



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

Análisis estructural de la Formina C de *Trypanosoma cruzi*

TESIS
QUE PARA OPTAR POR EL GRADO DE
Maestro en Ciencias

PRESENTA:
RICARDO ALEXIS MORA RINCÓN

TUTOR PRINCIPAL
DRA. ANA MARÍA CEVALLOS GAOS
Instituto de Investigaciones Biomédicas

MIEMBROS DEL COMITÉ TUTOR
Dra. Maria Soledad Funes
Instituto de Fisiología Celular

Dr. Diego González Halphen
Instituto de Fisiología Celular

Ciudad de México. Junio, 2023



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

RECONOCIMIENTOS

Al Posgrado en Ciencias Bioquímicas de la Universidad Nacional Autónoma de México (UNAM).

Al Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT), por el donativo IN203921 Estudio celular y molecular de la Formina C de *Trypanosoma cruzi* otorgado a la Dra. Ana María Cevallos Gaos, que permitió financiar el presente proyecto.

Al Programa de Apoyo a los Estudios del Posgrado (PAEP), por el financiamiento otorgado para asistir a congresos de corte nacional e internacional

A la Coordinación General de Estudios de Posgrado (CGEP) de la UNAM por la beca otorgada para la realización de la maestría.

Al PAEP (Programa de Apoyo a los Estudios del Posgrado) por el apoyo económico para realizar el curso "Beginning Programming in Python" de la Universidad de Stanford.

Esta tesis se realizó bajo la dirección de la Dra. Ana María Cevallos Gaos en el Laboratorio de Biología Celular de *Trypanosoma cruzi*, a cargo del Dr. Roberto Hernández Fernández del Departamento de Biología Molecular y Biotecnología en el Instituto de Investigaciones Biomédicas de la UNAM

Al Dr. Martin González Andrade por su apoyo en los análisis de predicciones *in silico*, dinámica molecular, acoplamiento molecular, y por permitirme usar la supercomputadora barracuda del Laboratorio de Biosensores y Modelaje Molecular de la Facultad de Medicina de la UNAM.

A mi comité tutor, la Dra. Maria Soledad Funes y el Dr. Diego González Halphen, por sus comentarios y aportes durante el desarrollo de este proyecto.

A los integrantes del jurado: Dra. María Alicia González Manjarrez, Dr. Abraham Landa Piedra, Dr. Roberto Ortega Coria, Dra. Laura Domínguez Dueñas y Dra. Nuria Victoria Sánchez Puig por todos sus comentarios y correcciones las cuales me ayudaron a enriquecer esta tesis.

La identificación inicial de los dominios funcionales presentes en las tres Forminas de *T. cruzi* fueron realizados por el Dr. Juan Felipe Osorio Méndez durante su trabajo de doctorado gracias al apoyo financiero del proyecto CONACyT No. 239715 titulado "Caracterización de la expresión de profilina de *Trypanosoma cruzi* e identificación de sus ligandos".

AGRADECIMIENTOS

Gracias a María Paula, a mi madre biológica, María Isabel y a mi hermano Wilberg por ser mi soporte, incluso en la distancia durante esta etapa de mi vida.

DEDICATORIA

Para la Dra. Ana María Cevallos Gaos por brindarme la oportunidad de trabajar con ella durante mi estancia de pregrado y mi maestría. Gracias Dra. Ana María por su apoyo incondicional, su amistad, su paciencia y sobre todo por el conocimiento personal y profesional que me ha dado durante estos años de trabajo juntos.

TABLA DE CONTENIDO

1. INTRODUCCIÓN	15
1.1 CITOESQUELETO EUCARIONTE	16
1.2 ACTINA	17
1.2.1 Generalidades de actina	17
1.2.2 La polimerización in vitro de las actinas	18
1.2.3 Proteínas de unión a actina monomérica	20
1.2.4 Proteínas de corte	20
1.2.5 Nucleadores y elongadores de actina	21
1.2.6 Proteínas de unión a microfilamentos	22
1.3 FORMINAS	22
1.3.1 Generalidades de las Forminas	22
1.3.2 Estructura del dominio FH1	23
1.3.3 Estructura del dominio FH2	24
1.3.4 Otros dominios presentes en las Forminas	25
1.3.4 Forminas diáfanas	27
1.3.5 Funciones celulares de las Forminas	28
1.3.5 Ubicación y estructura 3D de las Forminas	28
1.3.2 Identificación de un inhibidor del dominio FH2	29
1.4 Modelaje de la estructura de proteínas	29
1.4.1 Modelado por homología	30
1.4.1.1 Modelado por homología en MODELLER	31
1.4.2 Modelado <i>ab-initio</i>	33
1.4.2.1 Modelado <i>ab-initio</i> en Rosetta	33
1.4.3 Dinámica molecular	35
1.4.3.1 Parámetros importantes en la dinámica molecular	35
1.5 <i>Trypanosoma cruzi</i> y la enfermedad de Chagas	36
1.5.1 Ciclo de vida de <i>Trypanosoma cruzi</i>	37
1.5.3 Citoesqueleto de <i>Trypanosoma cruzi</i>	39
3. HIPOTESIS	43
4. OBJETIVOS	44
4.1 Objetivo general	44
4.2 Objetivos particulares	44

5. METODOLOGIA.....	45
5.1 Búsqueda de secuencias de las Forminas A, B y C en <i>Trypanosoma cruzi</i> y otras especies de tripanosomátidos	45
5.2 Identificación de los dominios proteicos en las Forminas A, B y C	45
5.3 Elaboración del árbol filogenético de las Forminas A, B y C	46
5.4 Búsqueda de motivos adicionales en la Formina C de <i>Trypanosoma cruzi</i>	46
5.5 Análisis estructura primaria de la Formina C de <i>Trypanosoma cruzi</i>	46
5.6 Predicción de la estructura secundaria de la Formina C de <i>Trypanosoma cruzi</i>	47
5.7 Predicción de la estructura terciaria	47
5.7.1 Modelado por homología en MODELLER.....	47
5.7.2 Modelado ab-initio por Rosetta	48
5.8 Verificación del modelo de la Formina C de <i>Trypanosoma cruzi</i> obtenido por medio de la predicción <i>in silico</i>	49
5.9 Acoplamiento y dinámica molecular.....	49
5.9.1 Dinámica molecular para la preparación de la macromolécula.....	50
5.9.2 Acoplamiento molecular	51
5.9.2.1 Obtención del ligando	51
5.9.2.2 Avogadro y preparación del ligando	51
5.9.2.3 AutoDock4.....	52
5.9.3 Dinámica molecular.....	52
6. RESULTADOS	54
6.1 Análisis de las secuencias de las Forminas A, B y C codificadas en <i>Trypanosoma cruzi</i> y en otras especies de tripanosomátidos	54
6.2 Elaboración del árbol filogenético de las Forminas A, B y C de <i>Trypanosoma cruzi</i> y otros tripanosomátidos	56
6.3 Selección de la Formina de <i>Trypanosoma cruzi</i> para estudios posteriores	58
6.4 Búsqueda de motivos adicionales en la Formina C.....	58
6.5 Análisis de la estructura primaria de la Formina C.....	59
6.6 Análisis de la estructura secundaria de la Formina C.....	60
6.6 Predicción de la estructura terciaria de los dominios FH2 y GBD_FH3 la Formina C por homología	62
6.7 Modelaje <i>ab-initio</i> de la estructura terciaria de la región rica en prolinas del dominio FH1 por Rosetta	67
6.8 Modelado de la Formina C por MODELLER.....	68
6.9 Verificación de la estructura terciaria predicha para la Formina C de <i>Trypanosoma cruzi</i>	69

6.10 Acoplamiento y dinámica molecular del inhibidor SMIFH2 con la Formina C.....	71
6.11 Acoplamiento y dinámica molecular del inhibidor SMIFH2 con el dominio FH3_FH1	74
6.12 Acoplamiento y dinámica molecular del inhibidor SMIFH2 con el dominio FH2.....	75
7. CONCLUSIONES	78
8. PERSPECTIVAS	80
9. REFERENCIAS	81
10. ANEXOS	101

LISTA DE FIGURAS

- Figura 1 Estructura de la actina. En la figura se observa el dominio mayor y el dominio menor separados por la hendidura hidrofóbica y la hendidura de unión a nucleótido. Los cuatro subdominios se indican con diferentes colores: Subdominio 1 (rojo), Subdominio 2 (verde), Subdominio 3 (azul) y Subdominio 4 (amarillo). Tomado y modificado de Vizcaíno-Castillo et al., 2020..... 18
- Figura 2 Dominios de las Forminas reportadas en la literatura. Abreviaturas de dominio: GBD, dominio de unión a GTPasa; FH1, dominio de homología de Formina 1; FH2, dominio de homología 2 de Formina; FH3, dominio de homología 3 de Formina; DAD, dominio diáfano-autorregulador; PDZ, dominio presente en PSD-95, Dlg y ZO-1/2; WH2, región de homología del síndrome de Wiskott-Aldrich región 2; ZnF, dedo de zinc; SP, péptido señal; TMD, dominio transmembrana; C1, dominio de unión de ésteres de forbol / diacilglicerol; C2, módulo de orientación de membrana dependiente de Ca²⁺; FHA, dominio asociado a forkhead. Tomado y modificado de Chalkia et al., 2008. 23
- Figura 3 Diagrama en cintas que muestra la estructura cristalizada del homodímero del dominio FH2 de la Formina DAAM1 humana. Esta estructura es típica de los dominios FH2, los cuales constituyen la unidad básica estructural y funcional de la Forminas. En el diagrama se identifican los subdominios lazo, enlazador, botón, bobina y poste. Tomado y modificado de la estructura cristalográfica depositada en el PDB 2Z6E, reportado por Yamashita et al., 2007. 25
- Figura 4 Diagrama en cintas del dominio GBD/FH3. En la región amino se encuentra el dominio GBD y en la región carboxilo se encuentra el dominio FH3, que está conformada por una repetición armadillo (ARM) de cinco hélices alfa. Tomado y modificado de la estructura cristalográfica depositada en el PDB 3DAD, reportado por Schulte et al., 2008..... 26
- Figura 5 Arquitectura de los dominios de una Formina diáfana (Drf). Los límites del dominio se muestran a escala. Tomado y modificado de Chesarone et al., 2010 & What is the role of formin in actin polymerization?, s.f. 27
- Figura 6 Diagrama de flujo con las etapas para el modelado por homología en MODELLER. Imagen tomada y modificada de Bitencourt-Ferreira et al., 2019..... 32
- Figura 7 Distribución y taxonomía de los triatomas más importantes en Latino América. S.A: América del Sur; C.A: Centro América. Tomado de Vivas et al.1|, 2021. 36
- Figura 8 Distribución mundial de casos de la Enfermedad de Chagas. Tomado de Salazar, 2016. .. 37
- Figura 9 Ciclo de vida de *Trypanosoma cruzi*. Aproximaciones morfológicas comúnmente observadas en vertebrados e invertebrados. Elaboración propia. 38
- Figura 10 Localización diferencial de actina 1 y actina 2 en epimastigotes de *Trypanosoma cruzi*. Anticuerpos específicos que reconocen actina 1 (en rojo) o actina 2 en verde muestran que hay

una localización distinta de ambas proteínas sin evidencia de colocalización. n= núcleo; k= cinetoplasto. Scale bar= 5 µm. Tomada de Vizcaino-Castillo, 2020.....	41
Figura 11. Estrategia metodológica.....	45
<i>Figura 12 Pasos para el acoplamiento y dinámica molecular de Formina C y el ligando SMIFH2. A) Pasos generales para la preparación de la macromolécula por medio de dinámica molecular. B) Pasos generales para la preparación y parametrización del ligando, y posterior interacción con el modelado global de la macromolécula. B) Dinámica molecular del complejo macromolécula (modelo de la Formina C de T. cruzi) y ligando SMIFH2.</i>	50
Figura 13 Representación gráfica de los dominios de los 6 genes que codifican para Formina A, B y C en <i>Trypanosoma cruzi</i> CL Brener identificados por InterPro y los reportes en la literatura. ...	55
Figura 14 Árbol filogenético del dominio FH2 de las Forminas A, B y C ortólogas identificadas en <i>Trypanosoma cruzi</i> CL Brener y otros tripanosomátidos. Los valores de Bootstrap fueron calculados de 1,000 muestreos. Con flechas rojas se señalan las Forminas A, B y C de <i>Trypanosoma cruzi</i> CL Brener. Árbol construido con el método de máxima verosimilitud en MEGA 11.....	57
Figura 15 Gráfica de probabilidades de presentar dominios transmembranales en la Formina C de <i>Trypanosoma cruzi</i> CL Brener generada por el servidor TMHMM2.0. Rojo: residuos en hélices; Azul: residuos internos; Rosado: residuos externos.	59
Figura 16 Representación gráfica de la predicción de la estructura secundaria de la Formina C de <i>Trypanosoma cruzi</i> generada por el servidor Raptor X. Los dominios identificados por InterPro fueron añadidos a la gráfica. Naranja, dominio GBD_FH3; gris, dominio FH1; verde, dominio FH2.	61
Figura 17 Configuración abierta del dímero del dominio FH2 de la Formina C modelado a partir de la plantilla PDB 2Z6E por MODELLER (Modelo MOD3-REGIONFH2). En el modelo obtenido se observa la posición de los subdominios lazo, enlazador, botón, bobina y poste. El enlazador que actúa generando una unión flexible entre los dímeros se representa en líneas punteadas. En la parte inferior izquierda de la figura se mencionan las características de la plantilla (molde) empleada para generar el modelo.....	65
Figura 18 Análisis estructural del dominio GBD-FH3 de la Formina C. A) Plantilla (molde) PDB 3DAD con 379 aa empleado para el modelado por homología. B) Predicción del dominio GBD-FH3 de la Formina C realizado con MODELLER (modelo MOD1-MODELLER-SWISS-MODEL) con 339 aa. C) Comparación de las estructuras tridimensionales de la plantilla PDB 3DAD (estructura en cintas azules) y el modelo MOD1-MODELLER-SWISS-MODEL de la Formina C (estructura en cintas grises). dominios de hélice alfa única (SAH) estables se encuentra en la región carboxilo del modelo MOD1-MODELLER-SWISS-MODEL de la Formina C. D) Alineamiento de la secuencia del dominio FH3. Las cajas grises superiores corresponden a los motivos alfa-hélice conocidos como H1, H2 y H3. Los residuos de color amarillo son aminoácidos hidrofóbicos característicos de la región de repetición de armadillo (ARM), los residuos en color azul son hidrófilos, los	

- residuos en color verde son los 3 aminoácidos más conservados en un motivo ARM [D307, E331y K335], los residuos en color naranja corresponden a la predicción del motivo de dimerización. 66
- Figura 19 Predicción del dominio FH1 de la Formina C. La secuencia de aminoácidos empleada para el modelado se encuentra en la parte inferior izquierda de la figura. El modelo del dominio FH1 se presenta en cinta y los residuos de prolina se presentan en lazos. 68
- Figura 20 Predicción de la estructura terciaria completa de la Formina C de *Trypanosoma cruzi* CL Brener (937 aa). Este es el modelo con menor energía predicho empleando MODELLER. En la región amino (naranja) se enumeran los 5 motivos que contiene la región armadillo seguido por una alfa hélice que comunica con el primer dominio descrito de la Formina rico en prolinas FH1, en la región carboxilo (de color verde) se encuentra la región FH2 descrita como el posible dominio de unión con las actinas 69
- Figura 21 Resultados del acoplamiento molecular entre la macromolécula Formina C predicha en este estudio y el ligando SMIFH2. A) Modelo estructural de la Formina C unida al ligando SMIFH2 con vista detallada del sitio de unión del ligando. La estructura de la Formina C se presenta en cintas, los aminoácidos que participan en la unión al ligando se presentan en bastones y el inhibidor SMIFH2 se presenta en bastones de color verde. B) Ligando SMIFH2 unido a los 10 ns a la Formina C. C) Ligando SMIFH 2 unido a los 30 ns a la Formina C. D) Ligando SMIFH unido a los 50 ns a Formina C. Para las imágenes B, C y D, la estructura de la Formina C se presenta en cintas, los aminoácidos que participan en la unión al ligando se presentan en bastones y el inhibidor SMIFH2 se presenta en color celeste. 72
- Figura 22 Validación del acoplamiento molecular. A) RMSD de la interacción dominio FH2 y ligando SMIFH2; B) RMSD de la interacción dominio FH3-FH1 y ligando SMIFH2; C) RMSD de la interacción Formina C y ligando SMIFH2; D) RMSF de la interacción dominio FH2 y ligando SMIFH2; E) RMSF de la interacción dominio FH3-FH1 y ligando SMIFH2; F) RMSF de la interacción Formina C y ligando SMIFH2. 73
- Figura 23 Resultados de la interacción de los residuos de la Formina C con el ligando SMIFH2. En el eje de las Y se observan los residuos de aminoácidos de la macromolécula que interacción con el ligando a medida que pasa el tiempo hasta 4500 Ps. 74
- Figura 24 Esquema del sitio de unión del ligando SMIFH2 en el dominio FH2 de la Formina C. La estructura dl dominio FH2 se presenta en cintas y el ligando se presenta en bastones. 75
- Figura 25 Resultados de la dinámica molecular del dominio FH2 de la Formina C y el ligando SMIFH2 durante 50 ns. 76
- Figura 26 Resultados de la interacción de los residuos del dominio FH2 de la Formina C con el ligando SMIFH2. En el eje de las Y se observan los residuos de aminoácidos de la macromolécula que interacción con el ligando a medida que pasa el tiempo hasta 4500 Ps. 77

LISTA DE TABLAS

Tabla 1 Parámetros empleados para la búsqueda de proteínas cristalizadas homólogas a la Formina C en el servidor NCBIblast.	48
Tabla 2 Genes identificados en <i>Trypanosoma cruzi</i> CL Brener que codifican para Forminas empleando la base de datos de TriTrypDB. <i>Trypanosoma cruzi</i> CL Brener haplotipo Esmeraldo (S), <i>Trypanosoma cruzi</i> CL Brener haplotipo no Esmeraldo (P).	54
Tabla 3 Composición de la Formina C de <i>Trypanosoma cruzi</i> CL Brener en términos de aminoácidos.	59
Tabla 4 Plantillas que podrían ser empleadas para el modelado de la estructura tridimensional de la Formina C. A) Plantillas obtenidas por BLAST en NCBI a partir de la secuencia TcCLB.511393.30. B) Plantillas obtenidas por HHlist en SWISS-MODEL NCBI a partir de la secuencia TcCLB.511393.30. Flecha verde, plantilla utilizada para modelar el dominio FH2; flecha roja, plantilla utilizada para modelar el amino terminal; flecha naranja, plantilla utilizada para modelar la estructura de dominio bobina espiral. Con asterisco se encuentran demarcadas las plantillas identificadas empleando ambos algoritmos.	63
Tabla 5 Verificación de la estructura terciaria de la Formina C completa predicha. 1, Resultados de la verificación por el servidor SAVES v.6.0; 2, Resultados de la verificación por el servidor Rosetta Desing.....	71

TABLA DE ABREVIATURAS

ABREVIATURA	DEFINICIÓN
Å	Ángstrom
aa	Aminoácido
ADN	Ácido desoxirribonucleico
ADP	Adenosín difosfato
ATP	Adenosín trifosfato
AMBER	Assisted Model Building with Energy Refinement
BLAST	Herramienta básica de búsqueda de alineación local
ExPASy	Sistema de análisis de proteínas experto, del inglés Expert Protein Analysis System
FH	Dominio con homología a Forminas, del inglés Formin Homology domain
GBD	Dominio de unión a GTPasa
GTPasa	Enzima de la superfamilia guanosina trifosfatasa
kcal	Kilocaloría
kDa	Kilodalton
ns	Nanosegundos
PDB	Protein Data Bank
ARN	Ácido ribonucleico
RMSD	Distancia media cuadrática mínima
SMIFH2	Molécula pequeña inhibidora de FH2

RESUMEN

Trypanosoma cruzi es un eucariota parasítico causante de la enfermedad de Chagas. En el análisis del genoma de este parásito se detectó una amplia gama de proteínas identificadas como pertenecientes a la familia de las actinas (4 actinas, 5 proteínas similares a actina y 3 proteínas relacionadas a actina), proteínas motoras (10 miosinas), y proteínas de unión a actina (como profilina, el complejo ARP2/3, y Forminas A, B y C). La presencia de un sistema tan complejo sugiere que el citoesqueleto de actina estaría participando en múltiples procesos celulares. De manera particular, esta investigación se centró en el estudio de la Formina C, debido a que análisis genómicos demuestran que esta proteína está ausente en *T. brucei* y *Leishmania* sp. lo cual nos sugiere que podríamos estar frente a una proteína que esté involucrada en la nucleación y elongación de microfilamentos en *T. cruzi*. En este trabajo se realizó una caracterización *in silico* de la Formina C de *T. cruzi* por medio la química computacional para obtener información sobre su estructura, su posible función, y su posible interacción con un inhibidor. Para esto, nuestro proyecto se dividió en 2 etapas. En la primera etapa se realizó el análisis *in silico* de la estructura primaria y secundaria de la Formina C. En la segunda etapa se realizó el análisis de la estructura terciaria y de las posibles interacciones moleculares de la Formina C con el inhibidor SMIFH2.

En la secuencia deducida de aminoácidos, en el extremo carboxilo terminal se identificó el dominio FH2 de las Forminas precedido por un dominio FH1 presente en la mayoría de las proteínas de esta familia. En el extremo amino, se identificó un dominio GBD_FH3, el cual podría ser un dominio con una región reguladora típica de Forminas diáfanas. Los dominios FH2 y FH3 de la Formina C fueron modelados utilizando MODELLER a partir de plantillas de proteínas homologas a la proteína de interés. Por otra parte, por medio de Rosetta se realizó un análisis *ab-initio* del dominio FH1, el cual es un dominio rico en prolinas. Tras modelar los dominios FH1, FH2 y FH3 por separado, se integraron los fragmentos en una sola estructura y se realizó una predicción computacional completa de la Formina C utilizando MODELLER. Seguido, se realizó una dinámica molecular a la estructura de Formina C obtenida por medio de SAVES v.6.0, lo que nos permitió verificar que la predicción realizada tiene una calidad estereoquímica aceptable, por lo que el modelo obtenido es aceptable. Finalmente, se realizó un análisis de Acoplamiento Molecular con la estructura predicha de la Formina C y el ligando SMIFH2, lo cual nos permitió sugerir que el ligando SMIFH2 se une a la Formina C a través del dominio FH2.

ABSTRACT

Trypanosoma cruzi is a parasitic eukaryote that causes Chagas disease. In the analysis of the genome of this parasite, a wide range of proteins identified as actins (4 actins, 5 actin-like proteins and 3 actin-related proteins), motor proteins (10 myosins), and actin-binding proteins (profilin, the ARP2/3 complex, and Formins A, B, and C) were detected. The presence of such a complex system suggests that the actin cytoskeleton participates in multiple cellular processes. This research focused on the study of Formin C, since genomic analyzes show that this protein is absent in *T. brucei* and *Leishmania* sp. which suggests that we could be dealing with a protein that is involved in the nucleation and elongation of microfilaments in *T. cruzi*. Therefore, this research is based on the *in-silico* characterization of *T. cruzi* Formin C by means of computational chemistry to obtain information about its structure, its possible function, and its possible interaction with an inhibitor. For this, our project was divided into 2 stages. In the first stage, the *in-silico* analysis of the primary and secondary structure of Formin C was carried out. In the second stage, the analysis of the tertiary structure and the possible molecular interactions of Formin C with the SMIFH2 inhibitor was carried out.

In the deduced amino acid sequence, the pathognomonic FH2 domain of Formins was identified at the carboxyl terminus, preceded by an FH1 domain present in most proteins of this family. At the amino terminus, a GBD_FH3 domain was identified, which could be a domain with a regulatory region typical of diaphanous Formins. The FH2 and FH3 domains of Formin C were modeled using MODELER from protein templates homologous to the protein of interest. On the other hand, using Rosetta, an ab-initio analysis of the FH1 domain was performed, which is a proline-rich domain. After modeling the FH1, FH2 and FH3 domains separately, the fragments were joined, and a full computational prediction of Formin C was performed using MODELLER. Next, molecular dynamics was performed on the Formin C structure obtained by means of SAVES v.6.0, which allowed us to verify that the prediction made has an acceptable stereochemical quality, so that the model obtained is acceptable. Finally, a Molecular Docking analysis was performed with the predicted structure of form C and the SMIFH2 ligand, which allowed us to suggest that the SMIFH2 ligand binds to form C through the FH2 domain.

1. INTRODUCCIÓN

Los eucariotas poseen un citoesqueleto, que no solo cumple una función estructural, sino que también participa en múltiples procesos celulares como la endocitosis, la motilidad celular, la división celular, el transporte de vesículas y el movimiento de organelos dentro de la célula (Fletcher *et al.*, 2010). En la mayoría de los eucariontes, el citoesqueleto depende de tres tipos de filamentos: 1) Los filamentos de actina, 2) los microtúbulos y 3) los filamentos intermedios. Todas estas estructuras se coordinan y complementan para dar forma a la célula y, para controlar actividades como el transporte de vesículas y organelos al interior de la célula (Pegoraro *et al.*, 2017).

Dentro de los eucariotas encontramos a los protozoos *Leishmania* sp., *Trypanosoma brucei* y *Trypanosoma cruzi*, causantes respectivamente de leishmaniasis, la enfermedad del sueño y la enfermedad de Chagas (también conocida como tripanosomiasis americana), quienes pertenecen al orden Trypanosomatida y la clase Kinetoplastea (Stuart, 2008).

Estudios sobre la ultraestructura de los tripanosomátidos indican que estos protozoos flagelados altamente polimórficos poseen un citoesqueleto poco común, pues la mayor parte está constituida por microtúbulos y no se ha podido detectar la presencia de los filamentos de actina en su citoesqueleto. No obstante, si se ha reportado que todos los genomas de tripanosomátidos liberados a la fecha contienen los genes que codifican para estos filamentos, y que la única función que se les atribuye es su participación en el proceso endocitosis y tránsito vesicular.

Específicamente, en el análisis del genoma de *T. cruzi* se identificó la presencia de 12 loci que codifican para proteínas pertenecientes a la familia de las actinas (cuatro actinas, cinco proteínas similares a actina y tres proteínas relacionadas a actina [ARP2, ARP3 y ARP6]) (Cevallos *et al.*, 2011; Vizcaíno *et al.*, 2020). Además, se han identificado genes codificadores para proteínas reguladoras del ensamblaje y desensamble de actina como la profilina, el complejo ARP2/3, y tres Forminas

(Formina A, B y C), y también proteínas motoras (diez miosinas, nueve de las cuales son específicas de tripanosomátidos) (De Souza *et al.*, 2018; Vizcaíno Castillo, 2020). La presencia de un sistema complejo de actinas en el genoma de *T. cruzi* guía esta investigación en búsqueda del mecanismo de su regulación a través de proteínas de unión a actina como las Forminas.

En este proyecto se realizó un modelaje estructural de la Formina C y se evaluó su capacidad potencial de interactuar con el inhibidor genérico de Forminas SMIFH2, como un abordaje inicial para el estudio de su función en *T. cruzi*.

1.1 CITOESQUELETO EUKARIOTE

El citoesqueleto de los eucariotes está constituido por una red compleja de filamentos que se clasifican en microfilamentos, microtúbulos y filamentos intermedios. Cada uno de estos filamentos está constituido por polímeros de proteínas específicas, actina en microfilamentos, alfa y beta tubulina en los microtúbulos, y una familia de varias proteínas como queratina o láminas nucleares en el caso de los filamentos intermedios (Muñoz-Lasso *et al.*, 2020).

Los dos elementos más conservados a lo largo de la evolución son los microfilamentos y los microtúbulos. Su distribución subcelular y función varía entre cada tipo de célula y en diferentes momentos del ciclo celular. Por ejemplo, en las células musculares, los filamentos de actina al interactuar con miosina determinan su contracción. En células no musculares, los filamentos de actina asisten a la formación de protuberancias conocidas como lamelipodios o filopodios, que ayudan con la migración celular. Los lamelipodios utilizan mecanismos para regular su propia dinámica de desplazamiento por medio de estructuras llamadas filopodios, los cuales contribuyen a extender sus propios bordes celulares más allá de los mismos lamelipodios. Los filopodios generan un movimiento de traslación y se observa como haces de filamentos, siendo la actina el componente principal de esta organización (Faix & Rottner, 2006). Las funciones de las actinas están controladas

a través de proteínas de unión de actina (ABPs, *por sus siglas en inglés*) las cuales regulan el ensamblaje, desensamblaje y estabilización de los filamentos de actina.

1.2 ACTINA

1.2.1 Generalidades de actina

La actina es una proteína globular con un peso molecular aproximado de 42 kDa y es el componente básico de los microfilamentos. En su forma no polimerizada recibe el nombre de actina globular (Actina G). Su estructura se ha determinado mediante cristalografía de rayos X y se ha demostrado que es una molécula relativamente plana que cabría en un prisma rectangular de un tamaño de 55 Å × 55 Å × 35 Å (Domínguez *et al.*, 2011). La cadena polipeptídica de aproximadamente 375 aminoácidos se dobla de tal manera que forma dos dominios principales conocidos como dominio mayor y dominio menor debido a su tamaño diferente estimado mediante microscopía electrónica de barrido (SEM) (Domínguez *et al.*, 2004). Cada dominio se divide a su vez en dos subdominios que tradicionalmente han sido numerados como subdominios 1, 2, 3 y 4. Los dominios mayor y menor se encuentran separados por dos hendiduras, una que contiene el sitio de unión a ADP/ATP y un catión Mg²⁺ o Ca²⁺ y la otra, menos profunda, de naturaleza hidrofóbica a la que se unen muchas de las proteínas de unión a actina que regulan su función (Domínguez *et al.*, 2011) (Figura 1).

La actina polimerizada recibe el nombre de actina filamentosa (actina F). Su estructura se ha determinado mediante microscopía electrónica y se ha descrito como una hélice formada por una única hebra con subunidades que rotan de izquierda a derecha (Lee *et al.*, 2010). Las subunidades que forman el filamento se orientan de cabeza a cola y rotan aproximadamente 166° sobre el eje de la actina anterior (Cooper, 2000). Puesto que todas las subunidades de un microfilamento tienen la misma orientación dentro del mismo, se dice que el polímero presenta polaridad en su estructura. Por convención, se nombra al extremo que posee una subunidad de actina exponiendo la hendidura de unión a nucleótido como “extremo

(-)” mientras que tiene la subunidad que tiende exponiendo la hendidura hidrofóbica constituye el “extremo (+)” (Fuji *et al.*, 2010).

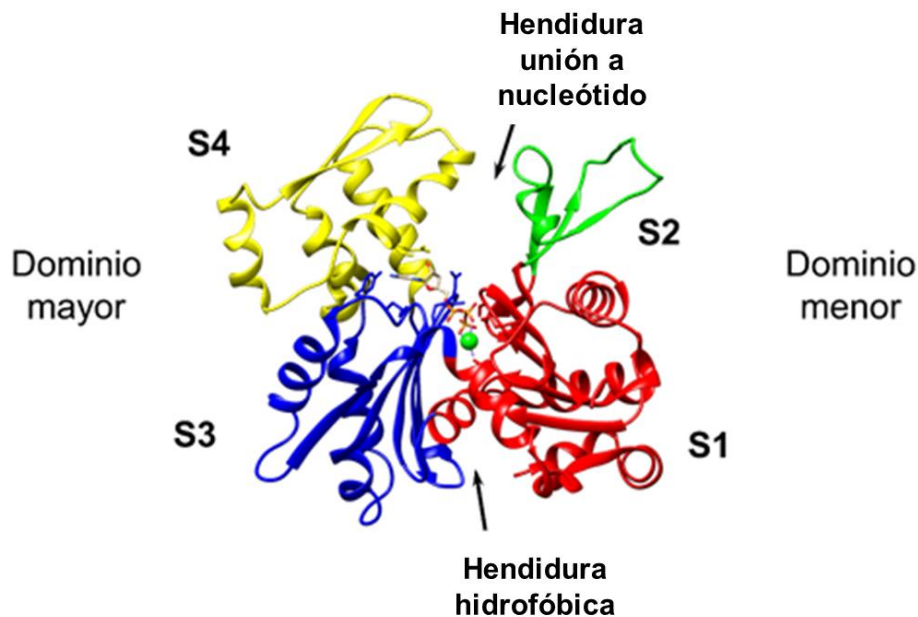


Figura 1 Estructura de la actina. En la figura se observa el dominio mayor y el dominio menor separados por la hendidura hidrofóbica y la hendidura de unión a nucleótido. Los cuatro subdominios se indican con diferentes colores: Subdominio 1 (rojo), Subdominio 2 (verde), Subdominio 3 (azul) y Subdominio 4 (amarillo). Tomado y modificado de Vizcaíno-Castillo *et al.*, 2020.

1.2.2 La polimerización *in vitro* de las actinas

La actina puede polimerizarse espontáneamente *in vitro*, siendo posible adicionar nuevos monómeros en ambos extremos del filamento. Sin embargo, la adición y despolimerización de subunidades de actina ocurre a velocidades diferentes, siendo la tasa de polimerización más alta en el extremo positivo. Aunque la actina monomérica puede polimerizar tanto si está unida a ATP como a ADP, ATP-actina es la forma principal de la proteína que interviene en la polimerización. Poco después de la adición de ATP-actina a un microfilamento, se produce la hidrólisis del ATP en ADP más un fosfato inorgánico (Pi). Debido a que el Pi tarda en ser liberado, la actina puede estar unida a ATP, ADP o ADP+Pi. Por otra parte, la subunidad de ADP-actina se separa más rápidamente del filamento que ATP-actina (Ampe *et al.*, 2017).

En condiciones *in vitro*, la polimerización se puede dividir en varias etapas. Primero ocurre la nucleación de dímeros y trímeros que sirven como núcleos para la formación de nuevos filamentos. Después hay un alargamiento del filamento, mediante la adición reversible de monómeros en ambos extremos del filamento. Finalmente, los monómeros de actina se intercambian en los extremos del microfilamento sin que varíe la longitud total del mismo (Pollard *et al.*, 2000). Este recambio de subunidades no significa un aumento en la masa de la actina filamentosa, sino que la proporción de actina filamentosa con respecto a la monomérica aumenta hasta llega a un punto de equilibrio (Luxenburg *et al.*, 2017).

La concentración de actina G en la cual la dinámica de adición y eliminación de monómeros no produce un cambio en la longitud del microfilamento se conoce como concentración crítica. La concentración crítica se define como la relación entre las constantes de ensamblaje y desensamblaje, es decir la concentración crítica se trata de una constante de disociación (Pollard *et al.*, 2017).

El proceso de polimerización y despolimerización descrito *in vitro* no refleja el comportamiento *in vivo* ya que la concentración de la actina dentro del organismo está por arriba de la concentración crítica. Por lo tanto, el ensamblaje espontáneo de monómeros de actina es desfavorable debido a la inestabilidad de los dímeros y trímeros de actina (Mullins & Hansen, 2013; Carlier & Pantaloni, 1997). Debido a esto, es necesaria la intervención de proteínas que interaccionan directa o indirectamente con actina, llamadas ABPs (los cuales en eucariotas superiores existen más de cien) (dos Remedios *et al.*, 2003). Una vez que comienza la polimerización, los filamentos crecen rápidamente. Por otra parte, las proteínas de unión a actina no afectan solo la velocidad sino la organización tridimensional de los filamentos, su polimerización y su despolarización (Lee *et al.*, 2010). La cantidad y características de las ABPs varía en cada tipo de célula y especie, y su función como proteínas efectoras en vías de señalización es distinto según la respuesta a diferentes estímulos. Para fines informativos en esta tesis se describirán

brevemente las principales proteínas de unión a actina que se han identificado en *T. cruzi*.

1.2.3 Proteínas de unión a actina monomérica

Una de las proteínas que se une a actina es la profilina, con un peso aproximado de 14 a 16 kDa, que se encuentra en casi todos los organismos eucariontes estudiados hasta el momento. La profilina se puede unir a 3 diferentes ligandos: actina, fosfatidil inositoles y regiones ricas en poli-prolinas (Lu & Pollard, 2001). La profilina lleva a la actina a sitios de polimerización activos y facilita el intercambio de ATP a ADP en el monómero (Krishnan *et al.*, 2009). La unión de profilina a fosfatidilinositol 1,4 bisfosfato (PI(4,5)P2) actúa como un inhibidor, ya que el sitio de unión se sobre lapa con la región de unión a actina y por lo tanto compite con ella (Sun *et al.*, 2013). El tercer ligando son proteínas con regiones ricas en prolinas como lo son las Forminas o las proteínas de homología Ena/VASP que promueven la nucleación y elongación de microfilamentos (Lu & Pollard, 2001).

1.2.4 Proteínas de corte

Las (ADF)/cofilina son proteínas pequeñas que se encuentran en casi todos los vertebrados, que se pueden clasificar en tres familias de proteínas pequeñas: cofilina 1, cofilina 2 y destrina [ADF, factor despolimerizante de actina]. Las (ADF)/cofilina trabajan junto a otras proteínas para reorganizar los filamentos de actina en la célula (Bernstein & Bamburg, 2010; Maciver *et al.*, 2002). Estas proteínas ayudan al desensamblaje de la actina filamentosa acelerando el paso de ATP-actina a ADP-actina en los extremos negativos de los filamentos para que la actina se disocie en forma de actina monomérica. Algunos estudios revelan que al aumentar las concentraciones de ADF/cofilina se puede promover la polimerización de las actinas a microfilamentos ya existentes (Andrianantoandro *et al.*, 2006).

Una de las principales proteínas que trabaja con las ADF/cofilina es la proteína de interacción con actina 1 (AIP1, *por sus siglas en inglés*), también conocida como proteína de repetición WDR1. Esta proteína modula la interacción entre la actina y

la cofilina mediante una comunicación física que evita la unión a la actina en ausencia del complejo (ADF/cofilina-AIP1) (Ono *et al.*, 2018). Esta proteína altamente conservada en eucariotas se describió por primera vez en 1995 en levaduras, usando un sistema indicador de interacción de doble híbrido (Amberg *et al.*, 1995). WDR1 tiene un peso de aproximado 66 kDa, una superficie cóncava y un motivo estructural de 40 aminoácidos con 2 residuos conservados y repetidos 14 veces (triptófano y ácido aspártico) al final. Por esta razón se conoce a este dominio WD40 como un dominio importante en las funciones del desensamblaje del filamento de actina (Ono S *et al.*, 2018).

La proteína asociada a adenilil ciclase (CAP, *por sus siglas en inglés*) está compuesta por 450-550aa (Ono S *et al.*, 2013) y regula tanto el citoesqueleto de actina como la vía Ras/cAMP. CAP ayuda al reciclaje de cofilina y de actina. Así mismo, mejora la unión de la cofilina-actina para promover su ruptura e interactúa en esta unión en ambos extremos del filamento de actina (Jin *et al.*, 2018).

Por otra parte, la twinfilina (TWF) contiene dos dominios ADF/cofilina conectados entre sí. La TWF tiene de aproximadamente 300 aa y los estudios bioquímicos clasifican a TWF como un factor despolimerizante del filamento de actina debido a su alta afinidad por la actina unida a ADP y su gran capacidad para inhibir la adición de subunidades en cualquier extremo del filamento (Mwangangi & Robinson, 2021). La interacción tanto de cofilina como de TWF ayuda a incrementar o mantener la concentración de monómeros de actina y generar nuevos extremos barbados para que se unan nuevas subunidades de actina, ayudando a que la célula tenga monómeros y extremos barbados siempre disponibles (Hakala *et al.*, 2018).

1.2.5 Nucleadores y elongadores de actina

Entre las proteínas que ayudan a la nucleación y elongación de los microfilamentos los más estudiados son el complejo ARP2/3 y las Forminas. Estas proteínas no solo controlan el sitio de polimerización, sino también la velocidad (Lodish *et al.*, 2000).

El complejo ARP2/3 está formado por las proteínas ARP2, ARP3 y otras cinco proteínas. Este complejo constituye el principal mecanismo de nucleación de los filamentos de actina ramificados. Por otra parte, la nucleación de nuevos filamentos de actina en el citoesqueleto está dado principalmente por las proteínas Forminas (Pollard, 2016), las cuales se analizarán a detalle más adelante.

1.2.6 Proteínas de unión a microfilamentos

La coronina es una proteína de 445 aa y una masa molecular relativa de 55 kDa que contiene varios motivos WD (como los previamente descritos para AIP1). Coronina interactúa con cofilina para ayudar a cortar los filamentos de actina (Gandhi *et al.*, 2008). En el protozoo *Leishmania* se ha visto que coronina ayuda a la remodelación de microtúbulos y participa en la citocinesis (Sahasrabudhe *et al.*, 2009). Coronina también se puede unir al complejo ARP2/3 inactivo, para activarlo y reclutarlo al unirlo a filamentos ya existentes, ayudando a la nucleación y ramificación de los filamentos de actina. Por lo tanto, la coronina tiene la capacidad única de controlar espacialmente la actividad del complejo Arp2/3 y promover selectivamente el crecimiento, tanto en la expansión de las redes existentes como en su corte (Rybakin & Clemen, 2005).

1.3 FORMINAS

1.3.1 Generalidades de las Forminas

Las Forminas son una familia de proteínas multidominio que tienen un peso aproximado de 100 kDa (Courtemanche *et al.*, 2018) y se encuentran en la mayoría de los eucariotas. Estas proteínas se describieron por primera vez en ratones, moscas y levaduras como genes que al ser mutados causaban defectos graves en la biosíntesis de membranas, la citocinesis y la estructura de la célula (Breitsprecher & Goode, 2013; Schönichen & Geyer, 2010).

El número de Forminas, el tamaño y la presencia de dominios específicos varía según la especie (Figura 2). Por ejemplo, los hongos contienen de una a tres

variantes de Forminas, mientras que las plantas y los vertebrados contienen más de 15 variantes de Forminas. El dominio que define a las Forminas es el dominio de homología a las Forminas 2 (FH2), y en la mayoría de las Forminas se identifica una región rica en prolina conocido como dominio de homología a la Formina 1 (FH1) en el extremo amino del dominio FH2.

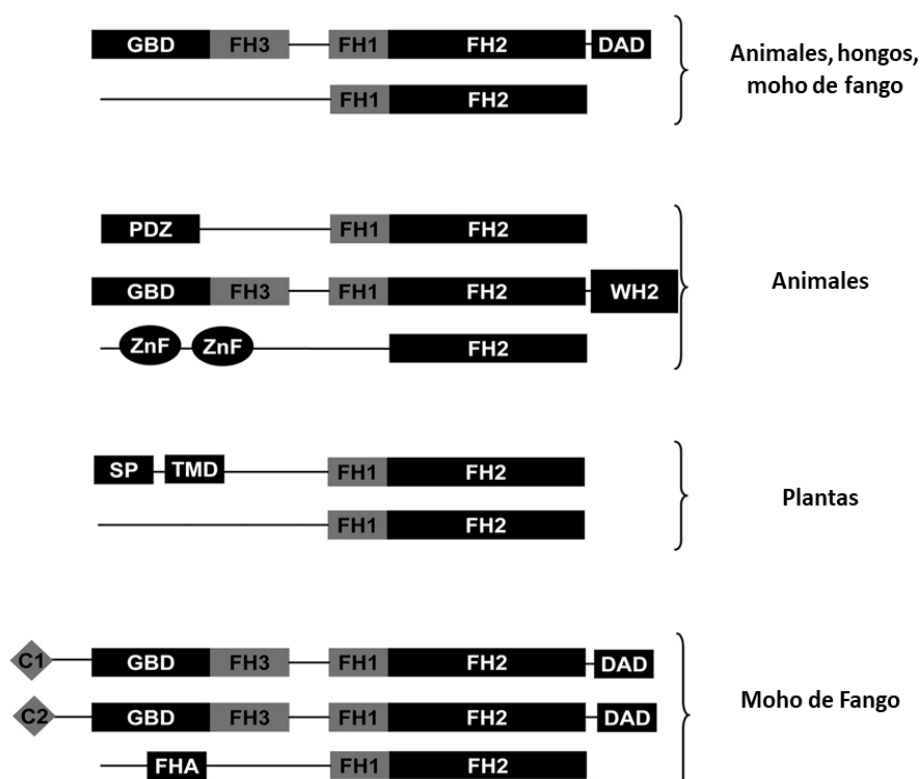


Figura 2 Dominios de las Forminas reportadas en la literatura. Abreviaturas de dominio: GBD, dominio de unión a GTPasa; FH1, dominio de homología de Formina 1; FH2, dominio de homología 2 de Formina; FH3, dominio de homología 3 de Formina; DAD, dominio diáfano-autorregulador; PDZ, dominio presente en PSD-95,Dlg y ZO-1/2; WH2, región de homología del síndrome de Wiskott-Aldrich región 2; ZnF, dedo de zinc; SP, péptido señal; TMD, dominio transmembrana; C1, dominio de unión de ésteres de forbol / diacilglicerol; C2, módulo de orientación de membrana dependiente de Ca²⁺; FHA, dominio asociado a forkhead. Tomado y modificado de Chalkia et al., 2008.

1.3.2 Estructura del dominio FH1

El dominio FH1 es una región rica en prolina que tiene la capacidad de unirse a profilina. La afinidad de profilina por regiones ricas en poli-prolina es hasta 11 veces mayor cuando la profilina está unida a actina que cuando está sola (Ferron *et al.*, 2007). Esta característica favorece la interacción de los complejos profilina-actina y

la entrega de nuevos monómeros de actina para ser incorporados en el filamento de actina. Generalmente, FH1 se encuentra localizado en el extremo amino del dominio FH2 y no posee una secuencia consenso clara (Lu *et al.*, 2007).

La mayoría de los artículos científicos han reportado que el dominio FH1 se predice como una región con múltiples tramos desordenados de prolinas que necesita estar unido al dominio FH2 para que la Formina pueda acelerar el alargamiento y nuclear a las actinas (Paul & Pollard, 2008; Breitsprecher & Goode, 2013; Szikora *et al.*, 2017). Sin embargo, otros artículos indican que la Formina es funcional aun cuando carece del dominio FH1 (Kitayama *et al.*, 2003). En los últimos años, se ha intentado estudiar la estructura secundaria que forman las poli-prolinas por medio de Dinámicas Moleculares (DM), Resonancia Magnética Nuclear (RMN) y Espectroscopia de rayos X. Los resultados de estos estudios describen la estructura secundaria del dominio FH1 como una alfa hélice que puede generar flexibilidad a uniones de proteína-proteína actuando como una especie de lazo capaz de unir un dominio con otro dentro de la misma proteína (Adzhubei *et al.*, 2013).

1.3.3 Estructura del dominio FH2

El dominio principal presente en todas las Forminas es el FH2 de un tamaño aproximado de 400 residuos de aminoácidos, que se localiza hacia el extremo carboxilo terminal de la mayoría de las Forminas, aunque puede variar mucho forzando la alineación en bloques completos de aminoácidos en lugar de secuencias (Higgs, 2005). A pesar de la poca conservación de estructura primaria, la estructura terciaria de las Forminas está altamente conservada.

El dominio FH2 de las Forminas forma homodímeros que se entrelazan, en forma antiparalela (región amino de un dímero con el carboxilo del otro), formando un anillo como se puede observar en la figura 3. Este anillo rodea el filamento de actina ayudando a la elongación de los microfilamentos mediante un movimiento procesivo en los extremos de punta (Breitsprecher & Goode, 2013).

En cada dominio FH2 se distinguen varias regiones que, del extremo amino al extremo carboxilo son el lazo, el enlazador, el botón, la bobina enrollada y el poste. El lazo interactúa con el poste del hemidímero junto a él y el enlazador es una región flexible y desordenada. El lazo y el enlazador cooperan para estabilizar la conformación en el movimiento gradual del dominio FH2, a lo largo de los filamentos de actina, y evitan la unión de dos moléculas diferentes (Lu *et al.*, 2007). El botón, la bobina enrollada y el poste conforman una estructura alargada en forma de puente o media luna. Finalmente, el poste es la porción que interactúa con el lazo del dominio FH2 de la otra Formina (Figura 3).

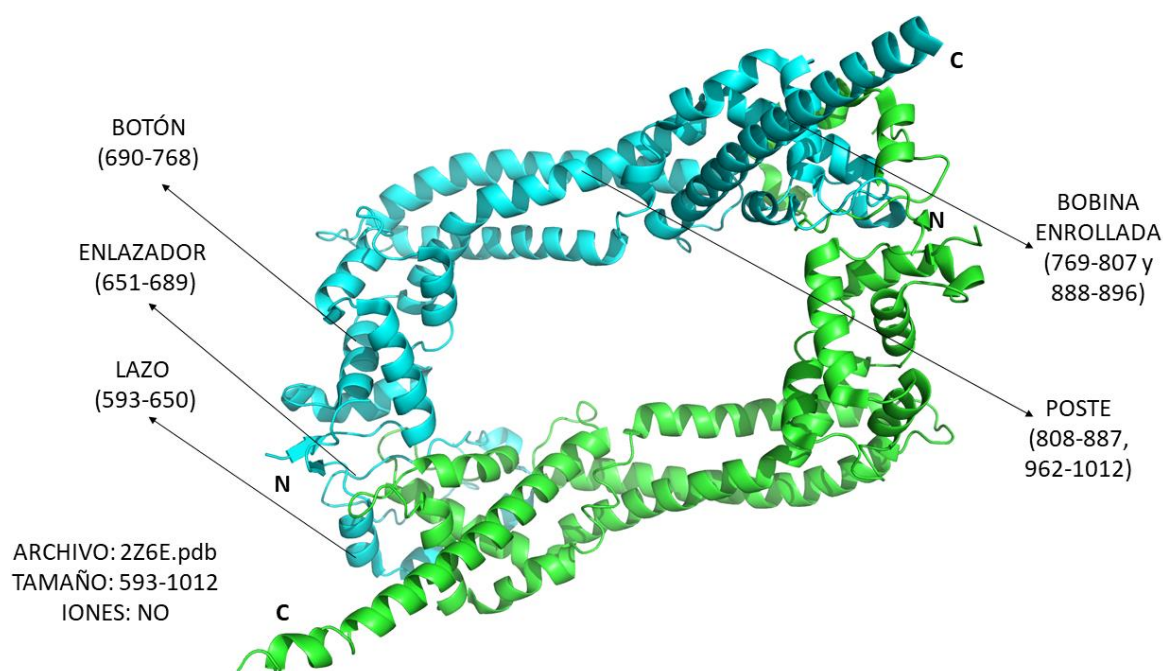


Figura 3 Diagrama en cintas que muestra la estructura cristalizada del homodímero del dominio FH2 de la Formina DAAM1 humana. Esta estructura es típica de los dominios FH2, los cuales constituyen la unidad básica estructural y funcional de la Forminas. En el diagrama se identifican los subdominios lazo, enlazador, botón, bobina y poste. Tomado y modificado de la estructura cristalográfica depositada en el PDB 2Z6E, reportado por Yamashita *et al.*, 2007.

1.3.4 Otros dominios presentes en las Forminas

Además de los dominios FH1 y FH2 característicos de todas las Forminas, estas proteínas pueden tener otros dominios que participan tanto en su localización

subcelular como en su regulación funcional. Dentro de estos dominios adicionales los mejor caracterizados son el dominio GBD/FH3 y el dominio PDZ.

El dominio GBD/FH3 (Figura 4) está presente en el extremo amino del subgrupo de las Forminas conocidas como “Forminas relacionadas a las diafanosas o Forminas diáfanas” que incluyen, las Forminas diáfanas (mDia1, mDia2, mDia3), las Forminas activadoras de morfogénesis asociado a la generación de filamentos (DAAM1 y DAAM2, *por sus siglas en inglés*) y las proteínas relacionadas a las Forminas identificadas en leucocitos (FRLs1, FRLs2 y FRLs3) (Rivero *et al.*, 2005). El dominio GBD/FH3 (IPR014768) tiene una longitud aproximada de 380 residuos y está compuesto por varios dominios que se traslapan: la parte amino terminal de un dominio de unión a GTPasas (GBD IPR010473), una región de repetición de armadillo (ARM IPR016024) y una hélice coiled coil o hélice superenrollada (Rose *et al.*, 2005). Se ha planteado la hipótesis de que este dominio tiene una función doble, por un lado, participa en la localización subcelular de las Forminas y por el otro participa en la regulación de su activación mediante un mecanismo de auto inhibición (Choi & Weis, 2005; Eng *et al.*, 2017).

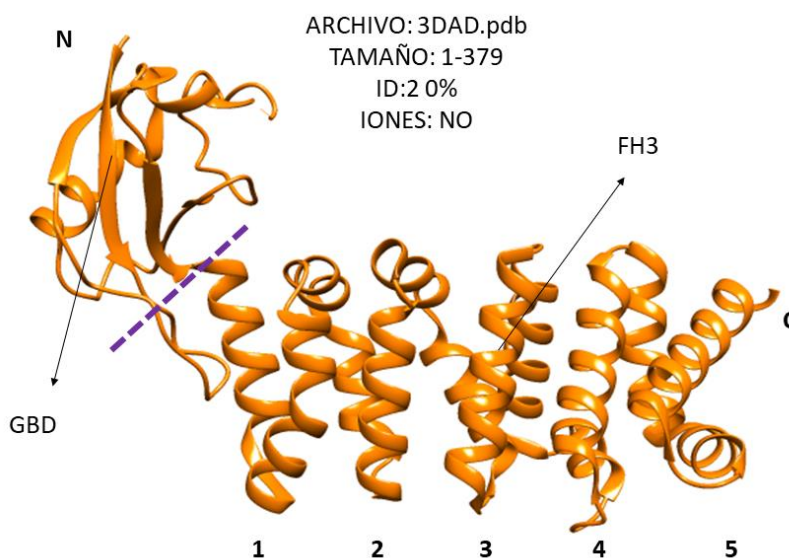


Figura 4 Diagrama en cintas del dominio GBD/FH3. En la región amino se encuentra el dominio GBD y en la región carboxilo se encuentra el dominio FH3, que está conformada por una repetición armadillo (ARM) de cinco hélices alfa. Tomado y modificado de la estructura cristalográfica depositada en el PDB 3DAD, reportado por Schulte *et al.*, 2008.

El dominio PDZ recibe su nombre por las tres primeras proteínas en las que fue identificado, las cuales fueron “*post synaptic density protein*” (PSD95), *Drosophila disc large tumor suppressor* (Dlg1) y *zónula occludens-1 protein* (zo-1). Este dominio tiene las funciones de organizar y mantener el andamiaje de complejos proteicos complejos y regular vías celulares permitiendo el acercamiento entre varias moléculas (Kennedy *et al.*, 1995) (Figura 2).

1.3.4 Forminas diáfanas

Una de las estructuras de las Forminas más estudiadas es la del subgrupo de las Forminas diáfanas. En la figura 5 se presenta la estructura de una de las 15 Forminas diáfanas reportadas en la literatura (Campellone *et al.*, 2010). Esta Formina está compuesta en el extremo amino por un dominio de sitio de unión a GTPasa (GBD), seguido de un dominio inhibidor diáfano (DID) y del dominio de dimerización (DD). Posteriormente, se encuentra el dominio FH1 y hacia el extremo carboxilo encontramos el dominio FH2 y el dominio autorregulador diáfano (DAD). Cuando la Formina está inactiva, hay una unión física entre el dominio DID y el dominio DAD. Por el contrario, cuando la Formina está activa se observa como un dímero unido por el dominio DD (Breitsprecher & Goode, 2013; Ergin & Zheng, 2020; Gorelik *et al.*, 2011).

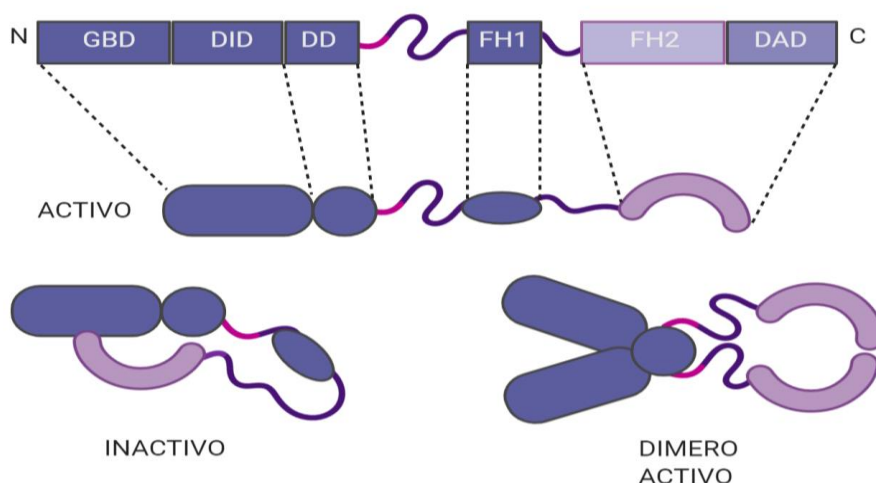


Figura 5 Arquitectura de los dominios de una Formina diáfana (*Drf*). Los límites del dominio se muestran a escala. Tomado y modificado de Chesarone *et al.*, 2010 & *What is the role of formin in actin polymerization?*, s.f.

1.3.5 Funciones celulares de las Forminas

Las Forminas se movilizan y se activan en diferentes sitios de la célula ayudando en la generación de estructuras fagocíticas, uniones celulares, estructuras endocíticas, fibras de estrés y membranas tubulares, y están involucrados activamente en la formación de lamelipodios y filopodios (Campellone *et al.*, 2010).

Como se nombró anteriormente las Forminas son una maquinaria de polimerización de actinas. Sin embargo, para muchos organismos se ha descrito que las Forminas tienen actividades más allá del ensamblaje de microfilamentos, trabajando como mediadores potenciales entre la actina y los microtúbulos durante la motilidad celular, observándose de manera estable estas uniones que se extienden hacia la corteza celular (Foldi *et al.*, 2017a).

1.3.5 Ubicación y estructura 3D de las Forminas

Las Forminas se movilizan y se activan en diferentes sitios de la célula ayudando en la generación de estructuras fagocíticas, uniones celulares, estructuras endocíticas, fibras de estrés, membranas tubulares. Así mismo, están involucradas activamente en la formación de lamelipodios y filopodios (Campellone *et al.*, 2010).

Las Forminas diáfanas se conocen como proteínas activadoras de GTPasa (GAP) y son reclutadas por una amplia variedad de proteínas Rho, siendo los factores reguladores Rho, Rac y Cdc42 los más característicos de este grupo, los cuales son activados por proteínas quinasas (Sit *et al.*, 2011). No obstante, otros factores también pueden regular las Forminas. Como ejemplos se encuentra la Formina Cdc12 de *Schizosaccharomyces pombe* que es activada por las MAPK y la Formina Bnr1 de *Saccharomyces cerevisiae* que es fosforilada en sus extremos amino y carboxilo por la quinasa Ark1/Prk1, lo que facilita la liberación de la auto inhibición (Gómez *et al.*, 2020; Smythe *et al.*, 2003). En los mamíferos, la Formina 2 invertida (INF2) y la proteína 2 similar a la Formina, conocida como FMNL2 o FRL3, es fosforilada por una GTPasa (Cheng *et al.*, 2011). Adicionalmente, se ha reportado

que algunas Forminas diáfanos e incluso sus GTPasas asociadas pueden sufrir modificaciones postraduccionales irreversibles por la adición de lípidos, las cuales ayudan a que estas proteínas cumplan diversas funciones en la reorganización citoesquelética. En el caso de la Formina mDia3, estas modificaciones le confieren una ventaja pues ayuda a que ella se ubique en la membrana de las células y a que su GTPasa blanco transforme sus células diana (Chhabra *et al.*, 2009; Block *et al.*, 2012).

1.3.2 Identificación de un inhibidor del dominio FH2

En el año 2009, mediante estudios *in vitro* en la búsqueda de moléculas inhibitorias de la polimerización de actina mediada por la región FH1-FH2 de las Forminas, se identificó un compuesto llamado “*Small Molecule Inhibitor of Formin FH2 domain*” (SMIFH2, *por sus siglas en inglés*), el cuál es derivado de 2-tiooxodihidropirimidina-4,6-diona. Rizvi *et al.* (2009) demostraron que este compuesto inhibe *in vitro* el ensamblaje de profilina-actina mediado por la Formina de ratón mDia1 (FH1FH2) y mDia2 (FH1FH2) de una manera dependiente de la concentración. A concentraciones de saturación de SMIFH2, la velocidad de ensamblaje era similar a la velocidad de ensamblaje sin Formina. No obstante, en estos experimentos no se determinó el sitio activo del inhibidor ni la región de la Formina mDia1 y mDia2 a la que se une.

1.4 Modelaje de la estructura de proteínas

Las proteínas se pliegan para conformar estructuras tridimensionales específicas según la secuencia de aminoácidos de la cadena polipeptídica. Esta estructura tridimensional sienta las bases para su interacción con otras moléculas, lo que es indispensable para ejercer su función. Por tanto, si se modifica sin una actividad específica la estructura tridimensional de una proteína podría perder su actividad.

A través del modelaje de la estructura terciaria de una proteína es posible predecir su papel dentro de la célula e identificar regiones y/o residuos de aminoácidos importantes para su función. El modelaje de la estructura de una proteína de interés

se puede abordar empleando programas que predicen un modelo mediante la comparación de la secuencia en estudio con modelos estructurales ya establecidos para otras proteínas (modelado por homología) o empleando programas que realizan un modelo a partir de deducciones de la estructura terciaria teniendo como base la secuencia de aminoácidos de la cadena polipeptídica (modelado *ab-initio* o *de novo*).

1.4.1 Modelado por homología

El objetivo principal de todos los programas de modelado por homología, al igual que los de *ab initio*, aunque sin una plantilla, es encontrar el modelo estructural conocido más parecido al plegamiento de la estructura nativa de la proteína a estudiar. Pero, aun cuando el resultado tenga una energía baja, no se garantiza que la predicción generada sea el plegamiento encontrado en la naturaleza. Este problema es más común de lo que se piensa. Por ejemplo, una proteína tendrá que atravesar el paisaje energético de plegamiento en su búsqueda de estabilidad estructural, entre más profundo llegue en este paisaje menos energía potencial tendrá y por lo tanto estará en equilibrio. En este “viaje”, la proteína tendrá que sortear varias barreras energéticas las cuales pueden funcionar como trampas que simularán un resultado final. Sin embargo, ese no sería el plegamiento que se encuentra en la naturaleza porque tal vez se puede ir más profundo (en cuestión de estabilidad) y por lo tanto tener un plegamiento diferente (Balchin *et al.*, 2020). En el caso de actina y tubulina, dentro de la célula requieren ser ayudadas en su plegamiento por chaperoninas, proceso que no es considerado al momento de realizarse el modelado por homología pero que afecta altamente la estructura tridimensional que toma la proteína.

Los servidores de modelaje por homología más empleados son ITASSER, PHYRE2, SWISS-MODEL y MODELLER. Estos programas comienzan demostrando que la proteína predicha es homóloga a la proteína usada como modelo y por lo tanto que van a tener un plegamiento y actividad similar, y que hay una mayor conservación de la estructura secundaria que de los residuos de aminoácidos (Webb *et al.*, 2021).

El modelado por homología se basa en 4 etapas: 1) identificación de las proteínas homólogas y determinación de la similitud entre la proteína molde y la proteína problema; 2) alineamiento de secuencias; 3) generación de las coordenadas de las regiones conservadas para construir el modelo; y 4) obtención del modelo (Webb *et al.*, 2021), siendo esta la metodología básica usada por la mayoría de los servidores como ITASSER, PHYRE2, SWISS-MODEL y MODELLER. Sin embargo, el uso de programas locales similares a MODELLER ayudan a generar predicciones a partir de un modelo comparativo con restricciones espaciales más específicas que en los servidores web, debido a que en estos últimos no se les puede modificar ningún parámetro (Figura 6). La identificación de las estructuras a partir de las cuales se realiza el modelado no necesariamente abarca una sola proteína. Adicionalmente, estos servidores permiten hacer modelado desde regiones o dominios homólogos hasta proteínas completas.

1.4.1.1 Modelado por homología en MODELLER

MODELLER es un programa de acceso libre que genera predicciones de estructura terciarias con base en la homología de la estructura entre la proteína problema (predicción) y la proteína empleada como molde (plantilla). Utiliza un método llamado “satisfacción de restricciones espaciales” que ubica la estructura predicha basándose en criterios probabilísticos y en la geometría de los átomos que están presentes en las plantillas obtenidas por cristalografía (Sali *et al.*, 1993). Estas restricciones espaciales están dadas principalmente por dos fuentes. En primer lugar, por las distancias de los enlaces y los ángulos diedros resultantes de la alineación entre la plantilla y la predicción. En segundo lugar, por las restricciones estereoquímicas como la longitud de enlace y las preferencias del ángulo de enlace, las cuales son extraídas del campo de fuerza de la mecánica molecular conocida como CHARMM-22 (Webb *et al.*, 2016).

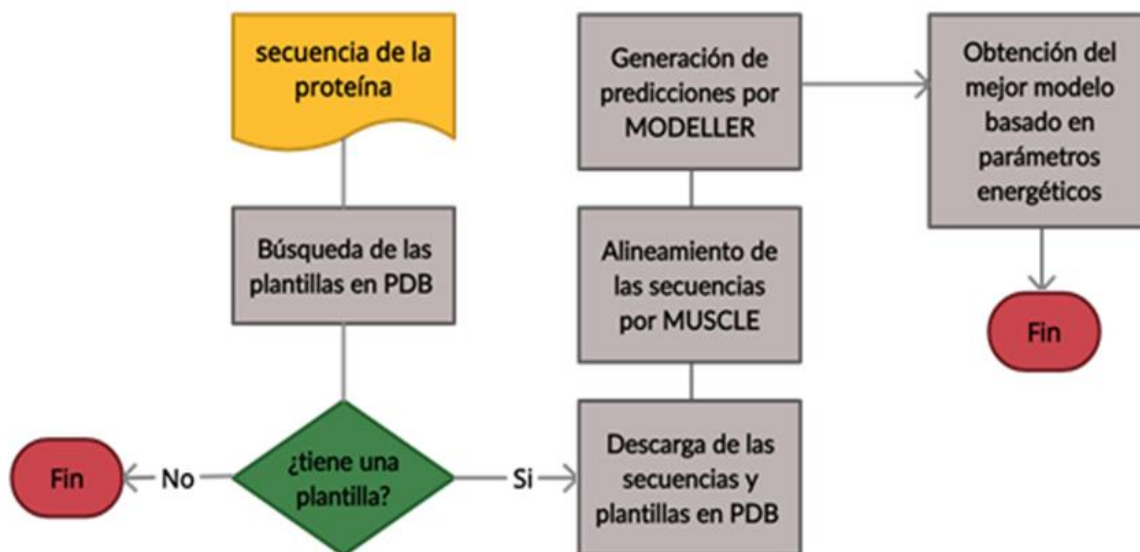


Figura 6 Diagrama de flujo con las etapas para el modelado por homología en MODELLER. Imagen tomada y modificada de Bitencourt-Ferreira *et al.*, 2019.

Cómo se observa en la figura 6, el primer paso para el modelado por homología es la búsqueda de la plantilla a partir de la secuencia de la proteína problema. Esta búsqueda se basa en el análisis de bibliotecas de proteínas cristalizadas depositadas en el Protein Data Bank (PDB, *por sus siglas en inglés*) por medio de la plataforma NCBIblast (National Center for Biotechnology information), que posee una herramienta de búsqueda de secuencias conocida como PSI-BLAST o por medio del servidor de SWISS-MODEL, el cual trabaja con un paquete de software conocido como HHblits (Deshpande *et al.*, 2005; Waterhouse *et al.*, 2018; Bienert *et al.*, 2017; Guex *et al.*, 2009; Studer *et al.*, 2020; Bertoni *et al.*, 2017).

Luego se hace una alineación en MUSCLE para escoger de todas las secuencias encontradas cual podría ser la mejor plantilla o más parecida a la proteína problema. Con la finalidad de generar un archivo de alineamiento el cual se va a utilizar para generar la predicción en MODELLER (Bitencourt-Ferreira *et al.*, 2019; Edgar, 2004; Deshpande *et al.*, 2005). Adicionalmente, este programa necesita un script de Python el cual ayuda ejecutar el modelado por homología y se especifica cuantas predicciones se desean hacer. El criterio para seleccionar cuantas predicciones realizará el programa se basa en la calidad de la plantilla usada. Cuando se tiene

un porcentaje de identidad bajo (menor del 30%) se necesita un mayor número de repeticiones de la predicción para asegurar que la energía de la estructura es la mínima esperada.

Al terminar su ejecución, MODELLER genera el número de predicciones seleccionado y un resumen de cada cálculo en la parte inferior del archivo, el cual posee el valor de la pseudoenergía conocido como DOPE (Discrete Optimized Protein Energy). Este valor se emplea como puntaje para establecer el mejor modelo comparando estadísticamente las distancias atómicas del modelo generado y de la plantilla (Eswar *et al.*, 2008).

Dado que el modelado mediante la satisfacción de las restricciones espaciales por homología puede utilizar diferentes tipos de información sobre la secuencia problema, es quizás la técnica de modelado comparativo más prometedora. Pero, si no existe una proteína cristalizada que pueda ser utilizada como plantilla, no se podría generar una predicción confiable. Es por esto por lo que para las proteínas o para los dominios o regiones de proteínas que no poseen un homólogo comparativo, se utiliza una metodología de modelaje *ab-initio* realizado por el programa computacional denominado Rosetta (Ovchinnikov *et al.*, 2018).

1.4.2 Modelado *ab-initio*

1.4.2.1 Modelado *ab-initio* en Rosetta

Rosetta es un programa que realiza predicciones *ab-initio* (no utiliza moldes o plantillas) mediante una búsqueda a través del espacio conformacional de la mejor posición de los átomos hasta encontrar la estructura con la menor energía libre (Bradley *et al.*, 2005). Este programa fue descrito por primera vez en 1999 en una comunidad científica que evalúa las predicciones de las estructuras conocida como Critical Assessment of protein Structure Prediction (CASP, *por sus siglas en inglés*) desde entonces han explicado sus avances y hasta limitaciones sobre las predicciones en Rosetta (Simons *et al.*, 1999).

Las estrategias para el modelaje *ab-initio* que utiliza Rosetta se basan en las leyes de la física y en las secuencias de aminoácidos de la proteína problema, en un proceso de ensamblaje de una biblioteca de fragmentos cortos (3 y 9 fragmentos) de proteínas construida en Robetta (Ovchinnikov *et al.*, 2018). Las interacciones de los enlaces de las cadenas laterales se reacomodan y son reemplazadas, para generar una predicción completa cercana a la proteína nativa, utilizando un método de minimización llamado Monte Carlo y apoyado bajo el criterio de metrópolis el cual puede aceptar o rechazar la nueva estructura reteniendo los ángulos que hacen que aumente la energía (Abbass & Nebel, 2020; Roterman-Konieczna, 2012). Estos métodos generan un análisis amplio pero ordinario basados en la energía libre (movimiento) de la estructura y la torsión de la columna vertebral que mide el cambio de los ángulos de los enlaces phi (Φ) y psi (ψ) (André, 2018).

Los métodos computacionales de predicción de estructuras pueden proporcionar modelos basados en la energía del sistema, siendo estas proteínas esenciales para comprender la función biológica y los procesos celulares. Pero, como se mencionó anteriormente; la naturaleza del viaje de una proteína a través del valle de energía libre es mucho más “complejo” y “movido”. Cuando se habla de “movido” se hace referencia al “movimiento atómico” descrito por Richard Feynman en una lectura sobre física en 1963, en donde se refiere a que los átomos están en un movimiento perpetuo y que es fundamental el estudio desde esta perspectiva. Entonces, desde el punto de vista del campo de fuerza atómico clásico en el cual se basan muchos programas como MODELLER o Rosetta, el estudio sobre la estructura e *interacción in silico* de las proteínas estaría incompleto o hasta errado, debido a que este campo clásico habla de cómo se conoce con exactitud donde se encuentran los átomos, reduciendo así el movimiento de la estructura y penalizando un cambio conformacional (Torda, 2004). Sin embargo, las simulaciones de dinámica molecular (DM) ayudan a predecir, describir y entender estos movimientos, mediante una suite que permite calcular las trayectorias en un periodo de tiempo, basados en la física Newtoniana.

1.4.3 Dinámica molecular

Los primeros conceptos sobre la dinámica molecular (DM) nacieron en la década de los 50's relatando la estructura de líquidos uniformes descrito por Alder y Wainwright. En 1977 se describió la primera simulación de DM para una proteína pequeña conocida como PTI (inhibidor de tripsina pancreática de vaca) (McCammon *et al.*, 1977). La evolución de esta técnica computacional, relativamente nueva, muestra una metodología dividida en dos sistemas. El primero trabaja con la mecánica molecular, en el cual se toma en cuenta las “fuerzas Newtonianas” para simular los movimientos de los átomos y los posibles enlaces o movimientos a futuro. La segunda metodología es la mecánica cuántica, la cual es una dinámica microscópica donde los movimientos son dados por los electrones que se comportan como “ondículas” (ondas y partículas) y su trayectoria es basada en probabilidades (Durrant & McCammon, 2011).

1.4.3.1 Parámetros importantes en la dinámica molecular

Para entender los movimientos moleculares de las fuerzas Newtonianas, es importante simular un sistema biológico parecido a la realidad. El programa empleado para la DM prepara el ambiente con parámetros como la temperatura, la presión, el pH, la solvatación y la unión a ligando o cofactores, factores importantes para analizar simulaciones en procesos biológicos como plegamientos de proteínas, reacciones catalíticas, cambios conformacionales e interacciones entre proteína-ADN, proteína-proteína o proteína-ligando, siendo esta última simulación una sección importante en este estudio para la identificación *in silico* de la Formina y el ligando SIMFH2 (Hollingsworth & Dror, 2018; Durrant & McCammon, 2011; Meena & Sulpizi, 2013; Jost Lopez *et al.*, 2020).

Las simulaciones en una DM utilizan campos de fuerza (pertenecientes a la mecánica clásica o molecular) que tiene información estructural y parámetros estandarizados, que ayudan a definir, dividir o estudiar cada tipo de átomo o molécula, y que junto a algoritmos computacionales describen la trayectoria de la

molécula a través del tiempo. Al observar la DM se ven a los átomos como “bolas sobre resortes” y a estos átomos se les asigna una “carga parcial fija” sin embargo, no se tienen en cuenta a los electrones desde el punto de vista cuántico (Priya *et al.*, 2016; Durrant & McCammon, 2011). Otra característica sobre el enlace de los átomos hace referencia a que estos métodos evitan los enlaces covalentes, en otras palabras; para el sistema no se crean ni se destruyen enlaces y el tiempo de corrida se puede medir hasta en una escala de milisegundos (Senn & Thiel, 2009; Hollingsworth & Dror, 2018; Pierce *et al.*, 2012).

1.5 *Trypanosoma cruzi* y la enfermedad de Chagas

T. cruzi es un protozoo unicelular flagelado causante de la tripanosomiasis americana o enfermedad de Chagas, el cuál es transmitido por insectos hemípteros de la familia Reduviidae (Figura 7). El nombre de la enfermedad deriva del apellido de su descubridor, Carlos Chagas, quien además describió el parásito en 1909 en Brasil. El nombre del género del parásito se deriva del griego (τρύπανον, trýpanon, que significa taladro, y σῶμα, soma, que significa cuerpo). El término griego hace referencia a la manera en que el microorganismo penetra las células eucariontes a las que parasita. Por otra parte, el nombre de la especie fue asignado por el Dr. Chagas en honor a su jefe y mentor Oswaldo Cruz (Guhl *et al.*, 2007).

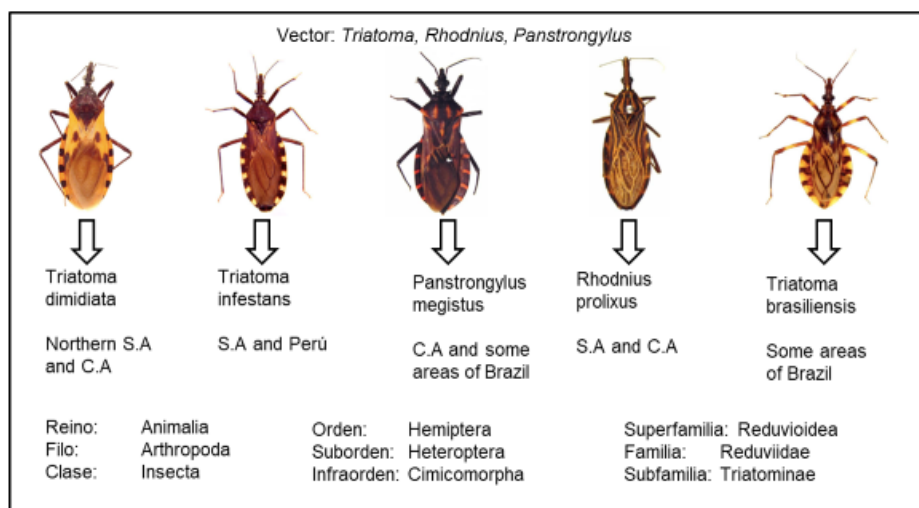


Figura 7 Distribución y taxonomía de los triatomas más importantes en Latino América. S.A: América del Sur; C.A: Centro América. Tomado de Vivas *et al.*1], 2021.

La Enfermedad de Chagas es endémica de América Latina, sin embargo, en los últimos años la movilidad de la población y los cambios climáticos han iniciado una variación epidemiológica de la enfermedad, generando un riesgo mundial, no solo por el transporte de los vectores sino por el aumento de donadores de sangre infectada con *T. cruzi*. Debido a esto, ya se considera una enfermedad emergente en países no endémicos como Estados Unidos, Canadá, Australia, Japón, Francia, España, Bélgica, Portugal, Suiza, Gran Bretaña, Irlanda del Norte, Italia, Alemania, Austria, Croacia, Dinamarca, Luxemburgo, Noruega, los Países Bajos, Rumania y Suecia (Dujardin *et al*, 2002; Salazar *et al.*, 2016) (Figura 8). Según datos de la Organización Mundial de la Salud (OMS), se estima que hay entre 6 y 7 millones de personas infectadas por *Trypanosoma cruzi* en el mundo y es considerada una de las enfermedades tropicales más olvidadas en América Latina (WHO, 2020).



Figura 8 Distribución mundial de casos de la Enfermedad de Chagas. Tomado de Salazar, 2016.

1.5.1 Ciclo de vida de *Trypanosoma cruzi*

El protozoo *Trypanosoma cruzi* es un organismo digenético con un ciclo biológico complejo con alternancia entre un hospedero vertebrado y un insecto vector. El ciclo

de vida de este parásito flagelado inicia cuando el estadio infectante, el tripomastigote metacíclico, se encuentra presente en las deyecciones del triatomino y se introduce en el huésped a través de las heridas y excoriaciones causadas al momento de la comezón generada por la picadura del vector. También, puede ingresar por contacto directo de las deyecciones con la mucosa ocular, nasal o bucal (Botero *et al.*, 2012).

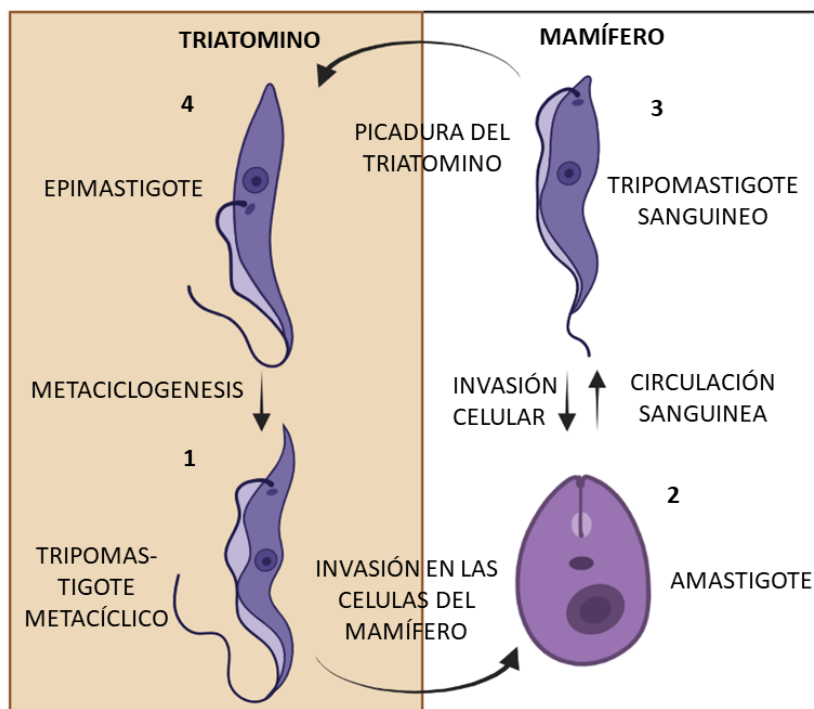


Figura 9 Ciclo de vida de *Trypanosoma cruzi*. Aproximaciones morfológicas comúnmente observadas en vertebrados e invertebrados. Elaboración propia.

Cuando el tripomastigote metacíclico (Figura 9.1) ingresa en una célula de mamífero se transforma en amastigote (Figura 9.2), el cual se replica de manera seriada acumulándose dentro de la célula. Posteriormente, los amastigotes se convierten en tripomastigotes sanguíneos rompiendo la célula, lo que los libera al torrente sanguíneo. El tripomastigote sanguíneo vuelve a infectar células, repitiéndose el ciclo de infección. Cuando el insecto se alimenta de un huésped infectado, los tripomastigotes sanguíneos circulantes pasan a él (Figura 9.3). En el intestino del insecto los tripomastigotes se transforman en epimastigotes (Figura 9.4), los cuales constituyen una segunda etapa reproductiva, replicándose dentro del intestino. Al

llegar al intestino posterior del insecto, los epimastigotes se convierten en tripomastigotes metacíclicos y se evacúan a través de las heces (Figura 9.1). Si las heces ingresan al hospedero, le causarán infección, repitiéndose así el ciclo de vida.

1.5.3 Citoesqueleto de *Trypanosoma cruzi*

Este parásito tiene la habilidad de adaptarse al entorno en el que se encuentra, llevando a cabo su ciclo de diferenciación celular por medio de transformaciones en la morfología y cambios en el metabolismo durante las diferentes etapas de su ciclo de vida. Estas transformaciones requieren cambios en la expresión de ARN mensajeros y proteínas, y en la reorganización de sus organelos, para lo cual requiere la remodelación de su citoesqueleto (Hamilton & Stevens, 2010).

El citoesqueleto de *T. cruzi* es atípico ya que está primordialmente constituido por microtúbulos. Estudios de microscopía electrónica de transmisión y microscopía electrónica de barrido de haz de iones enfocado (UPV), han permitido la reconstrucción tridimensional de toda la célula completa. Estos análisis demostraron que *T. cruzi* tiene en su citoesqueleto una red de microtúbulos estables, organizados y unidos a la membrana plasmática, conocidos como corsé subpelicular (Souza *et al.*, 2009). Este corsé le confiere rigidez al cuerpo del protozoario y determina su forma. El único sitio no recubierto por microtúbulos es la región por la que emerge el flagelo conocido como bolsillo flagelar. El flagelo tiene un axonema típico compuesto de dos pares de microtúbulos centrales y 9 pares periféricos, junto al cual se encuentra una estructura peculiar de tripanosomátidos conocida como bastón paraflagelar (Souza *et al.*, 2009).

Otra estructura del citoesqueleto de este protozoo es el citostoma-citofarínge, que se encuentra en las etapas proliferativas del parásito (epimastigote y amastigote) y es el sitio principal para la endocitosis. Esta estructura consiste en una abertura en la superficie de la membrana plasmática cerca del bolsillo flagelar, llamada citostoma, seguida de una invaginación profunda de la membrana, llamada citofarínge.

Hasta el momento no se ha podido identificar por estudios de microscopía de luz o electrónica la presencia de microfilamentos en tripanosomátidos ni tampoco se ha podido teñir microfilamentos utilizando faloidina, una toxina presente en setas de la familia Amanita, que se une a actina filamentosa y que se utiliza ampliamente para decorarla (Mortara *et al.*, 1989). Sin embargo, todos los tripanosomátidos reportados a la fecha codifican actina en su genoma la cual está involucrada en el proceso de endocitosis (Souza *et al.*, 2013; Vizcaíno-Castillo *et al.*, 2020; Bogitsh *et al.*, 1995; Chasen *et al.*, 2020). Por otra parte, no se ha podido demostrar su participación en otros procesos celulares como citocinesis, adherencia y motilidad. Estos reportes sugerirían que el citoesqueleto de actina es rudimentario en este parásito. No obstante, el análisis del genoma de *T. cruzi* ha demostrado que es más complejo de lo esperado ya que contiene 12 loci anotados como codificantes para proteínas de la familia de las actinas: dos actinas convencionales (Vizcaíno-Castillo *et al.*, 2020), siete proteínas parecidas a actina y tres ARP (ARP2, ARP3 y ARP6). Adicionalmente, se identificaron genes codificadores para proteínas reguladoras del ensamblaje y desensamble de actina como la profilina, las proteínas del complejo ARP2/3, y tres Forminas (Forminas A, B y C) (Melo *et al.*, 2008). Así mismo, se identificaron diez miosinas, nueve de las cuales son específicas de tripanosomátidos (de Souza *et al.*, 2018).

De manera interesante, estudios filogenéticos demostraron que las actinas 2 y 3 de *T. cruzi* están presentes solo en tripanosomátidos del grupo estercoreario, pero no en *T. brucei* o en especies del género *Leishmania* (Vizcaíno-Castillo *et al.*, 2019). Por otra parte, la actina 4 se encuentra presente en el género *Leishmania*, pero no en *T. brucei*. De manera similar, el número de proteínas de unión a actina es mayor en tripanosomátidos estercorearios que en los salivarios y *Leishmania* sp. Por ejemplo, mientras que los tripanosomátidos estercorearios tienen diez miosinas, los salivarios tienen dos miosinas y el género *Leishmania* tiene tres miosinas. Además, los tripanosomas estercorearios tienen tres Forminas mientras que los tripanosomas salivarios y el género *Leishmania* tienen dos Forminas.

Hasta el momento solo se ha estudiado la expresión de las dos actinas convencionales, Actina 1 y Actina 2. Ambas proteínas son expresadas de manera constitutiva en todos los estadios de *T. cruzi*, pero tienen localizaciones diferentes dentro de la célula (Vizcaíno-Castillo *et al.*, 2019) (Figura 10). Actina 1 aunque en su mayoría se encuentra en el citoplasma, también está asociada a membrana y citoesqueleto. En contraste, Actina 2 es exclusivamente citosólica (Vizcaíno-Castillo *et al.*, 2019).

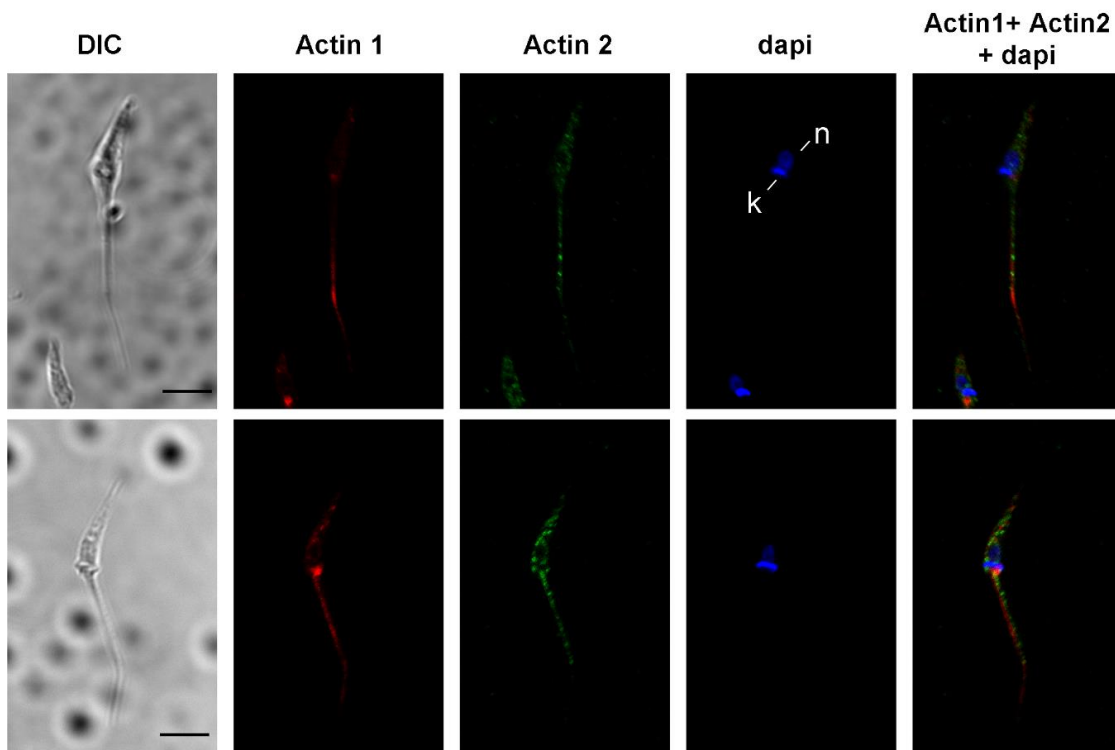


Figura 10 Localización diferencial de actina 1 y actina 2 en epimastigotes de *Trypanosoma cruzi*. Anticuerpos específicos que reconocen actina 1 (en rojo) o actina 2 en verde muestran que hay una localización distinta de ambas proteínas sin evidencia de colocalización. n= núcleo; k= cinetoplasto. Scale bar= 5 μ m. Tomada de Vizcaino-Castillo, 2020.

En cuanto a las proteínas de unión a actina, la única que se ha empezado a estudiar en *T. cruzi* es profilina. El análisis de su expresión demostró que esta proteína es primordialmente citosólica y que es capaz de interactuar con Actina 1, pero no se identificó interacción con otras actinas (Osorio-Méndez, 2017). Ya que las Forminas son un ligando natural de la profilina y hasta el momento poco se ha reportado sobre su función, se decidió iniciar con un estudio de caracterización de estas proteínas.

Un análisis bioinformático preliminar realizado en este laboratorio demostró que las tres Forminas de *T. cruzi* contienen los dominios FH1 y FH2, característicos de las Forminas, pero que en sus extremos amino terminal contienen dominios diferentes. En la Formina B se identificó un sitio de unión a PI(4,5)P2 llamado PTEN-like, en la Formina C se identificó un dominio auto inhibidor GBD-FH-3, uno de los motivos mejor caracterizados de las Forminas de metazoarios, y en la Formina A no se identificaron motivos específicos (Osorio Méndez, 2017).

3. HIPOTESIS

Ya que la utilidad biológica de una proteína y su capacidad de interactuar con otras proteínas o moléculas está dictada por su estructura tridimensional, se postula que la obtención de un modelo estructural de la Formina C nos permitirá generar información sobre cómo funciona, cómo se regula y permitirá diseñar mutaciones dirigidas a las regiones específicas con la intención de interferir con su localización y función.

4. OBJETIVOS

4.1 Objetivo general

Obtener un modelo estructural por homología de la Formina C de *Trypanosoma cruzi* que permita delimitar sus regiones funcionales.

4.2 Objetivos particulares

- ✓ Identificar y caracterizar los motivos conservados de las Forminas A, B y C de *Trypanosoma cruzi*.
- ✓ Determinar la relación filogenética de la Formina C de *Trypanosoma cruzi* con Forminas A y B de otros tripanosomátidos.
- ✓ Determinar la estructura terciaria *in silico* de la Formina C de *Trypanosoma cruzi*.
- ✓ Analizar la interacción por dinámica molecular entre la Formina C de *Trypanosoma cruzi* y el inhibidor SMIFH2.

5. METODOLOGIA

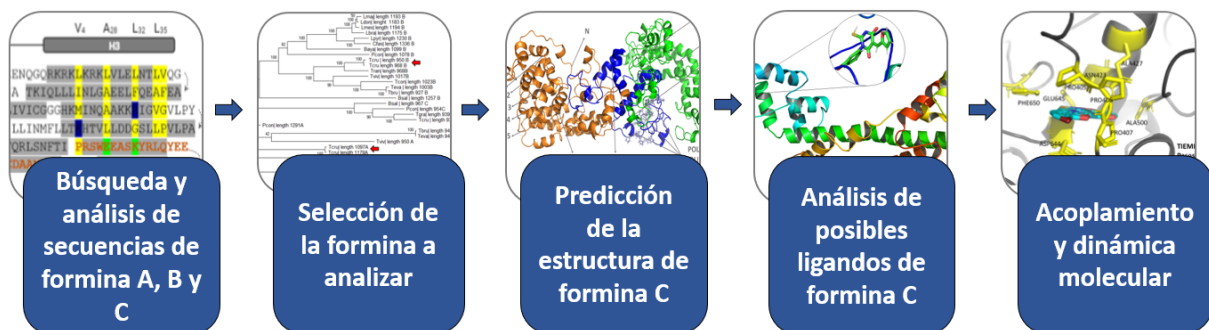


Figura 11. Estrategia metodológica.

5.1 Búsqueda de secuencias de las Forminas A, B y C en *Trypanosoma cruzi* y otras especies de tripanosomátidos

La búsqueda de genes ortólogos de las Forminas de *Trypanosoma cruzi* en otras especies de tripanosomátidos se realizó por medio de la función de identificador de ortólogos y parálogos presente en el servidor TriTrypDB (<https://tritrypdb.org/tritrypdb/app/>) parte de VEUPATHDB (The Eukaryotic Pathogen, Vector and Host Informatics Resource, <https://veupathdb.org/veupathdb/app/>). Para la búsqueda se utilizó como plantilla el número de identificación de la secuencia de las regiones codificadoras de Formina A (TcCLB.511755.80) y la Formina B (TcCLB.508641.320) del organismo *Trypanosoma cruzi* CL Brener No-Esmeraldo-like ubicados en el cromosoma 17-P (en antisentido, nucleótidos 349,795 a 353,088) y 26-P (en antisentido, nucleótidos 490,239 a 493,145), respectivamente. Para la búsqueda de la Formina C se empleó como plantilla el gen de la Formina C (TcCLB.511393.30) de *Trypanosoma cruzi* CL Brener Esmeraldo-like ubicado en el cromosoma 19-S (en antisentido, nucleótidos 517,371 a 520,184).

5.2 Identificación de los dominios proteicos en las Forminas A, B y C

Por medio del programa InterPro (<https://www.ebi.ac.uk/interpro/>) se hizo una búsqueda de los dominios que se pueden encontrar en las Forminas A, B y C

identificadas previamente en *T. cruzi* CL Brener y otras especies de tripanosomátidos.

5.3 Elaboración del árbol filogenético de las Forminas A, B y C

Una vez obtenidas las secuencias codificadoras de los ortólogos y parálogos de las Forminas A, B y C de otros tripanosomátidos, se realizó un alineamiento múltiple por medio de Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) (Madeira *et al.*, 2019) y se descartaron las secuencias incompletas. Con la finalidad de establecer la relación filogenética solo se tomó el dominio FH2 (Interpro número IPR029021) de las Forminas A, B y C, debido a que en la literatura se ha reportado que esta región ocurrió solo en una ocasión en la evolución (Chalkia *et al.* 2008). Para ello, se obtuvo la secuencia que correspondía al dominio FH2 de las Forminas a evaluar y con ella se construyó un árbol filogenético en MEGA 11 (Kumar *et al.* 2016) por el método de máxima verosimilitud con un Bootstrap de 1,000 replicados.

5.4 Búsqueda de motivos adicionales en la Formina C de *Trypanosoma cruzi*

Además de la identificación de dominios utilizando InterPro, se buscaron motivos de localización celular y transmembranales para predecir si estas proteínas cuentan con motivos de unión a membrana, nucleares y mitocondriales, mediante la herramienta TMHMM2.0 (<http://www.cbs.dtu.dk/services/TMHMM-2.0/>). Esta herramienta posee un algoritmo de búsqueda conocido como N-best que analiza las predicciones más probables a partir de las hélices que forma la secuencia problema, teniendo como guía las hélices transmembranales que se encuentran almacenadas en bibliotecas siendo descritas por análisis experimentales.

5.5 Análisis estructura primaria de la Formina C de *Trypanosoma cruzi*

Para hacer el análisis de la estructura primaria de la Formina C se utilizó la herramienta ProtParam (<https://www.expasy.org/resources/protparam>) contenida en Expasy, la cual revisa en la secuencia analizada diferentes factores físicos y químicos para calcular el peso molecular, el punto isoeléctrico, el porcentaje de

aminoácidos y el coeficiente de extinción molar, con la finalidad de obtener una caracterización inicial de la secuencia.

5.6 Predicción de la estructura secundaria de la Formina C de *Trypanosoma cruzi*

El análisis de la estructura secundaria se realizó en RaptorX (<http://raptorx.uchicago.edu/>) (Källberg *et al.*, 2012). Este servidor genera una correlación entre la secuencia de aminoácidos y su potencial para estructurarse en alfa-hélices, láminas β o no estructurarse y formar lazos. Para ello identifica la accesibilidad que tiene cada uno de los residuos a solventes (ACC o ASA) debido a 3 posiciones de los residuos (enterrado, intermedio y expuesto) y las regiones de desorden (DISO) que clasifica como desordenado (*) a un aminoácido. El programa califica a un residuo desordenado, si este residuo está en un segmento de más de tres residuos en donde faltan coordenadas atómicas en la estructura de rayos X, estando esta información almacenada en bases de datos de proteínas.

5.7 Predicción de la estructura terciaria

El modelaje bioinformático de la estructura terciaria de la Formina C de *Trypanosoma cruzi* CL Brener se hizo con la asesoría del Dr. Martin González Andrade, investigador del Laboratorio de Biosensores y Modelaje Molecular de la Facultad de Medicina de la UNAM (Para conocer las características de los equipos empleados para la ejecución de los modelados, el acoplamiento y la dinámica molecular, ver anexo 1).

5.7.1 Modelado por homología en MODELLER

El modelado por homología se dividió en cuatro pasos:

1. Se identificaron las proteínas cristalizadas homologas a la Formina C en el PDB de NCBIblast y en SWISS-MODEL (<https://swissmodel.expasy.org/>) (<https://blast.ncbi.nlm.nih.gov/>). Los criterios empleados para la búsqueda de las plantillas en el NCBIblast se describen en la tabla 1, mientras que los criterios empleados en SWISS-MODEL fueron los predeterminados por el programa en el servidor (Tabla 1).

Tabla 1 Parámetros empleados para la búsqueda de proteínas cristalizadas homólogas a la Formina C en el servidor NCBIblast.

Numero de acceso en TriTrypDB	TcCLB.511393.30
Base de datos	Protein Data Bank
Algoritmo	protein-protein BLAST
Matriz de sustitución	BLOSUM62
Valor esperado €	0,05

2. Se realizó un alineamiento de las plantillas seleccionadas con la región correspondiente de la Formina C en el programa ClustalW (<https://www.expasy.org/resources/clustalw>) generando un archivo de alineamiento (.txt).
3. Se diseñó un script de Python en donde se encontraban las directrices para ejecutar MODELLER (<https://salilab.org/MODELLER/tutorial/basic.html>) junto con la plantilla seleccionada. En este caso se hicieron 1000 modelos de cada región a modelar. El script de Python contenía aparte de las instrucciones de como modelar estas regiones, una biblioteca de las coordenadas generales de cada átomo, la nomenclatura, la topología, las coordenadas x, y, z, y el factor B (movilidad de cada átomo en el esqueleto proteico) teniendo en cuenta la metodología descrita por Yuan & Teasdale, (2005).
4. A partir de las 1000 predicciones resultantes y teniendo en cuenta el valor DOPE se escogió la estructura con menor energía para hacer los posteriores estudios de dinámica molecular.

5.7.2 Modelado ab-initio por Rosetta

Las regiones o dominios para los cuales no se les encontró una plantilla o se les encontró una plantilla con menos del 20% de identidad, se modelaron *ab-initio* empleando Rosetta. Este modelado se realizó siguiendo los 3 pasos descritos a continuación:

1. Se colocó la secuencia de aminoácidos de la región no estructurada de la Formina C en el servidor Robetta (<http://old.rosetta.org/>) para generar una biblioteca de archivos con 3 y 9 residuos de aminoácidos.
2. Se generó un archivo .flag el cual contenía el nombre de los archivos y su ubicación, el nombre del PDB que contiene la estructura de la proteína problema y el número de predicciones a generar, que en este caso fueron 10.000 predicciones para garantizar un modelo con la menor energía posible.
3. Posteriormente el archivo .flag y los archivos asociados a este se colocaron en el mismo directorio donde se corrió Rosetta.
4. Se generaron 10.000 modelos y el modelo con menor energía se seleccionó para la construcción del modelo global de la estructura de Formina C y para sus posteriores ensayos de dinámica molecular.

5.8 Verificación del modelo de la Formina C de *Trypanosoma cruzi* obtenido por medio de la predicción *in silico*

Como una estrategia para verificar si el modelo de la Formina C predicho anteriormente tiene una estructura terciaria satisfactoria, que pueda ser empleada para estudios posteriores, se evaluó el modelo construido en formato .pdb por medio de los servidores SAVES v.6.0 y Rosetta Design empleando los parámetros que trae por defecto cada uno de los servidores.

5.9 Acoplamiento y dinámica molecular

Para simular la interacción entre la macromolécula Formina C (estructura tridimensional modelada previamente) y el ligando SMIFH2 por medio de una DM, se realizó previamente una DM de la macromolécula y un acoplamiento molecular del ligando. Todos los ensayos de acoplamiento y dinámica molecular se realizaron por triplicado. La metodología de estos tres ensayos (DM de la macromolécula, acoplamiento del ligando y DM del complejo) se esquematiza en la figura 12 y se describe a continuación.

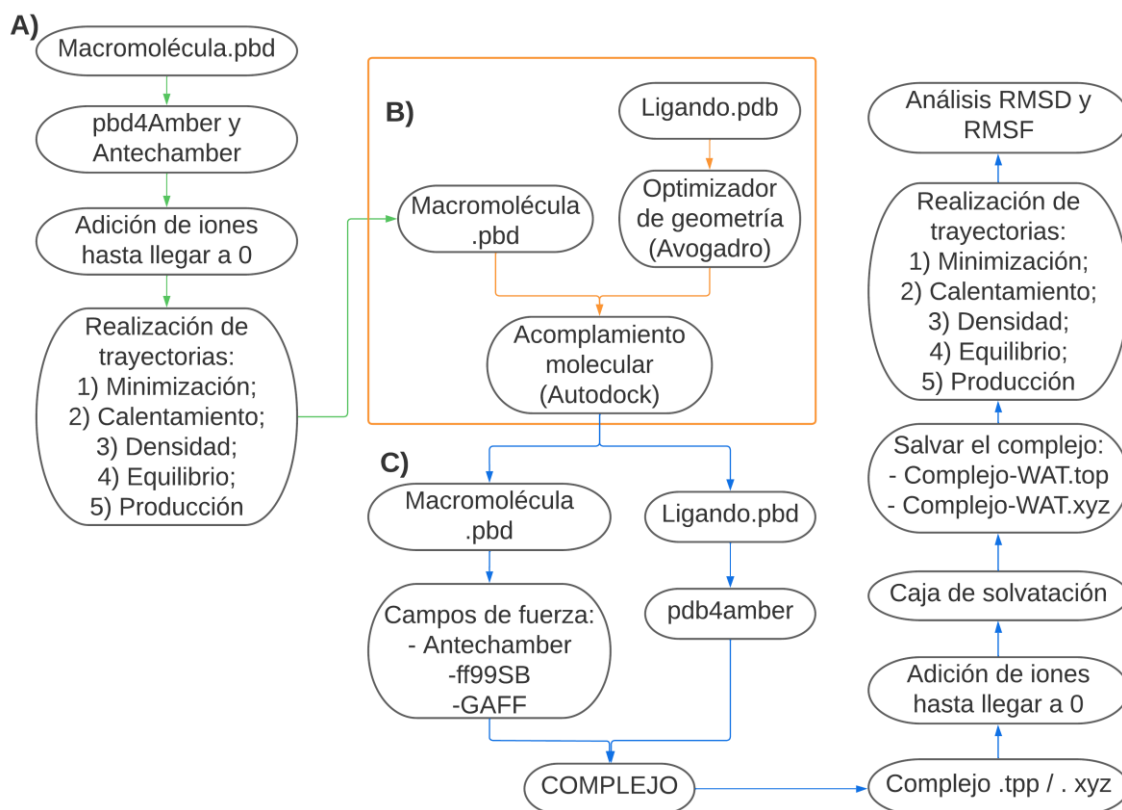


Figura 12 Pasos para el acoplamiento y dinámica molecular de Formina C y el ligando SMIFH2. A) Pasos generales para la preparación de la macromolécula por medio de dinámica molecular. B) Pasos generales para la preparación y parametrización del ligando, y posterior interacción con el modelado global de la macromolécula. B) Dinámica molecular del complejo macromolécula (modelo de la Formina C de *T. cruzi*) y ligando SMIFH2.

5.9.1 Dinámica molecular para la preparación de la macromolécula

Tras la obtención y verificación del modelo de la Formina C (macromolécula) se realizó una DM a esta estructura con la finalidad de prepararla para el acoplamiento y la dinámica molecular junto con el ligando.

La DM de la macromolécula se realizó por medio de un grupo de programas unidos bajo el nombre de AMBER (<http://ambermd.org/>) (Figura 12A). Para este ensayo se empleó el modelo de la estructura obtenido previamente en formato .pdb.

Los archivos de entrada necesarios para hacer la minimización y la dinámica molecular se crearon por pmemd (maquinaria para ejecutar dinámicas moleculares) de la siguiente forma:

1. Se enumeraron los átomos presentes en la molécula de manera consecutiva utilizando el programa pdb4amber.
2. Se usó Antechamber a manera de antesala, el cual generó archivos de topología fáciles de leer para LEaP.
3. Se ejecutó el comando parmchk2 donde se generó un archivo .frmod que después se utilizó como molde.
4. Se salvó la topología y coordenadas de la macromolécula y se revisó su carga.
5. Se adicionaron iones hasta llegar a una carga neta de 0 se guardó su topología y coordenadas en formato .pdb.
6. Se realizó un protocolo de minimización, calentamiento, densidad, equilibrio y producción.
7. Finalmente se analizaron los resultados por cpptraj.

5.9.2 Acoplamiento molecular

Se realizaron ensayos de acoplamiento molecular para parametrizar el ligando, así como para realizar un acoplamiento molecular entre la macromolécula y el ligando siguiendo los pasos descritos en la Figura 12B.

5.9.2.1 Obtención del ligando

Las coordenadas y estructura química del ligando SMIFH2 se obtuvieron a través de la base de datos PubChem (<https://pubchem.ncbi.nlm.nih.gov/>) en formato .pdb.

5.9.2.2 Avogadro y preparación del ligando

Con la finalidad mejorar la geometría, el ligando SMIFH2 se parametrizó de manera automática por medio del programa de edición de moléculas Avogadro (<https://sourceforge.net/projects/>). Para la parametrización se utilizó un el campo de fuerza UFF (Universal Force Field) el cual tiene un reconocimiento universal que mejora la estructura de los átomos de los compuestos analizados (Figura 12B).

5.9.2.3 AutoDock4

Una vez se mejoraron las características estructurales y de energía del ligando, se hizo un acoplamiento molecular, del ligando parametrizado con contra el modelo de la macromolécula Formina C previamente preparado por dinámica molecular, por medio de AutoDock4 (<http://autodock.scripps.edu/>):

1. Se crearon dos archivos PDBQT que contenían las coordenadas necesarias para llevar a cabo el acoplamiento de las moléculas, uno para la Formina C y el otro para SMIFH2. En estos archivos se incluyeron las cargas parciales atómicas, los grados de libertad de torsión del ligando y los tipos de los átomos.
2. Se creó una caja alrededor de la macromolécula como el posible espacio de unión del ligando y se guardó el archivo resultante en formato .gpf.
3. Se realizó el cálculo anticipado de afinidades entre los átomos por medio de las energías en AutoGrid4.
4. Se utilizó el algoritmo genético de Lamarck que es el algoritmo más usado para la búsqueda y optimización global y local de la unión entre el ligando y el receptor (Fuhrmann *et al.*, 2010). Se realizaron 1000 simulaciones de acoplamiento cuyos resultados se guardaron en formato .dlg.
5. Se analizaron los resultados de las interacciones y la agrupación de los clústeres por medio del rango de energía para seleccionar la mejor posición del acoplamiento, es decir la posición con menor energía.

5.9.3 Dinámica molecular

Una vez obtenida la mejor posición del acoplamiento entre la molécula y el ligando, se procedió a hacer la DM por medio de un grupo de programas unidos bajo el nombre de AMBER (<http://ambermd.org/>) (Figura 12C) con la finalidad de simular la interacción entre la macromolécula y el ligando. Para estos ensayos se empleó el ligando previamente parametrizado y el modelo de la macromolécula previamente preparado por acoplamiento molecular, ambos en formato .pdb.

Los archivos de entrada necesarios para hacer la minimización y la dinámica molecular se crearon por pmemd de la siguiente forma:

1. Se enumeraron los átomos presentes en la molécula de manera consecutiva utilizando el programa pdb4amber.
2. Se usó Antechamber a manera de antesala, el cual generó archivos de topología fáciles de leer para LEaP.
3. Se ejecutó el comando parmchk2 donde se generó un archivo .frcmod que después se utilizó como molde.
4. Se utilizó el programa LEaP que lee el campo de fuerza y se cargó el campo de fuerza GAFF para analizar las variables de las moléculas nuevas que no están en la base de datos de AMBER.
5. Por medio de la herramienta autodock se colocó el ligando dentro de la macromolécula.
6. Se analizó el complejo para determinar si la parametrización fue satisfactoria.
7. Se salvó la topología y coordenadas del complejo, y se revisó la carga del complejo.
8. Se adicionaron iones hasta llegar a una carga neta de 0 y se hizo una caja de solvatación.
9. Se guardó la topología y las coordenadas del complejo solvatado en formato .pdb.
10. Se realizó un protocolo de minimización, calentamiento, densidad, equilibrio y producción.
11. Finalmente se analizaron los resultados por cpptraj.

6. RESULTADOS

6.1 Análisis de las secuencias de las Forminas A, B y C codificadas en *Trypanosoma cruzi* y en otras especies de tripanosomátidos

Con la finalidad de determinar las secuencias que se codifican para las Forminas en el genoma de *T. cruzi* CL Brener, se realizó una búsqueda de estos genes en el servidor TriTrypDB. Se identificó que esta cepa tiene anotados 6 genes que podrían codificar para 3 Forminas, de los cuales 3 corresponden al haplotipo Esmeraldo (S) y 3 al haplotipo No-Esmeraldo (P) (Tabla 2).

Tabla 2 Genes identificados en Trypanosoma cruzi CL Brener que codifican para Forminas empleando la base de datos de TriTrypDB. Trypanosoma cruzi CL Brener haplotipo Esmeraldo (S), Trypanosoma cruzi CL Brener haplotipo no Esmeraldo (P).

Formina	Gene ID	Haplotipo	Número de nucleótidos	Número de aminoácidos
A	TcCLB.511313.30	S	3537	1178
	TcCLB.511755.80	P	3294	1097
B	TcCLB.506203.80	S	2853	950
	TcCLB.508641.320	P	2907	968
C	TcCLB.511393.30	S	2814	937
	TcCLB.503565.10	P	687	228

En la tabla 2 se puede ver que, los genes de las Forminas anotadas en los haplotipos S y P varían en el número de nucleótidos y por ende en el número de aminoácidos. Un análisis detallado de los dominios presentes en estos 6 genes evidenció que la variación de los tamaños se debe principalmente a diferencias en el dominio FH1 (Figura 13). Por otra parte, de estos 6 genes resalta el gen de la Formina C del haplotipo P puesto que es el de menor tamaño, con solo 228 residuos de aminoácidos.

Ya que los algoritmos con los que se identifican los dominios presentes en las proteínas son actualizados regularmente conforme aumenta el número de secuencias que ingresan a las bases de datos, se consideró importante determinar

si existían diferencias en los dominios de las Forminas identificados por Osorio Méndez (2017), con los identificados en este trabajo. Así mismo, se analizó con detalle el gen de la Formina C del haplotipo P con la finalidad de sugerir si codifica para algún dominio proteico reportado en la literatura.

El análisis de la secuencia de aminoácidos de los 6 genes de Forminas identificados en *T. cruzi* CL Brener por medio del programa InterPro demostró que, 5 de los 6 genes se encuentran completos y el gen que se encuentra incompleto corresponde al gen Formina C del haplotipo P puesto que solo tiene una parte de la secuencia del dominio FH2 (Figura 13). Por lo anterior se descartó el uso de esta secuencia en los análisis posteriores.

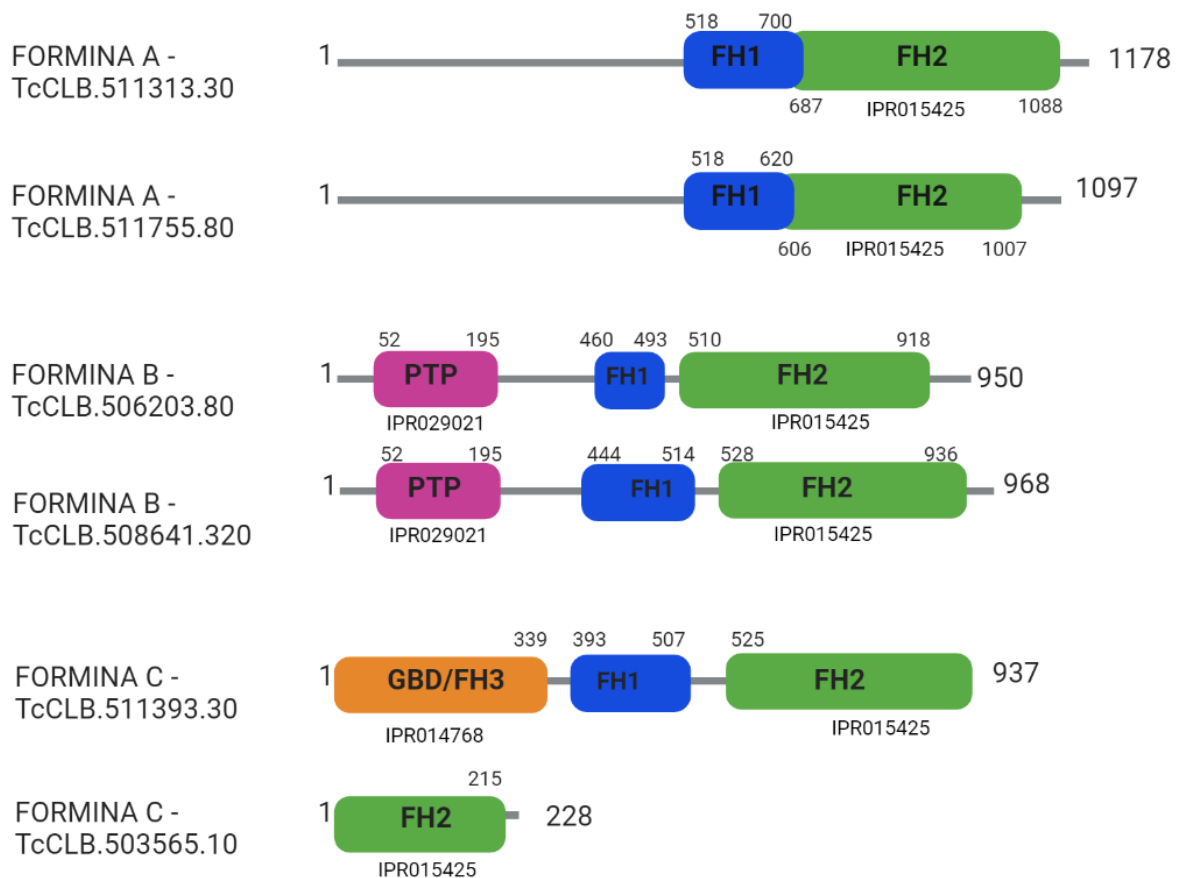


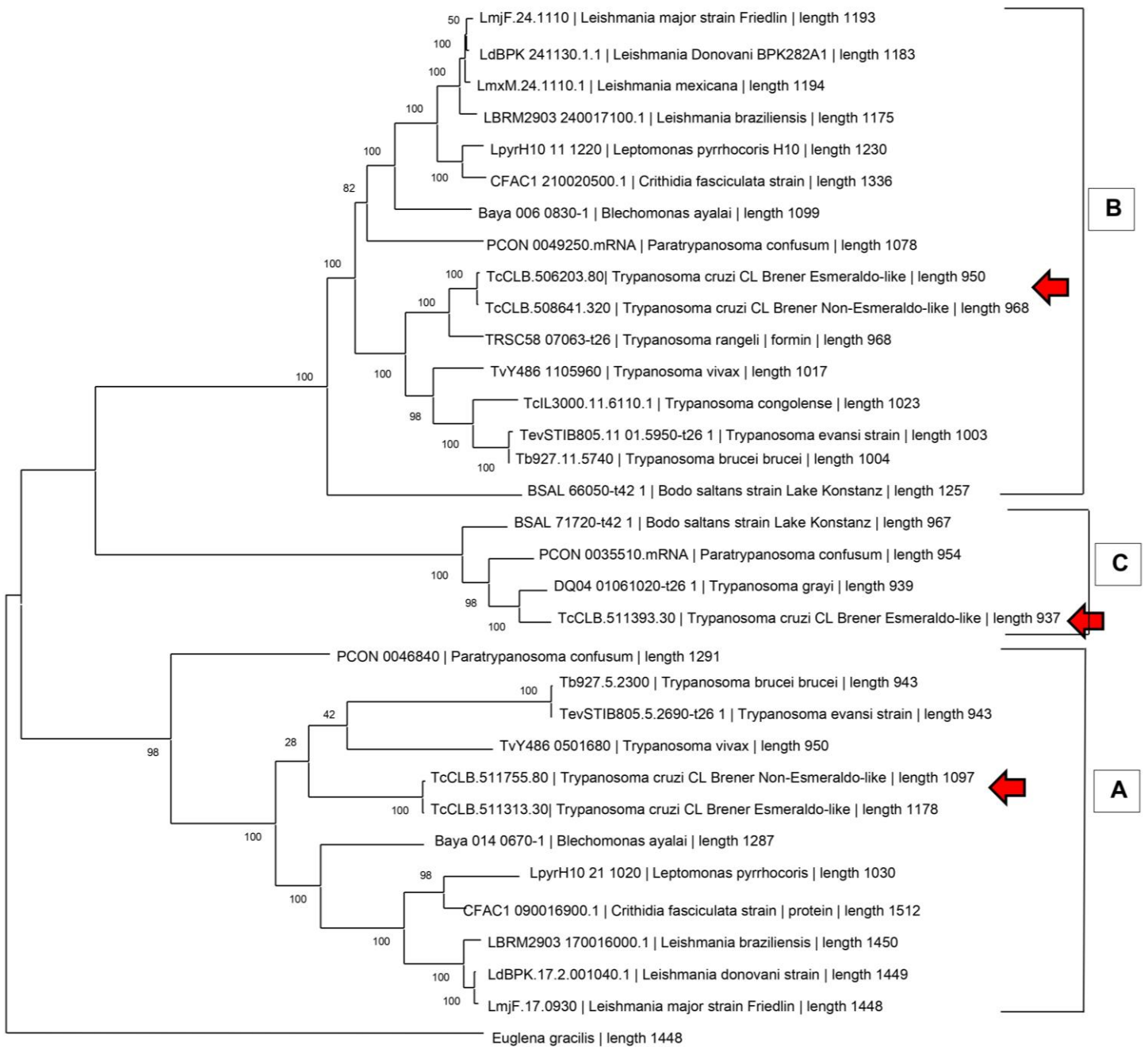
Figura 13 Representación gráfica de los dominios de los 6 genes que codifican para Formina A, B y C en *Trypanosoma cruzi* CL Brener identificados por InterPro y los reportes en la literatura.

Cabe resaltar que los 5 genes de Formina que se encuentran completos contienen todo el dominio FH2 (IPR015425) en el extremo carboxilo terminal del gen y, junto a él, en dirección al extremo amino terminal, se contienen un dominio FH1 (no incluido en la base de datos de InterPro, pero distinguible por la presencia de poliprolinas) (Figura 13). En extremo amino terminal de las Forminas A no se identificaron dominios que estuvieran presentes en la base de datos de InterPro. En contraste, las Forminas B poseen un dominio parecido a PTP (Proteínas Tirosina Fosfatasa, IPR029021), importante en la traducción de señales al reconocer las proteínas tirosín cinasas y, por lo tanto, se podría sugerir que podría ser una región regulatoria de estas Forminas (Tsarouhas *et al.*, 2019). En el caso de la Formina C, InterPro identificó en el extremo amino terminal el dominio GBD_FH3 (IPR014768) típico de las Forminas diafonosas, dentro del cual se encontró una región sobrelapante de tipo armadillo como se ha reportado en la literatura (Rivero *et al.*, 2005; Schulte *et al.*, 2008).

Adicionalmente, en esta búsqueda en el servidor TriTrypDB se identificaron 125 genes en 51 cepas de tripanosomátidos que codifican para Formina A (52 genes), B (57 genes) y C (16 genes) (Anexo 2).

6.2 Elaboración del árbol filogenético de las Forminas A, B y C de *Trypanosoma cruzi* y otros tripanosomátidos

De los 125 genes de Forminas identificados en los tripanosomátidos, se seleccionaron 32 secuencias, las cuales se encontraban completas, para realizar un árbol filogenético. Estas 32 secuencias se encuentran anotadas en los principales representantes de los clados taxonómicos de los tripanosomátidos y corresponden a 12 genes de Forminas A, 16 genes de Forminas B y 4 genes de Forminas C (Anexo 3). Ya que las Forminas son proteínas multidominio y el número y tipo de dominios varían entre ellas, se seleccionó solamente el dominio FH2 para el estudio filogenético debido a que está presente en todas las Forminas descritas hasta el momento y se ha reportado que se ha conservado a través de la evolución (Chalkia *et al.*, 2008).



1

Figura 14 Árbol filogenético del dominio FH2 de las Forminas A, B y C ortólogas identificadas en *Trypanosoma cruzi* CL Brener y otros tripanosomátidos. Los valores de Bootstrap fueron calculados de 1,000 muestreos. Con flechas rojas se señalan las Forminas A, B y C de *Trypanosoma cruzi* CL Brener. Árbol construido con el método de máxima verosimilitud en MEGA 11. Como secuencia raíz se empleó la secuencia del mRNA de la actina codificada por *Euglena gracilis*.

En la figura 14 se observa el árbol filogenético del dominio FH2 de las 32 Forminas completas que son ortólogas en los tripanosomátidos (Anexo 3). En este árbol se identificaron claramente dos clados. En el primer clado se agruparon las Forminas B y C, y en un segundo clado se agruparon solamente las Forminas A. Partiendo de este árbol y teniendo en cuenta la longitud de la rama de la Formina C, se podría plantear la hipótesis de que la Formina C está más relacionada con la Formina B que con la Formina A y que, a su vez la Formina C tiene más variación genética en comparación con la Formina B y A.

6.3 Selección de la Formina de *Trypanosoma cruzi* para estudios posteriores

Teniendo en cuenta que: 1) la Formina C solo está en tripanosomátidos estercorearios y 2) la Formina C del haplotipo S contiene el dominio GBD-FH3, uno de los motivos mejor caracterizados de las Forminas de metazoarios, se decidió iniciar con el estudio de la Formina C del haplotipo S de *T. cruzi* CL Brener como un acercamiento para sugerir su funcionalidad a partir de estudios *in silico*.

6.4 Búsqueda de motivos adicionales en la Formina C

Una vez que se seleccionó la Formina C como objeto de estudio y con el fin de complementar la identificación de los motivos presentes en esta proteína, se analizó su secuencia en el servidor TMHMM2.0 para identificar la presencia de dominios transmembranales. Como se observa en la Figura 15, el análisis probabilístico de la Formina C en este servidor indicó que la proteína no contiene regiones transmembranales, debido a que no se observan picos a lo largo de la secuencia analizada que lleven a una probabilidad cerca o mayor de 1.

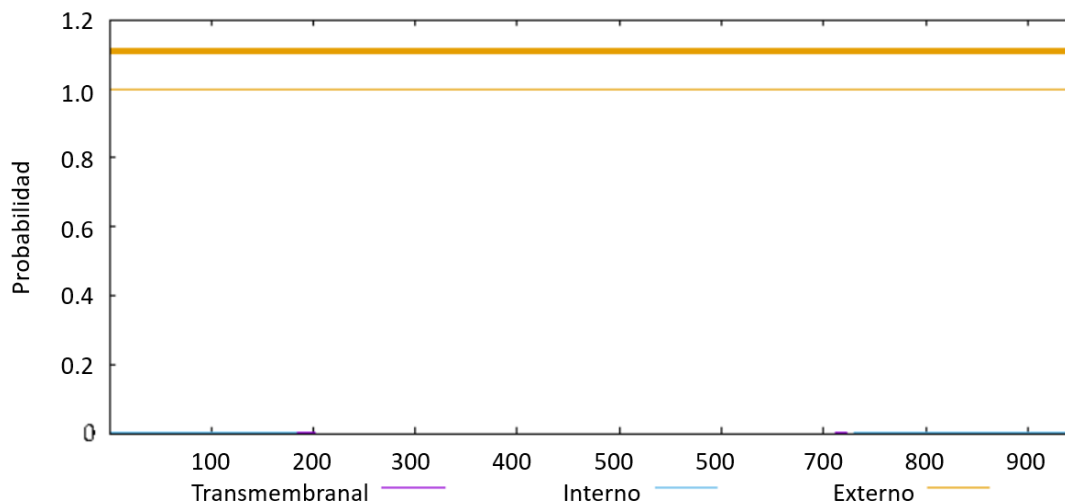


Figura 15 Gráfica de probabilidades de presentar dominios transmembranales en la Formina C de *Trypanosoma cruzi* CL Brener generada por el servidor TMHMM2.0. Líneas moradas: residuos transmembranales; Líneas azules: residuos internos; Líneas amarillas: residuos externos.

6.5 Análisis de la estructura primaria de la Formina C

Como parte inicial de la caracterización de la estructura de la Formina C de *T. cruzi* CL Brener, se analizó su secuencia primaria en el servidor ProtParam. Se identificó que esta proteína tiene 937 residuos de aminoácidos (Tabla 3), un peso molecular de 104606.69 Da, un punto isoeléctrico teórico de 8.24, y un coeficiente de extinción molar que puede variar entre 65570 y 65320 $M^{-1} cm^{-1}$ a 280 nm al ser medido en agua dependiendo del número de cisteínas reducidas.

Tabla 3 Composición de la Formina C de *Trypanosoma cruzi* CL Brener en términos de aminoácidos.

Aminoácido	Número de aminoácidos	Porcentaje del aminoácido
Alanina (A)	82	8.8%
Arginina (R)	46	4.9%
Asparagina (N)	37	3.9%
Aspartato (D)	44	4.7%
Cisteína (C)	4	0.4%
Glutamina (Q)	54	5.8%
Ácido glutámico (E)	72	7.7%
Glicina (G)	50	5.3%
Histidina (H)	19	2.0%

Isoleucina (I)	33	3.5%
Leucina (L)	103	11.0%
Lisina (K)	73	7.8%
Metionina (M)	32	3.4%
Fenilalanina (F)	34	3.6%
Prolina (P)	77	8.2%
Serina (S)	50	5.3%
Treonina (T)	36	3.8%
Triptófano (W)	7	0.7%
Tirosina (Y)	18	1.9%
Valina (V)	66	7.0%
TOTAL	937	100 %

6.6 Análisis de la estructura secundaria de la Formina C

El análisis de la estructura secundaria de la Formina C por medio de RaptorX dividió su secuencia en ventanas de 10 residuos de aminoácidos cada una. Todos los residuos fueron evaluados individualmente y gracias a modelos probabilísticos de los ángulos que forman estos aminoácidos se generó una predicción de su estructura secundaria. Como se puede observar en la figura 16, la Formina C tiene 60% alfa-hélices, 34% asas y 1% de láminas β .

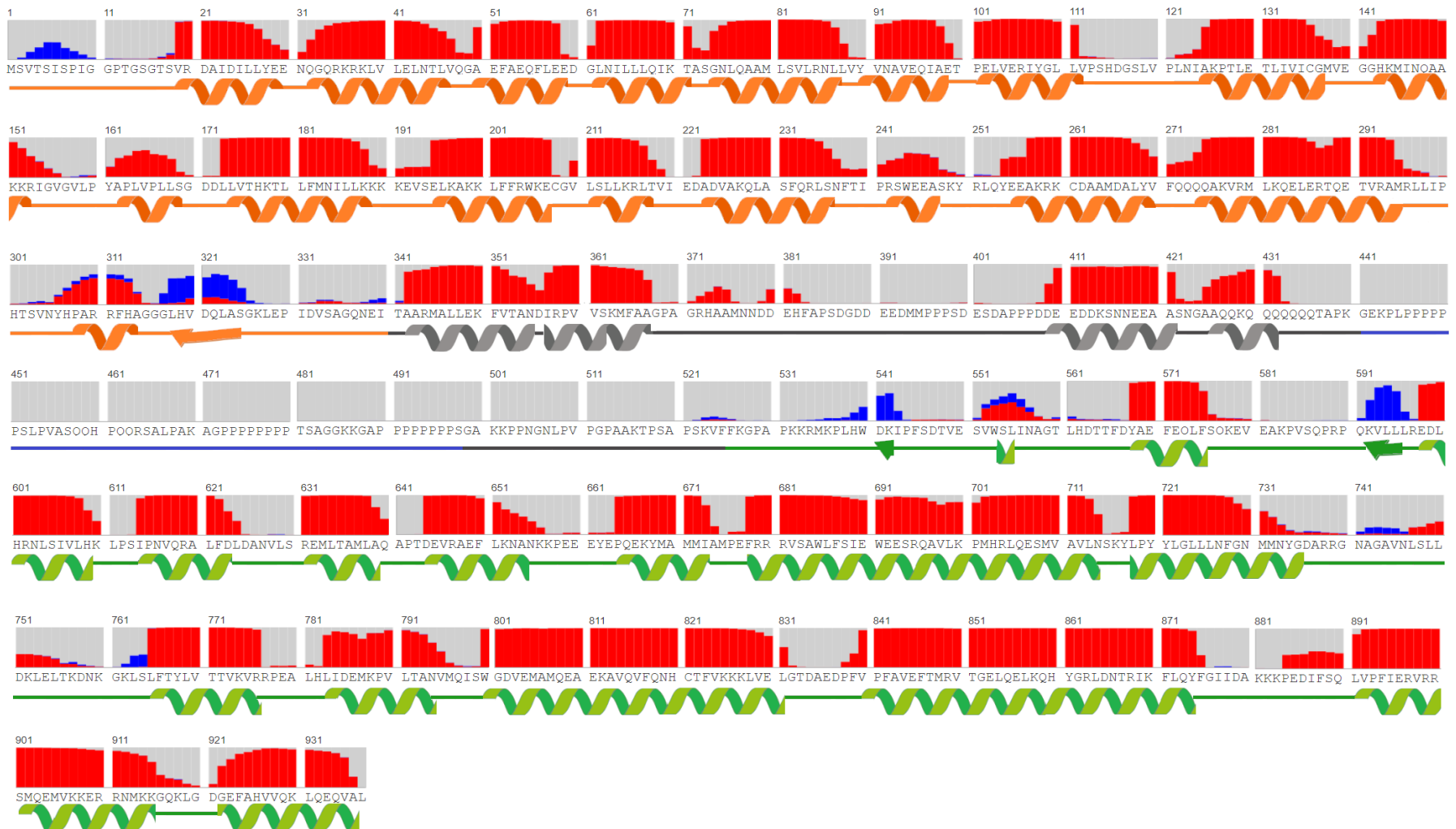


Figura 16 Representación gráfica de la predicción de la estructura secundaria de la Formina C de *Trypanosoma cruzi* generada por el servidor Raptor X. Los dominios identificados por InterPro fueron añadidos a la gráfica. Naranja, dominio GBD_FH3; gris, dominio FH1; verde, dominio FH2.

En la figura 16 se observa que la zona comprendida entre los aminoácidos 431 a 570, la cual que incluye una región rica en poliprolinas (región del dominio FH1), es fundamentalmente deseestructurada como se ha reportado en la literatura (Courtemanche & Pollard, 2012; Paul & Pollard, 2008). Así mismo, se observa que, el dominio FH2 (aminoácidos 525 al 937) está constituido primordialmente de alfa-hélices como se ha reportado en las estructuras cristalizadas de estas regiones (Lu *et al.*, 2007; Shimada *et al.*, 2004; Yamashita *et al.*, 2007). En cuanto al dominio GBD-FH3 ubicado entre los aminoácidos 1 al 339, se observó que está conformado por alfa-hélices y una sola lámina β , lo cual concuerda con lo mencionado en la literatura donde se describe que el dominio GBD, ubicado en el extremo amino del dominio GBD-FH3, puede estar conformado por láminas β (Dames *et al.*, 2011; Schulte *et al.*, 2008) o por alfa-hélices (Rose *et al.*, 2005; Kühn *et al.*, 2015).

Posteriormente, se continuó con la predicción de la estructura terciaria de cada uno de los dominios de la Formina C.

6.6 Predicción de la estructura terciaria de los dominios FH2 y GBD_FH3 la Formina C por homología

La predicción de la estructura terciaria de los dominios FH2 y GBD_FH3 de la Formina C se realizaron por homología empleando el servidor MODELLER. Para esto se buscaron estructuras cristalizadas que pudieran servir como plantillas utilizando el algoritmo de BLAST en NCBI y el algoritmo HHblits en SIWSS-MODEL. Las estructuras que podrían servir como plantillas se encuentran en la tabla 4.

La búsqueda realizada por BLAST en NCBI únicamente identificó estructuras que podrían servir como plantilla para el dominio FH2 (PDB: 1Y64, 1UX5, 2Z6E, 2J1D y 1UX4) con porcentajes de identidad entre el 20,62% y 23,56% (Tabla 4A).

Tabla 4 Plantillas que podrían ser empleadas para el modelado de la estructura tridimensional de la Formina C. A) Plantillas obtenidas por BLAST en NCBI a partir de la secuencia TcCLB.511393.30. B) Plantillas obtenidas por HHBlit en SWISS-MODEL NCBI a partir de la secuencia TcCLB.511393.30. Flecha verde, plantilla utilizada para modelar el dominio FH2; flecha roja, plantilla utilizada para modelar el amino terminal; flecha naranja, plantilla utilizada para modelar la estructura de dominio bobina espiral. Con asterisco se encuentran demarcadas las plantillas identificadas empleando ambos algoritmos.

A) Datos obtenidos por BLAST en NCBI a partir de TcCLB.511393.30									
PDB	Cadena	Descripción de la estructura cristalina	Ubicación	Cobertura	Valor de E	I.D %	Aminoácidos	REGIÓN	
1Y64	B	FH2 (BNI1)	<i>Saccharomyces cerevisiae</i>	31%-C	7.00E-16	22,16%	443	FH2 (526-937)	
1UX5	A	FH2 (BNI1)	<i>Saccharomyces cerevisiae</i>	30%-C	2.00E-14	21,54%	411	FH2 (526-937)	
2Z6E	A	FH2, (DAAM1)	<i>Homo sapiens</i>	40%-C	1.00E-13	23,56%	419	FH2 (526-937)	
2J1D	G	FH2, (DAAM1)	<i>Homo sapiens</i>	40%-C	2.00E-13	23,56%	483	FH2 (526-937)	
1UX4	A	FH2 (BNI1)	<i>Saccharomyces cerevisiae</i>	30%-C	2.00E-13	20,62%	410	FH2 (526-937)	

B) Datos obtenidos por HHBlit en SWISS-MODEL a partir de TcCLB.511393.30									
PDB	Cadena	Descripción de la estructura cristalina	Ubicación	Cobertura	I.D%	GMQE	QSQE	Aminoácidos	REGIÓN
2z6e *	A	FH2 (DAAM1)	<i>Homo sapiens</i>	40%-C	25.62	0.24	0.34	383	FH2 (526-937)
2j1d *	A	FH2 (DAAM1)	<i>Homo sapiens</i>	41%-C	25.62	0.25	0.32	375	FH2 (526-937)
1ux5 *	A	FH2 (BNI1)	<i>Saccharomyces cerevisiae</i>	30%-C	24.56	0.13	0.33	289	FH2 (526-937)
1y64 *	B	FH2 (BNI1)	<i>Saccharomyces cerevisiae</i>	31%-C	23.53	0.12	/	298	FH2 (526-937)
1ux4 *	A	FH2 (BNI1)	<i>Saccharomyces cerevisiae</i>	40%-C	17.21	0.22	0.32	388	FH2 (526-937)
3o4x	C	FH2 (mDia1)	<i>Mus musculus</i>	40%-C	20.75	0.29	0.3	381	FH2 (526-937)
4eah	B	FH2 (FMNL3)	<i>Mus musculus</i>	39%-C	21.92	0.29	0.37	383	FH2 (526-937)
1v9d	A	FH2 (mDia1)	<i>Mus musculus</i>	34%-C	21.38	0.21	0	327	FH2 (526-937)
4ydh	C	N-terminal (FMNL1)	<i>Homo sapiens</i>	20%-N	22.11	0.1	0.03	205	FH3 (1-237)
3dad	A	N-terminal (FHOD1)	<i>Homo sapiens</i>	23%-N	23.01	0.1	0.03	220	FH3 (1-237)
6ys4	A	Estructura de dominio bobina espiral	<i>Homo sapiens</i>	46%-N	21.43	0.2	0.2	55	FH3 (238-360)
4rnd	C	Estructura de bomba ATP	<i>Saccharomyces cerevisiae</i>	44%-N	24.07	0.19	0.18	53	FH3 (238-360)
5voy	M	Estructura de bomba ATP	<i>Saccharomyces cerevisiae</i>	44%-N	24.07	0.17	/	53	FH3 (238-360)

Por otra parte, utilizando el algoritmo Hhblits en SIWSS-MODEL no solo se identificaron las mismas 5 platillas identificadas en NCBI, sino también se encontraron dos estructuras cristalizadas (PDB: 4YDH y 3DAD) que podrían servir como plantilla para el modelado de la región amino de la Formina C entre los aminoácidos 1 a las 237, y tres estructuras cristalizadas (PDB: 6YS4, 4RND y 5VOY) que servirían como plantilla para la región comprendida entre los aminoácidos 238 a 360 (Tabla 4B).

Teniendo en cuenta la calidad de la estructura cristalizada y los porcentajes de cobertura y de identidad de cada plantilla identificado, se seleccionaron las estructuras cristalizadas PDB 2Z6E, 3DAD y 6YS4 para ser empleadas como plantilla en los ensayos de modelado del dominio FH2 y las regiones correspondientes a los aminoácidos 1 al 237 y 238 al 360, respectivamente.

Cada una de estas plantillas fue utilizada en MODELLER para generar 10,000 predicciones de manera individual de cada región a modelar. Se seleccionó la mejor predicción de cada región teniendo en cuenta los parámetros energéticos obtenidos.

El modelo óptimo del dominio FH2 que se obtuvo se denominó MOD3-REGIONFH2 (Figura 17). En este modelo se pueden identificar los subdominios característicos de esta región que son el lazo, el poste, la bobina enrollada y el botón, los cuales se organizan de manera canónica (Lu *et al.*, 2007). Adicionalmente, se generó una estructura dimérica, la cual sugiere que la unión entre los hemidímeros de FH2 se da predominantemente a través del lazo y el enlazador de cada monómero, como se ha encontrado en las estructuras cristalizadas reportadas en la literatura (Lu *et al.*, 2007; Shimada *et al.*, 2004; Yamashita *et al.*, 2007). En el dímero del dominio FH2 (Figura 17) también se observa una configuración “abierta”, la cual podría permitir la unión de subunidades de actina en su interior.

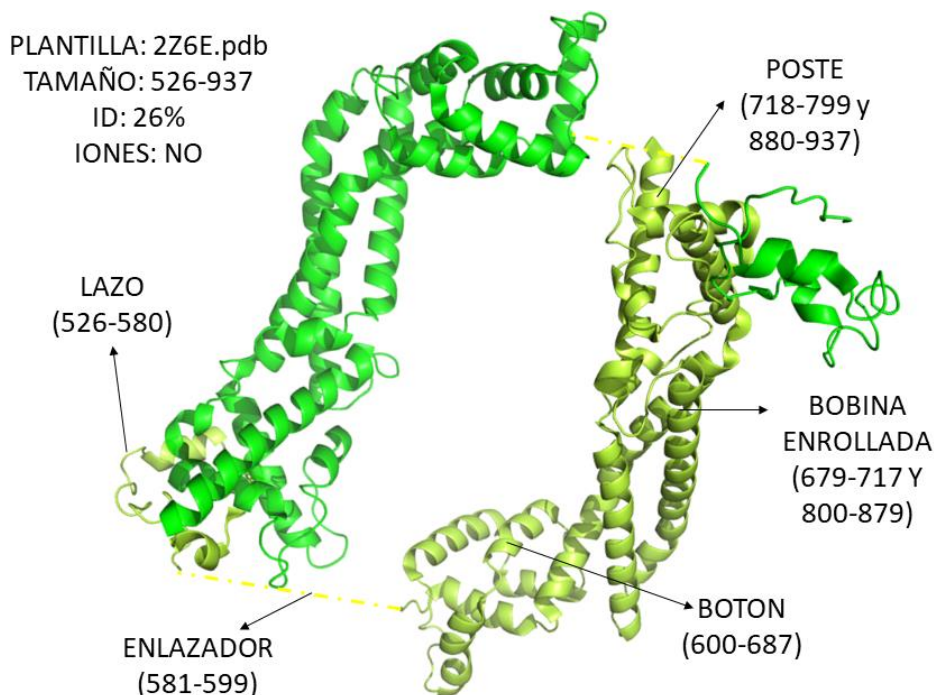


Figura 17 Configuración abierta del dímero del dominio FH2 de la Formina C modelado a partir de la plantilla PDB 2Z6E por MODELLER (Modelo MOD3-REGIONFH2). En el modelo obtenido se observa la posición de los subdominios lazo, enlazador, botón, bobina y poste. El enlazador que actúa generando una unión flexible entre los dímeros se representa en líneas punteadas. En la parte inferior izquierda de la figura se mencionan las características de la plantilla (molde) empleada para generar el modelo.

Con respecto al dominio GBD/FH3 presente en las Forminas diáfanas, se ha reportado que está compuesto por un dominio de unión a Rho GTPasa ubicado en el extremo amino terminal (GBD) y un dominio FH3, conformado por una región de repetición de armadillo (ARM), que se unen de alguna manera al extremo carboxilo terminal del Dominio Autorregulador Diáfano (DAD) para mantener la proteína en una conformación inactiva (Rose *et al.*, 2005).

El modelado del dominio GBD-FH3 de la Formina C se realizó utilizando como plantilla la estructura cristalizada PDB 3DAD (región amino terminal de la Formina FHOD1 humana) (Figura 18A), la cual tiene un dominio de unión a GTPasa y una región de repetición de armadillo, la cual podría estar participando en procesos de localización subcelular y regulación de la actividad de la Formina (Choi & Weis, 2005; Eng *et al.*, 2017).

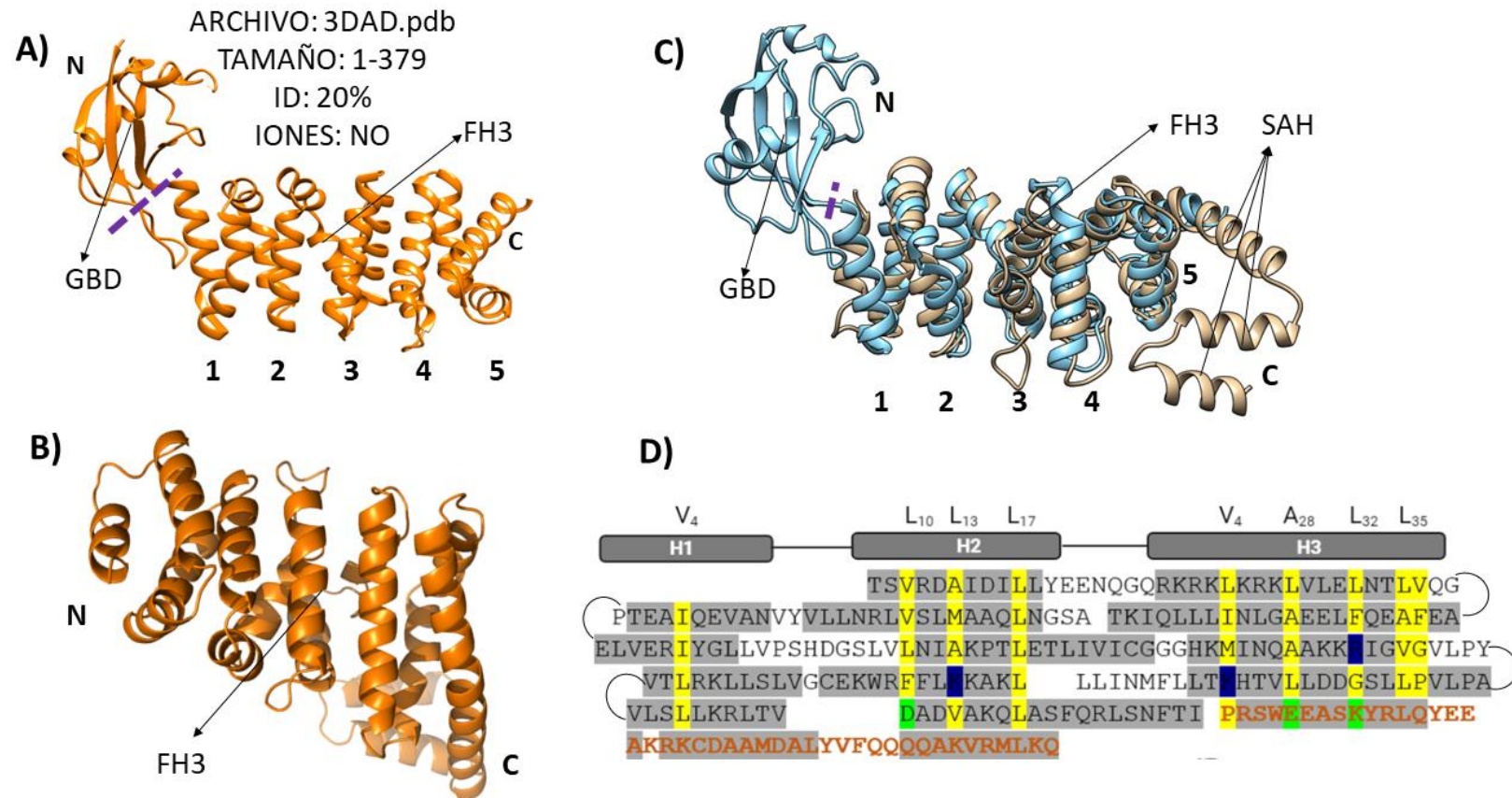


Figura 18 Análisis estructural del dominio GBD-FH3 de la Formina C. A) Plantilla (molde) PDB 3DAD con 379 aa empleado para el modelado por homología. B) Predicción del dominio GBD-FH3 de la Formina C realizado con MODELLER (modelo MOD1-MODELLER-SWISS-MODEL) con 339 aa. C) Comparación de las estructuras tridimensionales de la plantilla PDB 3DAD (estructura en cintas azules) y el modelo MOD1-MODELLER-SWISS-MODEL de la Formina C (estructura en cintas grises). dominios de hélice alfa única (SAH) estables se encuentra en la región carboxilo del modelo MOD1-MODELLER-SWISS-MODEL de la Formina C. D) Alineamiento de la secuencia del dominio FH3. Las cajas grises superiores corresponden a los motivos alfa-hélice conocidos como H1, H2 y H3. Los residuos de color amarillo son aminoácidos hidrofóbicos característicos de la región de repetición de armadillo (ARM), los residuos en color azul son hidrófilos, los residuos en color verde son los 3 aminoácidos más conservados en un motivo ARM [D307, E331y K335], los residuos en color naranja corresponden a la predicción del motivo de dimerización.

En el modelo óptimo que se obtuvo del dominio GBD-FH3 de la Formina C (Figura 18B) se identificó el motivo de repetición de armadillo (ARM), presente en el dominio FH3, pero no el dominio GBD, como era de esperarse. El motivo armadillo de la Formina C contiene 5 repeticiones individuales y superpuestas de alfa-hélices como se ha reportado para la Formina FHOD1 humana (PDB: 3DAD) (Figura 18C). La ausencia del dominio GBD podría sugerir que la regulación de la Formina C en estudio no depende de su interacción con una GTPasa, sino que podría depender de su interacción con regiones desordenadas del citoesqueleto (Gull *et al.*, 2017). Estos resultados evidencian la importancia de realizar análisis experimentales más detallados al estudiar una proteína de interés ya que, aunque los servidores identifiquen la presencia de dominios proteicos a partir del análisis de la estructura primaria, estos dominios pueden no estar completos o pueden no ser funcionales.

En la figura 18D se esquematiza un alineamiento de las 5 alfa-hélices que conforman el dominio FH3 de la Formina C, en el que se identifican los aminoácidos hidrofóbicos y conservados de un motivo ARM. Teniendo en cuenta los resultados obtenidos en este alineamiento y lo reportado en la literatura (Schulte *et al.*, 2008; Higgs & Peterson, 2005), se sugiere verificar experimentalmente si el dominio FH3 es funcional sólo al estar conformado por el motivo ARM.

6.7 Modelaje *ab-initio* de la estructura terciaria de la región rica en prolinas del dominio FH1 por Rosetta

Teniendo en cuenta que no se encontró ninguna plantilla para realizar el modelado de la estructura terciaria del dominio FH1, se realizó el modelaje *ab-initio* de esta región por medio de Rosetta. Para esto se modeló una región de 60 residuos de aminoácidos (aa 440 - 499) y se obtuvo la mejor predicción basados en la menor energía de un resultado de 6000 predicciones (Figura 19). Al ser una región con aminoácidos repetidos (región rica en prolinas) la predicción evidenció una zona desordenada, generando un motivo en forma de lazo.

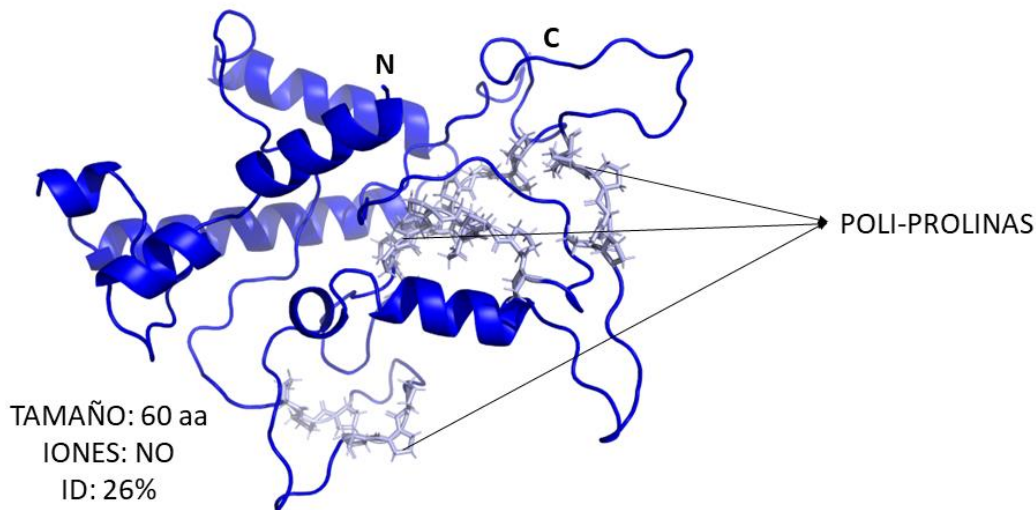


Figura 19 Predicción del dominio FH1 de la Formina C. La secuencia de aminoácidos empleada para el modelado se encuentra en la parte inferior izquierda de la figura. El modelo del dominio FH1 se presenta en cinta y los residuos de prolina se presentan en lazos.

6.8 Modelado de la Formina C por MODELLER

Una vez se obtuvo por separado las predicciones de las 3 regiones de Formina C modeladas (FH2, GBD-FH3 y FH1), se procedió a unir los fragmentos por medio de MODELLER para diseñar un solo modelo completo. Para este análisis se realizaron 10,000 predicciones con el fin de aumentar la probabilidad de identificar la estructura de la Formina C con el valor energético más bajo. Este ensayo es similar a los realizados con los fragmentos predichos anteriormente, puesto que se desea que la macromolécula tenga una alta estabilidad estructural viéndose reflejada en un modelo con una mínima energía cinética.

La predicción de la estructura terciaria de la Formina C con el valor energético más bajo se observa en la figura 20. En la estructura predicha se observa en color naranja el dominio GBD-FH3 con la región ARM, la cual está formada por 5 alfa-hélices. En color azul se observa el dominio FH1 con la región de poli-prolinas. En verde se encuentra señalado el dominio FH2, en el cual se demarcan las regiones de lazo, enlazador y poste.

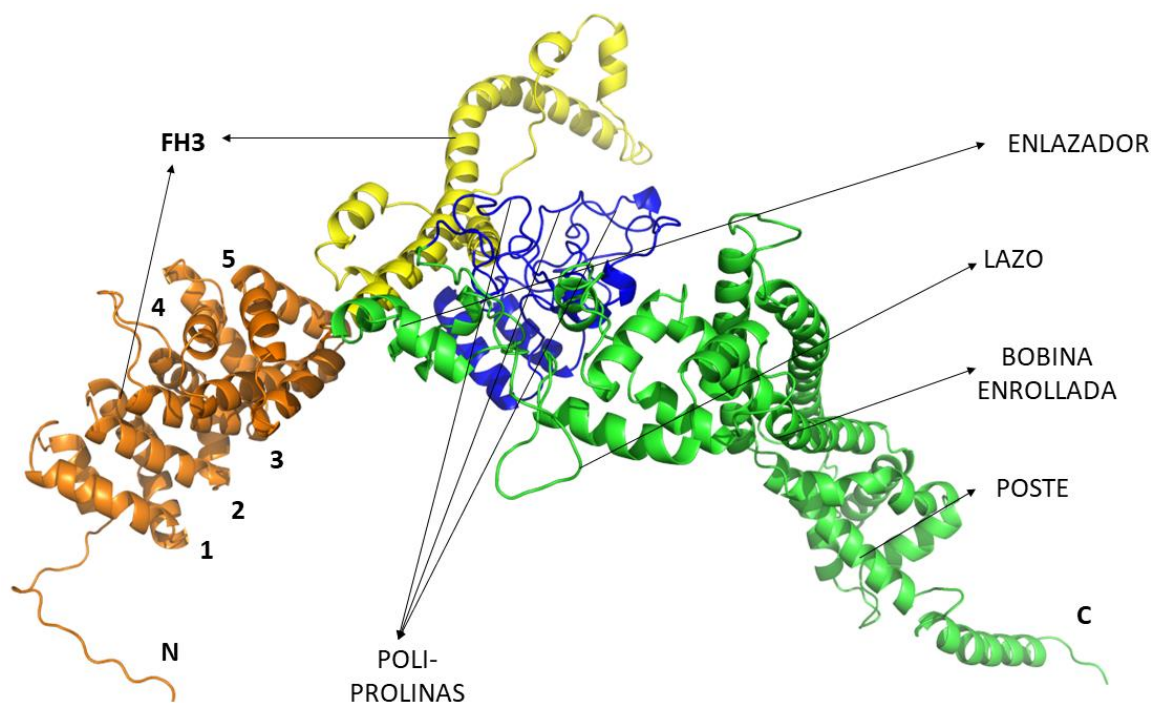


Figura 20 Predicción de la estructura terciaria completa de la Formina C de *Trypanosoma cruzi* CL Brener (937 aa). Este es el modelo con menor energía predicho empleando MODELLER. En la región amino (naranja) se enumeran los 5 motivos que contiene la región armadillo seguido por una alfa hélice que comunica con el primer dominio descrito de la Formina rico en prolinas FH1, en la región carboxilo (de color verde) se encuentra la región FH2 descrita como el posible dominio de unión con las actinas

6.9 Verificación de la estructura terciaria predicha para la Formina C de *Trypanosoma cruzi*

Con la finalidad de mejorar el plegado del modelo completo de la Formina C y acercarlo a una posible conformación nativa se realizó una dinámica molecular. Los resultados de este ensayo evidenciaron que la estructura tridimensional de la Formina C es estable.

Seguido, se verificó la calidad de la estructura terciaria de la Formina C completa por medio de los servidores SAVES v.6.0 y Rosetta Desing. Los resultados de estos análisis se presentan en la tabla 5.

El servidor SAVES v.6.0 evalúa la calidad de una estructura terciaria a partir de 5 programas (ERRAT, VERIFY 3D, PROVE, PROCHECK Y WHATCHECK) en las

que se analiza la geometría de la proteína y de cada uno de los residuos. Los programas ERRAT y WHATCHECK aceptaron el modelo de la Formina C, mientras que los programas VERIFY y PROVE lo rechazan. Por otra parte, aunque el programa PROCHECK generó una advertencia al evaluar la calidad de la predicción, este modelo fue aceptado (Tabla 5-1).

De manera general, 3 de 5 programas aprobaron la calidad del modelo evaluado, sin embargo, la aceptación de la estereoquímica y geometría de los residuos de aminoácidos por parte de los 3 programas no permite concluir que sea un modelo de Formina, pues puede existir un modelo de estructura terciaria con una estereoquímica aceptable pero que no corresponde a una Formina. Por esta razón se utilizó Rosetta Desing, el cual permite establecer si el modelo de Formina C evaluado corresponde realmente a una Formina como se reporta en la literatura.

En contraste, el análisis del modelo de Formina C por el servidor Rosetta Desing generó un puntaje de 77 sobre 100, sugiriendo que el modelo predicho tiene una calidad aceptable y podría corresponder a una Formina DAAM1 (Disheveled-associated activator of morphogenesis 1) (Tabla 5-2).

Teniendo en cuenta los resultados de los estos dos servidores se sugiere que el modelo predicho de la Formina C de *T. cruzi* CL Brener tiene una calidad estereoquímica aceptable y, por tanto, este modelo podrá ser empleado para los siguientes ensayos de acoplamiento y dinámica molecular.

Tabla 5 Verificación de la estructura terciaria de la Formina C completa predicha. 1, Resultados de la verificación por el servidor SAVES v.6.0; 2, Resultados de la verificación por el servidor Rosetta Desing.

1. Verificación por SAVES

TRABAJO	ERRAT	VERIFY	PROVE	WHATCHECK	PROCHECK
ForminaC_cruz i.pdb	89.67	ERROR	ERROR	ACEPTADO	ADVERTENCIA

2. Verificación para Rosetta Desing

NOMBRE DE SECUENCIA:	Formin_C_amber			TAMAÑO:	937
SIMILITUD	E-VALUE	PUNTAJE	SECUENCIA	DESCRIPCIÓN	
MUS MUSCULUS	4.40E-19	77.2	Q8BPM0/DAAM1_MOUSE	DISHEVELED ASSOCIATED ACTIVATOR OF MORPHOGENESIS 1	
HOMO SAPIENS	1.6 E-18	75.3	Q9Y4D1/DAAM1_HUMAN	DISHEVELED ASSOCIATED ACTIVATOR OF MORPHOGENESIS 1	
XENOPUS LAEVIS	2.90E-14	61.2	B0DOB5/DAA1A_XENLA	DISHEVELED ASSOCIATED ACTIVATOR OF MORPHOGENESIS 1	
HOMO SAPIENS	2.80E-13	58	Q86T65/DAAM2_HUMAN	DISHEVELED ASSOCIATED ACTIVATOR OF MORPHOGENESIS 2	
MUS MUSCULUS	1.30E-12	55.8	Q80V19/DAAM2_MOUSE	DISHEVELED ASSOCIATED ACTIVATOR OF MORPHOGENESIS 2	
GALLUS GALLUS	9.90E-12	52.8	AQA1D5P556/DAAM_CHICK	DISHEVELED ASSOCIATED ACTIVATOR OF MORPHOGENESIS 2	

6.10 Acoplamiento y dinámica molecular del inhibidor SMIFH2 con la Formina C

Al modelo de la Formina C generado en este estudio se le realizó un acoplamiento molecular con el ligando SMIFH2 en el programa AutoDock4, con la finalidad de determinar que residuos de aminoácidos se podrían unir entre estas moléculas. Para el análisis por AutoDock4, inicialmente se hizo un espacio de exploración en ciego entre el ligando y la macromolécula, debido a que el conocimiento previo sobre la interacción de los átomos de estas moléculas es nulo. Esta exploración en ciego se realizó por medio de una rejilla o GRID con ayuda de la herramienta de AutoDokTools donde se generaron 1000 conformaciones diferentes. Tras realizar el acoplamiento, se seleccionó la interacción más estable teniendo en cuenta la cantidad de confómeros y la mayor energía de unión. Posteriormente, se realizó una dinámica molecular para evaluar la estabilidad de la unión.

En la figura 21 se presentan los resultados del acoplamiento y la dinámica molecular entre la Formina C y el ligando SMIFH2, donde se observa que la interacción se mantiene estable hasta los 50 ns (Figura 21B-D). En la figura 21A se observa a una distancia de 4 Å que el ligando interacciona con el dominio FH2 (cintas verdes) y el dominio FH1 azul. En la figura 21B-D se sugiere que los aminoácidos de la Formina C que podrían estar participando en su unión con el ligando son: Pro-405, Pro-406, Pro-407, Asn-423, Ala- 427, Glu-465, Ala-500, Asp-644, Phe-650.

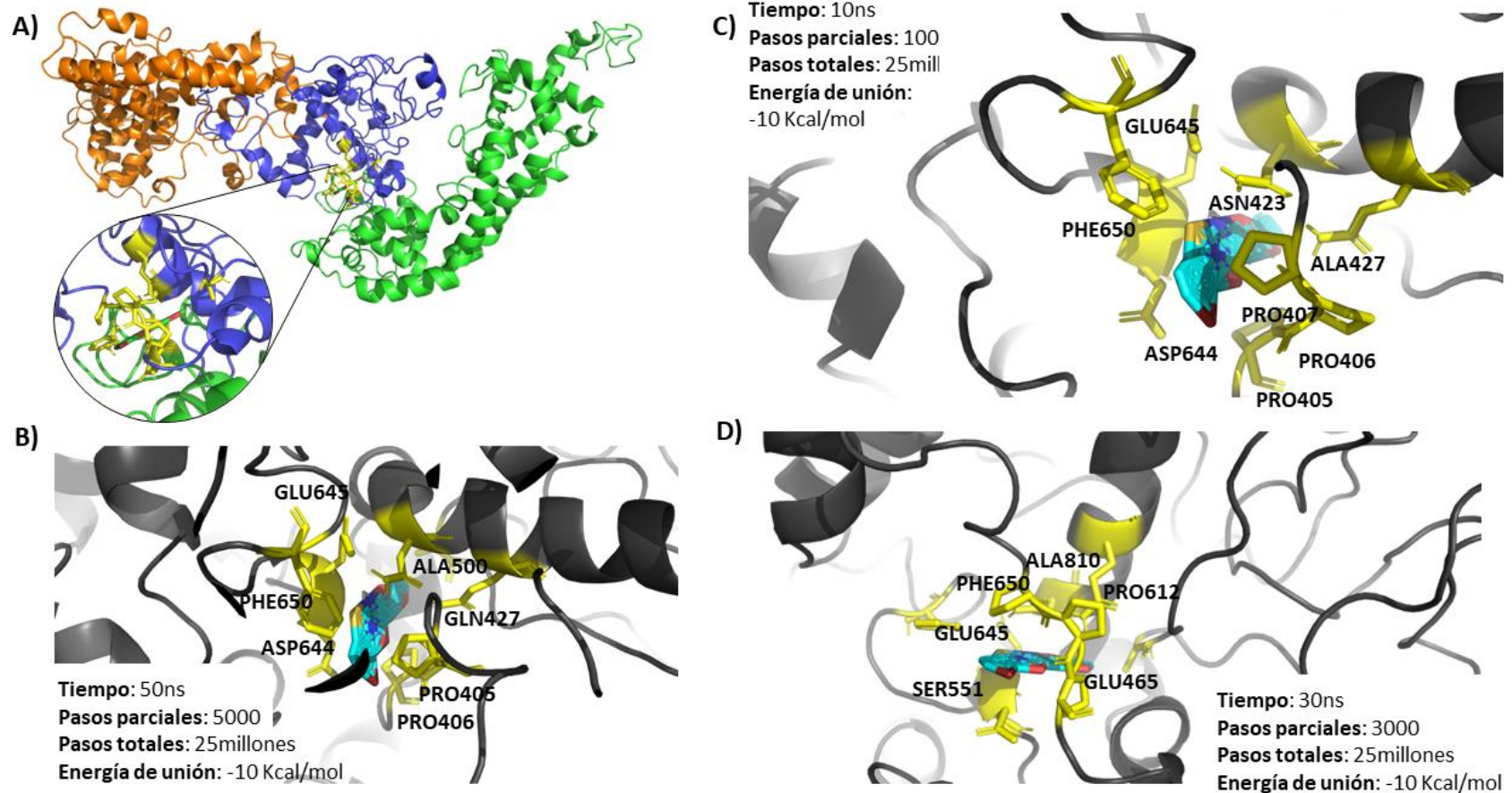


Figura 21 Resultados del acoplamiento molecular entre la macromolécula Formina C predicha en este estudio y el ligando SMIFH2. A) Modelo estructural de la Formina C unida al ligando SMIFH2 con vista detallada del sitio de unión del ligando. La estructura de la Formina C se presenta en cintas, los aminoácidos que participan en la unión al ligando se presentan en bastones y el inhibidor SMIFH2 se presenta en bastones de color verde. B) Ligando SMIFH2 unido a los 10 ns a la Formina C. C) Ligando SMIFH 2 unido a los 30 ns a la Formina C. D) Ligando SMIFH unido a los 50 ns a la Formina C. Para las imágenes B, C y D, la estructura de la Formina C se presenta en cintas, los aminoácidos que participan en la unión al ligando se presentan en bastones y el inhibidor SMIFH2 se presenta en color celeste.

Para el análisis de la RMSD (Raíz de la Desviación Media Cuadrática), RMSF (Raíz de la Fluctuación Cuadrática Media) y de contactos entre residuos, se utilizaron las herramientas de análisis AmberTools20 y CPPTRAJ. El RMSD de la unión Formina C – Ligando SMIFH2 muestra que al parecer alcanzó el equilibrio en el periodo de tiempo simulado (50 ns) y que el ligando permaneció en la región amino del dominio FH2 con contactos en la región FH1 (Figura 22C). Al examinar el RMSF se identificaron regiones móviles que fueron descritas como lazos que unen las alfa-hélices, las cuales están presentes principalmente en la región FH1 y en la región enlazadora del dominio FH2 (Figura 22F).

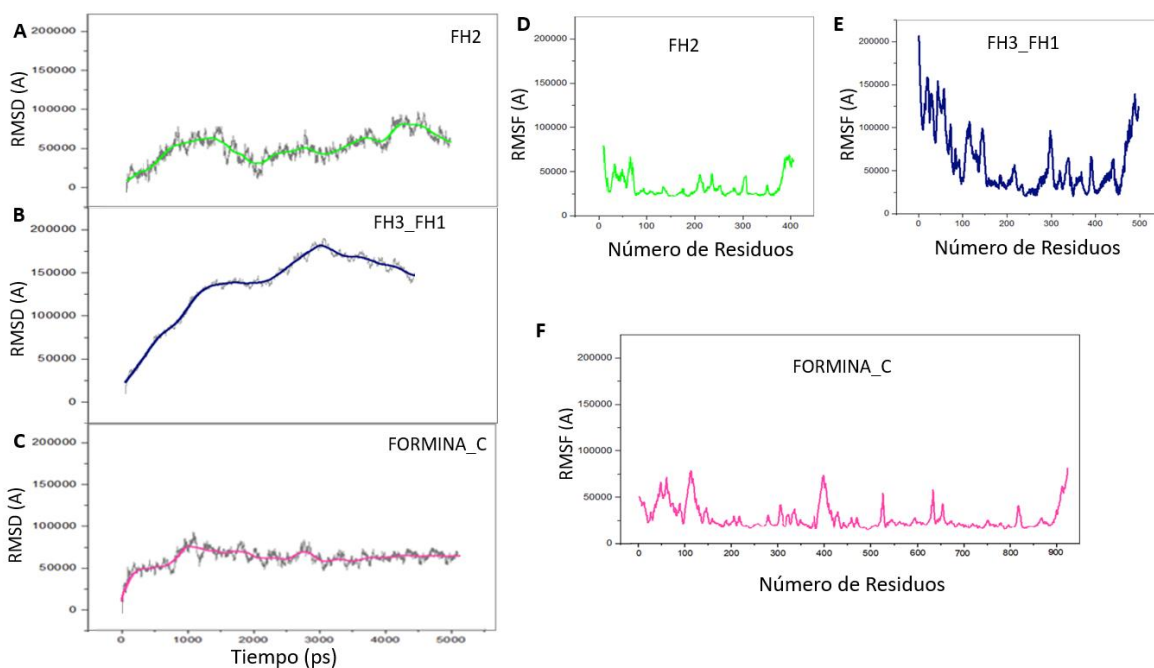


Figura 22 Validación del acoplamiento molecular. A) RMSD de la interacción dominio FH2 y ligando SMIFH2; B) RMSD de la interacción dominio FH3-FH1 y ligando SMIFH2; C) RMSD de la interacción Formina C y ligando SMIFH2; D) RMSF de la interacción dominio FH2 y ligando SMIFH2; E) RMSF de la interacción dominio FH3-FH1 y ligando SMIFH2; F) RMSF de la interacción Formina C y ligando SMIFH2.

Con respecto al análisis de contactos entre las dos moléculas, en la figura 23 se presentan los resultados de la interacción de los residuos de aminoácidos de la Formina C con el ligando SMIFH2 durante 4500 Ps. En el eje de las Y se observa que, los residuos de la Formina C que interactúan principalmente con el ligando se

encuentran ubicados en el dominio FH2 y FH1. Las interacciones constantes entre los residuos se evidencian en la figura por una línea continua en el eje de la X, en tanto las pequeñas interacciones se representan por una cruz.

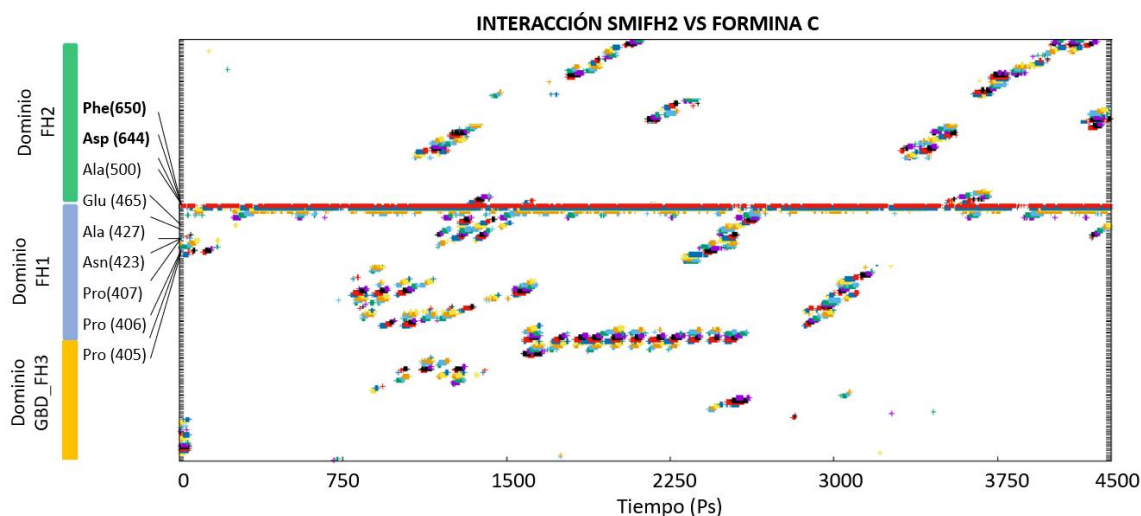


Figura 23 Resultados de la interacción de los residuos de la Formina C con el ligando SMIFH2. En el eje de las Y se observan los residuos de aminoácidos de la macromolécula que interacción con el ligando a medida que pasa el tiempo hasta 4500 Ps.

6.11 Acoplamiento y dinámica molecular del inhibidor SMIFH2 con el dominio FH3_FH1

Al identificar que en la Formina C el posible lugar de unión del Ligando SMIFH2 es el dominio FH1, se decidió realizar un acoplamiento dirigido entre el dominio FH3_FH1 y el ligando por medio de las herramientas del paquete de AutoDockTools, en el cual se generaron 1000 conformaciones diferentes. Tras el acoplamiento y según la cantidad de confórmeros y a la mayor energía de unión, se seleccionó la interacción más estable para ser corroborada por una dinámica molecular. Los resultados de la dinámica molecular arrojaron que, a pesar de que el ligando es capaz de unirse a los residuos de aminoácidos del dominio FH1 al comienzo de la simulación, esta interacción no es suficientemente estable para continuar con la unión durante los 5000 Ps evaluados. El criterio de evaluación del RMSD evidencia una variación en la energía y una poca estabilidad entre la unión a medida que pasa

el tiempo (Figura 22B). Con respecto al RMSF, se muestra que el dominio FH3_FH1 presentan una alta flexibilidad. Esto se corrobora con la estructura terciaria de este dominio puesto que la región FH1 es una estructura desordenada y genera mayor vibración de los átomos en la secuencia de aminoácidos.

6.12 Acoplamiento y dinámica molecular del inhibidor SMIFH2 con el dominio FH2

Con la finalidad de analizar con más detalle la interacción que ocurre entre el ligando SMIFH2 y la Formina C, se realizó un acoplamiento y dinámica molecular empleando únicamente el ligando y el dominio FH2. El acoplamiento realizado fue dirigido hacia este dominio y se generaron 1000 conformaciones diferentes. En la figura 24 se observa los resultados del acoplamiento donde el ligando se une al dominio FH2 por el extremo amino terminal donde se encuentran los subdominios lazo y enlazador.

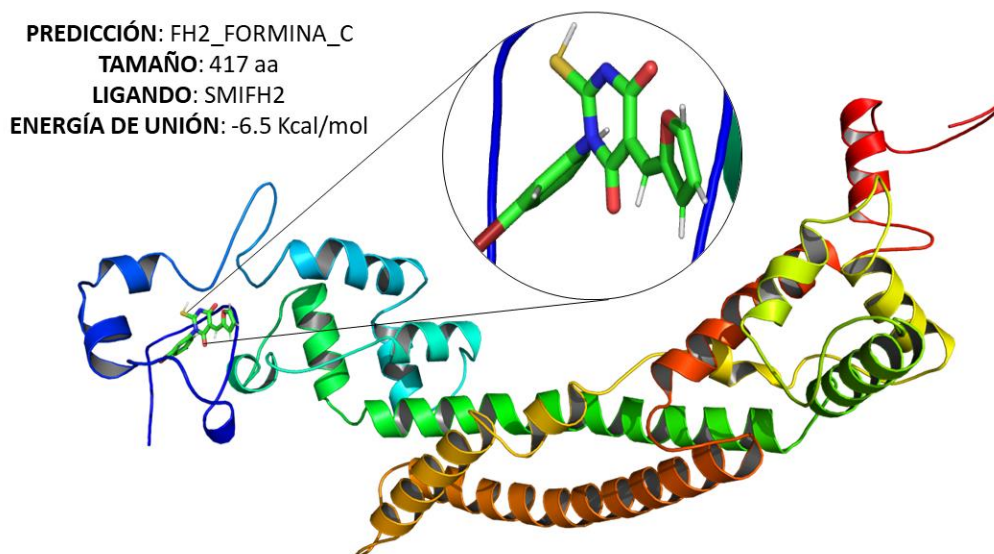


Figura 24 Esquema del sitio de unión del ligando SMIFH2 en el dominio FH2 de la Formina C. La estructura del dominio FH2 se presenta en cintas y el ligando se presenta en bastones.

Tras obtener los resultados de AutoDock4 y según la cantidad de conformeros y la mayor energía de unión, se seleccionó la interacción más estable para ser corroborada por una dinámica molecular.

En la figura 25 se evidencia que la unión ligando – dominio FH2 se mantiene estable hasta los 50 ns, lo que sugiere que tanto la estructura de la Formina C como la interacción entre el ligando SMIFH2 y el dominio FH2 son estables.

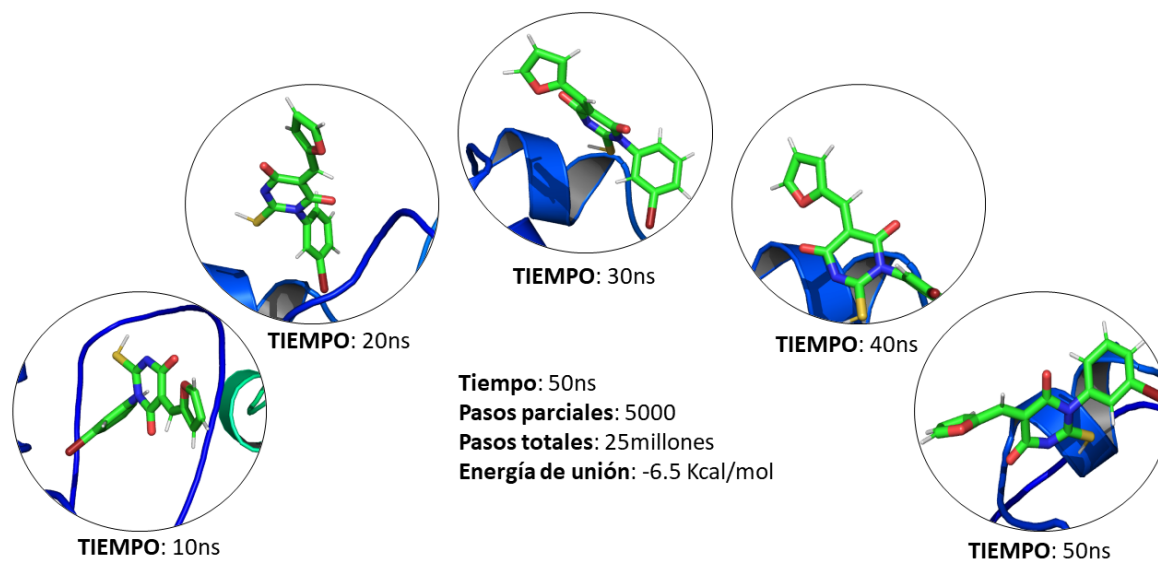


Figura 25 Resultados de la dinámica molecular del dominio FH2 de la Formina C y el ligando SMIFH2 durante 50 ns.

En el análisis de RMSD se observó una variación de energía sin un mínimo local aparente. Esta variación hace inferir en la necesidad de aumentar el tiempo de la dinámica molecular, ya que al parecer 5000 Ps no son suficientes para ver la estabilidad de la unión (Figura 22A). Así mismo, la RMSD calculada para la interacción FH2 y residuo Prolina_FH2 con el inhibidor SMIFH2 indicó que el complejo permanecía unido durante los 50 ns de simulación. Por otra parte, el análisis de RMSF muestra la flexibilidad y la amplia vibración que se presentan en los lazos que unen las alfa-hélices que se encuentran en la región enlazadora y el poste del dominio FH2 (Figura 22D).

Los resultados del análisis de contactos entre el dominio FH2 y el ligando, representados en la figura 26, evidencian los aminoácidos que interactúan principalmente con el ligando a lo largo del tiempo (4500 Ps).

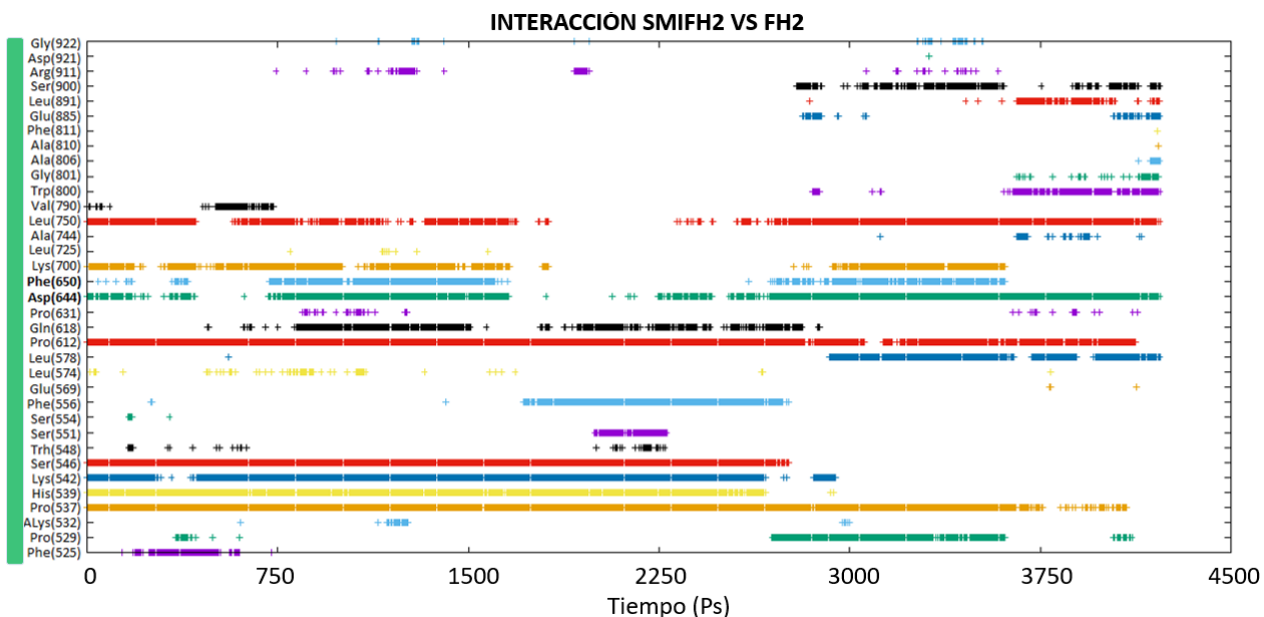


Figura 26 Resultados de la interacción de los residuos del dominio FH2 de la Formina C con el ligando SMIFH2. En el eje de las Y se observan los residuos de aminoácidos de la macromolécula que interacción con el ligando a medida que pasa el tiempo hasta 4500 Ps.

Al correlacionar los resultados del análisis de contactos entre el dominio FH2 y el ligando contra la Formina C y el ligando, se identificó que los aminoácidos que se mantienen constantes en las interacciones son Asp (644) y Phe (650), y que el sitio de la localización de la unión entre las dos moléculas permanece constante. Estos resultados soportan la hipótesis de que, el ligando SMIFH2 se une al dominio FH2 de la Formina C, pero necesita el dominio FH1 para su estabilidad.

7. CONCLUSIONES

Aunque el avance en la biotecnología ha permitido que en los últimos años se haya logrado secuenciar gran cantidad de genes y/o genomas, sigue siendo un reto el comprender cómo las proteínas interactúan para crear y mantener la vida. Los estudios *in vitro* forman una parte esencial para predecir cómo funcionan, se comportan e interactúan las diferentes macromoléculas que conforman una célula.

Los resultados obtenidos hasta el momento sugieren por primera vez, que la unión entre el inhibidor SMIFH2 y la Formina C se da mediante el dominio FH2. No obstante, se deben realizar ensayos experimentales, como mutaciones dirigidas, para establecer exactamente cuáles son los aminoácidos que juegan un papel importante en el sitio de unión. Por otra parte, el tener un modelo aceptable de la estructura terciaria de la Formina C facilitará realizar ensayos *in silico* para sugerir posibles interacciones entre la Formina C y otras proteínas importantes para el funcionamiento y la regulación del citoesqueleto, con la finalidad de obtener información que ayude a entender el papel biológico de esta proteína en el citoesqueleto de los tripanosomátidos estercorearios.

En este trabajo demostramos que hacer una búsqueda de los dominios y estructuras homologas de Formina C fue esencial porque se demostró que el dominio GBD/FH3 identificado en InterPro, carece de la región de unión a GTPasas (dominio GBD), y que por lo tanto su regulación dependerá de otras vías de señalización. Se identificaron una región armadillo en el modelo de la Formina C predicho, los cuales podrían estar involucrados en procesos de localización subcelular, regulación de su actividad y dimerización de la proteína. Se confirmó que la región del dominio FH 1 rica en prolinas es una región intrínsecamente desordenada sin una estructura tridimensional definida. Este tipo de organización sugiere que la función biológica de este dominio o región “desordenado” es dinámica y que está siendo moldeada u ordenada gracias a las interacciones moleculares con otras proteínas o receptores y, por tanto, por eso no posee una estructura en 3D definida. Adicionalmente, a

partir de los análisis *in silico* realizados se sugiere que el dominio FH2 podría ser estable y posiblemente funcional. Finalmente, el obtener el modelado de la única Formina C completa de *Trypanosoma cruzi* CL Brener, así como la predicción del sitio de unión del inhibidor SMIFH2 en esta proteína abre el campo de investigación para ensayos experimentales e *in silico* sobre la interacción las Forminas no sólo con sus inhibidores sino con otras proteínas presentes en el citoesqueleto del parásito como una forma de explicar su papel biológico así como el funcionamiento y regulación del citoesqueleto, e incluso evaluar si estas proteínas podrían ser empleadas como blancos terapéuticos para el desarrollo de tratamientos para las enfermedades causadas por este microorganismo.

Aunque la región hemidimérica FH2 posee la misma cantidad de residuos de aminoácidos, es relevante analizar más adelante si existe alguna interferencia estérica que impida la unión del sub-monómero con las subunidades terminales del filamento de actina.

8. PERSPECTIVAS

A corto plazo:

- Expresar la Formina C como proteína de fusión con EGFP en los diferentes estadios del parásito para determinar su localización subcelular
- Obtener mutantes en las que se eliminen diferentes regiones y buscar las regiones que determinan su localización subcelular.

A largo plazo:

- Determinar el papel funcional de la Formina C en *T. cruzi* CL Brener, mediante la obtención de parásitos *knockout* por recombinación homóloga o por CRISPR/Cas9.
- Caracterizar *in silico* las Forminas A y B de *T. cruzi*.

9. REFERENCIAS

- Alas-Guardado, S. de J., Rojo A., & Merino, G. (2010). La paradoja de Levinthal: cuando una contradicción se vuelve lógica. *Educación Química*, 22(1), 51-54. [http://dx.doi.org/10.1016/S0187-893X\(18\)30114-9](http://dx.doi.org/10.1016/S0187-893X(18)30114-9)
- Abbass, J., & Nebel, J-C. (2020) Enhanced Rosetta-based Protein Structure Prediction for Non-Beta Sheet Dominated Targets. *IEEE*. <https://10.1109/MECBME47393.2020.9265175>
- Adzhubei, A. A., Sternberg, M. J., & Makarov, A. A. (2013). Polyproline-II helix in proteins: structure and function. *Journal of Molecular Biology*, 425(12), 2100–2132. <https://doi.org/10.1016/j.jmb.2013.03.018>
- Alves, C. R., Albuquerque-Cunha, J. M., Mello, C. B., Garcia, E. S., Nogueira, N. F., Bourguignon, S. C., de Souza, W., Azambuja, P., & Gonzalez, M. S. (2007). *Trypanosoma cruzi*: attachment to perimicrovillar membrane glycoproteins of *Rhodnius prolixus*. *Experimental Parasitology*, 116(1), 44–52. <https://doi.org/10.1016/j.exppara.2006.11.012>
- Amberg, D. C., Basart, E., & Botstein, D. (1995). Defining protein interactions with yeast actin in vivo. *Nature Structural Biology*, 2(1), 28–35. <https://doi.org/10.1038/nsb0195-28>
- Ampe, C., & Van Troys, M. (2017). Mammalian Actins: Isoform-Specific Functions and Diseases. *Handbook of experimental pharmacology*, 235, 1–37. https://doi.org/10.1007/164_2016_43
- André I. (2018). Modeling the Structure of Helical Assemblies with Experimental Constraints in Rosetta. *Methods In Molecular Biology (Clifton, N.J.)*, 1764, 475–489. https://doi.org/10.1007/978-1-4939-7759-8_30
- Andrianantoandro, E., & Pollard, T. D. (2006). Mechanism of actin filament turnover by severing and nucleation at different concentrations of ADF/cofilin. *Molecular Cell*, 24(1), 13–23. <https://doi.org/10.1016/j.molcel.2006.08.006>

Anfinsen C. B. (1973). Principles that govern the folding of protein chains. *Science* (New York, N.Y.), 181(4096), 223–230. <https://doi.org/10.1126/science.181.4096.223>

Aydin, F., Courtemanche, N., Pollard, T. D., & Voth, G. A. (2018). Gating mechanisms during actin filament elongation by formins. *eLife*, 7, e37342. <https://doi.org/10.7554/eLife.37342>

Balchin, D., Hayer-Hartl, M., & Hartl, F. U. (2020). Recent advances in understanding catalysis of protein folding by molecular chaperones. *FEBS letters*, 594(17), 2770–2781. <https://doi.org/10.1002/1873-3468.13844>

Bernstein, B. W., & Bamburg, J. R. (2010). ADF/cofilin: a functional node in cell biology. *Trends in Cell Biology*, 20(4), 187–195. <https://doi.org/10.1016/j.tcb.2010.01.001>

Bertoni, M., Kiefer, F., Biasini, M., Bordoli, L., & Schwede, T. (2017). Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Scientific Reports*, 7(1), 10480. <https://doi.org/10.1038/s41598-017-09654-8>

Bienert, S., Waterhouse, A., de Beer, T. A., Tauriello, G., Studer, G., Bordoli, L., & Schwede, T. (2017). The SWISS-MODEL Repository-new features and functionality. *Nucleic Acids Research*, 45(D1), D313–D319. <https://doi.org/10.1093/nar/gkw1132>

Bitencourt-Ferreira, G., & de Azevedo, W. F., Jr (2019). Homology Modeling of Protein Targets with MODELLER. *Methods in Molecular Biology* (Clifton, N.J.), 2053, 231–249. https://doi.org/10.1007/978-1-4939-9752-7_15

Bittencourt-Cunha, P. R., Silva-Cardoso, L., Oliveira, G. A., Silva, J. R., Silveira, A. B., Kluck, G. E., Souza-Lima, M., Gondim, K. C., Dansa-Petretsky, M., Silva, C. P., Masuda, H., Silva Neto, M. A., & Atella, G. C. (2013). Perimicrovillar membrane assembly: the fate of phospholipids synthesised by the midgut of *Rhodnius prolixus*. *Memorias do Instituto Oswaldo Cruz*, 108(4), 494–500. <https://doi.org/10.1590/S0074-0276108042013016>

Block, J., Breitsprecher, D., Kühn, S., Winterhoff, M., Kage, F., Geffers, R., Duwe, P., Rohn, J. L., Baum, B., Brakebusch, C., Geyer, M., Stradal, T. E., Faix, J., & Rottner, K. (2012). FMNL2 drives actin-based protrusion and migration downstream of Cdc42. *Current biology: CB*, 22(11), 1005–1012. <https://doi.org/10.1016/j.cub.2012.03.064>

Bogitsh BJ, Ribeiro-Rodrigues R, Carter CE. In vitro effects of mannan and cytochalasin B on the uptake of horseradish peroxidase and [¹⁴C] sucrose by *Trypanosoma cruzi* epimastigotes. (1995) *J Parasitol*, 81(2), 144-8.

Botero, D., & Restrepo, M. (2012). Parasitosis Humanas. En Botero, D., & Restrepo, M (Eds), *Parasitosis Humanas* (5th ed). Medellín, Colombia: Corporación para investigaciones biológicas.

Bradley, P., Malmström, L., Qian, B., Schonbrun, J., Chivian, D., Kim, D. E., Meiler, J., Misura, K. M., & Baker, D. (2005). Free modeling with Rosetta in CASP6. *Proteins*, 61 Suppl 7, 128–134. <https://doi.org/10.1002/prot.20729>

Breitsprecher, D., & Goode, B. L. (2013). Formins at a glance. *Journal of cell science*, 126(Pt 1), 1–7. <https://doi.org/10.1242/jcs.107250>

Campellone, K. G., & Welch, M. D. (2010). A nucleator arms race: cellular control of actin assembly. *Nature reviews. Molecular cell biology*, 11(4), 237–251. <https://doi.org/10.1038/nrm2867>

Carlier, M. F., & Pantaloni, D. (1997). Control of actin dynamics in cell motility. *Journal of molecular biology*, 269(4), 459–467. <https://doi.org/10.1006/jmbi.1997.1062>

Castro, D. P., Seabra, S. H., Garcia, E. S., de Souza, W., & Azambuja, P. (2007). *Trypanosoma cruzi*: ultrastructural studies of adhesion, lysis and biofilm formation by *Serratia marcescens*. *Experimental Parasitology*, 117(2), 201–207. <https://doi.org/10.1016/j.exppara.2007.04.014>

Chalkia, D., Nikolaidis, N., Makalowski, W., Klein, J., & Nei, M. (2008). Origins and evolution of the formin multigene family that is involved in the formation of actin

filaments. *Molecular Biology and Evolution*, 25(12), 2717–2733. <https://doi.org/10.1093/molbev/msn215>

Chasen NM, Coppens I, Etheridge RD. Identification and Localization of the First Known Proteins of the *Trypanosoma cruzi* Cytostome Cytopharynx Endocytic Complex. (2020) *Front Cell Infect Microbiol*, 17, 9:445. doi: 10.3389/fcimb.2019.00445

Cheng, L., Zhang, J., Ahmad, S., Rozier, L., Yu, H., Deng, H., & Mao, Y. (2011). Aurora B regulates formin mDia3 in achieving metaphase chromosome alignment. *Developmental Cell*, 20(3), 342–352. <https://doi.org/10.1016/j.devcel.2011.01.008>

Chesarone MA, DuPage AG, Goode BL. Unleashing formins to remodel the actin and microtubule cytoskeletons. *Nat Rev Mol Cell Biol*. 2010 Jan;11(1):62-74. doi: 10.1038/nrm2816. Epub 2009 Dec 9. PMID: 19997130

Chhabra, E. S., Ramabhadran, V., Gerber, S. A., & Higgs, H. N. (2009). INF2 is an endoplasmic reticulum-associated formin protein. *Journal of cell science*, 122(Pt 9), 1430–1440. <https://doi.org/10.1242/jcs.040691>

Choi, H. J., & Weis, W. I. (2005). Structure of the armadillo repeat domain of plakophilin 1. *Journal of molecular biology*, 346(1), 367–376. <https://doi.org/10.1016/j.jmb.2004.11.048>

Cooper, G.M. (2000). Protein Folding and Processing. In *The Cell: A Molecular Approach* (2 nd edition). Sunderland (MA): Sinauer Associates

Corrêa, J. R., Atella, G. C., Batista, M. M., Soares, M. J. (2008). Transferrin uptake in *Trypanosoma cruzi* is impaired by interference on cytosome-associated cytoskeleton elements and stability of membrane cholesterol, but not by obstruction of clathrin-dependent endocytosis. *Exp Parasitol*, 119, 58–66. <https://doi.org/10.1016/j.exppara.2007.12.010>

Courtemanche, N. (2018). Mechanisms of formin-mediated actin assembly and dynamics. *Biophysical Reviews*, 10(6), 1553–1569. <https://doi.org/10.1007/s12551-018-0468-6>

Courtemanche, N., & Pollard, T. D. (2012). Determinants of Formin Homology 1 (FH1) domain function in actin filament elongation by formins. *The Journal of biological chemistry*, 287(10), 7812–7820. <https://doi.org/10.1074/jbc.M111.322958>

Dames, S. A., Junemann, A., Sass, H. J., Schönichen, A., Stopschinski, B. E., Grzesiek, S., Faix, J., & Geyer, M. (2011). Structure, dynamics, lipid binding, and physiological relevance of the putative GTPase-binding domain of Dictyostelium formin C. *The Journal of biological chemistry*, 286(42), 36907–36920. <https://doi.org/10.1074/jbc.M111.225052>

de Hostos, E. L., Bradtke, B., Lottspeich, F., Guggenheim, R., & Gerisch, G. (1991). Coronin, an actin binding protein of *Dictyostelium discoideum* localized to cell surface projections, has sequence similarities to G protein beta subunits. *The EMBO Journal*, 10(13), 4097–4104

de Souza, D., Pavoni, D.P., Krieger, M.A., & Ludwig, A. (2018). Evolutionary analyses of myosin genes in trypanosomatids show a history of expansion, secondary losses and neofunctionalization. *Scientific reports*, 8 (1), 1376. <https://doi.org/10.1038/s41598-017-18865-y>

Deshpande, N., Address, K. J., Bluhm, W. F., Merino-Ott, J. C., Townsend-Merino, W., Zhang, Q., Knezevich, C., Xie, L., Chen, L., Feng, Z., Green, R. K., Flippen-Anderson, J. L., Westbrook, J., Berman, H. M., & Bourne, P. E. (2005). The RCSB Protein Data Bank: a redesigned query system and relational database based on the mmCIF schema. *Nucleic Acids Research*, 33(Database issue), D233–D237. <https://doi.org/10.1093/nar/gki057>

Dill, K. A., Ozkan, S. B., Shell, M. S., & Weikl, T. R. (2008). The protein folding problem. *Annual Review of Biophysics*, 37, 289–316. <https://doi.org/10.1146/annurev.biophys.37.092707.153558>

Dinner, A. R., Sali, A., Smith, L. J., Dobson, C. M., & Karplus, M. (2000). Understanding protein folding via free-energy surfaces from theory and experiment. *Trends in Biochemical Sciences*, 25(7), 331–339. [https://doi.org/10.1016/s0968-0004\(00\)01610-8](https://doi.org/10.1016/s0968-0004(00)01610-8)

Dominguez R. (2004). Actin-binding proteins--a unifying hypothesis. *Trends in biochemical sciences*, 29(11), 572–578. <https://doi.org/10.1016/j.tibs.2004.09.004>

Dominguez, R., & Holmes, K. C. (2011). Actin structure and function. *Annual review of biophysics*, 40, 169–186. <https://doi.org/10.1146/annurev-biophys-042910-155359>

dos Remedios, C. G., Chhabra, D., Kekic, M., Dedova, I. V., Tsubakihara, M., Berry, D. A., & Nosworthy, N. J. (2003). Actin binding proteins: regulation of cytoskeletal microfilaments. *Physiological reviews*, 83(2), 433–473. <https://doi.org/10.1152/physrev.00026.2002>

Dujardin, J. P., Schofield, C.J., & Panzera, F. (2002). Los vectores de la Enfermedad de Chagas. Vol. 25. Bruselas: Académie Royale des Sciences d'Outre-Mer; 189 p

Durrant, J. D., & McCammon, J. A. (2011). Molecular dynamics simulations and drug discovery. *BMC biology*, 9, 71. <https://doi.org/10.1186/1741-7007-9-71>

Edgar R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5), 1792–1797. <https://doi.org/10.1093/nar/gkh340>

Eisinger, J. (1957). Adsorption of CO on Tungsten and Its Effect on the Work Function. *Journal of Chemical Physics*, 27, 1206-1208. <https://doi.org/10.1063/1.1743955>

Eng, R. C., Halat, L. S., Livingston, S. J., Sakai, T., Motose, H., & Wasteney, G. O. (2017). The ARM Domain of ARMADILLO-REPEAT KINESIN 1 is Not Required for Microtubule Catastrophe But Can Negatively Regulate NIMA-RELATED KINASE 6 in *Arabidopsis thaliana*. *Plant & cell physiology*, 58(8), 1350–1363. <https://doi.org/10.1093/pcp/pcx070>

Englander, S. W., & Mayne, L. (2017). The case for defined protein folding pathways. *Proceedings of the National Academy of Sciences of the United States of America*, 114(31), 8253–8258. <https://doi.org/10.1073/pnas.1706196114>

Englander, S. W., & Mayne, L. (2014). The nature of protein folding pathways. *Proceedings of the National Academy of Sciences of the United States of America*, 111(45), 15873–15880. <https://doi.org/10.1073/pnas.1411798111>

Ergin V, Zheng S. Putative Coiled-Coil Domain-Dependent Autoinhibition and Alternative Splicing Determine SHTN1's Actin-Binding Activity. *J Mol Biol*. 2020 Jun 26;432(14):4154-4166. doi: 10.1016/j.jmb.2020.04.025. Epub 2020 May 1. PMID: 32371045; PMCID: PMC7418779

Eswar, N., Eramian, D., Webb, B., Shen MY., & Sali A. (2008). Protein Structure Modeling with MODELLER. In: Kobe, B., Guss, M., & Huber T. (Eds.), *Structural Proteomics. Methods in Molecular Biology™* (pp. 145-159). Humana Press. https://doi.org/10.1007/978-1-60327-058-8_8

Faix, J., & Rottner, K. (2006). The making of filopodia. *Current opinion in cell biology*, 18(1), 18–25. <https://doi.org/10.1016/j.ceb.2005.11.002>

Ferron F, Rebowski G, Lee SH, Dominguez R. Structural basis for the recruitment of profilin-actin complexes during filament elongation by Ena/VASP. (2007) *EMBO J*, 26(21):4597-606. doi: 10.1038/sj.emboj.7601874

Feynman, R. P., & Vernon, F. L. (1963). The theory of a general quantum system interacting with a linear dissipative system. *Annals of Physics*, 24, 118-173.

[https://doi.org/10.1016/0003-4916\(63\)90068-x](https://doi.org/10.1016/0003-4916(63)90068-x)Fletcher, D. A., & Mullins, R. D. (2010). Cell mechanics and the cytoskeleton. *Nature*, 463(7280), 485–492. <https://doi.org/10.1038/nature08908>

Foldi, I., Szikora, S., & Mihály, J. (2017). Formin' bridges between microtubules and actin filaments in axonal growth cones. *Neural Regeneration Research*, 12(12), 1971–1973. <https://doi.org/10.4103/1673-5374.221148>

Fuhrmann, J., Rurainski, A., Lenhof, H. P., & Neumann, D. (2010). A new Lamarckian genetic algorithm for flexible ligand-receptor docking. *Journal of Computational Chemistry*, 31(9), 1911–1918. <https://doi.org/10.1002/jcc.21478>

Fujii, T., Iwane, A. H., Yanagida, T., & Namba, K. (2010). Direct visualization of secondary structures of F-actin by electron cryomicroscopy. *Nature*, 467(7316), 724–728. <https://doi.org/10.1038/nature09372>

Gandhi, M., & Goode, B. L. (2008). Coronin: the double-edged sword of actin dynamics. *Sub-cellular Biochemistry*, 48, 72–87. https://doi.org/10.1007/978-0-387-09595-0_7

Gómez-Gil, E., Martín-García, R., Vicente-Soler, J., Franco, A., Vázquez-Marín, B., Prieto-Ruiz, F., Soto, T., Pérez, P., Madrid, M., & Cansado, J. (2020). Stress-activated MAPK signaling controls fission yeast actomyosin ring integrity by modulating formin For3 levels. *eLife*, 9, e57951. <https://doi.org/10.7554/eLife.57951>

Gorelik R, Yang C, Kameswaran V, Dominguez R, Svitkina T. Mechanisms of plasma membrane targeting of formin mDia2 through its amino terminal domains. *Mol Biol Cell*. 2011 Jan 15;22(2):189-201. doi: 10.1091/mbc.E10-03-0256. Epub 2010 Nov 30. PMID: 21119010; PMCID: PMC3020915.

Guhl, F., & Lazdins-Helds, J. (2007). Reporte sobre la enfermedad de Chagas. Reporte del grupo de trabajo científico sobre la enfermedad de Chagas. http://apps.who.int/iris/bitstream/handle/10665/69724/TDR_SWG_09_spa.pdf?sequence=1

Gul, I. S., Hulpiau, P., Saeys, Y., & van Roy, F. (2017). Metazoan evolution of the armadillo repeat superfamily. *Cellular and molecular life sciences: CMLS*, 74(3), 525–541. <https://doi.org/10.1007/s00018-016-2319-6>

Guex, N., Peitsch, M. C., & Schwede, T. (2009). Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis*, 30 Suppl 1, S162–S173. <https://doi.org/10.1002/elps.200900140>

Hakala, M., Kalimeri, M., Enkavi, G., Vattulainen, I., & Lappalainen, P. (2018). Molecular mechanism for inhibition of twinfilin by phosphoinositides. *The Journal of Biological Chemistry*, 293(13), 4818–4829. <https://doi.org/10.1074/jbc.RA117.000484>

Hamilton, P. B., & Stevens, J. R. (2010). Classification and Phylogeny of *Trypanosoma cruzi*. In Telleria, J. & Tibayrenc, M (Eds.), *American Trypanosomiasis Chagas Disease One Hundred Years of Research* (pp. 321-338). Elsevier. <https://doi.org/10.1016/B978-0-12-384876-5.00013-7>

Higgs H. N. (2005). Formin proteins: a domain-based approach. *Trends in Biochemical Sciences*, 30(6), 342–353. <https://doi.org/10.1016/j.tibs.2005.04.014>

Higgs, H. N., & Peterson, K. J. (2005). Phylogenetic analysis of the formin homology 2 domain. *Molecular Biology of the Cell*, 16(1), 1–13. <https://doi.org/10.1091/mbc.e04-07-0565>

Hilton, D. M., Aguilar, R. M., Johnston, A. B., & Goode, B. L. (2018). Species-Specific Functions of Twinfilin in Actin Filament Depolymerization. *Journal of Molecular Biology*, 430(18 Pt B), 3323–3336. <https://doi.org/10.1016/j.jmb.2018.06.025>

Hollingsworth, S. A., & Dror, R. O. (2018). Molecular Dynamics Simulation for All. *Neuron*, 99(6), 1129–1143. <https://doi.org/10.1016/j.neuron.2018.08.011>

Jin, Z. L., Jo, Y. J., Namgoong, S., & Kim, N. H. (2018). CAP1-mediated actin cycling via ADF/cofilin proteins is essential for asymmetric division in mouse oocytes. *Journal of cell science*, 131(23), jcs222356. <https://doi.org/10.1242/jcs.222356>

Jost Lopez, A., Quoika, P. K., Linke, M., Hummer, G., & Köfinger, J. (2020). Quantifying Protein-Protein Interactions in Molecular Simulations. *The journal of Physical Chemistry. B*, 124(23), 4673–4685. <https://doi.org/10.1021/acs.jpcc.9b11802>

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, 10(6), 845–858. <https://doi.org/10.1038/nprot.2015.053>

Kennedy M. B. (1995). Origin of PDZ (DHR, GLGF) domains. *Trends in Biochemical Sciences*, 20(9), 350. [https://doi.org/10.1016/s0968-0004\(00\)89074-x](https://doi.org/10.1016/s0968-0004(00)89074-x)

Kirkbride, K. C., Sung, B. H., Sinha, S., & Weaver, A. M. (2011). Cortactin: a multifunctional regulator of cellular invasiveness. *Cell Adhesion & Migration*, 5(2), 187–198. <https://doi.org/10.4161/cam.5.2.14773>

Kitayama, C., & Uyeda, T. Q. (2003). ForC, a novel type of formin family protein lacking an FH1 domain, is involved in multicellular development in *Dictyostelium discoideum*. *Journal of Cell Science*, 116(Pt 4), 711–723. <https://doi.org/10.1242/jcs.00265>

Knodler LA, Ibarra JA, Pérez-Rueda E, Yip CK, Steele-Mortimer O. Coiled-coil domains enhance the membrane association of Salmonella type III effectors. *Cell Microbiol.* 2011 Oct;13(10):1497-517. doi: 10.1111/j.1462-5822.2011.01635.x. Epub 2011 Jul 11. PMID: 21679290; PMCID: PMC3418822.

Krishnan, K., & Moens, P. (2009). Structure and functions of profilins. *Biophysical Reviews*, 1(2), 71–81. <https://doi.org/10.1007/s12551-009-0010-y>

Kühn, S., Erdmann, C., Kage, F., Block, J., Schwenkmezger, L., Steffen, A., Rottner, K., & Geyer, M. (2015). The structure of FMNL2-Cdc42 yields insights into the mechanism of lamellipodia and filopodia formation. *Nature communications*, 6, 7088. <https://doi.org/10.1038/ncomms8088>

Lee, S. H., & Dominguez, R. (2010). Regulation of actin cytoskeleton dynamics in cells. *Molecules and cells*, 29(4), 311–325. <https://doi.org/10.1007/s10059-010-0053-8>

Lodish, H., Berk, A., Zipursky, S.L., *et al.* (2000). Section 18.2, The Dynamics of Actin Assembly. In Freeman, W.H. (Ed.), *Molecular Cell Biology* (4th ed). New York.

Lodish, H., Berk, A. and Zipursky, S.L. (2000) Section 24.2 Proto-Oncogenes and Tumor-Suppressor Genes. *Molecular Cell Biology*. 4th Edition, W. H. Freeman, New York.

Lu, J., Meng, W., Poy, F., Maiti, S., Goode, B. L., & Eck, M. J. (2007). Structure of the FH2 domain of Daam1: implications for formin regulation of actin assembly.

Journal of Molecular Biology, 369(5), 1258–1269.
<https://doi.org/10.1016/j.jmb.2007.04.002>

Lu, J., & Pollard, T. D. (2001). Profilin binding to poly-L-proline and actin monomers along with ability to catalyze actin nucleotide exchange is required for viability of fission yeast. *Molecular biology of the cell*, 12(4), 1161–1175.
<https://doi.org/10.1091/mbc.12.4.1161>

Luxenburg, C., & Geiger, B. (2017). Multiscale View of Cytoskeletal Mechano-regulation of Cell and Tissue Polarity. *Handbook of experimental pharmacology*, 235, 263–284. https://doi.org/10.1007/164_2016_34

Maciver, S. K., & Hussey, P. J. (2002). The ADF/cofilin family: actin-remodeling proteins. *Genome Biology*, 3(5), reviews3007. <https://doi.org/10.1186/gb-2002-3-5-reviews3007>

Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A., Potter, S. C., Finn, R. D., & Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research*, 47(W1), W636–W641. <https://doi.org/10.1093/nar/gkz268>

McCammon, J. A., Gelin, B. R., & Karplus, M. (1977). Dynamics of folded proteins. *Nature*, 267(5612), 585–590. <https://doi.org/10.1038/267585a0>

Meena, S. K., & Sulpizi, M. (2013). Understanding the microscopic origin of gold nanoparticle anisotropic growth from molecular dynamics simulations. *Langmuir: the ACS Journal of Surfaces and Colloids*, 29(48), 14954–14961. <https://doi.org/10.1021/la403843n>

Melo, L., Sant'anna, C., Reis, S., Lourenço, D., Souza, W., Lopes, U., & Cunha-E-Silva, N. (2008). Evolutionary conservation of actin-binding proteins in *Trypanosoma cruzi* and unusual subcellular localization of the actin homologue. *Parasitology*, 135(8), 955–965. doi:10.1017/S0031182008004496

Mortara RA. Studies on trypanosomatid actin. I. Immunochemical and biochemical identification. (1989) J Protozool, 36(1):8-13. doi: 10.1111/j.1550-7408.1989.tb02666.x.

Mullins, R. D., & Hansen, S.D. (2013). In vitro studies of the dynamics of actin filaments and networks. Current Opinion in Cell Biology, 25 (1), 6–13. <https://doi.org/10.1016/j.ceb.2012.11.007>

Muñoz-Lasso, D. C., Romá-Mateo, C., Pallardó, F. V., & Gonzalez-Cabo, P. (2020). Much More Than a Scaffold: Cytoskeletal Proteins in Neurological Disorders. Cells, 9(2), 358. <https://doi.org/10.3390/cells9020358>

Mwangangi, D. M., Manser, E., & Robinson, R. C. (2021). The structure of the actin filament uncapping complex mediated by twinfilin. Science advances, 7(5), eabd5271. <https://doi.org/10.1126/sciadv.abd5271>

Ono S. (2013). The role of cyclase-associated protein in regulating actin filament dynamics - more than a monomer-sequestration factor. Journal of Cell Science, 126(Pt 15), 3249–3258. <https://doi.org/10.1242/jcs.128231>

Ono S. (2018). Functions of actin-interacting protein 1 (AIP1)/WD repeat protein 1 (WDR1) in actin filament dynamics and cytoskeletal regulation. Biochemical and Biophysical Research Communications, 506(2), 315–322. <https://doi.org/10.1016/j.bbrc.2017.10.096>

Osorio-Méndez, J. F., Vizcaíno-Castillo, A., Manning-Cela, R., Hernández, R., & Cevallos, A. M. (2016). Expression of profilin in *Trypanosoma cruzi* and identification of some of its ligands. Biochemical and Biophysical Research Communications, 480(4), 709–714. <https://doi.org/10.1016/j.bbrc.2016.10.044>

Osorio Méndez, J. F. (2012). Caracterización de profilina en *trypanosoma cruzi*. Tesis para optar por el grado de Doctorado. Universidad Nacional Autónoma de México

Ovchinnikov, S., Park, H., Kim, D. E., DiMaio, F., & Baker, D. (2018). Protein structure prediction using Rosetta in CASP12. *Proteins*, 86 Suppl 1(Suppl 1), 113–121. <https://doi.org/10.1002/prot.25390>

Pacella, M. S., Koo, d., Thottungal, R. A., & Gray, J. J. (2013). Using the RosettaSurface algorithm to predict protein structure at mineral surfaces. *Methods in Enzymology*, 532, 343–366. <https://doi.org/10.1016/B978-0-12-416617-2.00016-3>

Parsons, M., Worthey, E. A., Ward, P. N., & Mottram, J. C. (2005). Comparative analysis of the kinomes of three pathogenic trypanosomatids: *Leishmania major*, *Trypanosoma brucei* and *Trypanosoma cruzi*. *BMC genomics*, 6, 127. <https://doi.org/10.1186/1471-2164-6-127>

Paul, A. S., & Pollard, T. D. (2008). The role of the FH1 domain and profilin in formin-mediated actin-filament elongation and nucleation. *Current biology: CB*, 18(1), 9–19. <https://doi.org/10.1016/j.cub.2007.11.062>

Pearson W. R. (2013). An introduction to sequence similarity ("homology") searching. *Current protocols in bioinformatics*, Chapter 3, Unit3.1. <https://doi.org/10.1002/0471250953.bi0301s42>

Pegoraro, A. F., Janmey, P., & Weitz, D. A. (2017). Mechanical Properties of the Cytoskeleton and Cells. *Cold Spring Harbor Perspectives in Biology*, 9(11), a022038. <https://doi.org/10.1101/cshperspect.a022038>

Pernier, J., Shekhar, S., Jegou, A., Guichard, B., & Carlier, M. F. (2016). Profilin Interaction with Actin Filament Barbed End Controls Dynamic Instability, Capping, Branching, and Motility. *Developmental Cell*, 36(2), 201–214. <https://doi.org/10.1016/j.devcel.2015.12.024>

Pierce, L. C., Salomon-Ferrer, R., Augusto F de Oliveira, C., McCammon, J. A., & Walker, R. C. (2012). Routine Access to Millisecond Time Scale Events with Accelerated Molecular Dynamics. *Journal of Chemical Theory and Computation*, 8(9), 2997–3002. <https://doi.org/10.1021/ct300284c>

Pollard, T. D., Blanchoin, L., & Mullins, R. D. (2000). Molecular mechanisms controlling actin filament dynamics in nonmuscle cells. *Annual Review of Biophysics and Biomolecular Structure*, 29, 545–576. <https://doi.org/10.1146/annurev.biophys.29.1.545>

Pollard T. D. (2016). Actin and Actin-Binding Proteins. *Cold Spring Harbor Perspectives in Biology*, 8(8), a018226. <https://doi.org/10.1101/cshperspect.a018226>

Pollard T. D. (2017). What We Know and Do Not Know About Actin. *Handbook of Experimental Pharmacology*, 235, 331–347. https://doi.org/10.1007/164_2016_44

Priya, P., Kesheri, M., Sinha, R. P., & Kanchan, S. (2016). Molecular Dynamics Simulations for Biological Systems. In Karâa, W & Dey, N (Eds.), *Biomedical Image Analysis and Mining Techniques for Improved Health Outcomes* (pp. 286-313). IGI Global. <http://doi:10.4018/978-1-4666-8811-7.ch014>

Prostak, S. M., Robinson, K. A., Titus, M. A., & Fritz-Laylin, L. K. (2021). The actin networks of chytrid fungi reveal evolutionary loss of cytoskeletal complexity in the fungal kingdom. *Current biology: CB*, 31(6), 1192–1205.e6. <https://doi.org/10.1016/j.cub.2021.01.001>

Rivero, F., Muramoto, T., Meyer, A. K., Urushihara, H., Uyeda, T. Q., & Kitayama, C. (2005). A comparative sequence analysis reveals a common GBD/FH3-FH1-FH2-DAD architecture in formins from *Dictyostelium*, fungi and metazoa. *BMC genomics*, 6, 28. <https://doi.org/10.1186/1471-2164-6-28>

Rizvi, S. A., Neidt, E. M., Cui, J., Feiger, Z., Skau, C. T., Gardel, M. L., Kozmin, S. A., & Kovar, D. R. (2009). Identification and characterization of a small molecule inhibitor of formin-mediated actin assembly. *Chemistry & biology*, 16(11), 1158–1168. <https://doi.org/10.1016/j.chembiol.2009.10.006>

Rose R, Weyand M, Lammers M, et al. Structural and mechanistic insights into the interaction between Rho and mammalian Dia. (2005). *Nature*, 435(7041):513-518. DOI: 10.1038/nature03604

Roterman-Konieczna, I. (2012). A short description of other selected ab initio methods for protein structure prediction. In Roterman-Konieczna, I. (Ed.), *Protein Folding in Silico, Protein Folding Versus Protein Structure Prediction*, 165-189. Woodhead Publishing. <https://doi.org/10.1533/9781908818256.165>

Russell, A. J., & Torda, A. E. (2002). Protein sequence threading: Averaging over structures. *Proteins*, 47(4), 496–505. <https://doi.org/10.1002/prot.10088>

Rybakin, V., & Clemen, C. S. (2005). Coronin proteins as multifunctional regulators of the cytoskeleton and membrane trafficking. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, 27(6), 625–632. <https://doi.org/10.1002/bies.20235>

Sahasrabuddhe, A. A., Nayak, R. C., & Gupta, C. M. (2009). Ancient Leishmania coronin (CRN12) is involved in microtubule remodeling during cytokinesis. *Journal of cell science*, 122(Pt 10), 1691–1699. <https://doi.org/10.1242/jcs.044651>

Salazar-Schettino, P. M., Bucio-Torres, M. I., Cabrera-Bravo, M., Alba-Alvarado, M. C. de, Castillo-Saldaña, D. R., Zenteno-Galindo, E. A., Rojo-Medina, J., Fernández-Santos, N. A., & Perera-Salazar, M. G. (2016). Enfermedad de Chagas en México. *Revista de la Facultad de Medicina (México)*, 59(3), 6-16. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0026-17422016000300006&lng=es&tlng=es

Sali, A., & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology*, 234(3), 779–815. <https://doi.org/10.1006/jmbi.1993.1626>

Schnoor, M., Stradal, T. E., & Rottner, K. (2018). Cortactin: Cell Functions of A Multifaceted Actin-Binding Protein. *Trends in Cell Biology*, 28(2), 79–98. <https://doi.org/10.1016/j.tcb.2017.10.009>

Schönichen, A., & Geyer, M. (2010). Fifteen formins for an actin filament: a molecular view on the regulation of human formins. *Biochimica et Biophysica Acta*, 1803(2), 152–163. <https://doi.org/10.1016/j.bbamcr.2010.01.014>

Schulte, A., Stolp, B., Schönichen, A., Pylypenko, O., Rak, A., Fackler, O. T., & Geyer, M. (2008). The human formin FHOD1 contains a bipartite structure of FH3 and GTPase-binding domains required for activation. *Structure (London, England: 1993)*, 16(9), 1313–1323. <https://doi.org/10.1016/j.str.2008.06.008>

Sebti S. M. (2005). Protein farnesylation: implications for normal physiology, malignant transformation, and cancer therapy. *Cancer Cell*, 7(4), 297–300. <https://doi.org/10.1016/j.ccr.2005.04.005>

Senn, H. M., & Thiel, W. (2009). QM/MM methods for biomolecular systems. *Angewandte Chemie (International ed. in English)*, 48(7), 1198–1229. <https://doi.org/10.1002/anie.200802019>

Shimada, A., Nyitrai, M., Vetter, I. R., Köhlmann, D., Bugyi, B., Narumiya, S., Geeves, M. A., & Wittinghofer, A. (2004). The core FH2 domain of diaphanous-related formins is an elongated actin binding protein that inhibits polymerization. *Molecular cell*, 13(4), 511–522. [https://doi.org/10.1016/s1097-2765\(04\)00059-0](https://doi.org/10.1016/s1097-2765(04)00059-0)

Sigrist, C. J., de Castro, E., Cerutti, L., Cuče, B. A., Hulo, N., Bridge, A., Bougueleret, L., & Xenarios, I. (2013). New and continuing developments at PROSITE. *Nucleic Acids Research*, 41(Database issue), D344–D347. <https://doi.org/10.1093/nar/gks1067>

Simons, K. T., Bonneau, R., Ruczinski, I., & Baker, D. (1999). Ab initio protein structure prediction of CASP III targets using ROSETTA. *Proteins, Suppl 3*, 171–176. [https://doi.org/10.1002/\(sici\)1097-0134\(1999\)37:3+<171::aid-prot21>3.3.co;2-q](https://doi.org/10.1002/(sici)1097-0134(1999)37:3+<171::aid-prot21>3.3.co;2-q)

Sit, S. T., & Manser, E. (2011). Rho GTPases and their role in organizing the actin cytoskeleton. *Journal of cell science*, 124(Pt 5), 679–683. <https://doi.org/10.1242/jcs.064964>

Smythe, E., & Ayscough, K. R. (2003). The Ark1/Prk1 family of protein kinases. Regulators of endocytosis and the actin skeleton. *EMBO reports*, 4(3), 246–251. <https://doi.org/10.1038/sj.embor.embor776>

Souza W. d. (2009). Structural organization of *Trypanosoma cruzi*. Memorias do Instituto Oswaldo Cruz, 104 Suppl 1, 89–100. <https://doi.org/10.1590/s0074-02762009000900014>

Souza LC, Pinho RE, Lima CV, Fragoso SP, Soares MJ. Actin expression in trypanosomatids (Euglenozoa: Kinetoplastea). (2013). Mem Inst Oswaldo Cruz, 108(5):631-6. doi: 10.1590/0074-0276108052013015.

Stuart, K., Brun, R., Croft, S., Fairlamb, A., Gürtler, R. E., McKerrow, J., Reed, S., & Tarleton, R. (2008). Kinetoplastids: related protozoan pathogens, different diseases. The Journal of Clinical Investigation, 118(4), 1301–1310. <https://doi.org/10.1172/JCI33945>

Studer, G., Rempfer, C., Waterhouse, A. M., Gumienny, R., Haas, J., & Schwede, T. (2020). QMEANDisCo-distance constraints applied on model quality estimation. Bioinformatics (Oxford, England), 36(6), 1765–1771. <https://doi.org/10.1093/bioinformatics/btz828>

Sun, T., Li, S., & Ren, H. (2013). Profilin as a regulator of the membrane-actin cytoskeleton interface in plant cells. Frontiers in Plant Science, 4, 512. <https://doi.org/10.3389/fpls.2013.00512>

Szikora, S., Földi, I., Tóth, K., Migh, E., Vig, A., Bugyi, B., Maléth, J., Hegyi, P., Kaltenecker, P., Sanchez-Soriano, N., & Mihály, J. (2017). The formin DAAM is required for coordination of the actin and microtubule cytoskeleton in axonal growth cones. Journal of Cell Science, 130(15), 2506–2519. <https://doi.org/10.1242/jcs.203455>

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Molecular Biology and Evolution, 38(7), 3022–3027. <https://doi.org/10.1093/molbev/msab120>

Tanaka K. (2000). Formin family proteins in cytoskeletal control. Biochemical and biophysical research communications, 267(2), 479–481. <https://doi.org/10.1006/bbrc.1999.1707>

Torda A. E. (2004). Protein Sequence Optimization—Theory, Practice, and Fundamental Impossibility. *Soft Materials*, 2(1), 1-10. <https://10.1081/SMTS-120030139>

Truebestein L, Leonard TA. Coiled-coils: The long and short of it. *Bioessays*. 2016 Sep;38(9):903-16. doi: 10.1002/bies.201600062. Epub 2016 Aug 5. PMID: 27492088; PMCID: PMC5082667

Tsarouhas V, Liu D, Tsikala G, Fedoseienko A, Zinn K, Matsuda R, Billadeau DD, Samakovlis C. WASH phosphorylation balances endosomal versus cortical actin network integrities during epithelial morphogenesis. *Nat Commun*. 2019 May 16;10(1):2193. doi: 10.1038/s41467-019-10229-6. PMID: 31097705; PMCID: PMC6522504

Tyler, K. M., & Engman, D. M. (2001). The life cycle of *Trypanosoma cruzi* revisited. *International Journal for Parasitology*, 31(5-6), 472–481. [https://doi.org/10.1016/s0020-7519\(01\)00153-9](https://doi.org/10.1016/s0020-7519(01)00153-9)

Vadillo Mata, J. (2009). Clonación y caracterización de un gene atípico de actina de *Trypanosoma Cruzi*. Tesis para optar por el grado de Maestría. Universidad Nacional Autónoma de México.

Vivas, R. J., García, J. E., Guhl, F., Hernández, C., Velásquez, N., Ramírez, J. D., Carranza, J. C., & Vallejo, G. A. (2021). Systematic review on the biology, ecology, genetic diversity and parasite transmission potential of *Panstrongylus geniculatus* (Latreille 1811) in Latin America. *Memorias do Instituto Oswaldo Cruz*, 116, e200528. <https://doi.org/10.1590/0074-02760200528>

Vizcaíno-Castillo, A., Osorio-Méndez, J. F., Rubio-Ortiz, M., Manning-Cela, R. G., Hernández, R., & Cevallos, A. M. (2019). *Trypanosoma cruzi* actins: Expression analysis of actin 2. *Biochemical and Biophysical Research Communications*, 513(2), 347–353. <https://doi.org/10.1016/j.bbrc.2019.04.007>

Vizcaíno-Castillo, A., Osorio-Méndez, J. F., Ambrosio, J. R., Hernández, R., & Cevallos, A. M. (2020). The complexity and diversity of the actin cytoskeleton of

trypanosomatids. *Molecular and Biochemical Parasitology*, 237, 111278. <https://doi.org/10.1016/j.molbiopara.2020.111278>

Vizcaíno-Castillo, A. (2020). Caracterización de la expresión celular de TcAcT2, una actina divergente de *Trypanosoma cruzi*. Tesis que para optar por el grado de doctorado. Universidad Nacional Autónoma de México

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T., Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research*, 46(W1), W296–W303. <https://doi.org/10.1093/nar/gky427>

Webb, B., & Sali, A. (2021). Protein Structure Modeling with MODELLER. *Methods in Molecular Biology* (Clifton, N.J.), 2199, 239–255. https://doi.org/10.1007/978-1-0716-0892-0_14

Webb, B., & Sali, A. (2016). Comparative Protein Structure Modeling Using MODELLER. *Current Protocols in Bioinformatics*, 54, 5.6.1–5.6.37. <https://doi.org/10.1002/cpbi.3>

What is the role of formin in actin polymerization?. (s.f.). MBINFO Defining Mechanobiology. <https://www.mechanobio.info/cytoskeleton-dynamics/what-is-the-cytoskeleton/what-are-actin-filaments/what-is-the-role-of-formin/>

Williams, T. and Kelley, C. Gnuplot: an interactive plotting program. 2016

World Health Organization (WHO). 2020. Chagas disease (American trypanosomiasis). Acceso en: 1 de julio de 2022. Disponible en: https://www.who.int/health-topics/chagas-disease#tab=tab_1

Yamashita M, Higashi T, Suetsugu S, Sato Y, Ikeda T, Shirakawa R, Kita T, Takenawa T, Horiuchi H, Fukai S, Nureki O. (2007). Crystal structure of human DAAM1 formin homology 2 domain. *Genes Cells*, (11):1255-65. doi: 10.1111/j.1365-2443.2007.01132.x

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

Yuan, Z., Bailey, T. L., & Teasdale, R. D. (2005). Prediction of protein B-factor profiles. *Proteins*, 58(4), 905–912. <https://doi.org/10.1002/prot.20375>

10. ANEXOS

ANEXO 1. Características de los equipos empleados para la ejecución de los modelados, el acoplamiento y la dinámica molecular.

```
Linux barracuda 2.6.32-754.35.1.el6.x86_64 #1 SMP Sat Nov 7 12:42:14 UTC 2020
x86_64 x86_64 x86_64 GNU/Linux
[martin@barracuda ~]$ lscpu
Architecture:          x86_64
CPU op-mode(s):        32-bit, 64-bit
Byte Order:             Little Endian
CPU(s):                 20
On-line CPU(s) list:   0-19
Thread(s) per core:    1
Core(s) per socket:    10
Socket(s):              2
NUMA node(s):          2
Vendor ID:              GenuineIntel
CPU family:             6
Model:                  62
Model name:             Intel(R) Xeon(R) CPU E5-2670 v2 @ 2.50GHz
Stepping:               4
CPU MHz:                2501.000
BogoMIPS:               4999.29
Virtualization:        VT-x
L1d cache:              32K
L1i cache:              32K
L2 cache:               256K
L3 cache:               25600K
NUMA node0 CPU(s):     0-9
NUMA node1 CPU(s):     10-19
```

```
NVIDIA-SMI 340.29 Driver Version: 340.29 |
+-----+-----+-----+-----+
| GPU Name Persistence-M| Bus-Id  Disp.A | Volatile Uncorr. ECC |
| Fan  Temp  Perf  Pwr:Usage/Cap|      Memory-Usage | GPU-Util  Compute M. |
+=====+=====+=====+=====+
| 0 Tesla K20c      Off | 0000:02:00.0  Off |          0 |
| 56%  62C  Po    49W / 225W | 1698MiB / 4799MiB |  0%    Default
```

Anexo 2. Genes ortólogos de la Formina A, B y C identificados en cepas de tripanosomátidos empleando la función de identificador de ortólogos y parálogos presente en el servidor TriTrypDB y utilizando como plantilla la secuencia de las regiones codificadoras de Formina A (TcCLB.511755.80) y la Formina B (TcCLB.508641.320) del organismo *Trypanosoma cruzi* CL Brener No-Esmeraldo-like y la secuencia de la secuencia de la Formina C (TcCLB.511393.30) de *Trypanosoma cruzi* CL Brener Esmeraldo-like.

NOMBRE	ESPECIE	FORMINA A		FORMINA B		FORMINA C	
		GEN I.D	A.A	GEN I.D	A.A	GEN I.D	A.A
Blechomonas							
Bodo	B. ayalai B08-376	Baya_014_0670	1287	Baya_006_0830	1099	BSAL_71720	967
	B. saltans strain Lake Konstanz	BSAL_06335	679	BSAL_66050	1257		
		BSAL_16265	477				
Crithidia	C. fasciculata strain Cf-Cl	CFAC1_090016800	1106	CFAC1_210020500	1336		
		CFAC1_090016900	1512	EMOLV88_240011100	953		
	P. confusum CUL13					PCON_0035510	954
Endotrypanum	E. monterogeii strain LV88	EMOLV88_170013900	1239	EMOLV88_240011200	248		
Leishmania	L. aethiopica L147			LAEL147_000384900	1195		
	L. amazonensis MHOM/BR/71973/M2269	LAMA_000267900	584	LAMA_000464800	599		
	L. arabica strain LEM1108	LARLEM1108_170015000	1439	LARLEM1108_240016100	1195		
	L. braziliensis MHOM/BR/75/M2903	LBRM2903_170016000	1450	LBRM2903_240017100	1175		
	L. braziliensis MHOM/BR/75/M2904	LbrM.17.0950	1450	LbrM.24.1120	1175		
	L. braziliensis MHOM/BR/75/M2904 2019	LbrM.17.2.000950	1450	LbrM.24.2.001120	1175		

	L. donovani BPK282A1	LdBPK_171040.1	1436	LdBPK_241130.1	1183		
	L. donovani CL-SL	LdCL_170016300	1288	LdCL_240016300	1183		
	L. donovani strain LV9	LdBPK.17.2.001040	1449	LdBPK.24.2.001130	1183		
	L. enriettii strain LEM3045	LENLEM3045_170015000	1456	LENLEM3045_240016300	1190		
	L. gerbilli strain LEM452	LGELEM452_170011200	1447	LGELEM452_240017000	1197		
	L. infantum JPCM5	LINF_170016400	1288	LINF_240016300	1183		
	L. major strain Friedlin	LmjF.17.0930	1448	LmjF.24.1110	1193		
	L. major strain LV39c5	LMJLV39_170016500	1446	LMJLV39_240016900	1193		
	L. major strain SD 75.1	LMJSD75_170016100	1451	LMJSD75_240016400	1193		
	L. mexicana MHOM/GT/2001/U1103	LmxM.17.0930	1291	LmxM.24.1110	1194		
	L. panamensis MHOM/COL/81/L13	LPAL13_170015100	862	LPAL13_240016900	1174		
		LPAL13_170015200	564				
	L. panamensis strain MHOM/PA/94/PSC-1	LPMP_170920	1406	LPMP_241090	1174		
	L. sp. MAR LEM2494	LMARLEM2494_170014600	1469	LMARLEM2494_240016800	1198		
	L. tarentolae Parrot-TarII	LtaP17.1100	548	LtaP24.1200	568		
	L. tropica L590	LTRL590_170014500	1449	LTRL590_240016900	1200		
	L. turanica strain LEM423	LTULEM423_170014800	1449	LTULEM423_240016500	1197		
Leptomonas	L. pyrrhocoris H10	LpyrH10_21_1020	1030	LpyrH10_11_1220	1230		
		LpyrH10_21_1030	876				
		LpyrH10_21_1040	476				
	L. seymouri ATCC 30220	Lsey_0872_0010	428	Lsey_0152_0150	1222		
Paratrypanosoma	-	-	-	PCON_0049250	1078		
Trypanosoma	T. brucei Lister strain 427	Tb427.05.2300	943	Tb427tmp.02.3470	1004		
	T. brucei Lister strain 427 2018	Tb427_050029200	943	Tb427_110062500	996		
	T. brucei brucei TREU927	Tb927.5.2300	943	Tb927.11.5740	1004		

T. brucei gambiense DAL972	Tbg972.5.3240	943	Tbg972.11.6480	987		
T. congolense IL3000			TcIL3000.11.6110	1023		
			TcIL3000_0_32260	582		
T. congolense IL3000 2019			TcIL3000.A.H_000916500:pseudogene	1005		
T. cruzi Brazil A4	TcBrA4_0134890	1153	TcBrA4_0009510	973	TcBrA4_0000490	935
T. cruzi CL Brener Esmeraldo-like	TcCLB.511313.30	1178	TcCLB.506203.80	950	TcCLB.503565.10	228
T. cruzi Dm28c 2014	TCDM_09235	1057	TCDM_00772	880	TCDM_07992	1009
T. cruzi Dm28c 2017	BCY84_16838	1089	BCY84_02309	973	BCY84_12807	935
T. cruzi Dm28c 2018	C4B63_25g186	1076	C4B63_28g151	514	C4B63_87g26	935
			C4B63_28g152	172		
			C4B63_28g153	227		
T. cruzi Sylvio X10/1			TcSYL_0047600	334	TcSYL_0202050	523
					TcSYL_0202060	135
T. cruzi Sylvio X10/1-2012	TCSYLVIO_008994	1042	TCSYLVIO_009903	271		
T. cruzi TCC	C3747_23g262	1097	C3747_11g455	950	C3747_220g67	936
	C3747_24g92	1178	C3747_212g35	968	C3747_15g64	937
	TcYC6_0048650	1126				
T. cruzi Y C6			TcYC6_0082080	962	TcYC6_0099270	935
T. cruzi marinkellei strain B7	Tc_MARK_7659	1072	Tc_MARK_307	977	Tc_MARK_2235	935
T. rangeli SC58			TRSC58_07063	968	TRSC58_01689	269
T. evansi strain STIB 805	TevSTIB805.5.2690	943	TevSTIB805.11_01.5950	1003		
T. grayi ANR4	DQ04_02931060	685	DQ04_08191000	450	DQ04_01061020	939
	DQ04_02931070	458	DQ04_10221000	501		
T. theileri isolate Edinburgh	TM35_000201660	1232	TM35_000151360	881	TM35_000371130	961
			TM35_002551000	506		
T. vivax Y486	TvY486_0501680	950	TvY486_1105960	1017		

ANEXO 3. Identificador de la secuencia de las Forminas A, B y C completas que se encuentran codificadas en el genoma de especies de tripanosomátidos empleadas para el árbol filogenético.

Espece	Cepa	Abreviatura	ID Formina A	ID Formina B	ID Formina C
<i>Trypanosoma cruzi</i>	CL Brener S	<i>Tcru</i>	TcCLB.511313.30	TcCLB.506203.80	TcCLB.511393.30
<i>Trypanosoma rangeli</i>	Tran	<i>Tran</i>		TRSC58_07063	TRSC58_01689
<i>Trypanosoma grayi</i>	ANR4	<i>Tgra</i>			DQ04_01061020
<i>Trypanosoma theileri</i>	Edinburgh	<i>Tthe</i>	TM35_000201660		TM35_000371130
<i>Trypanosoma brucei</i>	TREU927	<i>Tbru</i>	Tb927.5.2300	Tb927.11.5740	
<i>Trypanosoma congolense</i>	Tcon	<i>Tcon</i>		TcIL3000.11.6110	
<i>Trypanosoma evansi</i>	STIB 805	<i>Teva</i>	TevSTIB805.5.2690	TevSTIB805.11_01.5950	
<i>Trypanosoma vivax</i>	Y486	<i>Tviv</i>	TvY486_0501680	TvY486_1105960	
<i>Leishmania major</i>	Friedlin	<i>Lmaj</i>	LmjF.17.0930	LmjF.24.1110	
<i>Leishmania donovani</i>	LV9	<i>Ldon</i>	LdBPK_171040.1	LdBPK_241130.1	
<i>Leishmania mexicana</i>	MHOM/GT/2001/U1103	<i>Lmex</i>		LmxM.24.1110	
<i>Leishmania braziliensis</i>	MHOM/BR/75/M2903	<i>Lbra</i>	LBRM2903_170016000	LBRM2903_240017100	
<i>Leptomonas pyrrocoris</i>	H10	<i>Lpyr</i>		LpyrH10_11_1220	
<i>Bleptomonas ayalai</i>	B08-376	<i>Baya</i>	Baya_014_0670	Baya_006_0830	
<i>Crithidia fasciculata</i>	Cf-CI	<i>Cfas</i>	CFAC1_090016900	CFAC1_210020500	
<i>Endotrypanum monterogeii</i>	LV88	<i>Emon</i>			
<i>Paratrypanosoma confusum</i>	CUL13	<i>Pcon</i>	PCON_0046840	PCON_0049250	PCON_0035510
Bodo saltans	Lake Konstanz	<i>Bsal</i>		BSAL_66050	BSAL_71720