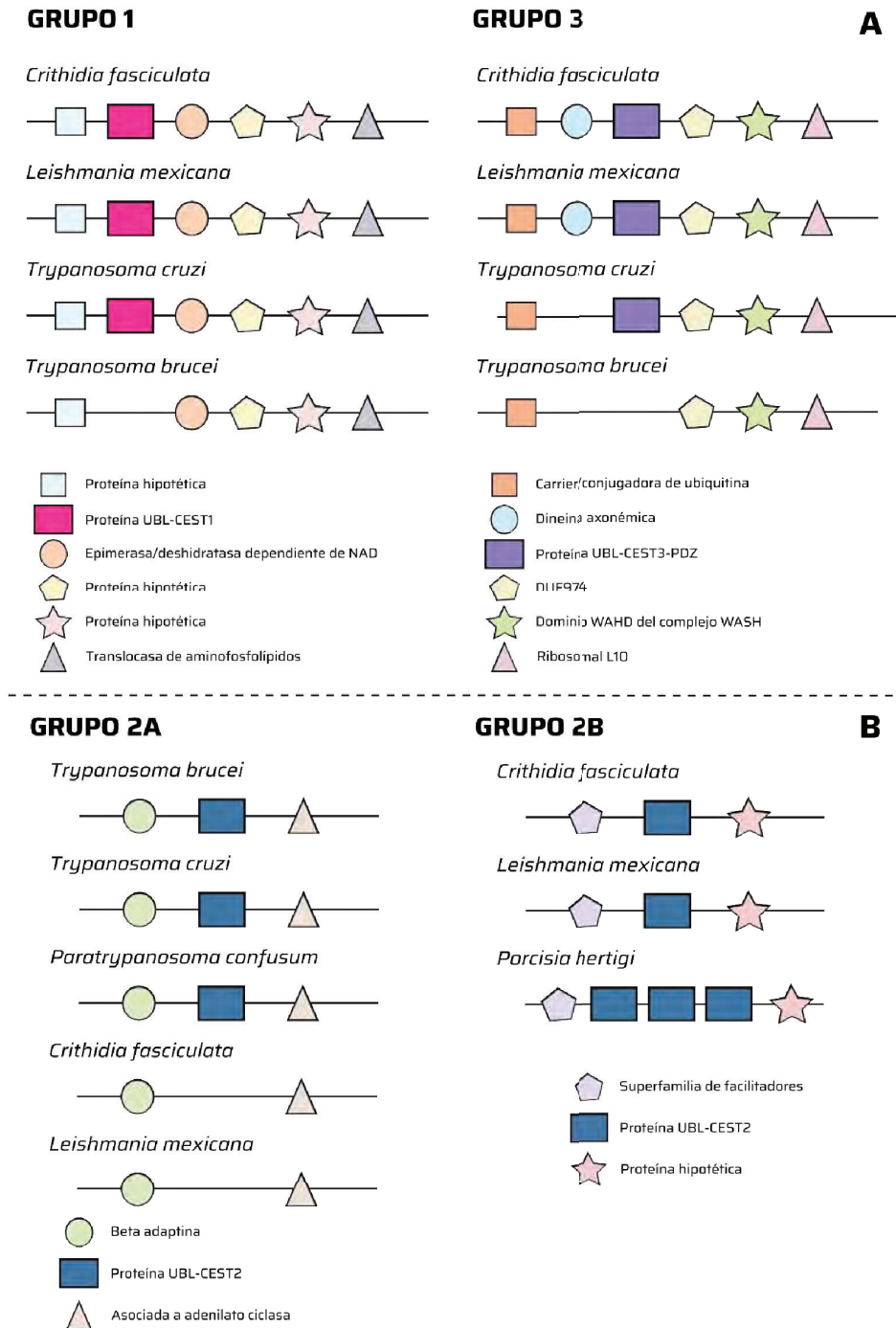
**Figura 11. Relación evolutiva de los genes CEST en kinetoplastidos.** [A] Árbol filogenético no enraizado de dominios CEST. Para cada nodo se indica el valor de bootstrap. Los grupos de homología se indican en rosa (CEST 1), celeste (CEST 2) y violeta (CEST 3). Para cada uno, se indican las ramas correspondientes al clado *Trypanosoma* (claro) y *Leishmania* (oscuro). [B] Esquema a escala de las proteínas CEST 1, 2 y 3.

### Caracterización genómica

Para cada uno de los 3 tipos de genes CEST encontrados, se evaluó la distribución (el contexto genómico del *locus*) y el dosaje (número de copias/haplotipo) en genomas de miembros selectos del clado ‘*Leishmania*’ (*L. mexicana* y *C. fasciculata*) y ‘*Trypanosoma*’ (*T. cruzi* y *T. brucei*). El estudio de los genes flanqueantes indica, por un lado, una localización genómica similar (sinténica) para CEST 1 y CEST 3 en todo el linaje de los tripanosomátidos (Fig 12A). El análisis de estas mismas regiones genómicas en *T. brucei* permite incluso mapear las huellas de la delección de ambos genes en el grupo salivaria (Fig 12A). Más interesante, uno de los genes río abajo de CEST 3 codifica para una enzima conjugadora de ubiquitina. Teniendo en cuenta la transcripción policistrónica en tripanosomátidos [65] y la presencia del dominio UBL en CEST 3, este hallazgo podría tener relevancia funcional.

El gen CEST 2, por otro lado, muestra localización diferencial entre *Trypanosoma* y *Leishmania* (Fig 12B), lo que sugiere eventos independientes de aparición en ambos clados. La región genómica de CEST 2 de *Trypanosoma* pudo ser mapeada en *Leishmania* (tanto en *L. major* como en *C. fasciculata*) (Fig 12B, izquierda). En cambio, el contexto genómico del *locus* CEST 2 de *Leishmania* no pudo ser identificado en *Trypanosoma* (Fig 12B, derecha), sugiriendo

que corresponde a un rearrreglo específico del clado *Leishmania*. Cabe resaltar que *P. confusum*, de localización evolutiva incierta [62], presenta una disposición genómica para CEST 2 similar a la descrita para el clado *Trypanosoma* (Fig 12B, izquierda), aunque su dominio CEST tiene mayor identidad de secuencia con los presentes en los genes CEST 2 del clado *Leishmania* (Fig 11).



**Figura 12. Disposición de los genes CEST en genomas de tripanosomátidos.** Análisis de sintenia para los genes CEST 1 y 3 [panel A] y CEST 2 [panel B]. Para CEST 2 se muestran las regiones genómicas diferenciales en el clado *Trypanosoma* (izquierda) y *Leishmania* (derecha).

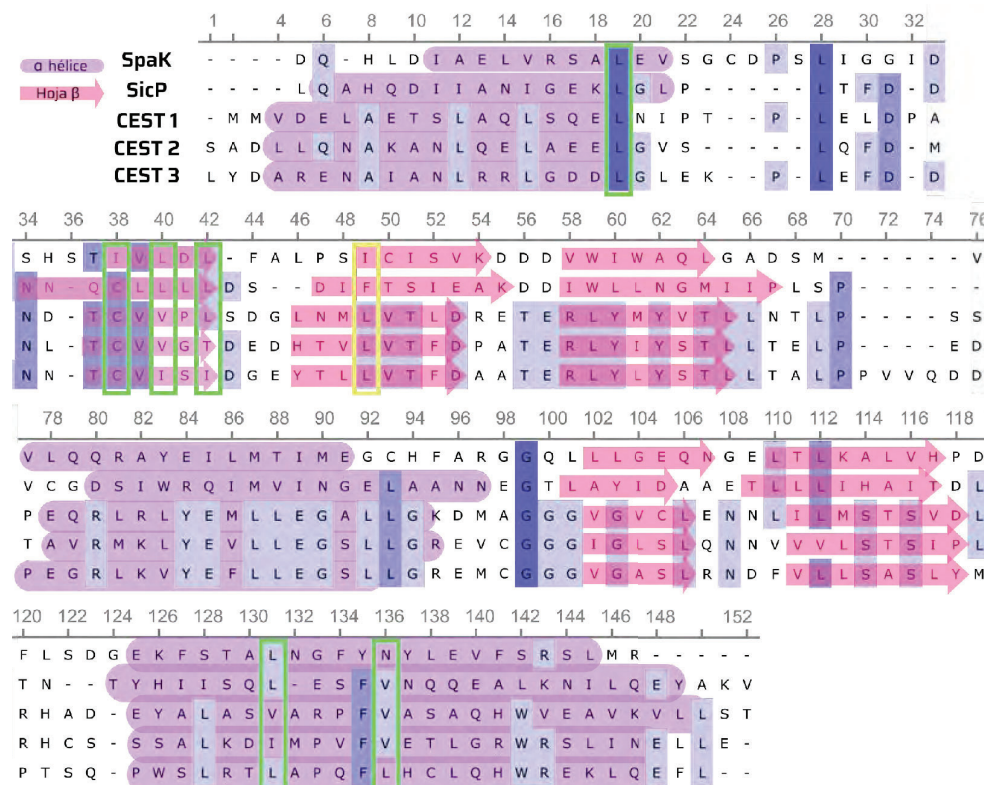


Por último, vale la pena mencionar que *Parcisia hertigi* (un patógeno de puercoespín perteneciente al clado *Leishmania*) presenta 3 copias en tándem del gen CEST 2 (Fig 12B). Estas 3 copias comparten un 96% de identidad entre ellas y constituyen el único caso de amplificación de dosaje que pudimos identificar para los genes CEST de tripanosomátidos. Lamentablemente, este tipo de análisis no pudo replicarse para *B. saltans* por deficiencias en el ensamblado y resolución de su genoma.

## Capítulo 2: Caracterización *in silico* de dominios CEST en tripanosomátidos

### Análisis de la estructura secundaria

Considerando los grupos de homología definidos en el capítulo anterior, se alinearon las secuencias de tripanosomátidos (excluyéndose las de *P. confusum*) de cada uno de ellos (12 de CEST 1, 18 de CEST 2 y 10 de CEST 3, S Tabla 2) y se generaron secuencias consenso, asignándose el residuo más representado en cada posición. Estas secuencias consenso se alinearon con las secuencias completas de SicP y SpaK, ambas de *S. Typhimurium* (Fig 13). Se eligieron estas chaperonas CestT como referencia por la existencia de un cuerpo robusto de estudios estructurales y funcionales, y porque representan modelos de proteínas asociadas a un mismo proceso (secreción de efectores a través del TIISS), pero con propiedades moleculares divergentes: mientras SicP es monoespecífica para el efector SptP, SpaK es multiespecífica, asistiendo en la translocación de distintos factores de virulencia (SopE, SopE2, SopA, etc). A partir de la predicción de JPRED 4 (Fig 13), se observó que, en los 3 casos, estas secuencias consenso, y por inferencia cada una de las 40 secuencias individuales, mantienen conservada la disposición de estructuras secundarias típica de las CestT bacterianas (del N- al C-terminal: 1  $\alpha$  hélice, 3 hojas  $\beta$ , 1  $\alpha$  hélice, 2 hojas  $\beta$ , 1  $\alpha$  hélice). Cada una de estas estructuras secundarias, a su vez, presenta conservación respecto del tamaño (Fig 13).



**Figura 13. Estructura secundaria de dominios CEST.** Alineamiento de las secuencias de aa consenso de los 3 grupos de homología CEST y las chaperonas SicP y SpaK. Se grafican además las predicciones de las estructuras secundarias para cada caso. Notar que las secuencias SpaK y SicP no comienzan con un residuo Metionina (M). En verde se encuentran resaltados los residuos clave para la unión entre SpaK y sus efectores. En amarillo se resalta la posición de la Phe indispensable para la unión de SicP y SptP.

### Selección de secuencias

Para profundizar los estudios estructurales, y para realizar ensayos de caracterización funcional, es necesario trabajar con un número discreto de secuencias individuales. Con este objetivo, priorizamos las 44 proteínas CEST identificadas (S Tabla 2) siguiendo distintos criterios: i) representatividad de los 3 grupos de homología; ii) representatividad de la diversidad de especies en las que fueron encontradas, y iii) disponibilidad de cultivos celulares y/o material genético de esas especies en el laboratorio. Este último punto explica por qué las proteínas seleccionadas de *Leishmania* corresponden a *L. mexicana* y no a *L. major* (subespecie 'de referencia'). Un alineamiento de secuencias mostrando la gran homología entre los dominios CEST de ambas subespecies de *Leishmania* spp se muestra en la S Fig 2. Los dominios finalmente elegidos se encuentran detallados en la Tabla 4.

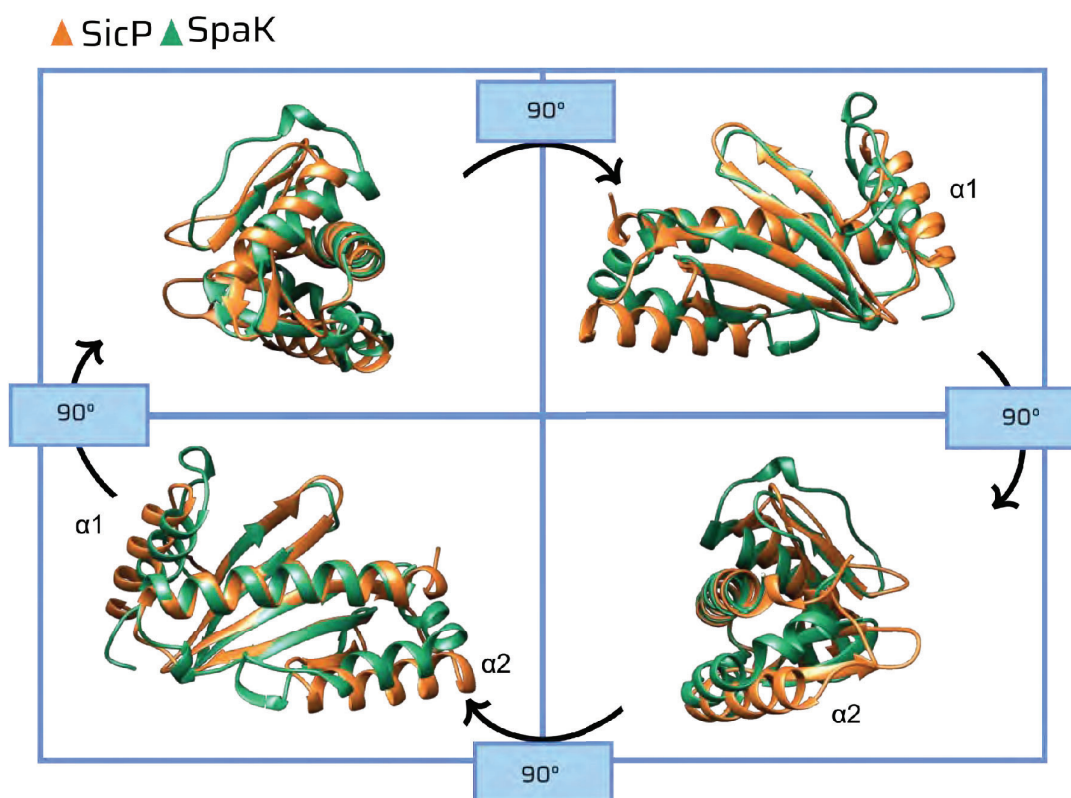
<b>Tabla 4: Proteínas CEST de tripanosomátidos evaluados.</b>			
<b>Dominio CEST</b>	<b>Organismo</b>	<b>Grupo de homología</b>	<b>ID</b>
LmxMG1	<i>Leishmania mexicana</i>	CEST 1	LmxM.33.3170
CfG2	<i>Crithidia fasciculata</i>	CEST 2	CFAC1_270017100
TbG2	<i>Trypanosoma brucei</i>	CEST 2	Tb927.10.9240
TcG2	<i>Trypanosoma cruzi</i>	CEST 2	TcCLB.504137.70
LmxMG3	<i>Leishmania mexicana</i>	CEST 3	LmxM.04.0710
TcG3*	<i>Trypanosoma cruzi</i>	CEST 3	TcCLB.510241.10

\* Corresponde a la misma secuencia TCLP1 definida en ([21]).

### Análisis de la estructura terciaria

Como primera medida, se realizó un alineamiento de las estructuras cristalinas de SicP y SpaK (obtenidas experimentalmente por difracción de rayos X, con una resolución de 1,9 Å [25] y 2.20 Å [34], respectivamente). El análisis de este alineamiento reveló un gran solapamiento entre ambas estructuras, con un RMSD global de 2,196 Å, muy similar al previamente reportado para la comparación de pares de chaperonas CesT de *S. Typhimurium* [66]. En la Fig 14 se muestran distintas perspectivas de este alineamiento. Se destacan las  $\alpha$  hélices 1 y 2 por ser las estructuras que peor alinean, con RMSD locales de 3,347 Å y 3,158 Å, respectivamente (Fig 14).

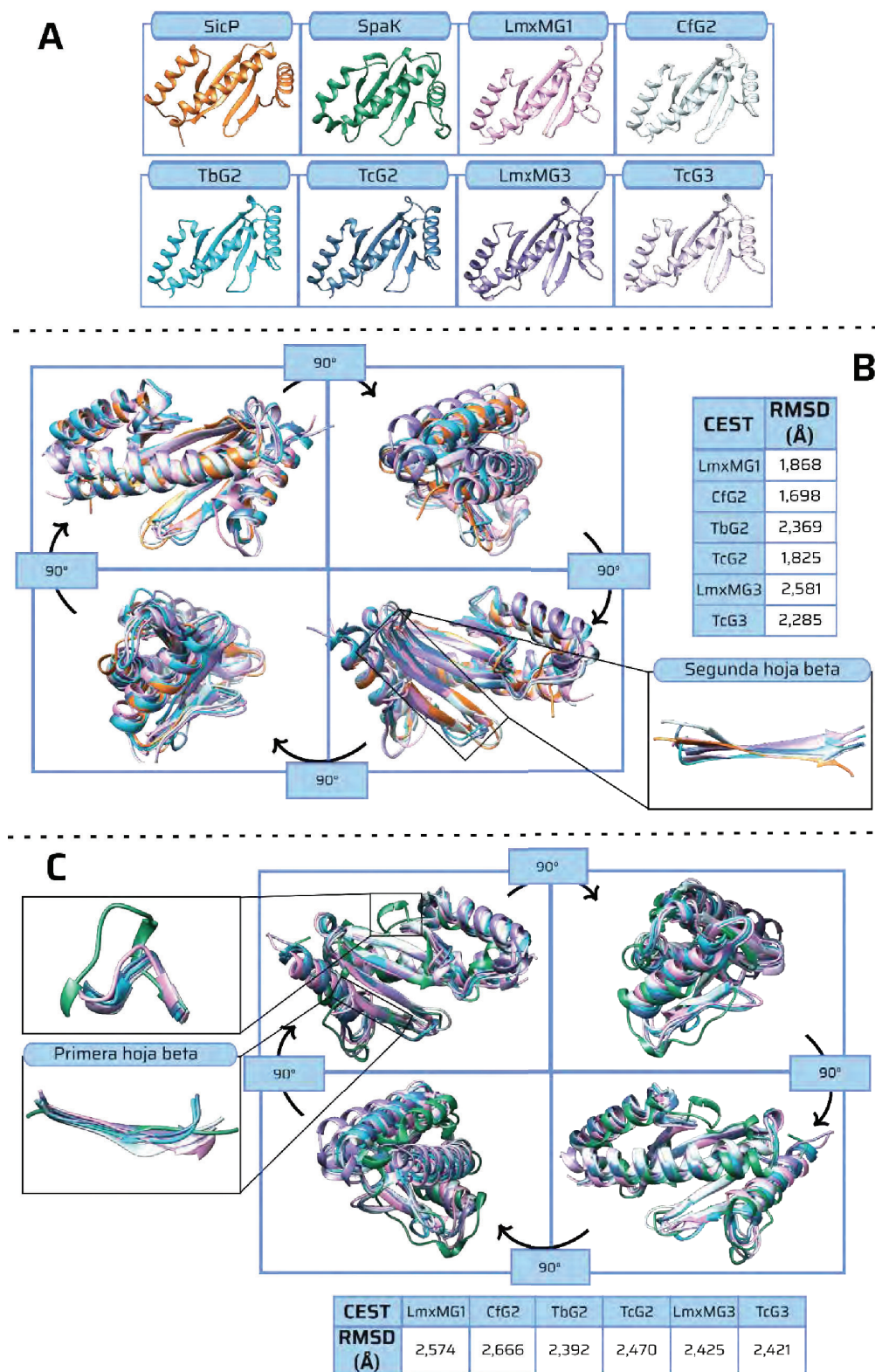
Posteriormente, se emplearon las secuencias curadas de los 6 dominios CEST seleccionados (Tabla 4) para predecir su estructura terciaria mediante AlphaFold2. Como resultado de este ejercicio, se obtuvieron modelos confiables (valores de pLLDT mayores a 70 [52]) para cada uno de los dominios CEST (S Fig 3). El alineamiento de estas estructuras mostró en todos los casos buena superposición tanto con SicP (Fig 15B) como con SpaK (Fig 15C), particularmente entre las hojas  $\beta$  hidrofóbicas internas y las  $\alpha$  hélices externas. Las mayores desviaciones se observaron para las secuencias que conectan entre estas estructuras secundarias. Los valores de RMSD globales para cada dominio CEST respecto de SicP y SpaK se detallan en la Fig 15. Debe mencionarse que los valores correspondientes al alineamiento con SicP se encuentran dentro del rango observado para alineamientos entre chaperonas CesT (1,8 Å a 2,2 Å), mientras que los correspondientes al alineamiento con SpaK son un poco mayores, aunque todavía menores a los 3 Å (valor máximo de RMSD que se obtiene para ortólogos [67]).



**Figura 14. Alineamiento estructural de SicP y SpaK.** Cada panel se corresponde con una vista del alineamiento. Naranja: SicP; Verde: SpaK;  $\alpha$ :  $\alpha$  hélice.

Una estructura sobre la que focalizamos es la segunda hoja  $\beta$  de SicP, en la que se encuentra la Phe involucrada directamente en la unión con el efector SptP (Fig 13) [25]. A la derecha de la Fig 15B se observa un acercamiento hacia dicha región de los modelos apilados. Las segundas hojas  $\beta$  de los dominios CEST, si bien no muestran una Phe o un residuo aromático en posición conservada, se encuentran mayormente alineadas con la de SicP. En la chaperona SpaK, en cambio, la interacción con sus múltiples efectores está mayormente mediada por residuos presentes en la primera hoja  $\beta$  (Fig 13) [34]. Los dominios CEST de tripanosomátidos muestran una gran superposición con dicha región de la proteína SpaK (izquierda de la Fig 15C, panel inferior). Además, debe destacarse que entre la cuarta y quinta hoja  $\beta$ , SpaK presenta un fragmento diferencial al resto de los dominios CEST (y también de SicP), que determina sus mayores valores de RMSD (izquierda de la Fig 15, panel superior).

En conjunto, estos resultados sugieren que, tal lo reportado para TCLP1 [21], los dominios CEST de tripanosomátidos muestran una similitud estructural significativa con chaperonas TIISS de origen bacteriano. Esta similitud estructural es mayor respecto de SicP (y quizás para otras chaperonas monoespecíficas) pero resulta también significativa para chaperonas promiscuas, representadas por SpaK.



**Figura 15. Análisis estructural de dominios CEST.** [A] Estructuras cristalinas de SicP y SpaK y predicción de estructuras terciarias para los dominios CEST. [B y C] Alineamiento estructural de SicP [B] y SpaK [C] con los dominios CEST. A la derecha de [B] se observa un acercamiento hacia la segunda hoja  $\beta$  de SicP. A la izquierda de [C] se observan dos acercamientos: uno hacia la secuencia que conecta la cuarta y quinta hoja  $\beta$  (arriba) y otro hacia la primera hoja  $\beta$  (abajo).

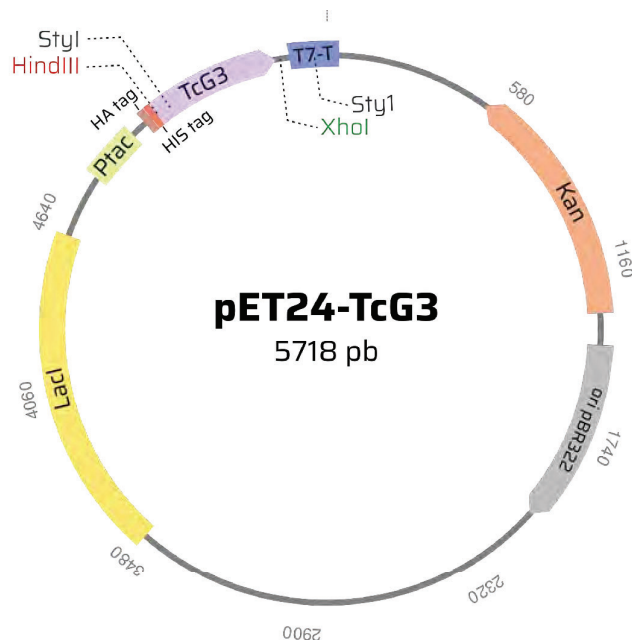
### Capítulo 3: Generación de herramientas para la evaluación funcional de los dominios CEST

Para evaluar funcionalmente a los dominios CEST optamos por realizar experimentos de complementación funcional en 2 modelos de chaperonas asociadas a la secreción de efectores por medio del TIISS de *S. Typhimurium*: SicP/Sptp y SpaK/SopE2-SopE. Para lograr implementar dichos modelos comenzamos por desarrollar las herramientas celulares y moleculares necesarias: cepas de *S. Typhimurium* KO para cada una de estas chaperonas, plásmidos de expresión en *S. Typhimurium* conteniendo los dominios CEST y vectores que permitan monitorear la expresión y localización de los efectores SptP y SopE/SopE2, para usar de reporteros del funcionamiento de uno y otro modelo.

#### Dominios CEST

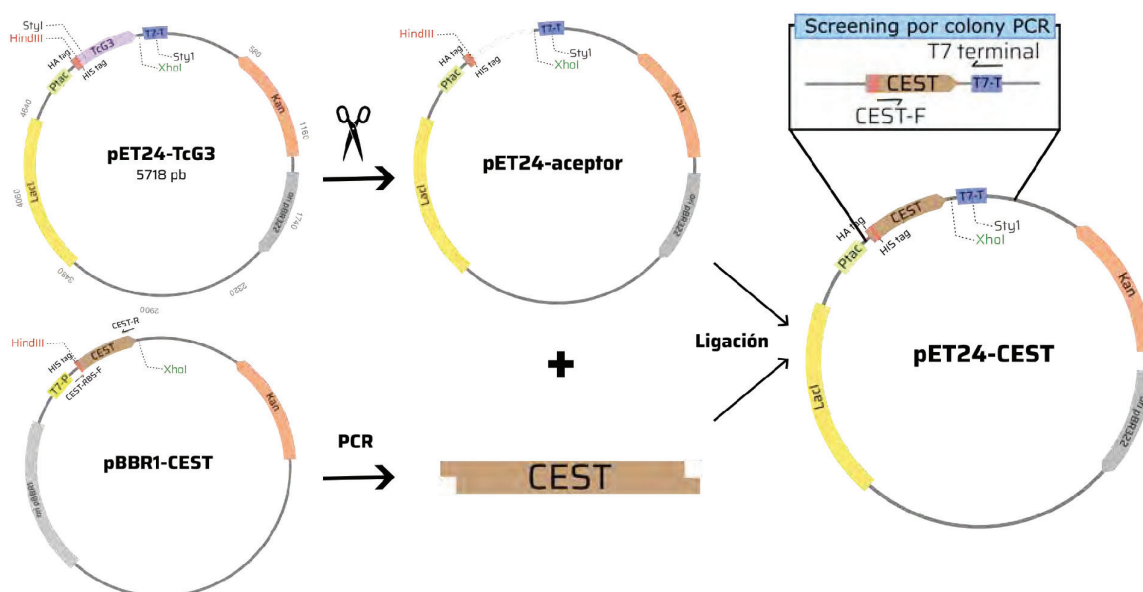
Se diseñó un vector de expresión en bacterias coliformes inducible por IPTG conteniendo el dominio CEST curado de TcG3 (Fig 16) y se solicitó la síntesis a GenScript. El mismo consistió en:

- Promotor tac: híbrido funcional derivado de los promotores trp y lac [68]
- Cassette de resistencia a Kanamicina (Kan)
- Gen lacI: codifica y sobreproduce la proteína represora del sistema Lac
- Sitio iniciador de la traducción (codón ATG, Met). Cabe destacar que, al tratarse de secuencias internas (Fig 11), los dominios CEST de tripanosomátidos no comienzan necesariamente en Met
- Tag 6xHis: situado en el N-terminal del dominio CEST
- Tag HA: situado en el N-terminal del dominio CEST
- Dominio CEST de TcG3 flanqueado por los sitios HindIII (N-terminal) y XhoI (C-terminal)



**Figura 16. Esquema del vector de expresión para los dominios CEST en bacterias.** *Ptac*: promotor tac; *Kan*: gen que otorga resistencia a kanamicina; *ori pBR322*: origen de replicación *pBR322*; *LacI*: gen que codifica para el represor *LacI*. Con línea punteada se indican los sitios de corte para las enzimas de restricción *HindIII*, *XhoI* y *StyI*.

Este vector, llamado pET24-TcG3, fue utilizado para el subclonado de los demás dominios CEST priorizados (Tabla 4) usando el diseño experimental detallado en la Figura 17.



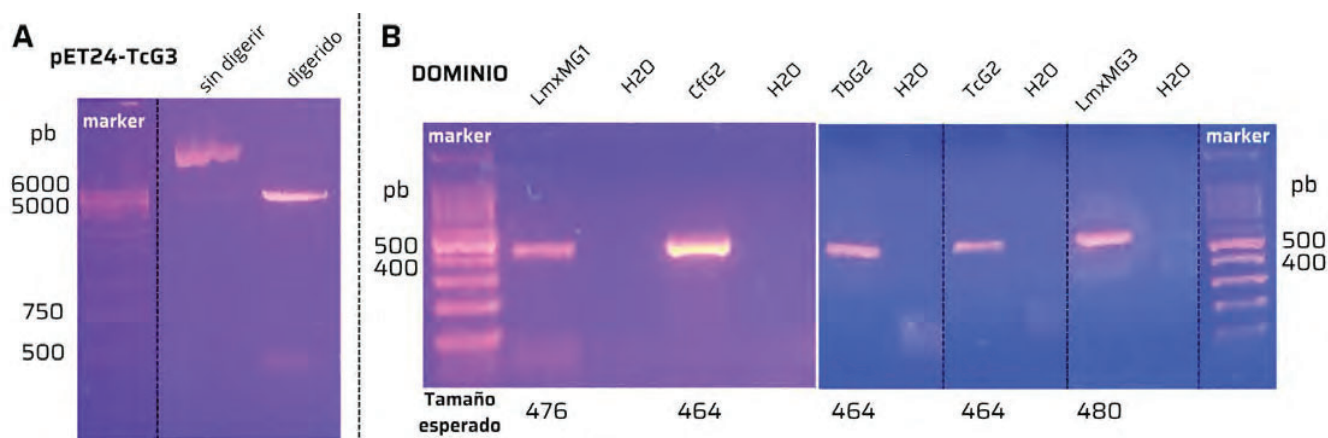
**Figura 17. Diseño experimental para generar vectores de expresión conteniendo los dominios CEST.**

Por un lado, se emplearon los sitios de restricción lindantes al inserto TcG3, HindIII y XhoI, para eliminar este dominio y generar el vector pET24 acepto de dominios CEST (Fig 17, arriba izquierda). Luego de una digestión con ambas enzimas, se resolvieron las 2 bandas obtenidas (Fig 18A), correspondientes al inserto TcG3 (444 bp) y al esqueleto del vector (5274 pb), y se purificó esta última. Por otro lado, se generaron los amplicones conteniendo las secuencias de los dominios CEST priorizados (excepto TcG3) mediante PCR (Fig 17, abajo izquierda), usando de templado los vectores pBBR1-CEST generados previamente en el laboratorio [69] (S Fig 4).

El oligonucleótido *forward* (CEST-F) hibrida sobre una región común a todos los vectores pBBR1-CEST, conteniendo parte del sitio de unión a ribosoma (RBS) y el codón de iniciación de la traducción (ATG) (S Fig 4 y S Tabla 1). Los oligonucleótidos *reverse* (CEST-R) se diseñaron *ad hoc* sobre las secuencias de cada dominio CEST y se les incluyó hacia el 5' un codón de finalización de la traducción (TAG) y el sitio de restricción de XhoI (S Fig 4 y S Tabla 1). La temperatura de hibridación fue de 50° C y el tiempo de extensión para todas las reacciones fue de 30 seg. Las bandas correspondientes a cada amplicón, y sus respectivos controles sin templado (H<sub>2</sub>O), se observan en la Fig 18B. El tamaño de los amplicones (entre 464 y 488 pb) coincidió con el esperado en todos los casos. Cada amplicón fue precipitado, sometido a una digestión doble con HindIII y XhoI y clonado en el vector pET24 acepto (Fig 17, derecha).

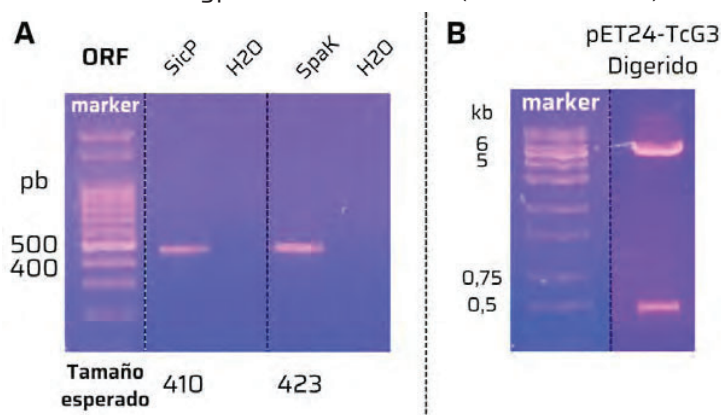
El rastreo de colonias positivas en las distintas ligaciones se hizo mediante *colony* PCR, usando un oligonucleótido *forward* (CEST-específico, que aparea sobre el 5' de cada dominio CEST), y uno *reverse* (común, que aparea sobre el terminador T7, Fig 17). Los clones positivos fueron identificados por movilidad electroforética diferencial (resultados no mostrados) y enviados a secuenciar a Macrogen. Siguiendo esta metodología, se obtuvieron los siguientes

vectores, denominados de ahora en adelante 'plásmidos chaperona': pET24-LmxMG1, pET24-CfG2, pET24-TbG2, pET24-TcG2, pET24-LmxMG3 y pET24-TcG3.



**Figura 18. Clonado de los dominios CEST en pET24.** [A] Vector pET24-TcG3 sin digerir y digerido con HindIII y XhoI, visualizado mediante una corrida electroforética en un gel de agarosa 1%. [B] PCR de los dominios CEST priorizados con sus correspondientes controles sin templado ( $H_2O$ ), visualizados mediante una corrida electroforética en geles de agarosa 1,5%. El marker se corresponde con el marcador de peso molecular 1 kb plus [A] y 100 pb [B].

También se generaron 'plásmidos chaperona' conteniendo CesT bacterianos (pET24-SicP y pET24-SpaK). Para esto se realizaron PCRs sobre ADN genómico de *S. Typhimurium* LT2, y se obtuvo el ORF completo de SpaK y SicP (Fig 19A), los cuales fueron digeridos por HindIII y XhoI (sitios provistos en los oligonucleótidos usados para la amplificación, S Tabla 1) y clonados en el vector pET24 aceptor (previamente digerido con HindIII y XhoI), quedando etiquetados de manera análoga a los CEST de tripanosomátidos (Fig 17). El rastreo de colonias positivas se realizó mediante *colony* PCR (resultados no mostrados). La secuenciación de estos constructos (S Fig 5 y 6), mostró identidad absoluta con los genes SpaK y SicP anotados en *Salmonella enterica* subsp. *enterica* serovar Typhimurium str. LT2 (ID: AE006468).



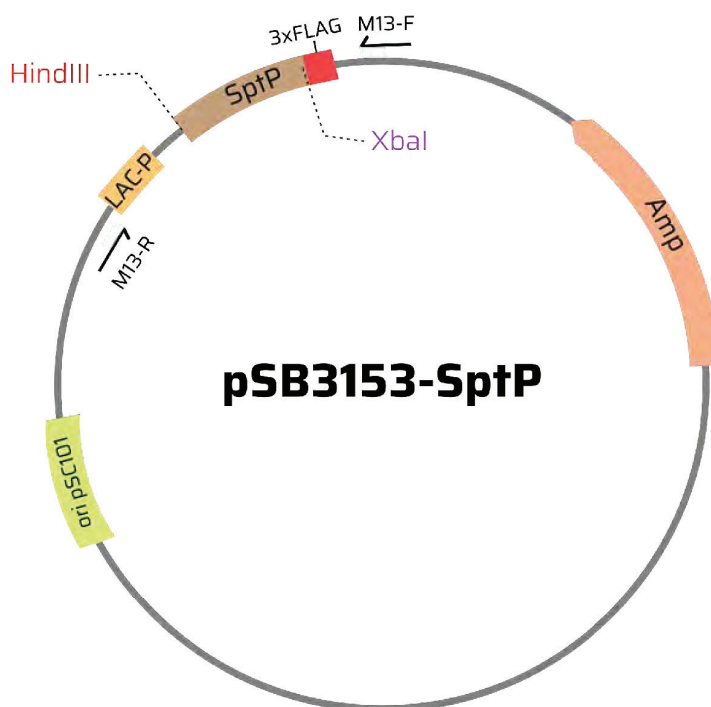
**Figura 19. Generación de controles experimentales en pET24.** [A] PCRs realizadas sobre ADN genómico de *S. Typhimurium* LT2 para obtener SicP y SpaK, con sus correspondientes controles sin templado ( $H_2O$ ), visualizados mediante una corrida electroforética en geles de agarosa 1,5%. [B] Vector pET24-TcG3 digerido con StyI visualizado mediante una corrida electroforética en un gel de agarosa 1%. El marker se corresponde con el marcador de peso molecular de 100 pb [A] y 1 kb plus [B].



Como control adicional, se generó el plásmido pET24 vacío mediante la digestión del plásmido pET24-TcG3 con la enzima *Syl*I (Fig 16). Se resolvieron las bandas correspondientes al inserto (conteniendo la mayor parte de TcG3 y algo de secuencia 3' no codificante, de 486 bp) y al esqueleto del vector (5232 pb) (Fig 19B), y se purificó esta última. Este constructo (flanqueado por sitios *Syl*I) se recircularizó mediante la adición de T4 ADN ligasa. Se realizó un rastreo de colonias positivas mediante sitios de restricción diagnósticos (*Hind*III y *Syl*I, Fig 16) y los clones positivos se secuenciaron.

### Efectores bacterianos

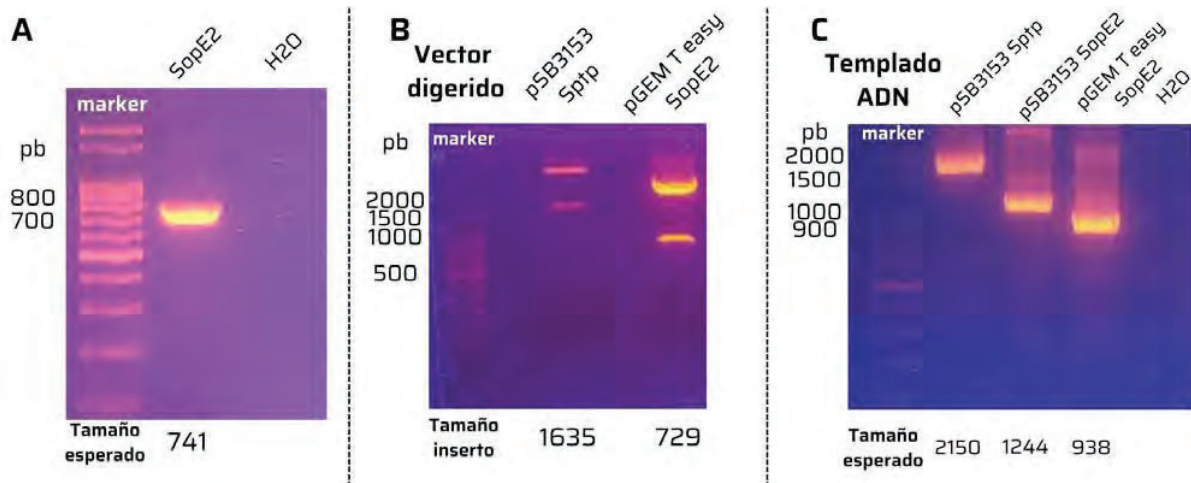
Previamente en el laboratorio se había generado un plásmido que expresa el efector SptP de *S. Typhimurium* [70] (pSB3153-SptP, Fig 20). Este consta del esqueleto de un vector de expresión constitutivo en bacterias coliformes (pSB, Amp<sup>r</sup>), en cuyo MCS se clonó el ORF completo de SptP, con un epitope 3xFLAG fusionado en fase en su C-terminal. En este trabajo, se generó un vector con estructura similar, pero expresando el ORF completo del efector SopE2.



**Figura 20. Esquema del vector pSB3153-SptP.** LAC-P: promotor derivado del operón lactosa; Amp: gen que otorga resistencia a ampicilina; ori pSC101: origen de replicación pSC101. Los sitios de restricción *Hind*III y *Xba*I se indican con líneas punteadas, y los oligonucleótidos, con flechas.

Por un lado, se realizó una PCR utilizando ADN genómico de *S. Typhimurium* LT2 como templado y oligonucleótidos diseñados *ad hoc* (S Tabla 1) para obtener el ORF completo de SopE2 (741 pb, Fig 21A). El amplicón fue precipitado y clonado en el pGEM T easy. Se rastrearon colonias positivas por PCR utilizando los oligonucleótidos universales *M13 forward* y *M13 reverse* (S Tabla 1) y los clones positivos se enviaron a secuenciar. A partir de un clon positivo que mostró un 99,6% identidad respecto del gen SopE2 (ID: AF200952) anotado en

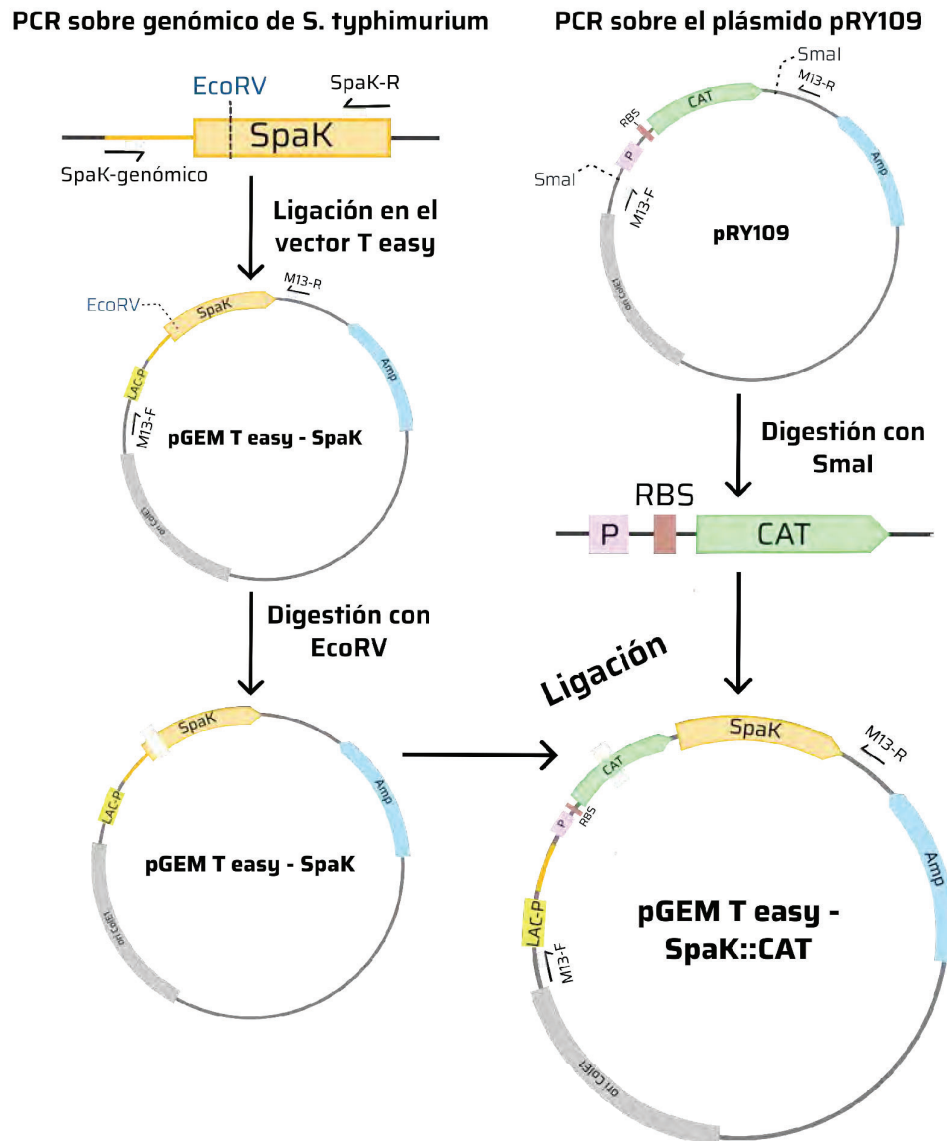
*Salmonella enterica* subsp. *enterica* serovar Typhimurium F98 (S Fig 7) se purificó el plásmido de interés. El mismo fue digerido con HindIII y XbaI (los sitios de restricción son provistos por los oligonucleótidos usados para la amplificación, S Tabla 1), se resolvieron las bandas correspondientes al inserto SopE2 (729 bp) y al esqueleto del vector pGEM T easy (3027 pb) (Fig 21B, calle derecha), y se purificó la primera de acuerdo a lo descrito. Por otro lado, para generar el vector pSB3153 aceptor, se digirió el vector pSB3153-SptP con HindIII y XbaI (Fig 20). Se resolvieron las bandas correspondientes al inserto SptP (1635 bp) y al esqueleto del vector pSB3153 (5688 pb) (Fig 21B, calle izquierda), y se purificó esta última. Luego, se clonó SopE2 en el vector pSB3153 aceptor y se rastrearon colonias positivas mediante *colony* PCR. En la Fig 21C se muestran los productos de amplificación obtenidos utilizando un oligonucleótido *reverse* universal, que pega tanto en pSB3153 como en pGEM T easy y un oligonucleótido *forward* específico sobre diferentes templados: el vector pSB3153-Sptp (2150 pb), el vector pGEM T easy-SopE2 (1244 pb), una colonia proveniente de la placa de ligación correspondiente a pSB3153-SopE2 (938 pb), y un control sin templado (H<sub>2</sub>O). Todos los tamaños coincidieron con los esperados. El set de plásmidos efectores evaluados se completó con el ya mencionado pSB3153-Sptp (Fig 20) y un plásmido pSB3153 conteniendo el CBD del efector SopE (también dependiente de SpaK para su secreción vía TIIIS), etiquetado de igual manera (pSB3153-SopE<sup>1-120</sup>), cedido gentilmente por el Dr. Gabriel Briones [71].



**Figura 21. Generación del plásmido pSB3153-SopE2.** Carridas electroforéticas en geles de agarosa 1,5% (A y C) o 1% (B), con los marcadores de peso molecular 100 pb (A y C) o 1 kb plus (B). [A] PCRs realizadas sobre ADN genómico *S. Typhimurium* LT2 con oligonucleótidos específicos para *SopE2*. [B] Vectores pSB3153-Sptp y pGEM T easy-SopE2 digeridos con HindIII y XbaI. [C] PCRs realizadas para el rastreo de colonias positivas para la ligación de pSB3153-SopE2. Se incluyeron dos controles (pSB3153-Sptp y pGEM T easy-SopE2). H<sub>2</sub>O: control sin templado.

Líneas de *S. Typhimurium* genéticamente deficiente para SpaK

- Cepa LT2

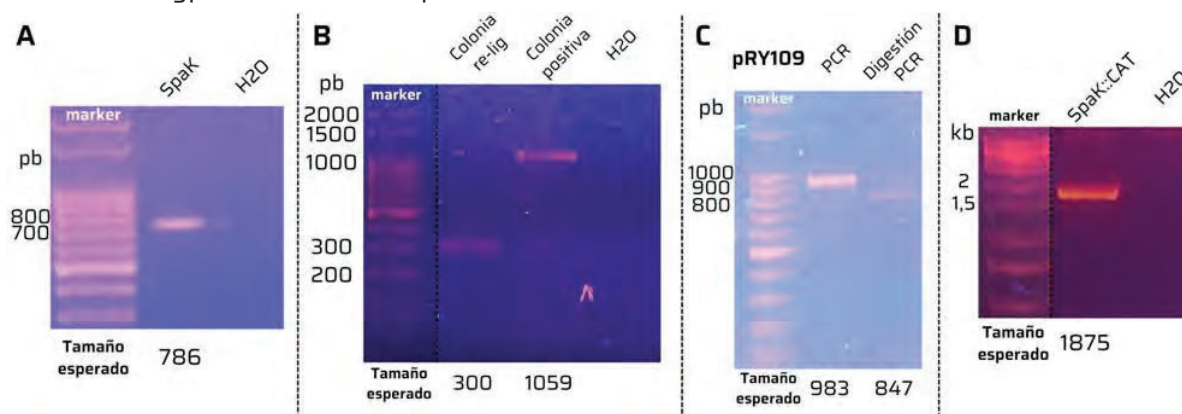


**Figura 22. Esquema de la metodología empleada para generar el alelo *SpaK* mutante.**

Previamente en el laboratorio se había generado una cepa de *S. Typhimurium* LT2 genéticamente deficiente para SicP [70]. En este trabajo, se utilizó una estrategia similar para intentar generar una *S. Typhimurium* LT2 mutante para el gen *SpaK* ( $\Delta$ *SpaK*). En primer lugar, se generó un alelo mutante *SpaK* siguiendo el esquema mostrado en la Fig 22. Brevemente, se amplificó el ORF completo de *SpaK* y parte de sus regiones flanqueantes utilizando como templado ADN genómico de *S. Typhimurium* LT2. El oligonucleótido *forward* aparece 393 pb río abajo del inicio del gen (*SpaK*-genómico), y como oligonucleótido *reverse* se eligió el mismo que se usó para generar el plásmido *pET24-SpaK* (*SpaK-R*) (Fig 22, arriba izquierda y S Tabla 1). El amplicón generado (786 pb, Fig 23A) se clonó en el vector *pGEM T easy* y se rastrearon colonias positivas mediante *colony* PCR (Fig 23B). A partir de un clon positivo, se purificó el plásmido de interés y se lo digirió con *EcoRV*, generando el 'vector aceptor' (Fig 22).

Paralelamente, se realizó una PCR sobre el plásmido pRY109 [72] (S Fig 8) utilizando los oligonucleótidos M13F y M13R para obtener el cassette de resistencia a cloranfenicol (Fig 22, arriba derecha). Se digirió el amplicón con *Sma*I (que al igual que *Eco*RV, genera sitios romos) y se lo ligó al 'vector aceptor' para obtener el plásmido pGEM T easy-SpaK::CAT (Fig 22). El rastreo de colonias positivas se realizó por PCR y los clones positivos se secuenciaron (S Fig 9).

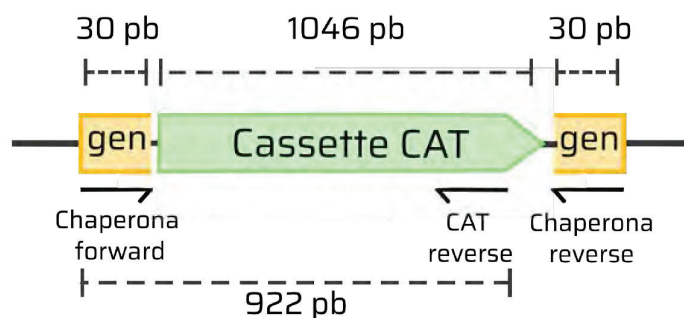
Para generar la recombinación homóloga del alelo mutante SpaK::CAT en *S. Typhimurium* LT2, bacterias WT electrocompetentes fueron primero electroporadas con el plásmido pKD46 [20], un vector de expresión regulado por el promotor pAraBad, que es termosensible y contiene el sistema  $\lambda$  Red. A las bacterias transformadas, seleccionadas en Amp, se les realizó nuevamente un protocolo de electrocompetencia y durante esta etapa se indujo la expresión del sistema  $\lambda$  Red. Las bacterias fueron electroporadas con el alelo SpaK::CAT, el cual se obtuvo mediante PCR sobre el plásmido pGEM T easy-SpaK::CAT con los oligonucleótidos universales M13 *forward* y *reverse* (Fig 23D), se recuperaron durante una hora a 30 °C y se sembraron en una placa de LB complementada con cloranfenicol. Lamentablemente, en sucesivos intentos no creció ninguna colonia, por lo que se procedió a modificar, de a uno por vez, múltiples parámetros del proceso. Brevemente, se probó i) inducir el sistema  $\lambda$  Red con diferentes concentraciones de arabinosa (0,2% p/v y 0,4% p/v); ii) aumentar la masa electroporada del alelo SpaK::CAT, progresivamente desde 200 ng finales hasta 1  $\mu$ g; iii) usar diferentes concentraciones de cloranfenicol en las placas de selección (30  $\mu$ g/ml, 15  $\mu$ g/ml, 7,5  $\mu$ g/ml, 3,75  $\mu$ g/ml); iv) aumentar el tiempo de recuperación posterior a la última electroporación, desde 1 hora hasta 4 horas; v) usar distintos medios de recuperación de las bacterias luego de la electroporación (LB, SOB y TSB); y vi) generar el alelo SpaK::CAT mediante una digestión del plásmido con la enzima *Not*I (presente en el MCS del vector pGEM T easy, a ambos lados del inserto SpaK::CAT) en vez de por PCR. No obstante, ninguno de los cambios permitió obtener bacterias *S. Typhimurium* LT2  $\Delta$ SpaK.



**Figura 23. Generación del alelo mutante SpaK::CAT.** Corridas electroforéticas en geles de agarosa 1,5%, con los marcadores de peso molecular 100 pb (A, B y C) o 1 kb plus (D). [A] PCR realizadas sobre ADN genómico *S. Typhimurium* para obtener SpaK y parte de sus regiones flanqueantes. [B] Colony PCR realizada sobre una colonia de la placa de re-ligado y una colonia de la placa de ligación del vector pGEM T easy-SpaK. [C] En la segunda calle se visualiza el amplicón correspondiente a la PCR realizada sobre el pRY109; en la tercera calle se observa dicha PCR digerida. [D] PCR realizada sobre el plásmido pGEM T easy-SpaK::CAT para obtener el alelo mutante. H<sub>2</sub>O: control sin templado.

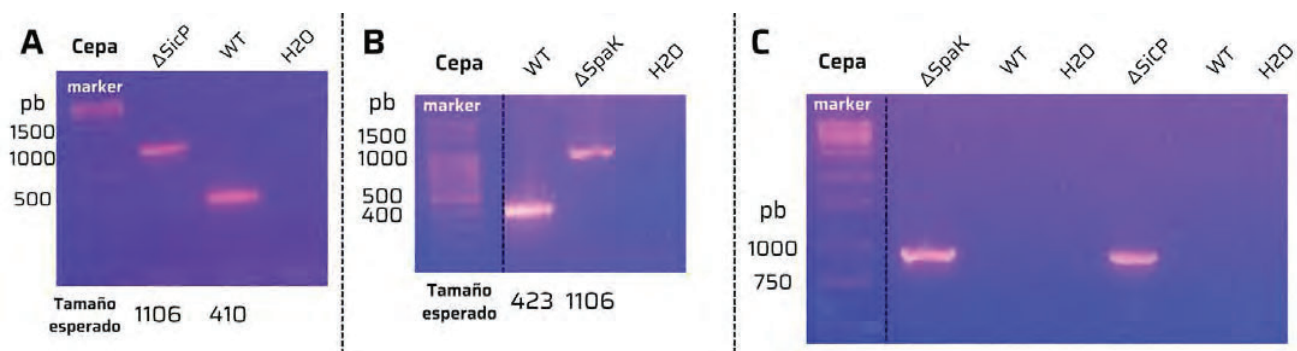
- Cepa 14028S

Como alternativa, se solicitaron bacterias  $\Delta$ SpaK al Dr. Michael McClelland, cuyo laboratorio generó una biblioteca de mutantes *S. Typhimurium*, cepa 14028s, mediante el sistema  $\lambda$  Red [73]. También solicitamos la mutante para SicP ( $\Delta$ SicP), con la idea de evaluar la complementación funcional de los dominios CEST sobre esta chaperona en dos *backgrounds* genéticos distintos (LT2 y 14028s). La estructura de las mutantes de esta biblioteca consiste en 30 pb del 5' del gen blanco, el cassette de resistencia a cloranfenicol, y 30 pb correspondientes al 3' de gen blanco (Fig 24).



**Figura 24. Estructura genómica de las cepas *S. Typhimurium* 14028s mutantes.** El gen blanco de delección se indica en naranja. Se indican los oligonucleótidos utilizados para los análisis genotípicos.

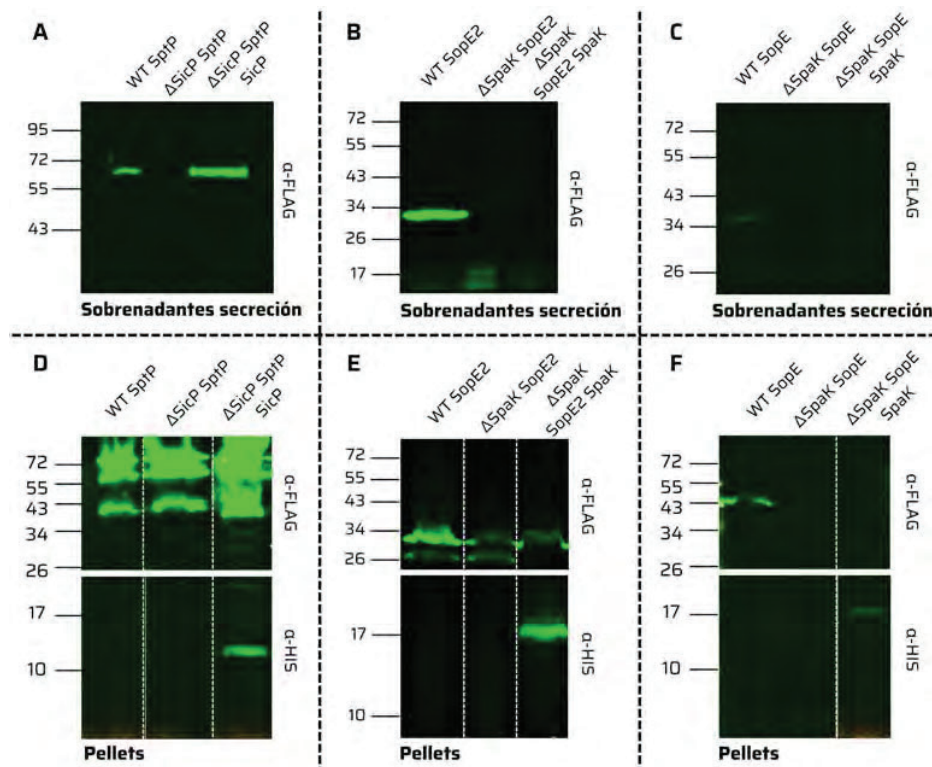
Las mutantes  $\Delta$ SicP y  $\Delta$ SpaK de 14028s se tipificaron mediante PCR, utilizando los oligonucleótidos *Chaperona forward* y *Chaperona reverse* (Fig 24 y S Tabla 1) y ADN genómico purificado de estas líneas. En los paneles A y B de la Fig 25 se visualiza, respectivamente, la diferencia de tamaño entre los amplicones generados a partir de las líneas  $\Delta$ SicP,  $\Delta$ SpaK y WT (usada de control). Además, para completar la caracterización se realizaron amplificaciones adicionales empleando el oligonucleótido *Chaperona forward* en combinación con el oligonucleótido *CAT reverse* (Fig 24 y S Tabla 1). Solo en las bacterias mutantes ( $\Delta$ SicP y  $\Delta$ SpaK) se observa una banda correspondiente a un producto de amplificación, cuyo tamaño se corresponde con el esperado (Fig 25C).



**Figura 25. Caracterización genotípica de las cepas *S. Typhimurium* 14028s mutantes.** Corridas electroforéticas en gels de agarosa 1,5%, con los marcadores de peso molecular 100 pb (A y B) o 1 kb plus (C). PCR realizada sobre ADN genómico de *S. Typhimurium* 14028s  $\Delta$ SpaK,  $\Delta$ SicP o WT, utilizando los oligonucleótidos: [A] SicP específicos, [B] SpaK específicos, [C] *Chaperona forward* y *CAT reverse* (ver Fig 24). H<sub>2</sub>O: control sin templado.

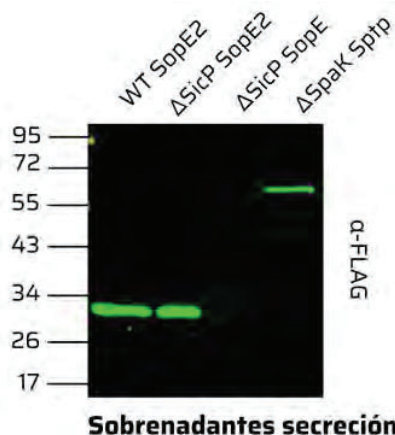
Las mutantes  $\Delta$ SicP y  $\Delta$ SpaK de la cepa 14028s se caracterizaron fenotípicamente realizando un análisis de secreción *in vitro* (Fig 9). Brevemente, los cultivos bacterianos fueron adicionados con NaCl para promover la expresión y el ensamblaje del TIIS y se crecieron a 37°C en agitación suave. Luego de separar los sobrenadantes, se sembraron alícuotas de los pellets en geles SDS-PAGE 15% y se procesaron por Western blot. Para el revelado se utilizaron anticuerpos  $\alpha$ -FLAG para visualizar los efectores (Sptp, SopE2 y SopE) etiquetados con este epitope (Fig 20) y  $\alpha$ -HIS para visualizar las chaperonas (Fig 16). A los sobrenadantes se les agregó TCA para precipitar las proteínas secretadas, se resolvieron alícuotas en geles SDS-PAGE 10% y se procesaron por Western blot, utilizando el mAb  $\alpha$ -FLAG.

Para la caracterización inicial de  $\Delta$ SicP se realizó un análisis de secreción *in vitro* sobre las cepas WT,  $\Delta$ SicP y  $\Delta$ SicP transformada con pET24-SicP. Todas ellas, además, transformadas con pSB3153-SptP. Como se observa en la Fig 26A, tanto la WT como la  $\Delta$ SicP complementada con SicP secretan al medio una banda inmunoreactiva para  $\alpha$ -FLAG, de un tamaño compatible con el esperado para el efector SptP etiquetado con 3xFLAG (~63 kDa). Esta banda no se observa en la mutante  $\Delta$ SicP sin complementar. Como control, se evaluó la secreción de dos efectores independientes de SicP (SopE2 y SopE<sup>1-120</sup>). En los sobrenadantes de secreción de la cepa  $\Delta$ SicP (Fig 27, calles 2 y 3) se observan bandas inmunoreactivas para  $\alpha$ -FLAG de los tamaños esperados, 29 y 34 kDa respectivamente. Estos resultados, en conjunto, indican que la deficiencia en la secreción de la mutante  $\Delta$ SicP es específica para el efector SptP, y que esta deficiencia puede ser revertida por la expresión ectópica de SicP.



**Figura 26. Caracterización fenotípica de las cepas *S. Typhimurium* 14028s mutantes.** Western blot de los sobrenadantes del ensayo de inducción del TIIS revelado con mAb  $\alpha$ -FLAG [A, B, C]. Western blot de los pellets del ensayo de inducción del TIIS, revelado con mAb  $\alpha$ -FLAG (arriba) y antisuero  $\alpha$ -6xHIS (abajo) [D, E, F]. Se indican los marcadores de masa molecular (en kDa) a la izquierda.

Para la caracterización fenotípica de la mutante  $\Delta$ SpaK se procedió de igual manera, aunque las cepas analizadas inicialmente fueron la WT,  $\Delta$ SpaK y  $\Delta$ SpaK transformada con pET24-SpaK. En este caso, todas ellas fueron transformadas con pSB3153-SopE2. Como puede verse en la Fig 26B, la cepa WT, pero no la  $\Delta$ SpaK, secreta al medio una banda inmunoreactiva para  $\alpha$ -FLAG, de un tamaño compatible con el esperado para SopE2 etiquetado con 3xFLAG (~29 kDa). Esta banda, sin embargo, no se observa en la fracción secretada por la mutante  $\Delta$ SpaK complementada con el plásmido pET24-SpaK (Fig 26B). Resultados muy similares se obtuvieron al hacer ensayos de secreción sobre las mismas cepas, ahora transformadas con pSB3153-SopE<sup>1-120</sup> (Fig 26C). Como control de especificidad, se evaluó la secreción de SptP en la mutante para SpaK. Como se muestra en la Fig 27, el sobrenadante de  $\Delta$ SpaK presenta una banda inmunoreactiva para el mAb  $\alpha$ -FLAG que se corresponde con el tamaño esperado para SptP-3xFLAG, indicando que el defecto de secreción observado en esta cepa no era el resultado de un efecto global sobre la secreción mediada por TIIIS. En las fracciones correspondientes a los pellets (geles inferiores de la Fig 26D, E y F) de los diferentes ensayos se visualiza la correcta expresión de las chaperonas evaluadas, como bandas inmunoreactiva para  $\alpha$ -HIS que se corresponden con los tamaños esperados para SicP (~15 kDa) y Spak (~17 kDa). La presencia de los distintos efectores en el citoplasma de las bacterias, independientemente de su secreción o no, también se confirmó en todos los casos (geles superiores de la Fig 26D, E y F). Estos resultados, en conjunto, indican que la deficiencia en la secreción de la mutante  $\Delta$ SpaK es específica para sus efectores (SopE, SopE2), y que esta deficiencia no puede ser revertida por la expresión ectópica de SpaK.

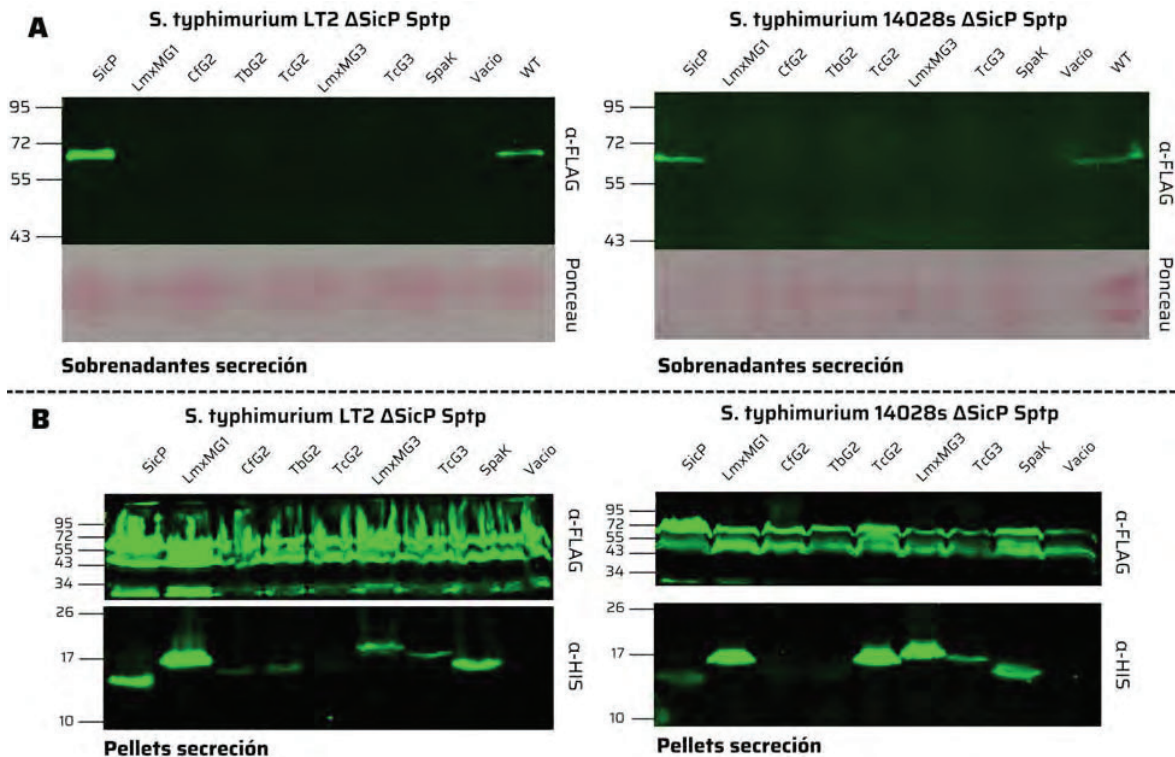


**Figura 27. Ensayo de secreción de efectores no específicos en cepas mutantes.** Western blot de los sobrenadantes del ensayo de inducción del TIIIS revelado con mAb  $\alpha$ -FLAG. De izquierda a derecha, la calle 1 corresponde a la cepa WT 14028s transformada con SopE2; las calles 2 y 3 a la cepa  $\Delta$ SicP 14028s transformada con el efector SopE2 o SopE; y la calle 4, a la cepa  $\Delta$ SpaK 14028s transformada con SptP. Se indican los marcadores de masa molecular (en kDa) a la izquierda.

## Capítulo 4: Evaluación funcional de los dominios CEST

### Sistema SicP/SptP

Por un lado, evaluamos la presencia de SptP en el sobrenadante de distintas cepas por Western Blot utilizando mAb  $\alpha$ -FLAG. Como se observa en la Fig 28A, en ambos *backgrounds* genéticos tanto la cepa WT como la mutante  $\Delta$ SicP complementada con SicP secretan al medio una banda inmunoreactiva para  $\alpha$ -FLAG, de un tamaño compatible con el esperado para el efector SptP etiquetado con 3xFLAG (~63 kDa). Esta banda no se observa en la mutante  $\Delta$ SicP complementada con el plásmido pET vacío o con los distintos dominios CEST de tripanosomátidos (Fig 28A) indicando falta de complementación funcional. La complementación con SpaK tampoco logra restituir la secreción de SptP en la mutante  $\Delta$ SicP, reafirmando la especificidad del sistema.



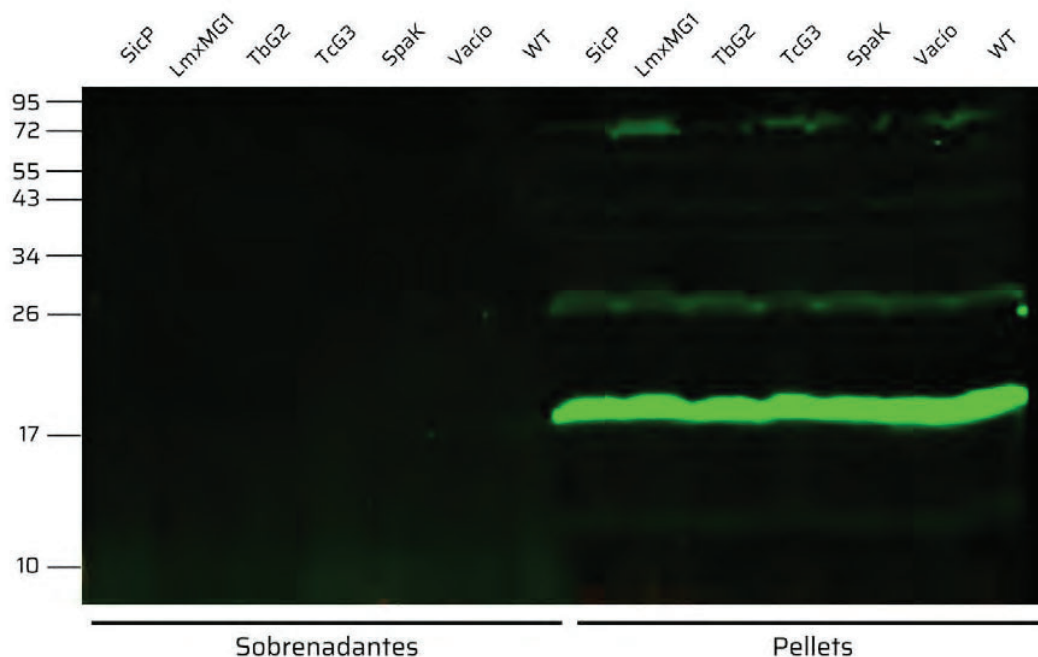
**Figura 28. Ensayos de complementación funcional para SicP.** [A] Western blot de los sobrenadantes del ensayo de inducción del TIIS revelado con mAb  $\alpha$ -FLAG. De izquierda a derecha, las calles 1 a 9 corresponden a muestras de la cepa  $\Delta$ SicP LT2 (izquierda) o 14028s (derecha) co-transformadas con el efector SptP y el 'plásmido chaperona' indicado. La calle 10 corresponde a la cepa WT transformada con SptP. [B] Western blot de los pellets del ensayo de inducción del TIIS, revelado con mAb  $\alpha$ -FLAG (arriba) y antisero  $\alpha$ -6xHIS (abajo). Las calles 1 a 9 corresponden a muestras de la cepa  $\Delta$ SicP LT2 (izquierda) o 14028s (derecha) co-transformadas con el efector SptP-FLAG y el 'plásmido chaperona' indicado. Se indican los marcadores de masa molecular (en kDa) a la izquierda.

Todos los dominios CEST (a excepción de TcG2 y TbG2) mostraron niveles intracelulares similares o incluso mayores que los de SicP, lo que permite descartar deficiencias en su expresión como causa de la no complementación funcional observada (Fig 28B, panel inferior). Interesantemente, el monitoreo de SptP en el citoplasma de las distintas cepas reveló bandas



de menor intensidad respecto de la complementada con SicP, e incluso la aparición de bandas adicionales de menor tamaño, resultados compatibles con una mayor degradación intracelular del efector.

Para descartar el efecto de la lisis espontánea durante la inducción y posterior procesamiento de las bacterias en la secreción de SptP (Fig 28), se determinó la presencia de Ecotin [74], un marcador periplasmático de *S. Typhimurium* en muestras de pellets y sobrenadantes. En ninguno de los casos, ni siquiera en las cepas capaces de secretar SptP (WT y mutante  $\Delta$ SicP complementada con SicP), se observó Ecotin en el sobrenadante (Fig 29). En todos los pellets, en cambio, sí se observó una banda inmunoreactiva para  $\alpha$ -ecotin, de un tamaño compatible con el esperado para esta molécula (~16 kDa). Las bandas minoritarias reveladas por el antisuero  $\alpha$ -ecotin se corresponden con las reportadas en literatura [75]. A partir de los ensayos realizados, concluimos que los dominios CEST de tripanosomátidos no son capaces de complementar la función de la chaperona SicP *in vivo*.



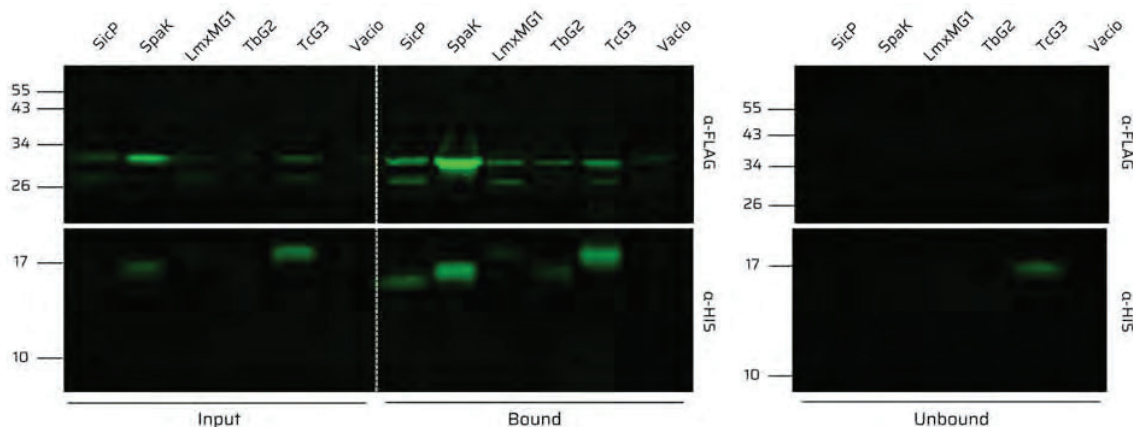
**Figura 29. Control de lisis sobre el ensayo de secreción de SptP.** Western blot de los sobrenadantes y pellets del ensayo de inducción del TIIIS revelado con suero  $\alpha$ -ecotin. De izquierda a derecha, las calles 1-6 y 8-13 se corresponden a muestras de la cepa  $\Delta$ SicP 14028s co-transformadas con el efector SptP y el 'plásmido chaperona' indicado. Las calles 7 y 14 corresponden a la cepa WT transformada con SptP. Se indican los marcadores de masa molecular (en kDa) a la izquierda.

### Sistema SpaK/SopE2

En la sección anterior mostramos que la deficiencia en la secreción de la mutante  $\Delta$ SpaK, si bien es específica para sus efectores (SopE, SopE2), no puede ser revertida por la complementación con SpaK. En este marco, para evaluar de manera indirecta la complementación funcional de los dominios CEST sobre esta chaperona, implementamos un ensayo de interacción *in vivo* (en el citoplasma bacteriano) entre estos dominios y el efector SopE2. Brevemente, se realizó un ensayo de *pull-down* en la cepa *S. Typhimurium* 14028s

$\Delta$ SpaK co-transformada con el efector SopE2 y distintos 'plásmidos chaperona'. La idea general es co-expresar el efector y las chaperonas en una misma bacteria, lisarlas, capturar las chaperonas por su etiqueta 6xHis mediante cromatografía de afinidad y evaluar la presencia del efector SopE2 en la fase retenida (*bound*) mediante Western blot usando el mAb  $\alpha$ -FLAG.

En la Fig 30 se observa la retención diferencial del efector SopE2 (banda inmunoreactiva para  $\alpha$ -FLAG de ~29 kDa) cuando la bacteria se encuentra transformada con SpaK respecto a los controles negativos (SicP y el plásmido vacío). A su vez, puede observarse en el *input* que la complementación con SpaK correlaciona con cierta estabilización (mayor nivel de expresión y menor nivel de degradación) del efector (Fig 30). La sobre-expresión de ciertos dominios CEST de tripanosomátidos (LmxMG1, TbG2, TccG3) no generó una retención diferencial de SopE2 en el *bound* respecto a los controles negativos. Además, el monitoreo de SopE2 en las fracciones *input* de las distintas cepas reveló bandas de menor intensidad respecto de la complementada con SpaK, e incluso la aparición de una banda adicional de menor tamaño, resultados compatibles con una mayor degradación intracelular del efector. A su vez, analizamos los niveles de los dominios CEST (o chaperonas Cest bacterianas) en las diferentes fracciones mediante Western Blot utilizando mAb  $\alpha$ -6xHis. Todos los dominios CEST (a excepción de TcG2 y TbG2) mostraron niveles similares o incluso mayores que los de SpaK, lo que permite descartar deficiencias en la expresión de los dominios CEST como causa de la falta de interacción observada (Fig 30, panel inferior). A partir de los ensayos realizados, concluimos que los dominios CEST de tripanosomátidos no son capaces de interactuar con el efector SopE2, lo que sugiere que, al igual que lo observado para la chaperona SicP, no serían capaces de complementar funcionalmente a SpaK.



**Figura 30. Ensayo de pull-down realizados en *S. Typhimurium* 14028s  $\Delta$ SpaK transformada con SopE2.** Western blot revelado con mAb  $\alpha$ -FLAG (paneles superiores) y mAb  $\alpha$ -HIS (paneles inferiores). Se analizó las 3 fracciones (*input*, *bound* y *unbound*) de la cepa  $\Delta$ SpaK co-transformada con SopE2 y el 'plásmido chaperona' indicado. Se indican los marcadores de masa molecular (en kDa) a la izquierda.

## Discusión

El objetivo de esta Tesis de Licenciatura fue caracterizar tanto estructural como funcionalmente los dominios CEST presentes en kinetoplástidos. Como primera aproximación, se realizó un análisis exhaustivo de la estructura molecular, distribución genómica y evolución de los dominios CEST (y proteínas que los contienen) mediante herramientas bioinformáticas.

Descartando posibles sesgos asociados a problemas de cobertura y/o ensamblaje de los genomas y/o a la disparidad de la información genómica disponible para los kinetoplástidos (particularmente en ciertos clados), la hipótesis más parsimoniosa para explicar nuestros hallazgos genómicos indica que la familia de genes CEST es un desarrollo evolutivo propio de los kinetoplástidos, de distribución (aparentemente) restringida a dos de sus linajes: los tripanosomátidos y los bodónidos. Esta familia génica está compuesta por 3 miembros, a los que llamamos CEST 1, 2 y 3, que se distinguen por contener un dominio CEST, típicamente procariota. Si bien existen variaciones lógicas entre los alelos de cada CEST identificados en las diferentes especies, la estructura general de cada uno de estos genes se encuentra muy conservada. Los genes CEST 1 y 3 serían los ancestrales, los cuales habrían sido adquiridos (posiblemente por transferencia horizontal a partir de una chaperona CesT de una bacteria del grupo de las Chlamydiaceae) y combinados con los dominios UBL y/o PDZ en un ancestro común de los kinetoplástidos, posterior a su segregación de los Euglénidos. Si bien no encontramos organismos conteniendo solo CEST 1 o CEST 3, lo más plausible es que estos dos genes se hayan originado por duplicación, a partir de un único evento de adquisición. Para explicar la ausencia de CEST 1 y 3 en los *Trypanosoma* salivaria, proponemos la ocurrencia de eventos de delección genómica en un ancestro común a este grupo. CEST 2, por otro lado, es evolutivamente más reciente, y parece haberse generado a partir de una duplicación (y posterior diversificación) del gen CEST 1. La ausencia de CEST 2 en *B. saltans* y su presencia en *P. confusum* sugieren que el origen de este gen es posterior a la separación de tripanosomátidos y bodónidos. Además, la falta de sintenia para CEST 2 entre los clados *Trypanosoma* y *Leishmania* sugiere 2 eventos de generación independientes para este gen.

Un punto a destacar es la presencia del dominio auxiliar UBL en los 3 genes CEST. Este tipo de dominios se encuentran en una gran variedad de proteínas que están implicadas en el reciclaje de proteínas mediada por el proteosoma, pueden participar en el reconocimiento de sustratos, la regulación de laclado de proteínas se encuentra asociado al bolsillo flagelar [77], región donde se localiza actividad del proteosoma o en el ensamblaje del mismo [76]. En tripanosomátidos, el recizan las proteínas CEST tanto de *T. cruzi* [21] como de *T. brucei* [78]. Más aún, la sobreexpresión homóloga de esta proteína condujo a una reducción de la capacidad endocítica de los epimastigotes de *T. cruzi* y a una mortalidad exacerbada durante la fase estacionaria tardía de los cultivos axénicos [21], fenotipos compatibles con la alteración del turnover de moléculas asociadas a la captación de macromoléculas necesarias para la nutrición. Por lo tanto, considerando su función inferida y localización, resulta interesante estudiarlos estructural y funcionalmente.

A pesar de la gran divergencia en la secuencia aminoacídica entre los dominios CEST y las chaperonas TIISS de origen bacteriano, estos dominios muestran una significativa similitud

estructural a nivel global. Esta similitud es más pronunciada en comparación con SicP (y posiblemente otras chaperonas mono-específicas), pero también es notable para chaperonas multiespecíficas, como SpaK. Al examinar las regiones implicadas en la interacción de las Cest con sus respectivos efectores, sin embargo, se observa que los dominios CEST conservan sólo alguno de los aa claves para la actividad chaperona de SpaK; mientras que no contienen el residuo Phe clave para la actividad chaperona de SicP. La evaluación en detalle de las regiones del HBG de los dominios CEST, además, revela solo leves divergencias tanto con SicP como con SpaK.

En una segunda parte, se evaluó funcionalmente a los dominios CEST identificados, para lo cual se desarrolló un modelo experimental basado en el TIISS de *S. Typhimurium*. Esto implicó la generación de diferentes herramientas moleculares *ad hoc*, tales como vectores de expresión conteniendo los dominios CEST (y chaperonas bacterianas usadas de control) etiquetados con 6xHIS, vectores de expresión conteniendo distintos efectores bacterianos (SptP, SopE2, SopE) etiquetados con 3xFLAG y, sobre todo, cepas de *S. Typhimurium* KO para chaperonas específicas. Sin embargo, no se logró generar un cepa mutante para SpaK en el background genético LT2, es necesario que el sistema de KO basado en lambda red siga poniéndose a punto para obtener la eficiencia necesaria. De todas maneras, en el background genético 14028s se determinó la correcta mutación del gen de interés (SicP o SpaK) mediante geno- y fenotipificación, donde se confirmó que el defecto de secreción observado en las cepas no fue el resultado de un efecto global sobre la secreción mediada por TIISS.

Los dominios CEST no lograron complementar la función de la chaperona SicP *in vivo*, no fueron capaces de secretar el efector Sptp al medio, ni lograron interactuar con el efector SopE2. Para lograr secretar un efector al medio extracelular es necesario no solo que se produzca una interacción con dicha proteína, sino que también se requiere de la interacción del complejo chaperona-efector con el inyectisoma [31]. La gran divergencia en la secuencia aminoacídica observada en los dominios CEST podría impedir la unión con el inyectisoma evitando así la translocación del efector Sptp. A su vez, los dominios CEST no poseen todos los residuos conservados en el sitio de unión de la chaperona, por lo tanto se podría ver afectada la interacción directa con los efectores por la variación de la secuencia de aminoácidos.

La cepa  $\Delta$ SpaK en el background genético 14028s que sobreexpresa la proteína SpaK no secretó los efectores SopE2 ni SopE al medio extracelular. Dado que la mutante  $\Delta$ SpaK no tiene afectada la secreción mediada por TIISS, este resultado contradice lo observado previamente en mutantes  $\Delta$ SpaK [35], [36], [55]. Por lo tanto, el defecto de secreción podría estar asociado a la manera en la que fue generada la mutante. Otra posible explicación es que la mutación sobre el gen en cuestión haya generado un ligero efecto polar del cassette de cloranfenicol sobre genes posteriores que codifican componentes regulatorios de TIISS, afectando el transporte de forma dependiente de SpaK.

En definitiva, los genes CEST han sido conservados a lo largo de la evolución de los tripanosomátidos y se encuentran ampliamente distribuidos. Considerando su ubicación, la presencia del dominio UBL y su estructura conservada, nuestros resultados sugieren que los dominios CEST podrían desempeñar un papel en la secreción de factores de virulencia. Este

trabajo abre nuevos interrogantes sobre la función de estas proteínas, será necesario utilizar las herramientas generadas para explorar en profundidad los dominios a nivel funcional.

## Conclusiones

- Los genomas de tripanosomátidos contienen una gran cantidad de secuencias CEST, en las que pueden delimitarse 3 grupos de homología (CEST 1-3). La alta conservación estructural entre ellas (y también respecto de sus contrapartes bacterianas) las convierte en prometedores candidatos para la identificación de nuevos blancos terapéuticos.
- Se generaron herramientas celulares (cepas mutantes de *S. Typhimurium*) y moleculares ('plásmidos efectores', 'plásmidos chaperonas', etc) apropiadas que permitieron evaluar la complementación funcional de los dominios CEST de tripanosomátidos sobre distintas chaperonas del TIIS en *S. Typhimurium in vivo*.
- Los dominios CEST de tripanosomátidos, si bien se expresan correctamente en *S. Typhimurium*, no complementan la función de las chaperonas bacterianas SicP y SpaK.

## Bibliografía

- [1] M. L. Povelones, «Beyond replication: division and segregation of mitochondrial DNA in kinetoplastids», *Mol. Biochem. Parasitol.*, vol. 196, n.º 1, pp. 53-60, ago. 2014, doi: 10.1016/j.molbiopara.2014.03.008.
- [2] A. G. B. Simpson, J. R. Stevens, y J. Lukeš, «The evolution and diversity of kinetoplastid flagellates», *Trends Parasitol.*, vol. 22, n.º 4, pp. 168-174, abr. 2006, doi: 10.1016/j.pt.2006.02.006.
- [3] J. Harmer, V. Yurchenko, A. Nenarokova, J. Lukeš, y M. L. Ginger, «Farming, slaving and enslavement: histories of endosymbioses during kinetoplastid evolution», *Parasitology*, vol. 145, n.º 10, pp. 1311-1323, sep. 2018, doi: 10.1017/S0031182018000781.
- [4] D. A. S. de Souza, D. P. Pavoni, M. A. Krieger, y A. Ludwig, «Evolutionary analyses of myosin genes in trypanosomatids show a history of expansion, secondary losses and neofunctionalization», *Sci. Rep.*, vol. 8, n.º 1, p. 1376, ene. 2018, doi: 10.1038/s41598-017-18865-y.
- [5] A. Kaufer, J. Ellis, D. Stark, y J. Barratt, «The evolution of trypanosomatid taxonomy», *Parasit. Vectors*, vol. 10, n.º 1, p. 287, jun. 2017, doi: 10.1186/s13071-017-2204-7.
- [6] D. A. Maslov, F. R. Opperdoes, A. Y. Kostygov, H. Hashimi, J. Lukeš, y V. Yurchenko, «Recent advances in trypanosomatid research: genome organization, expression, metabolism, taxonomy and evolution», *Parasitology*, vol. 146, n.º 1, pp. 1-27, ene. 2019, doi: 10.1017/S0031182018000951.
- [7] K. Katsuno *et al.*, «Hit and lead criteria in drug discovery for infectious diseases of the developing world», *Nat. Rev. Drug Discov.*, vol. 14, n.º 11, pp. 751-758, nov. 2015, doi: 10.1038/nrd4683.
- [8] B. Amos *et al.*, «VEuPathDB: the eukaryotic pathogen, vector and host bioinformatics resource center», *Nucleic Acids Res.*, vol. 50, n.º D1, pp. D898-D911, ene. 2022, doi: 10.1093/nar/gkab929.
- [9] A. Butenko, M. Hammond, M. C. Field, M. L. Ginger, V. Yurchenko, y J. Lukeš, «Reductionist Pathways for Parasitism in Euglenozoans? Expanded Datasets Provide New Insights», *Trends Parasitol.*, vol. 37, n.º 2, pp. 100-116, feb. 2021, doi: 10.1016/j.pt.2020.10.001.
- [10] N. M. El-Sayed *et al.*, «The Genome Sequence of *Trypanosoma cruzi*, Etiologic Agent of Chagas Disease», *Science*, vol. 309, n.º 5733, pp. 409-415, jul. 2005, doi: 10.1126/science.1112631.
- [11] M. Berriman *et al.*, «The Genome of the African Trypanosome *Trypanosoma brucei*», *Science*, vol. 309, n.º 5733, pp. 416-422, jul. 2005, doi: 10.1126/science.1112642.
- [12] A. C. Ivens *et al.*, «The genome of the kinetoplastid parasite, *Leishmania major*», *Science*, vol. 309, n.º 5733, pp. 436-442, jul. 2005, doi: 10.1126/science.1112680.
- [13] N. M. El-Sayed *et al.*, «Comparative genomics of trypanosomatid parasitic protozoa», *Science*, vol. 309, n.º 5733, pp. 404-409, jul. 2005, doi: 10.1126/science.1112181.
- [14] F. R. Opperdoes y P. A. M. Michels, «Horizontal gene transfer in trypanosomatids», *Trends Parasitol.*, vol. 23, n.º 10, pp. 470-476, oct. 2007, doi: 10.1016/j.pt.2007.08.002.
- [15] A. Kaufer, D. Stark, y J. Ellis, «A review of the systematics, species identification and diagnostics of the Trypanosomatidae using the maxicircle kinetoplast DNA: from past to present», *Int. J. Parasitol.*, vol. 50, n.º 6-7, pp. 449-460, jun. 2020, doi: 10.1016/j.ijpara.2020.03.003.
- [16] K. Gull, «Host-parasite interactions and trypanosome morphogenesis: a flagellar pocketful of goodies», *Curr. Opin. Microbiol.*, vol. 6, n.º 4, pp. 365-370, ago. 2003, doi: 10.1016/s1369-5274(03)00092-4.
- [17] W. de Souza, «Structural organization of *Trypanosoma cruzi*», *Mem. Inst. Oswaldo Cruz*, vol. 104 Suppl 1, pp. 89-100, jul. 2009, doi: 10.1590/s0074-02762009000900014.
- [18] M. Engstler *et al.*, «Hydrodynamic flow-mediated protein sorting on the cell surface of trypanosomes», *Cell*, vol. 131, n.º 3, pp. 505-515, nov. 2007, doi: 10.1016/j.cell.2007.08.046.
- [19] O. Campetella, C. A. Buscaglia, J. Mucci, y M. S. Leguizamón, «Parasite-host glycan interactions during

- Trypanosoma cruzi infection: trans-Sialidase rides the show», *Biochim. Biophys. Acta Mol. Basis Dis.*, vol. 1866, n.º 5, p. 165692, may 2020, doi: 10.1016/j.bbadis.2020.165692.
- [20] M. C. Field y M. Carrington, «The trypanosome flagellar pocket», *Nat. Rev. Microbiol.*, vol. 7, n.º 11, pp. 775-786, nov. 2009, doi: 10.1038/nrmicro2221.
- [21] I. M. Durante, M. D. L. M. Cámara, y C. A. Buscaglia, «A Novel Trypanosoma cruzi Protein Associated to the Flagellar Pocket of Replicative Stages and Involved in Parasite Growth», *PLOS ONE*, vol. 10, n.º 6, p. e0130099, jun. 2015, doi: 10.1371/journal.pone.0130099.
- [22] C. Grabbe y I. Dikic, «Functional Roles of Ubiquitin-Like Domain (ULD) and Ubiquitin-Binding Domain (UBD) Containing Proteins», *Chem. Rev.*, vol. 109, n.º 4, pp. 1481-1494, abr. 2009, doi: 10.1021/cr800413p.
- [23] A. Y. Hung y M. Sheng, «PDZ Domains: Structural Modules for Protein Complex Assembly», *J. Biol. Chem.*, vol. 277, n.º 8, pp. 5699-5702, feb. 2002, doi: 10.1074/jbc.R100065200.
- [24] M. F. Feldman y G. R. Cornelis, «The multitaled type III chaperones: all you can do with 15 kDa», *FEMS Microbiol. Lett.*, vol. 219, n.º 2, pp. 151-158, feb. 2003, doi: 10.1016/S0378-1097(03)00042-9.
- [25] C. E. Stebbins y J. E. Galán, «Maintenance of an unfolded polypeptide by a cognate chaperone in bacterial type III secretion», *Nature*, vol. 414, n.º 6859, pp. 77-81, nov. 2001, doi: 10.1038/35102073.
- [26] D. S. A. Beeckman, T. Geens, J.-P. Timmermans, P. Van Oostveldt, y D. C. G. Vanrompay, «Identification and characterization of a type III secretion system in Chlamydomonas reinhardtii», *Vet. Res.*, vol. 39, n.º 3, p. 27, 2008, doi: 10.1051/vetres:2008002.
- [27] R. S. Gupta y E. Griffiths, «Chlamydiae-specific proteins and indels: novel tools for studies», *Trends Microbiol.*, vol. 14, n.º 12, pp. 527-535, dic. 2006, doi: 10.1016/j.tim.2006.10.002.
- [28] D. C. F. Silva, «Examining marginal sequence similarities between bacterial type III secretion system components and Trypanosoma cruzi surface proteins: horizontal gene transfer or convergent evolution?», *Front. Genet.*, 2013, doi: 10.3389/fgene.2013.00143.
- [29] J. E. Galán y H. Wolf-Watz, «Protein delivery into eukaryotic cells by type III secretion machines», *Nature*, vol. 444, n.º 7119, pp. 567-573, nov. 2006, doi: 10.1038/nature05272.
- [30] J. E. Button y J. E. Galán, «Regulation of Chaperone/Effector Complex Synthesis in a Bacterial Type III Secretion System», *Mol. Microbiol.*, vol. 81, n.º 6, pp. 1474-1483, sep. 2011, doi: 10.1111/j.1365-2958.2011.07784.x.
- [31] W. Deng *et al.*, «Assembly, structure, function and regulation of type III secretion systems», *Nat. Rev. Microbiol.*, vol. 15, n.º 6, pp. 323-337, jun. 2017, doi: 10.1038/nrmicro.2017.20.
- [32] S. Ho Lee y J. E. Galán, «InvB Is a Type III Secretion-Associated Chaperone for the Salmonella enterica Effector Protein SopE», *J. Bacteriol.*, vol. 185, n.º 24, pp. 7279-7284, dic. 2003, doi: 10.1128/JB.185.24.7279-7284.2003.
- [33] Y. Fu y J. E. Galán, «Identification of a Specific Chaperone for SptP, a Substrate of the Centisome 63 Type III Secretion System of Salmonella Typhimurium», *J. Bacteriol.*, vol. 180, n.º 13, pp. 3393-3399, jul. 1998, doi: 10.1128/JB.180.13.3393-3399.1998.
- [34] M. Lilic, M. Vujanac, y C. E. Stebbins, «A Common Structural Motif in the Binding of Virulence Factors to Bacterial Secretion Chaperones», *Mol. Cell*, vol. 21, n.º 5, pp. 653-664, mar. 2006, doi: 10.1016/j.molcel.2006.01.026.
- [35] P. A. Bronstein, E. A. Miao, y S. I. Miller, «InvB Is a Type III Secretion Chaperone Specific for SspA», *J. Bacteriol.*, vol. 182, n.º 23, pp. 6638-6644, dic. 2000, doi: 10.1128/JB.182.23.6638-6644.2000.
- [36] K. Ehrbar, S. Hapfelmeier, B. Stecher, y W.-D. Hardt, «InvB Is Required for Type III-Dependent Secretion of SopA in Salmonella enterica Serovar Typhimurium», *J. Bacteriol.*, vol. 186, n.º 4, pp. 1215-1219, feb. 2004, doi: 10.1128/JB.186.4.1215-1219.2004.
- [37] P. Wattiau y G. R. Cornelis, «SycE, a chaperone-like protein of Yersinia enterocolitica involved in Ome



- secretion of YopE», *Mol. Microbiol.*, vol. 8, n.º 1, pp. 123-131, abr. 1993, doi: 10.1111/j.1365-2958.1993.tb01209.x.
- [38] R. Dewoody, P. M. Merritt, y M. M. Marketon, «Regulation of the Yersinia type III secretion system: traffic control», *Front. Cell. Infect. Microbiol.*, vol. 3, feb. 2013, doi: 10.3389/fcimb.2013.00004.
- [39] G. Horna y J. Ruiz, «Type 3 secretion system of *Pseudomonas aeruginosa*», *Microbiol. Res.*, vol. 246, p. 126719, may 2021, doi: 10.1016/j.micres.2021.126719.
- [40] Y.-S. Chen *et al.*, «The Chlamydia trachomatis Type III Secretion Chaperone Slc1 Engages Multiple Early Effectors, Including TepP, a Tyrosine-phosphorylated Protein Required for the Recruitment of CrkI-II to Nascent Inclusions and Innate Immune Signaling», *PLoS Pathog.*, vol. 10, n.º 2, p. e1003954, feb. 2014, doi: 10.1371/journal.ppat.1003954.
- [41] S. L. Slater, A. M. Sångfors, D. J. Pollard, D. Ruano-Gallego, y G. Frankel, «The Type III Secretion System of Pathogenic *Escherichia coli*», en *Escherichia coli, a Versatile Pathogen*, vol. 416, G. Frankel y E. Z. Ron, Eds., en Current Topics in Microbiology and Immunology, vol. 416., Cham: Springer International Publishing, 2018, pp. 51-72. doi: 10.1007/82\_2018\_116.
- [42] R. Manon *et al.*, «The Different Strategies Used by Salmonella to Invade Host Cells», en *Salmonella - Distribution, Adaptation, Control Measures and Molecular Technologies*, IntechOpen, 2012. doi: 10.5772/29979.
- [43] J. E. Galán, «Salmonella interactions with host cells: type III secretion at work», *Annu. Rev. Cell Dev. Biol.*, vol. 17, pp. 53-86, 2001, doi: 10.1146/annurev.cellbio.17.1.53.
- [44] F. Madeira *et al.*, «Search and sequence analysis tools services from EMBL-EBI in 2022», *Nucleic Acids Res.*, vol. 50, n.º W1, pp. W276-W279, jul. 2022, doi: 10.1093/nar/gkac240.
- [45] K. Okonechnikov, O. Golosova, M. Fursov, y the UGENE team, «Unipro UGENE: a unified bioinformatics toolkit», *Bioinformatics*, vol. 28, n.º 8, pp. 1166-1167, abr. 2012, doi: 10.1093/bioinformatics/bts091.
- [46] K. Tamura, G. Stecher, y S. Kumar, «MEGA11: Molecular Evolutionary Genetics Analysis Version 11», *Mol. Biol. Evol.*, vol. 38, n.º 7, pp. 3022-3027, jun. 2021, doi: 10.1093/molbev/msab120.
- [47] D. T. Jones, W. R. Taylor, y J. M. Thornton, «The rapid generation of mutation data matrices from protein sequences», *Bioinformatics*, vol. 8, n.º 3, pp. 275-282, 1992, doi: 10.1093/bioinformatics/8.3.275.
- [48] «FigTree». Accedido: 29 de enero de 2024. [En línea]. Disponible en: <http://tree.bio.ed.ac.uk/software/figtree/>
- [49] T. Paysan-Lafosse *et al.*, «InterPro in 2022», *Nucleic Acids Res.*, vol. 51, n.º D1, pp. D418-D427, ene. 2023, doi: 10.1093/nar/gkac993.
- [50] A. Drozdetskiy, C. Cole, J. Procter, y G. J. Barton, «JPred4: a protein secondary structure prediction server», *Nucleic Acids Res.*, vol. 43, n.º W1, pp. W389-W394, jul. 2015, doi: 10.1093/nar/gkv332.
- [51] H. Wickham, «ggplot2: Elegant Graphics for Data Analysis». Springer-Verlag New York, 2016. [En línea]. Disponible en: <https://ggplot2.tidyverse.org>
- [52] J. Jumper *et al.*, «Highly accurate protein structure prediction with AlphaFold», *Nature*, vol. 596, n.º 7873, pp. 583-589, ago. 2021, doi: 10.1038/s41586-021-03819-2.
- [53] Promega, «pGEM®-T and pGEM®-T Easy Vector Systems». Promega Corporation.
- [54] H. Inoue, H. Nojima, y H. Okayama, «High efficiency transformation of *Escherichia coli* with plasmids», *Gene*, vol. 96, n.º 1, pp. 23-28, nov. 1990, doi: 10.1016/0378-1119(90)90336-p.
- [55] K. Ehrbar, A. Friebel, S. I. Miller, y W.-D. Hardt, «Role of the *Salmonella* Pathogenicity Island 1 (SPI-1) Protein InvB in Type III Secretion of SopE and SopE2, Two *Salmonella* Effector Proteins Encoded Outside of SPI-1», *J. Bacteriol.*, vol. 185, n.º 23, pp. 6950-6967, dic. 2003, doi: 10.1128/JB.185.23.6950-6967.2003.
- [56] J. E. Karlinsey, «lambda-Red genetic engineering in *Salmonella enterica* serovar Typhimurium»,

- Methods Enzymol.*, vol. 421, pp. 199-209, 2007, doi: 10.1016/S0076-6879(06)21016-4.
- [57] S. K. Sharan, L. C. Thomason, S. G. Kuznetsov, y D. L. Court, «Recombineering: a homologous recombination-based method of genetic engineering», *Nat. Protoc.*, vol. 4, n.º 2, Art. n.º 2, feb. 2009, doi: 10.1038/nprot.2008.227.
- [58] J. E. Karlinsey, «λ-Red Genetic Engineering in *Salmonella enterica* serovar Typhimurium», en *Methods in Enzymology*, vol. 421, Elsevier, 2007, pp. 199-209. doi: 10.1016/S0076-6879(06)21016-4.
- [59] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, y D. J. Lipman, «Basic Local Alignment Search Tool».
- [60] P. Flegontov *et al.*, «Genome of *Leptomonas pyrrocoris*: a high-quality reference for monoxenous trypanosomatids and new insights into evolution of *Leishmania*», *Sci. Rep.*, vol. 6, n.º 1, Art. n.º 1, mar. 2016, doi: 10.1038/srep23704.
- [61] G. C. Mitchell, J. H. Baker, y M. A. Sleigh, «Feeding of a Freshwater Flagellate, *Bodo saltans*, on Diverse Bacteria<sup>1</sup>», *J. Protozool.*, vol. 35, n.º 2, pp. 219-222, may 1988, doi: 10.1111/j.1550-7408.1988.tb04327.x.
- [62] T. Skalický *et al.*, «Extensive flagellar remodeling during the complex life cycle of *Paratrypanosoma*, an early-branching trypanosomatid», *Proc. Natl. Acad. Sci.*, vol. 114, n.º 44, pp. 11757-11762, oct. 2017, doi: 10.1073/pnas.1712311114.
- [63] J. Haag, «The molecular phylogeny of trypanosomes: evidence for an early divergence of the *Salivaria*», *Mol. Biochem. Parasitol.*, vol. 91, n.º 1, pp. 37-49, mar. 1998, doi: 10.1016/S0166-6851(97)00185-0.
- [64] P. Ghosh, «Process of Protein Transport by the Type III Secretion System», *Microbiol. Mol. Biol. Rev.*, vol. 68, n.º 4, pp. 771-795, dic. 2004, doi: 10.1128/MMBR.68.4.771-795.2004.
- [65] C. Clayton, «Regulation of gene expression in trypanosomatids: living with polycistronic transcription», *Open Biol.*, vol. 9, n.º 6, p. 190072, jun. 2019, doi: 10.1098/rsob.190072.
- [66] C. A. Cooper *et al.*, «Structural and Biochemical Characterization of SrcA, a Multi-Cargo Type III Secretion Chaperone in *Salmonella* Required for Pathogenic Association with a Host», *PLoS Pathog.*, vol. 6, n.º 2, p. e1000751, feb. 2010, doi: 10.1371/journal.ppat.1000751.
- [67] M. M. Konaté, G. Plata, J. Park, D. R. Usmanova, H. Wang, y D. Vitkup, «Molecular function limits divergent protein evolution on planetary timescales», *eLife*, vol. 8, p. e39705, sep. 2019, doi: 10.7554/eLife.39705.
- [68] H. A. de Boer, L. J. Comstock, y M. Vasser, «The *tac* promoter: a functional hybrid derived from the *trp* and *lac* promoters», *Proc. Natl. Acad. Sci. U. S. A.*, vol. 80, n.º 1, pp. 21-25, ene. 1983, doi: 10.1073/pnas.80.1.21.
- [69] C. Centeno Camean, M. D. L. M. Camara, y C. Buscaglia, «Caracterización funcional del motivo CEST en *Trypanosomatidos*», Universidad Nacional de General San Martín, Instituto de Investigaciones Biotecnológicas "Dr. Rodolfo A. Ugalde", 2019.
- [70] R. Takada, «Caracterización funcional del motivo cest en la proteína TCLP1 de *Trypanosoma cruzi*», Universidad Nacional de General San Martín, Instituto de Investigaciones Biotecnológicas "Dr. Rodolfo A. Ugalde", 2017.
- [71] G. Briones, D. Hofreuter, y J. E. Galán, «Cre Reporter System To Monitor the Translocation of Type III Secreted Proteins into Host Cells», *Infect. Immun.*, vol. 74, n.º 2, pp. 1084-1090, feb. 2006, doi: 10.1128/iai.74.2.1084-1090.2006.
- [72] R. Yao, R. A. Alm, T. J. Trust, y P. Guerry, «Construction of new *Campylobacter* cloning vectors and a new mutational *cat* cassette», *Gene*, vol. 130, n.º 1, pp. 127-130, ago. 1993, doi: 10.1016/0378-1119(93)90355-7.
- [73] S. Porwollik *et al.*, «Defined Single-Gene and Multi-Gene Deletion Mutant Collections in *Salmonella*

- enterica sv Typhimurium», *PLoS ONE*, vol. 9, n.º 7, p. e99820, jul. 2014, doi: 10.1371/journal.pone.0099820.
- [74] F. De Meyer y A. Carlier, «Ecotin: A versatile protease inhibitor of bacteria and eukaryotes», *Front. Microbiol.*, vol. 14, p. 1114690, ene. 2023, doi: 10.3389/fmicb.2023.1114690.
- [75] C. T. Eggers, I. A. Murray, V. A. Delmar, A. G. Day, y C. S. Craik, «The periplasmic serine protease inhibitor ecotin protects bacteria against neutrophil elastase», *Biochem. J.*, vol. 379, n.º Pt 1, pp. 107-118, abr. 2004, doi: 10.1042/BJ20031790.
- [76] G. A. Collins y A. L. Goldberg, «Proteins containing ubiquitin-like (Ubl) domains not only bind to 26S proteasomes but also induce their activation», *Proc. Natl. Acad. Sci. U. S. A.*, vol. 117, n.º 9, pp. 4664-4674, mar. 2020, doi: 10.1073/pnas.1915534117.
- [77] M. Engstler *et al.*, «Kinetics of endocytosis and recycling of the GPI-anchored variant surface glycoprotein in *Trypanosoma brucei*», *J. Cell Sci.*, vol. 117, n.º Pt 7, pp. 1105-1115, mar. 2004, doi: 10.1242/jcs.00938.
- [78] K. Billington *et al.*, «Genome-wide subcellular protein map for the flagellate parasite *Trypanosoma brucei*», *Nat. Microbiol.*, vol. 8, n.º 3, pp. 533-547, mar. 2023, doi: 10.1038/s41564-022-01295-6.

## Anexos

Tabla Suplementaria 1: Oligonucleótidos usados en este trabajo			
Oligonucleótido	Secuencia	Sitio de restricción	Temperatura de hibridación (°C)
CEST-F	5'- AGAGGTAC <b>CAAGAGGAAAGTAAA</b> ATG -3'	-	55
LmxMG1-F	5'- AGA <b>AAGCTT</b> GAGTTGGCGGAAACGGCCTT -3'	HindIII	64
LmxMG1-R	5'- AGACT <b>CTGAGCTA</b> GGTGTGAGCAAAACGT -3'	XhoI	63
CfG2-F	5'- AGA <b>AAGCTT</b> GCAAACGCGAAGGCAAGTCT -3'	HindIII	65
CfG2-R	5'- AGACT <b>CTGAGCTA</b> CGCAAGCAGCTCACTCA -3'	XhoI	66
TbG2-F	5'- AGA <b>AAGCTT</b> CAGAGTGCTAAATCCAATCT -3'	HindIII	58
TbG2-R	5'- AGACT <b>CTGAGCTA</b> GTAAAGCAGTTCATTTA -3'	XhoI	58
TcG2-F	5'- AGA <b>AAGCTT</b> CAGAATGCCAAGTCTAATCT -3'	HindIII	58
TcG2-R	5'- AGACT <b>CTGAGCTA</b> TCAAGAAGTTCGTGTA -3'	XhoI	60
LmxMG3-F	5'- AGA <b>AAGCTT</b> GAGAACGCCGAGGAGAACCT -3'	HindIII	65
LmxMG3-R	5'- AGACT <b>CTGAGCTA</b> CAGGGTCTGCAAGAAGTTCGGT -3'	XhoI	68
SicP-F	5'- AGA <b>AAGCTT</b> TTGCAAGCACACCAGGAT -3'	HindIII	63
SicP-R	5'- AGACT <b>CTGAGCTA</b> CTACTTTAGCATATTCTGCAGTAT -3'	XhoI	62
SpaK-F	5'- AGA <b>AAGCTT</b> CAACATTTGGATATCGCTGA -3'	HindIII	59
SpaK-R	5'- AGACT <b>CTGAGCTA</b> TCTCATTAGCGACCG -3'	XhoI	61
SopE2-F	5'- AGA <b>AAGCTT</b> ATGACTAACATAACACTATCC -3'	HindIII	55
SopE2-R	5'- AGAT <b>CTAGAGC</b> AGGAGGCATTCTGAAGATAC -3'	XbaI	60
T7 terminal	5'- GCTAGTTATTGCTCAGCGG -3'	-	53
M13F	5'- CGCCAGGGTTTTCCAGTCACGAC -3'	-	60
M13R	5'- TCACACAGGAAACAGCTATGAC -3'	-	60
SpaK-genoma	5'- GCATATTCGTGGAGCAATGG -3'	-	55
CAT	5'- CAAACGGCATGATGCACTTG -3'	-	55

**ATG**: codón de iniciación (metionina); **CAAGAGGAAAGTAAA**: RBS; **AAGCTT**: HindIII; **CTCGAG**: XhoI; **CTA**, **TCA**: reversa complementaria del codón de stop TAG/TGA; **TCTAGA**: XbaI

Tabla Suplementaria 2: Genes y proteínas CEST en tripanosomátidos				
ID	Organismo	Grupo	Largo (aa)	CEST (aa inicio - aa final)
BSAL_53315	<i>Bodo saltans</i>	-	462	312-452
PCON_0036280	<i>Paratrypanosoma confusum</i>	-	579	397-533
ADEAN_000234100	<i>Angomonas deanei</i>	1	394	251-394
Baya_022_0590	<i>Blechomonas ayalai</i>	1	531	389-531
CFAC1_290061200	<i>Crithidia fasciculata</i>	1	528	387-528
DQ04_01831060	<i>Trypanosoma grayi</i>	1	442	301-442
EMOLV88_000030800	<i>Endotrypanum monterogeii</i>	1	527	388-527
JKF63_01882	<i>Parcisia hertigi</i>	1	559	424-559
LmjF.34.3170	<i>Leishmania major</i>	1	568	428-568
LpyrH10_05_3380	<i>Leptomonas pyrrocoris</i>	1	579	439-579
LSM04_005563	<i>Trypanosoma melophagium</i>	1	454	312-454
TcCLB.504057.50	<i>Trypanosoma cruzi</i>	1	446	304-443
TM35_000015940	<i>Trypanosoma theileri</i>	1	453	312-451
TRSC58_06400	<i>Trypanosoma rangeli</i>	1	446	307-446
PCON_0029740	<i>Paratrypanosoma confusum</i>	2	384	244-384
ADEAN_000139900	<i>Angomonas deanei</i>	2	301	162-301
Baya_114_0220	<i>Blechomonas ayalai</i>	2	321	183-321
CFAC1_270017100	<i>Crithidia fasciculata</i>	2	328	196-328
DQ04_02321100	<i>Trypanosoma grayi</i>	2	362	224-362
EMOLV88_310012100	<i>Endotrypanum monterogeii</i>	2	332	196-332
JKF63_03809	<i>Parcisia hertigi</i>	2	335	199-335
JKF63_03813	<i>Parcisia hertigi</i>	2	335	199-335
JKF63_03814	<i>Parcisia hertigi</i>	2	335	199-335
LmjF.31.0740	<i>Leishmania major</i>	2	319	182-319
LpyrH10_32_0240	<i>Leptomonas pyrrocoris</i>	2	333	197-333
LSM04_000363	<i>Trypanosoma melophagium</i>	2	354	216-354
Tb927.10.9240	<i>Trypanosoma brucei</i>	2	365	227-365
TcCLB.504137.70	<i>Trypanosoma cruzi</i>	2	353	215-353
TcIL3000.A.H_000801100	<i>Trypanosoma congolense</i>	2	368	230-368
TevSTIB805.10.9820	<i>Trypanosoma evansi</i>	2	365	227-365
TM35_000081200	<i>Trypanosoma theileri</i>	2	361	223-361
TvY486_1009120	<i>Trypanosoma vivax</i>	2	367	229-367

Baya_003_0330	<i>Blechomonas ayalai</i>	3	747	421-564
CFAC1_030016400	<i>Crithidia fasciculata</i>	3	799	481-627
DQ04_00791080	<i>Trypanosoma grayi</i>	3	729	432-574
EMOLV88_040011800	<i>Endotrypanum monterogeii</i>	3	777	443-589
JKF63_07263	<i>Parcisia hertigi</i>	3	784	445-579
LmjF.04.0710	<i>Leishmania major</i>	3	794	452-598
LpyrH10_24_0830	<i>Leptomonas pyrrocoris</i>	3	782	435-581
LSM04_004211	<i>Trypanosoma melophagium</i>	3	709	426-566
TcCLB.510241.10 (TCLP1)	<i>Trypanosoma cruzi</i>	3	776	494-631
TM35_000351540	<i>Trypanosoma theileri</i>	3	710	426-568
BSAL_61300	<i>Bodo saltans</i>	-	671	396-537
PCON_0067470	<i>Paratrypanosoma confusum</i>	-	934	468-610

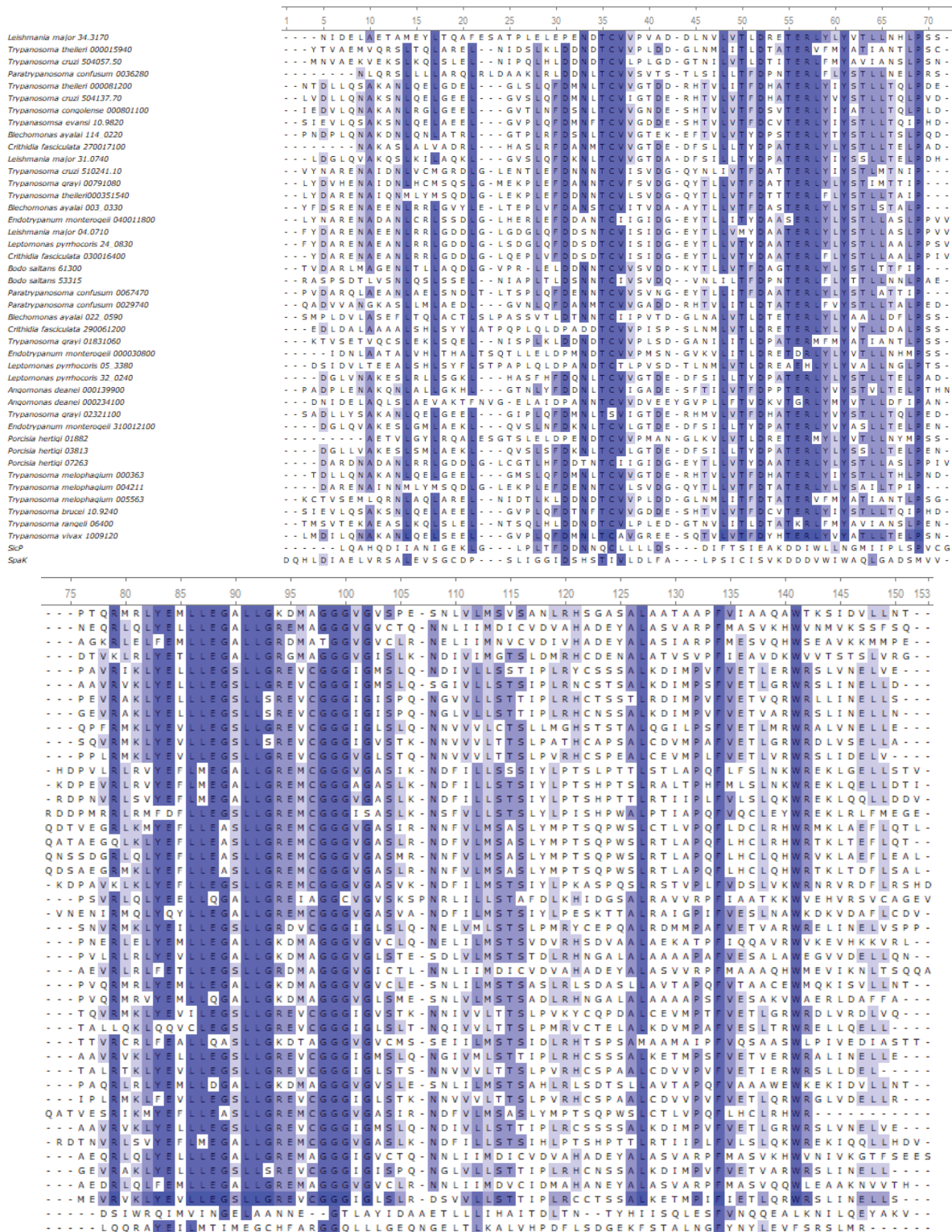
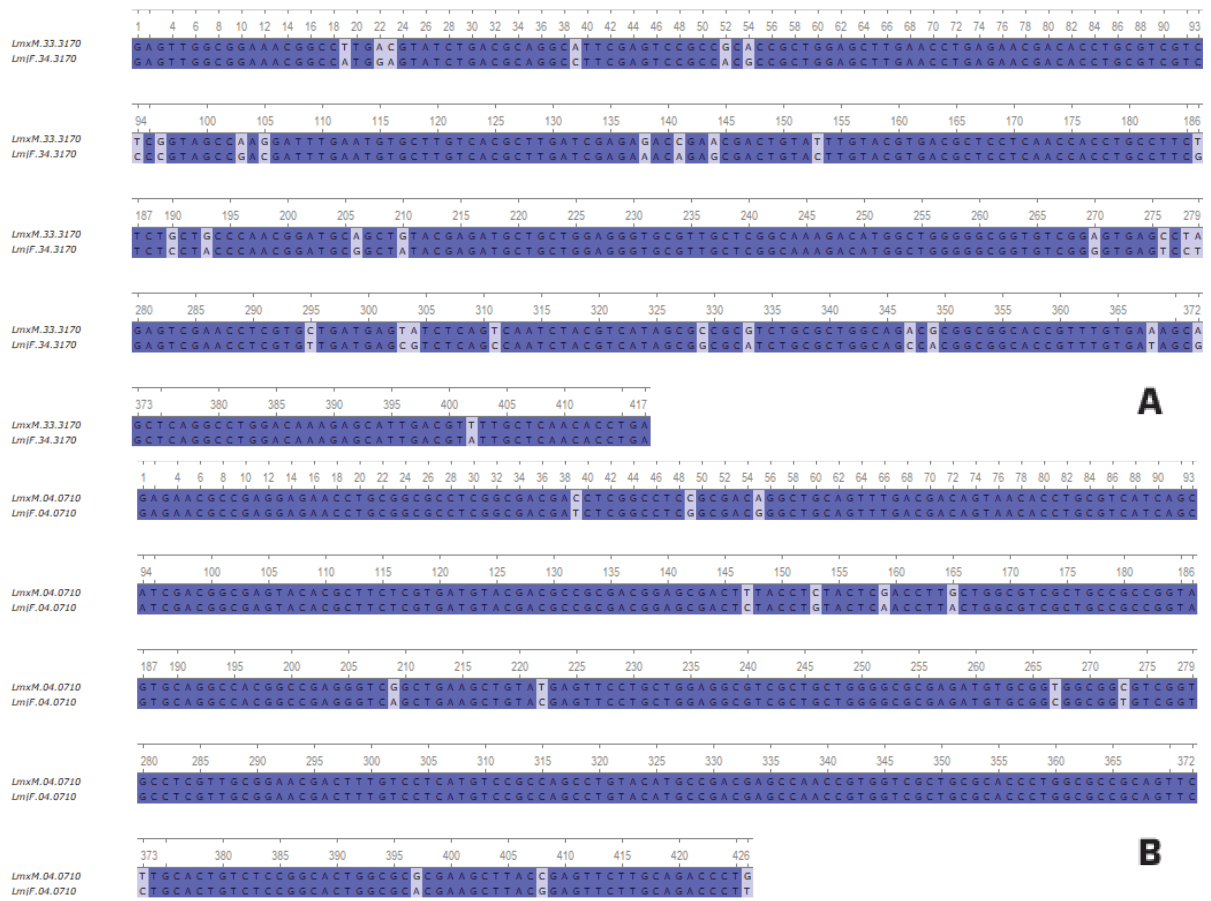
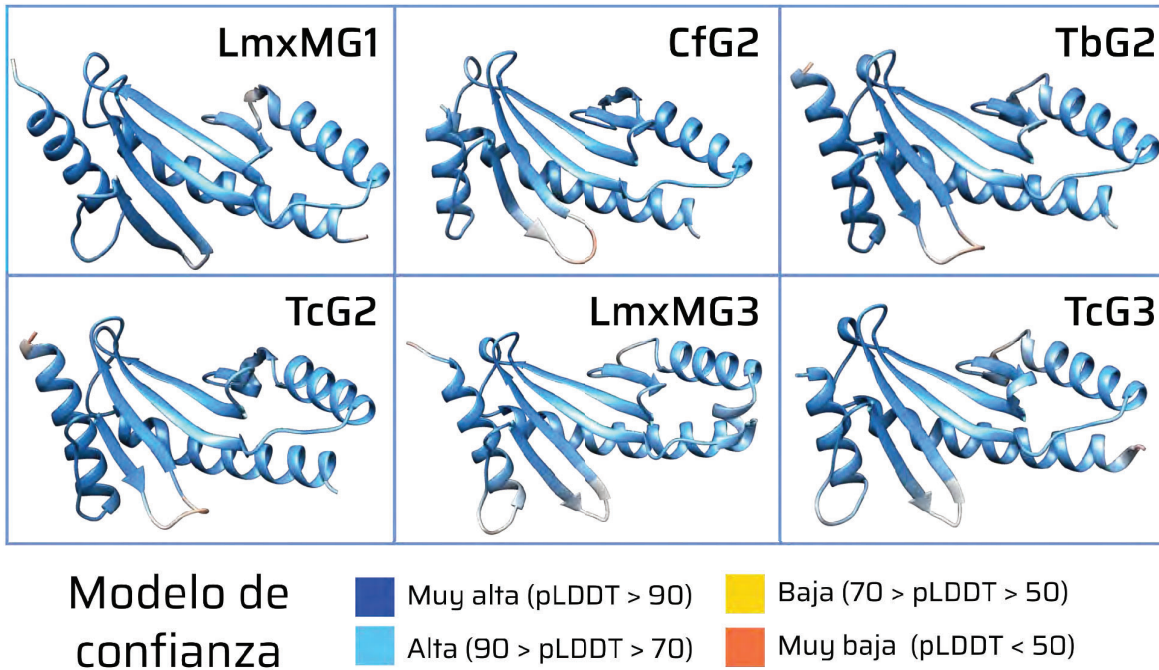


Figura Suplementaria 1. Alineamiento de secuencias de los 44 dominios CEST no redundantes identificados en las proteínas de kinetoplastos. También se incluyen las secuencias de SicP y SpaK.

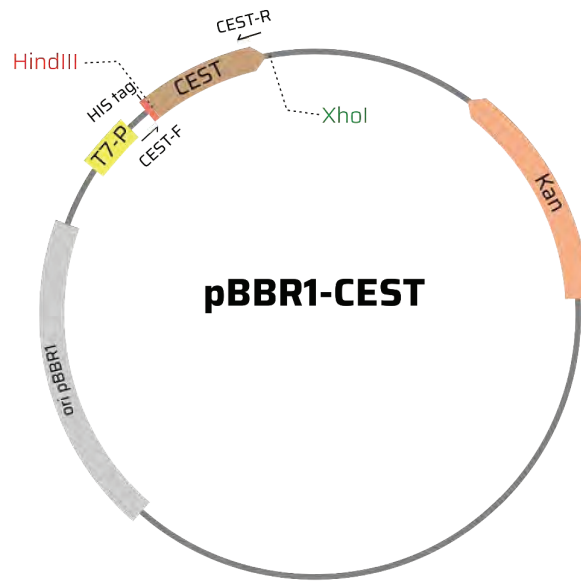


**Figura Suplementaria 2. Alineamiento de secuencias de ADN de los dominios CEST 1 [A] y 3 [B] de *L. major* y *L. mexicana*.**





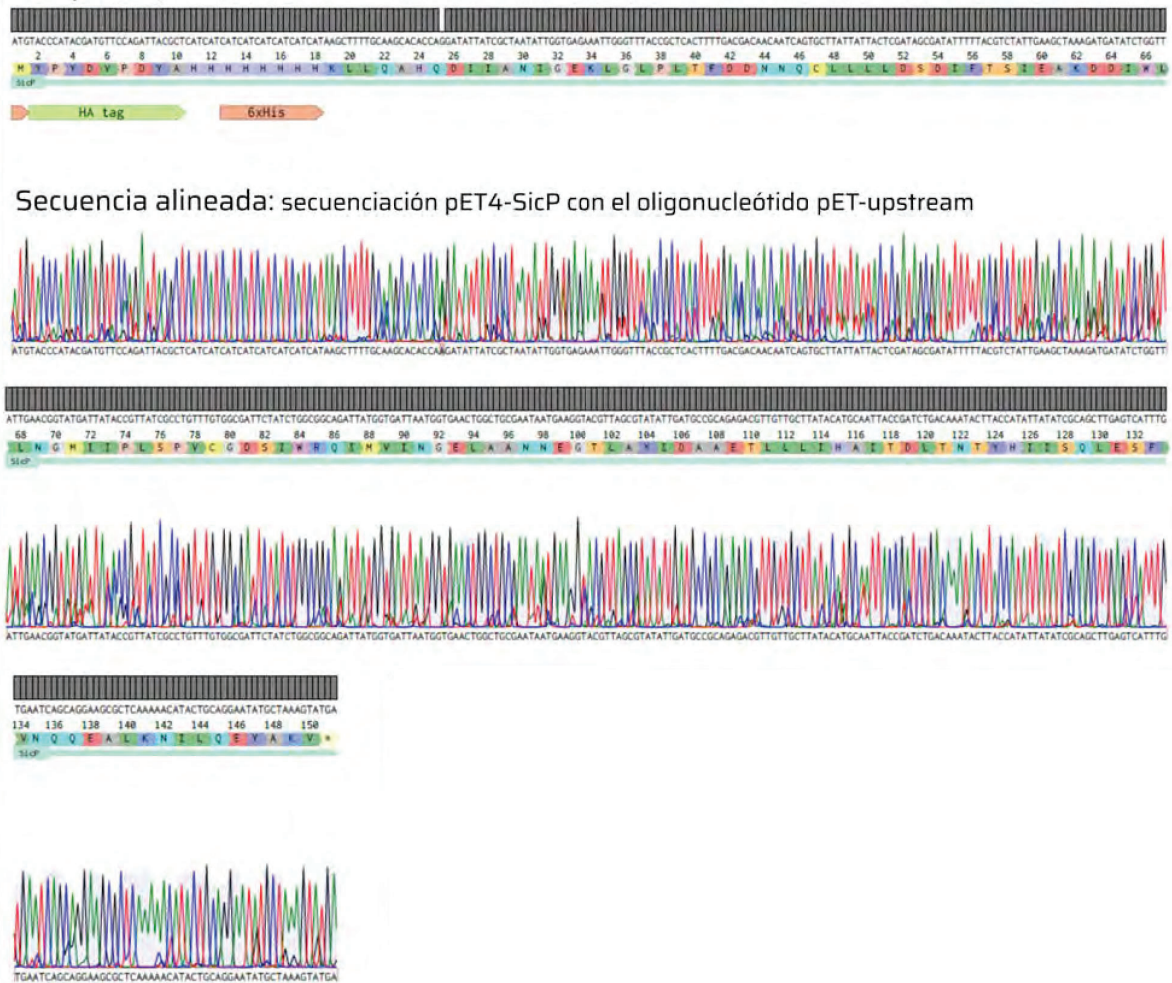
**Figura Suplementaria 3. Modelos obtenidos para los diferentes dominios CEST por AlphaFold2. Los valores de pLDDT se muestran en cada modelo mediante una escala de colores.**



**Figura Suplementaria 4. Esquema del vector pBBR1 que contiene los dominios CEST.** T7-P: promotor T7; Kan: gen que otorga resistencia a kanamicina; ori pBBR1: origen de replicación pBBR1. Los sitios de corte para las enzimas de restricción HindIII y XhoI se indican con líneas punteadas, y los oligonucleótidos, con flechas.

## Secuenciación pET24-SicP

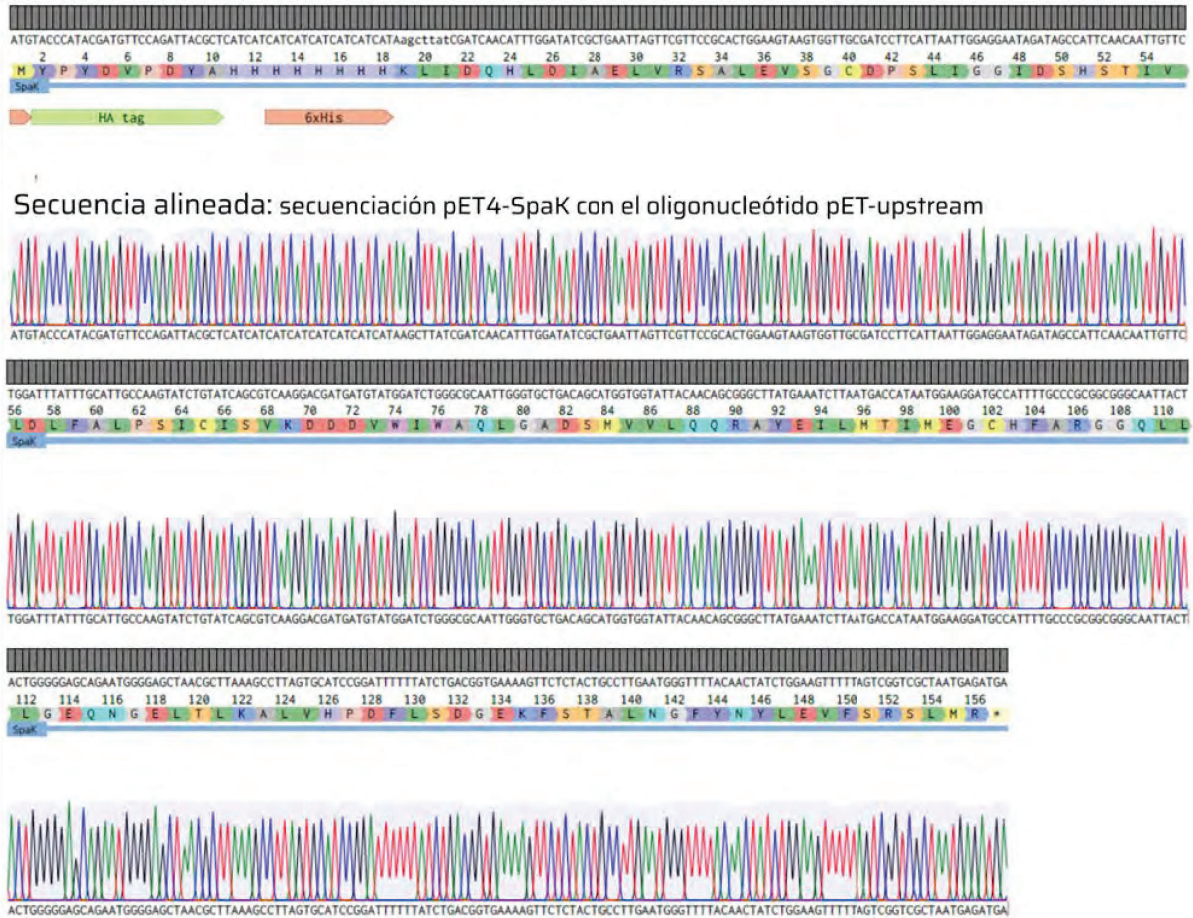
Templado: *pET4-SicP in silico*



**Figura Suplementaria 5. Secuenciación del constructo pET24-SicP.** Se empleó el oligonucleótido pET-upstream. Alineado con el gen *SicP* de *Salmonella enterica subsp. enterica serovar Typhimurium str. LT2* (ID: AE006468).

## Secuenciación pET24-SpaK

Templado: pET4-SpaK *in silico*



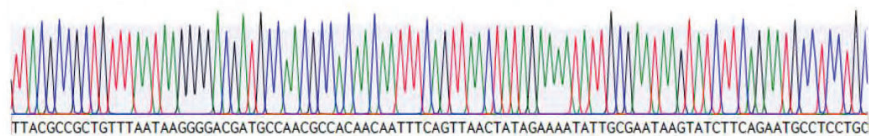
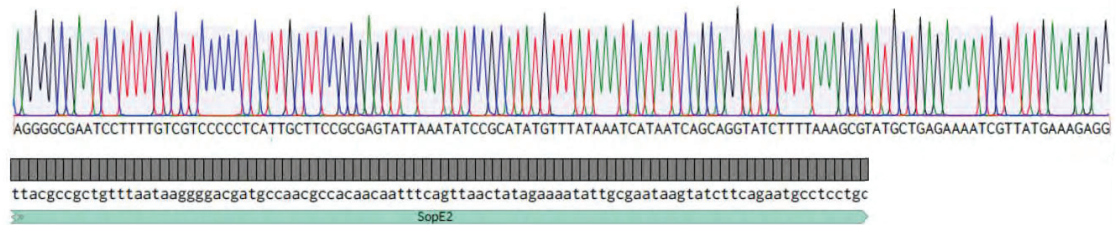
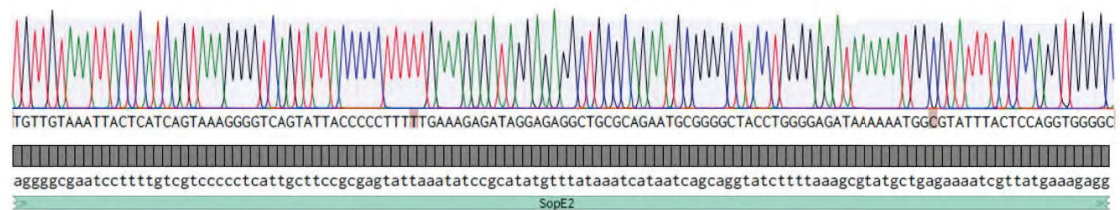
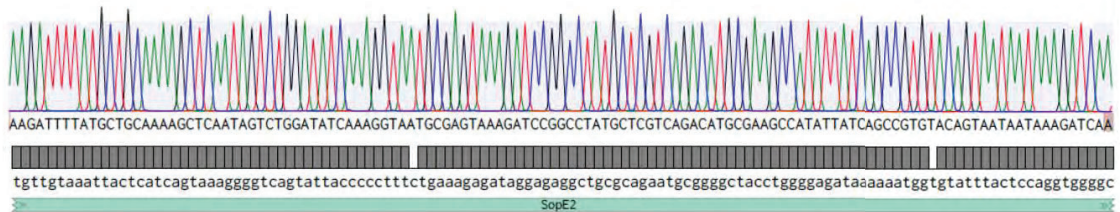
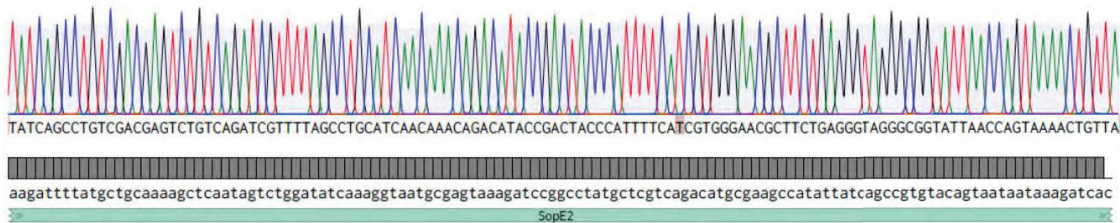
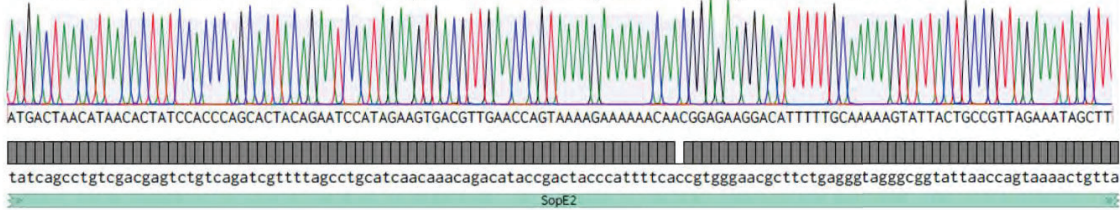
**Figura Suplementaria 6. Secuenciación del constructo pET24-SpaK.** Se empleó el oligonucleótido pET-upstream. Alineado con el gen SpaK de *Salmonella enterica subsp. enterica serovar Typhimurium str. LT2* (ID: AE006468).

## Secuenciación pGEM T easy - SopE2

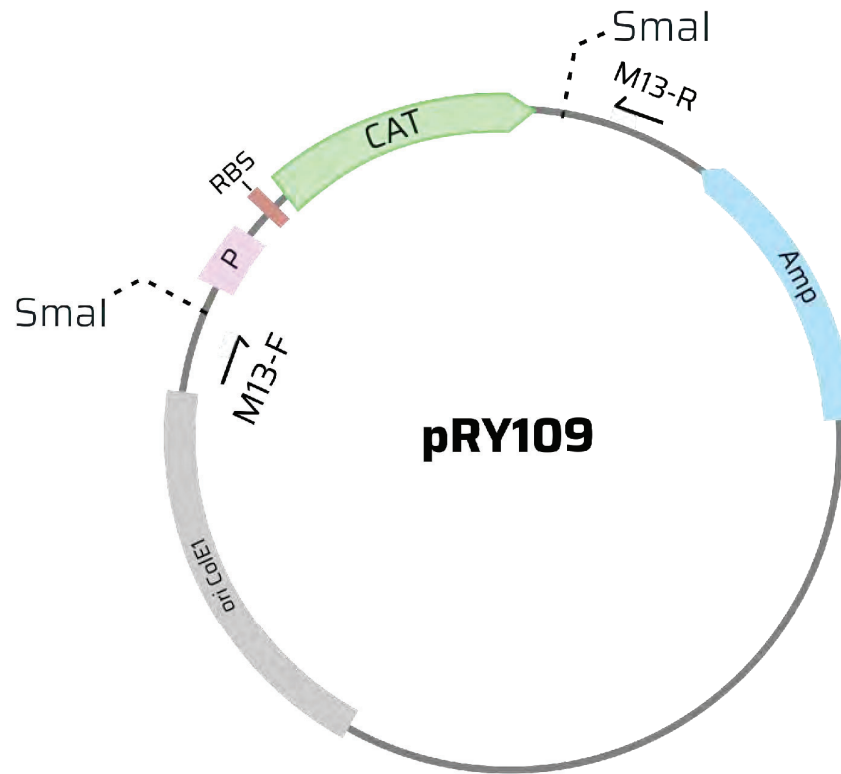
Templado: pGEM T easy - SopE2

atgactaacataacactatccaccgcactacagaatccatagaagtgacgttgaaccagtaaaagaaaaaacacggagaaggacatTTTTGCAAAAAGTATTACTGCCGTTAGAAAAGCTT  
SopE2

Secuencia alineada: secuenciación pGEM T easy - SopE2 con el oligonucleótido M13F



**Figura Suplementaria 7. Secuenciación del constructo pGEM T easy-SopE2.** Se empleó el oligonucleótido pET-upstream y alineó con el gen SopE2 (ID: AF200952).



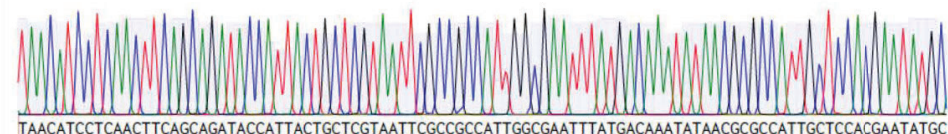
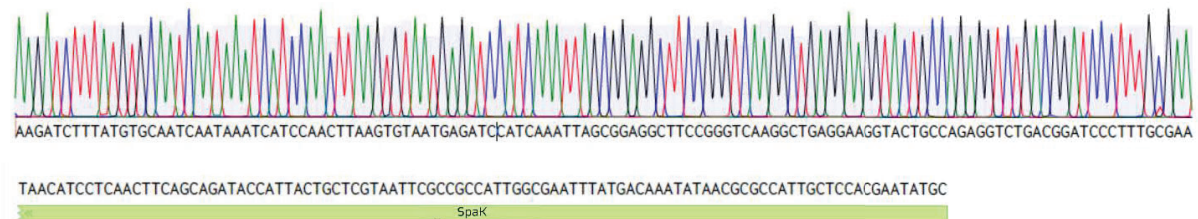
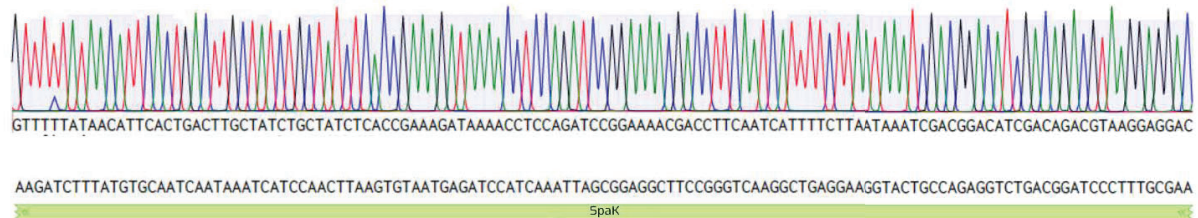
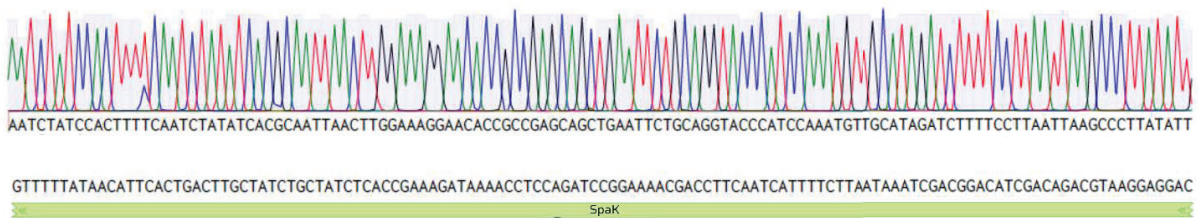
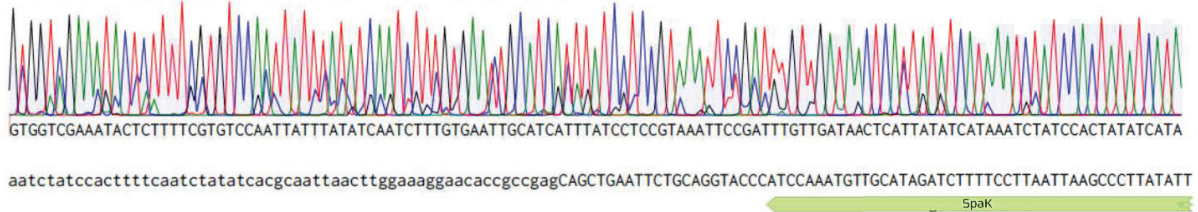
**Figura Suplementaria 8. Esquema del vector pRY109 que contiene cassette de resistencia a cloranfenicol.** P: promotor; Amp: gen que otorga resistencia a ampicilina; ori ColE1: origen de replicación ColE1. Los sitios de corte para la enzima de SmaI se indican con líneas punteadas, y los oligonucleótidos, con flechas.

## Secuenciación pGEM T easy-SpaK::CAT

Templado: pGEM T easy-SpaK::CAT

gtggtcgaataactcttttcgtgtccaattatttatatcaatctttgtgaattgcatcatttatcctccgtaaattccgatttggataactcattatatcataaactatccactatcata  
 mRNA (chloramphenicol acetyltransferase mRNA)  
 chloramphenicol acetyltransferase (cat) promoter


Secuencia alineada: secuenciación pGEM T easy-SpaK::CAT



**Figura Suplementaria 9. Secuenciación del constructo pGEM T easy-SpaK::CAT.** Se empleó el oligonucleótido SP6 y se alineó con el gen SpaK de *Salmonella enterica* subsp. *enterica* serovar *Typhimurium* str. LT2 (ID: AE006468) y una fracción del cassette de resistencia de cloranfenicol.

## Última Versión de la Tesis

Este manuscrito corresponde a la última versión de la tesis, la cual incluye todas las correcciones y sugerencias proporcionadas por los jurados.

Firma del autor: \_\_\_\_\_

Lucia Celay

Firma del director: \_\_\_\_\_

Carlos A. Buscaglia