



**UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS**

UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS
INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS

CENTRO DE INVESTIGACIÓN EN DINÁMICA CELULAR

***“Caracterización de la superficie molecular de TBPs
de parásitos de humanos, hacia la identificación de
sitios para el diseño o repropósito de fármacos”***

TESIS

QUE PARA OBTENER EL GRADO DE

DOCTOR EN CIENCIAS

PRESENTA

JOSÉ ÁNGEL SANTIAGO TERRONES

DIRECTOR DE TESIS

Dra. Carmen Nina Pastor Colón

CUERNAVACA, MORELOS

Noviembre, 2019

Jurado Evaluador

Dr. Lorenzo Segovia Forcella

IBT-UNAM

Presidente

Dr. Carlos Daniel Amero Tello

CIDC-UAEM

Secretario

Dra. Laura Domínguez Dueñas

FQ-UNAM

Vocal

Dr. Samuel Enoch Estrada Soto

FF-UAEM

Vocal

Dr. Martín González Andrade

FM-UNAM

Vocal

Dra. Carmen Nina Pastor Colón

CIDC-UAEM

Suplente

Dr. Rodrigo Said Razo Hernández

CIDC-UAEM

Suplente

Agradecimientos

Al laboratorio de Dinámica de Proteínas, a CONACyT por la beca de doctorado (becario 292986) y al proyecto INFR-2014-02-231509. Al Centro Nacional de Supercómputo (CNS, IPICYT-San Luis Potosí), al clúster híbrido de Supercómputo (Cinvestav, Ciudad de México) y al Laboratorio Nacional de Supercómputo (LNS, Puebla) por los recursos computacionales y la asistencia técnica.

Índice

Jurado Evaluador	II
Agradecimientos	III
Resumen.....	V
Lista Siglas y Abreviaturas.....	VI
Lista de Figuras.....	VII
Lista de Tablas	VIII
1. Introducción.....	1
1.1 Búsqueda de sitios de unión a fármacos en superficies de proteínas	3
1.2 La familia de TBPs como un posible blanco molecular.	5
1.3 Ventajas de usar TBP como blanco molecular de enfermedades parasitarias.....	7
2. Fundamento teórico	8
2.1 Justificación	8
2.2 Hipótesis	8
2.3 Objetivos	8
2.3.1 Objetivo general.....	8
2.3.2 Objetivos particulares	9
3. Estrategia Computacional.....	9
3.1 Selección de TBPs y generación de modelos	9
3.2 Simulaciones de Dinámica molecular.....	11
3.3 Cálculo de parámetros estructurales de la superficie.	11
3.4 Acoplamiento molecular de compuestos sobre la superficie de las TBPs.	13
4. Resultados.....	16
4.1 Diferencias en secuencia de las TBPs de parásitos comparadas a la TBP humana.	16
4.2 Análisis de las propiedades de superficie en las TBPs.	18
4.3 Presencia de cavidades sobre las TBPs.....	25
4.4 Evaluación computacional de la unión de compuestos tipo fármaco en las TBPs.....	29
5. Discusión	41
6. Conclusiones	45
7. Perspectivas	45
8. Bibliografía	46
9. Anexos	VIII

Resumen

Las parasitosis son enfermedades de gran interés a nivel mundial, debido a su rápido contagio y a la adquisición de resistencia por los parásitos. La búsqueda de blancos moleculares novedosos y su caracterización estructural por métodos computacionales, ha sido aplicado para encontrar nuevos sitios y proponerlos como potenciales opciones de tratamientos contra diferentes enfermedades parasitarias. La proteína de unión a cajas TATA (TBP, por sus siglas en inglés) es uno de los factores de transcripción centrales en organismos eucariotas y juega un papel substancial durante el ensamble del complejo de pre-inicio de la transcripción por las tres ARN polimerasas. Debido que TBP interactúa con un gran número de factores de transcripción, interferir con estas interacciones sería letal para la sobrevivencia del parásito. En este trabajo presentamos el primer estudio computacional de las propiedades de superficie del dominio de unión a ADN de TBP para un listado de parásitos involucrados en enfermedades parasitarias de interés mundial. Las diferencias estructurales y de secuencia de estas TBPs con respecto a la TBP de humano fueron propuestas para seleccionar ensambles de estructuras representativos generados por simulaciones de dinámica molecular (MD, por sus siglas en inglés). Para evaluar el potencial de estas TBP como blanco molecular, llevamos a cabo acoplamiento molecular de ensamble usando una biblioteca de compuestos tipo-fármaco, seleccionados en base a una combinación de las reglas de Lipinski y Veber. Encontramos que las principales diferencias estructurales y dinámicas están presentes en el sitio de unión de NC2, el extremo N-terminal, la hélice H2 y la región interdominio. Estos sitios corresponden a potenciales sitios de unión a fármacos en estas TBPs. Encontramos una buena selectividad de los compuestos para las TBPs de *Plasmodium falciparum*, *Necator americanus*, *Candida albicans*, *Entamoeba histolytica* y *Taenia solium*, en comparación a la TBP humana. Los mejores compuestos comparten similitud estructural entre ellos y con constantes de unión calculadas (K_d) en el intervalo nM a μ M. Estos compuestos pueden ser propuestos como andamios estructurales para pruebas experimentales y optimización adicional. Con los resultados obtenidos, proponemos a TBP como un blanco molecular novedoso para el tratamiento de parasitosis humanas.

Lista Siglas y Abreviaturas

2D	Dos dimensiones
3D	Tres dimensiones
μM	Micromolar
Å	Ångström
AAS	Área accesible al solvente
APBS	<i>Adaptative Poisson-Boltzmann Solver</i>
ARN	Ácido Ribonucleico
ARNpol	ARN polimerasa
CHARMM	<i>Chemistry and HARvard Molecular Mechanics</i>
C-ter	Carboxilo Terminal
fs	femtosegundo
I-TASSER	<i>Iterative Implementation of Threading ASSEmbly Refinement</i>
K	Grados Kelvin
Kd	Constante de unión
LogP	Coefficiente de Partición Octanol/Agua
MD	Dinámica Molecular
NAMD	<i>Not (just) Another Molecular Dynamics</i>
NCBI	<i>National Center for Biotechnology Information</i>
nM	Nanomolar
ns	nanosegundo
N-ter	Amino Terminal
PDB	<i>Protein Data Bank</i>
ps	picosegundo
RMN	Resonancia Magnética Nuclear
RMSD	<i>Root-Mean-Square Deviation</i> (raíz de la desviación cuadrática media)
TBP	TATA-Binding Protein
VdW	Van der Waals
VMD	<i>Visual Molecular Dynamics</i>

Lista de Figuras

Figura 1. Estructura del dominio de unión a ADN de TBP (PDB ID 1NVP).	6
Figura 2. Diagrama del procedimiento computacional usado para seleccionar ensambles de estructuras de las MD con mayores diferencias respecto a la TBP de humano y llevar a cabo acoplamiento molecular..	14
Figura 3. Diferencias a nivel de secuencia en las TBPs de parásitos.	18
Figura 4. Polaridad de la superficie de TBPs de parásitos.	19
Figura 5. Diferencias en AAS relativa por residuo con respecto a hsaTBP.	20
Figura 6. Electroestática en TBPs.	21
Figura 7. Superficie de potencial electrostático en TBPs.	22
Figura 8. Interacciones polares en las TBPs. Interacciones polares en las TBPs de parásitos.	23
Figura 9. Puentes salinos de cadena lateral en las TBPs.	24
Figura 10. Cavidades calculadas con Metapocket sobre las TBPs.	26
Figura 11. Dinámica de apertura en las TBPs calculada con POVME 3.0.	27
Figura 12. Dinámica de apertura en las TBPs calculadas con POVME 3.0.	28
Figura 13. Sitios potenciales de unión a compuestos sobre las TBPs de parásitos.	32
Figura 14. Compuestos con la mejor selectividad obtenidos del acoplamiento flexible.	37
Figura 15. Dinámica de apertura del extremo N-terminal en las TBPs.	38
Figura 16. Compuesto con la más alta selectividad sobre el estribo N-ter en pfaTBP.	38
Figura 17. Top seis de los compuestos con más alta selectividad sobre los sitios de TBPs de parásitos.	40

Lista de Tablas

Tabla 1. Enfermedades causadas por parásitos eucariotas en México.....	1
Tabla 2. Lista de organismos y sus enfermedades asociadas.	9
Tabla 3. Lista de TBPs analizadas en este trabajo.	10
Tabla 4. Resultados del agrupamiento por RMSD-2D.....	12
Tabla 5. Número de estructuras usadas para el acoplamiento de compuestos tipo fármaco.	15
Tabla 6. Mejores cinco compuestos con selectividad en las TBPs de parásitos con respecto a hsaTBP.....	30
Tabla 7. Compuestos con mejor selectividad con respecto a hsaTBP, obtenidos del acoplamiento flexible sobre los sitios seleccionados para cada TBP.	32

Anexos

Anexo 1. Validación estructural de los modelos de las TBPs usados para MD.....	VIII
Anexo 2. RMSD contra la estructura inicial para cada TBP.....	IX
Anexo 3. Diferencias en fluctuaciones de carbonos alfa con respecto a hsaTBP.	X
Anexo 4. Puente salino entre el asa S3-S4 y la hélice H2 en TBPs.....	XI
Anexo 5. Cavidad en el extremo N-terminal en calTBP, ovoTBP y tsoTBP.	XII
Anexo 6. Resultados del acoplamiento molecular flexible sobre el sitio extremo N-terminal de tsoTBP y ovoTBP.	XII

1. Introducción

A pesar del surgimiento de medicamentos y nuevos programas de control, las parasitosis continúan siendo un gran problema de salud pública a nivel mundial, debido a la resistencia exhibida por los parásitos, rápida transmisión, cambios en los hábitos alimenticios y cambio climático, entre otros.¹ Estas enfermedades afectan fundamentalmente a la población de países en vías de desarrollo, incluyendo las parasitosis más comunes que promueven enfermedades gastrointestinales leves (causadas por algunos protozoarios y helmintos), y enfermedades mortales que son de importancia mundial (ej. enfermedad de Chagas, tripanosomiasis africana, leishmaniasis, malaria, filariasis linfática). La infecciones simultáneas con múltiples parásitos pertenecientes a diferentes filos también aumentan su propagación (por ejemplo en pacientes con helmintos aumenta la morbilidad por infecciones con malaria).² A su vez, las zoonosis emergentes (transmitidas por vectores y por alimentos) generadas por el desplazamiento de humanos a ecosistemas silvestres, el contacto con animales silvestres y domésticos, también potencia las infecciones parasitarias. En el caso de México, algunas de las enfermedades más comunes causadas por parásitos incluyen al Mal de Chagas, amibiasis, giardiasis, ascariasis, oncocercosis, necatoriasis, helmintiasis y trichuriasis, entre otras (Tabla 1).^{3,4}

Tabla 1. Enfermedades causadas por parásitos eucariotas en México. Casos de enfermedades por parásitos eucariotas reportados en todo el 2018 y hasta junio de 2019 de acuerdo a la Dirección general de epidemiología en México (<https://www.gob.mx/salud/acciones-y-programas/direccion-general-de-epidemiologia-boletin-epidemiologico>). Instituciones de las que se reportan casos: IMSS, ISSTE, IMSS PROSPERA, DIF, PEMEX, SEDENA, SEMAR, SSP, otras.

Enfermedad parasitaria	Número de casos acumulados en 2018 y hasta la semana 26 del 2019
Amebiasis	193 576
Giardiasis	11 257
Paludismo- <i>Plasmodium vivax</i>	568
Mal de Chagas	890
Otros protozoarios	70 462
Ascariasis	41 576
Oncocercosis	5

Helmintos	130 232
Candidiasis	167 750
Teniasis	392
Cisticercosis	192

Con la concomitante resistencia a fármacos, muchos de los medicamentos antiparasitarios y antibióticos muestran ya una reducida efectividad. Un ejemplo es la enfermedad de malaria en África, donde el fármaco Artemisinina es el principal medicamento usado, en combinación de pabellones impregnados con insecticidas peritroides para el control del vector. Esto ha promovido una baja en la mortalidad y de la infección durante los años 2000-2015.⁵ Sin embargo, el amplio uso de Artemisinina ha llevado a la resistencia del parásito, por lo que existe la necesidad de búsqueda de nuevos blancos terapéuticos y nuevos fármacos, que actúen directamente en contra del parásito o de su vector (en el caso de aquellos que sean transmitidos por vector).

La mayoría de los fármacos antiparasitarios actúan contra ciertas etapas del desarrollo de los parásitos. Los parásitos del filo Apicomplexa (especies de *Plasmodium*, *Toxoplasma* y *Cryptosporidium*) y Microsporidia (*Encephalitozoon cuniculi*) son parásitos intracelulares obligados, por lo que los fármacos disponibles deben entrar a la célula. Por ejemplo, el mecanismo de acción farmacológica de la mayoría de los fármacos contra la malaria actúa contra el estado intra-eritrocítico de *Plasmodium falciparum*.⁶ Sin embargo también es de interés atacar otros estados (esquizontes en hígado) y las formas infectivas presentes en los mosquitos (esporozoítos), para el control sobre el propio vector. Ejemplo de esto último es usar los medicamentos disponibles contra el parásito impregnados sobre los cubrecamas en lugar de insecticidas.⁷

En el caso de *Trypanosoma cruzi*, que causa la Enfermedad Chagas, los fármacos existentes son efectivos sólo si son administrados en las etapas tempranas de la infección, además de que son muy tóxicos,⁸ por lo que al igual que con *Plasmodium* es necesario buscar otras alternativas. Parásitos como *Giardia intestinalis* y *Entamoeba histolytica*, que muestran un metabolismo anaeróbico que los distingue de los demás eucariotas, han sido tratados con fármacos derivados de nitroimidazoles que inhiben la síntesis de ácidos nucleicos en estos parásitos y que sólo son activados por su reducción en células anaeróbicas. Sin embargo, los derivados de nitroimidazoles muestran alta toxicidad y reducida efectividad debido a la resistencia de estos organismos.⁹

Para la mayoría de gusanos nemátodos y tremátodos, los derivados del benzimidazol han mostrado ser efectivos en su eliminación; por ejemplo el albendazol inhibe la

polimerización de microtúbulos en el parásito, lo que resulta en alteraciones en las células del tegumento y del intestino del gusano, que posteriormente conlleva a una deficiente captación de glucosa por los parásitos, disminuyendo así los depósitos de energía del gusano.¹⁰ Otros fármacos son efectivos en la eliminación de estos parásitos y son usados como alternativa; tal es el caso de la ivermectina, la cual se une a canales de cloro activados por glutamato en células nerviosas y de músculo de parásitos, y causan un incremento en la permeabilidad celular a iones cloro, resultando en la parálisis y muerte de los parásitos.¹¹ Sin embargo, parásitos resistentes han surgido debido al uso de estos fármacos.

Esta reducida efectividad ha llevado a la propuesta de nuevas proteínas como blancos terapéuticos. En relación a esto, la selección de nuevos blancos ha sido principalmente orientada hacia la búsqueda de moléculas específicas existentes en el parásito y no en el organismo huésped, por ejemplo la inhibición de proteínas que participan en procesos esenciales para el parásito. Un ejemplo es el tratamiento de la infección de *Trypanosoma cruzi* (Mal de Chagas) con el fármaco Nifurtimox, el cual inhibe la ruta de formación de aminos como la espermina y espermidina, los cuales son esenciales para el parásito pero no para el ser humano;¹² por otra parte el fármaco antifúngico Posaconazol ha mostrado una limitada actividad contra *Trypanosoma cruzi*, inhibiendo la formación de ergosteroles esenciales para el parásito.¹³

Aunque en la búsqueda de blancos moleculares se prefiera dirigir fármacos contra proteínas específicas de los parásitos, algunas veces resulta difícil encontrarlos debido a la estrecha relación filogenética en eucariotes, por lo que otra de las aproximaciones usadas es buscar diferencias de secuencia o estructurales en proteínas conservadas. Estas diferencias pueden ser propuestas como sitios de unión a fármacos. Con respecto a esto último podemos encontrar algunos ejemplos: existe una amplia cantidad de compuestos selectivos dirigidos aminoacil-tRNAs sintetasas de algunos parásitos (*Plasmodium*, *Trypanosoma brucei*, *Brugia malayi*) con el fin de afectar el proceso de síntesis de proteínas selectivamente en parásitos.⁸ Otro ejemplo incluyen compuestos dirigidos contra cinasas de *Leishmania*, donde al igual que las aminoacil-tRNAs sintetasas, sus estructuras presentan cavidades exclusivas en comparación a las de humano.¹⁴

1.1 Búsqueda de sitios de unión a fármacos en superficies de proteínas

Para que una proteína sea un blanco molecular, ésta necesita contener un sitio de unión que posea las características estructurales y propiedades fisicoquímicas que favorezcan la unión de un compuesto (fármaco) con alta afinidad y especificidad.^{15,16}

Una alta afinidad garantiza mejor potencia del fármaco y así dosis más pequeñas mostrarán el efecto deseado. La especificidad garantizará una unión preferente a la proteína del parásito en comparación a la del hospedero, y por tanto no afectará en gran medida las funciones celulares del hospedero.

Diferentes aproximaciones experimentales han sido usadas para evaluar sitios potenciales en una proteína blanco, como resonancia magnética nuclear (RMN), calorimetría de titulación isotérmica y espectrometría de masas.¹⁷ El tamizaje de alto rendimiento o tamizaje de fragmentos seguido por RMN ha sido una de las aproximaciones ampliamente usadas para evaluar la accesibilidad farmacológica (*drugability* en inglés), ya que se pueden probar una gran cantidad de compuestos contra la proteína blanco y se seleccionan aquellos que interaccionan con buena afinidad. Sin embargo, estos métodos son costosos y requieren mucho tiempo, además de que dependen de la calidad y cantidad de proteína pura y de la diversidad de los compuestos probados.¹⁵

Otra aproximación es utilizar métodos computacionales, donde los sitios blanco pueden ser identificados si se dispone de la estructura de la proteína – determinada por RMN o cristalografía de rayos X y almacenadas en bases de datos como el Protein Data Bank (PDB)¹⁸ o de un modelo – mediante algoritmos de detección de cavidades.¹⁹ Una vez identificado el sitio, una biblioteca de compuestos es muestreada sobre la superficie de la proteína. Esto se conoce como tamizaje virtual o tamizaje *in silico*.¹⁷ Al igual que el tamizaje experimental, el tamizaje virtual puede ser usado para evaluar la accesibilidad farmacológica e identificar los sitios con mayor afinidad y selectividad. Una de las ventajas del tamizaje virtual es que permite reducir el número de fármacos a evaluar experimentalmente. La evaluación de la afinidad de cada compuesto se lleva a cabo con algoritmos de acoplamiento molecular (Docking). Cada algoritmo de acoplamiento está compuesto por una función de búsqueda y una función de puntaje. La función de búsqueda está relacionada con la generación de conformaciones o poses, mientras que la función de puntaje está relacionada con la afinidad del compuesto.

Hay múltiples algoritmos de acoplamiento disponibles (Autodock4, Autodock Vina, Dock, Glide, FlexX, Phdoc, etc.), incluyendo programas de acoplamiento rígido y flexible, cada uno con sus funciones de puntaje basadas en diferentes criterios (complementaridad de forma, interacciones de enlace de hidrógeno, de van der Waals, e hidrofóbicas).^{20,21} Una lista de diferentes algoritmos y programas de acoplamiento se pueden encontrar en la página <http://www.click2drug.org/index.html>.

Por su parte, la biblioteca de fármacos puede ser obtenida de diferentes bases de datos comerciales y gratuitas (como la base de datos ZINC15²²), seleccionando

aquellos ligandos con las características deseadas, si es que se tiene idea del tipo de ligando que interacciona con la proteína blanco.

Dado que la búsqueda computacional de sitios blanco hace uso de estructuras 3D de las proteínas, es de crucial importancia investigar la influencia de variaciones estructurales en la predicción de dichos sitios. Tales variaciones estructurales van desde ligeros ajustes de cadenas laterales, hasta grandes movimientos que pueden dar lugar a cavidades transitorias o cavidades crípticas, no identificadas experimentalmente.^{17, 19} Por lo tanto, no es suficiente usar una sola estructura para el acoplamiento, y así el reto actual es analizar ensambles de conformaciones de las proteínas.^{17, 19} Tales ensambles pueden ser obtenidos de estructuras de RMN, uso de múltiples estructuras cristalográficas, o como alternativa usar ensambles generados por MD.

Las combinación de MD, acoplamiento molecular y algoritmos de predicción de cavidades han sido aplicados para estudiar la dinámica de apertura de cavidades transitorias y ha permitido la caracterización de potenciales sitios de unión a fármacos usando estos métodos computacionales.^{15,21,23,24,25,26} Esto permite proponer ensambles representativos de estructuras para evaluar su accesibilidad farmacológica por algoritmos de acoplamiento molecular de ensamble (conocido como *ensamble-based docking*).^{27, 28} El uso de estas técnicas computacionales ofrece una buena herramienta para la búsqueda y caracterización de sitios blanco en la superficie de proteínas, que no tienen un sitio de unión definido como las enzimas o receptores, y los cuales son más difíciles de encontrar.

1.2 La familia de TBPs como un posible blanco molecular.

Los fármacos disponibles contra parásitos están implicados en inhibir procesos esenciales (tales como la síntesis de ADN, metabolismo, traducción de proteínas etc.), por lo que atacar un blanco molecular involucrado en estos procesos podría ser letal para el parásito. En ese sentido la transcripción de genes es un proceso esencial para la supervivencia en todo ser vivo. En organismos eucariotas la maquinaria basal está ampliamente conservada entre diferentes especies.^{29, 30} La transcripción es llevada a cabo por tres ARN polimerasas (ARNpol). Todas las ARNpol interaccionan con distintos conjuntos de proteínas llamados factores de transcripción.

Uno de los factores de transcripción centrales es la TBP ("TATA binding protein"- en inglés), el cual está también presente en arqueas.³¹ TBP reconoce la caja TATA – localizada aproximadamente a 30pb río arriba del sitio de inicio de la transcripción – y promueve el ensamble del complejo de pre-inicio de la transcripción por las tres

ARNpol. Mientras el extremo N-ter es desordenado, altamente divergente en secuencia y longitud, la estructura del dominio de unión a ADN está ampliamente conservada entre diferentes especies, con una longitud de alrededor de 180 residuos (Figura 1a),^{32, 33} por lo que en este estudio no consideraremos el extremo N-ter. Estructuras para este dominio han sido resueltas por rayos X o RMN para las TBP de *Saccharomyces cerevisiae*,³⁴ *Arabidopsis thaliana*,³⁵ *Homo sapiens*,³⁶ *Encephalitozoon cuniculi*,³⁷ *Pyrococcus woesei*,³⁸ *Methanococcus jannaschii*,³² y *Sulfolobus acidocaldarius*.³⁹ Prácticamente, toda la superficie de TBP establece interacciones con otros factores en el contexto del complejo de pre-inicio de la transcripción, tales como TFIIA y TFIIB, el regulador transcripcional MOT1, el activador Brf1 (del complejo de la ARNpol III), el factor asociado a TBP de 250 kDA (TAF1), y el inactivador transcripcional NC2, entre otros (Fig. 1b), para los cuales las estructuras cristalográficas en complejo con TBP están disponibles.^{29, 40, 41}

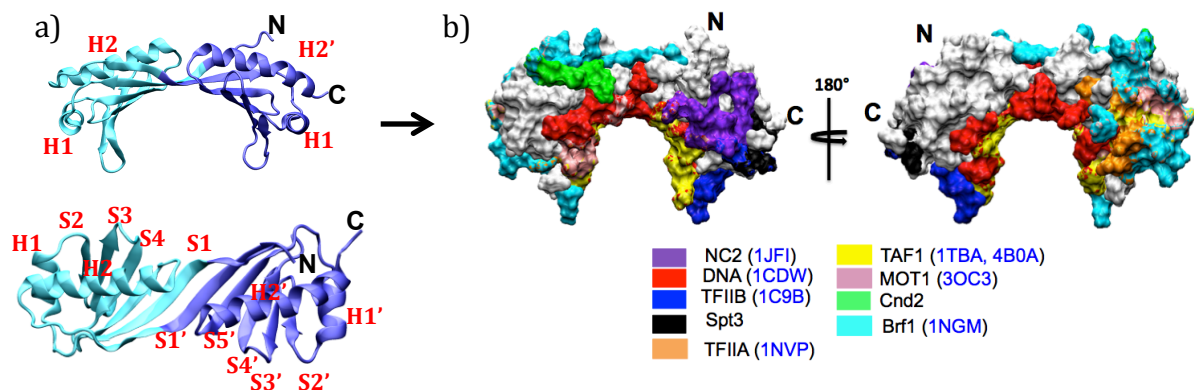


Figura 1. Estructura del dominio de unión a ADN de TBP (PDB ID 1NVP). a) La estructura muestra los dominios simétricos en cian y en azul, así como los diferentes elementos de estructura secundaria: S: hebra beta, H: hélice. b) Vista frontal (izquierda) y trasera (derecha) de los sitios de interacción de TBP con otras proteínas de acuerdo a las estructuras cristalográficas depositadas en el PDB e identificados experimentalmente. El código de colores se indica debajo de la figura, así como el PDB:ID.

Otros interactores de TBP sin estructura reportada son la proteína Spt3 del complejo SAGA, el dominio de activación Gal4 y el C-ter de la subunidad Cnd2 del complejo de condensinas.⁴² El complejo supramolecular resultante representa un blanco molecular atractivo para el tratamiento de enfermedades parasitarias, ya que suficiente divergencia existe entre la TBP humana y la del parásito.

A pesar de la conservación del dominio de unión a ADN de TBP (Fig. 1a), las TBPs de parásitos eucariotas muestran divergencia en su composición de aminoácidos,⁴³ y diferente preferencia por sus cajas TATA respecto a la TBP humana (como las TBPs de *Entamoeba histolytica*, *Plasmodium falciparum* y *Cryptosporidium parvum*), lo que

sugiere diferencias estructurales y en sus propiedades de superficie.^{43, 44, 45, 46} Por ejemplo, diferencias a nivel de secuencia en la hélice H2 (que muestra un mayor número de aminoácidos positivos) han sido reportadas en las TBPs de *Plasmodium falciparum* y *Cryptosporidium parvum*, así como también una asa más larga entre las hebras S2' y S3'.^{45, 46} Respecto a la información estructural disponible de TBPs de parásitos, sólo la estructura de la TBP del parásito microsporidio *Encephalitozoon cuniculi* ha sido reportada.^{37, 47} Sin embargo, un gran número de secuencias de TBPs de diferentes parásitos están disponibles en el NCBI (*Plasmodium falciparum*,⁴⁸ *Cryptosporidium parvum*,⁴⁶ *Trypanosoma brucei*,⁴⁹ *Trypanosoma cruzi*,⁵⁰ *Candida albicans*,⁵¹ *Taenia solium*,⁵² *Entamoeba histolytica*,^{44, 53} *Oncocerca volvulus*,⁵⁴ *Encephalitozoon cuniculi*,^{37, 47} *Pneumocystis carinii*,⁵⁵ entre otros), permitiendo generar modelos estructurales con buena confianza. La búsqueda de diferencias en la superficie de proteínas de parásitos respecto a humano ha sido usada para proponer sitios potenciales, por lo que un análisis de las propiedades de superficie de TBPs de parásitos podría resultar en su propuesta como blanco molecular novedoso.

1.3 Ventajas de usar TBP como blanco molecular de enfermedades parasitarias

Los argumentos para proponer a TBP como un nuevo blanco molecular son: 1) TBP es una proteína esencial y por tanto inhibirla sería letal para el parásito; 2) a pesar de la similitud estructural, las diferencias en secuencia darían lugar a diferencias en la superficie de la proteína, que podrían ser propuestas como sitios blancos. Esta estrategia ha funcionado en otras proteínas conservadas;^{8, 14, 56} 3) la extensa superficie molecular disponible abre la opción de utilizar más de un fármaco simultáneamente en terapias de combinación, ya sea para tratar infecciones con un solo agente infeccioso o provocadas por más de uno a la vez; 4) como la tasa de mutación de TBP es baja,⁵⁷ no se esperaría una tasa alta de resistencia de los parásitos; 5) a pesar de que anteriormente los factores de transcripción han sido clasificados como blancos con baja accesibilidad farmacológica (debido a que las interacciones que establecen son muy dinámicas, involucran un gran número de interacciones con proteínas y carecen de cavidades o son poco profundas), algunos caso de éxito han sido reportados.^{58, 59}

A la fecha no existe un análisis estructural de TBPs de parásitos y tampoco se han reportado inhibidores de sus interacciones con otras proteínas, lo cual lo hace un sistema atractivo de estudio para evaluar su accesibilidad farmacológica. Nuestro estudio pretende determinar si TBP es un blanco molecular potencial, en el contexto de sus múltiples interacciones proteína-proteína, buscando moléculas pequeñas que

puedan interrumpir o modificar las interacciones del complejo, resultando en un complejo de pre-inicio no funcional. En este trabajo llevamos a cabo el primer estudio de las propiedades de superficie de un conjunto de TBPs de parásitos usando métodos computacionales, buscando las principales diferencias con respecto a la TBP humana, con el fin de proponerlas como sitios blanco de unión a fármacos. Para este fin, seleccionamos sitios basados en las variaciones de propiedades de superficie con respecto a la TBP humana y la posibilidad de formación de cavidades. Para estos sitios, creamos ensamblajes representativos de estructuras obtenidas por simulaciones de dinámica molecular y llevamos a cabo acoplamiento molecular de ensamblaje, usando una biblioteca de compuestos tipo fármaco obtenidos de la base de datos de ZINC12.⁶⁰

2. Fundamento teórico

2.1 Justificación

Dada la creciente resistencia a fármacos por parásitos eucariotas que infectan humanos, existe la necesidad de la búsqueda de nuevos blancos moleculares y/o compuestos antiparasitarios. Nosotros proponemos a TBP como blanco terapéutico de enfermedades parasitarias que aquejan a humanos, debido a que salvo en estados de animación suspendida (esporas o quistes), todos los organismos realizan transcripción y por lo tanto, los fármacos atacarían casi cualquier estado del desarrollo del parásito, resultándoles letales. A pesar de la alta conservación estructural, las TBPs de parásitos muestran diferencias en sus secuencias que podrían resultar en sitios blanco novedosos. La identificación de estos sitios es clave para la propuesta de armazones iniciales, esperando mejorar la afinidad y especificidad respecto a la TBP de humano, mediante el diseño y análisis experimental a futuro.

2.2 Hipótesis

Las TBPs de parásitos muestran diferencias en sus propiedades de superficie, comparadas con la TBP de humano. Tales sitios pueden ser propuestos como sitios blanco de unión a fármacos.

2.3 Objetivos

2.3.1 Objetivo general

Caracterizar la superficie molecular de las TBPs de parásitos e identificar sitios blanco de unión a fármacos, partiendo de las diferencias en superficie respecto a la TBP de humano.

2.3.2 Objetivos particulares

1. Realizar el análisis bioinformático de las secuencias del dominio de unión a ADN de TBPs de parásitos seleccionados, modelar la estructura del dominio de unión a ADN, y hacer la comparación con la TBP de humano.
2. Caracterizar la superficie molecular de TBPs de humano y parásitos, calculando parámetros estructurales y electrostáticos de la superficie a partir de ensamblajes de estructuras generados por dinámica molecular para las TBPs. Identificar diferencias estructurales y cavidades sobre la superficie de TBPs de parásitos y proponerlas como sitios blanco de unión a fármacos.
3. Generar ensamblajes de estructuras para el acoplamiento molecular de fármacos a partir de simulaciones por dinámica molecular de las TBPs de parásitos y de humano, y analizar su plasticidad conformacional.
4. Se realizará el acoplamiento molecular de una biblioteca de compuestos tipo fármaco, sobre las TBP de humano y de parásitos.

3. Estrategia Computacional

3.1 Selección de TBPs y generación de modelos

Nosotros seleccionamos a la TBP de humano y una lista de TBPs de parásitos que incluye a diferentes filos (dos de hongos, una de amiba, dos de nemátodos, una de un platelminto y dos protozoarios), los cuales están involucrados en enfermedades de importancia mundial (Tabla 2).

Tabla 2. Lista de organismos y sus enfermedades asociadas.

Organismo	Enfermedad asociada	Filo
<i>Encephalitozoon cuniculi</i>	Encefalitozoonosis	Microsporidia
	Pneumocitosis	
<i>Pneumocystis carinii</i>	Candidiasis	Ascomycota
<i>Candida albicans</i>	Taeniasis	Ascomycota
<i>Taenia solium</i>	Oncocercosis	Platyhelminthes
<i>Onchocerca volvulus</i>	Necatoriasis	Nematoda

<i>Necator americanus</i>	Amibiasis	Nematoda
<i>Entamoeba histolytica</i>	Malaria	Amoebozoa
<i>Plasmodium falciparum</i>		Apicomplexa

Todas las secuencias de TBPs cubren el dominio de unión a ADN (~180 residuos) y muestran diferente identidad en secuencia respecto a la TBP de humano (Tabla 3). Para **hsa**TBP y **ecu**TBP, seleccionamos tres estructuras cristalográficas del PDB.^{36, 37, 47, 61, 62} Tomando ventaja de la buena conservación en secuencia entre TBPs, generamos modelos estructurales para el resto de las TBPs, utilizando tres diferentes predictores: I-TASSER,⁶³ MODELLER⁶⁴ y SWISS-MODEL.⁶⁵

Tabla 3. Lista de TBPs analizadas en este trabajo.

Organismo	PDB:ID	Abreviación	Identidad*
<i>Homo sapiens</i>	1nvp, 1ngm, 1c9b	hsa	
<i>Encephalitozoon cuniculi</i>	3eik, 3oc3, 4wzs	ecu	76.0
<i>Pneumocystis carinii</i>		pnc	82.2
<i>Candida albicans</i>		cal	79.4
<i>Taenia solium</i>		tso	76.6
<i>Onchocerca volvulus</i>		ovo	82.1
<i>Necator americanus</i>		nam	81.0
<i>Entamoeba histolytica</i>		ehi	54.4
<i>Plasmodium falciparum</i>		pfa	38.3

*El porcentaje de identidad es respecto a la TBP de humano.

Todos los modelos mostraron una estructura adecuada como calculado con RAMPAGE⁶⁶ (Anexo 1). Para el caso de **pfa**TBP, sólo los modelos generados por I-TASSER y MODELLER fueron seleccionados, debido a que SWISS-MODEL falló en generar un modelo estructuralmente estable. Todas las secuencias de TBPs sin estructuras fueron obtenidas del NCBI.⁶⁷

Partiendo de un alineamiento de secuencias obtenido por Clustal omega,⁶⁸ creamos un alineamiento estructural usando el plugin Multiseq de VMD.⁶⁹ Este alineamiento fue usado para mapear las diferencias en secuencia respecto a **hsa**TBP. Este alineamiento fue también enviado al servidor ConSurf para ver las diferencias en superficie respecto a la TBP de humano.⁷⁰ ConSurf calcula una calificación de conservación por posición, partiendo de un alineamiento múltiple de al menos cinco secuencias; en nuestro caso las TBPs de parásitos son comparadas contra la TBP de

humano. La estructura de la **hsaTBP** (PDBID 1NVP) fue usada para mapear las diferencias en superficie.

3.2 Simulaciones de Dinámica molecular.

Todas las estructuras fueron previamente optimizadas: se les añadieron hidrógenos, se minimizaron 50 pasos usando el esquema *Steepest descent* sobre los hidrógenos, manteniendo el resto de la estructura rígida, seguida de una minimización de 50 pasos sobre todos los átomos usando el esquema *Conjugate gradient*. Esta optimización se hizo con el potencial CHARMM36,⁷¹ y el programa charmm38b1.⁷² Las estructuras finales se usaron como semillas para las simulaciones de MD.

Los sistemas incluyen explícitamente todos los átomos (~30000 átomos totales), en una caja de agua ortorrómbica de 92x64x60 Å³ (modelo de agua TIP3P), y una concentración de sal de 0.15 M de NaCl. Los extremos N-terminal y C-terminal fueron considerados como neutros. Las cajas fueron preparadas con el servidor CHARMM-GUI.⁷³

Las simulaciones se realizaron en el ensamble NPT, por 101 ns para cada estructura inicial a 323 K para mejorar el muestreo conformacional. El primer ns fue utilizado como periodo de equilibración para cada réplica, rindiendo 300 ns de tiempo acumulado de simulación para cada TBP. Para la **pfaTBP**, extendimos cada modelo a 151 ns para obtener el mismo tiempo neto de simulación.

Las dinámicas fueron propagadas con *NAMD2.11*,⁷⁴ usando el potencial *CHARMM36*.⁷¹ Las interacciones no covalentes se suavizaron con una función *switch* para la de Van der Waals y de *shift* para las de Colulomb, usando un corte de 12 Å. La lista de vecinos se recalculó cada 10 pasos con un radio de corte en 14 Å. El tiempo de integración fue de 2 fs y la frecuencia de guardado de 500 pasos.

3.3 Cálculo de parámetros estructurales de la superficie.

Para evaluar la diversidad estructural de los ensambles generados por MD, llevamos a cabo un agrupamiento por RMSD en dos dimensiones (RMSD-2D), usando los 300 ns para cada TBP, ya que el RMSD contra las estructuras iniciales mostró ligeras variaciones (Anexo 2).

El RMSD-2D fue calculado sobre los átomos de cadena principal, con un corte de 2.0 Å para el análisis de agrupamiento. El programa *SIMULAID* fue usado para llevar a cabo este análisis.⁷⁵ Obtuvimos un solo grupo con más del 90% de las estructuras muestreadas para cada TBP, sugiriendo una buena convergencia de las MD. La Tabla

4 muestra el número de estructuras para el grupo más grande. Las estructuras representativas corresponden a los centros de este grupo.

Tabla 4. Resultados del agrupamiento por RMSD-2D. Se obtuvo un sólo grupo representativo con más del 90% de las estructuras muestreadas (3000 estructuras analizadas) para cada TBP.

TBP	Número de estructuras del grupo más grande
hsa	2888
ecu	2964
pnc	2941
cal	2859
tso	2722
nam	2900
ovo	2977
ehi	2960
pfa	2716

Los parámetros de superficie calculados sobre las trayectorias de MD se obtuvieron con el programa charmm38b1 y fueron las siguientes: RMSD sobre átomos pesados con respecto a la estructura inicial, fluctuaciones promedio de carbonos alfa, energía electrostática, ángulos diedros de cadena lateral (posteriormente clasificados como rotámeros con un *script* local), el total de puentes neutros (interacciones entre las cadenas laterales de Ser, Thr, Cys, Asn, Gln, His, Tyr), el total de puentes salinos de cadena lateral (interacciones entre las cadenas laterales de los residuos cargados Arg y Lys con Asp y Glu), el total de puentes cargados (interacciones entre las cadenas laterales de los residuos cargados con residuos neutros), el porcentaje de ocupación promedio de puentes salinos, el área accesible al solvente (AAS) polar y no polar, y área accesible al solvente por residuo, utilizando una esfera de prueba de 1.4 Å. El cálculo de AAS polar y no polar incluye todos los átomos polares y no polares sobre la superficie respectivamente, independiente de su distribución espacial. Los valores de AAS por residuo fueron convertidos a su valor relativo utilizando los valores de máxima área accesible al solvente por residuo,⁷⁶ y la diferencia respecto a la TBP de humano fue calculada.

La apertura de cavidades fue computada con el programa POVME 3.0,⁷⁷ usando los parámetros: a espaciamiento de 1.0 Å, una distancia de corte d 1.09 Å y una esfera de inclusión de 8 Å para el sitio de TFIIA y de 6 Å para los sitios NC2, Interdominio, N-ter y hélice H2. La esfera de inclusión fue centrada en cada sitio respectivamente. Todas

estas propiedades fueron calculadas en bloques de 100 ps sobre los 300 ns de simulación.

Adicionalmente, calculamos el potencial electrostático de las estructuras representativas usando el servidor PDB2PQR-APBS,⁷⁸ con el esquema de cargas de CHARMM27, a pH 7 y con una concentración de sal de 0.15 M. Para la comparación de las superficies de potencial electrostático, se utilizó el servidor webPIPSA,⁷⁹ el cual calcula índices de similitud entre pares de superficies electrostáticas, que van de valores de -1 (potenciales anticorrelacionados), pasan por 0 (no correlacionados) y llegan a +1 (potenciales idénticos). Los índices son luego convertidos a distancias mediante la fórmula $\sqrt{2-2SI}$, donde SI es el índice respectivo de similitud, tomando así valores de 0 (potenciales idénticos) a 2 (potenciales anticorrelacionados). Las distancias son finalmente representadas en un tipo de árbol filogenético (epograma) y con un mapa de calor, los cuales indican qué tan similares son las superficies.

Todas las propiedades calculadas se graficaron con KaleidaGraph 4.5, y las imágenes se generaron con el visualizador VMD 1.9.2.⁶⁹ Todas las representaciones 2D de las interacciones de TBP con los mejores compuestos fueron generadas con el software Maestro 11.09.010.⁸⁰

3.4 Acoplamiento molecular de compuestos sobre la superficie de las TBPs.

Para la selección de estructuras para el acoplamiento molecular de ensamble, utilizamos la estrategia que se muestra en la Figura 2. El objetivo de la estrategia es obtener un ensamble de estructuras de las MD que represente cambios a nivel de cadena principal y cadena lateral, que son importantes para la apertura/formación de cavidades transitorias.

1) Las estructuras de los centros de grupo, se enviaron al metaservidor Metapocket 2.0.⁸¹ Metapocket calcula cavidades con base en diferentes aproximaciones, usando los siguientes programas: LIGSITE, PASS, Q-Site Finder, SURFNET, Fpocket, GHECOM, ConCavity y POCASA. Las cavidades finales corresponden a los centros de masa de las cavidades agrupadas obtenidas por cada programa.

2) Tomando las cavidades predichas, las diferencias en secuencia y propiedades estructurales, seleccionamos los sitios de unión a NC2, TFIIA, la región interdominio y el sitio de la hélice H2.

3) Para estos sitios, seleccionamos los residuos alrededor de 4 Å, y las conformaciones de cadenas laterales con alta población durante las MD fueron identificadas. Seleccionamos las conformaciones de cadenas laterales con una

población mayor o igual al 10%; para los residuos de Lys, Arg, Glu, Gln y Met, sólo los ángulos diedros de cadena lateral χ_1 y χ_2 fueron utilizados, mientras que para el resto de los residuos usamos todas los ángulos diedros para clasificar rotámeros de acuerdo a Dunbrack.⁸² Para cada sitio, las estructuras seleccionados de las MD corresponden al 50% de todas las combinaciones muestreadas.

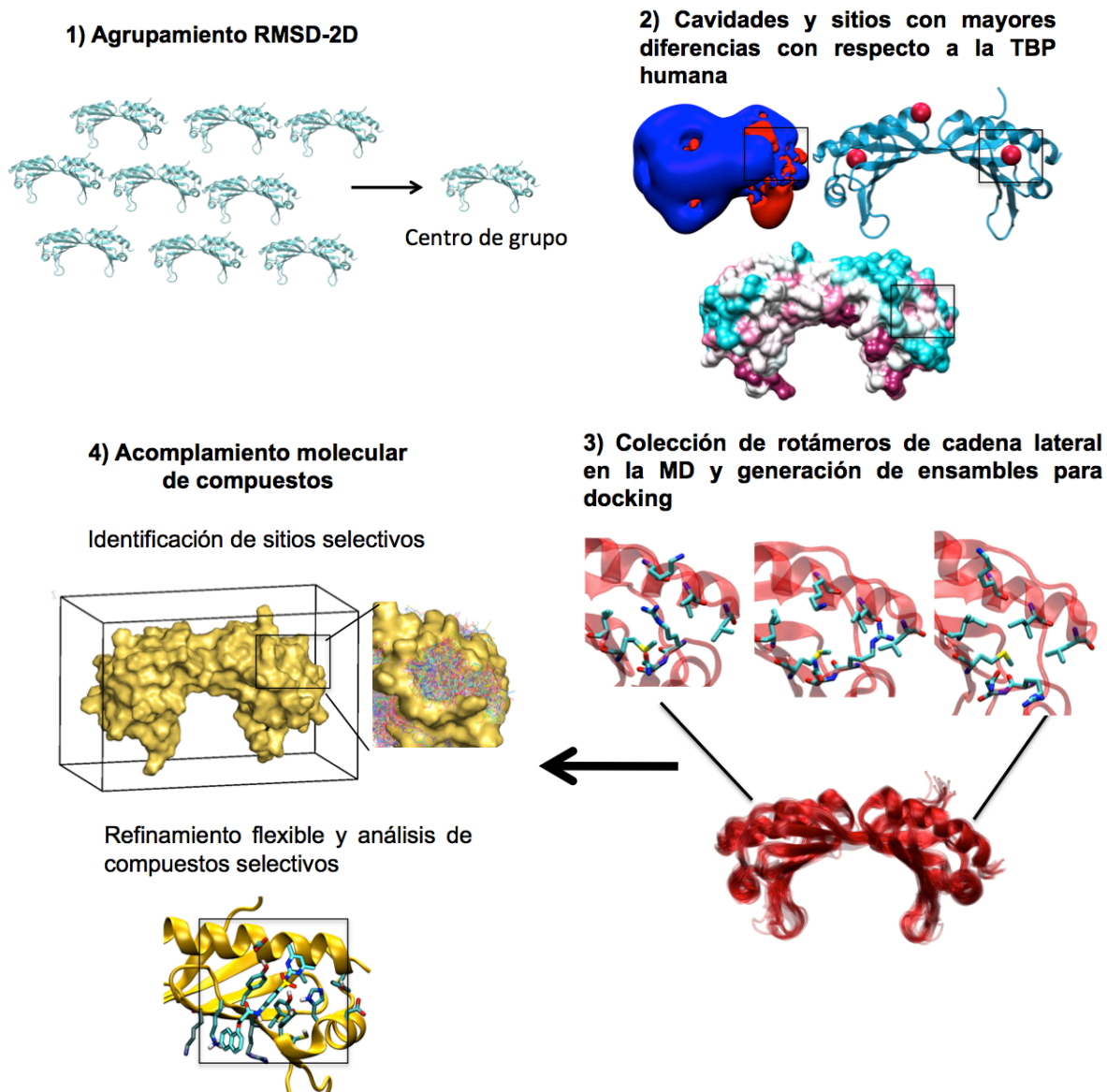


Figura 2. Diagrama del procedimiento computacional usado para seleccionar ensambles de estructuras de las MD con mayores diferencias respecto a la TBP de humano y llevar a cabo acomplamiento molecular. 1) Un agrupamiento por RMSD en 2D se llevó a cabo para seleccionar los centros de grupo. Estos centros de grupo fueron luego usados para predecir cavidades en la superficie de TBPs. 2) Las cavidades predichas, junto con las diferencias estructurales con respecto a **hsaTBP** (tales como el potencial electrostático, área accesible al solvente), fueron luego propuestas como potenciales sitios de unión a fármacos. 3) Para cada uno de estos sitios, los

residuos alrededor de 4 Å fueron seleccionados y las estructuras con las combinaciones de rotámeros de cadena lateral más frecuentes durante en las MD fueron usadas para crear el ensamble final para acoplamiento. 4) Una primera ronda de acoplamiento rígido fue llevada a cabo sobre toda la superficie del ensamble para cada TBP, para identificar los sitios selectivos. Una vez que los sitios con mejor selectividad con respecto a **hsa**TBP fueron identificados, una segunda ronda de acoplamiento flexible restringido en cada sitio se llevó a cabo, permitiendo a los residuos alrededor de 4 Å ser flexibles. Los compuestos más selectivos fueron luego analizados.

Para los ensambles de las cavidades de TFIIA y NC2, sólo las estructuras con un volumen mayor o igual a 50 Å³ calculado con MOLEGRO 6.0 fueron seleccionadas.⁸³ Los ensambles finales de estructuras para cada TBP se muestran en la Tabla 5 y fueron usados para el acoplamiento.

Tabla 5. Número de estructuras usadas para el acoplamiento de compuestos tipo fármaco.

TBP	Ensamble final
hsa	36
ecu	42
pnc	39
cal	48
tso	43
nam	40
ovo	45
ehi	40
pfa	48

4) La selección de compuestos tipo fármaco se hizo con base en una combinación de las reglas de Lipinski y Veber.^{84, 85} moléculas con peso molecular entre 160 y 500 g/mol; con coeficiente de partición octanol-agua (log P, que indica la lipofilidad del ligando) entre 0 y 5, para seleccionar aquellos que puedan penetrar al interior celular; con un número de enlaces rotables menor o igual a 7; con una área polar menor de 140 Å²; con un número máximo número de cinco grupos aceptores y diez donadores de enlace de hidrógeno respectivamente; que fueran neutros y que tuvieran funcionalidad benigna. Este filtrado se hizo sobre la base de datos de ZINC12,⁶⁰

usando <http://zinc.docking.org/search/properties>. Las estructuras 3D de 5887 compuestos fueron obtenidas en formato .mol2, y fueron posteriormente transformados a formato .pdbqt con el plugin de Vina en Pymol 1.7.7.2, para su uso en el acoplamiento molecular.

5) Para determinar la accesibilidad farmacológica de TBPs de parásitos, el acoplamiento molecular fue realizado con el programa *AutoDock Vina*,⁸⁶ usando un script hecho en casa. *AutoDock Vina* ha rendido buenos resultados en la literatura.^{87,88,89} Como primer paso, un acoplamiento rígido (flexible para el ligando) se llevó a cabo sobre cada ensamble de TBP. Todas las estructuras de las TBPs fueron transformadas a pdbqt con el plugin de Vina en Pymol 1.7.7.2. La caja usada para el acoplamiento cubrió toda la superficie de TBP, usando un espaciamiento de 1 Å, un valor exhaustividad de 20 y generando cinco poses para cada compuesto. De las mejores poses generadas (mejor puntaje), los compuestos con las más grandes diferencias respecto a la **hsa**TBP (**selectividad** = puntaje pose TBPparásito – puntaje TBPhumana), fueron listados (Tabla 6). Una selectividad menor o igual a -1.5 kcal/mol (que corresponde a una diferencia de 10 veces en Kds a 298 K) fue tomado como corte para una buena selectividad. Estos sitios, junto con aquellos que mostraron diferencias con respecto a **hsa**TBP, fueron seleccionados para una segunda ronda de acoplamiento flexible usando los mismos compuestos. La caja para el acoplamiento fue centrada en cada sitio, usando los mismos parámetros que en el acoplamiento rígido y los residuos equivalentes alrededor de 4 Å con respecto a **hsa**TBP fueron seleccionados como flexibles. Los compuestos con mejor selectividad fueron también listados (Tabla 7). Las Kds estimadas fueron obtenidas con la fórmula $Kd = e^{-\frac{\text{puntaje pose}}{RT}}$, donde T es la temperatura (298 K) y R es la constante de los gases (0.001985 kcal/K.mol).

4. Resultados

4.1 Diferencias en secuencia de las TBPs de parásitos comparadas a la TBP humana.

Las firmas principales del plegamiento de TBP pueden ser apreciadas en el alineamiento estructural (Figura 3a): la mayoría de las TBPs conservan los cuatro residuos de Phe necesarios para el doblamiento del DNA, con excepción de **pfa**TBP que tiene sólo tres (reemplazando una Phe por un residuo de isoleucina). Los sitios de interacción con TFIIA y TFIIIB están muy conservados, así como la Pro de la hélice H2'

Figura 3. Diferencias a nivel de secuencia en las TBPs de parásitos. a) Alineamiento estructural de las TBPs analizadas. Los residuos en azul y rojo corresponden a aquellos residuos 100% conservados y conservados más del 70% respectivamente; los asteriscos indican los residuos de unión a ADN en las estructuras cristalográficas reportadas; los elementos de estructura secundaria, residuos de Phe que participan en el doblamiento del ADN (F), cis-Pro (&) y la Pro conservada en hélice H2' (\$), los sitios de unión a TFIIA y TFIIIB, también se indican. b) Resultado de ConSurf mostrando las diferencias en superficie con respecto a **hsaTBP**: la imagen frontal y posterior se muestran en la figura (los sitios de unión a NC2 y a TFIIA son encerrados en un cuadro respectivamente), coloreados de variable (cian) a conservados (púrpura). c) Vista frontal y d) posterior de la comparación de secuencias por pares con respecto a **hsaTBP**: los residuos conservados y no conservados se muestran en rojo y azul respectivamente. La vista frontal se muestra como en b).

El sitio de unión a TFIIA está mejor conservado que su sitio simétrico – el sitio de unión a NC2 – el cual está rodeado de muchos residuos variables. El sitio de unión a TFIIIB está también conservado (estribo C-terminal, localizada entre las hebras S2'-S3') en todas las TBPs. La región interdominio (que es el sitio de unión a Brf1 del complejo de pre-inicio con la RNAPol III) y la hélice H2 son los sitios que mostraron mayores diferencias.

ConSurf muestra las diferencias respecto a un consenso de secuencias, así que para analizar estas diferencias para cada TBP contra **hsaTBP**, mapeamos los residuos conservados y no conservados contra ésta, usando el plugin multiseq de VMD (Figura 3c y Figura 3d).⁶⁹ Similar a lo visto con ConSurf, las principales diferencias yacen sobre la superficie convexa de la proteína, siendo **ehiTBP** y **pfaTBP** aquellas con los más grandes cambios en la región interdominio, la hélice H2 y sitio de unión a NC2. Con base en este análisis, se encontró que las principales diferencias respecto a **hsaTBP** están presentes en el sitio de unión a NC2, la hélice H2 y la región interdominio.

4.2 Análisis de las propiedades de superficie en las TBPs.

Las diferencias en secuencia en las TBPs pueden resultar en cambios estructurales que afecten su función y reconocimiento por otras proteínas. Para determinar cómo estas diferencias en secuencia influyen en cambios estructurales, calculamos una colección de parámetros estructurales sobre las trayectorias de las simulaciones de MD, y las comparamos con aquellas de **hsaTBP**.

Como un paso inicial analizamos el área accesible al solvente (AAS) no polar y polar para las diferentes TBPs (Figura 4). Esto es importante por las interacciones que puede establecer TBP con otras moléculas.

Con respecto al AAS no polar encontramos ligeras diferencias principalmente sobre la **pfa**TBP (Figura 4a), sugiriendo que algunos parches hidrofóbicos pueden ser importantes para el reconocimiento de otras moléculas. Las **tso**TBP, **ehi**TBP y **pfa**TBP muestran los mayores cambios en AAS no polar, mientras que el resto de TBPs mostraron similar AAS no polar con respecto a **hsa**TBP.

Las diferencias del AAS fueron mejor apreciadas cuando el AAS polar fue analizada: estas diferencias son igualmente mayores para las TBPs de **tso**, **ehi** y **pfa**, comparadas con **hsa**TBP (Figura 4b). Cuando los residuos polares, hidrofóbicos y cargados fueron mapeados en la superficie de las TBPs, las principales diferencias con respecto a **hsa**TBP están presentes en el sitio de unión a NC2, la región interdominio y la hélice H2 (Figura 4c).

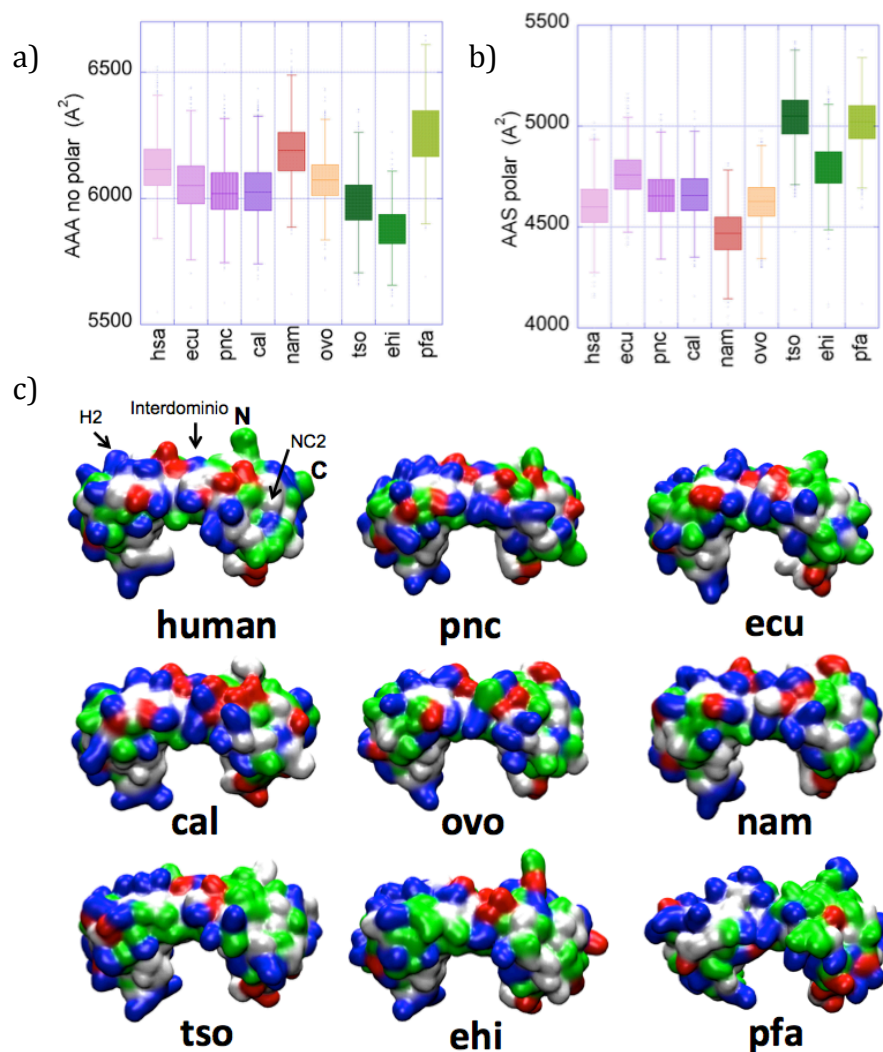


Figura 4. Polaridad de la superficie de TBPs de parásitos. a) AAS polar y b) no polar en las TBPs analizadas. Los valores de los límites y medianas se muestran en el diagrama de caja y bigotes. c) Representación de superficie de TBPs, donde se muestran los residuos positivos, negativos, polares

e hidrofóbicos en azul, rojo, verde y blanco, respectivamente. Los sitios con mayores diferencias se indican sobre **hsaTBP**.

Si los cambios en AAS relativa por residuo con respecto a **hsaTBP** son mapeados sobre la superficie de TBP (Ver Figura 5), encontramos que las diferencias también están presentes sobre el sitio de unión a NC2, interdominio y la hélice H2. Por otra parte también evaluamos las diferencias en fluctuaciones de carbonos alfa con respecto a **hsaTBP** (Anexo 3). Sin embargo a nivel de carbonos alfa solo encontramos ligeras diferencias en las asas de la estructura.

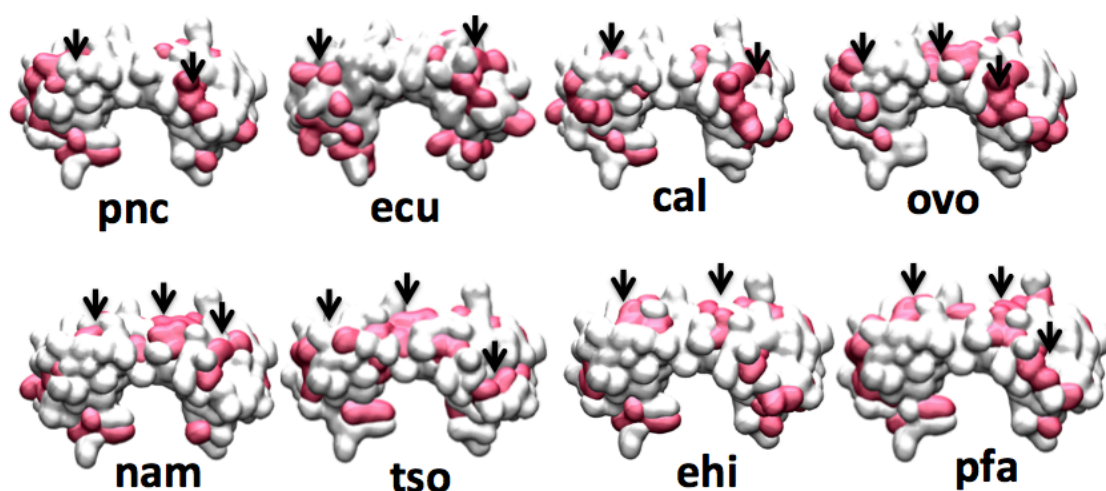


Figura 5. Diferencias en AAS relativa por residuo con respecto a **hsaTBP.** Diferencias en AAS relativa por residuo respecto a **hsaTBP**. Una representación de superficie se muestra para las TBPs. Los residuos con mayor AAS relativa se muestran en color rojo para cada TBP. Las flechas indican los residuos que concuerdan con los sitios de NC2, la hélice H2 y el interdominio.

Las diferencias en polaridad se apreciaron de mejor forma al calcular la energía electrostática de toda la proteína (Figura 6a): encontramos los cambios más notables en la TBP de **pfa**, la cual mostró la energía electrostática menos favorable; **ehi**TBP y **ecu**TBP mostraron valores intermedios, mientras que la **tso**TBP mostró la energía electrostática más favorable, cuando fueron comparadas con respecto a **hsaTBP**. **nam**TBP y **ovo**TBP mostraron similar energía electrostática, mientras sólo ligeros cambios fueron apreciados en **pnc**TBP, **ecu**TBP y **cal**TBP. Las diferencias en electrostática están relacionadas con la presencia de cargas en la superficie de la proteína.

Con respecto a esto, **pfa**TBP mostró un menor número de residuos negativos que **hsaTBP**, pero un mayor número de cargas en la hélice H2 (Figura 6b y 6d). Este exceso de cargas positivas presente en la superficie de **pfa**TBP puede estar causando

una repulsión en su superficie, promoviendo así una energía electrostática menos favorable. Curiosamente, el número total de residuos cargados fue similar entre todas las TBPs (Figura 6c), sugiriendo que lo relevante es que su distribución espacial es diferente.

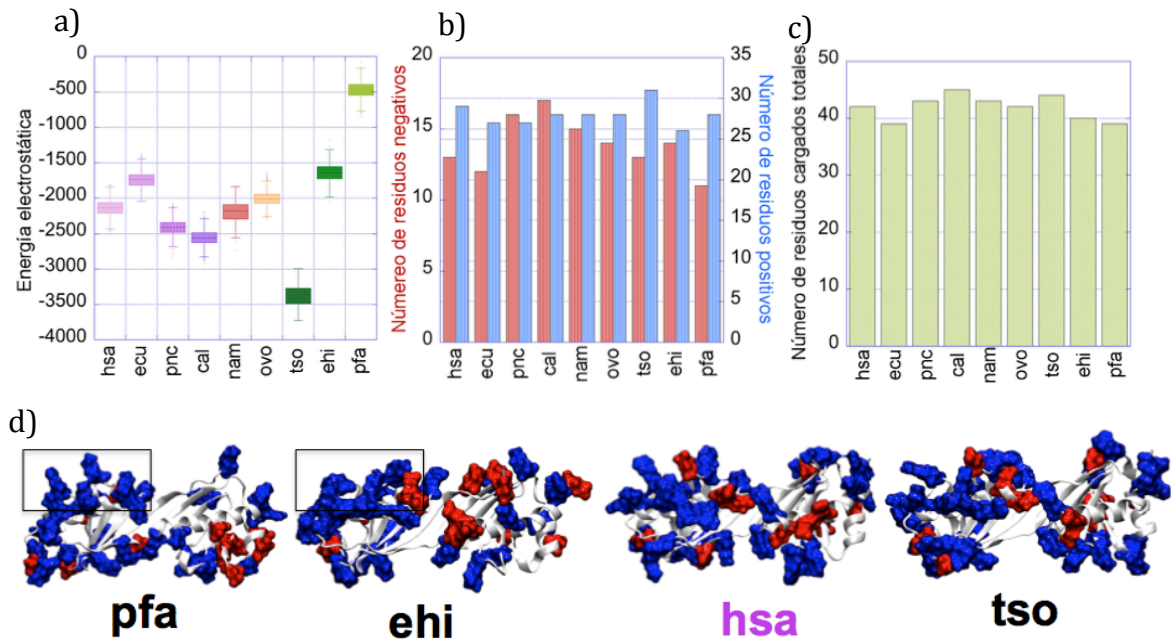


Figura 6. Electrostática en TBPs. a) Energía electrostática de las TBP. Los valores de los límites y medianas se muestran en el diagrama de caja y bigotes. b) Número de residuos positivos y negativos presentes en las TBPs analizadas. c) Número de residuos cargados totales en las TBPs. d) Para **pfa**TBP, **ehi**TBP, **hsa**TBP y **tso**TBP, se muestran los residuos positivos y negativos sobre la estructura en color azul y rojo respectivamente. El cuadro indica la posición de la hélice H2.

Para tener un mejor entendimiento de las consecuencias de estas diferencias en electrostática, calculamos el potencial electrostático sobre la superficie de las TBPs (Figura 7). Como previamente descrito,⁹⁰ las TBPs de eucariotas muestran su característico potencial electrostático positivo, indicativo de una proteína de unión a DNA (Figura 7a).

PfaTBP mostró un mayor potencial electrostático positivo en la superficie convexa de la proteína (debido principalmente a la hélice H2 cargada positivamente y parte del interdominio). Este mayor potencial electrostático positivo explica la energía electrostática desfavorable en **pfa**TBP.

OvoTBP, **ecu**TBP, **pnc**TBP y **nam**TBP mostraron ligeros cambios en su distribución electrostática, relacionado con una mayor conservación en secuencia con respecto a **hsa**TBP. Sin embargo, **cal**TBP, **ehi**TBP, **tso**TBP y **pfa**TBP mostraron marcadas diferencias, principalmente en los sitios de unión de NC2, la región interdominio, la hélice H2 y en el extremo N-terminal.

Para un análisis menos cualitativo de las diferencias a nivel de potencial electrostático, usamos el servidor WebPIPSA para compararlos con respecto al de **hsa**TBP (Figura 7b). Nuevamente **ecu**TBP, **ovo**TBP, **pnc**TBP y **nam**TBP mostraron superficies similares de potencial electrostático con respecto a **hsa**TBP, relacionado nuevamente con una mayor conservación en secuencia (las ramas del cladograma son más cortas para estas TBPs comparadas con **hsa**TBP). **cal**TBP, **tso**TBP, **ehi**TBP y **pfa**TBP mostraron las superficies de potencial electrostático más distintas a **hsa**TBP.

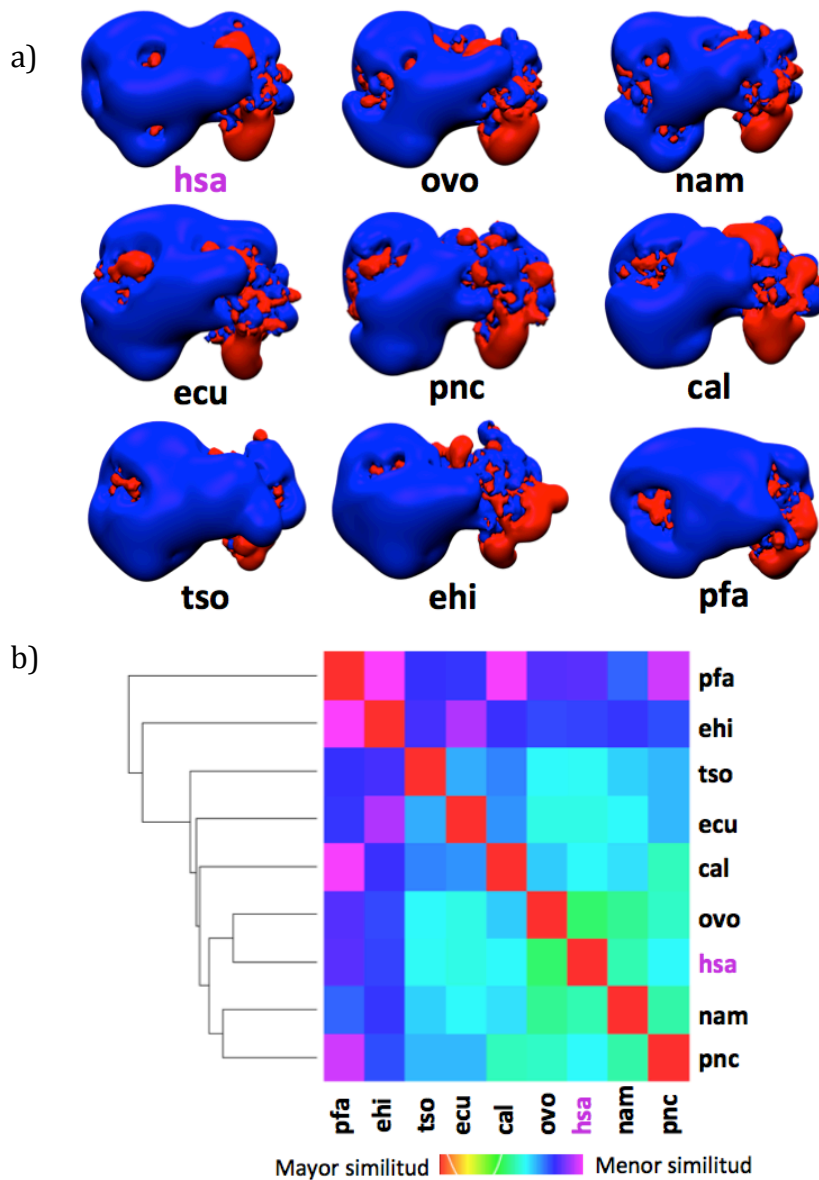


Figura 7. Superficie de potencial electrostático en TBPs. a) El potencial electrostático fue mapeado sobre las estructuras representativas de cada TBP. Los isovalores usados fueron -2 kT/e (en rojo) y $+2$ kT (en azul). b) Comparación de superficies de potencial electrostático con WebPIPSA. El código de colores está indicado en la figura.

En el caso de **tso**TBP, el potencial electrostático en la superficie no explicó su energía electrostática favorable, sin embargo, esta podría ser analizada con las interacciones formadas en la superficie de la proteína. Para esto, calculamos las interacciones de cadena lateral entre residuos polares (Ser, Thr, Cys, Asn, Gln, His, Tyr), entre residuos cargados (Asp, Glu con Lys, Arg, interacciones conocidas como puentes salinos), y entre residuos polares-cargados que incluyen la combinación de las dos anteriores. Con respecto a los puentes polares y polares-cargados, no encontramos diferencias entre las TBPs de parásitos comparadas con la TBP de humano (Figura 8), sin embargo, encontramos diferencias en los puentes salinos de cadena lateral (Figura 9a).

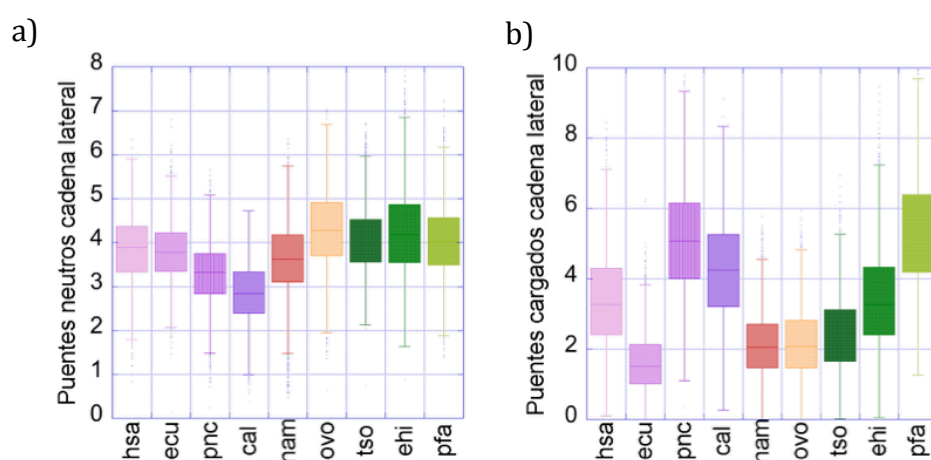


Figura 8. Interacciones polares en las TBPs. Interacciones polares en las TBPs de parásitos. a) Puentes de hidrógeno de las cadenas laterales de residuos Ser, Thr, Cys, Asn, Gln, His y Tyr. b) Puentes de hidrógeno entre las cadenas laterales de residuos Ser, Thr, Cys, Asn, Gln, His y Tyr y los residuos cargados, Asp, Glu, Lys y Arg. Los valores de los límites y medianas se muestran en el diagrama de caja y bigotes.

Efectivamente encontramos que **tso**TBP mostró el mayor número de estas interacciones, cuando fue comparado con el resto de TBPs, mientras que **pfa**TBP mostró el menor número de ellas. Esto correlaciona con las propiedades electrostáticas de estas TBPs, siendo la **tso**TBP aquella con mejor electrostática. Cuando sólo los puentes salinos más estables (aquellos con una ocupación mayor al 75% del tiempo simulación) se grafican, las diferencias entre **tso**TBP, **pfa**TBP y **hsa**TBP se aprecian mejor (Figura 9b). El número de puentes salinos fue similar para el resto de las TBPs cuando fueron comparadas contra **hsa**TBP, lo cual correlaciona con su energía electrostática similar.

Con excepción de **ehi**TBP y **pfa**TBP, todas las TBPs muestran un puente salino muy estable en la región interdominio (Figura 9c). En el caso de **nam**TBP y **ovo**TBP, esta

interacción se mantuvo presente solo el 50% del tiempo de simulación (Figura 9d). Estas TBPs tienen un residuo de Lys al inicio de la H2', en lugar de un residuo de Arg como en **tsoTBP**, **pncTBP**, **calTBP** y **hsaTBP**. El residuo de Lys promueve un puente salino menos estable, como es apreciado con un bajo tiempo de ocupación para **namTBP** y **ovoTBP**, mientras que la presencia de una Arg promueve una interacción más estable.

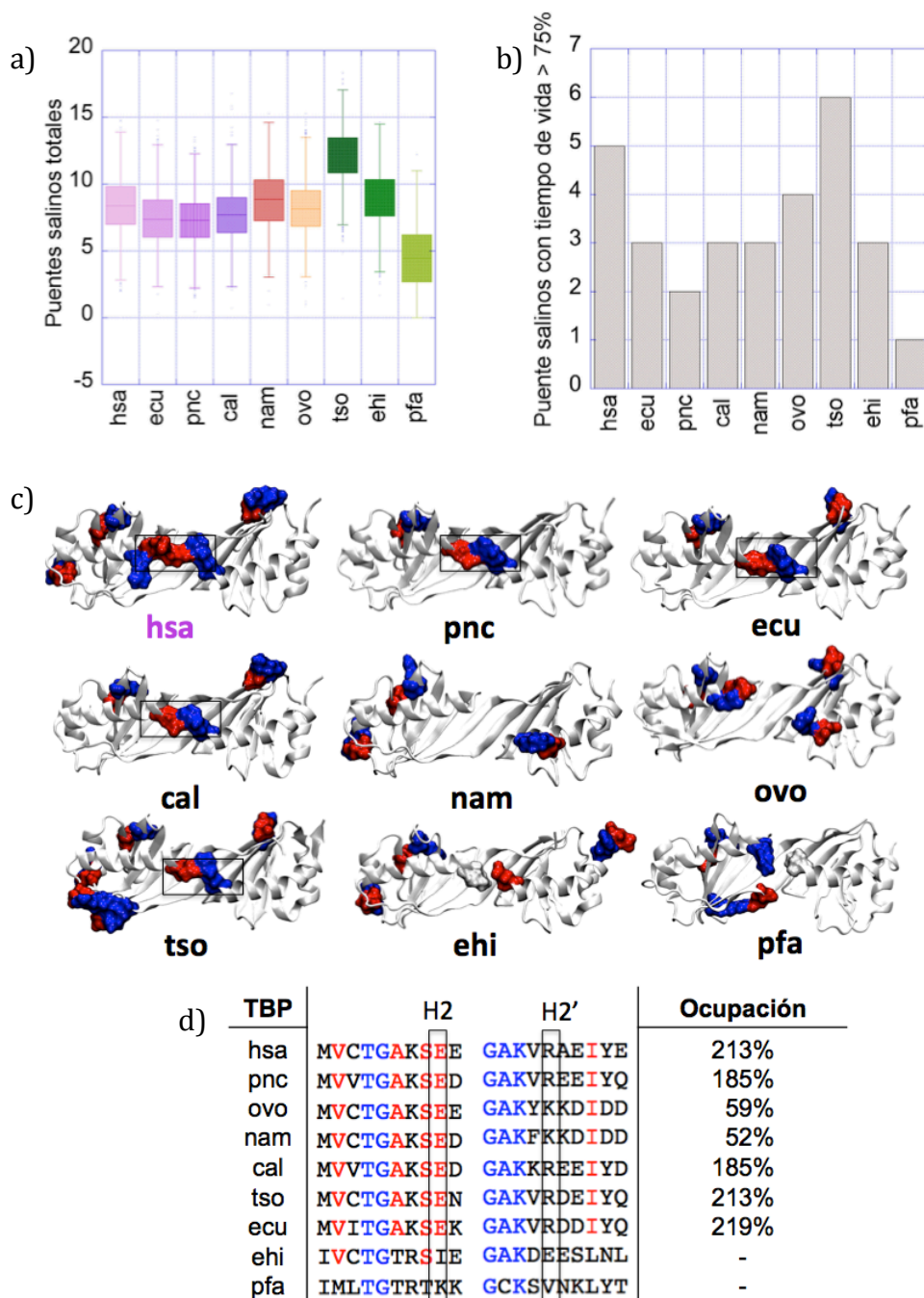


Figura 9. Puentes salinos de cadena lateral en las TBPs. a) Puentes salinos totales y b) puentes salinos presentes >75% del tiempo de simulación. Los valores de los límites y medianas se muestran en el diagrama de caja y bigotes. c) Puentes salinos presentes el 75% de simulación

mapeados sobre la superficie de TBPs. Los residuos negativos y positivos se muestran en esferas de VDW de color rojo y azul respectivamente. El cuadro indica el puente salino en el interdominio. d) Conservación en secuencia de residuos involucrados en el puente salino en la región interdominio (encerrados en un cuadro) y ocupación durante la MD, sumada sobre todos los átomos involucrados en un puente salino por cada par de residuos. Los residuos conservados y parcialmente conservados se muestran en azul y rojo respectivamente.

En el caso de **ehi**TBP y **pfa**TBP, el par de residuos cargados no está conservado. Sin embargo, los residuos en posiciones equivalentes hacen ~5 contactos entre carbonos en promedio para estabilizar este sitio. Esto hace al interdominio de **nam**TBP, **ovo**TBP, **ehi**TBP y **pfa**TBP un sitio atractivo para la interacción con compuestos tipo fármaco.

En resumen, los parámetros de superficie presentados en esta sección indican que los sitios con mayores diferencias a la TBP de humano son el sitio de unión a NC2, la región interdominio y la hélice H2. Estos sitios parecen ser interesantes para la propuesta de potenciales sitios de unión a fármacos. La ventaja de proponer estos sitios es que se conocen las proteínas que reconocen cada sitio, así el mecanismo de inhibición por el fármaco estaría basado en la pérdida de interacción, lo cual puede ser analizado experimentalmente.

4.3 Presencia de cavidades sobre las TBPs.

Una característica de un sitio de unión a un compuesto es que, en general, éste corresponde a una cavidad en la superficie de la proteína donde un compuesto tipo fármaco puede interactuar.^{15, 16, 58} Para evaluar la presencia de cavidades en la superficie de las TBPs, enviamos las estructuras representativas de cada TBP al servidor Metapocket (Figura 10).

Encontramos que algunas de las cavidades concuerdan con los sitios en TBPs que mostraron diferencias al nivel de secuencia y estructura: el sitio de unión a NC2, la región interdominio y la hélice H2. El sitio de unión a TFIIA fue predicho como cavidad en todas las TBPs (y es un sitio conservado en éstas). Sin embargo, otros sitios no fueron predichos en algunas TBPs. Por ejemplo, el sitio de unión a NC2 no fue predicho en **ecu**TBP, **ehi**TBP, **cal**TBP ni **tso**TBP. El sitio del interdominio y la hélice H2 tampoco fueron predichos para la mayoría de TBPs. Esto puede deberse a que ambos sitios son planos o que las estructuras representativas tengan tales cavidades cerradas. Sin embargo, la predicción por consenso fue útil para la identificación de estos sitios. Esto sugiere que es importante tomar en cuenta la flexibilidad de cadena lateral y principal muestreadas en las simulaciones de MD, ya que esto puede promover la apertura de cavidades.

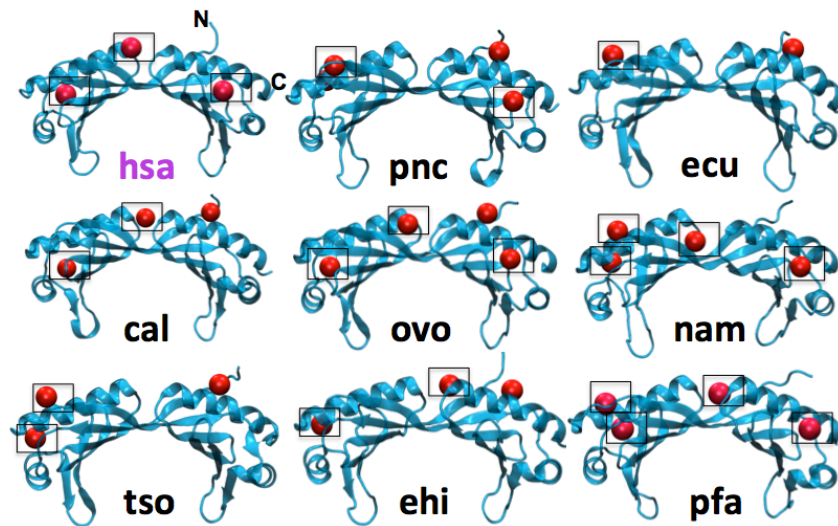


Figura 10. Cavidades calculadas con Metapocket sobre las TBPs. Cavidades calculadas con Metapocket sobre las TBPs. Las cavidades son representadas con esferas rojas. Las cavidades encerradas en un cuadro corresponden a los sitios de unión a NC2 y TFIIA, la región interdominio y la hélice H2.

Para determinar la dinámica de apertura seleccionamos estas cavidades y calculamos los volúmenes con el programa POVME 3.0 (Figura 11 y Figura 12). En el caso del sitio de unión a TFIIA, éstos fueron más grandes en **tso**TBP y **pfa**TBP comparados con los de **hsa**TBP, mientras que el resto de las TBPs mostraron una apertura similar o menor apertura, siendo el sitio de unión a TFIIA de **ehi**TBP el sitio menos abierto durante las trayectorias de MD. Aparte del sitio de unión a DNA, el sitio de unión a TFIIA es la cavidad más grande presente sobre la superficie de TBP, lo cual puede estar relacionado con sus interactores, tales como MOT1, TAF1, Brf1 (del complejo con RNApolIII), TFIIA y el dominio de activación GAL4. Aunque este sitio está altamente conservado de acuerdo al análisis de secuencia, la mayor apertura del sitio en las TBPs de parásitos podría hacerlo un sitio atractivo.

Con respecto al sitio de unión a NC2, los volúmenes fueron más pequeños en comparación con los del sitio de unión a TFIIA, a pesar de estar relacionado por simetría con este último; interesantemente todas las TBPs de parásitos mostraron una mayor apertura de la cavidad en comparación con **hsa**TBP (Figura 11b y 12b).

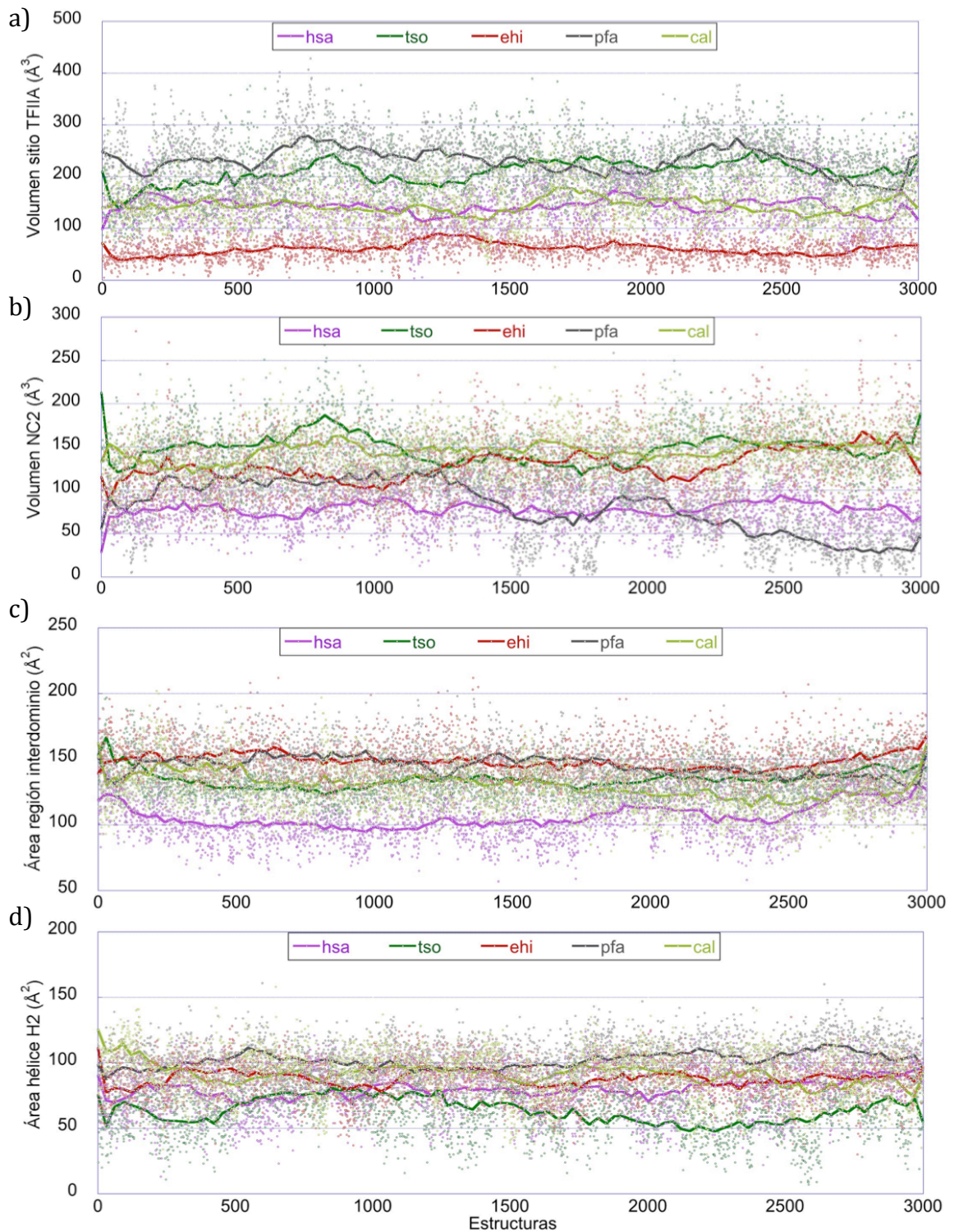


Figura 11. Dinámica de apertura en las TBP calculada con POVME 3.0. De a) a d) se muestra la dinámica de apertura de los sitios de unión a TFIIA y NC2, así como para la región interdominio y la hélice H2. En el caso de estos dos últimos, se muestra el área superficial debido a que estos sitios corresponden a superficies planas. El ajuste suavizado se muestra como líneas en cada gráfico.

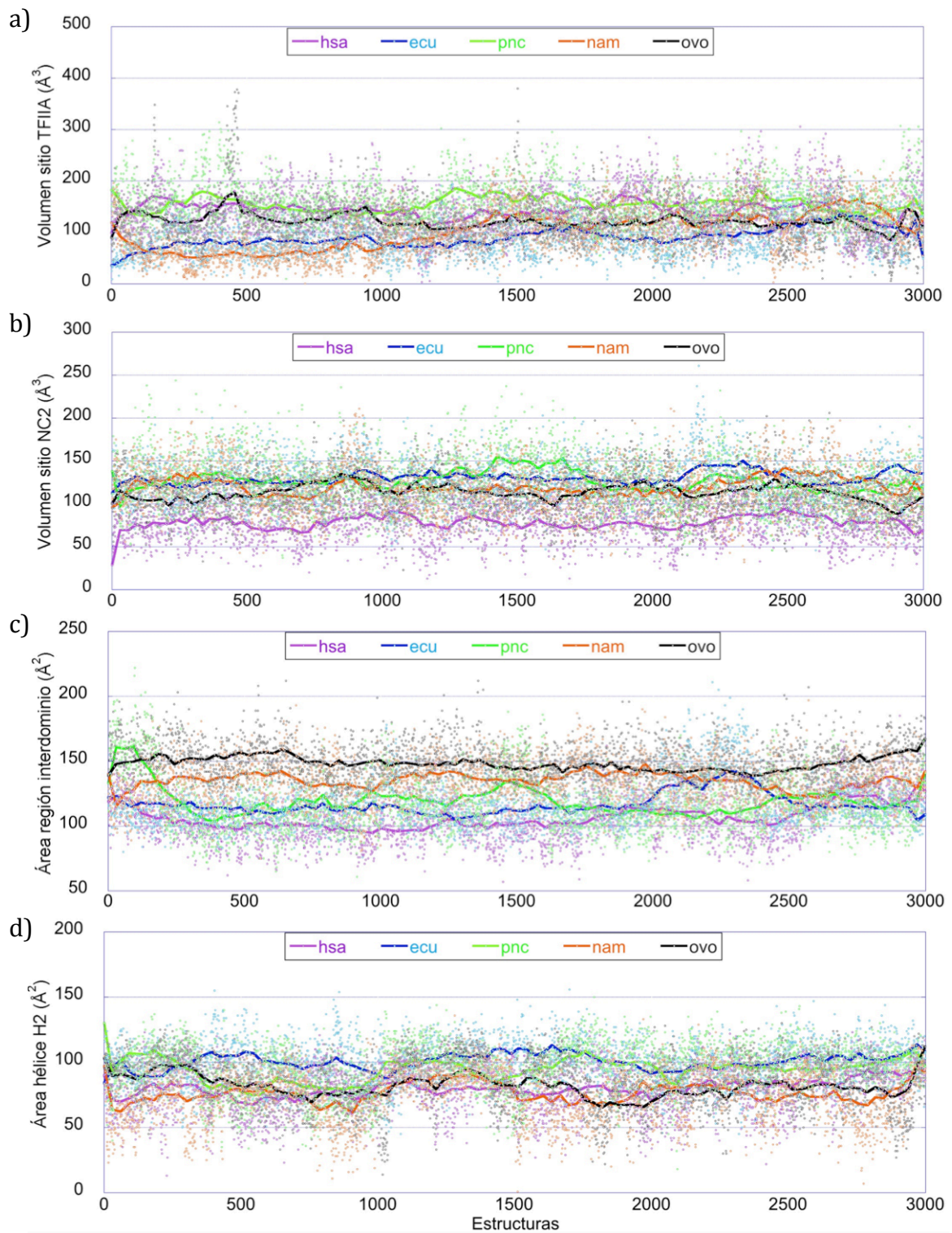


Figura 12. Dinámica de apertura en las TBPs calculadas con POVME 3.0. De a) a d) se muestra la dinámica de apertura de los sitios de unión a TFIIA y NC2, así como para la región interdominio y la hélice H2. En el caso de estos dos últimos, se muestra el área superficial debido a que estos sitios corresponden a superficies planas. El ajuste suavizado se muestra como líneas en cada gráfico.

La interacción con NC2 es sólo a través de un residuo de Phe y uno de Gln que se unen a esta cavidad en TBP.⁹¹ De acuerdo a nuestro análisis, el sitio de unión a NC2 está menos conservado con respecto al de TFIIA y mostró diferencias estructurales. El factor de transcripción NC2 es un regulador negativo de la transcripción, sugiriendo que su sitio de unión en TBP es un sitio atractivo para evaluar su accesibilidad farmacológica.

Por otra parte, la región interdominio y la hélice H2 son sitios más expuestos al solvente, a diferencia de los sitios de unión a TFIIA y NC2; por tanto, computamos el área superficial para estos sitios (Figura 11c-d y 12c-d). La región interdominio tuvo una área superficial más grande en todas las TBPs cuando fue comparada con **hsa**TBP. Sin embargo, para la hélice H2 todas las TBPs mostraron un área similar, con excepción para **tso**TBP y **pfa**TBP. Ambos sitios son plataformas de unión para el sitio de unión a Brf1 y TFIIA, y junto con las diferencias, haciéndolos buenos candidatos como sitios blanco.

Tomando estos cuatro sitios construimos ensamblajes de estructuras para cada TBP, basados en las combinaciones de rotámeros de cadenas laterales más frecuentes que se muestrearon durante las simulaciones de MD (Tabla 5). El ensamblaje tiene información de cambios estructurales en cadena principal y lateral. Para los sitios de TFIIA y NC2, los ensamblajes incluyen estructuras con estas cavidades abiertas. Estos ensamblajes fueron usados para evaluar su potencial de unión a compuestos tipo fármaco.

4.4 Evaluación computacional de la unión de compuestos tipo fármaco en las TBPs.

Como paso inicial, llevamos a cabo un acoplamiento molecular de ensamblaje rígido con Autodock Vina⁸⁶ sobre la superficie entera del ensamblaje de estructuras de cada TBP, usando 5887 compuestos tipo fármaco, obtenidos de la base de datos de ZINC12.⁶⁰ Esta estrategia fue usada para identificar los sitios con mayor potencial de unión a compuestos tipo fármaco en las TBPs de parásitos, al compararlos contra los de **hsa**TBP.

La Tabla 6 muestra los compuestos con mayor selectividad con respecto a **hsa**TBP para cada TBP. Encontramos buena selectividad (valores ≤ -1.5 kcal/mol) para **nam**TBP, **tso**TBP, **cal**TBP y **pfa**TBP, mientras que encontramos reducida selectividad para **ecu**TBP, **ovo**TBP y **ehi**TBP. Esto indica una mejor accesibilidad farmacológica para **nam**TBP, **tso**TBP, **cal**TBP y **pfa**TBP.

Todas las TBPs de parásitos muestran valores de afinidad (puntaje de Autodock Vina) que están en el orden micromolar-nanomolar, de acuerdo con sus valores estimados de Kds a 298. Interesantemente, los compuestos en **pfaTBP** y **calTBP** tuvieron la mejor selectividad con respecto a **hsaTBP**, con diferencias de -2.0 kcal/mol y -1.8 Kcal/mol respectivamente. **pncTBP** y **ehiTBP** fueron las TBPs del conjunto que mostraron los valores de selectividad más bajos (alrededor de ~-0.9 Kcal/mol), a pesar de que nuestro análisis estructural mostró que **ehiTBP** es diferente de **hsaTBP**.

Tabla 6. Mejores cinco compuestos con selectividad en las TBPs de parásitos con respecto a hsaTBP. La selectividad corresponde a la diferencia del valor de energía entre la mejor pose en la TBP de parásitos y aquella en **hsaTBP** para el mismo compuesto. El código del compuesto corresponde al de la base de datos de ZINC12.

Acoplamiento de ensamble rígido			
Compuesto	Selectividad	Energía (Kcal/mol)	Kd estimada 298 K para la pose en la TBPs de parásitos
ZINC00386658	-0.9	ehi -7.9/ hsa -7.0	1.5x10 ⁻⁶
ZINC04656177	-1.0	ehi -7.2/ hsa -6.3	5.1x10 ⁻⁶
ZINC08439353	-1.0	ehi -8.1/ hsa -7.1	1.1x10 ⁻⁶
ZINC00678144	-1.0	ehi -8.3/ hsa -7.3	8.0x10 ⁻⁷
ZINC00687495	-1.0	ehi -8.1/ hsa -7.1	1.1x10 ⁻⁶
ZINC00633806	-1.1	ecu -9.0/ hsa -7.9	2.4x10 ⁻⁷
ZINC00655012	-1.2	ecu -8.3/ hsa -7.1	8.0x10 ⁻⁷
ZINC02124174	-1.2	ecu -8.4/ hsa -7.2	6.8x10 ⁻⁷
ZINC08400467	-1.2	ecu -8.5/ hsa -7.3	5.7x10 ⁻⁷
ZINC02124172	-1.4	ecu -8.4/ hsa -6.9	6.8x10 ⁻⁷
ZINC08384212	-1.3	tso -8.8/ hsa -7.5	3.4x10 ⁻⁷
ZINC08439353	-1.3	tso -8.4/ hsa -7.1	6.8x10 ⁻⁷
ZINC00623808	-1.4	tso -8.9/ hsa -7.5	2.9x10 ⁻⁷
ZINC02077897	-1.4	tso -9.2/ hsa -7.8	1.7x10 ⁻⁷
ZINC00707055	-1.5	tso -8.6/ hsa -7.1	4.8x10 ⁻⁷
ZINC00654142	-1.5	cal -8.1/ hsa -6.6	1.1x10 ⁻⁶
ZINC00654525	-1.5	cal -8.4/ hsa -6.9	6.8x10 ⁻⁷
ZINC08400463	-1.5	cal -9.3/ hsa -7.8	1.4x10 ⁻⁷

ZINC00681612	-1.8	cal -8.3/ hsa -6.6	8.0×10^{-7}
ZINC00681612	-1.8	cal -8.4/ hsa -6.6	6.8×10^{-7}
ZINC01020433	-0.9	pnc -8.4/ hsa -7.5	6.8×10^{-7}
ZINC08396924	-0.9	pnc -8.9/ hsa -8.0	2.9×10^{-7}
ZINC08436921	-0.9	pnc -7.8/ hsa -6.9	1.8×10^{-6}
ZINC00627455	-1.0	pnc -8.9/ hsa -7.9	2.9×10^{-7}
ZINC00684897	-1.2	pnc -8.6/ hsa -7.4	4.8×10^{-7}
ZINC08440929	-1.1	ovo -8.8/ hsa -7.7	3.4×10^{-7}
ZINC02876856	-1.1	ovo -8.6/ hsa -7.5	4.8×10^{-7}
ZINC08386942	-1.1	ovo -9.6/ hsa -8.5	8.9×10^{-8}
ZINC06446067	-1.2	ovo -9.2/ hsa -8.0	1.7×10^{-7}
ZINC00703977	-1.3	ovo -9.2/ hsa -7.9	1.7×10^{-7}
ZINC00663824	-1.3	nam -7.8/ hsa -6.5	1.8×10^{-6}
ZINC00704367	-1.4	nam -8.9/ hsa -7.5	2.9×10^{-7}
ZINC04363229	-1.4	nam -9.2/ hsa -7.8	1.7×10^{-7}
ZINC04176965	-1.5	nam -8.8/ hsa -7.3	3.4×10^{-7}
ZINC19796971	-1.6	nam -8.0/ hsa -6.4	1.3×10^{-6}
ZINC00654241	-1.7	pfa -8.1/ hsa -6.4	1.1×10^{-6}
ZINC08439353	-1.8	pfa -8.9/ hsa -7.1	2.9×10^{-7}
ZINC00624433	-1.8	pfa -10.3/ hsa -8.5	2.6×10^{-8}
ZINC00972546	-1.8	pfa -10.1/ hsa -8.3	3.8×10^{-8}
ZINC00655804	-2.0	pfa -9.4/ hsa -7.4	1.2×10^{-7}

Si todo el conjunto de compuestos con mejor selectividad mostrado en la Tabla 6 es mapeado sobre la superficie de cada TBP, encontramos que los principales sitios de unión corresponden a los sitios de unión a TFIIA y NC2, la región interdominio, el estribo N-terminal (que es el asa S3-S4) y el extremo N-terminal. Estos sitios corresponden a los sitios con mayor potencial de unión a compuestos tipo fármaco en estas TBPs (Figura 13).

Como se muestra en la Figura 13, encontramos otros dos sitios potenciales además de los identificados con el análisis estructural. Estos sitios corresponden al extremo N-terminal en **cal**TBP y el estribo N-terminal (asa S3-S4) en **pfa**TBP, y estos mostraron buena selectividad con respecto a **hsa**TBP. En el caso de la región interdominio, encontramos solo compuestos sobre la **nam**TBP. Sin embargo, este compuesto mostró baja selectividad (-1.3 kcal/mol) con respecto a **hsa**TBP.

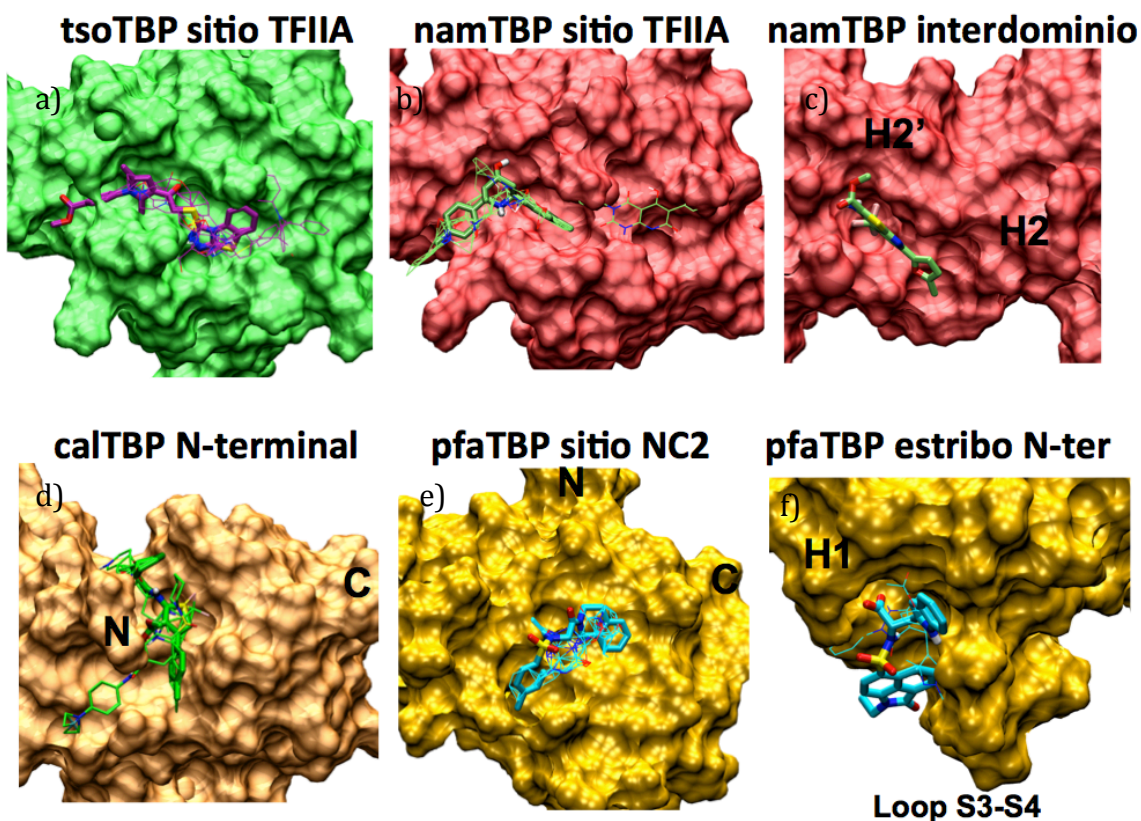


Figura 13. Sitios potenciales de unión a compuestos sobre las TBPs de parásitos. Los compuestos más selectivos mostrados en la Tabla 3 fueron mapeados sobre las superficies de a) **tsoTBP**, b) **namTBP**, c) **calTBP** y c) **pfaTBP**. El compuesto con la más alta selectividad se muestra con varillas, mientras que el resto de los compuestos se muestran como líneas. Para el caso del interdominio de **namTBP**, sólo se encontró un compuesto, aunque con baja selectividad (-1.3 Kcal/mol).

Dado que la consideración de la flexibilidad en la función de puntaje es importante, una segunda ronda de acoplamiento flexible sobre las estructuras representativas fue llevado a cabo para cada uno de los sitios identificados en el acoplamiento rígido: para **pfaTBP** éste fue hecho sobre el sitio de unión a NC2 y el estribo N-terminal, para **calTBP** sobre el extremo N-terminal y para **namTBP** y **tsoTBP** sobre el sitio de unión a TFIIA. También incluimos el sitio interdominio de **namTBP** y **ehiTBP** dado que encontramos diferencias estructurales y dinámicas con respecto a **hsaTBP**, a pesar de su baja selectividad. La Tabla 7 muestra los resultados del acoplamiento flexible sobre las estructuras representativas para estas TBPs.

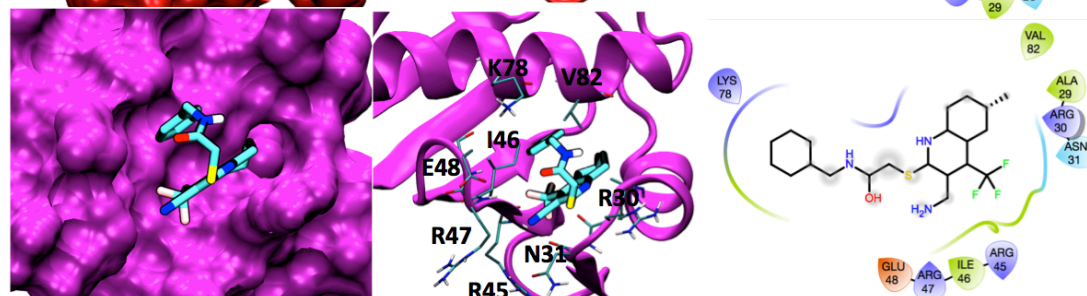
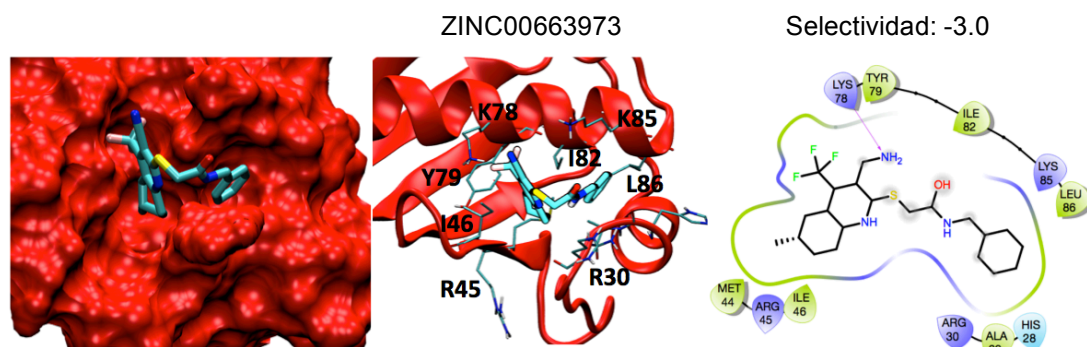
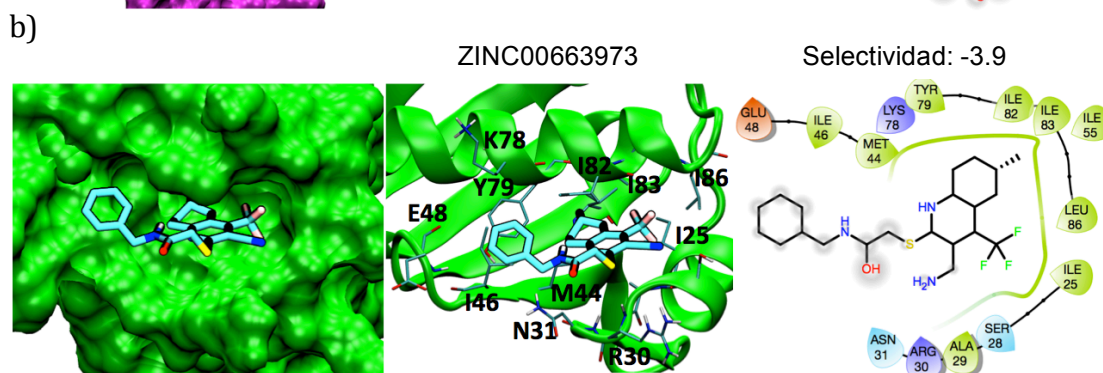
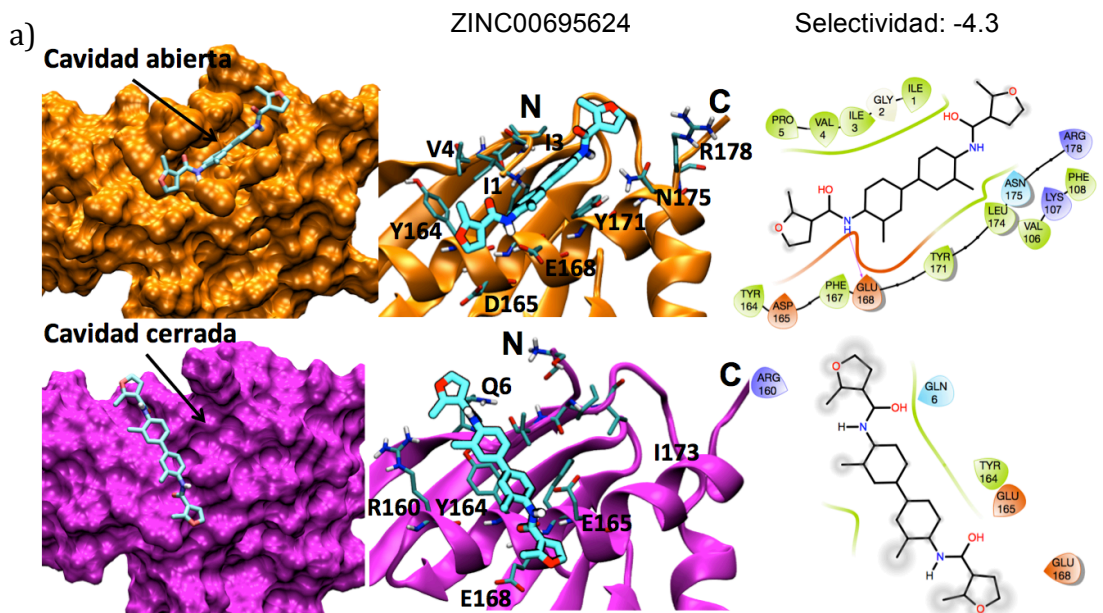
Tabla 7. Compuestos con mejor selectividad con respecto a **hsaTBP**, obtenidos del acoplamiento flexible sobre los sitios seleccionados para cada TBP. La selectividad corresponde a la diferencia del valor de energía entre la mejor pose en la TBP de parásitos y aquella

en **hsa**TBP para el mismo compuesto. El código del compuesto corresponde al de la base de datos de ZINC12.

Acoplamiento flexible sobre las estructuras representativas			
Compuesto para cada sitio	Selectividad	Energía (Kcal/mol)	Kd estimada 298 K para la pose en la TBP de parásitos
Extremo N-ter			
ZINC00695624	-4.3	cal -11.1/ hsa -6.8	7.0x10 ⁻⁹
ZINC02497152	-3.9	cal -10.4/ hsa -6.5	2.3x10 ⁻⁸
ZINC08385471	-3.9	cal -10.5/ hsa -6.5	1.9x10 ⁻⁸
ZINC01427920	-3.7	cal -10.4/ hsa -6.7	2.3x10 ⁻⁸
ZINC04656082	-3.7	cal -10.4/ hsa -6.7	2.3x10 ⁻⁸
ZINC00478694	-3.5	cal -11.2/ hsa -7.7	5.9x10 ⁻⁹
ZINC00677022	-3.4	cal -11.3/ hsa -7.9	5.0x10 ⁻⁹
Sitio TFIIA			
ZINC01758808	-3.5	nam -9.9/ hsa -6.4	5.3x10 ⁻⁸
ZINC00663973	-3.0	nam -8.7/ hsa -5.7	4.7x10 ⁻⁷
ZINC04656123	-2.7	nam -9.1/ hsa -6.4	2.0x10 ⁻⁷
ZINC00663975	-2.6	nam -8.5/ hsa -5.9	5.7x10 ⁻⁷
ZINC09365179	-2.5	nam -9.4/ hsa -6.9	1.2x10 ⁻⁷
ZINC00983556	-2.3	nam -9.2/ hsa -6.9	1.7x10 ⁻⁷
ZINC08442128	-2.1	nam -10.2/ hsa -8.1	3.2x10 ⁻⁸
Interdominio			
ZINC00644211	-3.7	nam -9.1/ hsa -5.4	2.0x10 ⁻⁷
ZINC02252328	-3.6	nam -9.9/ hsa -6.3	5.3x10 ⁻⁸
ZINC00884274	-3.3	nam -10.2/ hsa -6.9	3.2x10 ⁻⁸
ZINC08431489	-3.0	nam -10.1/ hsa -7.1	3.8x10 ⁻⁸
ZINC00883987	-2.9	nam -9.8/ hsa -6.9	6.3x10 ⁻⁸
ZINC00676209	-2.8	nam -9.9/ hsa -7.1	5.3x10 ⁻⁸
ZINC00281924	-2.8	nam -9.8/ hsa -7.0	6.3x10 ⁻⁸
Interdomain			
ZINC00663973	-4.0	ehi -9.3/ hsa -5.3	1.4x10 ⁻⁷
ZINC00663970	-3.7	ehi -8.9/ hsa -5.2	2.9x10 ⁻⁷
ZINC00712044	-3.6	ehi -8.6/ hsa -6.9	4.8x10 ⁻⁷

ZINC00703575	-3.3	ehi -9.8/ hsa -6.5	6.3×10^{-8}
ZINC02077887	-3.3	ehi -9.5/ hsa -6.2	1.0×10^{-7}
ZINC08399657	-3.2	ehi -9.4/ hsa -6.2	1.2×10^{-7}
ZINC08442127	-3.2	ehi -8.9/ hsa -5.7	2.9×10^{-7}
Sitio TFIIA			
ZINC01758808	-4.2	tso -10.6/ hsa -6.4	1.6×10^{-8}
ZINC04656123	-4.1	tso -10.5/ hsa -6.4	1.9×10^{-8}
ZINC00971588	-3.9	tso -9.6/ hsa -7.3	8.9×10^{-8}
ZINC00663973	-3.9	tso -9.6/ hsa -5.7	8.9×10^{-8}
ZINC08439353	-3.8	tso -10.5/ hsa -6.7	1.9×10^{-8}
ZINC01000251	-3.7	tso -10.1/ hsa -6.4	3.8×10^{-8}
ZINC01000295	-3.6	tso -10.0/ hsa -6.4	4.5×10^{-8}
Estribo Nter			
ZINC00994448	-2.5	pfa -11.1/ hsa -8.5	7.0×10^{-9}
ZINC02158339	-2.4	pfa -8.1/ hsa -5.7	1.1×10^{-6}
ZINC06270551	-2.2	pfa -8.4/ hsa -6.2	6.8×10^{-7}
ZINC00996029	-2.2	pfa -8.8/ hsa -6.6	3.4×10^{-7}
ZINC01953156	-2.2	pfa -8.0/ hsa -5.8	1.3×10^{-6}
ZINC19796844	-2.2	pfa -8.1/ hsa -5.9	1.1×10^{-6}
ZINC00668078	-2.2	pfa -8.6/ hsa -6.4	4.8×10^{-7}
Sitio NC2			
ZINC08442277	-3.0	pfa -10.3/ hsa -7.3	2.7×10^{-8}
ZINC15989997	-2.9	pfa -8.8/ hsa -5.9	3.4×10^{-7}
ZINC00654357	-2.9	pfa -9.3/ hsa -6.4	1.4×10^{-7}
ZINC00654096	-2.9	pfa -9.5/ hsa -6.6	1.0×10^{-7}
ZINC00653348	-2.9	pfa -9.5/ hsa -6.6	1.0×10^{-7}
ZINC00654304	-2.8	pfa -9.3/ hsa -6.5	1.4×10^{-7}
ZINC00654006	-2.7	pfa -10.2/ hsa -7.5	3.2×10^{-8}

En todos los casos, el refinamiento flexible mejoró los valores de selectividad y afinidad, yendo de micromolar a nanomolar de acuerdo con los valores de Kd estimados. Las interacciones para los compuestos con mejor selectividad se muestran en la Figura 14 y son comparados con sus equivalentes en **hsa**TBP. Para todos los casos, se encontró un mejor ajuste de los compuestos en el sitio de las TBPs de parásitos, llevando a una mejor afinidad. Esto está principalmente relacionado a compuestos unidos a sitios más profundos, mientras que éstos fueron más expuestos en **hsa**TBP.



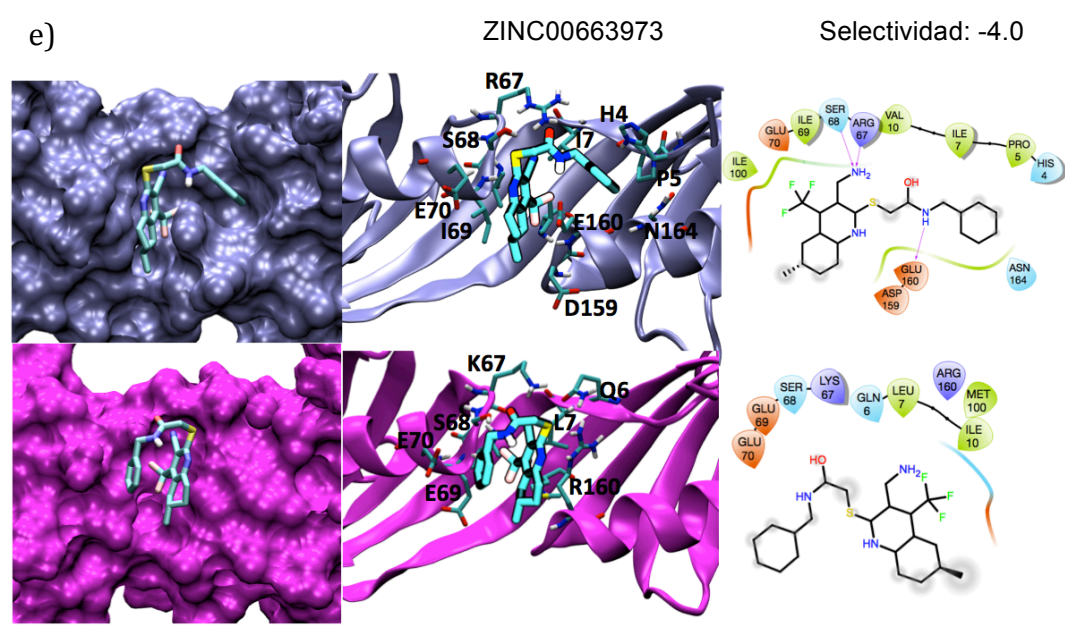
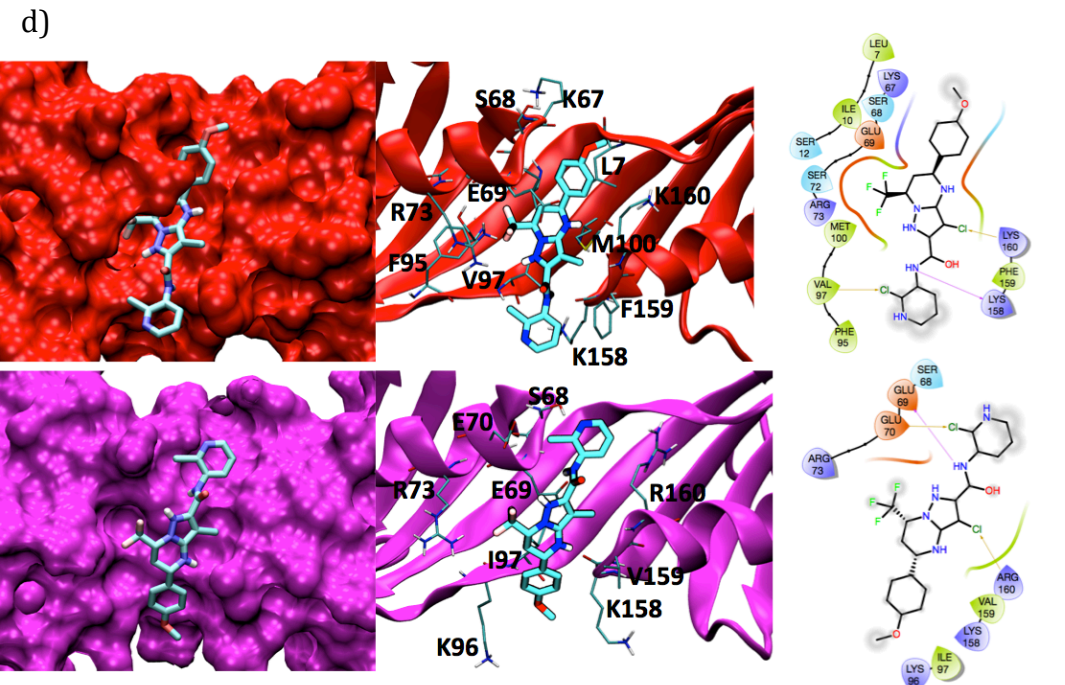
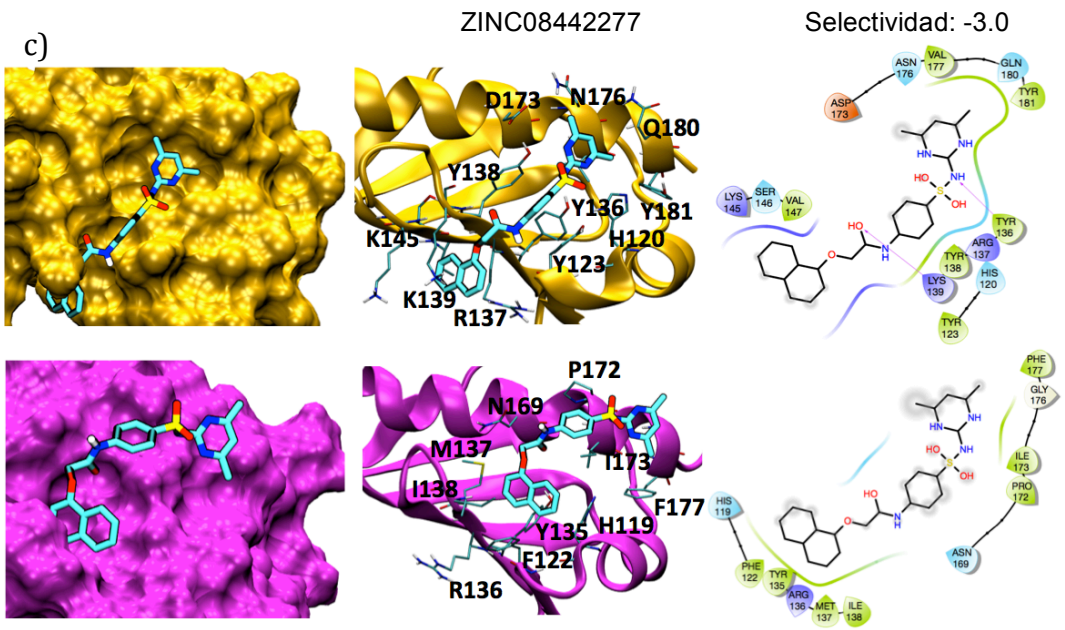


Figura 14. Compuestos con la mejor selectividad obtenidos del acoplamiento flexible. Los compuestos con mejor selectividad se muestran para **cal**TBP (naranja), **tso**TBP (verde), **nam**TBP (rojo), **ehi**TBP (azul metálico) y **pfa**TBP (amarillo), para a) el extremo N-terminal, b) sitio de unión a TFIIA, c) sitio de unión a NC2, y d) la región interdominio. Las representaciones de superficie, caricatura y mapa de interacciones en 2D se muestran para cada TBP. En la representación 2D, los átomos del compuesto con mayor accesibilidad al solvente están representados con sombras grises. En todos los casos se muestra la comparación con **hsa**TBP (magenta). El código de ZINC y los valores de selectividad se muestran en la parte superior.

Los mejores valores de selectividad se obtuvieron para el extremo N-terminal en **cal**TBP y para el sitio de unión a TFIIA en **tso**TBP. En el caso de **cal**TBP, una mejor interacción se debe a la mayor apertura de una cavidad en el extremo N-ter, mientras que ésta está cerrada en la estructura representativa de **hsa**TBP (Figura 14a). Cuando la apertura de esta cavidad fue calculada con POVME 3.0, una mayor apertura fue apreciada durante una parte de la trayectoria de MD para **cal**TBP comparada con **hsa**TBP (Figura 15).

TsoTBP y **ovo**TBP también mostraron esta cavidad con mayor apertura con respecto a **hsa**TBP. Sin embargo, estos sitios no fueron identificados durante el acoplamiento rígido. Esto puede estar relacionado con que el sitio N-ter no fue incluido para la creación de ensambles, y así la apertura de la cavidad no fue considerada. Esto es importante debido a que la predicción de cavidades con el servidor Metapocket indicó la presencia de la cavidad en los extremos N-ter de **tso**TBP y **ovo**TBP, sugiriendo otro sitio blanco novedoso.

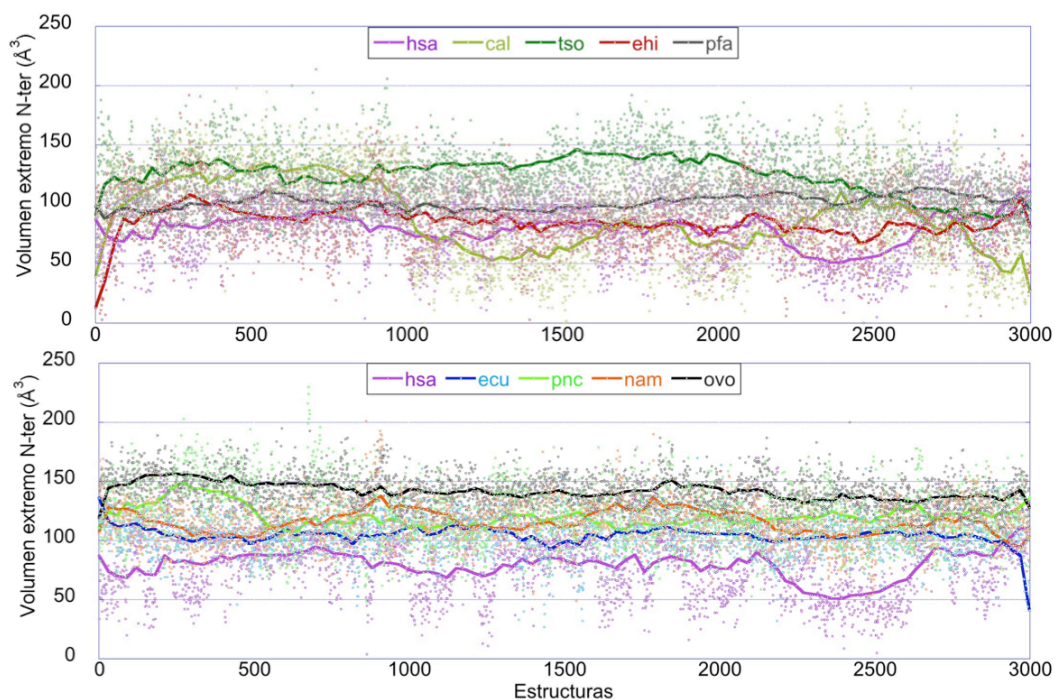


Figura 15. Dinámica de apertura del extremo N-terminal en las TBPs. Los volúmenes fueron calculados con POVME 3.0. El ajuste suavizado se muestra como líneas en cada gráfico.

Los resultados del acoplamiento sobre los sitios de unión a TFIIA para **namTBP** y **tsoTBP** mejoraron debido a un mayor número de interacciones en la cavidad en comparación con **hsaTBP** (Figura 14b). Esta cavidad fue más abierta durante las trayectorias de **tsoTBP** que en **hsaTBP**, explicando así su mejor afinidad y selectividad. En el caso de **namTBP**, esta cavidad mostró una apertura similar durante la MD que en **hsaTBP**. Sin embargo, la mejor orientación del compuesto promovió una mejor interacción en **namTBP**.

En el caso de **pfaTBP**, la interacción mejoró debido también a una mayor apertura de la cavidad del sitio de unión a NC2, comparado a **hsaTBP** (Figura 14c). **pfaTBP** presentó más diferencias en secuencia y en estructura, en comparación con **hsaTBP**, sugiriendo a este sitio como un potencial sitio. Con respecto al estribo N-ter de **pfaTBP**, encontramos moderada selectividad en comparación con **hsaTBP**. Este sitio fue diferente en las TBP de **pfaTBP** en comparación a **hsaTBP**, mientras que está conservado en el resto de las TBPs de parásitos. A diferencia de **hsaTBP**, la mejor pose fue localizada entre la hélice H1 y el estribo, mientras que **hsaTBP** ésta fue localizada sobre la cara del sitio de unión a DNA (Figura 16).

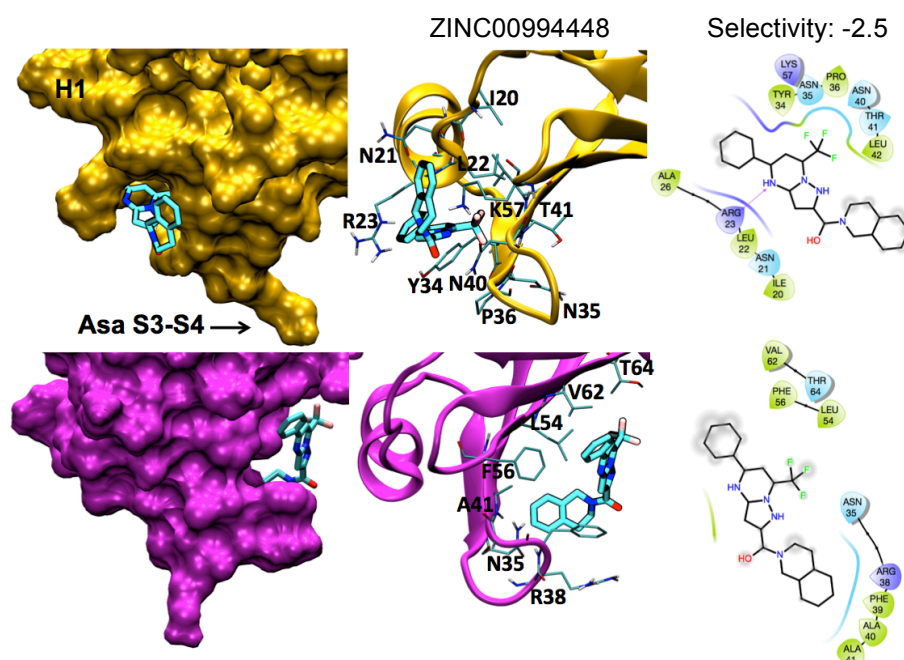


Figura 16. Compuesto con la más alta selectividad sobre el estribo N-ter en pfaTBP. Las representaciones de superficie, caricatura y 2D se muestran para **pfaTBP** y **hsaTBP**. En la representación 2D, los átomos del compuesto con mayor accesibilidad al solvente son

representados con sombras grises. En todos los casos se muestra la comparación con **hsaTBP** (magenta). El código de ZINC se muestra en la parte superior de la imagen.

Aunque los compuestos en la región interdominio de **namTBP** y **ehiTBP** mostraron baja selectividad en el acoplamiento rígido, el acoplamiento flexible mejoró la selectividad en ambas (Tabla 7). Esto puede estar relacionado con un pose más enterrada en este sitio y un mayor número de interacciones en **namTBP** y **ehiTBP** en comparación con **hsaTBP** (Figura 14d).

En resumen, para los compuestos con la mejor selectividad (Figura 14), las partes aromáticas establecen mejores interacciones de apilamiento e interacciones hidrofóbicas, debido a las cavidades más profundas o más abiertas en TBPs de parásitos, mientras que los grupos polares de los compuestos son estabilizados por enlaces de hidrógeno. Es interesante notar que los aminoácidos que decoran los sitios de unión a TFIIA y a NC2 son residuos de tipo aromático como Phe o Pro, que establecen interacciones de apilamiento con los compuestos con mejor selectividad. Esto sugiere que la búsqueda de compuestos aromáticos similares en estos sitios sería una buena estrategia. Con respecto a esto, los compuestos con mejor selectividad para el sitio de unión a TFIIA en **namTBP** y en **tsoTBP**, mostraron similitud estructural: estos contienen el grupo aromático pireno, e interesadamente los tres compuestos más selectivos son los mismos para ambas TBPs (Figura 17).

Para la región interdominio, el sitio de unión a NC2 y el extremo N-terminal, se encontró que los compuestos son más largos y presentaron conformaciones más extendidas en comparación a los compuestos del sitio de unión a TFIIA (Figura 17). Esto es interesante porque estos sitios están más expuestos en la superficie en comparación con el sitio de unión a TFIIA. Por tanto, los compuestos más largos y extendidos podrían ser más adecuados para establecer mejores interacciones. Los compuestos unidos al sitio de unión a NC2 de **pfaTBP** mantienen un grupo fenil seguido por una sulfona en tres de los compuestos con mejor selectividad, sugiriendo un armazón común para este sitio.

Con respecto a los compuestos en la región interdominio de **namTBP**, cuatro de estos compuestos mostraron el mismo centro aromático, un grupo tetrahidropirazolo pirimidina conectado a un tri fluorocarbono. Por su parte, los compuestos del extremo N-terminal mostraron mayor diversidad estructural, en comparación con los compuestos de otros sitios, sin embargo mantienen las partes aromáticos y conformaciones extendidas.

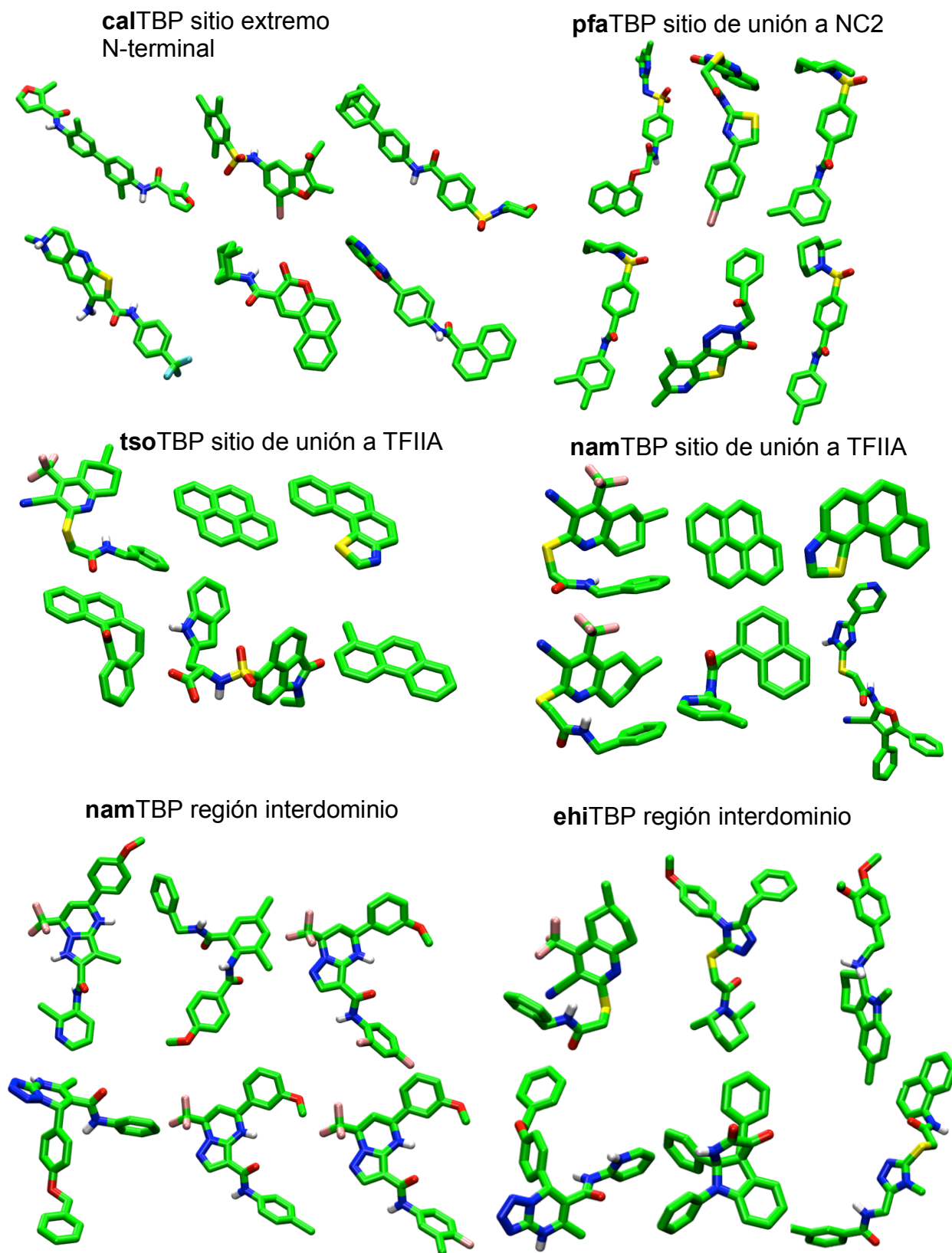


Figura 17. Top seis de los compuestos con más alta selectividad sobre los sitios de TBPs de parásitos. Todos los compuestos se muestran como varillas. Los átomos de carbono, oxígeno, nitrógeno, azufre y flúor, se muestran en color verde, azul, rojo, amarillo y rosa respectivamente.

Nuestro análisis muestra que las TBPs de parásitos son un blanco molecular atractivo, y que los compuestos identificados podrían ser considerados como armazones iniciales para su evaluación experimental.

5. Discusión

Este trabajo es el primer análisis estructural y dinámico de las TBPs de parásitos y su evaluación como posible blanco molecular de enfermedades parasitarias. Calculamos diferentes parámetros estructurales sobre las trayectorias de MD, encontrando sitios novedosos con las mayores diferencias con respecto a la TBP de humano, y evaluamos su potencial de unión a compuestos tipo fármaco por acoplamiento molecular de ensamble.

Con respecto al análisis de secuencias, las principales diferencias están presentes en la superficie convexa de TBPs (Figura 3b-d), un área de la proteína dedicada a las interacciones con múltiples proteínas. Sin embargo, poca información estructural está disponible para las TBPs de parásitos, y solo ensayos bioquímicos de unión a ADN y con algunos factores basales de la transcripción han sido reportados.^{44, 45, 46, 49, 50, 51, 52, 53, 54, 55} Sólo la estructura cristalográfica de **ecu**TBP unida al regulador transcripcional MOT1 está disponible.^{37, 47}

En contraste con la superficie convexa, la superficie de unión a ADN está ampliamente conservada en el conjunto de TBPs estudiadas, con excepción de las más divergentes **ehi**TBP y **pfa**TBP. De hecho los genomas de *Plasmodium* y *Entamoeba* son ricos en AT y diferentes preferencias por sus secuencias TATA han sido reportadas.^{44, 45} Con excepción de **pfa**TBP, los residuos de Phe que doblan el DNA están conservados en el sitio de unión a ADN, sugiriendo nuevamente diferencias en el reconocimiento de sus cajas TATA. Encontramos algunos compuestos que se unieron al sitio de unión a ADN de las TBPs de parásitos durante el acoplamiento de ensamble rígido. Sin embargo, éstos mostraron baja selectividad (relacionado con su alta conservación), por lo que este sitio no fue considerado en el acoplamiento flexible. Sólo una buena selectividad fue obtenida en el estribo N-terminal de **pfa**TBP, el cual es cercano al sitio de unión a ADN y mostró ser divergente comparado con **hsa**TBP (Figura 3c, 3d y Figura 16).

Por otra parte, el análisis de los parámetros estructurales indica diferencias en la polaridad de las TBPs: estas diferencias están relacionadas con la distribución de cargas en la superficie y no con el número total de cargas (Figuras 4, 6 y 7). La presencia de un “parche positivo” en la hélice H2 y un menor número de cargas

negativas en **pfa**TBP explica su peor electrostática (Figura 5, 6 y 7), mientras que la creación de un mayor número de puentes salinos en **tso**TBP explica su mejor electrostática (Figura 9). Adicionalmente encontramos un área superficial más pequeña para la hélice H2 en **tso**TBP (Figura 11d), que puede estar relacionado con el mayor número de interacciones de puente salino, y mejor electrostática. En contraste, el área de exposición para la hélice H2 en **pfa**TBP fue mayor y puede estar asociado con la repulsión de cargas positivas y por tanto su desfavorable electrostática (Figura 11d). Otra cosa interesante es que la hélice H2' de **pfa**TBP mostró un mayor número de residuos polares (Figura 4c). Sólo esta TBP mostró este parche polar, y en una menor extensión **tso**TBP y **ecu**TBP. Estos parches podrían estar asociados con interactores específicos en estas TBPs.

Los sitios de unión a TFIIA y NC2 son importantes para regular la función de TBP.^{29, 40} TFIIA es un activador transcripcional y NC2 es un regulador negativo de la transcripción, por lo que inhibir estos sitios podría ser letal para los parásitos. Encontramos que sus sitios de unión son cavidades sobre la superficie de la TBP, y para el caso del sitio de unión a NC2 fue más abierto en TBPs de parásitos en comparación a **hsa**TBP, mientras que el sitio de unión a TFIIA fue sólo ligeramente más abierto en **tso**TBP y **pfa**TBP (Figura 11a y 12a). El sitio de unión a TFIIA está muy conservado entre todas las TBPs (Figura 3), con excepción de **ehi**TBP y **pfa**TBP. A pesar de esta conservación, el acoplamiento de ensamble reveló buena selectividad para **nam**TBP y **tso**TBP. Encontramos que, durante las simulaciones, el rompimiento de un puente salino conservado entre el asa S3-S4 y la hélice H2, promueve una extensión de la cavidad del sitio de unión a TFIIA en **tso**TBP y **nam**TBP, en comparación con **hsa**TBP (Figura 11b y Anexo 4). La importancia de este particular puente salino para la unión a compuestos podría ser examinada por estudios de mutagénesis. Otro hallazgo interesante fue que el acoplamiento flexible sobre el sitio de unión a TFIIA de **tso**TBP y **nam**TBP seleccionó los mismos tres compuestos más selectivos (Tabla 7 y Figura 17), sugiriendo que la búsqueda de compuestos similares para estas dos TBPs dirigidos a este sitio, sería una buena opción, considerando además de que ambos son parásitos intestinales.

En contraste con el sitio de unión a TFIIA, el sitio de unión a NC2 fue menos conservado en todas las TBPs de parásitos, cuando fueron comparados con **hsa**TBP (Figuras 3b, 3c, 4c y 6a). **pfa**TBP mostró las mayores diferencias en este sitio. Una de estas diferencias es un grupo de residuos aromáticos (Y123, Y136 y Y138) que fue importante para la unión de compuestos selectivos (Figura 14c). Ninguna otra TBP mostró este grupo de residuos aromáticos en el sitio de unión a NC2. Interesantemente, todos los compuestos identificados por acoplamiento flexible en

pfaTBP (Tabla 4 y Figura 17), corresponden a compuestos aromáticos. La búsqueda de compuestos similares pueden ser una opción para mejorar la selectividad.

A pesar de que el sitio en el extremo N-terminal no fue considerado para crear los ensamblajes para el acoplamiento rígido, la unión de compuestos en este sitio mostró buena selectividad y afinidad para **calTBP** (Figura 14a). Con la excepción de **ehiTBP** y **pfaTBP**, todas las TBPs mostraron una ligera mayor apertura para esta cavidad en comparación con **hsaTBP** (Figura 15). A pesar de que **ovoTBP** y **tsoTBP** mostraron una mayor apertura comparada con **calTBP** (Figura 15 y Anexo 5), cuando un ensayo de acoplamiento flexible fue llevado a cabo sobre este sitio, la afinidad y selectividad fueron bajas en comparación de **calTBP** (Anexo 6). Diferentes bibliotecas podrían ser examinadas para estas TBPs. Este sitio podría ser un sitio blanco potencial en estas TBPs de parásitos. Además, el extremo N-terminal de TBP es altamente divergente en longitud y secuencia, y tiende a ser desordenado,⁹⁰ así la formación de esta cavidad podría ser muy dinámica en el contexto de la proteína completa.

Para la región interdominio encontramos un puente salino con diferente estabilidad en las TBPs estudiadas (Figura 9c y 9d). Este puente salino mostró una alta ocupación a través de las trayectorias, excepto para **namTBP** y **ovoTBP**. En el caso de **ehiTBP** y **pfaTBP**, este puente salino es ausente debido a que los residuos que participan en la formación no están conservados (Figura 9d). Aunque nosotros sólo encontramos un compuesto en este sitio para **namTBP** durante el acoplamiento de ensamble rígido, el acoplamiento flexible mejoró la selectividad. En este sentido, como la región de **ehiTBP** mostró diferencias con respecto a **hsaTBP**, un acoplamiento flexible fue llevado a cabo sobre este sitio en **ehiTBP** usando la misma biblioteca de compuestos. Encontramos que la selectividad mejoró con el docking flexible (Tabla 7 y Figura 14e), sugiriendo que la región interdominio es otro sitio potencial para ser considerado.

Una baja selectividad para el interdominio del resto de TBPs puede ser explicada por tres razones: 1) la biblioteca de compuestos usada para el acoplamiento es neutra, y la región interdominio mostró una carga negativa o positiva (en **ehiTBP**, **pfaTBP** y **calTBP**), sugiriendo que un esquema diferente de selección de compuestos podría ser requerida. Lo mismo aplica para la hélice H2, para la cual no encontramos compuestos selectivos.

2) Tanto la región interdominio, como la hélice H2, corresponden a sitios expuestos y no cavidades, y por tanto podrían resultar en sitios más difíciles. Con respecto a esto último, se ha mostrado que diferentes factores de transcripción son blancos moleculares difíciles, debido a sus superficies planas e interacciones dinámicas.^{58, 59}

3) La baja selectividad también puede estar asociada con el acoplamiento rígido, ya que el acoplamiento flexible mejoró la selectividad para los sitios evaluados. Esto

indica que nuestro paso de acoplamiento flexible es importante para evaluar el efecto inductivo en la formación del complejo y así mejorar la selectividad. Como perspectiva será necesario llevar a cabo el acoplamiento flexible en el resto de los sitios con diferencias estructurales y dinámicas con respecto a **hsa**TBP en otras TBPs.

Por otra parte, encontramos que la selectividad para **ecu**TBP, **pnc**TBP y **ovo**TBP fue moderada. Esto puede ser asociado con secuencias y superficies más conservadas con respecto a **hsa**TBP (Figura 3 y 4c). La comparación de la superficie electrostática usando webPIPSA también reveló superficies similares para **ecu**TBP, **pnc**TBP y **ovo**TBP (Figura 7b).

Durante nuestro análisis observamos un mejoramiento de la selectividad en el acoplamiento flexible usando la estructura representativa en sitios específicos en comparación con el acoplamiento de ensamble rígido. Esto puede deberse al parámetro de flexibilidad de la proteína incluido en la función de puntaje. En nuestro caso, el acoplamiento de ensamble rígido fue importante para determinar sitios potenciales. Algo en común entre los compuestos más selectivos identificados fue que se unieron con mejor afinidad y a cavidades más abiertas en las TBPs de parásitos, en comparación con **hsa**TBP. Esto indica la importancia de considerar la dinámica de la proteína, a pesar de la conservación en estos sitios.

Algunos de los compuestos selectivos identificados mostraron similitud estructural (Figura 17). Para aquellos localizados en el sitio de unión a TFIIA en **tso**TBP y **nam**TBP, los tres mejores compuestos selectivos fueron los mismos. Esto sugiere un armazón potencial para este sitio, donde la estructura central está relacionada con la estructura del pireno. En el caso del sitio de unión a NC2 en **pfa**TBP, los compuestos selectivos comparten un grupo fenil seguido de una sulfona, el cual fue importante para la interacción con el grupo de residuos de Tyr en **pfa**TBP. Los compuestos en la región interdominio de **nam**TBP mostraron el mismo centro aromático (tetrahidropirazolo pirimidina) conectado a un trifluorocarbono. Este grupo se enterró en una subcavidad en la región interdominio de **nam**TBP, mientras que en **hsa**TBP este grupo fue expuesto (Figura 14d). Sólo el compuesto más selectivo en la región interdominio de **ehi**TBP mantiene este grupo trifluorocarbono conectado a un centro aromático, pero el modo de unión fue diferente comparado con **nam**TBP. Este compuesto fue mejor acoplado comparado a **hsa**TBP, debido a que la parte aromática fue anclada en la cavidad del interdominio y parte del extremo N-terminal (Figura 14e). Por otra parte, los compuestos unidos al extremo N-terminal mostraron una mayor diversidad estructural, manteniendo una conformación extendida y anillos aromáticos que promueven interacciones aisladas del solvente.

Otro dato importante es que los compuestos más selectivos mostrados en la Tabla 7, no tiene actividad anotada y no están siendo utilizados en ensayos clínicos. Por lo que estos compuestos selectivos podrían ser propuestos como armazones iniciales para evaluar experimentalmente. El análisis estructural junto con los métodos de acoplamiento molecular sugieren a las TBPs de parásitos como posibles blancos moleculares de enfermedades parasitarias.

6. Conclusiones

Nuestro estudio de las propiedades de superficie en las TBPs de parásitos, junto con la evaluación de su potencial unión a compuestos tipo fármaco, hacen a TBP un blanco interesante para ser evaluado experimentalmente. Las propiedades estructurales y dinámicas calculadas revelan que las principales diferencias entre la TBP humana y las de parásitos corresponden al sitio de unión a NC2, la hélice H2, el extremo N-terminal y la región interdominio.

A pesar de la conservación entre las TBPs analizadas, encontramos buena selectividad para **tso**TBP, **cal**TBP, **nam**TBP, **ehi**TBP y **pfa**TBP. La mejor accesibilidad farmacológica en **cal**TBP, **tso**TBP, **ehi**TBP y **nam**TBP fue principalmente debido a cambios estructurales y dinámicos, como la apertura de cavidades y pérdida de interacciones en los sitios propuestos. Esto indica la importancia de tomar en cuenta cambios estructurales al nivel de cadena principal y lateral cuando se modelan complejos proteína-ligando. En el caso de **pfa**TBP, la mejor selectividad fue en el sitio de unión a NC2, debido a que su cavidad es más abierta, y a que se encontraron mayores diferencias estructurales y de secuencia con respecto a **hsa**TBP.

Los mejores compuestos selectivos identificados en este trabajo mostraron una similitud estructural y afinidades en el intervalo micromolar a nanomolar, indicando armazones potenciales que pueden servir como puntos de inicio en pruebas experimentales. Finalmente, la búsqueda de compuestos con estructuras similares a los compuestos identificados en nuestro análisis podría ser explorada para tener una biblioteca de compuestos más diversa, y así mejorar la selectividad de los compuestos que pueden ser propuestos como compuestos líderes para su evaluación experimental.

7. Perspectivas

1. Aumentar el número TBPs de parásitos a analizar, con el fin de proponer un set de andamios estructurales más amplio a evaluar experimentalmente.

2. Llevar a cabo el acoplamiento flexible en el resto de los sitios con diferencias estructurales y dinámicas con respecto a **hsa**TBP en otras TBPs.
3. Evaluar nuevas bibliotecas de compuestos por acoplamiento molecular, así como proponer la selección de ensamblajes de estructuras basados en la dinámica de apertura de las cavidades identificadas.
4. Utilizar otras funciones de puntaje para evaluar la selectividad de las poses generadas con Autodock Vina.
5. Establecer colaboraciones con grupos de investigación que trabajan directamente con los parásitos, con el fin de evaluar el potencial de los compuestos identificados en nuestro estudio.

8. Bibliografía

- (1) Hotez, P. J. Chapter Two - Human Parasitology and Parasitic Diseases: Heading Towards 2050. In *Advances in Parasitology*; Rollinson, D., Stothard, J. R., Eds.; Academic Press, 2018; Vol. 100, pp 29–38.
- (2) Lustigman, S.; Prichard, R. K.; Gazzinelli, A.; Grant, W. N.; Boatman, B. A.; McCarthy, J. S.; Basáñez, M.-G. A Research Agenda for Helminth Diseases of Humans: The Problem of Helminthiases. *PLoS Negl. Trop. Dis.* **2012**, *6* (4), e1582. <https://doi.org/10.1371/journal.pntd.0001582>.
- (3) Hotez, P. J.; Dumonteil, E.; Betancourt Cravioto, M.; Bottazzi, M. E.; Tapia-Conyer, R.; Meymandi, S.; Karunakara, U.; Ribeiro, I.; Cohen, R. M.; Pecoul, B. An Unfolding Tragedy of Chagas Disease in North America. *PLoS Negl Trop Dis* **2013**, *7* (10), e2300. <https://doi.org/10.1371/journal.pntd.0002300>.
- (4) Gutierrez-Jimenez, J.; Torres-Sanchez, M. G. C.; Fajardo-Martinez, L. P.; Schlie-Guzman, M. A.; Luna-Cazares, L. M.; Gonzalez-Esquinca, A. R.; Guerrero-Fuentes, S.; Vidal, J. E. Malnutrition and the Presence of Intestinal Parasites in Children from the Poorest Municipalities of Mexico. *J. Infect. Dev. Ctries.* **2013**, *7* (10), 741–747. <https://doi.org/10.3855/jidc.2990>.
- (5) Bhatt, S.; Weiss, D. J.; Cameron, E.; Bisanzio, D.; Mappin, B.; Dalrymple, U.; Battle, K. E.; Moyes, C. L.; Henry, A.; Eckhoff, P. A.; et al. The Effect of Malaria Control on Plasmodium Falciparum in Africa between 2000 and 2015. *Nature* **2015**, *526* (7572), 207–211. <https://doi.org/10.1038/nature15535>.
- (6) Burrows, J. N.; Burlot, E.; Campo, B.; Cherbuin, S.; Jeanneret, S.; Leroy, D.;

Spangenberg, T.; Waterson, D.; Wells, T. N.; Willis, P. Antimalarial Drug Discovery - the Path towards Eradication. *Parasitology* **2014**, *141* (1), 128–139. <https://doi.org/10.1017/S0031182013000826>.

(7) Hemingway, J. Battling Disease by Giving Mosquitoes an Antimalarial Drug. *Nature* **2019**, *567* (7747), 185. <https://doi.org/10.1038/d41586-019-00648-2>.

(8) Pham, J. S.; Dawson, K. L.; Jackson, K. E.; Lim, E. E.; Pasaje, C. F. A.; Turner, K. E. C.; Ralph, S. A. Aminoacyl-tRNA Synthetases as Drug Targets in Eukaryotic Parasites. *Int. J. Parasitol. Drugs Drug Resist.* **2014**, *4* (1), 1–13. <https://doi.org/10.1016/j.ijpddr.2013.10.001>.

(9) Upcroft, P.; Upcroft, J. A. Drug Targets and Mechanisms of Resistance in the Anaerobic Protozoa. *Clin. Microbiol. Rev.* **2001**, *14* (1), 150–164. <https://doi.org/10.1128/CMR.14.1.150-164.2001>.

(10) Lloyd, A. E.; Honey, B. L.; John, B. M.; Condren, M. Treatment Options and Considerations for Intestinal Helminthic Infections. *J. Pharm. Technol. JPT Off. Publ. Assoc. Pharm. Tech.* **2014**, *30* (4), 130–139. <https://doi.org/10.1177/8755122514533667>.

(11) Ottesen, E. A.; Campbell, W. C. Ivermectin in Human Medicine. *J. Antimicrob. Chemother.* **1994**, *34* (2), 195–203.

(12) Maya, J. D.; Repetto, Y.; Agosín, M.; Ojeda, J. M.; Tellez, R.; Gaule, C.; Morello, A. Effects of Nifurtimox and Benznidazole upon Glutathione and Trypanothione Content in Epimastigote, Trypomastigote and Amastigote Forms of *Trypanosoma Cruzi*. *Mol. Biochem. Parasitol.* **1997**, *86* (1), 101–106. [https://doi.org/10.1016/S0166-6851\(96\)02837-X](https://doi.org/10.1016/S0166-6851(96)02837-X).

(13) Francisco, A. F.; Lewis, M. D.; Jayawardhana, S.; Taylor, M. C.; Chatelain, E.; Kelly, J. M. Limited Ability of Posaconazole To Cure Both Acute and Chronic *Trypanosoma Cruzi* Infections Revealed by Highly Sensitive In Vivo Imaging. *Antimicrob. Agents Chemother.* **2015**, *59* (8), 4653–4661. <https://doi.org/10.1128/AAC.00520-15>.

(14) Rachidi, N.; Taly, J. F.; Durieu, E.; Leclercq, O.; Aulner, N.; Prina, E.; Pescher, P.; Notredame, C.; Meijer, L.; Späth, G. F. Pharmacological Assessment Defines *Leishmania Donovanii* Casein Kinase 1 as a Drug Target and Reveals Important Functions in Parasite Viability and Intracellular Infection. *Antimicrob. Agents Chemother.* **2014**, *58* (3), 1501–1515. <https://doi.org/10.1128/AAC.02022-13>.

- (15) Nisius, B.; Sha, F.; Gohlke, H. Structure-Based Computational Analysis of Protein Binding Sites for Function and Druggability Prediction. *J. Biotechnol.* **2012**, *159* (3), 123–134. <https://doi.org/10.1016/j.jbiotec.2011.12.005>.
- (16) Pérot, S.; Sperandio, O.; Miteva, M. A.; Camproux, A.-C.; Villoutreix, B. O. Druggable Pockets and Binding Site Centric Chemical Space: A Paradigm Shift in Drug Discovery. *Drug Discov. Today* **2010**, *15* (15), 656–667. <https://doi.org/10.1016/j.drudis.2010.05.015>.
- (17) Nero, T. L.; Morton, C. J.; Holien, J. K.; Wielens, J.; Parker, M. W. Oncogenic Protein Interfaces: Small Molecules, Big Challenges. *Nat. Rev. Cancer* **2014**, *14* (4), 248–262. <https://doi.org/10.1038/nrc3690>.
- (18) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28* (1), 235–242.
- (19) Henrich, S.; Salo-Ahen, O. M. H.; Huang, B.; Rippmann, F. F.; Cruciani, G.; Wade, R. C. Computational Approaches to Identifying and Characterizing Protein Binding Sites for Ligand Design. *J. Mol. Recognit. JMR* **2010**, *23* (2), 209–219. <https://doi.org/10.1002/jmr.984>.
- (20) Cross, J. B.; Thompson, D. C.; Rai, B. K.; Baber, J. C.; Fan, K. Y.; Hu, Y.; Humblet, C. Comparison of Several Molecular Docking Programs: Pose Prediction and Virtual Screening Accuracy. *J. Chem. Inf. Model.* **2009**, *49* (6), 1455–1474. <https://doi.org/10.1021/ci900056c>.
- (21) Yuriev, E.; Ramsland, P. A. Latest Developments in Molecular Docking: 2010-2011 in Review. *J. Mol. Recognit. JMR* **2013**, *26* (5), 215–239. <https://doi.org/10.1002/jmr.2266>.
- (22) Sterling, T.; Irwin, J. J. ZINC 15 – Ligand Discovery for Everyone. *J. Chem. Inf. Model.* **2015**, *55* (11), 2324–2337. <https://doi.org/10.1021/acs.jcim.5b00559>.
- (23) Porter, J. R.; Moeder, K. E.; Sibbald, C. A.; Zimmerman, M. I.; Hart, K. M.; Greenberg, M. J.; Bowman, G. R. Cooperative Changes in Solvent Exposure Identify Cryptic Pockets, Switches, and Allosteric Coupling. *Biophys. J.* **2019**, *116* (5), 818–830. <https://doi.org/10.1016/j.bpj.2018.11.3144>.
- (24) Stank, A.; Kokh, D. B.; Fuller, J. C.; Wade, R. C. Protein Binding Pocket Dynamics. *Acc. Chem. Res.* **2016**, *49* (5), 809–815. <https://doi.org/10.1021/acs.accounts.5b00516>.

- (25) Paramo, T.; East, A.; Garzón, D.; Ulmschneider, M. B.; Bond, P. J. Efficient Characterization of Protein Cavities within Molecular Simulation Trajectories: Trj_cavity. *J. Chem. Theory Comput.* **2014**, *10* (5), 2151–2164. <https://doi.org/10.1021/ct401098b>.
- (26) Pradhan, M. R.; Siau, J. W.; Kannan, S.; Nguyen, M. N.; Ouaray, Z.; Kwoh, C. K.; Lane, D. P.; Ghadessy, F.; Verma, C. S. Simulations of Mutant p53 DNA Binding Domains Reveal a Novel Druggable Pocket. *Nucleic Acids Res.* **2019**, *47* (4), 1637–1652. <https://doi.org/10.1093/nar/gky1314>.
- (27) Huang, S.-Y.; Zou, X. Ensemble Docking of Multiple Protein Structures: Considering Protein Structural Variations in Molecular Docking. *Proteins Struct. Funct. Bioinforma.* **2007**, *66* (2), 399–421. <https://doi.org/10.1002/prot.21214>.
- (28) Ellingson, S. R.; Miao, Y.; Baudry, J.; Smith, J. C. Multi-Conformer Ensemble Docking to Difficult Protein Targets. *J. Phys. Chem. B* **2015**, *119* (3), 1026–1034. <https://doi.org/10.1021/jp506511p>.
- (29) Sainsbury, S.; Bernecky, C.; Cramer, P. Structural Basis of Transcription Initiation by RNA Polymerase II. *Nat. Rev. Mol. Cell Biol.* **2015**, *16* (3), 129–143. <https://doi.org/10.1038/nrm3952>.
- (30) Duttke, S. H. C. Evolution and Diversification of the Basal Transcription Machinery. *Trends Biochem. Sci.* **2015**, *40* (3), 127–129. <https://doi.org/10.1016/j.tibs.2015.01.005>.
- (31) Rowlands, T.; Baumann, P.; Jackson, S. P. The TATA-Binding Protein: A General Transcription Factor in Eukaryotes and Archaeobacteria. *Science* **1994**, *264* (5163), 1326–1329. <https://doi.org/10.1126/science.8191287>.
- (32) Adachi, N.; Senda, M.; Natsume, R.; Senda, T.; Horikoshi, M. Crystal Structure of Methanococcus Jannaschii TATA Box-Binding Protein. *Genes Cells Devoted Mol. Cell. Mech.* **2008**, *13* (11), 1127–1140. <https://doi.org/10.1111/j.1365-2443.2008.01233.x>.
- (33) Adachi, N.; Natsume, R.; Senda, M.; Muto, S.; Senda, T.; Horikoshi, M. Purification, Crystallization and Preliminary X-Ray Analysis of Methanococcus Jannaschii TATA Box-Binding Protein (TBP). *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60* (Pt 12 Pt 2), 2328–2331.
- (34) Chasman, D. I.; Flaherty, K. M.; Sharp, P. A.; Kornberg, R. D. Crystal Structure of Yeast TATA-Binding Protein and Model for Interaction with DNA. *Proc. Natl.*

- Acad. Sci.* **1993**, *90* (17), 8174–8178. <https://doi.org/10.1073/pnas.90.17.8174>.
- (35) Nikolov, D. B.; Burley, S. K. 2.1 Å Resolution Refined Structure of a TATA Box-Binding Protein (TBP). *Nat. Struct. Biol.* **1994**, *1* (9), 621. <https://doi.org/10.1038/nsb0994-621>.
- (36) Bleichenbacher, M.; Tan, S.; Richmond, T. J. Novel Interactions Between the Components of Human and Yeast TFIIA/TBP/DNA Complexes. *J. Mol. Biol.* **2003**, *332* (4), 783–793. [https://doi.org/10.1016/S0022-2836\(03\)00887-8](https://doi.org/10.1016/S0022-2836(03)00887-8).
- (37) Wollmann, P.; Cui, S.; Viswanathan, R.; Berninghausen, O.; Wells, M. N.; Moldt, M.; Witte, G.; Butryn, A.; Wendler, P.; Beckmann, R.; et al. Structure and Mechanism of the Swi2/Snf2 Remodeller Mot1 in Complex with Its Substrate TBP. *Nature* **2011**, *475* (7356), 403–407. <https://doi.org/10.1038/nature10215>.
- (38) DeDecker, B. S.; O'Brien, R.; Fleming, P. J.; Geiger, J. H.; Jackson, S. P.; Sigler, P. B. The Crystal Structure of a Hyperthermophilic Archaeal TATA-Box Binding Protein. *J. Mol. Biol.* **1996**, *264* (5), 1072–1084. <https://doi.org/10.1006/jmbi.1996.0697>.
- (39) Koike, H.; Kawashima-Ohya, Y.; Yamasaki, T.; Clowney, L.; Katsuya, Y.; Suzuki, M. Origins of Protein Stability Revealed by Comparing Crystal Structures of TATA Binding Proteins. *Struct. Lond. Engl. 1993* **2004**, *12* (1), 157–168.
- (40) Pugh, B. F. Control of Gene Expression through Regulation of the TATA-Binding Protein. *Gene* **2000**, *255* (1), 1–14.
- (41) Anandapadamanaban, M.; Andresen, C.; Helander, S.; Ohyama, Y.; Siponen, M. I.; Lundström, P.; Kokubo, T.; Ikura, M.; Moche, M.; Sunnerhagen, M. High-Resolution Structure of TBP with TAF1 Reveals Anchoring Patterns in Transcriptional Regulation. *Nat. Struct. Mol. Biol.* **2013**, *20* (8), 1008–1014. <https://doi.org/10.1038/nsmb.2611>.
- (42) Iwasaki, O.; Tanizawa, H.; Kim, K.-D.; Yokoyama, Y.; Corcoran, C. J.; Tanaka, A.; Skordalakes, E.; Showe, L. C.; Noma, K.-I. Interaction between TBP and Condensin Drives the Organization and Faithful Segregation of Mitotic Chromosomes. *Mol. Cell* **2015**, *59* (5), 755–767. <https://doi.org/10.1016/j.molcel.2015.07.007>.
- (43) Fast, N. M.; Logsdon, J. M.; Doolittle, W. F. Phylogenetic Analysis of the TATA Box Binding Protein (TBP) Gene from *Nosema* Locustae: Evidence for a Microsporidia-Fungi Relationship and Spliceosomal Intron Loss. *Mol. Biol. Evol.*

1999, 16 (10), 1415–1419.

(44) de Dios-Bravo, G.; Luna-Arias, J. P.; Riverón, A. M.; Olivares-Trejo, J. J.; López-Camarillo, C.; Orozco, E. Entamoeba Histolytica TATA-Box Binding Protein Binds to Different TATA Variants in Vitro. *FEBS J.* **2005**, 272 (6), 1354–1366. <https://doi.org/10.1111/j.1742-4658.2005.04566.x>.

(45) McAndrew, M. B.; Read, M.; Sims, P. F. G.; Hyde, J. E. Characterisation of the Gene Encoding an Unusually Divergent TATA-Binding Protein (TBP) from the Extremely A+T-Rich Human Malaria Parasite Plasmodium Falciparum. *Gene* **1993**, 124 (2), 165–171. [https://doi.org/10.1016/0378-1119\(93\)90390-0](https://doi.org/10.1016/0378-1119(93)90390-0).

(46) Millership, J. J.; Waghela, P.; Cai, X.; Cockerham, A.; Zhu, G. Differential Expression and Interaction of Transcription Co-Activator MBF1 with TATA-Binding Protein (TBP) in the Apicomplexan Cryptosporidium Parvum. *Microbiol. Read. Engl.* **2004**, 150 (Pt 5), 1207–1213. <https://doi.org/10.1099/mic.0.26891-0>.

(47) Butryn, A.; Schuller, J. M.; Stoehr, G.; Runge-Wollmann, P.; Förster, F.; Auble, D. T.; Hopfner, K.-P. Structural Basis for Recognition and Remodeling of the TBP:DNA:NC2 Complex by Mot1. *eLife* **2015**, 4. <https://doi.org/10.7554/eLife.07432>.

(48) McAndrew, M. B.; Read, M.; Sims, P. F. G.; Hyde, J. E. Characterisation of the Gene Encoding an Unusually Divergent TATA-Binding Protein (TBP) from the Extremely A+T-Rich Human Malaria Parasite Plasmodium Falciparum. *Gene* **1993**, 124 (2), 165–171. [https://doi.org/10.1016/0378-1119\(93\)90390-0](https://doi.org/10.1016/0378-1119(93)90390-0).

(49) Ruan, J.-P.; Arhin, G. K.; Ullu, E.; Tschudi, C. Functional Characterization of a Trypanosoma Brucei TATA-Binding Protein-Related Factor Points to a Universal Regulator of Transcription in Trypanosomes. *Mol. Cell. Biol.* **2004**, 24 (21), 9610–9618. <https://doi.org/10.1128/MCB.24.21.9610-9618.2004>.

(50) Cribb, P.; Esteban, L.; Trochine, A.; Girardini, J.; Serra, E. Trypanosoma Cruzi TBP Shows Preference for C/G-Rich DNA Sequences in Vitro. *Exp. Parasitol.* **2010**, 124 (3), 346–349. <https://doi.org/10.1016/j.exppara.2009.11.003>.

(51) Leng, P.; Carter, P. E.; Brown, A. J. P. The TATA-Binding Protein (TBP) from the Human Fungal Pathogen Candida Albicans Can Complement Defects in Human and Yeast TBPs. *J. Bacteriol.* **1998**, 180 (7), 1771–1776.

(52) Rodríguez-Lima, O.; García-Gutierrez, P.; Jiménez, L.; Zarain-Herzberg, Á.; Lazzarini, R.; Landa, A. Molecular Cloning of a cDNA Encoding for Taenia Solium

- TATA-Box Binding Protein 1 (TsTBP1) and Study of Its Interactions with the TATA-Box of Actin 5 and Typical 2-Cys Peroxiredoxin Genes. *PloS One* **2015**, *10* (11), e0141818. <https://doi.org/10.1371/journal.pone.0141818>.
- (53) de Dios-Bravo, G.; López, C.; Luna-Arias, J. P.; Orozco, E. DNA Binding Activity and Predicted Tertiary Structure of the TATA Binding Protein of *Entamoeba Histolytica*. *Arch. Med. Res.* **2000**, *31* (4, Supplement 1), S299–S300. [https://doi.org/10.1016/S0188-4409\(00\)00126-0](https://doi.org/10.1016/S0188-4409(00)00126-0).
- (54) Li, S.; Donelson, J. E. The Gene for the TATA Box-Binding Protein of *Onchocerca Volvulus*. *Mol. Biochem. Parasitol.* **1993**, *61* (2), 321–324.
- (55) Sunkin, S. M.; Stringer, J. R. Transcription Factor Genes from Rat *Pneumocystis Carinii*. *J. Eukaryot. Microbiol.* **1995**, *42* (1), 12–19.
- (56) Walker, S. S.; Degen, D.; Nickbarg, E.; Carr, D.; Soriano, A.; Mandal, M.; Painter, R. E.; Sheth, P.; Xiao, L.; Sher, X.; et al. Affinity Selection–Mass Spectrometry Identifies a Novel Antibacterial RNA Polymerase Inhibitor. *ACS Chem. Biol.* **2017**, *12* (5), 1346–1352. <https://doi.org/10.1021/acscchembio.6b01133>.
- (57) Hoshiyama, D.; Kuma, K.; Miyata, T. Extremely Reduced Evolutionary Rate of TATA-Box Binding Protein in Higher Vertebrates and Its Evolutionary Implications. *Gene* **2001**, *280* (1), 169–173. [https://doi.org/10.1016/S0378-1119\(01\)00766-1](https://doi.org/10.1016/S0378-1119(01)00766-1).
- (58) Dang, C. V.; Reddy, E. P.; Shokat, K. M.; Soucek, L. Drugging the “Undruggable” Cancer Targets. *Nat. Rev. Cancer* **2017**, *17* (8), 502–508. <https://doi.org/10.1038/nrc.2017.36>.
- (59) Mapp, A. K.; Pricer, R.; Sturlis, S. Targeting Transcription Is No Longer a Quixotic Quest. *Nat. Chem. Biol.* **2015**, *11* (12), 891–894. <https://doi.org/10.1038/nchembio.1962>.
- (60) Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. ZINC: A Free Tool to Discover Chemistry for Biology. *J. Chem. Inf. Model.* **2012**, *52* (7), 1757–1768. <https://doi.org/10.1021/ci3001277>.
- (61) Juo, Z. S.; Kassavetis, G. A.; Wang, J.; Geiduschek, E. P.; Sigler, P. B. Crystal Structure of a Transcription Factor IIIB Core Interface Ternary Complex. *Nature* **2003**, *422* (6931), 534–539. <https://doi.org/10.1038/nature01534>.
- (62) Tsai, F. T.; Sigler, P. B. Structural Basis of Preinitiation Complex Assembly on Human Pol II Promoters. *EMBO J.* **2000**, *19* (1), 25–36.

- <https://doi.org/10.1093/emboj/19.1.25>.
- (63) Yang, J.; Yan, R.; Roy, A.; Xu, D.; Poisson, J.; Zhang, Y. The I-TASSER Suite: Protein Structure and Function Prediction. *Nat. Methods* **2015**, *12* (1), 7–8. <https://doi.org/10.1038/nmeth.3213>.
- (64) Fiser, A.; Sali, A. Modeller: Generation and Refinement of Homology-Based Protein Structure Models. *Methods Enzymol.* **2003**, *374*, 461–491. [https://doi.org/10.1016/S0076-6879\(03\)74020-8](https://doi.org/10.1016/S0076-6879(03)74020-8).
- (65) Biasini, M.; Bienert, S.; Waterhouse, A.; Arnold, K.; Studer, G.; Schmidt, T.; Kiefer, F.; Cassarino, T. G.; Bertoni, M.; Bordoli, L.; et al. SWISS-MODEL: Modelling Protein Tertiary and Quaternary Structure Using Evolutionary Information. *Nucleic Acids Res.* **2014**, *42* (Web Server issue), W252–W258. <https://doi.org/10.1093/nar/gku340>.
- (66) Lovell, S. C.; Davis, I. W.; Arendall, W. B.; de Bakker, P. I. W.; Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C. Structure Validation by C α Geometry: ϕ, ψ and C β Deviation. *Proteins Struct. Funct. Bioinforma.* **2003**, *50* (3), 437–450. <https://doi.org/10.1002/prot.10286>.
- (67) Link [Http://www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov).
- (68) McWilliam, H.; Li, W.; Uludag, M.; Squizzato, S.; Park, Y. M.; Buso, N.; Cowley, A. P.; Lopez, R. Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res.* **2013**, *41* (W1), W597–W600. <https://doi.org/10.1093/nar/gkt376>.
- (69) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5).
- (70) Ashkenazy, H.; Erez, E.; Martz, E.; Pupko, T.; Ben-Tal, N. ConSurf 2010: Calculating Evolutionary Conservation in Sequence and Structure of Proteins and Nucleic Acids. *Nucleic Acids Res.* **2010**, *38* (Web Server issue), W529–W533. <https://doi.org/10.1093/nar/gkq399>.
- (71) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; Mackerell, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ, ψ and Side-Chain $\chi(1)$ and $\chi(2)$ Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8* (9), 3257–3273. <https://doi.org/10.1021/ct300400x>.
- (72) Brooks, B. R.; Brooks, C. L.; MacKerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The Biomolecular

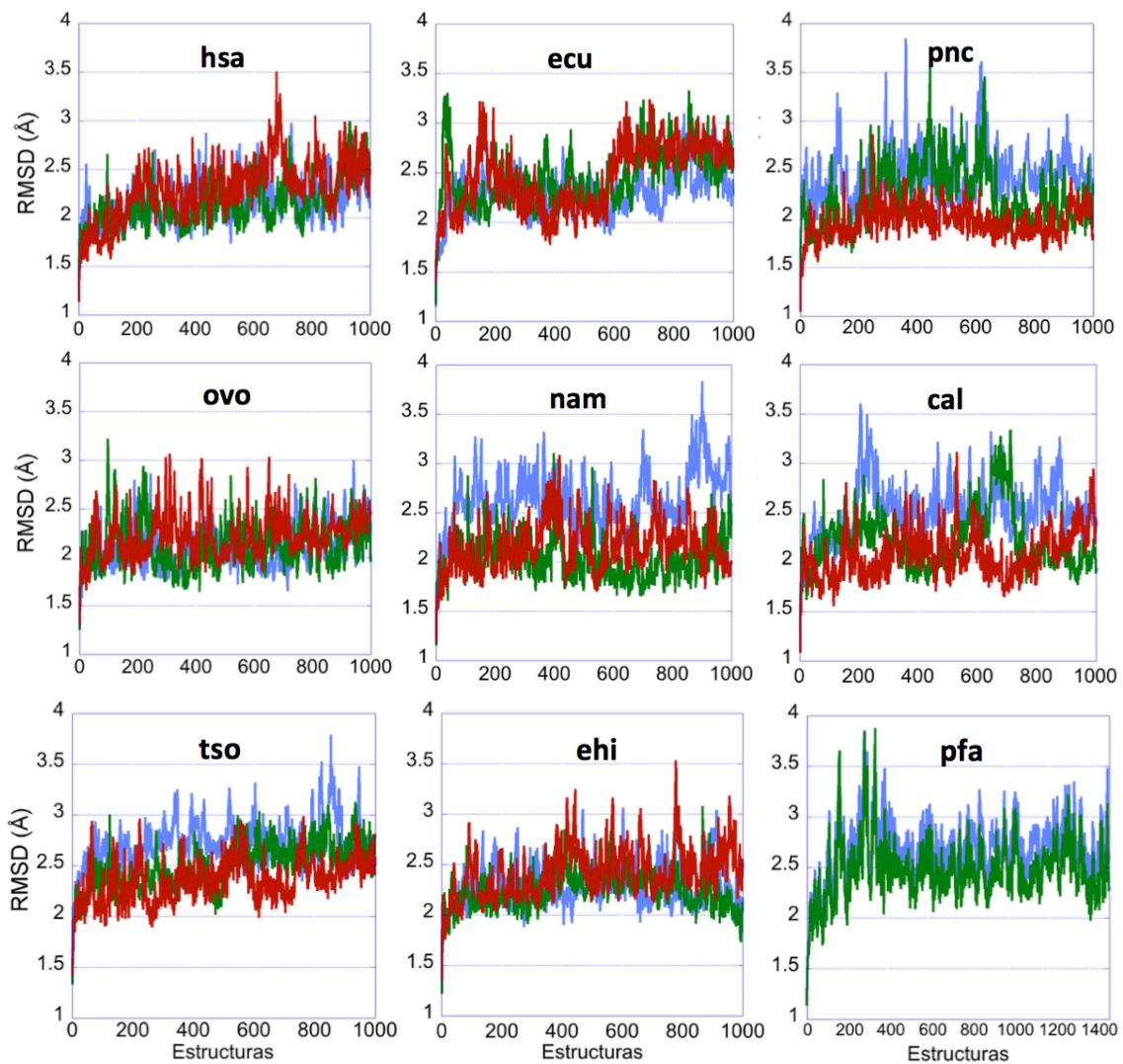
- Simulation Program. *J. Comput. Chem.* **2009**, *30* (10), 1545–1614. <https://doi.org/10.1002/jcc.21287>.
- (73) Jo, S.; Cheng, X.; Lee, J.; Kim, S.; Park, S.-J.; Patel, D. S.; Beaven, A. H.; Lee, K. I.; Rui, H.; Park, S.; et al. CHARMM-GUI 10 Years for Biomolecular Modeling and Simulation. *J. Comput. Chem.* **2017**, *38* (15), 1114–1124. <https://doi.org/10.1002/jcc.24660>.
- (74) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781–1802. <https://doi.org/10.1002/jcc.20289>.
- (75) Mezei, M. Simulaid: A Simulation Facilitator and Analysis Program. *J. Comput. Chem.* **2010**, *31* (14), 2658–2668. <https://doi.org/10.1002/jcc.21551>.
- (76) Tien, M. Z.; Meyer, A. G.; Sydykova, D. K.; Spielman, S. J.; Wilke, C. O. Maximum Allowed Solvent Accessibilities of Residues in Proteins. *PloS One* **2013**, *8* (11), e80635. <https://doi.org/10.1371/journal.pone.0080635>.
- (77) Wagner, J. R.; Sørensen, J.; Hensley, N.; Wong, C.; Zhu, C.; Perison, T.; Amaro, R. E. POVME 3.0: Software for Mapping Binding Pocket Flexibility. *J. Chem. Theory Comput.* **2017**, *13* (9), 4584–4592. <https://doi.org/10.1021/acs.jctc.7b00500>.
- (78) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. PDB2PQR: An Automated Pipeline for the Setup of Poisson–Boltzmann Electrostatics Calculations. *Nucleic Acids Res.* **2004**, *32* (suppl 2), W665–W667. <https://doi.org/10.1093/nar/gkh381>.
- (79) Richter, S.; Wenzel, A.; Stein, M.; Gabdoulline, R. R.; Wade, R. C. webPIPSA: A Web Server for the Comparison of Protein Interaction Properties. *Nucleic Acids Res.* **2008**, *36* (suppl_2), W276–W280. <https://doi.org/10.1093/nar/gkn181>.
- (80) Schrödinger Release 2019-1. *Maestro*, Schrödinger, LLC.; New York, NY, USA, 2019.
- (81) Huang, B. MetaPocket: A Meta Approach to Improve Protein Ligand Binding Site Prediction. *Omics J. Integr. Biol.* **2009**, *13* (4), 325–330. <https://doi.org/10.1089/omi.2009.0045>.
- (82) Dunbrack, R. L.; Cohen, F. E. Bayesian Statistical Analysis of Protein Side-Chain Rotamer Preferences. *Protein Sci. Publ. Protein Soc.* **1997**, *6* (8), 1661–1681.
- (83) Latest Improvements in Molegro Virtual Docker. *CLC bio*.

- (84) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug Deliv. Rev.* **2001**, *46* (1-3), 3–26.
- (85) Veber, D. F.; Johnson, S. R.; Cheng, H.-Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. *J. Med. Chem.* **2002**, *45* (12), 2615–2623.
- (86) Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, *31* (2), 455–461. <https://doi.org/10.1002/jcc.21334>.
- (87) Perryman, A. L.; Yu, W.; Wang, X.; Ekins, S.; Forli, S.; Li, S.-G.; Freundlich, J. S.; Tonge, P. J.; Olson, A. J. A Virtual Screen Discovers Novel, Fragment-Sized Inhibitors of Mycobacterium Tuberculosis InhA. *J. Chem. Inf. Model.* **2015**, *55* (3), 645–659. <https://doi.org/10.1021/ci500672v>.
- (88) Chang, M. W.; Ayeni, C.; Breuer, S.; Torbett, B. E. Virtual Screening for HIV Protease Inhibitors: A Comparison of AutoDock 4 and Vina. *PLOS ONE* **2010**, *5* (8), e11955. <https://doi.org/10.1371/journal.pone.0011955>.
- (89) Wang, Z.; Sun, H.; Yao, X.; Li, D.; Xu, L.; Li, Y.; Tian, S.; Hou, T. Comprehensive Evaluation of Ten Docking Programs on a Diverse Set of Protein–ligand Complexes: The Prediction Accuracy of Sampling Power and Scoring Power. *Phys. Chem. Chem. Phys.* **2016**, *18* (18), 12964–12975. <https://doi.org/10.1039/C6CP01555G>.
- (90) Adachi, N.; Senda, M.; Natsume, R.; Senda, T.; Horikoshi, M. Crystal Structure of Methanococcus Jannaschii TATA Box-Binding Protein. *Genes Cells Devoted Mol. Cell. Mech.* **2008**, *13* (11), 1127–1140. <https://doi.org/10.1111/j.1365-2443.2008.01233.x>.
- (91) Kamada, K.; Shu, F.; Chen, H.; Malik, S.; Stelzer, G.; Roeder, R. G.; Meisterernst, M.; Burley, S. K. Crystal Structure of Negative Cofactor 2 Recognizing the TBP-DNA Transcription Complex. *Cell* **2001**, *106* (1), 71–81. [https://doi.org/10.1016/S0092-8674\(01\)00417-2](https://doi.org/10.1016/S0092-8674(01)00417-2).

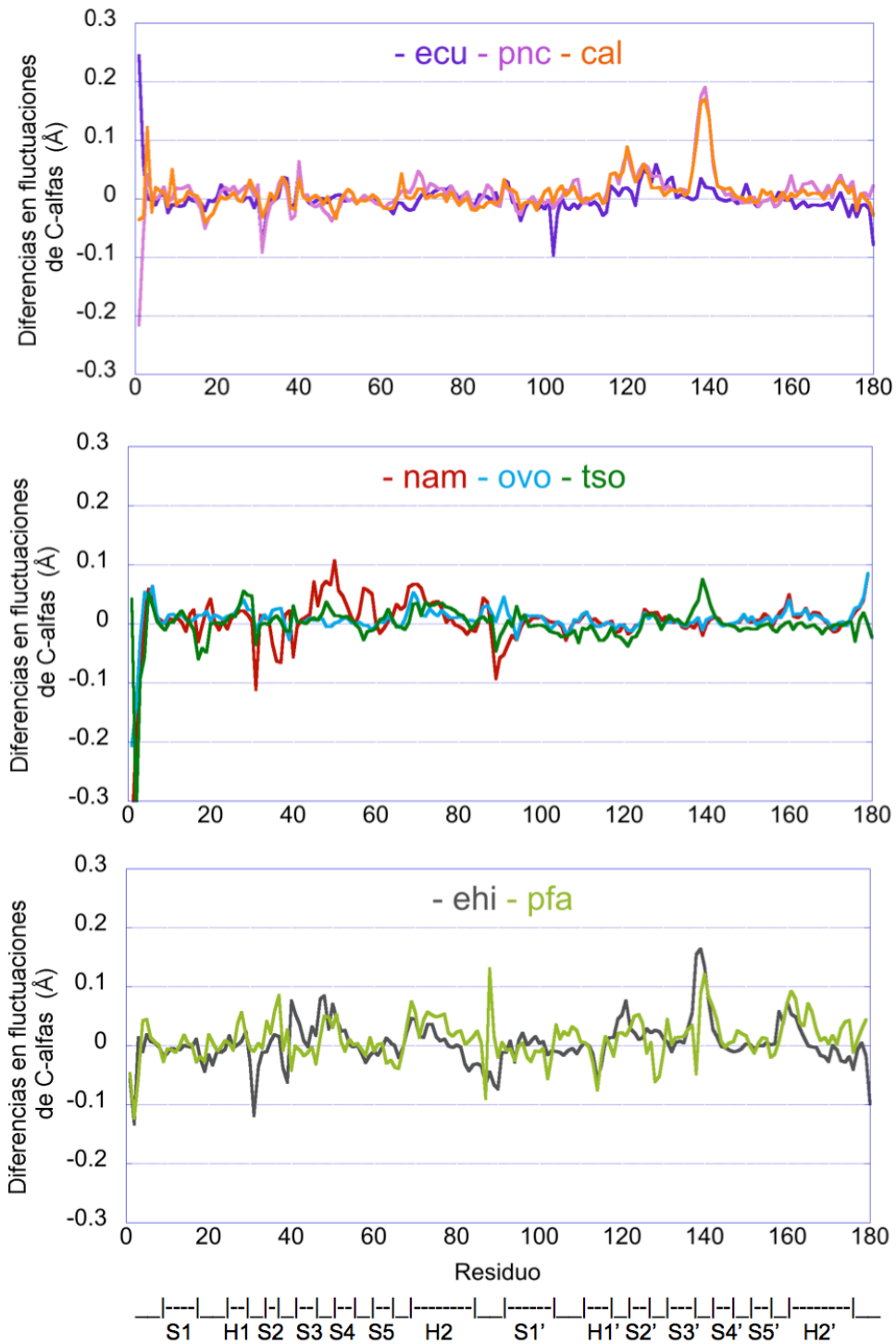
9. Anexos

Anexo 1. Validación estructural de los modelos de las TBPs usadas para MD. La tabla muestra los porcentajes de residuos en la región favorecida del mapa de Ramachandran para los modelos de las TBPs de acuerdo con la predicción de RAMPAGE. Un Valor cercano o por arriba de 90% indica un buen modelo.

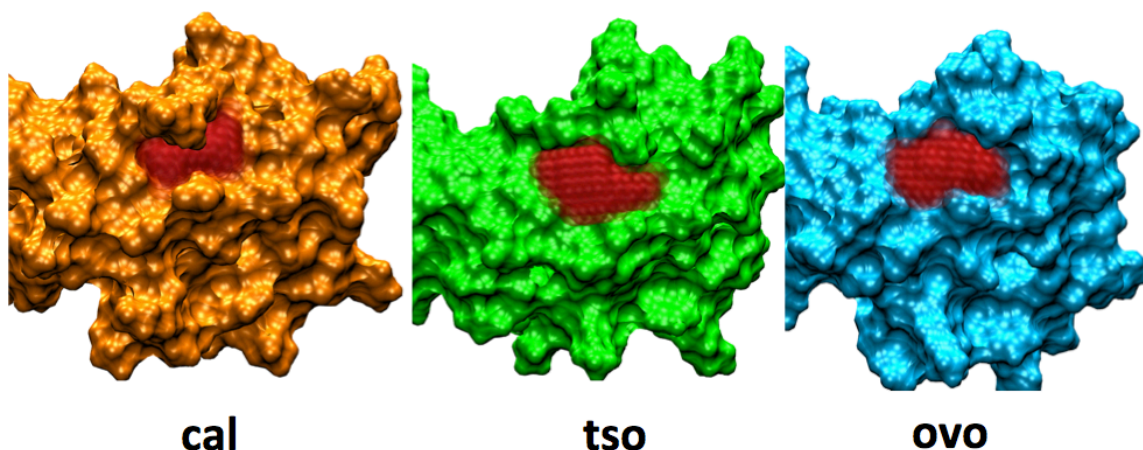
TBP	I-TASSER	MODELLER	SWISS-MODEL
pnc	92.0%	99.4%	97.7%
nam	90.6%	99.4%	97.6%
ovo	97.0%	99.4%	98.2%
cal	89.0%	100%	97.1%
ehi	90.6%	98.2%	97.7%
tso	89.0%	98.9%	96.0%
pfa	91.3%	93.0%	-



Anexo 2. RMSD contra la estructura inicial para cada TBP. Cada gráfico muestra la evolución del RMSD (en Å) en el tiempo contra la respectiva estructura inicial. Cada simulación fue iniciada de tres semillas distintas (ver *Selección de TBPs y generación de modelos*), que en el caso de las TBPs modeladas corresponden a I-TASSER (azul), MODELLER (verde) y SWISS-MODEL (rojo).



Anexo 3. Diferencias en fluctuaciones de carbonos alfa con respecto a hsaTBP. Las diferencias se muestran para cada TBP con respecto a **hsaTBP**, así como la ilustración de elementos de estructura secundaria. Valores positivos en la gráfica significan un mayor movimiento en las TBPs de parásitos.



Anexo 5. Cavity en el extremo N-terminal en calTBP, ovoTBP y tsoTBP. Para calTBP (naranja), tsoTBP (verde) y ovoTBP (azul), se muestra la cavity del N-terminal sobre las superficie. La cavity se indica con esferas rojas transparentes y fue calculada con POVME 3.0 sobre las estructuras representativas.

Anexo 6. Resultados del acoplamiento molecular flexible sobre el sitio extremo N-terminal de tsoTBP y ovoTBP. La selectividad corresponde a la diferencia del valor de energía entre la mejor pose en la TBP de parásitos y aquella en hsaTBP para el mismo compuesto. El código del compuesto corresponde al de la base de datos de ZINC12.

Acoplamiento flexible			
Compuesto	Selectividad	Energía (Kcal/mol)	Kd estimada a 298 K
ZINC02170910	-2.3	ovo -6.4/hsa -4.1	2.0×10^{-5}
ZINC00209227	-2.1	ovo -7.1/hsa -5.0	6.1×10^{-6}
ZINC00307043	-2.0	ovo -6.7/hsa -4.7	1.2×10^{-5}
ZINC06159424	-1.9	ovo -6.5/hsa -4.6	1.6×10^{-5}
ZINC00803769	-1.9	ovo -7.4/hsa -5.5	3.6×10^{-6}
ZINC04476398	-1.8	ovo -7.2/hsa -5.4	5.1×10^{-6}
ZINC00703555	-1.7	ovo -8.1/hsa -6.4	1.1×10^{-6}
ZINC08439724	-3.8	tso -9.5/hsa -5.7	1.0×10^{-7}
ZINC06159424	-3.7	tso -8.3/hsa -4.6	8.0×10^{-7}
ZINC01850581	-3.6	tso -8.0/hsa -4.4	1.3×10^{-6}
ZINC08384915	-3.4	tso -9.7/hsa -6.3	7.5×10^{-8}
ZINC00627460	-3.3	tso -9.3/hsa -6.0	1.4×10^{-7}
ZINC00627346	-3.3	tso -9.2/hsa -5.9	1.7×10^{-7}
ZINC00627492	-3.2	tso -9.1/hsa -5.9	2.0×10^{-7}



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS



INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS

Coordinación de Programas Educativos

Posgrado en Ciencias



**DR. VICTOR BARBA LÓPEZ
COORDINADOR DEL POSGRADO EN CIENCIAS
PRESENTE**

Atendiendo a la solicitud para emitir DICTAMEN sobre la revisión de la TESIS titulada “Caracterización de la superficie molecular de TBPs de parásitos de humanos, hacia la identificación de sitios para el diseño o repropósito de fármacos” que presenta el alumno José Ángel Santiago Terrones (5620131008) para obtener el título de **Doctor en Ciencias**.

Nos permitimos informarle que nuestro voto es:

NOMBRE	DICTAMEN	FIRMA
Dr. Lorenzo Patrick Segovia Forcella IBT-UNAM	Aprobado	
Dr. Carlos Daniel Amero Tello CIDC-UAEM	Aprobado	
Dra. Laura Domínguez Dueñas FQ-UNAM	Aprobado	
Dr. Samuel Enoch Estrada Soto FF-UAEM	Aprobado	
Dr. Martin González Andrade FM-UNAM	Aprobado	
Dra. Carmen Nina Pastor Colón CIDC-UAEM	Aprobado	
Dr. Rodrigo Said Razo Hernández CIDC-UAEM	Aprobado	