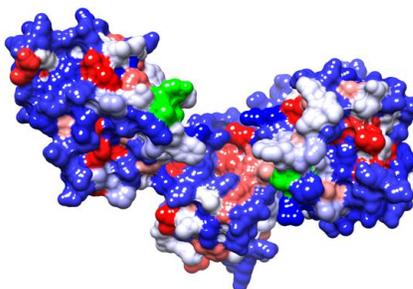


UNIVERSIDAD DEL QUINDÍO
FACULTAD DE CIENCIAS DE LA SALUD
DOCTORADO EN CIENCIAS BIOMÉDICAS

Identificación *in silico*, perfil de expresión y búsqueda de inhibidores de la familia de proteínas disulfuro isomerasa de *Toxoplasma gondii*



Tesis de Doctorado presentada como requisito parcial para optar al título de Doctor en Ciencias Biomédicas

Diego Mauricio Moncada Giraldo

Estudiante Programa de Doctorado en Ciencias Biomédicas

Dirigida por:

Jorge Enrique Gómez Marín. MD., PhD

Comité tutorial

Juan Carlos Sepúlveda. MD., PhD.

Nelson Enrique Arenas. MSc., PhD.

ARMENIA

Abril de 2018

Identificación *in silico*, perfil de expresión y búsqueda de inhibidores de la familia de proteínas disulfuro isomerasa de *Toxoplasma gondii*

Con un tercio de la población mundial infectada por *Toxoplasma gondii*, las infecciones oportunistas causadas por este parásito son una causa frecuente de complicaciones clínicas en personas inmunocomprometidas. Actualmente los tratamientos se enfocan en infecciones agudas, pero no son efectivos contra las demás etapas del ciclo de vida del parásito. Para identificar nuevos blancos farmacológicos, se realizó una comparación de niveles de expresión génica de múltiples etapas durante del ciclo de vida del parásito y se identificaron ortólogos en las especies de coccidios, *Neospora caninum*, *Sarcocystis neurona*, *Hammondia hammondi* y *Eimeria tenella*. Se identificaron 23 proteínas en el transcriptoma de *T. gondii* ME49, 23 con sintenia en *N. caninum*, *S. neurona*, *H. hammondi* y 3 en *E. tenella*. Estas proteínas contienen al menos un dominio de tiorredoxina, que es característico de la familia de proteínas disulfuro isomerasa (PDI), con frecuencias variables de dominios de tiorredoxina y diferentes arquitecturas que se comparten entre estos parásitos. Se describen los perfiles de expresión de proteínas estadio específico durante los estadios taquizoito, bradizoito y de ooquistes y los de proteínas multi-estadio. Se propone a la proteína codificada por la secuencia genómica TGME49_211680 como un blanco farmacológico dado que tiene expresión multi-estadio y es la proteína con más expresión en las cepas arquetípicas durante la infección tanto de células humanas como de ratones en estadios crónicos y agudos y por tener evidencia de ser esencial para el estadio taquizoito. Encontramos en ensayos de invasión en fibroblastos de prepucio humano con bacitracina, un inhibidor conocido para proteínas PDI, reducción de los porcentajes de invasión de los taquizoitos. Estos resultados sugieren que la familia PDI tiene un papel en la invasión de taquizoitos y podría ser un blanco para nuevas intervenciones terapéuticas. Además, debido a la alta conservación de las

secuencias de esta familia de proteínas entre los coccidios, la familia PDI podría ser un blanco farmacológico importante para estos parásitos.

In silico* Identification and Expression Profiling of the Protein Disulfide Isomerase Gene Family in *Toxoplasma gondii

With one third of the world's population infected by *Toxoplasma gondii*, opportunistic infections caused by this parasite are a frequent cause of clinical complications in immunocompromised individuals. Current treatments are focused on acute infections, but they are not effective against the many stages of the parasite's life cycle. To identify new pharmacological targets, we perform a comparison of multistage gene expression levels over the life cycle of the parasite by using *in silico* approaches. Comparison with other coccidian species, *Neospora caninum*, *Sarcocystis neurona*, *Hammondia hammondi* and *Eimeria tenella* revealed that the protein disulfide isomerase (PDI) family might be an important target for these parasites. We used a combination of approaches to evaluate PDIs as a viable target for pharmaceutical intervention. We identified 20 proteins in the proteome of *T. gondii* ME49, 20 with synteny in *N. caninum*, *S. neurona*, *H. hammondi* and 3 in *E. tenella*. The same domain architecture is shared among these parasites. These proteins contain at least one thioredoxin domain, as this is characteristic of the PDI family, with variable numbers of thioredoxin domains in different architectures and combinations. We identified proteins in the *T. gondii* life cycle that are expressed only during specific stages (tachyzoite and cat enteroepithelial stages). We found that the protein TGME49_211680 has a multistage expression and it is the most expressed in archetypal strains during infection of human cells as well as during mice infection in chronic and acute stages. Invasion assays in human foreskin fibroblast cells with bacitracin, a known inhibitor for PDI proteins, showed reduced invasion percentages by tachyzoites. These results suggest that the PDI family has a role in tachyzoite invasion and could be a target for new therapeutic interventions. Furthermore, due to the high sequence conservation between coccidians, PDIs could also be a target against other coccidia.

Agradecimientos

a los programas de maestría y doctorado en ciencias biomédicas de la universidad del Quindío. Agradecimientos especiales y una expresión admiración para el doctor Jorge Enrique Gómez Marín quien es un ejemplo que seguir. Además, gracias al Dr Gomez por su valioso aporte en mi formación profesional y su incondicional apoyo.

Al Doctor Jhon Carlos Castaño Osorio por su aporte a la formación de todos los estudiantes de La escuela en ciencias biomédicas y particularmente por todas las discusiones y el tiempo compartido conmigo que siempre fue enriquecedor y divertido.

A la doctora Olga Alicia Nieto Cárdenas, Gracias por todo su apoyo, honestidad y trabajo duro y comportamiento intachable. Gracias por sus enseñanzas investigativas y administrativas y su constante acompañamiento.

Al Doctor Aylan Farid arenas Soto.

Al Doctor German Alberto Tellez Ramirez.

Al Doctor Néstor Ivan Cardona Pérez.

Al Doctor Leonardo Padilla.

Al Magister Diego Alejandro Molina Lara.

Al Magister Alejandro Hernández De Los Ríos.

Al Magister Jhon Alejandro Acosta Dávila.

gracias por todo.

A todo el personal administrativo, a Luz Piedad Muriel Palacio, Olga Lucia R,
Yulieth por su paciencia y ayuda constante, agradecimientos infinitos.

Tabla de Contenidos

Introducción	13
Materiales y métodos.....	17
Identificación de la familia de proteínas disulfuro isomerasa (PDI).....	17
Obtención de transcriptomas de la familia PDI	17
Análisis de transcriptomas	18
Fenotipo del Knockout de la familia de proteínas disulfuro isomerasa	18
Producción de proteína recombinante.	19
Ensayos de inhibición	19
Ensayos Inmunofluorescencia	20
Cuantificación relativa de la PDI TGME49_211680 por RT-PCR	20
Búsqueda in silico de alto rendimiento de nuevos candidatos de inhibición de la proteína TGME49_211680.....	20
Resultados.....	21
Identificación de la familia de proteínas disulfuro isomerasa (PDI).....	21
Estadio entero epitelial de gato.....	31
Perfil de expresión de las cepas ME49, VEG, GT1 taquizoitos	27
Perfil de expresión en las formas clínicas agudo y crónico.....	35
Fenotipo del knockout de la familia de proteínas PDI	37
Ensayos de invasión e interacción proteína células hospedadoras	40
Análisis estructural de la proteína PDI TGME49_211680.....	44
Regiones funcionales.....	44
Dominios activos.....	46
Dominios inactivos	49

Búsqueda in silico de alto rendimiento de nuevos candidatos de inhibición de la proteína TGME49_211680.....	52
Dominio a'	52
Dominio b' y región comprendida entre dominio b' y enlazador x.....	55
Discusión	61
Perfiles transcriptómicos de la familia de proteínas disulfuro isomerasa	64
Fenotipo de los knockout producidos para la familia de proteínas disulfuro isomerasa de T gondii GT1.....	65
Análisis funcional de la proteína TGME49_211680	66
Análisis estructural y búsqueda in silico de alto rendimiento de moléculas con potencial de inhibición de la proteína TGME49_211680.....	67
Regiones funcionales.....	¡Error! Marcador no definido.
Dominio a'	¡Error! Marcador no definido.
Conclusiones	77
Bibliografía.....	79

Lista de figuras

Figura 1 Representación esquemática de los dominios y arquitecturas de las proteínas disulfuro isomerasa de <i>T gondii</i>	22
Figura 2 Alineamiento múltiple de secuencias de dominios tipo a de la familia PDI de <i>T gondii</i>	24
Figura 3 Consenso de los aminoácidos del centro activo. Obtenido de alineamiento múltiple de secuencias de dominios tipo a de la familia PDI de <i>T gondii</i>	25
Figura 4 Componentes principales de los perfiles de expresión de la familia de proteínas disulfuro isomerasa de <i>T gondii</i> desde los transcriptomas de las cepas GT1, ME49 y VEG durante infección a células HFF.	28
Figura 5 Mapas de calor de expresión de la familia de proteínas disulfuro isomerasa de <i>T gondii</i> desde los transcriptomas de las cepas GT1, ME49 y VEG durante infección a células HFF.....	29
Figura 6 A. Análisis de diferencia de medianas Kruskal-Wallis.....	30
Figura 7 Componentes principales de los perfiles de expresión de la familia de proteínas disulfuro isomerasa de <i>T gondii</i> , transcriptoma de estadios entero epiteliales de gato.....	32
Figura 8 Mapas de calor de los perfiles de expresión de la familia de proteínas disulfuro isomerasa de <i>T gondii</i> , transcriptoma de estadios entero epiteliales de gato.....	33
Figura 9 Series de tiempo en los días 3, 5 y 7 y control (Taquizoito como punto 9) de los transcritos de la familia de proteínas disulfuro isomerasa de <i>T gondii</i> en los estadios entero epiteliales de gato.	34
Figura 10 Componentes principales de los perfiles de expresión de la familia de proteínas disulfuro isomerasa de <i>T gondii</i> , transcriptoma de infección aguda y crónica en cerebro de ratón.	35

Figura 11 Mapa de calor de los perfiles de expresión de la familia de proteínas disulfuro isomerasa de <i>T gondii</i> , transcriptoma de infección aguda y crónica en cerebro de ratón.	36
Figura 12 Mapa de calor de los perfiles de expresión de la familia de proteínas disulfuro isomerasa de <i>T gondii</i> , transcriptoma de infección aguda y crónica en cerebro de ratón.	37
Figura 13 Fenotipo de los <i>knockout</i> producidos para la familia de proteínas disulfuro isomerasa de <i>T gondii</i>	38
Figura 14 Evidencia <i>in silico</i> y experimental de la localización sub celular de la familia de proteínas PDI de <i>T gondii</i>	39
Figura 15 Ontología de gen.	40
Figura 16 Electroforesis de poliacrilamida al 12% y western blot.	41
Figura 17 porcentaje de células infectadas durante la infección <i>in vitro</i> de células HeLa por Taquizoitos de <i>T gondii</i> en una multiplicidad de infección de 5:1.	42
Figura 18 Gráfico de barras de la cuantificación relativa de los transcritos de la proteína TGME49_ 211680.	43
Figura 19 Inmunofluorescencia de la proteína recombinante TGME49_ 211680 en células HeLa y Müller.	44
Figura 20 Estructura general de PDI TGME49_211680.	46
Figura 21 Aminoácidos conservados en los dominios activos de PDI TGME49_211680.	48
Figura 22 Comparación del perfil de hidrofobicidad de los dominios b y b' de PDI TGME49_211680 y las proteínas humanas PDI A1 y ERP 57.	50
Figura 23 Perfil de hidrofobicidad y potencial electrostático de PDI TGME49_211680.	50
Figura 24 Representación gráfica de las regiones objetivo de búsqueda de inhibidores para PDI TGME49_211680.	52
Figura 25 Acoplamientos moleculares de ligandos con potencial inhibitorio sobre el sitio activo CGYC de PDI TGME49_211680.	54

Figura 26 Diagrama de interacciones intermoleculares entre PDI TGME49_211680 y ligandos con potencial inhibitorio sobre centro activo CGYC.	54
Figura 27 Acoplamiento molecular de ligandos con potencial inhibitorio sobre el dominio b' de PDI TGME49_211680.....	55
Figura 28 Acoplamiento molecular de ligando con potencial inhibitorio sobre la región comprendida entre el dominio b' y el enlazador x de la proteína PDI TGME49_211680.....	56
Figura 29 Diagrama de interacciones intermoleculares entre PDI TGME49_211680 y ligandos con potencial inhibitorio sobre bolsillo hidrofóbico del dominio b' y región comprendida entre el dominio b' y enlazador x de PDI TGME49_211680.	58

Lista de tablas

Tabla 1. Resumen de las regiones funcionales de la familia PDI.	23
Tabla 2 Ubicación de la familia PDI por cromosoma en la cepa ME49 de <i>T gondii</i>	26
Tabla 3 Arquitectura conservada y sintenia entre los parásitos coccidios <i>N caninum</i> , <i>S neurona</i> , <i>H hammodi</i> y <i>E tenella</i> . Cada círculo representa la sintenia	27
Table 4 Energías de acoplamiento e interacciones establecidas por moléculas con potencial inhibitorio en el motivo catalítico CGYC de PDI TGME49_211680 y su comparación con las proteínas PDIA1 y ERP 57.	55
Tabla 5 Energías de acoplamiento de moléculas con potencial inhibitorio sobre sitio de interacción del sustrato y región comprendida entre el dominio b' y el enlazador x de las proteínas disulfuro isomerasa PDI TGME49_211680, PDIA1 y ERP 57.	59
Tabla 6 Características de venta de moléculas identificadas como potenciales inhibidores de la proteína PDI TGME49_211680 de <i>Toxoplasma gondii</i>	¡Error! Marcador no definido.

Introducción

Aislado por primera vez desde el roedor africano *Ctenodactylus gundii* en 1908 por Nicolle C y Manceaux, *Toxoplasma gondii* es el único miembro del género *Toxoplasma* (Hunter and Sibley, 2012) y hace parte del phylum Apicomplexa que contiene más de 5000 especies, donde la mayoría de estas especies tienen un rango estrecho de hospedares, sin embargo a diferencia de otros parásitos cercanamente relacionados como *Neospora caninum*, *Sarcocystis neurona* y *Hammondia hammondi* que presentan distribuciones espaciales reducidas y rangos de hospedador especializados, *T gondii* cuenta con una amplia distribución, y es uno de los parásitos más prevalentes en el mundo, infectando una gran cantidad de animales salvajes y domésticos, causando infecciones zoonóticas en humanos, siendo una infección común, que generalmente causa enfermedad en hospedares inmunocromprometidos o como resultado de una infección transplacentaria (Montoya and Liesenfeld, 2004; Lorenzi *et al.*, 2016)

Además, el ciclo de vida heteroxeno de *T. gondii* contiene diferentes estadios celulares que van desde el estadio merozoito, los estadios sexuales con macro y microgametas hasta ooquiste, taquizoito y bradizoito los cuales presentan perfiles de expresión diferencial de proteínas y morfología celulares particulares (Hunter and Sibley, 2012; Behnke *et al.*, 2014). Para cumplir esta diversidad de funciones en diferentes estadios, son necesarias la presencia de chaperonas e isomerasas multiestadío, las cuales permiten el correcto plegamiento de las proteínas que participan en el cumplimiento de las diversas funciones necesarias para la infección parasitaria, en cada etapa del ciclo de vida.

La proteína disulfuro isomerasa fue identificada por primera vez en los años 60's (Goldberger, Epstein and Anfinsen, 1963, 1964) y fue descrita como un producto que catalizaba la reactivación de enzimas reducidas como

ribonucleasa y lisozima, actualmente. Las proteínas disulfuro isomerasas se describen como proteínas multifuncionales que tienen diferentes papeles fisiológicos y celulares a través de la oxido-reducción de enlaces disulfuro, presentando actividad isomerasas y chaperona, en este contexto diferentes organismos desarrollaron subgrupos de proteínas tioredoxinas especializados que se encuentran en diferentes localizaciones extracelulares, como el núcleo, retículo endoplasmático, citosol, membrana celular o ser productos de secreción, (Turano *et al.*, 2002; Collet and Messens, 2010; Sambasivarao, 2013; Porter, Yuen and Christopher, 2015; Soares Moretti and Martins Laurindo, 2017) además de presentar diferentes rangos de actividad dependiendo del ambiente reductor en donde este ubicada lo cual también está relacionado con el sustrato sobre el cual actúa mostrando que son proteínas con múltiples sustratos y rangos de actividad (Collet and Messens, 2010; Krossa *et al.*, 2018).

Específicamente el orden de la alineación de puentes disulfuros es esencial para la estabilidad estructural y la integridad funcional de un amplio rango de proteínas siendo clave para las proteínas destinadas a la secreción (Safavi-Hemami *et al.*, 2016), La formación de enlaces disulfuro es principalmente mediada por oxido reductasas incluyendo la familia de proteínas disulfuro isomerasa (PDI) (Sevier and Kaiser, 2002; Haque, Majumdar and Barik, 2012; Bechtel, Weerapana and Hill, 2017). Las proteínas de familia PDI son miembros de la super familia de tioredoxinas tiol-oxidoreductasas (Soares Moretti and Martins Laurindo, 2017) formada por arquitecturas que contienen al menos un dominio con plegamiento tioredoxina ($\beta\alpha\ \beta\alpha\ \beta\alpha\ \beta\beta\alpha$), el cual puede variar en tamaño y organización (Kozlov *et al.*, 2010),

En eucariotas esta familia presenta múltiples localizaciones subcelulares que van desde el retículo endoplasmático y el aparato de Golgi en relación a la presencia de dominios KDEL en su extremo carboxilo terminal, hasta el citosol, membranas o localizaciones extracelulares, donde presenta

diferentes sustratos y dependiendo de su localización pueden ser chaperonas, isomerasas, oxidasas o reductasas de puentes disulfuro (Turano *et al.*, 2002; Soares Moretti and Martins Laurindo, 2017; Hettinghouse, Liu and Liu, 2018; Krossa *et al.*, 2018; Popielarski *et al.*, 2018)

En el contexto de las infecciones por parásitos como *N. caninum*, *Plasmodium falciparum*, *Entamoeba histolytica*, *Trypanosoma cruzi*, *Leishmania amazonensis*, *L. donovani*, *L. major* la formación reversible de enlaces disulfuro es clave para regular la estructura y función de un amplio rango de proteínas donde se incluyen factores de virulencia de parásitos protozoos, donde se ha descrito a las proteínas de la familia PDI como esenciales para el crecimiento del parásito, en algunos casos para su invasión, así como factores de virulencia y candidatos vacunales (Achour *et al.*, 2002; Mahajan *et al.*, 2006; Ramos *et al.*, 2008, 2011; Haque, Majumdar and Barik, 2012)

También se ha descrito a miembros de la familia PDI como candidatos blanco de inhibición a través de tamizajes experimentales para la identificación de moléculas inhibitoras y el posterior desarrollo de nuevas estrategias de intervención contra protozoos de los géneros *Leishmania*, *Plasmodium* y de las especies *N. caninum* y *Besnoitia besnoiti* (Mouray *et al.*, 2007; Müller *et al.*, 2008; Novo *et al.*, 2009; Ben Khalaf *et al.*, 2011, 2012)

Particularmente en parásitos apicomplexos se han descrito proteínas de la familia PDI en *N. caninum*, *Besnoitia besnoiti*, *Eimeria tenella* (Naguleswaran, Müller and Hemphill, 2003; Naguleswaran *et al.*, 2005; Müller *et al.*, 2008; Marcelino *et al.*, 2011; Han *et al.*, 2014): donde para en *T. gondii* se han descrito 5 proteínas con dominios tiorredoxina, una es la proteína tiorredoxina ATrx1 localizada en el apicoplasto que se sugiere está involucrada en el transporte de proteínas hacia este organelo (DeRocher *et al.*, 2008), con el código de identificación TGME49_312110. Dos de estas relacionadas con los microtúbulos corticales del parásito (Liu *et al.*, 2013) con

códigos de identificación TGME49_225790 y TGME49_232410. También hay dos proteínas de la familia que se describen como altamente inmunogénicas, una identificada como el mayor receptor de anticuerpos en productos de secreción del parásito (Saadatnia *et al.*, 2012; Tommy *et al.*, 2012) con código de identificación TGME49_293870 y la proteína TGME49_211680, la cual ha sido la más estudiada de esta familia para la cual se han identificado anticuerpos IgA en lágrimas. También se ha descrito reactividad cruzada de reconocimiento de anticuerpos en con su ortólogo en *N. caninum* y *B. besnoiti* (Marcelino *et al.*, 2011; Zhang *et al.*, 2011)

Además la proteína de la familia PDI TGME49_211680 se ha propuesto también como candidato vacunal y se ha descrito reactividad ante esta misma proteína de anticuerpos en sueros de ratón, conejo y humanos mostrando su inmonogenicidad en diferentes especies (Wang *et al.*, 2013). Funcionalmente se sugiere que esta proteína en *T. gondii* y su ortólogo en *N. caninum* esta relaciona con la interacción entre taquizoito y célula hospedadora (Naguleswaran *et al.*, 2005; Moncada, Arenas, Acosta, Molina, Hernández, *et al.*, 2016).

Este trabajo tuvo como objetivo describir las características de secuencia genómica y de aminoácidos y de arquitectura de las proteínas de la familia de proteínas disulfuro isomerasa de *T. gondii* y los perfiles de expresión de estas a través del análisis de transcriptomas multiestadio, los fenotipos de su correspondiente *knockout* y análisis *in silico*, con el fin de ofrecer un modelo conceptual de su bioquímica y funcionalidad para esta familia de proteínas en *T. gondii*, así como describir sus ortólogos en las especies de coccidios *N. caninum*, *S. neurona*, *H. hammondi* y *E. tenella*. Por último, se realizó un tamizaje computacional de alto de desempeño para la búsqueda de moléculas inhibitoras de las proteínas de importancia para *T. gondii* de la familia PDI.

Materiales y métodos

Identificación de la familia de proteínas disulfuro isomerasa (PDI)

Se realizaron alineamientos de proteínas en la base de datos *ToxoDB* versión 32 (Gajria *et al.*, 2008; Aurrecochea *et al.*, 2017) usando como molde las proteínas TGME49_211680 y TGME49_312110 de *T gondii* sobre el proteoma de la cepa ME49 y se realizó una búsqueda por texto con la palabra clave *Thioredoxin*. Las proteínas obtenidas por estas estrategias fueron analizadas en los identificadores de dominios SMART, PROSITE y NCBI CD, con el fin de verificar la presencia de dominios tiorredoxina, y se predijo la localización subcelular con Euk-mPLoc (Chou and Shen, 2008) Posteriormente se realizó una estrategia anidada de búsqueda en la base de datos, para identificar secuencias asociadas a la localización subcelular como péptido señal y regiones transmembranales. Así mismo, se identificó la localización genómica de cada secuencia, la sintenia y la relación entre secuencias de aminoácidos (Dereeper *et al.*, 2008) de estos genes en los coccidios *Neospora caninum*, *Eimeria tenella*, *Sarcocystis neurona* y *Hammondia hammondi*. Por último, se identificaron las secuencias con mayor porcentaje de identidad en la base de datos PDB.

Obtención de transcriptomas de la familia PDI

Se usaron en total 6 perfiles de transcriptomas asociados a cada proteína desde los experimentos: I- Series de tiempo de los transcriptomas de taquizoito de las cepa ME49, VEG y GT1 (Lorenzi *et al.*, 2016), II- transcriptomas de los estadios entero epiteliales de gato (Hehl *et al.*, 2015), III – transcriptoma de cerebros de ratón durante una infección aguda y crónica de *T gondii* (Pittman, Aliota and Knoll, 2014).

Análisis de transcriptomas

Se realizaron análisis de componentes principales para identificar los grupos de transcriptos con mayor variabilidad entre las cepas ME49, VEG y GT1, los estadios entero-epiteliales de gato y la infección a cerebros de ratón durante una infección aguda y crónica, utilizando el software *Infostat versión 12*. Posteriormente en *MatLab* y R se realizaron mapas de calor y rangos de expresión en gráficos de cajas para visualizar las diferencias de expresión entre la familia de proteínas para cada transcriptoma. Finalmente, se realizaron análisis de varianza para identificar la diferencia entre los perfiles de expresión de las cepas ME49, VEG y GT1, y se graficaron las series de tiempo de cada transcriptoma.

Fenotipo de cepas “Knockout” para genes de la familia de proteínas disulfuro isomerasa

A partir de los datos disponibles en ToxoDB (<http://toxodb.org>) se obtuvieron los datos del “*screening*” de genes esenciales para *T. gondii* usando CRISPR en taquizoitos modificados con CASP9. El genoma de estos parásitos mutados con metodología CRISP fueron secuenciados con la tecnología Illumina. El número de guías únicas de ARN (sgRNA) es equivalente al número de parásitos evaluados y se estableció que la pérdida de un gen reducía el *fitness* (capacidad de adaptación de la población parasitaria al medio ambiente celular) de manera significativa (o sea que es esencial) a partir del número de parásitos que sobrevivían a tres ciclos líticos en fibroblastos de prepucio humano (*Human Foreskin Fibroblast*, HFF) comparando el número de guías de las cepas *knockout* (mutadas) con la cepa original o no mutada y así se obtuvo un puntaje de fenotipo, utilizando la siguiente fórmula (Sidik *et al.*, 2016):

Puntaje de *fitness* del fenotipo = sgRNA de la cepa *knockout* / sgRNA de la librería original.

Un puntaje de fenotipo entre -6,89 a -2,96 indicaba que el gen era esencial.

Producción de proteína recombinante.

A través de análisis de transcriptomas se identificó el gen con más abundancia de transcritos y se seleccionó para la producción de una proteína recombinante (proteína de código TGME49_211680), usando los cebadores sentido (ATGCGAGCCGGGTTTTTC) y antisentido (CAGTTCTTCACCCTTGTCGTC) para su amplificación. Posteriormente se insertó la secuencia en el vector pEXP5-CT/TOPO® (Invitrogen, USA), y se transformaron las células químicamente competentes *Escherichia coli* One Shot® TOP10 para la replicación del plásmido. La proteína recombinante se produjo en *E coli* BL21 inducidas con 1,4 mM de IPTG durante 2 horas. La proteína fue purificada con el sistema PROBOND (Invitrogen, USA) y se realizó un cambio de solución para experimentos posteriores con la columna de desalinización HiTrap de General electric.(Moncada, Arenas, Acosta, Molina, Hernández, *et al.*, 2016).

Ensayos de inhibición

Se realizaron ensayos de invasión a células HeLa y Muller con taquizoitos de la cepa RH expresando constitutivamente proteína verde fluorescente (Green fluorescent Protein GFP), lo cual permitió la identificación del porcentaje de invasión por citometría de flujo al reconocerse células con y sin el parasito. Se realizaron tratamientos con concentraciones de la proteína recombinante desde 1 a 4 µg/ml. Además, se pre-trataron los taquizoitos con el inhibidor de PDI Bacitracina (Moncada, Arenas, Acosta, Molina, Hernández, *et al.*, 2016).

Ensayos Inmunofluorescencia

Para identificar si la proteína recombinante tenía interacción con la membrana celular de células HeLa, se incubó la proteína recombinante con las células durante 24 horas y posteriormente se realizó la identificación de la proteína con un anticuerpo primario anti-His y un anticuerpo secundario conjugado con GFP (Moncada, Arenas, Acosta, Molina, Hernández, *et al.*, 2016).

Cuantificación relativa de la PDI TGME49_211680 por RT-PCR

Para realizar la cuantificación relativa de la producción de transcritos para la proteína TGME49_211680 se realizó una RT-PCR con ARN de taquizoitos antes de la infección y 2 horas post- infección en células HeLa y Muller, utilizando sondas *Taqman* con el kit TaqMan® qRT-PCR en el termociclador en tiempo real Applied Biosystems 7900HT, para la detección de transcritos de la PDI y para el gen constitutivo factor de elongación (Moncada, Arenas, Acosta, Molina, Hernández, *et al.*, 2016).

Búsqueda de alto rendimiento in silico de nuevos candidatos de inhibición de la proteína TGME49_211680.

Se realizó un modelo por homología con el servidor I-Tasser (Roy, Kucukural and Zhang, 2010; Yang *et al.*, 2015) utilizando como molde la proteína con código PDB: 3F8U (Dong *et al.*, 2009). Posteriormente se caracterizaron las regiones funcionales de PDI TGME49_211680 usando las herramientas en línea NCBI CD, PROSITE, SMART, T-Coffee (Ridgeway and Nw, 2000; Sigrist *et al.*, 2002; Cucho *et al.*, 2013; Letunic and Bork, 2017; Marchler-bauer *et al.*, 2017) Una vez identificadas las regiones de importancia para la actividad de la proteína, a través del portal web DrugDiscovery@TACC se

evaluó masivamente la afinidad entre estas y 642.769 moléculas de la librería virtual ZINC con propiedades óptimas para ser utilizadas como medicamentos (Irwin and Shoichet, 2005), ejecutando el software de acoplamiento molecular AutoDock Vina 1.1.2 (Trott and Olson, 2009) en la supercomputadora “Lonestar” del centro de computación avanzada de la Universidad de Texas.

Resultados

Identificación de la familia de proteínas disulfuro isomerasa (PDI)

Se identificaron en la base de datos genómica de ToxoDB, 23 proteínas con dominios tiorredoxina en el proteoma de *T gondii* ME49, las cuales se presentan en diferentes arquitecturas que van desde 1 hasta 6 dominios tiorredoxina, en combinación con dominios de retención al retículo KDEL, Dnaj y glicosil transferasa. 11 de estas proteínas presentaron péptido señal y 6 región transmembranal y 6 de ellas presentaron dominios de tipo KDEL dos (figura 1 y Tabla 1). Todas las proteínas de la familia PDI evidenciaron en su dominio a los aminoácidos conservados necesarios para la actividad CxxC (figuras 2, 3). Para las 23 proteínas de la familia se identificó su posición genómica (Tabla 2) y su sintenia, encontrándose sintenia en las cepas GT1 y VEG de *T gondii* y en los coccidios analizados (Tabla 3).

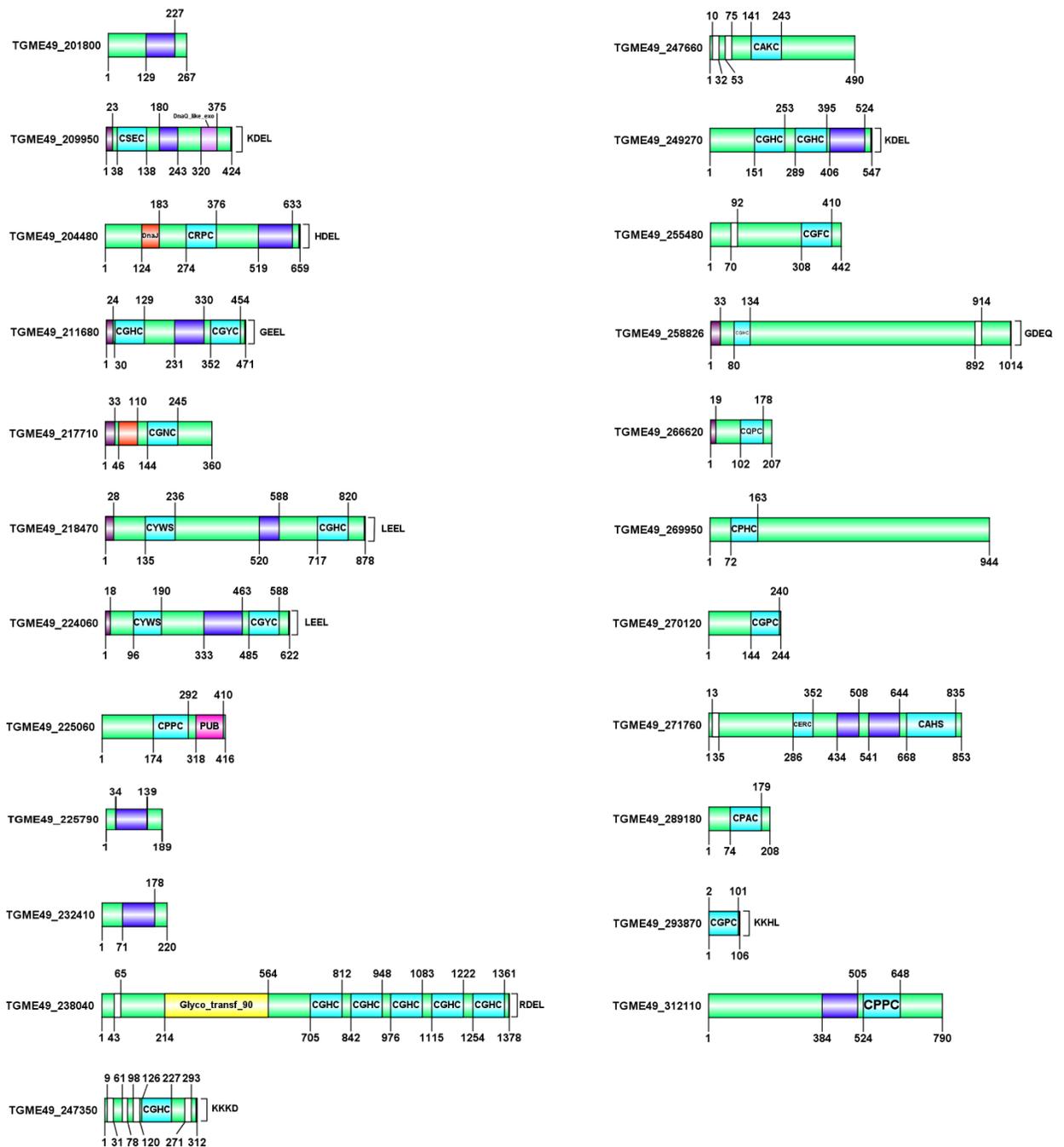


Figura 1. Representación esquemática de los dominios y arquitecturas de las proteínas disulfuro isomerasa de *T. gondii*. Se muestran los números de acceso en la base de datos ToxoDB versión 32. En azul claro se destacan los dominios activos a y en purpura los dominios inactivos b, en amarillo el péptido señal y en azul los dominios transmembranales. Además, se muestran las proteínas con dominios de retención al retículo KDEL.

Table 1. Resumen de las regiones funcionales de la familia PDI.

Numero de acceso a la secuencia en ToxoDB	Número aminoácidos	Composición de dominios	Sitio activo	Motivo de localización a reticulo endoplásmico	Arginina conservada	Péptido señal predicho
TGME49_211680	471	a-b-b'-a'	CGHC, CGYC	GEEL	No, Yes	SP
TGME49_224060	622	a-b-b'-a'	CYWS, CGYC	LEEL	No, Yes	SP
TGME49_218470	878	a-b-b'-a'	CYWS, CGHC	LEEL	No, Yes	SP
TGME49_271760	853	a-b-b'-a'	CERC, CAHS		No, No	SP
TGME49_247660	490	a-b-b'	CAKC		No	
TGME49_204480	659	a-b-b'	CRPC	HDEL	No	
TGME49_249270	547	a-a'-b	CGHG,CGHC		No	
TGME49_255480	442	b-b'-a'	CGFC		No	
TGME49_312110	790	b-a-b	CPPC		No	SP
TGME49_238040	1378	b-a-a-a'-a'-a'	CGHC, CGHC CGHC,CGHC CGHC,	RDEL	Yes, No, Yes, Yes, Yes	
TGME49_270120	244	A	CGPC		No	
TGME49_266620	207	A	CQPC	EKDK	No	SP
TGME49_247350	312	A	CGHC	KKKD	No	
TGME49_258826	1014	A	CGHC		No	SP
TGME49_293870	106	A	CGPC	KKHL	No	
TGME49_232410	220	A	DPKC		No	
TGME49_225790	189	A	HSKC		No	
TGME49_225060	416	a-PUB	CPPC		No	
TGME49_289180	208	A	CPAC		No	
TGME49_201800	267	B			No	SP
TGME49_209950	424	a-b-DNAQ	CSEC	KDEL	No	SP
TGME49_269950	944	A	CPHC		No	SP
TGME49_217710	360	DnaJ-a	CGNC		No	SP

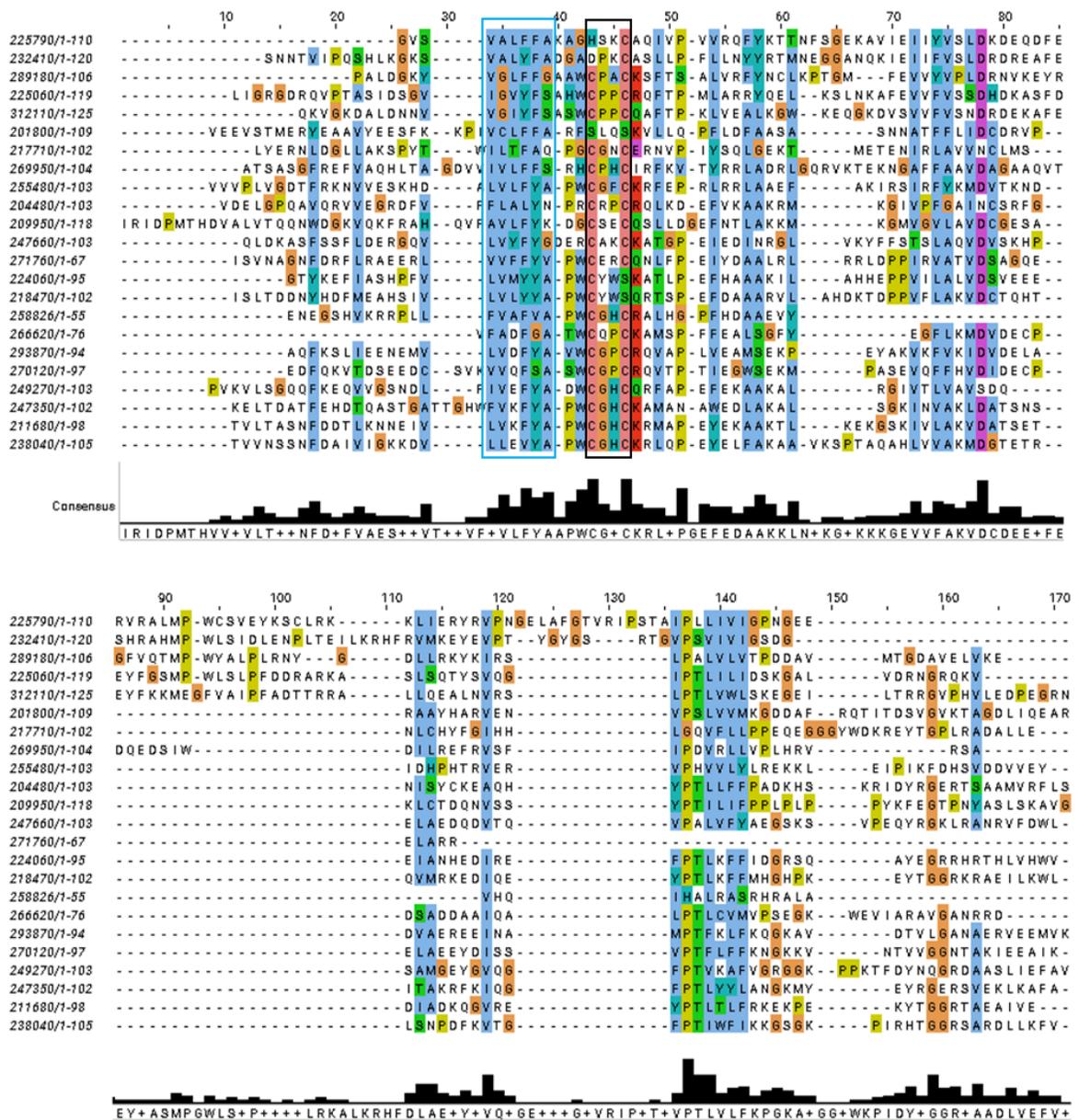


Figura 2 Alineamiento múltiple de secuencias de dominios tipo a de la familia PDI de *T. gondii*. En azul claro se muestran aminoácidos hidrofóbicos relacionados con la estabilidad del núcleo de la proteína como la lámina $\beta 2$ destacada en el recuadro azul, y en el recuadro negro se señalan los aminoácidos del centro activo CxxC.

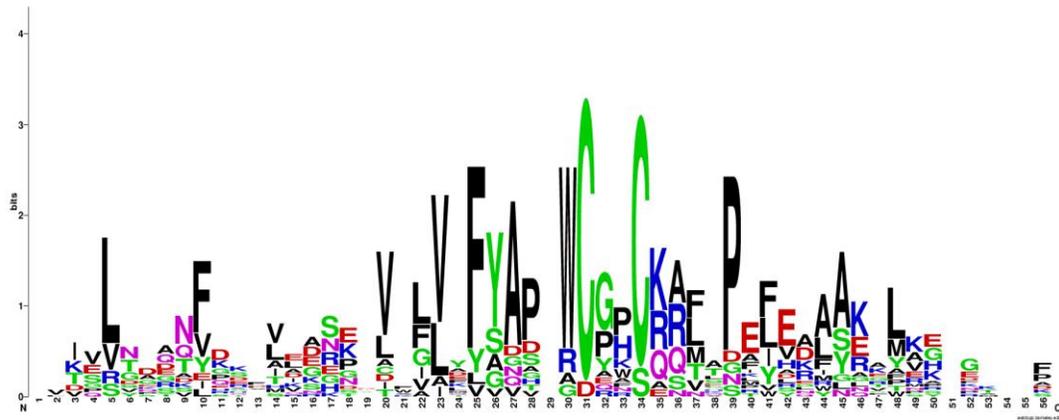


Figura 3. Aminoácidos consenso en la región del centro activo de la familia de proteínas PDI. Se realizó un alineamiento múltiple de secuencias de dominios tipo a de la familia PDI de *T gondii* a partir de las secuencias presentes en la base de datos genómica ToxoDB versión 32.

Tabla 2 Ubicación de la familia PDI por cromosoma en la cepa ME49 de *T gondii*.

Chromosome	#Genes	ToxoDB ID
TGME49_Chrla	1	TGME49_293870
TGME49_Chrlb	1	TGME49_209950
TGME49_ChrlI	0	No ID
TGME49_ChrlII	0	No ID
TGME49_ChrlV	1	TGME49_211680
TGME49_ChrlX	2	TGME49_266620, TGME49_289180
TGME49_ChrlV	0	
TGME49_ChrlVI	1	TGME49_238040
TGME49_ChrlVIIa	2	TGME49_204480, TGME49_201800
TGME49_ChrlVIIb	2	TGME49_258826, TGME49_255480
TGME49_ChrlVIII	4	TGME49_232410, TGME49_271760, TGME49_270120, TGME49_269950
TGME49_ChrlX	3	TGME49_225790, TGME49_225060, TGME49_224060
TGME49_ChrlXI	1	TGME49_312110
TGME49_ChrlXII	5	TGME49_218470, TGME49_217710, TGME49_247350, TGME49_247660, TGME49_249270

Tabla 3 Arquitectura conservada y sintenia entre los parásitos coccidios *N. caninum*, *S. neurona*, *H. hammodi* y *E. tenella*. Cada círculo representa la sintenia

ACCESSION	DOMAIN COMPOSITION	<i>Tg</i>	<i>Nc</i>	<i>Sn</i>	<i>Hm</i>	<i>Et</i>
TGME49_211680	a-b-b'-a'	•	•	•	•	•
TGME49_224060	a-b-b'-a'	•	•	•	•	
TGME49_218470	a-b-b'-a'	•	•	•	•	
TGME49_271760	a-b-b'-a'	•	•	•	•	
TGME49_247660	a-b-b'	•	•	•	•	
TGME49_204480	DnaJ a-b-b'	•	•	•	•	
TGME49_249270	a-a'-b	•	•	•	•	
TGME49_255480	b-b'-a'	•	•	•	•	
TGME49_312110	b-a-b	•	•	•	•	•
TGME49_238040	b-a -a-a'-a'-a'	•	•	•	•	•
TGME49_270120	a	•	•	•	•	
TGME49_266620	a	•	•	•	•	
TGME49_247350	a	•	•	•	•	
TGME49_258826	a	•	•	•	•	
TGME49_293870	a	•	•	•	•	
TGME49_232410	a	•	•	•	•	
TGME49_225790	a	•	•	•	•	
TGME49_225060	a-PUB	•	•	•	•	
TGME49_289180	a	•	•	•	•	
TGME49_217710	DnaJ-a	•	•	•	•	

Perfil de expresión de las cepas ME49, VEG, GT1 taquizoitos

El análisis de componentes principales del perfil de expresión de los genes de la familia PDI de cada una de las cepas arquetípicas (ME49 representativo del linaje clonal II, VEG representativo del linaje clonal III y GT1 representativo del linaje clonal I) muestra un 80% de variabilidad de los datos en el componente 1 y un 18 % en el componente 2, los que muestra que la variabilidad en las expresiones de estas proteínas es diferente entre las tres cepas, lo cual sugiere que los grupos de proteínas de esta familia están relacionados con los perfiles de expresión y la magnitud de expresión

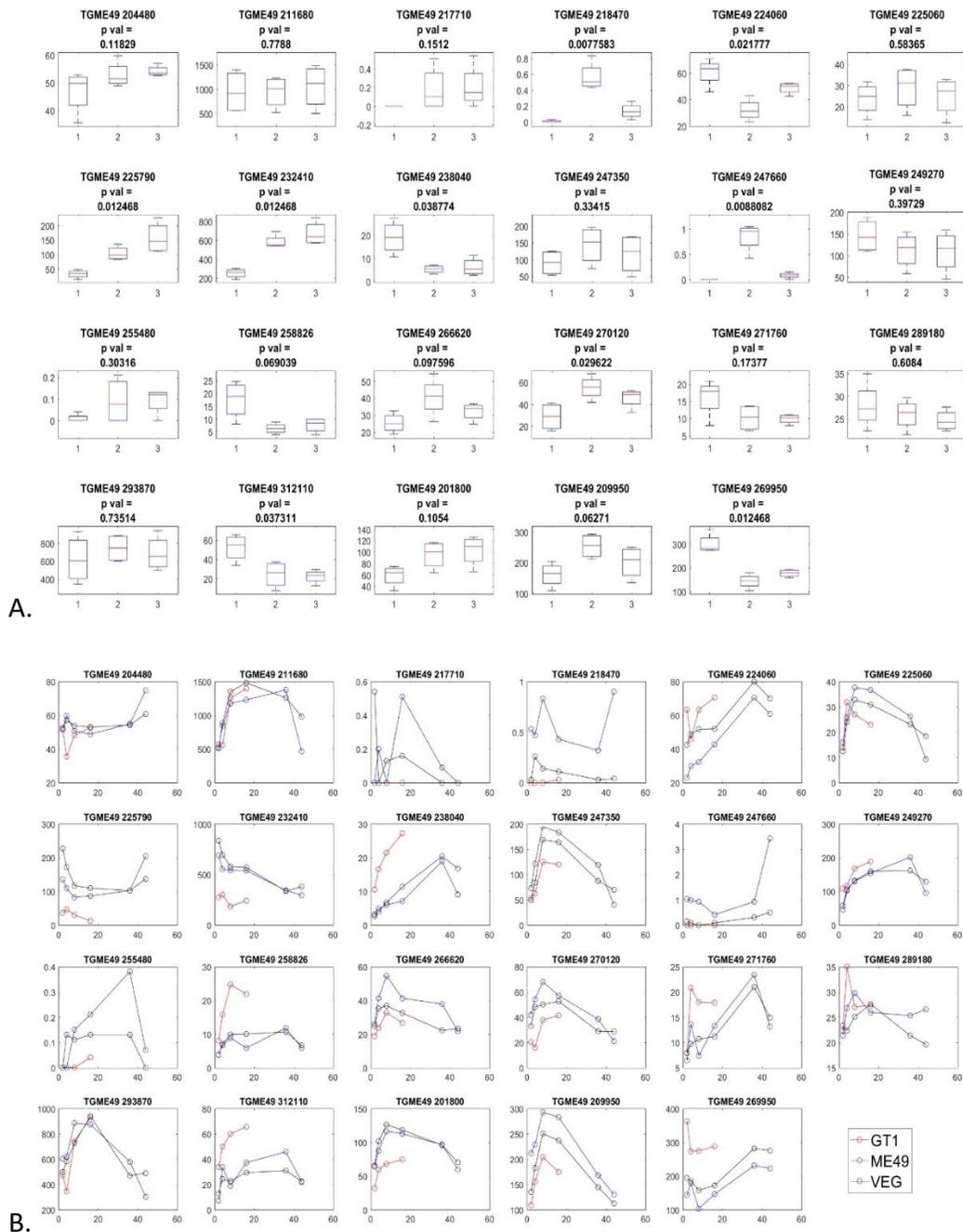


Figura 6 A. Análisis de diferencia de medianas de expresión. En la gráfica 1,2 y 3 son respectivamente, cepas GT1, ME49 y VEG. Se estableció si existían diferencias de significancia estadística para las medianas de los valores de expresión de los genes de la familia PDI con la prueba de Kruskal-Wallis. **B. Curvas de niveles de expresión de genes de la familia de PDI de *T gondii* de las cepas GT1, ME49 y VEG.** Se evaluaron los niveles

de expresión de genes luego de 2, 4, 8, 16, 36 y 44 horas post-infección a células HFF. Las proteínas de mayor expresión fueron TGME49_211680.232410 y 293870.

Expresión de genes durante infección a células entero epiteliales de gato (etapa de replicación sexuada de T. gondii)

El análisis de componentes principales muestra un 60% de variabilidad de los datos en el componente 1 y un 29% para el componente 2, lo que muestra que la expresión de los genes de estas proteínas cambia durante el estadio entero epitelial en el gato. En particular, la Figura 7 muestra que se expresan tanto en estadio entéricos como estadio taquizoito los genes TGME49_218470, TGME49_255480, TGME49_217710 y TGME49_247660. Se encontraron como genes exclusivos del estadio sexual las secuencias TGME49_238040, TGME49_258826 y TGME49_270120. Los genes con mayor expresión en estadios entero epiteliales de gato y una marcada reducción en sus niveles de expresión en el estadio taquizoito, tanto en cultivo celular como en el modelo murino de infección fueron TGME49_201800, TGME49_312110, TGME49_266620 y TGME49_247350 (Figuras 5, 6, 7, 8, 9, 11, 12 y S4). En el estadio taquizoito se encontraron las secuencias TGME49_225790, TGME49_209950, TGME49_269950, TGME49_271760, TGME49_224060, TGME49_225060 y particularmente la secuencia TGME49_225790 llegó a niveles indetectables en los estadios sexuales.

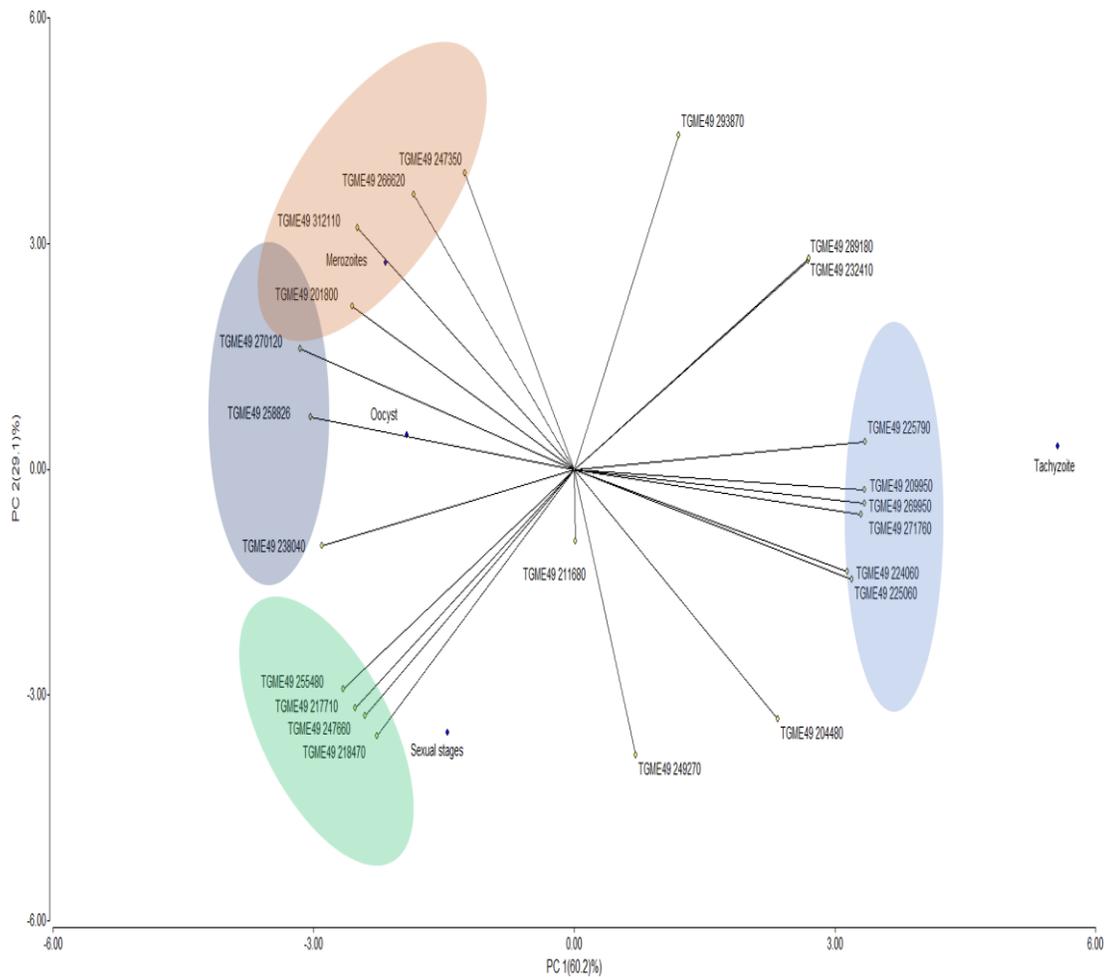


Figura 7 Componentes principales de los perfiles de expresión de la familia de proteínas disulfuro isomerasa de *T gondii*, transcriptoma de estadios entero-epiteliales de gato. 3 días post infección en gato (merozoito). 5 días post infección en gato (estadios sexuales), 7 días post infección en gato (ooquiste) y taquizoito.

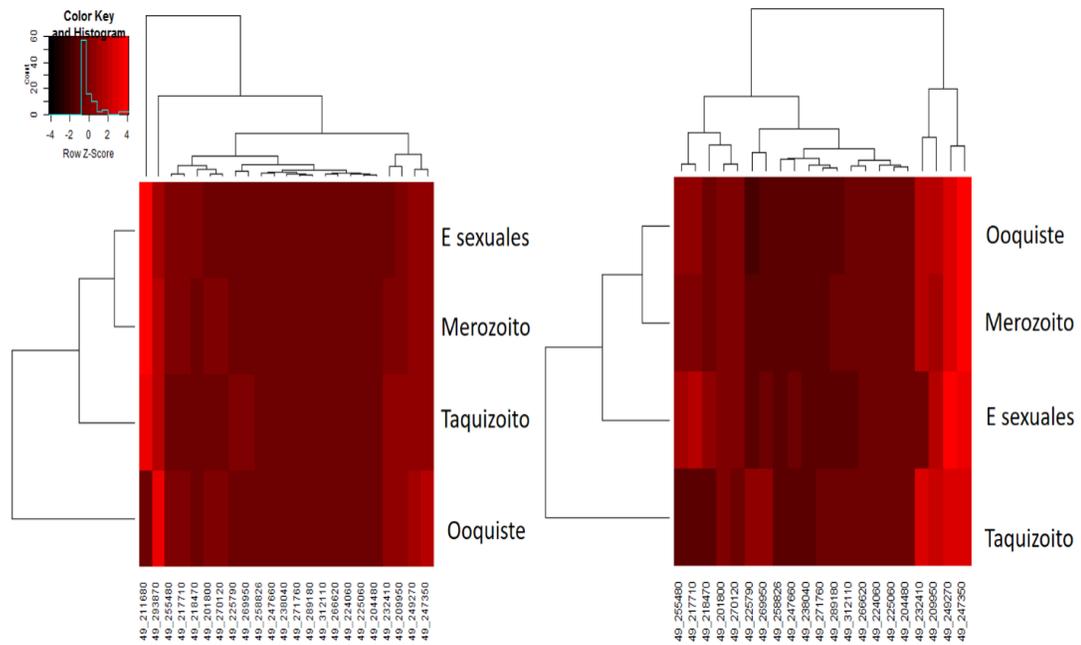


Figura 8 Mapas de calor de los perfiles de expresión de la familia de proteínas disulfuro isomerasa de *T gondii*, transcriptoma de estadios entero-epiteliales de gato. 3 días post infección en gato (Merozoito). 5 días post infección en gato (estadios sexuales), 7 días post infección en gato (Ooquiste) y Taquizoite. Al lado izquierdo el perfil completo de la familia PDI y a la derecha el perfil sin las proteínas de mayor expresión.

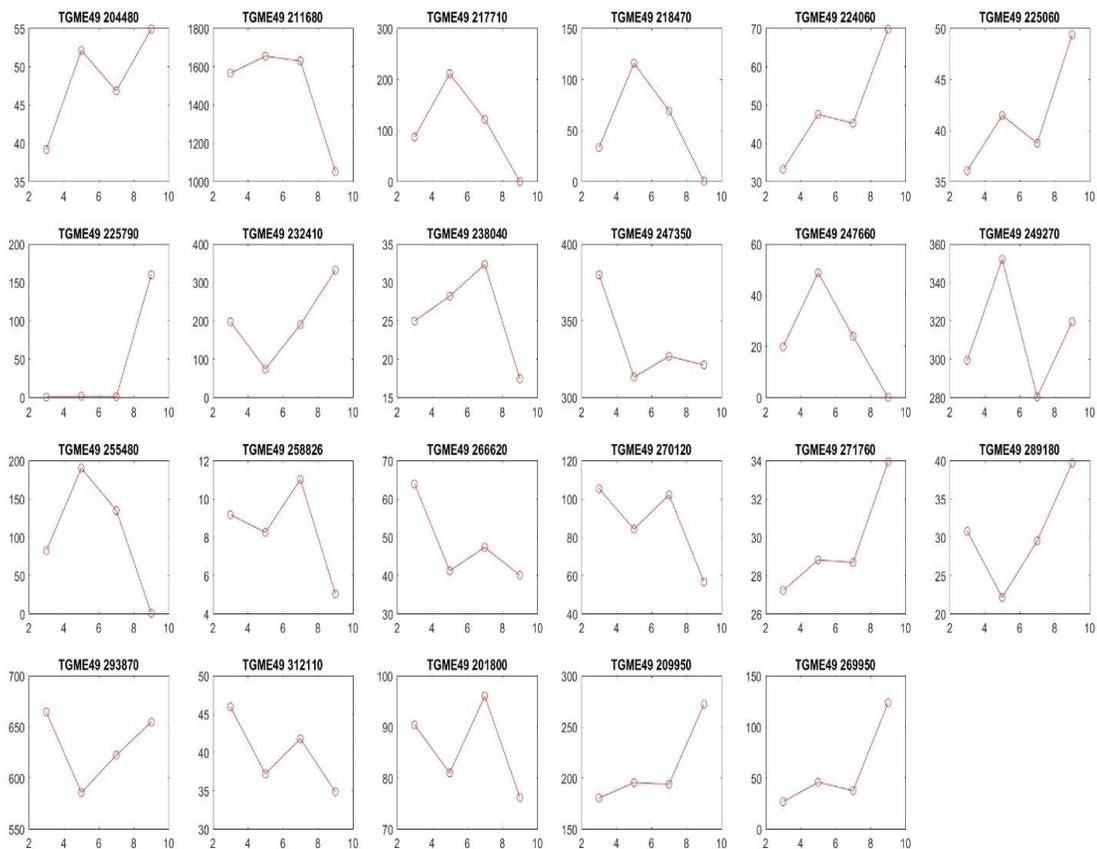


Figura 9. Series de tiempo (días 3, 5 y 7) de los perfiles de expresión de transcritos de la familia de proteínas disulfuro isomerasa de *T. gondii* post infección en gato. Luego de 3 días post infección en gato se encuentra merozoito, luego de 5 días post infección en gato se presentan los estadios sexuales (etapa de infección a células enteroepiteliales) y luego de 7 días post infección del gato se encuentra el ooquiste. El nivel de expresión en el taquizoito se muestra en el punto 9.

Perfil de expresión durante las fases de replicación aguda y crónica

El análisis de componentes principales encuentra un 100% de la variabilidad de los datos para el componente 1. Esto indica que la variabilidad en la expresión de estas proteínas de esta familia está relacionada con cada fase (Figura 10). Se encontraron once proteínas de fase aguda, ocho de fase crónica y cuatro sin relación con alguna fase (Figuras 10, 11 y 12).

Los resultados conjuntos evidenciaron que en todos los perfiles de transcriptomas la proteína TGME49_211680 presenta mayor expresión multiestadio seguida de las proteínas 232410 y 293870.

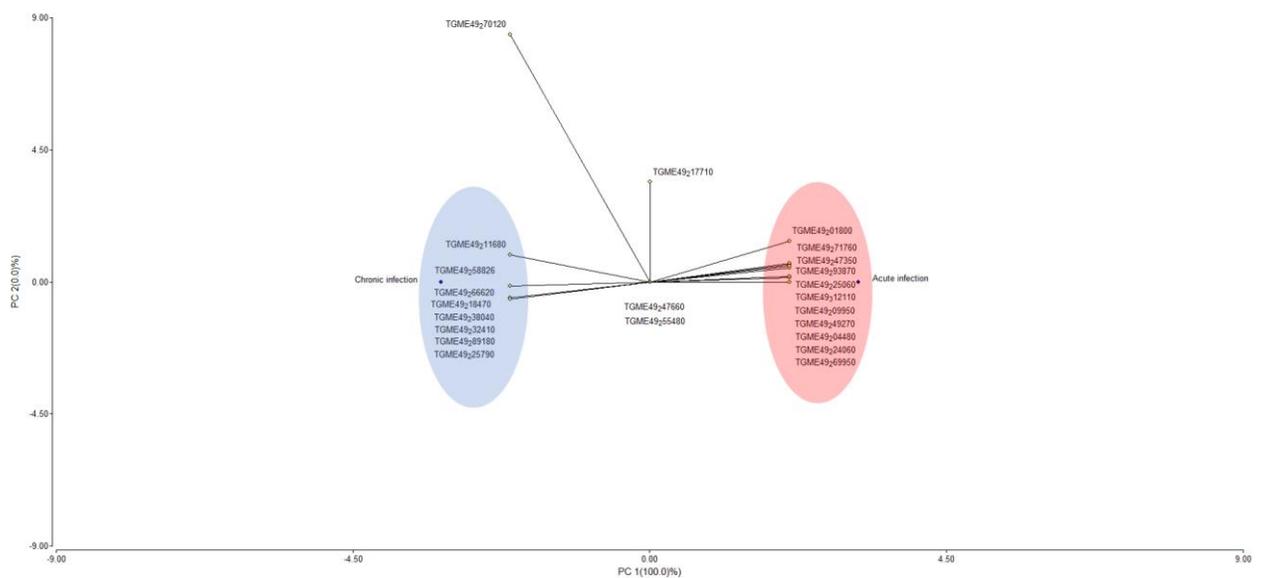


Figura 10. Componentes principales de los perfiles de expresión de la familia de proteínas disulfuro isomerasa de *T. gondii* durante fase aguda y fase crónica de infección en ratón. El perfil de expresión de la fase aguda se obtuvo a los 10 días post infección y el perfil de fase crónica se

obtuvo luego de 28 días post infección. Dentro del círculo azul se muestran las secuencias de fase aguda y en círculo rosado las de fase crónica

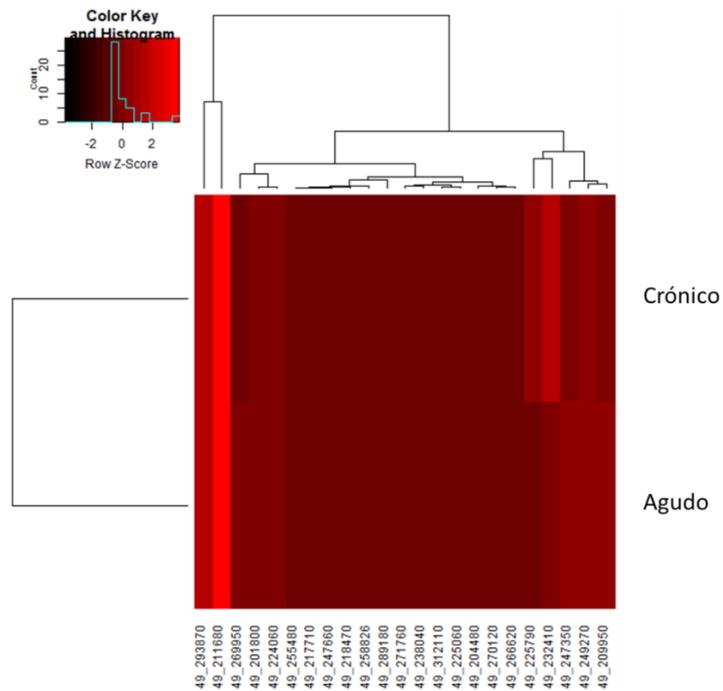


Figura 11. Mapa de calor de los perfiles de expresión en fases aguda y crónica de infección en ratón para la familia de genes de proteínas disulfuro isomerasa de *T. gondii*. El perfil de expresión de la fase aguda se obtuvo a los 10 días post infección y el perfil de fase crónica se obtuvo luego de 28 días post infección.

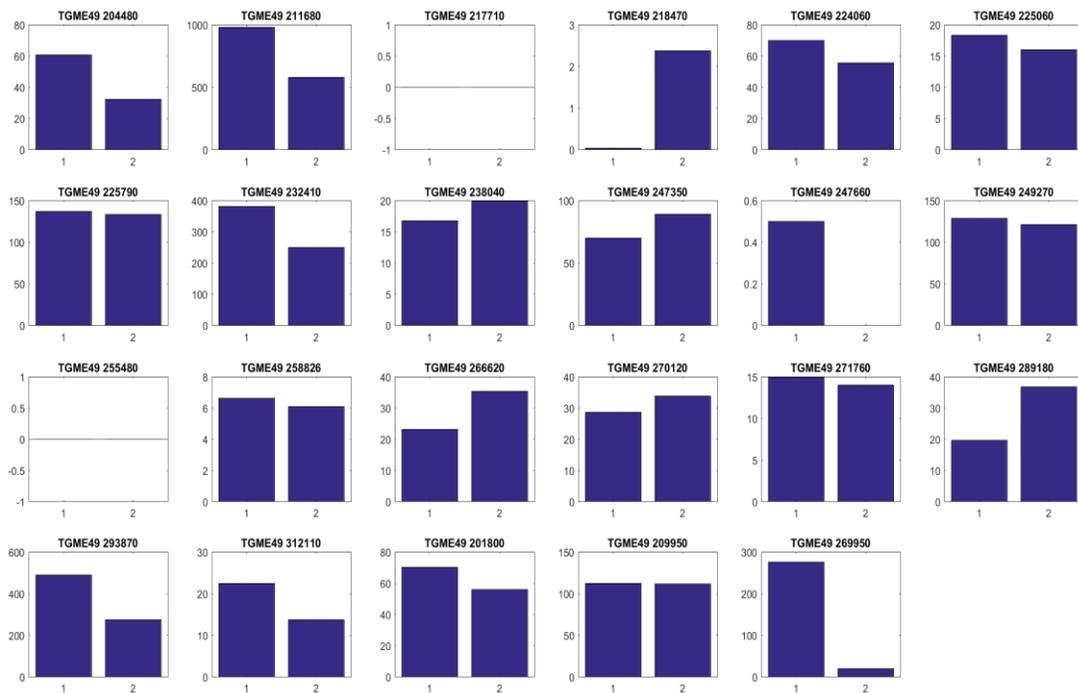


Figura 12. Histograma de los niveles de expresión de la familia de proteínas disulfuro isomerasa de *T. gondii*. Se muestra la columna 1 de estadio crónico y la 2 para el estadio agudo para cada secuencia luego de infección en ratón.

Importancia para el fitness parasitario de la familia de proteínas PDI

Luego de analizar el efecto sobre el fitness parasitario en el modelo de infección a células humanas HFF, se identificaron 15 proteínas con puntajes entre -0.53 hasta -5.65 y 8 proteínas desde -0.01 hasta 1.98 (Figura 13). De particular importancia fue que 5 proteínas tuvieron puntaje menor a -4, indicando que son proteínas esenciales para la supervivencia del parásito, dentro de ellas la secuencia para la proteína TGME49_211680 fue la que impactó en mayor medida la sobrevivencia del taquizoito de la cepa GT1 de *T. gondii*, seguido por las proteínas TGME49_209950, TGME49_249270 y

TGME49_293870 (Figura 13). Esto hace a la proteína TGME49_211680 un candidato atractivo para la evaluación de compuestos inhibidores. Esta secuencia es aún de mayor interés ya que posee altos niveles de expresión multiestadio y tiene una predicción de localización subcelular múltiple (Figura 14). Aunado a ello, un análisis de ontogenia para esta clase de proteínas encuentra que tiene una gran importancia por estar involucrada en múltiples procesos biológicos y moleculares de la fisiología parasitaria (Figura 15).

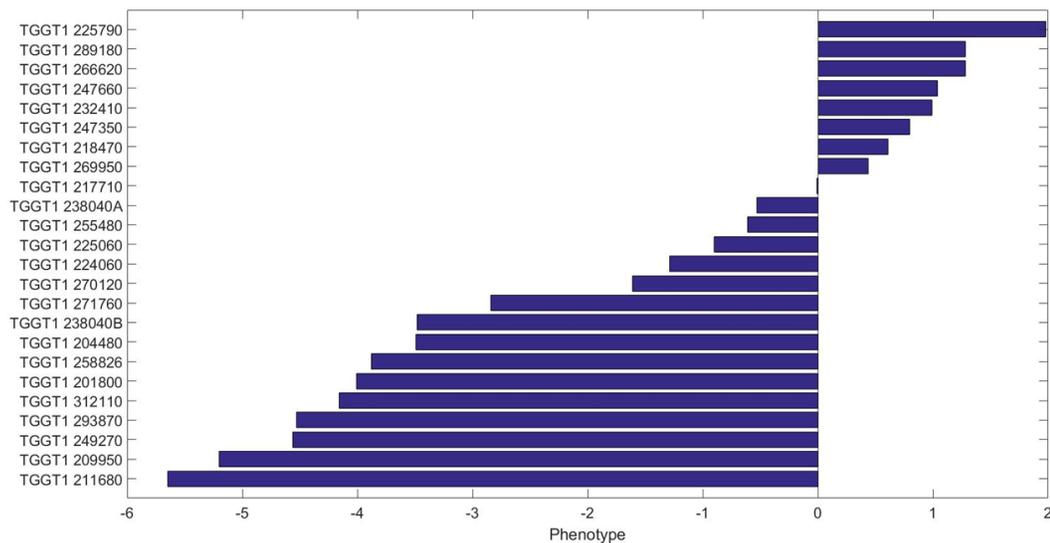


Figura 13. Fenotipo de *fitness* parasitario de los *knockout* producidos para la familia de proteínas disulfuro isomerasa de *T. gondii*. Un puntaje de fenotipo (*Phenotype*) <1 indica reducción en la capacidad de supervivencia (ajuste al medio ambiente celular o *fitness*). La barra vertical indica la secuencia de PDI afectada por el proceso de invalidación del gen (*knockout*) por el método de CRIS/Cas9.

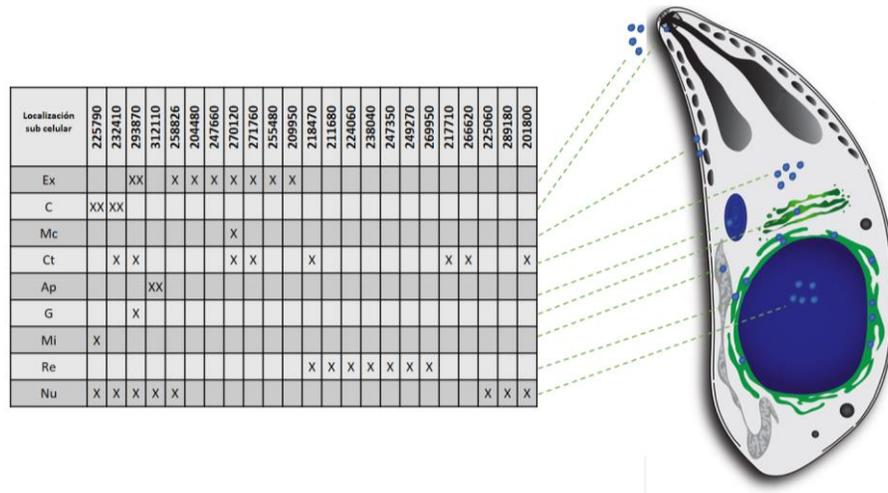


Figura 14. Evidencia *in silico* y experimental de la localización subcelular de la familia de proteínas PDI de *T. gondii*. Tabla de localizaciones subcelulares, marcadas con X para predicciones *in silico* y XX para evidencia experimental en taquizoito (Kremer *et al.*, 2013). Se muestra cada proteína como un punto azul en su respectiva localización subcelular. Nu: núcleo, ER: retículo endoplasmático, G: Aparato de Golgi, Ct: citoplasma, Mi: mitocondria, Ap: apicoplasto, C: conoide, Cm: Membrana celular, Ex: espacio extracelular.

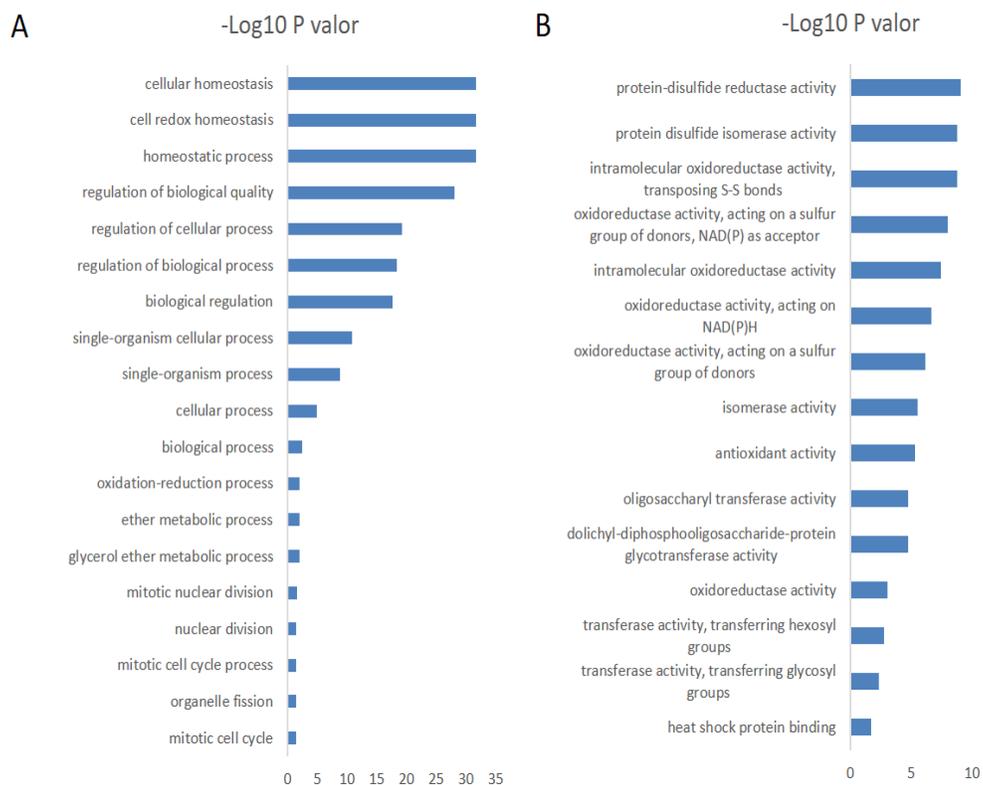


Figura 15. Ontología de gen. Muestra el Log P valor de los términos de ontología de gen asociados a la familia de proteínas PDI de *T gondii*. A- Proceso biológico, B- actividad molecular.

Ensayos de invasión e interacción proteína células hospedadoras

Se obtuvo una proteína recombinante con un peso molecular de 52,8 kDa (Figura 16), con la cual se realizaron experimentos de invasión a células HeLa, ensayos de interacción con células HeLa y Müller, donde se observó un porcentaje de reducción máximo de un 19% de la invasión con la proteína y con bacitracina Figura 17, Además, se pudo observar que la expresión de transcritos para esta proteína es mayor antes de la invasión a células HeLa figura 18, lo que sugiere que podría estar involucrada en procesos de

invasión. Además los ensayos de inmunofluorecencia sugieren que esta proteína recombinante puede interactuar con células HeLA y Muller (Figura 19). (Dickerhof *et al.*, 2011; Ramos *et al.*, 2011; Moncada, Arenas, Acosta, Molina, Hernandez, *et al.*, 2016)

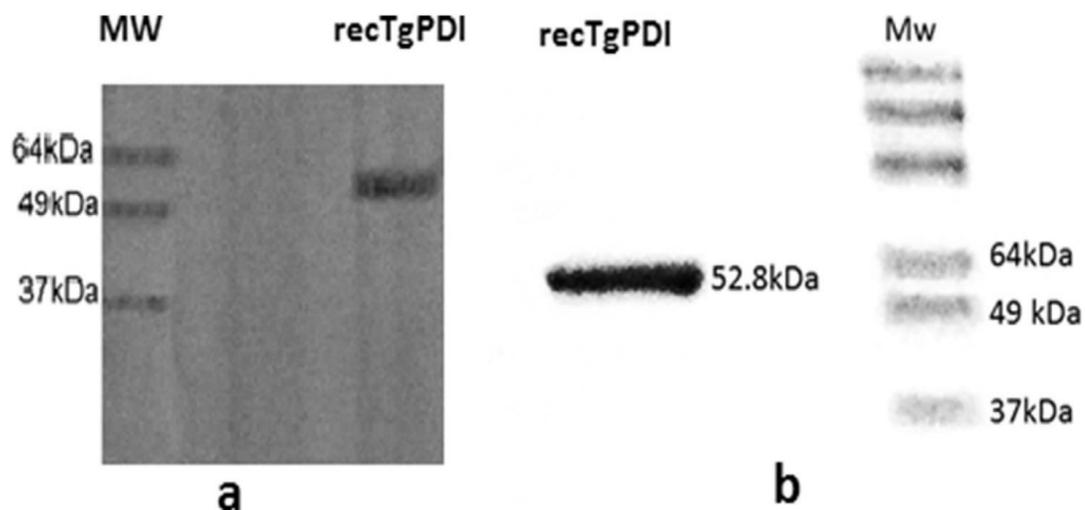


Figura 16 Electroforesis de poliacrilamida al 12% y western blot. a- izquierda electroforesis de poliacrilamida pozo 1 marcador de peso molecular. b- western blot de la proteína purificada recTgPDI (TGME49_211680)

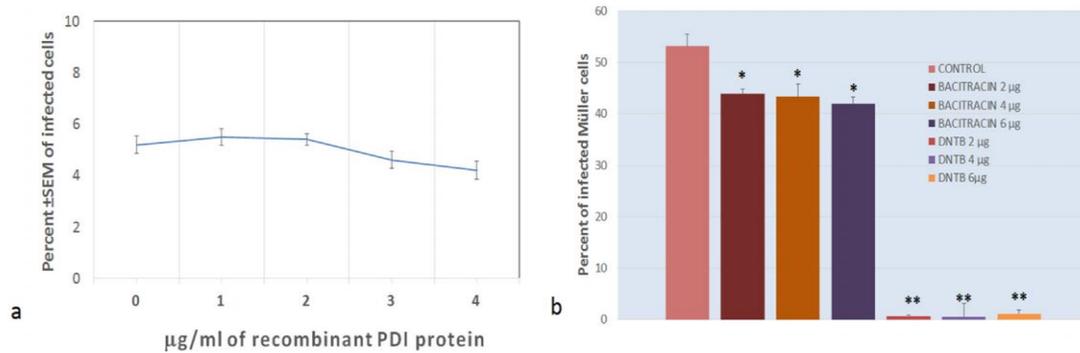


Figura 17 porcentaje de células infectadas durante la infección *in vitro* de células HeLa por Taquizoitos de *T gondii* en una multiplicidad de infección de 5:1. a- Se muestra el porcentaje de células infectadas por concentración de la proteína recombinante exógena. b-Porcentaje de invasión de células Müller con dos inhibidores bacitracina y DNTB en concentraciones de 2,4,6 µg.

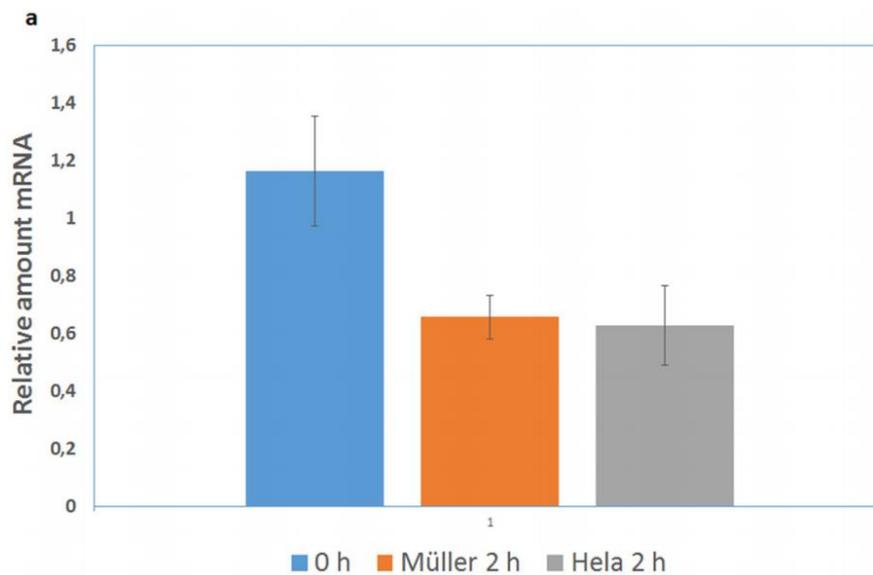


Figura 18 Gráfico de barras de la cuantificación relativa de los transcritos de la proteína TGME49_ 211680. Tiempo 0h (antes de la invasión) y Tiempo 2h (2 horas post infección).

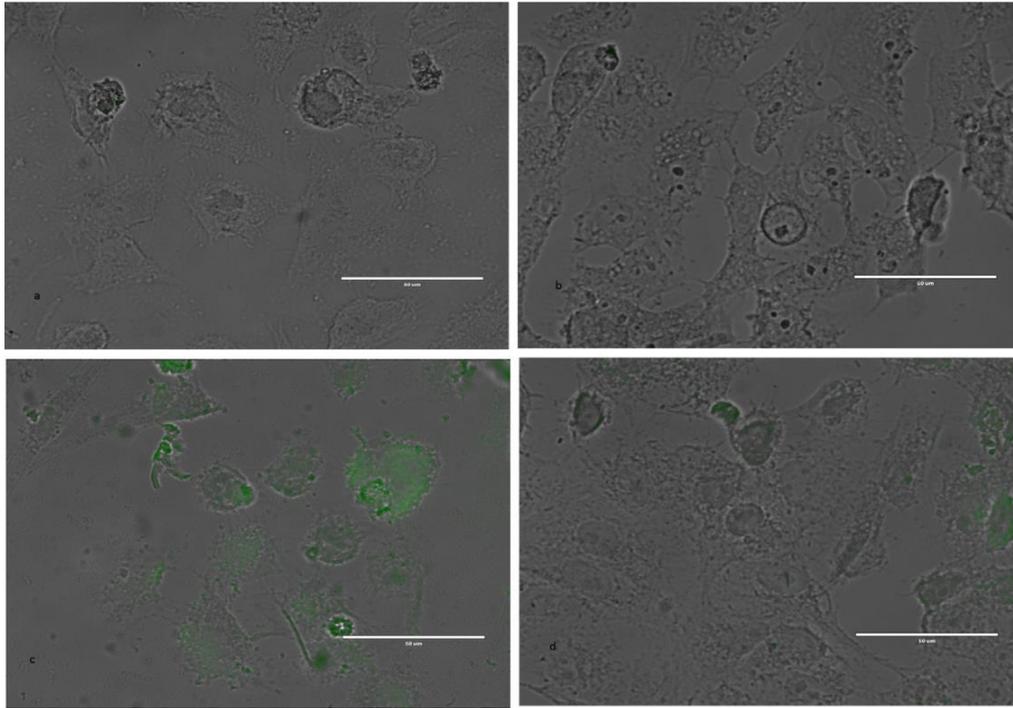


Figura 19. Inmunofluorescencia de la proteína recombinante TGME49_211680 en células HeLa y Müller. a y b: control de células sin proteína exógena. c y d: células Hela y Müller, respectivamente, después de ser adicionadas con de 4 µg/ml de proteína recombinante. En verde se muestra la fluorescencia asociada a la proteína recombinante.

Análisis estructural de la proteína PDI TGME49_211680

Regiones funcionales

TGME49_211680 es una proteína de 471 aminoácidos, cuya estructura secundaria se compone de 18 α hélices y 13 láminas β , ordenadas tridimensionalmente en 4 dominios. Estos dominios, denominados a, a', b, b',

se ubican espacialmente en forma de “U retorcida” y conservan un plegamiento tiorredoxina en el cual cuatro α hélices rodean a tres o cuatro láminas β de la forma $\alpha_1 \beta_1 \alpha_2 \beta_2 \alpha_3 \beta_3 \alpha_4$ o $\beta_1 \alpha_1 \beta_2 \alpha_2 \beta_3 \alpha_3 \beta_4 \alpha_4$, respectivamente. Posee un péptido señal de 24 aminoácidos y estructuralmente se encuentra ordenada del modo a-b-b'-x-a'-c, donde “x” es un enlazador de 21 residuos, situado entre los aminoácidos 331 y 352, y “c” es una extensión terminal de 28 aminoácidos que incluye una región de retención al retículo endoplasmático, conformada por los aminoácidos GEEL figura 20. Así mismo, los dominios a y a' representan dominios catalíticamente activos que contienen el motivo CGHC y CGYC, y se ubican entre los aminoácidos 29- 131 y 352- 459, respectivamente. Por el contrario, los dominios b y b' encontrados entre los aminoácidos 136- 223 y 232- 331, respectivamente, son dominios catalíticamente inactivos, al carecer del motivo catalítico CxxC.

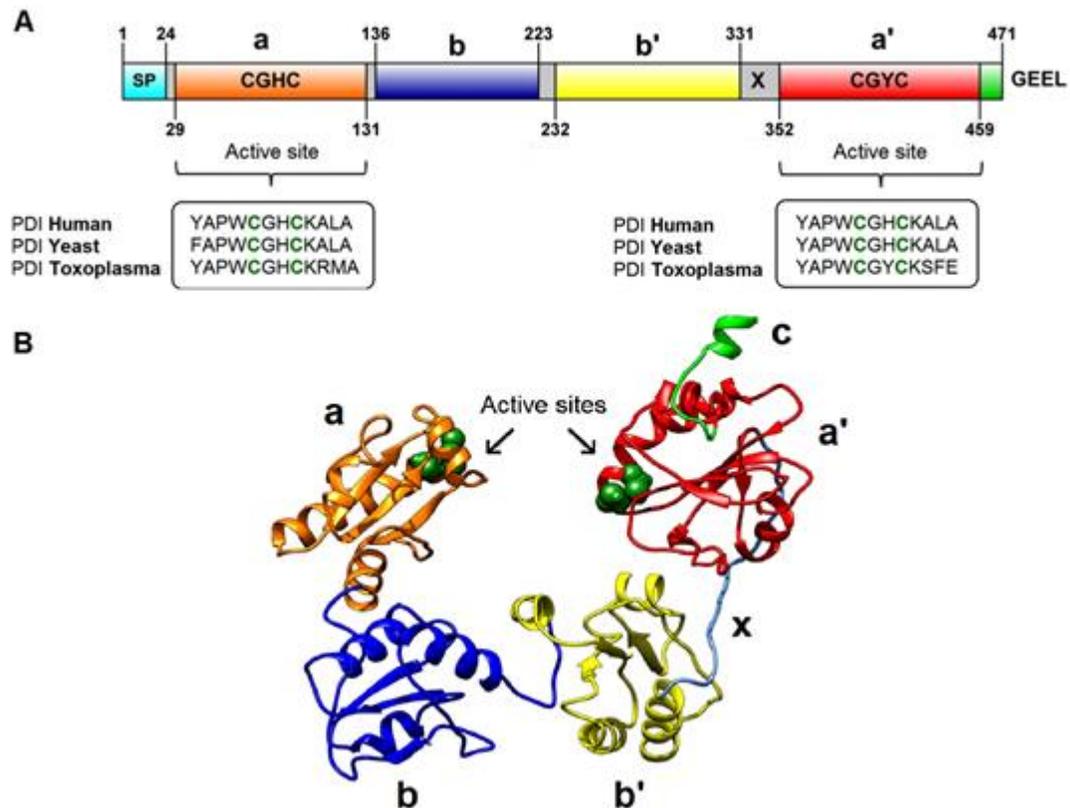


Figura 20. Estructura general de PDI TGME49_211680. A. Esquema de las regiones funcionales de PDI TGME49_211680. En color azul claro se resalta la región del péptido señal; en naranja y rojo los dominios catalíticamente activos; los colores azul y amarillo representan los dominios catalíticamente inactivos y en verde se muestra una extensión C-terminal, la cual posee una región de retención al retículo endoplasmático GEEL. **B. Estructura tridimensional de PDI TGME49_211680.** En esferas de color verde se resaltan los residuos de cisteína responsables de la actividad redox de la proteína y los colores que diferencian cada dominio corresponden con los colores descritos en A.

Dominios activos

A través de alineamientos de secuencias entre los dominios activos de la familia de proteínas disulfuro isomerasa humanas y PDI TGME49_211680,

se identificaron en los dominios a y a' de PDI TGME49_211680, aminoácidos altamente conservados, cuyas funciones ya han sido descritas para diferentes proteínas humanas (Figura 21).

Además, se encontraron residuos aromáticos conservados correspondientes a dos fenilalaninas en el dominio a, ubicadas en la parte N-terminal de la α -hélice 1 y al final de la lámina β - 1, en las posiciones 13 y 27, respectivamente. Estos residuos aromáticos en el dominio a' corresponden a una fenilalanina encontrada en la parte N-terminal de la α -hélice 1 y una isoleucina al final de la lámina β - 2, en las posiciones 336 y 351, respectivamente. También, en los dominios a y a' se evidencia un triptófano ubicado sobre la α -hélice 2, inmediatamente antes del motivo catalítico, en las posiciones 31 y 355, respectivamente, así como una alanina encontrada en la posición 29 entre la lámina β - 1 y α -hélice 2 en el dominio a, y en la posición 353 entre la lámina β - 2 y α -hélice 2 en el dominio a', y un ácido aspártico localizado en la α -hélice 3 de ambos dominios, en las posiciones 62 y 385, respectivamente.

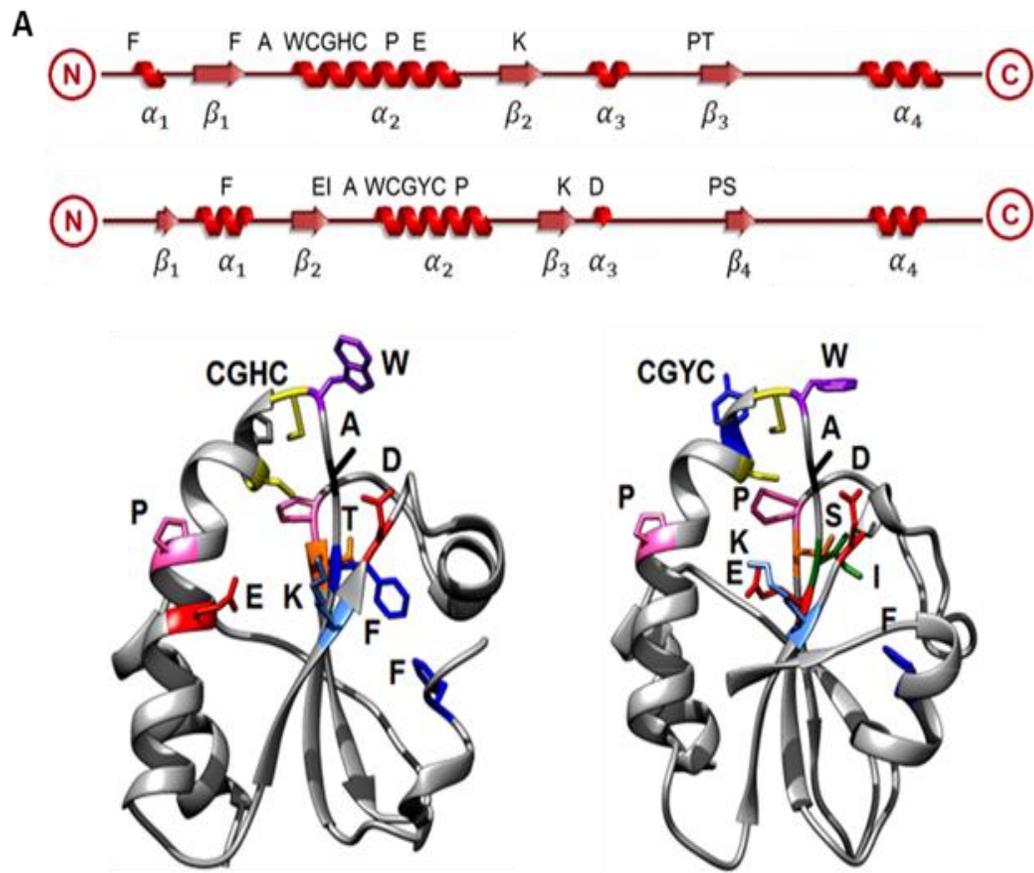


Figura 21. Aminoácidos conservados en los dominios activos de PDI TGME49_211680.
A. Estructura secundaria del dominio a de PDI TGME49_211680. B. Estructura secundaria del dominio a' de PDI TGME49_211680. C. Estructura terciaria del dominio a de PDI TGME49_211680. D. Estructura terciaria del dominio a' de PDI TGME49_211680. Los colores empleados para destacar los aminoácidos conservados corresponden al código de colores utilizado por el servidor I-TASSER para agrupar los aminoácidos según sus características fisicoquímicas (<http://zhanglab.cmb.med.umich.edu/I-TASSER/color.html>).

Los residuos que conforman el sitio catalítico y son esenciales para la actividad de la proteína, también se encuentran altamente conservados entre la familia de proteínas PDI. En PDI TGME49_211680 estos aminoácidos corresponden inicialmente a dos residuos de cisteína ubicados sobre la α -

hélice 2 de los dominios a y a', en las posiciones 32- 35 y 356- 359, respectivamente. Entre estas cisteínas se localiza un residuo de glicina en la posición 33 y una histidina en la posición 34 en el dominio a, y en el dominio a', aun cuando se conserva el residuo de glicina en la posición 357, en la posición 358 hay una variación de la histidina altamente conservada, por un residuo de tirosina. Así mismo, en los dominios a y a' permanecen conservadas dos prolinas. La primera, se halla en la α - hélice 2 en las posiciones 40 y 364, respectivamente, y la segunda, encontrada en posición *cis*, se ubica al lado opuesto del motivo catalítico en las posiciones 79 y 402, respectivamente. En esta región opuesta al motivo catalítico, también se evidencia una treonina contigua a la *cis*- prolina. Mientras en el dominio a de PDI TGME49_211680 este residuo es conservado y se localiza en la posición 80 sobre la lámina β - 3, en el dominio a' se encuentra una variación de la treonina por la serina 403.

Finalmente, se localizó una región cargada en el interior de los dominios, compuesta por un ácido glutámico y una lisina, ubicados en las posiciones 43 y 60 sobre la α -hélice 2 y lámina β -2 del dominio a, respectivamente, y en las posiciones 350 y 383 sobre la lámina β -2 y lámina β -3 del dominio a', respectivamente. A su vez, se encontró un residuo de arginina altamente conservado en la posición 96 para el dominio a y 421 para el dominio a'.

Dominios inactivos

Por medio de alineamientos de secuencias y estructurales entre los dominios catalíticamente inactivos de PDI TGME49_211680 y proteínas PDIs humanas, se identificó en la proteína del parásito un bolsillo hidrofóbico, ubicado entre la primera y la tercera α -hélice del dominio b' (Figuras 22,23 y S5), el cual, aunque es característico de esta familia de proteínas, posee

aminoácidos poco conservados que dan lugar a un perfil de hidrofobicidad diferente en esta región, respecto a las proteínas humanas .

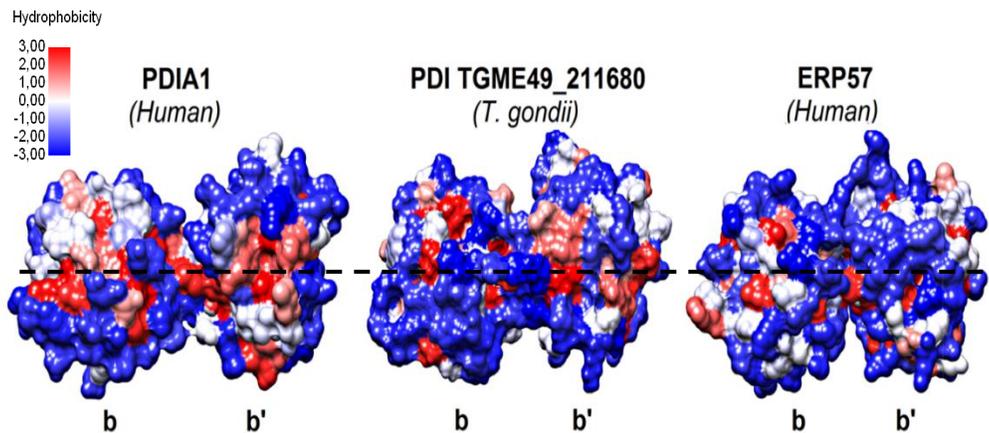


Figura 22. Comparación del perfil de hidrofobicidad de los dominios b y b' de PDI TGME49_211680 y las proteínas humanas PDI A1 y ERP 57. En azul se representan las regiones hidrófilas, en blanco las regiones neutras y en rojo las regiones hidrofóbicas.

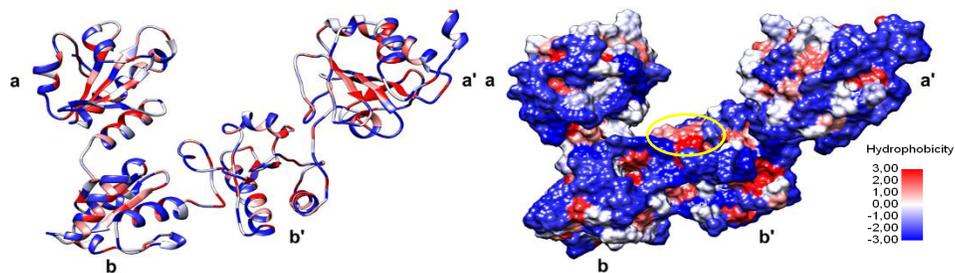


Figura 23 Perfil de hidrofobicidad y potencial electrostático de PDI TGME49_211680. En azul se representan las regiones hidrófilas, en blanco las regiones neutras, en rojo las regiones hidrofóbicas y en un óvalo amarillo se resalta la región altamente

Los aminoácidos que conforman el bolsillo hidrofóbico de PDI TGME49_211680 son alanina 208, fenilalanina 211 e isoleucina 215, que anteceden a la primera lámina β del dominio b' y se diferencian de proteínas humanas como PDI A1, debido a que esta última, conserva en estas posiciones espaciales relativos aminoácidos hidrófilos como glutamato 242 y glutamina 243. Así mismo, sobre la lámina β - 1 de PDI TGME49_211680 se encuentran los residuos leucina 219 y 221, y el triptófano 223, cuyos análogos en PDI A1 corresponden a residuos hidrófilos como histidina 256, de hidrofobicidad media como fenilalanina 260 e hidrofobicidad alta como leucina 258.

Por otra parte, precediendo a la α -hélice 2 del dominio b' de PDI TGME49_211680, se localizan la valina 252 y leucina 254, y sobre la hélice se ubican además la fenilalanina 259, alanina 263, alanina 266 y leucina 267. Aunque estos aminoácidos no son conservados en proteínas humanas, en PDI A1 se conserva en su mayoría el carácter hidrofóbico general de esta hélice. Finalmente la región hidrofóbica del dominio b' de PDI TGME49_211680 se compone de la valina 276 posicionada sobre la lámina β - 2 y la fenilalanina 284 y leucina 286, consecutivas a esta lámina β figura S5. Estos residuos representan una zona de gran variabilidad en potencial electrostático respecto a PDI A1, ya que esta contiene aminoácidos hidrófilos como arginina 316, neutros como tirosina 327 y de hidrofobicidad media como metionina 324.

Búsqueda in silico de alto rendimiento de nuevos candidatos de inhibición de la proteína TGME49_211680

Por medio de una búsqueda computacional de alto rendimiento, se encontraron seis moléculas con potencial inhibitorio sobre regiones diferentes respecto a proteínas humanas y consideradas importantes para la actividad de PDI TGME49_211680, como son el sitio activo del dominio a', el principal sitio de acoplamiento del sustrato en el dominio b' y la región comprendida entre el dominio b' y el enlazador x figura 24.

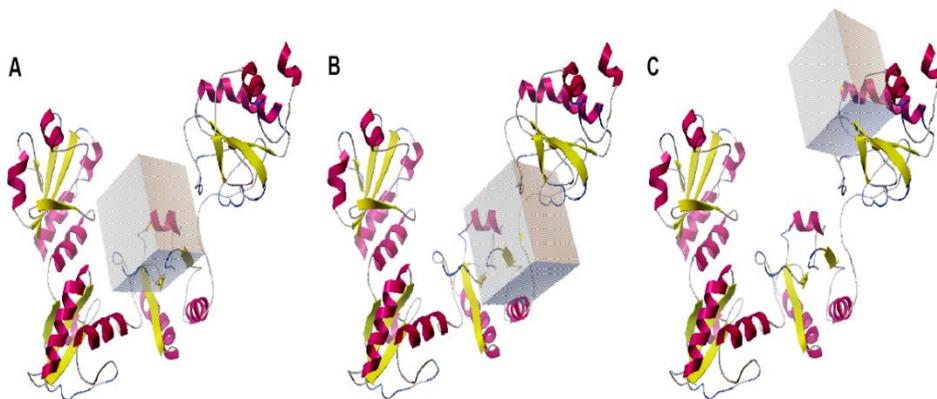


Figura 24. Representación gráfica de las regiones objetivo de búsqueda de inhibidores para PDI TGME49_211680. A. Espacio de búsqueda centrado el dominio b' correspondiente a las coordenadas X: 84.773 Y: 82.87 Z: 83.694. B. Región comprendida entre el dominio b' y el enlazador x, con centro en las coordenadas X: 94.825 Y: 87.684 Z: 81.872. C. Espacio de búsqueda centrado en el motivo catalítico **CGYC con coordenadas X: 106.037 Y: 60.725 Z: 93.266.**

En el centro activo las moléculas ZINC 08610024 y 09642074, no se presentaron interacciones con aminoácidos del motivo CxxC de las proteínas

humanas, y su energía de unión calculada fue menor en la proteína del parásito, con una diferencia de al menos 2 kcal/mol figura 25 Tabla 4.

Las interacciones ligando- proteína establecidas entre la molécula ZINC 08610024 y la proteína TGME49_211680 corresponden a interacciones de tipo hidrofóbico pi- pi, entre el anillo furano del ligando y la tirosina 358, y el grupo fenil del ligando y la fenilalanina 362. De igual forma, se generan contactos electrostáticos atractivos entre el ácido glutámico 418 y el azufre del grupo sulfonilo del ligando, y enlaces de hidrógeno con los aminoácidos glicina 357, fenilalanina 401, arginina 421 y prolina 402; con distancias de enlaces de 3,14, 3,33, 3,05 y 4,54 Å, respectivamente. Finalmente se evidencian interacciones de van der Waals con los aminoácidos cisteína 356 y serina 400 (Figura 26A).

De igual modo, las interacciones generadas entre el ligando ZINC 09642074 y la proteína PDI TGME49_211680 corresponden a interacciones de tipo hidrofóbico pi- pi, entre el anillo fenílico del ligando y los aminoácidos tirosina 358 y fenilalanina 362. Adicionalmente, se establecen contactos electrostáticos atractivos entre el ácido glutámico 418 y el azufre del grupo sulfonilo del ligando, y enlaces de hidrógeno con los aminoácidos fenilalanina 401, arginina 421 y prolina 402; con distancias de enlaces de 3,32, 3,13 y 4,47 Å, respectivamente. Así mismo, se evidencian interacciones de van der Waals con los aminoácidos triptófano 355, cisteína 356, glicina 357 y serina 400 (Figura 26B).

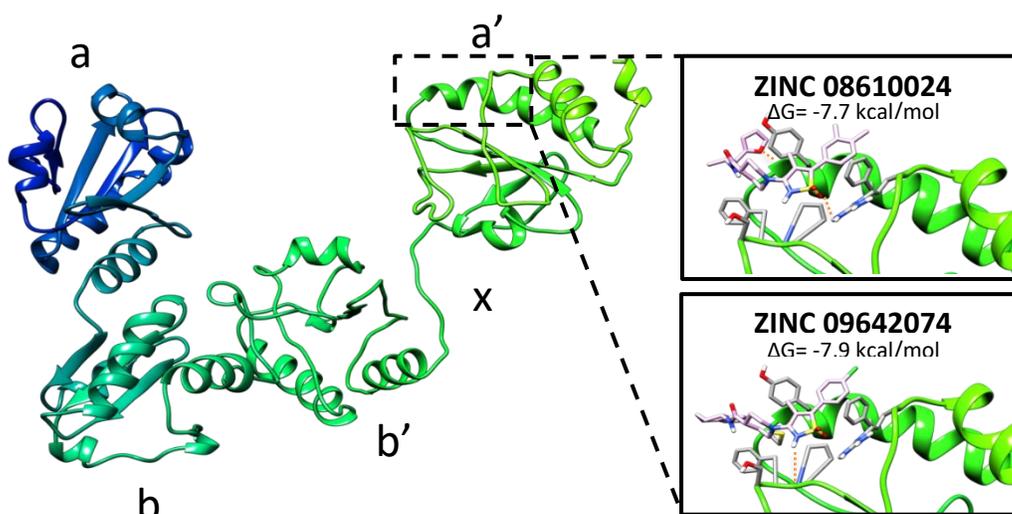


Figura 25 Acoplamiento molecular de ligandos con potencial inhibitorio sobre el sitio activo CGYC de PDI TGME49_211680. En color rojo se resaltan los átomos de oxígeno, en azul los átomos de nitrógeno, blanco átomos de hidrógeno, gris átomos de carbono y amarillo átomos de azufre

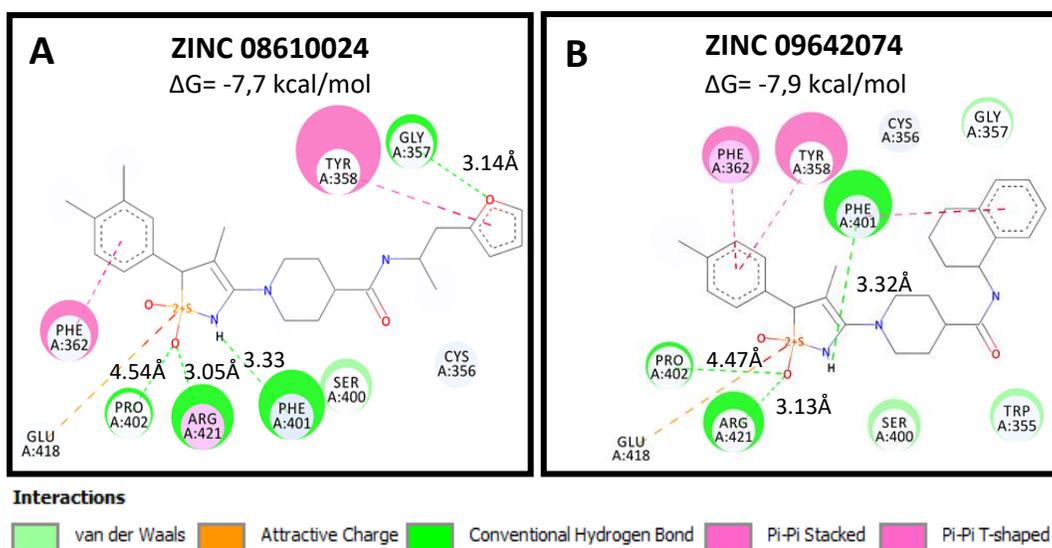


Figura 26 Diagrama de interacciones intermoleculares entre PDI TGME49_211680 y ligandos con potencial inhibitorio sobre centro activo CGYC. **A.** Interacciones de enlace entre PDI TGME49_211680 y ligando ZINC 08610024. **B.** Interacciones de enlace entre PDI TGME49_211680 y ligando ZINC 09642074. En verde eléctrico se destacan los enlaces de hidrógeno, en rosa las interacciones de tipo hidrofóbico, en verde pastel las interacciones de Van der Waals

Tabla 4. Energías de acoplamiento e interacciones establecidas por moléculas con potencial inhibitorio en el motivo catalítico CGYC de PDI TGME49_211680 y su comparación con las

Código ZINC	PDI TGME49_2116880		PDIA1				ERP 57			
	Dominio a'		Dominio a		Dominio a'		Dominio a		Dominio a'	
	ΔG (kcal/mol)	Interacción CGYC	ΔG (kcal/mol)	Interacción CGHC	ΔG (kcal/mol)	Interacción CGHC	ΔG (kcal/mol)	Interacción CGHA	ΔG (kcal/mol)	Interacción CGHC
08610024	-7,7	C356 G357 Y358	-7,5	Ninguna	-6	Ninguna	-5,4	Ninguna	-7	Ninguna
09642074	-7,9	C356 G357 Y358	-7,3	Ninguna	-5,8	H399	-5,1	H59	-6,8	Ninguna

Dominio b' y región comprendida entre dominio b' y enlazador x

Las búsquedas de moléculas inhibitoras sobre el dominio b' de PDI TGME49_211680, condujeron a identificar que los ligandos de códigos ZINC 12210072, 15784130 y 33010617, se acoplan en una de las zonas hidrofóbicas de este dominio, con energías de unión de -9, -9,1 y -9,7 kcal/mol, respectivamente figura 27. Debido a que las regiones hidrofóbicas del dominio b' pueden ser esenciales para el acoplamiento del sustrato a la proteína, los ligandos encontrados podrían ser potenciales competidores por este sitio de unión, gracias al establecimiento de diferentes tipos de interacciones, que pueden estabilizar energéticamente los complejos proteína- ligando.

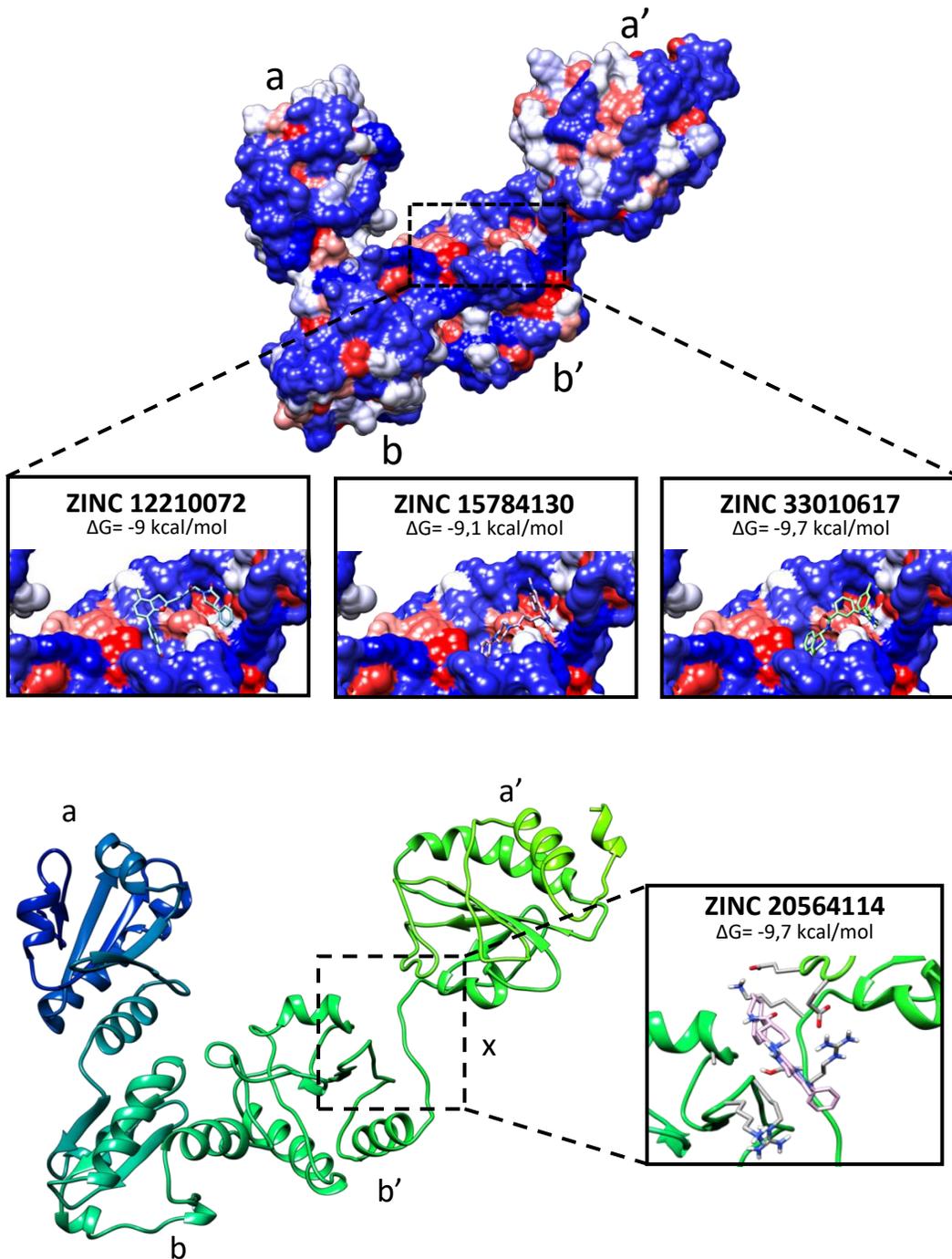


Figura 28 Acoplamiento molecular de ligando con potencial inhibitorio sobre la región comprendida entre el dominio *b'* y el enlazador *x* de la proteína PDI TGME49_211680. En color rojo se resaltan los átomos de oxígeno, en azul los átomos de nitrógeno, blanco átomos de hidrógeno y en

En particular, el ligando 12210072 presenta interacciones hidrofóbicas de tipo pi- pi con la histidina 262, enlaces de hidrógeno energéticamente significativos con los aminoácidos glutamina 278, arginina 217 y asparagina 265, y diferentes contactos de van der Waals, que favorecen su interacción con la proteína PDI TGME49_211680 figura 29A. Así mismo, los ligandos ZINC 33010617 y 15784130 establecen enlaces de hidrógeno de fuerza moderada y débil con la serina 314, interacciones hidrofóbicas tipo pi- pi con la fenilalanina 284, interacciones alquílicas con lisina 316 e isoleucina 215 y contactos de van der Waals con diferentes aminoácidos de este bolsillo hidrofóbico (Figura 29 B-C).

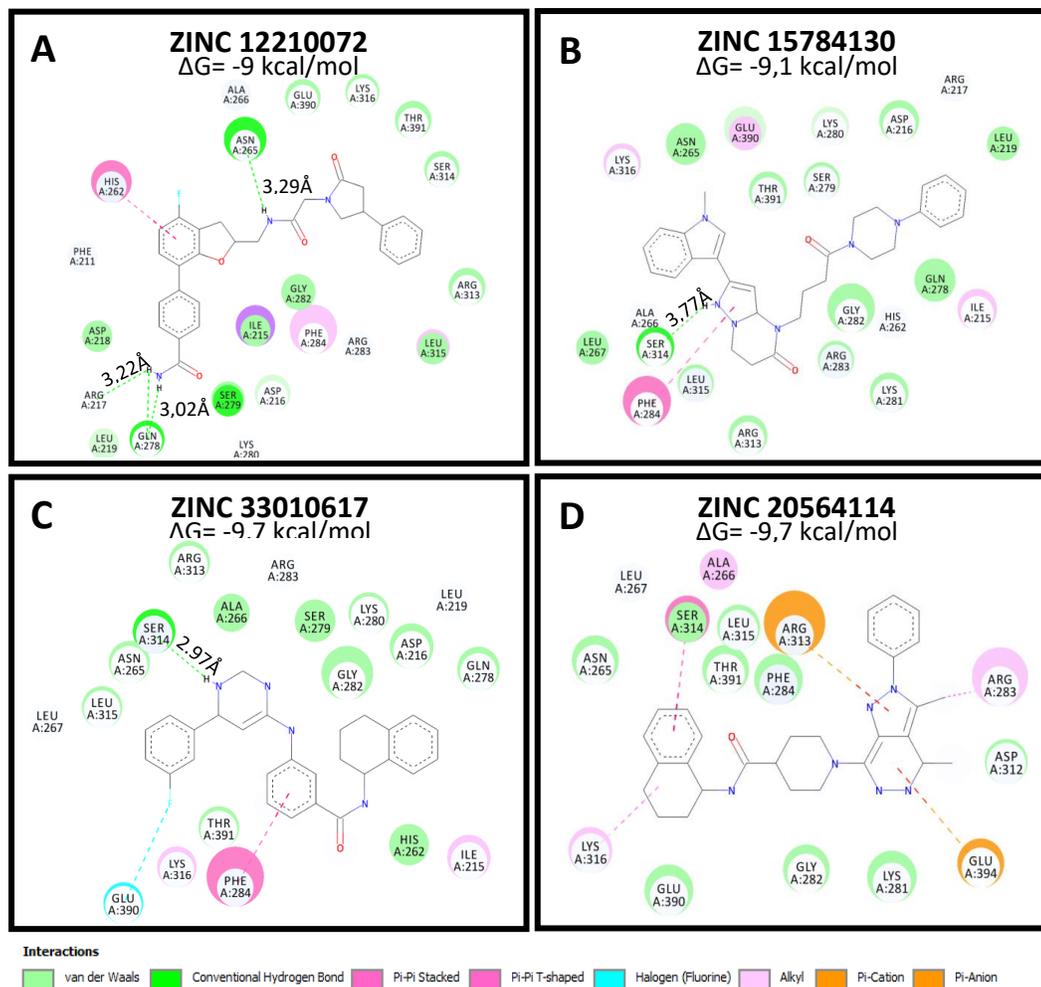


Figura 29 Diagrama de interacciones intermoleculares entre PDI TGME49_211680 y ligandos con potencial inhibitorio sobre bolsillo hidrofóbico del dominio b' y región comprendida entre el dominio b' y enlazador x de PDI TGME49_211680. **A.** Interacciones de enlace entre dominio b' de PDI TGME49_211680 y ligando ZINC 12210072. **B.** Interacciones de enlace entre dominio b' de PDI TGME49_211680 y ligando ZINC 15784130. **C.** Interacciones de enlace entre dominio b' de PDI TGME49_211680 y ligando ZINC 33010617. **D.** Interacciones de enlace entre región comprendida entre el dominio b' y enlazador x de PDI TGME49_211680 y ligando ZINC 20564114. En verde eléctrico se destacan los enlaces de hidrógeno, en rosa las interacciones de tipo hidrofóbico, en verde pastel las interacciones de van der Waals, en azul claro los enlaces de halógeno y en líneas punteadas color naranja, las interacciones electroestáticas.

Tabla 5. Energías de acoplamiento de moléculas con potencial inhibitorio sobre sitio de interacción del sustrato y región comprendida entre el dominio b' y el enlazador x de las proteínas disulfuro isomerasa PDI TGME49_211680, PDIA1 y ERP 57.

Sitio de acoplamiento	Código ZINC	PDI	PDIA1	ERP 57
		TGME49_211680		
		ΔG (kcal/mol)	ΔG (kcal/mol)	ΔG (kcal/mol)
Dominio b'	12210072	-9	-6,8	-6
Dominio b'	15784130	-9,1	-7	-6,4
Dominio b'	33010617	-9,7	-7,2	-7,4
Dominio b'- enlazador x	20564114	-9,7	-7,4	-6,7

En la Tabla 6 se muestran las moléculas candidatas para ser utilizadas como inhibidoras de PDI TGME49_211680.

Tabla 6. Características de venta de moléculas identificadas como potenciales inhibidores de la proteína PDI TGME49_211680 de *Toxoplasma gondii*.

Código ZINC	Actividad biológica	Toxicidad	Vendedores	%Pureza	Precio USD (1 mg)
08610024	No especificada	No	ChemDiv	90	76
09642074	No especificada	No	ChemDiv	90	76
12210072	No especificada	No	ChemBridge	85	51
15784130	No especificada	No	ChemDiv	90	76
33010617	No especificada	No	ChemDiv	90	76
20564114	No especificada	No	ChemDiv	90	76

Se encontró que las moléculas identificadas con los códigos ZINC 08610024 y 09642074, se acoplan en el centro activo del dominio a' de PDI TGME49_211680. En particular, el ligando ZINC 08610024 interactúa con residuos involucrados directamente en las reacciones de intercambio tiol-

disulfuro como la cisteína 356, o con posibles moduladores de los potenciales de oxidación- reducción del sitio activo como la glicina 357, donde se establece un enlace de hidrógeno de fuerza moderada, con distancia de enlace de 3,14 Å (Emsley and Bonds, 1968) y el aminoácido diferente en el sitio activo, tirosina 358, con el que se ejercen interacciones hidrofóbicas tipo pi- pi (Figura 26A).

Discusión

En este trabajo pudimos identificar la existencia de secuencias en el genoma de *T. gondii* que codifican para 23 proteínas con la presencia de al menos un dominio tiorredoxina, las cuales presentan el dominio característico CxxC o una secuencia homóloga. Estas secuencias son de interés ya que pueden ser usadas como blanco para inhibidores con mayor afinidad por los centros activos de estas proteínas. El dominio tiorredoxina de *T. gondii* en esta familia de proteínas es esencial para que ellas tengan un correcto plegamiento y disposición estructural (Collet and Messens, 2010). Dentro de la estructura que pudimos encontrar se encontró la prolina ubicada en la hélice alfa que precede al centro activo, la cual se encuentra conservada en el dominio **a** para el 69% de las proteínas analizadas y la cual está involucrada en la formación de una curvatura característica en la hélice alfa que precede al dominio catalítico.

Las secuencias de *T. gondii* muestran arquitecturas conservadas (ver datos suplementarios S1 y S2). También se identificó que esta familia de proteínas presentaba sintenia para todas las secuencias identificadas en los coccidios *H. hammondi* y *N. caninum* y en 18 de ellas con secuencias de *S. neurona* y *E. tenella*. El análisis filogenético muestra que las secuencias más cercanas a *T. gondii* son las pertenecientes a las secuencias de las especies *H. hammondi* y *N. caninum*, seguidas por *S. neurona* y las más lejanas, las de *E. tenella* (Morrison *et al.*, 2004; Lorenzi *et al.*, 2016).

A partir del análisis *in silico*, se identificaron en las secuencias de PDI de *T. gondii* características que sugieren donde puede estar localizadas a nivel subcelular. En 11 secuencias se encontró la presencia de péptido señal, por lo cual estas proteínas pueden translocar al retículo endoplásmico o RE

(Petersen *et al.*, 2011). En seis de ellas se encontraron regiones transmembranales y en seis se encontraron dominios KDEL de retención en RE. En otras dos, se observó un dominio KKxx, lo que sugiere que estas proteínas pueden estar relacionadas con la formación de puentes disulfuro en el RE y con el transporte retrogrado entre el RE y el aparato de Golgi (Sevier and Kaiser, 2002; Appenzeller-Herzog and Ellgaard, 2008). Esta localización indica que esta proteína puede participar en varios fenómenos particularmente importantes para el procesamiento de proteínas en eucariotas, como son las reacciones de glicosilación y la formación de puentes disulfuro en proteínas destinadas a la secreción (Hagiwara *et al.*, 2011). Esto es un proceso celular clave en el ciclo de vida de *T. gondii*, particularmente para la etapa de invasión, donde se reporta que inhibidores de proteínas PDI y de formación de enlaces disulfuro reducen el número de células invadidas por *T. gondii*. En *N. caninum* se obtuvieron hallazgos similares y las proteínas PDI estuvieron relacionadas con el proceso de interacción entre taquizoito y la célula hospedadora (Naguleswaran *et al.*, 2005; Moncada, Arenas, Acosta, Molina, Hernández, *et al.*, 2016).

En otros apicomplexos como *Plasmodium Falciparum* se ha relacionado que el factor de virulencia AMA-1 requiere la correcta disposición de puentes disulfuro en su conformación estructural activa (Mahajan *et al.*, 2006; Mouray *et al.*, 2007; Haque, Majumdar and Barik, 2012). Esto indica una relación con la actividad óxido reductora de esta familia de proteínas necesaria para el plegamiento de proteínas claves para la invasión y el crecimiento celular en el ciclo de vida del parásito. Existe evidencia adicional aportada por los tratamientos con inhibidores de proteínas PDI en *T. gondii*, *N. caninum* y *L. major* (Naguleswaran *et al.*, 2005; Müller *et al.*, 2008; Ben Khalaf *et al.*, 2012; Moncada, Arenas, Acosta, Molina, Hernández, *et al.*, 2016). Todos estos datos sugieren que las proteínas PDI son un blanco ideal para la búsqueda de nuevas drogas en parásitos coccidios.

En *T. gondii* experimentalmente se han descrito 6 proteínas de la familia PDI, de las cuales se identificó la localización subcelular de 3 de ellas y se reportó evidencia de inmunogenicidad de dos proteínas de la familia PDI. La proteína tiorredoxina ATrx1, con el código de identificación TGME49_312110, se encuentra localizada en el apicoplasto y esto sugiere que está involucrada en el transporte de proteínas hacia este organelo (DeRocher *et al.*, 2008). Así mismo dos proteínas ortologas de la cepa ME49 en GT1, con códigos de identificación TGME49_225790 y TGME49_232410 fueron localizadas y relacionadas a complejos proteínicos asociados con los microtúbulos corticales del parasito (Liu *et al.*, 2013).

Además dos proteínas de la familia las cuales se describen como altamente inmunogenicas la primera con código de identificación TGME49_293870 fue identificada como el mayor receptor de anticuerpos en productos de secreción del parasito lo cual sugiere una localización extracelular para esta proteína (Saadatnia *et al.*, 2012; Tommy *et al.*, 2012) y la proteína TGME49_211680, la cual ha sido la más estudiada de esta familia, se han identificado anticuerpos IgA en lágrimas, siendo el dominio tiorredoxina uno de los blancos para estos anticuerpos (Meek, Back, V. N. a Klaren, *et al.*, 2002; Meek, Back, V. N. A. Klaren, *et al.*, 2002). También se ha descrito reactividad cruzada de reconocimiento de anticuerpos en con su ortólogo en *N. caninum* y *B. besnoiti* (Marcelino *et al.*, 2011; Zhang *et al.*, 2011).

Además la proteína de la familia PDI TGME49_211680 se ha propuesto también como candidato vacunal y se ha descrito reactividad ante esta misma proteína de anticuerpos en sueros de ratón, conejo y humanos, mostrando su inmunogenicidad en diferentes especies (Wang *et al.*, 2013)

Funcionalmente esto sugiere que esta proteína en *T gondii* y su ortólogo en *N caninum* están relacionadas con la interacción entre taquizoito y célula hospedadora (Naguleswaran *et al.*, 2005; Moncada, Arenas, Acosta, Molina, Hernández, *et al.*, 2016).

Perfiles transcriptómicos de la familia de proteínas disulfuro isomerasa

Como es conocido en otros organismos (Han *et al.*, 2014; Zhu *et al.*, 2014; Porter, Yuen and Christopher, 2015) los perfiles transcriptómicos de la familia de PDI presentaron un perfil de expresión multiestadio y estadio específico en algunos casos, sugiriendo que la familia de proteínas PDI presenta picos de actividad específicos indicando una dependencia espaciotemporal para su expresión, lo cual es un fenómeno conocido para los perfiles de expresión de proteínas en *T. gondii* con respecto al ciclo de vida y su ciclo celular, con particularidades que dependen del estadio celular y del hospedador (Gaji *et al.*, 2012; Behnke *et al.*, 2014; Hehl *et al.*, 2015).

Cuando se analizan los perfiles de transcripción de algunas proteínas de la familia PDI agrupadas con respecto a los estadios entero-epiteliales de gato, se encuentra un perfil de expresión exclusivo para formas sexuales. Estas mismas proteínas tienen expresiones casi indetectables en los transcriptomas de taquizoitos de las cepas analizadas y en los estadios agudo y crónico en ratón, lo que sugiere que las proteínas TGME49_218470, TGME49_255480, TGME49_217710, TGME49_247660 tienen patrones de expresión restringidos a los estadios entéricos sexuales y no presentan expresión significativa durante la infección a células de hospedadores intermedios, tales como células humanas y de ratón, lo que sugiere que no tienen una labor esencial durante la invasión a células de hospedadores intermediarios

Además se observó que las proteínas con mayor nivel de expresión de la familia de proteínas PDI de *T. gondii* y cuyo patrón de expresión es multiestadio, fueron las proteínas de códigos TGME49_211680, 232410 y 293870 respectivamente, siendo la proteína TGME49_211680 la que presentaba los mayores niveles de expresión en todos los estadios estudiados y en 29 cepas de *T. gondii* (Figura S3 y S4). De esta manera la

proteína TGME49_211680 se posiciona como un candidato ideal de intervención, tal como se ha encontrado para secuencias homologas de otros parásitos intracelulares tales como *N. caninum*, *Plasmodium falciparum*, *Entamoeba histolytica*, *T. cruzi*, *L. amazonensis*, *L. donovani* y *L. major* (Achour *et al.*, 2002; Rubotham, 2005; Mahajan *et al.*, 2006; Mouray *et al.*, 2007; Müller *et al.*, 2008; Ramos *et al.*, 2008; Novo *et al.*, 2009; Ben Khalaf *et al.*, 2011, 2012; Haque, Majumdar and Barik, 2012)

No se escapa a nuestra atención que aunque la proteína TGME49_211680 o no tiene patrones de dependencia de expresión con un estadio particular o con una de las cepas analizadas, si presenta un aumento particular en la fase crónica de infección en el modelo murino, señalando posibles asociaciones con el estadio bradizoito, por lo cual sería importante evaluarla como blanco farmacéutico en manifestaciones crónicas de la enfermedad, tales como profilaxis en estados de inmunodeficiencia como el cáncer y el SIDA (Khan *et al.*, 2011; Biology *et al.*, 2012; Lee and Lee, 2017; Samanta *et al.*, 2017; Hettinghouse, Liu and Liu, 2018; Kyani *et al.*, 2018).

Fenotipo de los knockout producidos para la familia de proteínas disulfuro isomerasa de T. gondii GT1

Para una interpretación funcional que ayude a identificar un blanco candidato para inhibición, se contrastó la evidencia transcriptómica con el fenotipo de ajuste o *fitness* parasitario en cepas *knockout* para cada una de las secuencias de esta familia y se encontró que el *knockout* de la proteína TGME49_211680 fue el fenotipo que obtuvo el menor puntaje de capacidad de ajuste o sobrevivencia (puntaje -5,65), lo cual indica que esta proteína es esencial para el taquizoito, por lo tanto se puede concluir que esta proteína es un candidato claro para la inhibición como blanco terapéutico.

Tampoco se escapa a nuestra atención que la proteína TGME49_258826 a pesar de tener niveles de expresión multiestadio casi indetectables, pero cuyo valor de puntaje de adaptación o *fitness* fue de -3.88 sugiriendo que es importante para el parásito en el estadio taquizoito.

Análisis funcional de la proteína TGME49_211680

Como se señaló en la sección precedente, la proteína TGME49_211680 mostró la mayor producción de transcritos en todos los estadios analizados y en las tres cepas analizadas y con mayor expresión en estadios crónicos que en agudos. El *knockout* de esta proteína presenta el puntaje de fenotipo más bajo de la familia, lo cual sugiere que esta proteína es un blanco farmacéutico para *T. gondii*. Además se identificó que esta proteína tiene sintenia en las especies *N. caninum*, *E. tenella*, *S. neurona* y *H. hammondi*, las cuales son agentes causales de diferentes enfermedades en ganado bovino, producción avícola y la industria veterinaria (Dubey and Ferguson, 2015; McAllister, 2016; Pusterla and Tobin, 2017; Cheng *et al.*, 2018).

En este contexto y dada la conservación de esta secuencia en coccidios posiciona a la proteína TGME49_211680 como un candidato de inhibición en *T. gondii* debido a sus altos niveles de expresión multiestadio y el fenotipo inducido por su *knockout*, que sugieren la importancia de esta proteína en particular, para la cual se presentan ortólogos en otros parásitos de los géneros *Leishmania*, *Trypanosoma*, *Plasmodium*, para los cuales se han realizado tamizajes para la búsqueda de inhibidores y la descripción de esta como candidato vacunal, además de la posible relación de esta proteína en la invasión del parásito e interacción de esta proteína con la célula hospedadora en *T. gondii* y *N. caninum* (Naguleswaran *et al.*, 2005; Moncada, Arenas, Acosta, Molina, Hernández, *et al.*, 2016).

Además, la transcritos de la proteína TGME49_211680 presentó los mayores niveles de expresión antes de la invasión en células HeLa e inhibidores conocidos para proteínas PDI, reducen la invasión del taquizoito a las células humanas (Moncada, Arenas, Acosta, Molina, Hernandez, *et al.*, 2016). Los ensayos con proteína recombinante recTgTGME49_211680 mostraron que está involucrada en interacción con membranas de células hospedadoras. Lo cual también se ha sugerido para otros patógenos (Naguleswaran *et al.*, 2005; Stolf *et al.*, 2011; Moncada, Arenas, Acosta, Molina, Hernández, *et al.*, 2016). Un grupo de proteínas que pueden requerir la acción tiorreductora de las PDI y que puede explicar porqué son necesarias para la invasión, son las proteínas adhesinas de *T. gondii* y coccidios cercanos, las cuales poseen cisteínas en sus sitio de adhesión a la membrana celular del hospedero (Arenas *et al.*, 2012)

Análisis estructural y búsqueda in silico de alto rendimiento de moléculas con potencial de inhibición de la proteína TGME49_211680.

En la búsqueda de inhibidores con características que favorecieran la interacción espontánea con sitios funcionales de la PDI TGME49_211680 y sin afectar proteínas humanas de esta misma familia, los alineamientos de secuencias y estructurales entre PDI TGME49_211680 y proteínas disulfuro isomerasa humanas, permitieron identificar en la proteína del parásito diferentes regiones y aminoácidos conservados y diferentes respecto a esta familia de proteínas, importantes para la búsqueda de fármacos selectivos.

De esta forma, se identificó en los dominios a y a' de PDI TGME49_211680 una agrupación hidrofóbica altamente conservada entre la familia de proteínas disulfuro isomerasa, compuesta de residuos de fenilalanina para el dominio a y fenilalanina e isoleucina para el dominio a'. Dado que estos aminoácidos conservan además una ubicación espacial relativa similar a la encontrada en los dominios activos de proteínas humanas y que se han

descrito como esenciales para mantener una conformación adecuada del dominio tiorredoxina (Collet and Messens, 2010), es posible que en la proteína del parásito ejerciera esta misma función, siendo importantes a nivel estructural.

Otros aminoácidos de significancia estructural en los dominios activos de proteínas disulfuro isomerasa humanas, se conservan en identidad y ubicación espacial relativa en los dominios a y a' de PDI TGME 49_211680, donde podrían compartir esta función. Por ejemplo, las prolinas 40 y 364, podrían ejercer un efecto estabilizante de la estructura tiorredoxina de los dominios activos de PDI TGME49_211680, y los aminoácidos triptófano 31 y 355, alanina 29 y 353, y ácido aspártico 62 y 385, podrían hacer parte de un clúster estructural implicado en la estabilidad termodinámica de estos dominios, tal como se ha descrito en proteínas humanas (Collet and Messens, 2010).

Los residuos que conforman el centro activo y que son esenciales para la actividad catalítica de las proteínas disulfuro isomerasa, también se encuentran altamente conservados en PDI TGME49_211680. Estos aminoácidos corresponden a los residuos de cisteína 32, 35, 356 y 359, ubicados sobre la alfa hélice 2 de los dominios tiorredoxina a y a' (Figura 21), donde se llevan a cabo las reacciones de intercambio tiol-disulfuro a través de ciclos de oxidación- reducción (Hatahet *et al.*, 2009). Así mismo, entre las cisteínas 32 y 35 del dominio a de PDI TGME49_211680 se localizan los residuos glicina 33 e histidina 34, los cuales se consideran en otras proteínas como determinantes del poder reductor y potencial redox del sitio activo (Gruber *et al.*, 2006; Collet and Messens, 2010) y se encuentran conservados tanto en identidad como disposición espacial relativa respecto a dominios tiorredoxina humanos. En contraparte, aunque el dominio a' conserva una glicina entre las cisteínas 356 y 359 del sitio activo, en este dominio se presenta una variación con respecto a las proteínas humanas con

el motivo catalítico clásico CGHC, que consiste en la presencia de una tirosina en la posición 358 (Figura 20,21).

El significado funcional de esta mutación esta dado principalmente en que los aminoácidos ubicados entre las cisteínas del centro catalítico determinan la actividad de este a través de sus propiedades fisicoquímicas (Collet and Messens, 2010), de manera que la tendencia hacia la oxidación o reducción puede verse afectada por la presencia de un aminoácido hidrofóbico como tirosina, en vez de uno cargado positivamente como es la histidina altamente conservada en esta posición.

Hasta la fecha, el motivo CGYC solo ha sido encontrado en microorganismos patógenos como *Neospora caninum* (Naguleswaran *et al.*, 2005; Müller *et al.*, 2008), *Besnoitia besnoiti* (Marcelino *et al.*, 2011) y sus implicaciones se han visto en una de las proteínas oxidorreductasas de esta última conocida como DsbC, la cual desplaza su función hacia la isomerización o reducción principalmente (Rietsch and Beckwith, 1998, Rubotham *et al.*, 2005), en comparación a las proteínas que contienen el motivo clásico CGHC, que actúan más como oxidantes de enlaces disulfuro (Rietsch and Beckwith, 1998). Dado que en *T. gondii* podría presentarse este comportamiento, se destaca esta zona como una zona diferencial para blanco de inhibición, que puede inducir a la selectividad de moléculas inhibitoras hacia PDI TGME49_211680 respecto a proteínas humanas similares.

En relación al mantenimiento de la estructura y el potencial redox del centro catalítico de las proteínas disulfuro isomerasa humanas, también se han visto involucrados aminoácidos como prolina y treonina, que actúan principalmente en la disminución del pKa de la cisteína N-terminal del motivo CxxC (Kozlov *et al.*, 2010). Mientras el dominio a de PDI TGME49_211680 conserva estos aminoácidos tanto en identidad como disposición espacial relativa, en las posiciones 79 y 80 respectivamente, el dominio a' de PDI TGME49_211680 conserva la prolina en la posición 402 y presenta una

mutación de treonina por serina en la posición 403. Considerando que la función de la treonina en las proteínas humanas es establecer enlace de hidrógeno entre su grupo hidroxilo y el oxígeno de la cadena principal de la prolina antecesora (Eklund, Gleason and Holmgren, 1991) y el residuo de serina encontrado en PDI TGME49_211680 conserva la orientación de su grupo hidroxilo hacia esta prolina, y es un residuo propenso a formar puentes de hidrógeno; este aminoácido podría compartir esta función en la proteína del parásito.

Entre las regiones conservadas entre la familia de proteínas disulfuro isomerasa humanas y PDI TGME49_211680, se encuentra finalmente una región cargada en el interior de los dominios activos, compuesta por un ácido glutámico y una lisina, ubicados en las posiciones 43 y 60 del dominio a y 350 y 383 del dominio a'; los cuales podrían participar en la transferencia de protones durante la liberación del sustrato, como se ha descrito en proteínas humanas con dominio tiorredoxina (Kozlov *et al.*, 2010). En adición, ambos dominios activos de PDI TGME49_211680 conservan un residuo de arginina en las posiciones 96 y 421, que podría modular el pKa de las cisteínas catalíticas por su colocación en el sitio activo, a través el movimiento de su cadena lateral, dando lugar a la formación del tiolato necesario para el inicio de las reacciones de reducción (Kozlov *et al.*, 2010).

Aunque los aminoácidos con función conocida en las proteínas disulfuro isomerasa humanas son altamente conservados en PDI TGME49_211680, es de destacar que las mutaciones encontradas en la proteína del parásito se localizan en su dominio a', lo que hace a este dominio promisorio para la búsqueda de fármacos selectivos.(Cole *et al.*, 2018).

Es importante resaltar que, si PDI TGME49_211680 puede desplazar sus funciones hacia la reducción debido a la presencia del motivo catalítico CGYC, como ocurre en la proteína DsbC de *E. coli* (Rubotham *et al.*, 2005)(Rietsch and Beckwith, 1998), y estos cambios conformacionales

pueden tener efecto directamente sobre la actividad reductasa de las proteínas disulfuro isomerasa, este dominio sería un blanco ideal para el desarrollo selectivo de moléculas inhibitoras y que las moléculas acopladas en regiones que están involucradas en la flexibilidad de la proteína PDI TGME49_211680, podrían ser inhibidores promisorios de esta proteína (Lin *et al.*, 2015; Freedman *et al.*, 2017).

Por otra parte, los alineamientos estructurales y de secuencias entre PDI TGME49_211680 y proteínas disulfuro isomerasa humanas permitieron identificar, además, en los dominios inactivos de PDI TGME49_21168, regiones con posible importancia para la actividad de esta proteína. Así y acorde a lo reportado en diferentes PDI, TGME49_211680 muestra en la superficie de su dominio b' una zona hidrofóbica, conformada por aminoácidos poco conservados, que puede ser esencial para la unión del sustrato, al permitir el acoplamiento de proteínas desplegadas o parcialmente plegadas con residuos hidrofóbicos expuestos (Tian *et al.*, 2006) (Figura 23,24 S5).

Dado que el dominio b' presenta gran variabilidad a lo largo de la familia de proteínas disulfuro isomerasa, se ha reportado que la especificidad por el sustrato es controlada por el perfil hidrofóbico de esta región, el cual es determinante para la actividad de estas proteínas (Nguyen *et al.*, 2008). En efecto, la comparación entre los dominios b' de PDI TGME49_211680 y las proteínas humanas PDI A1 y ERP 57 evidencia diferentes perfiles de hidrofobicidad entre estas proteínas, sugiriendo el reconocimiento de diferentes sustratos para cada una de ellas.

Además de la importancia de los dominios a' y b' sobre la actividad de las proteínas disulfuro isomerasa, estudios de mutagénesis y dinámica molecular han demostrado que el enlazador x participa en la actividad catalítica de este tipo de enzimas, promoviendo cambios conformacionales particulares, fundamentales en el reconocimiento del sustrato y/o su liberación durante el

ciclo catalítico (Yang *et al.*, 2014; Freedman *et al.*, 2017). En consecuencia, en PDI A1 se ha evidenciado que la formación de puentes salinos entre los aminoácidos K308 - E359 y K326 - E431, induce a una configuración estructural con efectos significativos sobre su actividad reductasa (Yang *et al.*, 2014). De manera similar, en PDI TGME49_211680 se han encontrado aminoácidos cuya disposición espacial y características fisicoquímicas permitirían ejercer este tipo de interacciones, como la lisina 316 del enlazador x que podría formar un puente salino con el ácido glutámico 264 del dominio b' o un enlace de hidrógeno con la asparagina 265 de este dominio, y tres aminoácidos cargados positivamente, correspondientes a K280, K281 y R283 del dominio b', que podrían formar puentes salinos con los ácidos glutámicos 394 y 395 del dominio a', para promover los cambios conformacionales necesarios para la actividad de la proteína.

En consecuencia, la significancia energética de este tipo de interacciones contribuye a la estabilización del complejo proteína-ligando, con lo cual se podría dar lugar a cambios en los potenciales redox del sitio activo o impedimento a la desprotonación de la cisteína 356, que finalmente podría traducirse en incapacidad de la proteína para catalizar reacciones de oxidación-reducción (Collet and Messens, 2010).

En adición, otras interacciones importantes se establecen entre el ligando ZINC 08610024 y la proteína PDI TGME49_211680, correspondientes a un enlace de hidrógeno de fuerza moderada con la arginina 421 y un contacto electrostático con la prolina 402 (Figura 26A). Dado que estos aminoácidos podrían modular el pKa de la cisteína N-terminal del motivo catalítico, principalmente por el movimiento de la cadena lateral del residuo de arginina; un enlace de hidrógeno con este aminoácido podría conferirle rigidez e imposibilitar su movimiento, impidiendo el cambio de pKa necesario para

iniciar las reacciones de intercambio tiol-disulfuro (Collet and Messens, 2010; Depuydt, Messens and Collet, 2011; Lin *et al.*, 2015; Freedman *et al.*, 2017).

De forma análoga, el ligando ZINC 09642074 interactúa con el centro catalítico del dominio a' de PDI TGME49_211680 a través de contactos de Van der Waals con la cisteína 356 y glicina 357 e interacciones hidrofóbicas tipo pi-pi con la tirosina 358. Así mismo, se establece un enlace de hidrógeno de fuerza moderada con la arginina 421 y contactos electrostáticos con la prolina 402 (Figura 26B). Adicional a las interacciones reportadas para el ligando ZINC 08610024, el ligando ZINC 09642074 tiene contactos de Van der Waals con el triptófano 355, el cual podría estar implicado en la estabilidad termodinámica del dominio tiorredoxina de PDI TGME49_211680.

El acoplamiento de estos ligandos a las proteínas humanas PDIA1 y ERP57 como evaluación de su selectividad, permitió evidenciar que, aunque el ligando ZINC 08610024 se acopla en los dominios a y a' de las proteínas humanas con energías de unión cercanas a la obtenida en la proteína del parásito, esta molécula no establece ningún tipo de interacción con los aminoácidos del sitio activo de estas proteínas figura S6 y 21. Por consiguiente, el ligando ZINC 08610024 podría no tener efectos directamente sobre la ejecución de las reacciones redox catalizadas por las proteínas PDIA1 y ERP57, y es por lo tanto un candidato potencial para ensayos de inhibición experimental de PDI TGME49_211680.

Por otra parte, aunque el ligando ZINC 09642074 no interactúa con el centro catalítico del dominio a de PDIA1, ni el centro catalítico del dominio a' de ERP57; este evidencia interacciones con aminoácidos determinantes del poder reductor y potencial redox de los demás sitios activos de estas proteínas (Collet and Messens, 2010) (Gruber *et al.*, 2006). De esta forma, la molécula ZINC 09642074 establece un enlace de hidrógeno con la glicina 398 del dominio a' de PDIA1, e interacciones electrostáticas tipo pi-cación con la histidina 59 del dominio a de ERP57 (Figura S6,21,25 y 26). No

obstante, debido a que las energías de unión con estos dominios humanos son mayores a la energía de unión con PDI TGME49_211680, la interacción con la proteína del parásito estaría energéticamente más favorecida, y podría presentarse selectividad hacia esta. Además de los contactos establecidos dentro del bolsillo hidrofóbico del dominio b' de PDI TGME49_211680, los ligandos ZINC 12210072, 15784130 y 33010617 interactúan con aminoácidos que podrían ser importantes para la flexibilidad de esta proteína. Así, el posible enlace de hidrógeno entre la lisina 316 y la asparagina 265, requerido para un cambio conformacional durante la actividad de PDI TGME49_211680, puede ser impedido o desestabilizado por el enlace de hidrógeno establecido entre la molécula ZINC 12210072 y la asparagina 265 y el contacto de Van der Waals entre este ligando y la lisina 316. En adición, el ligando ZINC 12210072 podría restringir la flexibilidad del enlazador x, debido a las interacciones efectuadas con aminoácidos de esta región, como serina 314, arginina 313 y leucina 315, e impedir el acercamiento entre K280, K281 y R283 y los ácidos glutámicos 394 y 395, necesario para la formación de puentes salinos durante los cambios de configuración espacial que requiera la proteína.

Un efecto de restricción de los cambios conformacionales de PDI TGME49_211680 por impedimento en la formación de enlace de hidrógeno entre lisina 316 y asparagina 265, también podría ser logrado por los ligandos ZINC 33010617 y 15784130, los cuales interactúan con estos aminoácidos a través de interacciones alquílicas y contactos de Van der Waals, respectivamente. Así mismo, estas moléculas podrían limitar la flexibilidad del enlazador x, principalmente por la formación de enlaces de hidrógeno energéticamente significativos con la serina 314 e interacciones débiles con diferentes aminoácidos de esta región.

La búsqueda de inhibidores dirigida hacia la región comprendida entre el dominio b' y el enlazador x, permitió identificar que la molécula de código

ZINC 20564114 interactúa con esta región, con una energía de -9,7 kcal/mol figura 28. Como resultado de este acoplamiento, planteamos que la PDI TGME49_211680 podría tener una flexibilidad restringida, debido posiblemente al bloqueo de las interacciones entre los aminoácidos que pueden formar puentes salinos y de hidrógeno para promover cambios conformacionales en la proteína. En efecto, el ligando ZINC 20564114 se ubica entre la triada de aminoácidos cargados positivamente (K280, K281 y R283) y los ácidos glutámicos 394 y 395, impidiendo el contacto entre estos. Este acoplamiento se ve favorecido principalmente por el establecimiento de interacciones alquílicas con arginina 283, lisina 316 y serina 314 e interacciones de tipo pi- anión con el ácido glutámico 394 y pi- catión con arginina 313 (Figura 29 D).

El acoplamiento de los ligandos ZINC 12210072, 15784130, 33010617 y 20564114 a las proteínas humanas PDIA1 y ERP57 como evaluación de su selectividad, permitió evidenciar que, aunque la mayoría de estas moléculas interactúan con regiones análogas a las evaluadas en la proteína del parásito, las energías de estos acoplamientos son mayores que la energía de acoplamiento en PDI TGME49_211680. Estos resultados sugieren que el acoplamiento de los ligandos ZINC 12210072, 15784130, 33010617 y 20564114 con PDI TGME49_211680 es energéticamente más favorecido en comparación a las proteínas humanas, posiblemente por la interacción con un número mayor de aminoácidos que pueden estabilizar el complejo proteína- ligando figura 27, 28 y 29). En adición, estos ligandos parecen tener afinidad por regiones de potencial electrostático hidrofóbico- hidrófilo, el cual es una de las características diferenciales entre el dominio b' de PDI TGME49_211680, respecto a las proteínas humanas, que, sumado a menores energías de acoplamiento en el parásito, podría generar selectividad de estas moléculas hacia esta proteína.

En este acercamiento computacional se analizaron las características estructurales de la proteína PDI TGME49_211680, que pudieran dirigir una búsqueda restringida a las características estructurales diferenciales de esta con respecto a las proteínas PDI humanas con más identidades. Se encontraron diferencias en el centro activo del dominio a' y los perfiles de hidrofobicidad de la superficie de los dominios b y b' con respecto a las proteínas humanas analizadas y se identificaron seis moléculas que podrían inducir a conformaciones estructurales que restringen la flexibilidad de la proteína y la actividad redox, de la proteína TGME49_211680. Todos los ligandos encontrados por medio de este acercamiento computacional se han reportado como no tóxicos y actualmente se encuentran disponibles para la venta, por lo cual representan candidatos muy atractivos para realizar ensayos de inhibición de la proteína PDI TGME49_211680 de *T. gondii*.

También es de destacar que inhibidores de esta proteína podrían ser candidatos para intervenciones en diferentes patologías como enfermedad trombótica, SIDA, cáncer y en infecciones como las causadas por los virus influenza A y B, ya que se conoce la relación entre la progresión de estas enfermedades con la actividad y expresión de las proteínas PDI además de que en casos de periodos de reactivación de *T. gondii* en inmunodeficiencias como sida y cáncer inhibidores para las proteínas PDI podrían ser una alternativa (Gowthaman, Jayakanthan and Sundar, 2008; Khan *et al.*, 2011; Biology *et al.*, 2012; Bekendam and Flaumenhaft, 2016; Flaumenhaft, Furie and Zwicker, 2016; Bechtel, Weerapana and Hill, 2017; Samanta *et al.*, 2017; Cole *et al.*, 2018; Hettinghouse, Liu and Liu, 2018; Kim and Chang, 2018; Kyani *et al.*, 2018).

Conclusiones

En este trabajo se identificaron 23 proteínas PDI en *T. gondii*, en las cuales se pudo describir que son una familia de proteínas de expresión multiestadio, con múltiples localizaciones subcelulares y, dentro de ellas, 15 proteínas mostraron ser esenciales para la sobrevivencia del taquizoito. Esto demuestra que esta familia de proteínas tiene un papel importante en el ciclo de vida de *T. gondii*.

Nuestro análisis de predicción bioinformático demostró que estas proteínas tienen características estructurales relacionadas con el plegamiento adecuado de proteínas, permitiendo la disposición correcta de puentes disulfuro, el cual es un paso esencial para el funcionamiento de múltiples rutas de señalización intracelular para parásitos protozoarios.

El conjunto de datos de transcriptómica, del uso de inhibidores específicos de actividad tiorredoxina y la localización por inmunofluorescencia, así como la información obtenida por modelamiento estructural, muestran que la proteína PDI TGME49_211680 de *T. gondii* es un candidato para intervenciones con inhibidores específicos con bajo reconocimiento para proteínas humanas.

Nuestros resultados abren un panorama de amplio espectro para el estudio de esta familia de proteínas, que van desde su ontología de gen, su evaluación como blanco de candidatos vacunales o de compuestos inhibidores con fines terapéuticos para *T. gondii* y otros coccidios de importancia para la salud humana y animal.

Bibliografía

Achour, Y. Ben *et al.* (2002) 'Identification of a Disulfide Isomerase Protein of Leishmania major as a Putative Virulence Factor Identification of a Disulfide Isomerase Protein of Leishmania major as a Putative Virulence Factor', *Infection and immunity*, 70(7), pp. 3576–3585. doi: 10.1128/IAI.70.7.3576.

Ali Khan, H. and Mutus, B. (2014) 'Protein disulfide isomerase a multifunctional protein with multiple physiological roles.', *Frontiers in chemistry*, 2(August), p. 70. doi: 10.3389/fchem.2014.00070.

Appenzeller-Herzog, C. and Ellgaard, L. (2008) 'The human PDI family: Versatility packed into a single fold', *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1783(4), pp. 535–548. doi: 10.1016/j.bbamcr.2007.11.010.

Arenas, A. F. *et al.* (2012) 'Cluster analysis identifies aminoacid compositional features that indicate Toxoplasma gondii adhesin proteins Abstract : Background ':, 8(19).

Aurrecoechea, C. *et al.* (2017) 'EuPathDB : the eukaryotic pathogen genomics database resource', 45(November 2016), pp. 581–591. doi: 10.1093/nar/gkw1105.

Bechtel, T. J., Weerapana, E. and Hill, C. (2017) 'HHS Public Access', 17(6), pp. 1–34. doi: 10.1002/pmic.201600391.From.

Behnke, M. S. *et al.* (2014) 'Toxoplasma gondii merozoite gene expression analysis with comparison to the life cycle discloses a unique expression state during enteric development Toxoplasma gondii merozoite gene expression analysis with comparison to the life cycle discloses a unique expression state during enteric development'.

Bekendam, R. H. and Flaumenhaft, R. (2016) 'MiniReview Inhibition of

- Protein Disulfide Isomerase in Thrombosis', pp. 42–48. doi: 10.1111/bcpt.12573.
- Biology, C. *et al.* (2012) 'The International Journal of Biochemistry Thioredoxin-1 and protein disulfide isomerase catalyze the reduction of similar disulfides in HIV gp120', 44, pp. 556–562. doi: 10.1016/j.biocel.2011.12.015.
- Cheng, P. *et al.* (2018) 'Research in Veterinary Science Pharmacokinetics of a novel triazine ethanamizuril in rats and broiler chickens', *Research in Veterinary Science*. Elsevier, 117(November 2017), pp. 99–103. doi: 10.1016/j.rvsc.2017.11.001.
- Chou, K.-C. and Shen, H.-B. (2008) 'Cell-PLoc: a package of Web servers for predicting subcellular localization of proteins in various organisms', *Nature Protocols*. Nature Publishing Group, 3, p. 153. Available at: <http://dx.doi.org/10.1038/nprot.2007.494>.
- Cole, K. S. *et al.* (2018) 'Characterization of an A-Site Selective Protein Disulfide Isomerase A1 Inhibitor'. doi: 10.1021/acs.biochem.8b00178.
- Collet, J.-F. and Messens, J. (2010) 'Structure, function, and mechanism of thioredoxin proteins.', *Antioxidants & redox signaling*, 13(8), pp. 1205–1216. doi: 10.1089/ars.2010.3114.
- Cuche, A. *et al.* (2013) 'New and continuing developments at PROSITE', 41(November 2012), pp. 344–347. doi: 10.1093/nar/gks1067.
- Depuydt, M., Messens, J. and Collet, J.-F. (2011) 'How proteins form disulfide bonds.', *Antioxidants & Redox Signaling*, 15(1), pp. 49–66. doi: 10.1089/ars.2010.3575.
- Dereeper, A. *et al.* (2008) 'Phylogeny . fr : robust phylogenetic analysis for the', 36(April), pp. 465–469. doi: 10.1093/nar/gkn180.

- DeRocher, A. E. *et al.* (2008) 'A thioredoxin family protein of the apicoplast periphery identifies abundant candidate transport vesicles in *Toxoplasma gondii*', *Eukaryotic Cell*, 7(9), pp. 1518–1529. doi: 10.1128/EC.00081-08.
- Dickerhof, N. *et al.* (2011) 'Bacitracin inhibits the reductive activity of protein disulfide isomerase by disulfide bond formation with free cysteines in the substrate-binding domain', *FEBS Journal*, 278, pp. 2034–2043. doi: 10.1111/j.1742-4658.2011.08119.x.
- Dong, G. *et al.* (2009) 'Insights into MHC class I peptide loading from the structure of the tapasin/ERp57 heterodimer', *Immunity*, 30(1), pp. 21–32. doi: 10.1016/j.immuni.2008.10.018.
- Dubey, J. P. and Ferguson, D. J. P. (2015) 'Life Cycle of *Hammondia hammondi* (Apicomplexa : Sarcocystidae) in Cats', (2013), pp. 346–352. doi: 10.1111/jeu.12188.
- Eklund, H., Gleason, F. and Holmgren, A. (1991) 'Structural and functional relations among thioredoxins of different species', *Proteins*, 11, pp. 13–28.
- Emsley, B. J. and Bonds, S. H. (1968) 'Very Strong Hydrogen Bonding', *Chemical Society Reviews*, 9, pp. 91–124.
- Flaumenhaft, R., Furie, B. and Zwicker, J. I. (2016) 'NIH Public Access', 35(1), pp. 16–23. doi: 10.1161/ATVBAHA.114.303410.Therapeutic.
- Freedman, R. B. *et al.* (2017) 'BBA - Proteins and Proteomics " Something in the way she moves " : The functional significance of flexibility in the multiple roles of protein disulfide isomerase (PDI)', 1865(August), pp. 1383–1394. doi: 10.1016/j.bbapap.2017.08.014.
- Gaji, R. Y. *et al.* (2012) 'expression', 79(1), pp. 192–204. doi: 10.1111/j.1365-2958.2010.07441.x.Cell.
- Gajria, B. *et al.* (2008) 'ToxoDB : an integrated *Toxoplasma gondii* database

resource', 36(November 2007), pp. 553–556. doi: 10.1093/nar/gkm981.

Galligan, J. J. and Petersen, D. R. (2012) 'The human protein disulfide isomerase gene family', *Human Genomics*, 6(1), p. 6. doi: 10.1186/1479-7364-6-6.

Goldberger, R. F., Epstein, C. J. and Anfinsen, C. B. (1963) 'Acceleration of Reactivation of Reduced Bovine Pancreatic Ribonuclease by a Microsomal System from Rat Liver', *Journal of Biological Chemistry*, 238(2), pp. 628–635. Available at: <http://www.jbc.org/content/238/2/628.short>.

Goldberger, R. F., Epstein, C. J. and Anfinsen, C. B. (1964) 'Purification and Properties of a Microsomal Enzyme System Catalyzing the Reactivation of Reduced Ribonuclease and Lysozyme', *Journal of Biological Chemistry*, 239(5), pp. 1406–1410. Available at: <http://www.jbc.org/content/239/5/1406.short>.

Gowthaman, U., Jayakanthan, M. and Sundar, D. (2008) 'Molecular docking studies of dithionitrobenzoic acid and its related compounds to protein disulfide isomerase : computational screening of inhibitors to HIV-1 entry', 10, pp. 1–10. doi: 10.1186/1471-2105-9-S12-S14.

Gruber, C. W. *et al.* (2006) 'Protein disulfide isomerase: the structure of oxidative folding', *Trends in Biochemical Sciences*, 31(8), pp. 455–464. doi: 10.1016/j.tibs.2006.06.001.

Hagiwara, M. *et al.* (2011) 'Structural Basis of an ERAD Pathway Mediated by the ER-Resident Protein Disulfide Reductase ERdj5', *Molecular Cell*. Elsevier Inc., 41(4), pp. 432–444. doi: 10.1016/j.molcel.2011.01.021.

Han, H. *et al.* (2014) 'Molecular characterization and analysis of a novel protein disulfide isomerase-like protein of *Eimeria tenella*.', *PloS one*, 9(6), p. e99914. doi: 10.1371/journal.pone.0099914.

Haque, S. J., Majumdar, T. and Barik, S. (2012) 'Redox-Assisted Protein

Folding Systems in Eukaryotic Parasites', *Antioxidants & Redox Signaling*, 17(4), pp. 674–683. doi: 10.1089/ars.2011.4433.

Hatahet, F. *et al.* (2009) 'Protein Disulfide Isomerase : A Critical Evaluation of Its Function in Disulfide Bond Formation', *Antioxidants & Redox Signaling*, 11(11).

Hehl, A. B. *et al.* (2015) 'Asexual expansion of *Toxoplasma gondii* merozoites is distinct from tachyzoites and entails expression of non-overlapping gene families to attach, invade, and replicate within feline enterocytes', *BMC Genomics*, 16(1), p. 66. doi: 10.1186/s12864-015-1225-x.

Hettinghouse, A., Liu, R. and Liu, C. (2018) 'Pharmacology & Therapeutics Multifunctional molecule ERp57 : From cancer to neurodegenerative diseases', *Pharmacology and Therapeutics*. Elsevier Inc., 181, pp. 34–48. doi: 10.1016/j.pharmthera.2017.07.011.

Hunter, C. a. and Sibley, L. D. (2012) 'Modulation of innate immunity by *Toxoplasma gondii* virulence effectors', *Nature Reviews Microbiology*, 10(11), pp. 766–778. doi: 10.1038/nrmicro2858.

Irwin, J. J. and Shoichet, B. K. (2005) 'ZINC - A free database of commercially available compounds for virtual screening', *Journal of Chemical Information and Modeling*, 45(1), pp. 177–182. doi: 10.1021/ci049714+.

Kayum, A. *et al.* (2017) 'Genome-wide characterization and expression profiling of PDI family gene reveals function as abiotic and biotic stress tolerance in Chinese cabbage (*Brassica rapa* ssp . *pekinensis*)'. *BMC Genomics*, pp. 1–20. doi: 10.1186/s12864-017-4277-2.

Ben Khalaf, N. *et al.* (2011) 'A high-throughput turbidometric assay for screening inhibitors of *Leishmania major* protein disulfide isomerase', *J Biomol Screen*, 16(5), pp. 545–551. doi: 10.1177/1087057111401026.

Ben Khalaf, N. *et al.* (2012) '*Leishmania major* protein disulfide isomerase as

- a drug target: Enzymatic and functional characterization', *Parasitology Research*, 110(5), pp. 1911–1917. doi: 10.1007/s00436-011-2717-5.
- Khan, M. M. G. *et al.* (2011) 'Discovery of a Small Molecule PDI Inhibitor That Inhibits Reduction of HIV-1 Envelope Glycoprotein gp120', pp. 245–251. doi: 10.1021/cb100387r.
- Kim, Y. and Chang, K. (2018) 'Protein disulfide isomerases as potential therapeutic targets for influenza A and B viruses', 247(September 2017), pp. 26–33. doi: 10.1016/j.virusres.2018.01.010.
- Kozlov, G. *et al.* (2010) 'A structural overview of the PDI family of proteins', *FEBS Journal*, 277(19), pp. 3924–3936. doi: 10.1111/j.1742-4658.2010.07793.x.
- Kremer, K. *et al.* (2013) 'An Overexpression Screen of *Toxoplasma gondii* Rab- GTPases Reveals Distinct Transport Routes to the Micronemes', 9(3). doi: 10.1371/journal.ppat.1003213.
- Krossa, S. *et al.* (2018) 'Redundancy of protein disulfide isomerases in the catalysis of the inactivating disulfide switch in A Disintegrin and Metalloprotease 17', (December 2017), pp. 2–10. doi: 10.1038/s41598-018-19429-4.
- Kyani, A. *et al.* (2018) 'Discovery and Mechanistic Elucidation of a Class of Protein Disulfide Isomerase Inhibitors for the Treatment of Glioblastoma', pp. 164–177. doi: 10.1002/cmdc.201700629.
- Lee, E. and Lee, D. H. (2017) 'Emerging roles of protein disulfide isomerase in cancer', 50(June), pp. 401–410.
- Letunic, I. and Bork, P. (2017) '20 years of the SMART protein domain annotation resource'. Oxford University Press, (December), pp. 4–7. doi: 10.1093/nar/gkx922.

- Lin, L. *et al.* (2015) 'Quercetin-3-rutinoside Inhibits Protein Disulfide Isomerase by Binding to Its b' x Domain *', *The Journal of Biological Chemistry*, 290(39), pp. 23543–23552. doi: 10.1074/jbc.M115.666180.
- Liu, J. *et al.* (2013) 'Novel Thioredoxin-like proteins are components of a protein complex coating the cortical microtubules of *Toxoplasma gondii*', *Eukaryotic Cell*, 12(12), pp. 1588–1599. doi: 10.1128/EC.00082-13.
- Lorenzi, H. *et al.* (2016) 'Local admixture of amplified and diversified secreted pathogenesis determinants shapes mosaic *Toxoplasma gondii* genomes', *Nature Communications*, 7, p. 10147. doi: 10.1038/ncomms10147.
- Mahajan, B. *et al.* (2006) 'Protein disulfide isomerase assisted protein folding in malaria parasites', *International Journal for Parasitology*, 36(9), pp. 1037–1048. doi: 10.1016/j.ijpara.2006.04.012.
- Marcelino, E. *et al.* (2011) 'Besnoitia besnoiti protein disulfide isomerase (BbPDI): molecular characterization, expression and in silico modelling.', *Experimental parasitology*. Elsevier Inc., 129(2), pp. 164–74. doi: 10.1016/j.exppara.2011.06.012.
- Marchler-bauer, A. *et al.* (2017) 'CDD / SPARCLE : functional classification of proteins via subfamily domain architectures', 45(November 2016), pp. 200–203. doi: 10.1093/nar/gkw1129.
- Markikou-Ouni, W., Ben Achour-Chenik, Y. and Meddeb-Garnaoui, a. (2012) 'Effects of *Leishmania major* clones showing different levels of virulence on infectivity, differentiation and maturation of human dendritic cells', *Clinical and Experimental Immunology*, 169(3), pp. 273–280. doi: 10.1111/j.1365-2249.2012.04611.x.
- Mcallister, M. M. (2016) 'D i a g n o s i s a n d C o n t r o l o f B o v i n e N e o s p o r o s i s Neospora caninum Abortion Transmission Prevention Parasite Review', 32, pp. 443–463. doi: 10.1016/j.cvfa.2016.01.012.

Meek, B., Back, J. W., Klaren, V. N. A., *et al.* (2002) 'Conserved regions of protein disulfide isomerase are targeted by natural IgA antibodies in humans', *International Immunology*, 14(11), pp. 1291–1301.

Meek, B., Back, J. W., Klaren, V. N. a, *et al.* (2002) 'Protein disulfide isomerase of *Toxoplasma gondii* is targeted by mucosal IgA antibodies in humans', *FEBS Letters*, 522(1–3), pp. 104–108. doi: 10.1016/S0014-5793(02)02911-3.

Moncada, D., Arenas, A., Acosta, A., Molina, D., Hernández, A., *et al.* (2016) 'Role of the 52 KDa thioredoxin protein disulfide isomerase of *Toxoplasma gondii* during infection to human cells', *Experimental Parasitology*, 164, pp. 36–42. doi: 10.1016/j.exppara.2016.02.005.

Moncada, D., Arenas, A., Acosta, A., Molina, D., Hernandez, A., *et al.* (2016) 'Role of the 52 KDa thioredoxin protein disulfide isomerase of *Toxoplasma gondii* during infection to human cells', *Experimental Parasitology*.

Montoya, J. G. and Liesenfeld, O. (2004) 'Toxoplasmosis', 363, pp. 1965–1976.

Morrison, D. A. *et al.* (2004) 'The current status of the small subunit rRNA phylogeny of the coccidia (Sporozoa) q', 34, pp. 501–514. doi: 10.1016/j.ijpara.2003.11.006.

Mouray, E. *et al.* (2007) 'Biochemical properties and cellular localization of *Plasmodium falciparum* protein disulfide isomerase', *Biochimie*, 89(3), pp. 337–346. doi: 10.1016/j.biochi.2006.11.001.

Müller, J. *et al.* (2008) 'Neospora caninum: functional inhibition of protein disulfide isomerase by the broad-spectrum anti-parasitic drug nitazoxanide and other thiazolidines.', *Experimental parasitology*, 118(1), pp. 80–8. doi: 10.1016/j.exppara.2007.06.008.

Naguleswaran, A. *et al.* (2005) 'Neospora caninum protein disulfide

isomerase is involved in tachyzoite-host cell interaction.’, *International journal for parasitology*, 35(13), pp. 1459–72. doi: 10.1016/j.ijpara.2005.06.006.

Naguleswaran, A., Müller, N. and Hemphill, A. (2003) ‘Neospora caninum and Toxoplasma gondii: A novel adhesion/invasion assay reveals distinct differences in tachyzoite-host cell interactions’, *Experimental Parasitology*, 104(3–4), pp. 149–158. doi: 10.1016/S0014-4894(03)00137-1.

Nguyen, V. D. *et al.* (2008) ‘Alternative Conformations of the x Region of Human Protein Disulphide-Isomerase Modulate Exposure of the Substrate Binding b’ Domain’, *Journal of Molecular Biology*. Elsevier Ltd, pp. 1144–1155. doi: 10.1016/j.jmb.2008.08.085.

Novo, C. *et al.* (2009) ‘Gene sequencing, modelling and immunolocalization of the protein disulfide isomerase from Plasmodium chabaudi’, *International Journal of Biological Macromolecules*, 45(4), pp. 399–406. doi: 10.1016/j.ijbiomac.2009.07.001.

Pedone, E., Limauro, D. and Bartolucci, S. (2008) ‘The machinery for oxidative protein folding in thermophiles.’, *Antioxidants & redox signaling*, 10(1), pp. 157–169. doi: 10.1089/ars.2007.1855.

Petersen, T. N. *et al.* (2011) ‘correspondence SignalP 4.0: discriminating signal peptides from transmembrane regions’, *Nature Publishing Group*. Nature Publishing Group, 8(10), pp. 785–786. doi: 10.1038/nmeth.1701.

Pittman, K. J., Aliota, M. T. and Knoll, L. J. (2014) ‘Dual transcriptional profiling of mice and Toxoplasma gondii during acute and chronic infection.’, *BMC genomics*, 15(1), p. 806. doi: 10.1186/1471-2164-15-806.

Popielarski, M. *et al.* (2018) ‘Biochemical and Biophysical Research Communications The role of Protein Disulfide Isomerase and thiol bond modifications in activation of integrin subunit alpha11’, 495, pp. 1635–1641. doi: 10.1016/j.bbrc.2017.11.186.

Porter, B. W., Yuen, C. Y. L. and Christopher, D. a. (2015) 'Dual protein trafficking to secretory and non-secretory cell compartments: Clear or double vision?', *Plant Science*. Elsevier Ireland Ltd, 234, pp. 174–179. doi: 10.1016/j.plantsci.2015.02.013.

Pusterla, N. and Tobin, T. (2017) 'Therapeutics for E q u i n e P r o t o z o a l Myeloencephalitis', 33, pp. 87–97. doi: 10.1016/j.cveq.2016.12.001.

Ramos, M. a *et al.* (2011) 'Entamoeba histolytica: biochemical characterization of a protein disulfide isomerase.', *Experimental parasitology*. Elsevier Inc., 128(1), pp. 76–81. doi: 10.1016/j.exppara.2011.02.009.

Ramos, M. a. *et al.* (2008) 'In silico identification of the protein disulfide isomerase family from a protozoan parasite', *Computational Biology and Chemistry*, 32, pp. 66–70. doi: 10.1016/j.compbiolchem.2007.09.002.

Ridgeway, T. and Nw, L. (2000) 'T-Coffee : A Novel Method for Fast and Accurate Multiple Sequence Alignment'. doi: 10.1006/jmbi.2000.4042.

Rietsch, A. and Beckwith, J. (1998) 'The genetics of disulfide bond metabolism', 32, pp. 84–163.

Roy, A., Kucukural, A. and Zhang, Y. (2010) 'I-TASSER : a unified platform for automated protein structure and function prediction', 5(4), pp. 725–738. doi: 10.1038/nprot.2010.5.

Rubotham, J. (2005) 'Characterization of Two Protein Disulfide Isomerases from the Endocytic Pathway of Bloodstream Forms of Trypanosoma brucei', *Journal of Biological Chemistry*, 280(11), pp. 10410–10418. doi: 10.1074/jbc.M409375200.

Rubotham, J. *et al.* (2005) 'Characterization of Two Protein Disulfide Isomerases from the Endocytic Pathway of Bloodstream Forms of Trypanosoma brucei *', *The Journal of Biological Chemistry*, 280(11), pp. 10410–10418. doi: 10.1074/jbc.M409375200.

Saadatnia, G. *et al.* (2012) 'Toxoplasma gondii excretory secretory antigenic proteins of diagnostic potential', *Apmis*, 120(1), pp. 47–55. doi: 10.1111/j.1600-0463.2011.02810.x.

Safavi-Hemami, H. *et al.* (2016) 'Rapid expansion of the protein disulfide isomerase gene family facilitates the folding of venom peptides', *Proceedings of the National Academy of Sciences*, 113(12), pp. 3227–3232. doi: 10.1073/pnas.1525790113.

Samanta, S. *et al.* (2017) 'Expression of protein disulfide isomerase family members correlates with tumor progression and patient survival in ovarian cancer', 8(61), pp. 103543–103556.

Sambasivarao, S. V (2013) 'NIH Public Access', 18(9), pp. 1199–1216. doi: 10.1016/j.micinf.2011.07.011.Innate.

Sevier, C. S. and Kaiser, C. A. (2002) 'Formation and transfer of disulphide bonds in living cells', *Nature Reviews Molecular Cell Biology*, 3(11), pp. 836–847. doi: 10.1038/nrm954.

Sidik, S. M. *et al.* (2016) 'A Genome-wide CRISPR Screen in Toxoplasma Identifies Essential Apicomplexan Genes', *Cell*. Elsevier Inc., 166(6), p. 1423–1435.e12. doi: 10.1016/j.cell.2016.08.019.

Sigrist, C. J. A. *et al.* (2002) 'PROSITE : A documented database using patterns and profiles as motif descriptors', 3(3), pp. 265–274.

Soares Moretti, A. I. and Martins Laurindo, F. R. (2017) 'Protein disulfide isomerases: Redox connections in and out of the endoplasmic reticulum', *Archives of Biochemistry and Biophysics*. Elsevier Inc, 617, pp. 106–119. doi: 10.1016/j.abb.2016.11.007.

Stolf, B. S. *et al.* (2011) 'Protein Disulfide Isomerase and Host-Pathogen Interaction', *The Scientific World JOURNAL*, 11, pp. 1749–1761. doi: 10.1100/2011/289182.

- Tian, G. *et al.* (2006) 'The Crystal Structure of Yeast Protein Disulfide Isomerase Suggests Cooperativity between Its Active Sites', pp. 61–73. doi: 10.1016/j.cell.2005.10.044.
- Tommy, Y. B. W. *et al.* (2012) 'Theoretical investigation on structural, functional and epitope of a 12 kDa excretory-secretory protein from *Toxoplasma gondii*.', *BMC structural biology*. *BMC Structural Biology*, 12(1), p. 30. doi: 10.1186/1472-6807-12-30.
- Trott, O. and Olson, A. J. (2009) 'Software News and Update AutoDock Vina : Improving the Speed and Accuracy of Docking with a New Scoring Function , Efficient Optimization , and Multithreading'. doi: 10.1002/jcc.
- Turano, C. *et al.* (2002) 'Proteins of the PDI family: Unpredicted non-ER locations and functions', *Journal of Cellular Physiology*, 193(2), pp. 154–163. doi: 10.1002/jcp.10172.
- Wang, H. L. *et al.* (2013) 'Toxoplasma gondii Protein Disulfide Isomerase (TgPDI) Is a Novel Vaccine Candidate against Toxoplasmosis', *PLoS ONE*, 8(8), pp. 1–9. doi: 10.1371/journal.pone.0070884.
- Yang, J. *et al.* (2015) 'The I-TASSER Suite : protein structure and function prediction', *Nature Publishing Group*. *Nature Publishing Group*, 12(1), pp. 7–8. doi: 10.1038/nmeth.3213.
- Yang, S. *et al.* (2014) 'Compact conformations of human protein disulfide isomerase.', *PloS one*, 9(8), p. e103472. doi: 10.1371/journal.pone.0103472.
- Zhang, H. *et al.* (2011) 'Identification of the cross-reactive and species-specific antigens between *Neospora caninum* and *Toxoplasma gondii* tachyzoites by a proteomics approach', pp. 899–911. doi: 10.1007/s00436-011-2332-5.
- Zhu, C. *et al.* (2014) 'Molecular Characterization and Expression Profiling of the Protein Disulfide Isomerase Gene Family in *Brachypodium distachyon* L

.', 9(4). doi: 10.1371/journal.pone.0094704.

Datos suplementarios

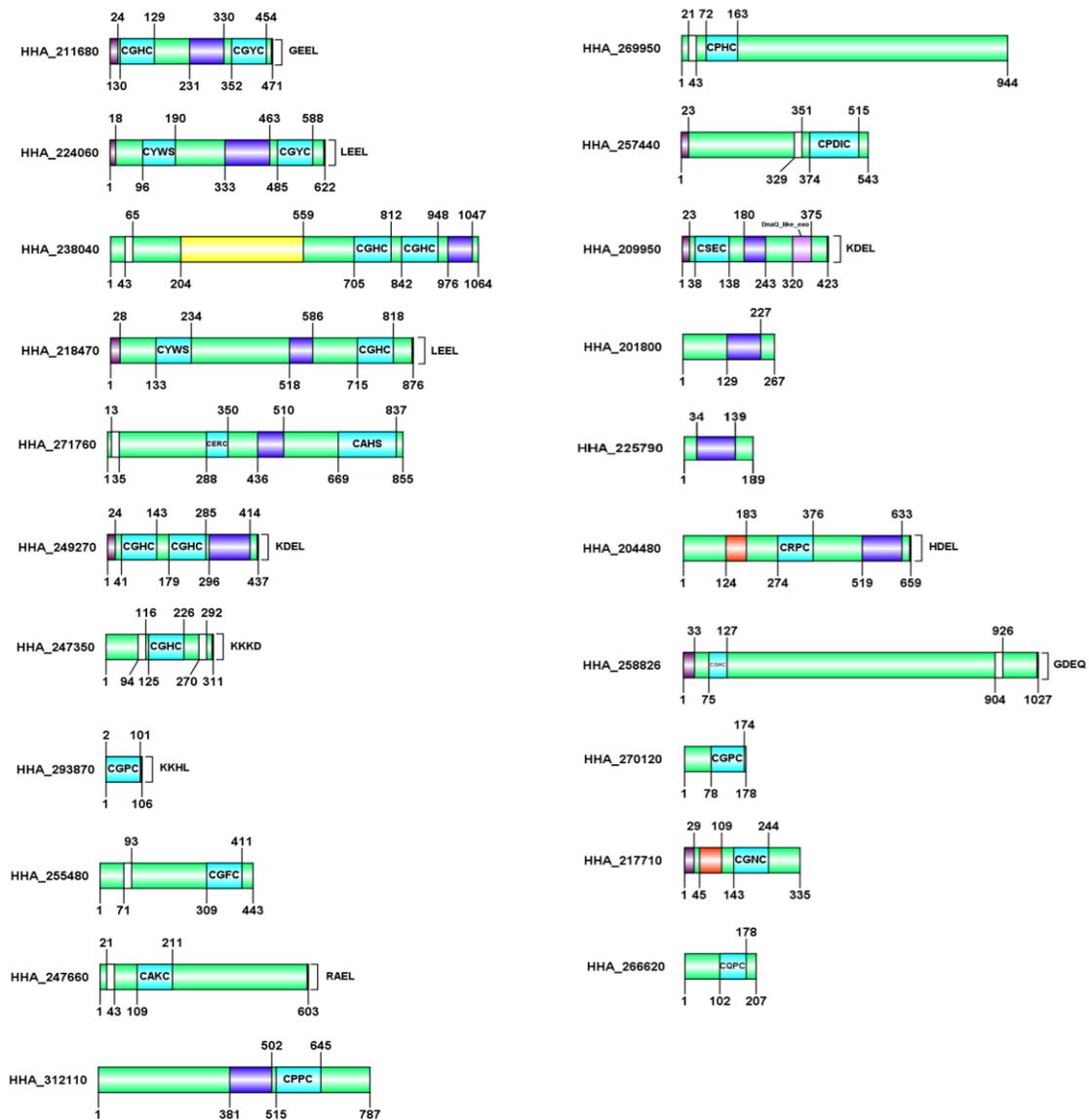


Figura S1a. Representación esquemática de los dominios y arquitecturas de las proteínas disulfuro isomerasa de *Hammondia hammondi*. Se muestran los números de acceso en la base de datos. En cian se destacan los dominios activos a y en purpura los dominios inactivos b, en amarillo el péptido señal y en azul los dominios transmembranales. Además, se muestran las proteínas con dominios de retención al retículo KDEL.

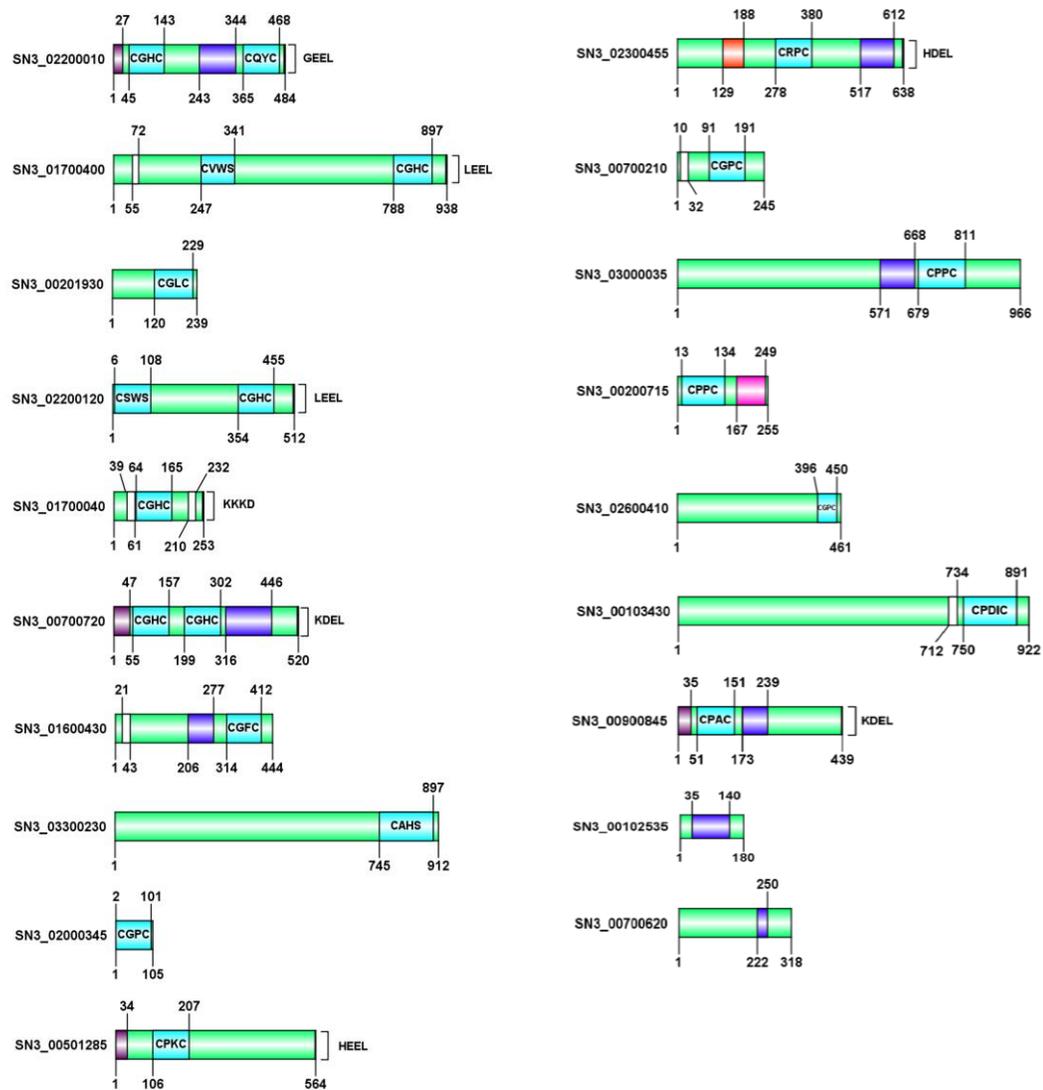


Figura S1b. Representación esquemática de los dominios y arquitecturas de las proteínas disulfuro isomerasa de *Sarcocystis neurona*. Se muestran los números de acceso en la base de datos. En cian se destacan los dominios activos a y en púrpura los dominios inactivos b, en amarillo el péptido señal y en azul los dominios transmembranales. Además, se muestran las proteínas con dominios de retención al retículo KDEL.

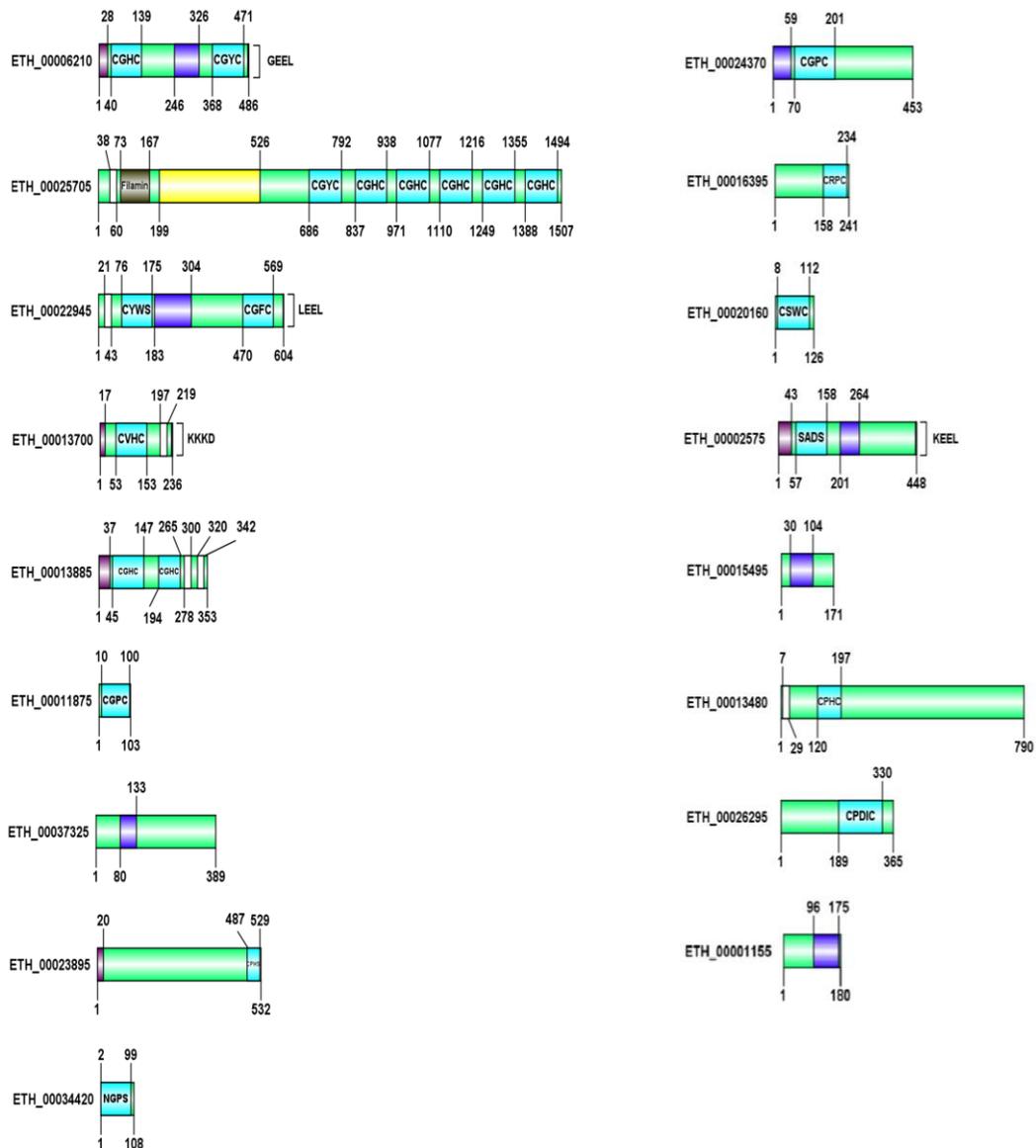


Figura S1c. Representación esquemática de los dominios y arquitecturas de las proteínas disulfuro isomerasa de *Eimeria tenella*. Se muestran los números de acceso en la base de datos. En cian se destacan los dominios activos a y en purpura los dominios inactivos b, en amarillo el péptido señal y en azul los dominios transmembranales. Además, se muestran las proteínas con dominios de retención al retículo KDEL.

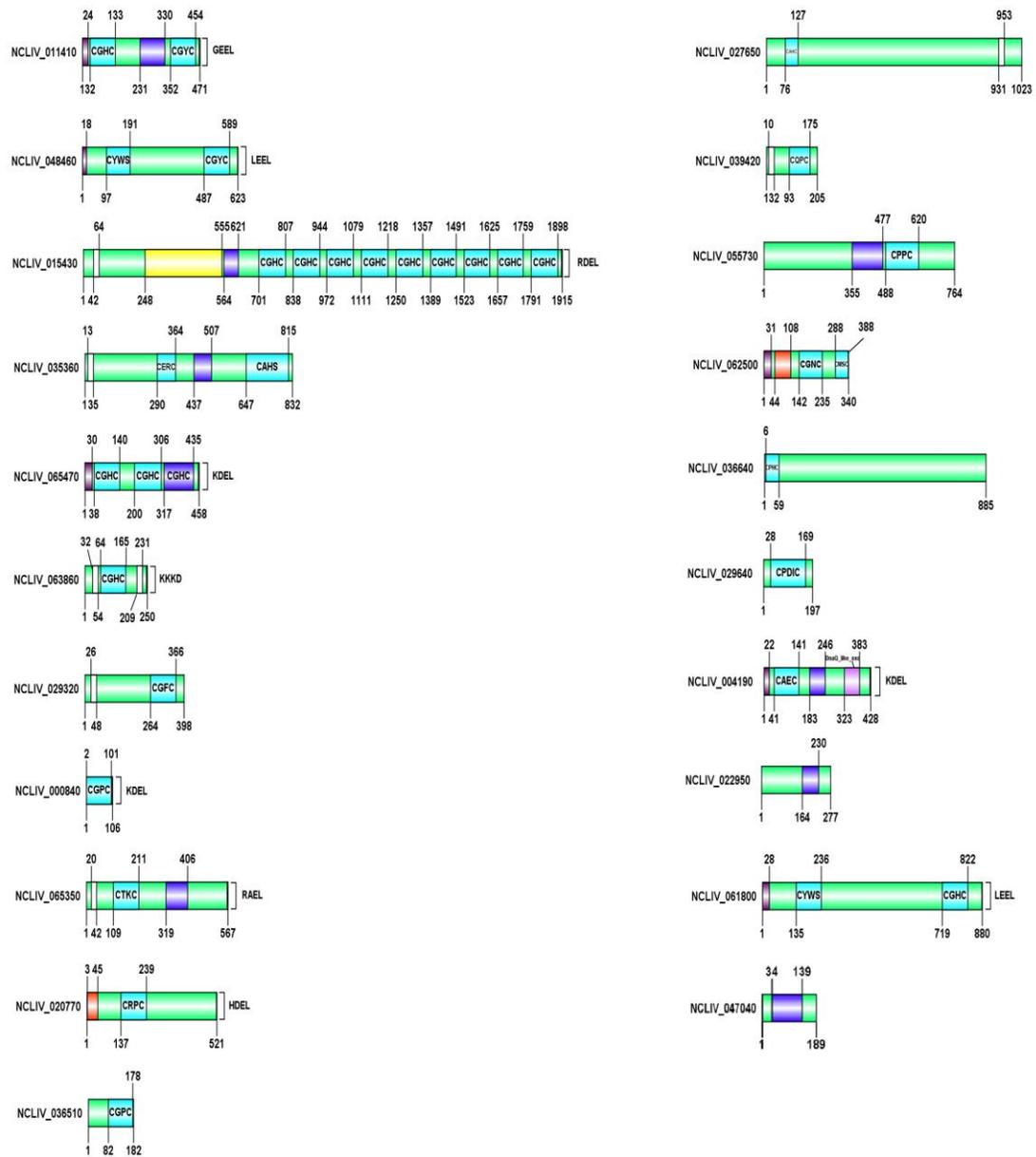


Figura S1d. Representación esquemática de los dominios y arquitecturas de las proteínas disulfuro isomerasa de *Neospora caninum*. Se muestran los números de acceso en la base de datos. En cian se destacan los dominios activos a y en púrpura los dominios inactivos b, en amarillo el péptido señal y en azul los dominios transmembranales. Además, se muestran las proteínas con dominios de retención al retículo KDEL.

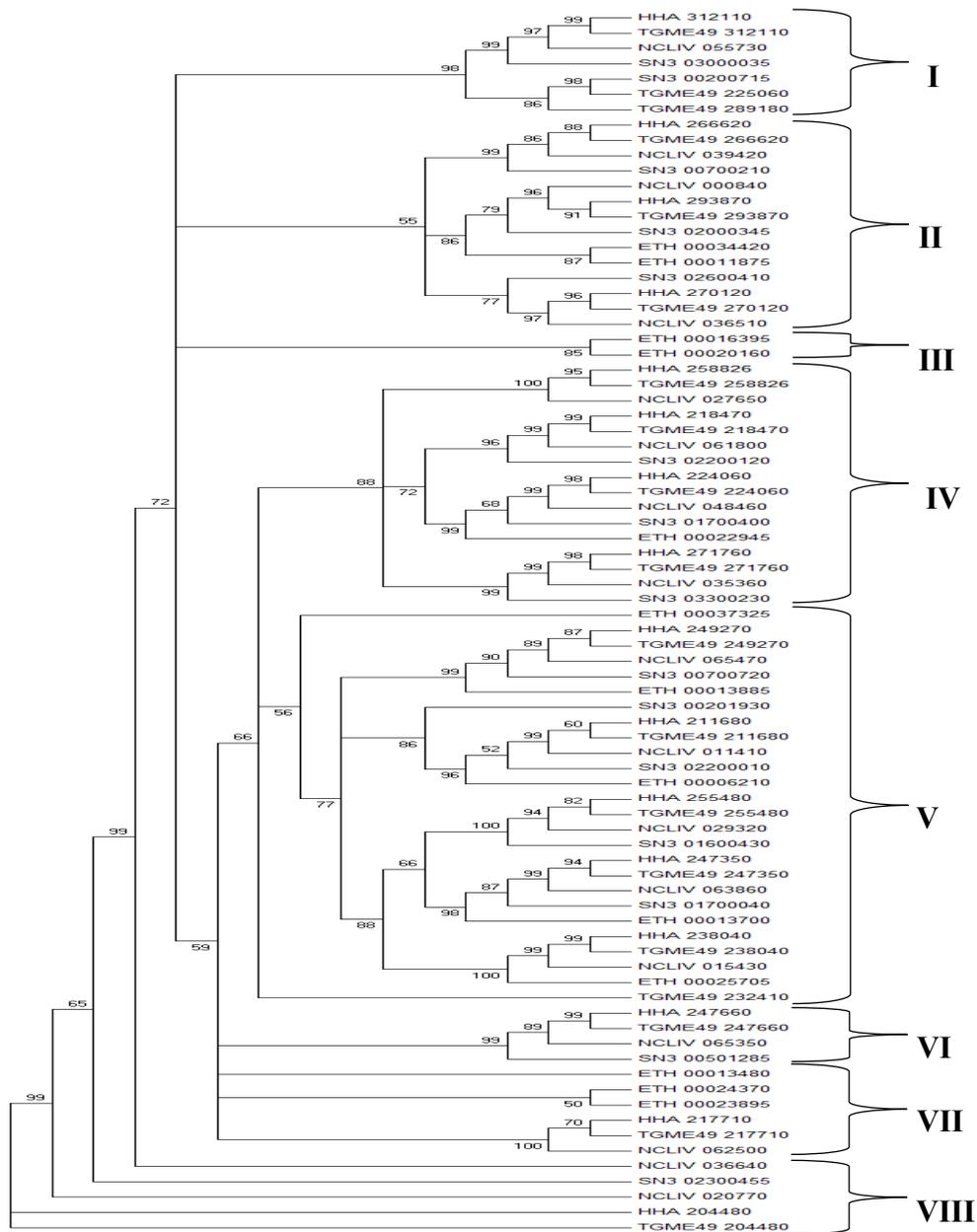


Figura S2a. Árbol filogenético de la familia de proteínas disulfuro isomerasa de los parásitos coccidios *T. gondii*, *N. caninum*, *H. hammondi*, *S. neurona* and *E. tenella*. muestra las agrupaciones de las proteínas PDI según su especie y código de EUPATHDB

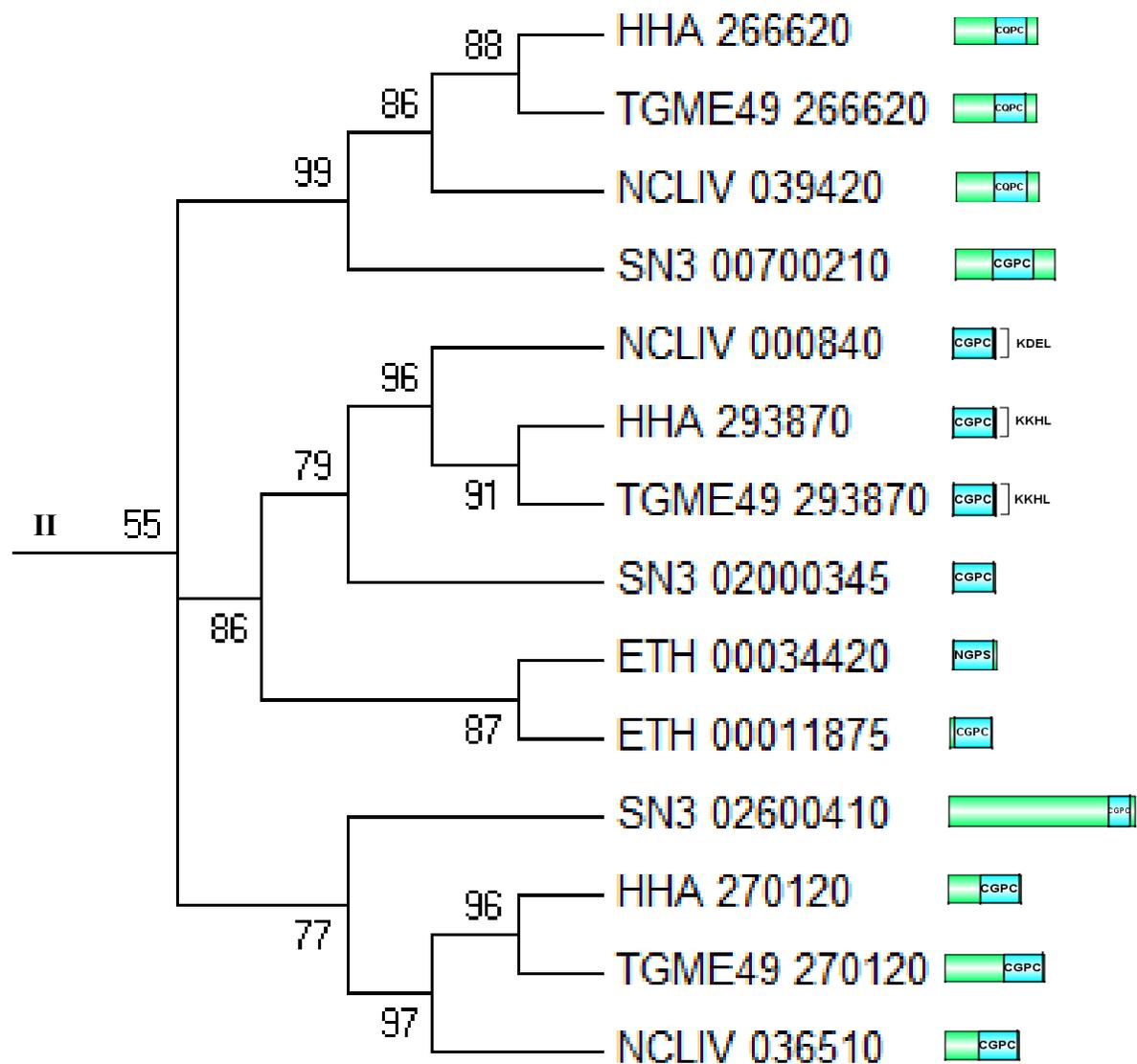
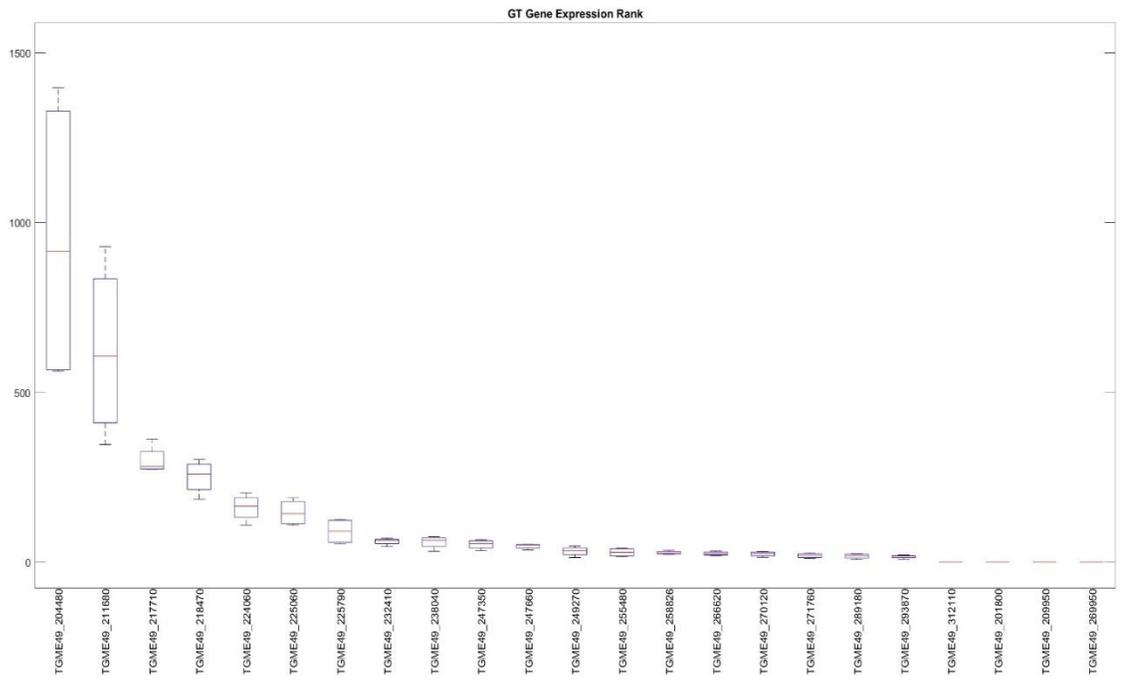
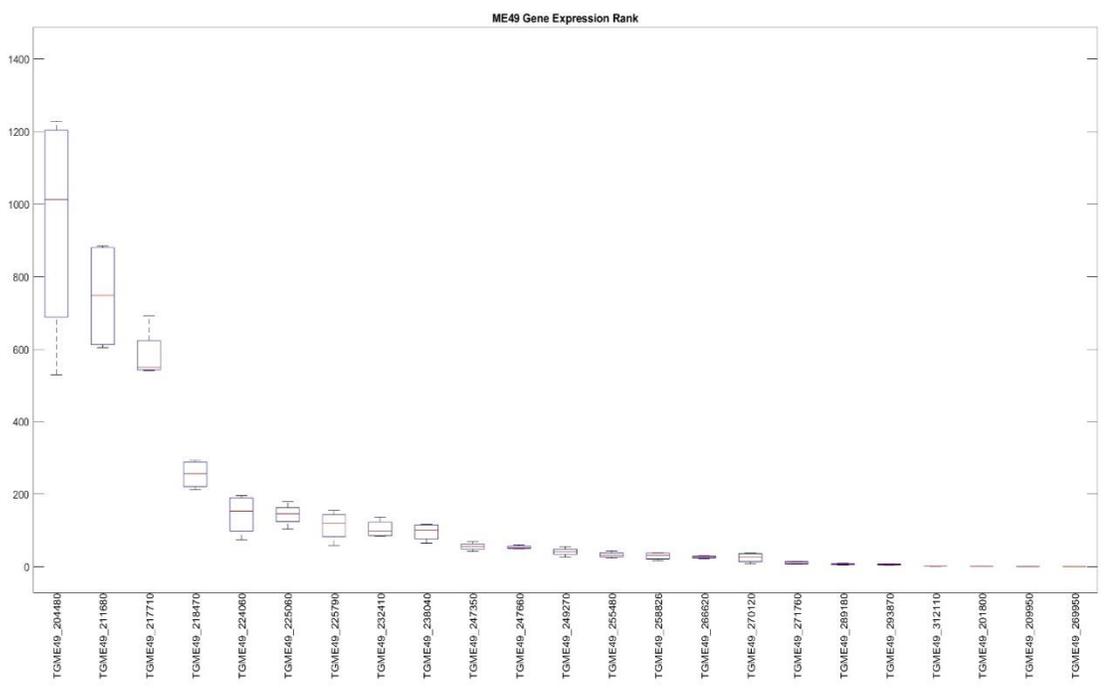


Figura S2b. Árbol filogenético subgrupo II de la familia de proteínas disulfuro isomerasa de los parásitos coccidios *T. gondii*, *N. caninum* *H. hammondi*, *S. neurona* and *E. tenella*.: los grupos formados por las proteínas de los coccidios analizados y las arquitecturas de dominios entre las agrupaciones

A-



B-



C-

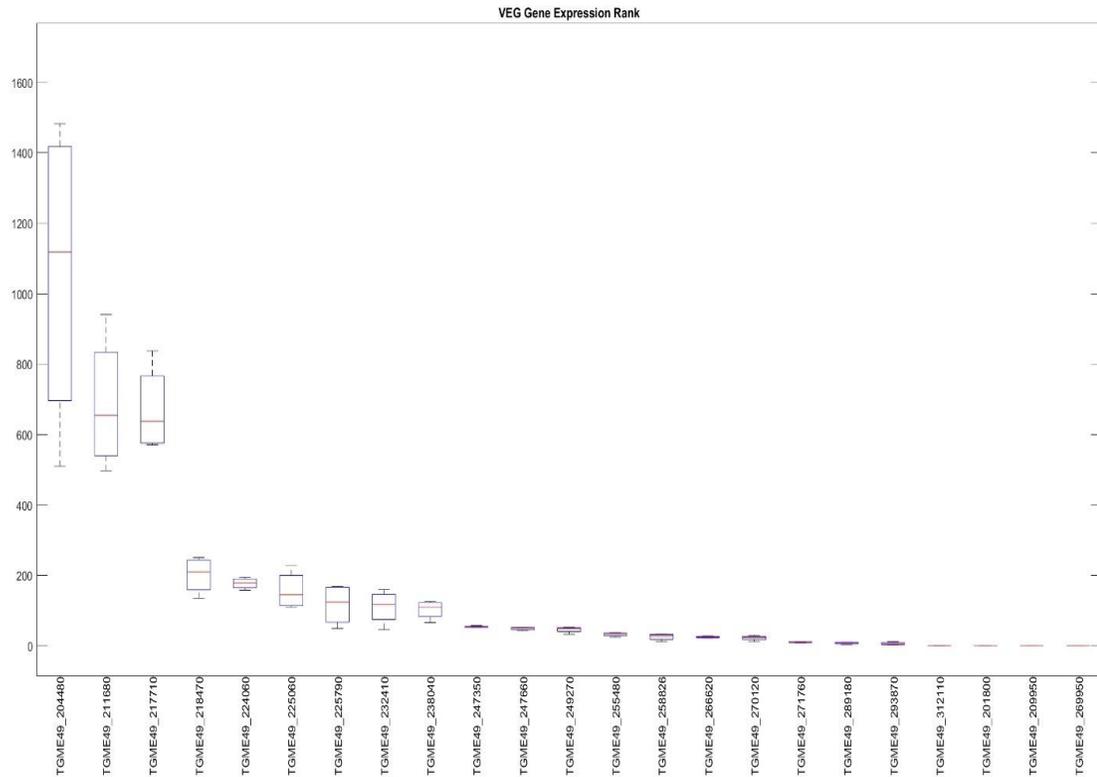


Figura S3 Grafico de barras para la expresión de la familia de proteínas PDI de *T gondii*. **A-** muestra la expresión de cada proteína identificada para la cepa GT1, **B-** muestra la expresión de cada proteína identificada para la cepa ME49, **C-** muestra la expresión de cada proteína identificada para la cepa VEG. En el eje X los códigos de identificación de la proteína y en el eje Y FPKM.

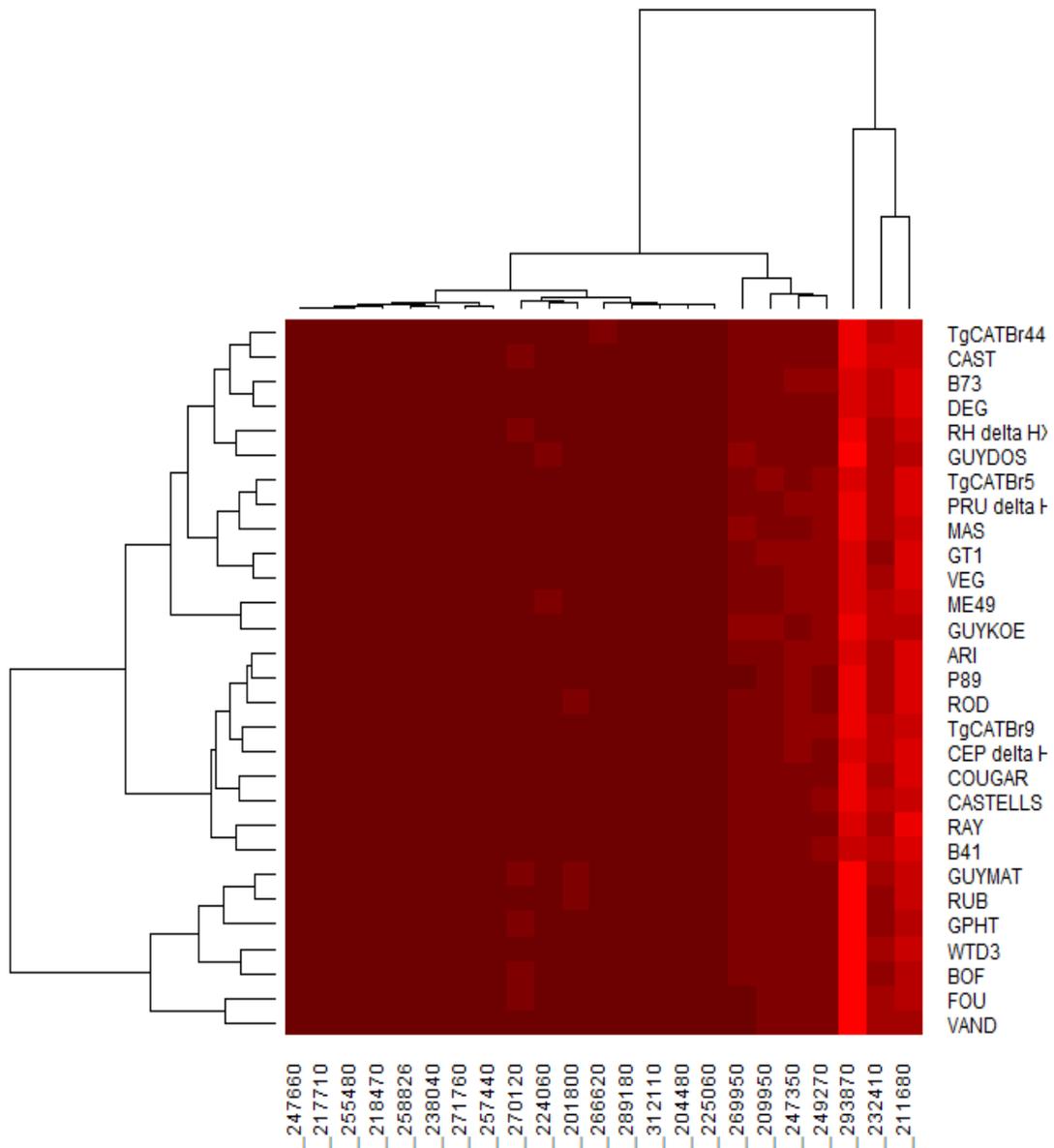


Figura S4. Mapas de calor de expresión de la familia de proteínas disulfuro isomerasa de *T. gondii* desde los transcriptomas de 26 cepas durante infección a células macrófagos de médula ósea murina. Mostrando como las proteínas de mayor expresión a TGME49_293870, 211680 y 232410.

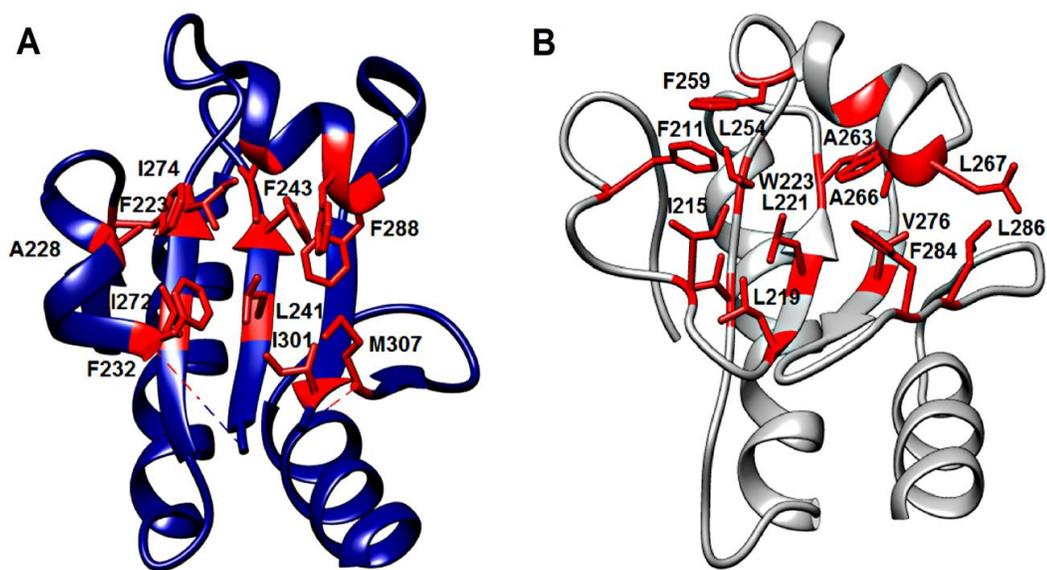


Figura S5. Comparación entre bolsillos hidrofóbicos del dominio b' de PDIA1 y PDI TGME49_211680. **A.** Bolsillo hidrofóbico del dominio b' de la proteína disulfuro isomerasa humana PDIA1. **B.** Bolsillo hidrofóbico del dominio b' de PDI TGME49_211680. En rojo se destacan los aminoácidos altamente hidrofóbicos que componen cada uno de los bolsillos de las proteínas.

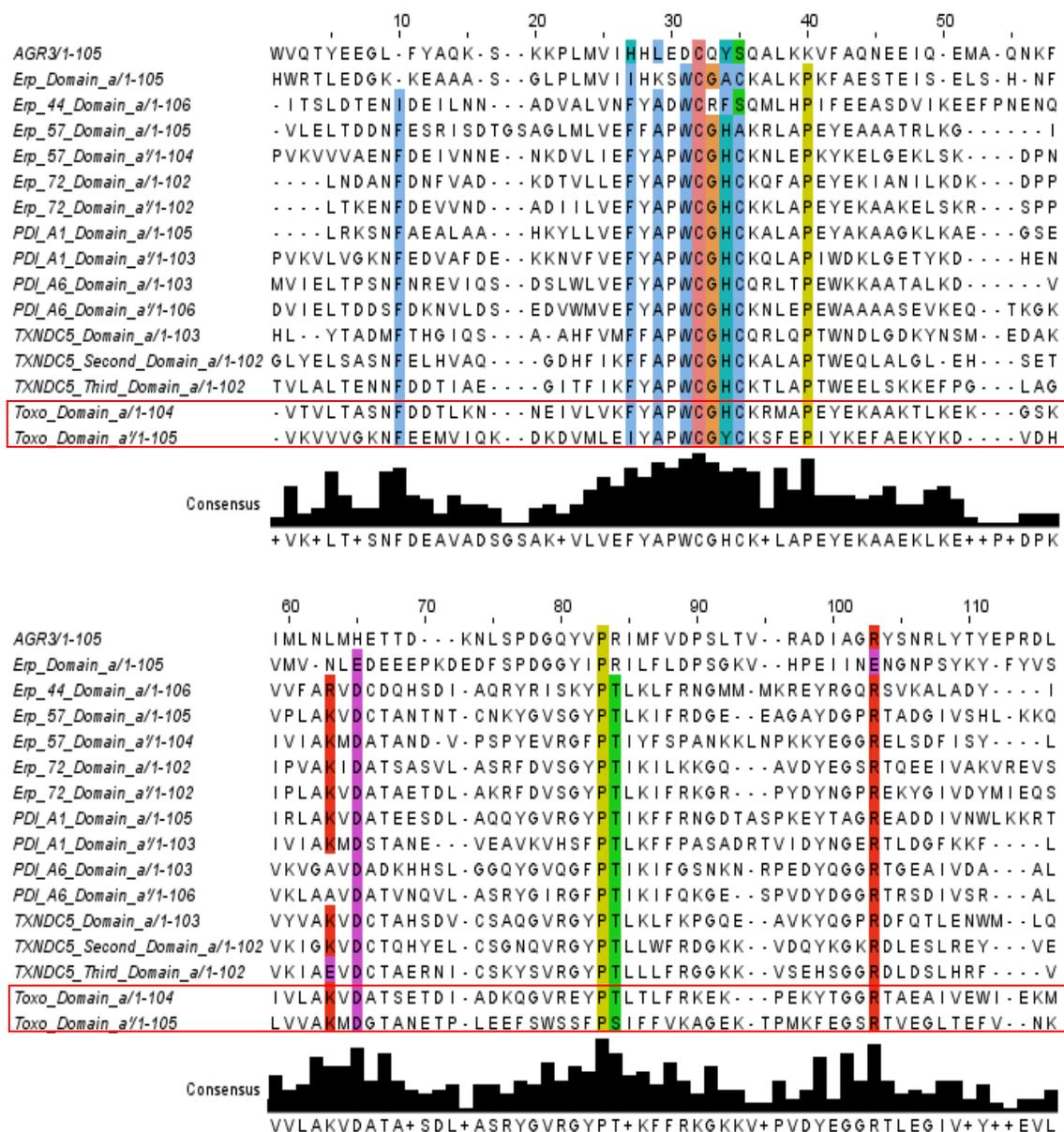


Figura S6. Alineamiento múltiple de los dominios activos de proteínas disulfuro isomerasa humanas y PDI TGME49_211680. En recuadros de color rojo se destacan los dominios activos de la proteína PDI TGME49_211680 y los aminoácidos conservados con funciones descritas en la familia PDI humana son resultados de acuerdo con el código de colores Clustal X.

Artículos y resúmenes en congresos derivados de esta tesis:

Moncada D, Arenas A, Acosta A, Molina D, Hernández A, Cardona N, Gomez Yepes M, Gomez-Marín JE. Role of the 52 KDa thioredoxin protein disulfide isomerase of *Toxoplasma gondii* during infection to human cells. Exp Parasitol 2016; 164: 36-42. Doi: 10.1016/j.exppara.2016.02.005

Presentaciones en congresos:

Moncada D, Gómez J, Arenas A, Cardona N, González K, Murillo M, Vargas M, Arenas J. Identificación de la familia de proteínas disulfuro isomerasa de *Toxoplasma gondii*. Resúmenes X Encuentro Nacional de Investigación en Enfermedades Infecciosas. Medellín 17 al 19 de Noviembre de 2016.

González K, **Moncada D**, Arenas A, Arenas J, Villegas, Gómez J, S. Identificación *in silico* de moléculas inhibidoras de la proteína disulfuro isomerasa TGME49_211680 de *Toxoplasma gondii*. Resúmenes X Encuentro Nacional de Investigación en Enfermedades Infecciosas. Medellín 17 al 19 de Noviembre de 2016.

Este trabajo recibió el Primer Premio para resúmenes en investigación básica.



X ENCUESTRO NACIONAL
DE INVESTIGACION
EN ENFERMEDADES INFECCIOSAS

17 al 19 de noviembre de 2016
Hotel Intercontinental
Medellín

La Asociación Colombiana de Infectología ACIN

En el marco del X Encuentro Nacional de Investigación en Enfermedades Infecciosas
otorga el primer puesto para resúmenes en investigación básica a:

**Identificación in silico de moléculas inhibidoras de la proteína disulfuro isomerasa
TGME49_211680 de *Toxoplasma gondii***

Maria Angelica Maya R.
MARIA ANGELICA MAYA R.
Presidente X Encuentro

Pío López
PÍO LÓPEZ
Presidente ACIN

Medellín, 19 de Noviembre 2016